

DOCTORADO EN CIENCIA Y TECNOLOGIA

Evaluado y acreditado por la Comisión Nacional de Evaluación y Acreditación
Universitaria (CONEAU). *Resolución N° 1178/11. Calificación "B".*

**Estimación de Parámetros en Imágenes
SAR Monopolarizadas Usando Distancias
Estocásticas y Núcleos Asimétricos**

TRABAJO DE TESIS PARA OPTAR POR EL TÍTULO DE
DOCTOR EN CIENCIA Y TECNOLOGÍA DE LA
UNIVERSIDAD NACIONAL DE GENERAL SARMIENTO

AUTOR: **JULIA ANALÍA CASSETTI**

DIRECTOR: DR. ALEJANDRO C. FRERY

Diciembre 2019

Niveles de acceso al documento autorizados por el autor.

El autor de la tesis puede elegir entre las siguientes posibilidades para autorizar a la UNGS a difundir el contenido de la tesis: a

- a) Liberar el contenido de la tesis para acceso público.
- b) Liberar el contenido de la tesis solamente a la comunidad universitaria de la UNGS.
- c) Retener el contenido de la tesis por motivos de patentes, publicación y/o derechos de autor por un lapso de cinco años.

- 1. Título completo del trabajo de Tesis: Estimación de Parámetros en Imágenes SAR Monopolarizadas Usando Distancias Estocásticas y Núcleos Asimétricos.
- 2. Presentado por (Apellido/s y Nombres completos del autor): Cassetti, Julia Analía.
- 3. E-mail del autor: julia.cassetti@gmail.com.
- 4. Estudiante del Posgrado (consignar el nombre completo del Posgrado): Doctorado en Ciencia y Tecnología.
- 5. Institución o Instituciones que dictaron el Posgrado (consignar los nombres desarrollados y completos): Universidad Nacional de General Sarmiento.
- 6. Para recibir el título de (consignar completo):
 - a) Grado académico que se obtiene: Doctor.
 - b) Nombre del grado académico: Doctor en Ciencia y Tecnología.
- 7. Fecha de la defensa: día / mes / año
- 8. Director de la Tesis (Apellidos y Nombres): Frery Orgambide, Alejandro César.

-
9. Tutor de la Tesis (Apellidos y Nombres): Borgna, Juan Pablo.
 10. Colaboradores con el trabajo de Tesis:
 11. Descripción física del trabajo de Tesis (cantidad total de páginas, imágenes, planos, videos, archivos digitales, etc.):
 - 140 páginas;
 - 13 tablas;
 - 61 figuras.
 12. Alcance geográfico y/o temporal de la Tesis: El alcance geográfico de la tesis es internacional.
 13. Temas tratados en la Tesis (palabras claves): Estimación de parámetros, Sensores Remotos, Imágenes SAR, Núcleos asimétricos, Distancias estocásticas, Estimadores de mínima distancia, Speckle.
 14. Resumen en español (hasta 1000 caracteres):

Resumen

El modelo \mathcal{G}^0 es una buena elección para explicar las características estadísticas de datos que provienen de imágenes de radar de apertura sintética (SAR), porque bajo este modelo se pueden caracterizar regiones con diferentes grado de textura a través de sus parámetros. Esta tesis propone una nueva estrategia para la estimación del parámetro de textura del modelo \mathcal{G}_0 para datos de intensidad, por medio de la minimización de distancias estocásticas entre la función de densidad teórica, y una estimación no paramétrica de la función de densidad subyacente que proviene de los datos observados utilizando núcleos asimétricos. Se comparará el desempeño de estos estimadores, en términos de sesgo y error cuadrático medio, con los obtenidos por el Método de Momentos, Máxima Verosimilitud y Logcumu-

lantes. Además se estudiarán propiedades de convergencia, como así también la robustez de los mismos bajo contaminación.

15. Resumen en portugués (hasta 1000 caracteres):

Resumo

O modelo \mathcal{G}^0 é uma boa opção para explicar as características estatísticas dos dados provenientes de imagens de radar de abertura sintética (SAR), porque nesse modelo regiões com diferentes graus de textura podem ser caracterizadas através de seus parâmetros. Esta tese propõe uma nova estratégia para a estimativa do parâmetro de textura do modelo \mathcal{G}_0 para dados de intensidade, através da minimização de distâncias estocásticas entre a função de densidade teórica e uma estimativa não paramétrica da função de densidade subjacente que vem dos dados observados usando núcleos assimétricos. O desempenho desses estimadores será comparado, em termos de viés e erro médio quadrático, aos obtidos pelo Método dos Momentos, Máxima Verossimilhança e Logcumulantes. Além disso, serão estudadas propriedades de convergência, bem como sua robustez sob contaminação.

16. Resumen en inglés (hasta 1000 caracteres):

Abstract

The \mathcal{G}^0 model is a good choice to explain the statistical characteristics of data that come from synthetic aperture radar (SAR) images, because under this model, areas with different degrees of texture can be characterized through their parameters. This thesis proposes a new strategy for the estimation of the texture parameter of the *mathcal{G}_0* model for intensity data, through the minimization of stochastic distances between the theoretical density function, and a non-parametric estimation of the underlying density function that comes from the observed data using asymmetric kernels. The performance of these estimators will be compared, in terms of

bias and mean square error, with those obtained by the Method of Moments, Maximum Likelihood and Logcumulantes. In addition, convergence properties will be studied, as well as their robustness under contamination.

17. Aprobado por (Apellidos y Nombres del Jurado):

Firma y aclaración de la firma del Presidente del Jurado:

Firma del autor de la tesis:

Publicaciones:

- En revistas con referato
 - M. Gambini, J. Cassetti, M. Lucini, A. Frery. *Parameter Estimation in SAR Imagery Using Stochastic Distances and Asymmetric Kernels* - IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 8, pp. 365-375, 2015.
- En actas de congresos
 - D. Chan, J. Cassetti, A. Frery - *Texture Parameter Estimation in Monopolarized SAR Imagery, for the Single Look Case* - IEEE International Geoscience and Remote Sensing Symposium - 2016 - ISSN: 2153-7003 - DOI: 10.1109/IGARSS.2016.7729845.
 - J. Cassetti, J. Gambini, A. Frery - *Parameter Estimation in SAR Imagery using Stochastic Distances* - The 4th Asia-Pacific Conference on Synthetic Aperture Radar, Tsukuba, Japón - pp. 573-576 - 2013- INSPEC Accession Number: 14026655
 - J. Cassetti, J. Gambini, A. Frery - *Estimación de Parámetros utilizando Distancias Estocásticas para Datos con Ruido Speckle*, Anales de la 42 JAIIO - 2013 - ISSN 1850-2776.

Trabajos presentados en congresos:

- IV JIAAIS Interdisciplinary Workshop on Advanced Signal and Image Analysis - *Gamma and Lognormal Kernels with Stochastic Distance for Texture Parameter Estimation under the Intensity Multilook \mathcal{G}_I^0 Law* - J. Cassetti, A. Frey - Maceió - Brasil - 2017 - Poster - Expositor.
- IEEE International Geoscience and Remote Sensing Symposium IGARSS - *Texture Parameter Estimation in Monopolarized SAR Imagery, for the Single Look Case* - D. Chan, J. Cassetti, A. Frery - Beijing, China 2016 - Presentación oral.

- Segundas Jornadas Interdisciplinarias de Análisis Avanzado de Imágenes y Señales - *Parameter Estimation in SAR Imagery using Stochastic Distances and Asymmetric Kernels* - J. Cassetti, A. Frery - Universidad Nacional de General Sarmiento - Argentina - octubre 2015 - Presentación oral - Expositor.
- Segundas Jornadas Interdisciplinarias de Análisis Avanzado de Imágenes y Señales - *Comparación de estimadores paramétricos para la distribución \mathcal{G}_I^0 para el caso $L = 1$* - D. Chan; J. Cassetti y A. Frery - Universidad Nacional de General Sarmiento - Argentina - 2015 - Presentación oral.
- IV Workshop en Ecuaciones de la Física Matemática - *Estimación del parámetro de textura de la distribución \mathcal{G}_I^0* - J. Cassetti, A. Frery, Universidad Nacional de General Sarmiento - Los Polvorines - Argentina - 2015. Presentación oral - Expositor.
- XVIII Conference on Nonequilibrium Statistical Mechanics and Nonlinear Physics - Medyfinol- *Texture Analysis in SAR Imagery using the \mathcal{G}_I^0 Distribution* - J. Gambini, M. Lucini, J. Cassetti, A. Frery - Maceió - Brasil - Poster - 2014.

Aportes Originales:

(Especificar cuales son los aportes originales o innovadores conseguidos en la realización de esta tesis. Indicar donde se encuentran. Máximo una carilla)

En el capítulo 4 de esta tesis, subsección 4.1.1 se presentan nuevos resultados teóricos de la familia de distribuciones \mathcal{G}_I^0 que modelan datos provenientes de radares de apertura sintética (SAR). Se prueba que esta distribución tiene colas pesadas y la convergencia uniforme de la función de densidad $f_{\mathcal{G}_I^0}$ sobre intervalos compactos $[z_1, z_2] \subset (0, +\infty)$, cuando el parámetro de textura se acerca a 0 y a $-\infty$.

En el capítulo 6 se propone un nuevo estimador para el parámetro de textura de la distribución \mathcal{G}_I^0 . Este estimador es el valor del argumento que minimiza una distancia estocástica entre el modelo teórico y una estimación no paramétrica de la función de densidad subyacente con núcleos asimétricos. El aporte de la propuesta es la combinación de utilizar funciones de densidad en vez de distribución en el estimador de mínima distancia, junto con distancias estocásticas y la estimación por núcleos asimétricos de la función de densidad subyacente. Se analiza el desempeño del estimador propuesto especialmente para el caso de muestras de pequeño tamaño ya que es de interés en el procesamiento de imágenes SAR. Este análisis se realizó para diferentes combinaciones de valores de los parámetros, tamaños de muestra y niveles de procesamiento comparando su performance con los estimadores clásicos que existen en la literatura. Se estudió también la performance del estimador frente a diferentes escenarios de contaminación.

En el capítulo 4 subsección 6.7 se aplica el estimador propuesto en esta tesis a una imagen real.

En el capítulo 7 se prueba que el estimador propuesto es fuertemente consistente del parámetro de textura.

Agradecimientos

Este trabajo no habría sido posible sin el apoyo y el estímulo

Índice general

1. Resumen	16
2. Introducción	18
3. Radar de Apertura Sintética	25
3.1. Sensores Remotos SAR	28
3.2. Sistema de iluminación coherente	30
3.3. Espectro Electromagnético	31
3.4. Polarización	33
3.5. Geometría SAR	35
3.5.1. Resolución	37
3.5.2. Ecuación del radar	39
3.6. Formación del ruido Speckle	40
3.7. Doble Bounce o Doble Rebote	42
3.8. Retorno Complejo	43
3.9. Conclusiones	44
4. Modelo \mathcal{G}_0	45
4.0.1. Modelo Multiplicativo	47
4.0.2. Modelo para el Ruido Speckle	49
4.0.3. Modelo para el Backscatter	51
4.0.4. Modelo para el Retorno	53
4.1. Modelo \mathcal{G}_I^0	55

4.1.1. Resultados Teóricos	57
4.2. Conclusiones	62
5. Metodología	64
5.1. Estimación Paramétrica	65
5.1.1. Método de los Momentos	66
5.1.2. Método de Máxima Verosimilitud	67
5.1.3. Logmomentos y Logcumulantes	69
5.2. Estimación No Paramétrica	73
5.2.1. Distancias Estocásticas	73
5.2.2. Distancia triangular generalizada	78
5.2.3. Núcleos	80
5.2.4. Estimadores de Mínima Distancia	93
5.3. Conclusiones	96
6. Simulaciones y aplicación a una imagen real	98
6.1. Generando muestras \mathcal{G}_I^0	100
6.2. Eligiendo distancias estocásticas	101
6.3. Incorporando núcleos y evaluando robustez	107
6.4. Mejorando la propuesta	123
6.5. Evaluación de la performance entre los distintos métodos de estimación	126
6.6. Función de influencia empírica	133
6.7. Aplicación a una imagen real	137
6.8. Conclusiones	143
7. Resultados Teóricos	145
8. Conclusiones	151
A. Información computacional	155

Índice de figuras

3.1. Resolución de un radar.	26
3.2. Diferentes tipos de sistemas de iluminación.	31
3.3. Espectro electromagnético.	32
3.4. Diferentes tipos de polarizaciones.	34
3.5. Geometría del Radar de Apertura Sintética SAR.	36
3.6. Dispersión de la señal emitida por un radar SAR.	37
3.7. Apertura Sintética en un SAR.	38
3.8. Radiación reflejada por distintos objetos en la superficie terrestre.	40
3.9. Coherencia de la señal reflejada.	41
3.10. Diferentes reflexiones de una señal emitida.	43
4.1. Modelo Multiplicativo, fuente Lee et.al [62].	48
4.2. Comportamiento asintótico da la distribución $N^{-1}(\alpha, \gamma, \lambda)$	52
4.3. Comportamiento asintótico de la distribución $\mathcal{G}_I(\alpha, \gamma, \lambda, L)$	54
4.4. Densidad \mathcal{G}_I^0 para $L = 3$ y $\alpha = -1.5, -5, -8$	56
4.5. Densidad \mathcal{G}_I^0 para $L = 1, 3, 8$ y $\alpha = -5$	56
4.6. Convergencia uniforme de $f_{\mathcal{G}_I^0}(z)$ a $\Gamma(3, 3)$ cuando $\gamma = \gamma^*$, $\alpha \rightarrow -\infty$, $L = 3$ y $\epsilon = 0.1$	61
4.7. Convergencia uniforme de $f_{\mathcal{G}_I^0}(z)$ a 0 cuando $\gamma = \gamma^*$, $\alpha \rightarrow -1$, $L = 3$ y $\epsilon = 0.2$	62
5.1. Función de verosimilitud para $L = 3$	69
5.2. Estimación de \mathcal{G}_I^0 para $L = 3$, $\alpha = -1.5$, $\gamma = 0.5$ y $n = 10$	86

5.3. Estimación de \mathcal{G}_I^0 para $L = 8$, $\alpha = -5$, $\gamma^* = 4$ y $n = 25$	87
5.4. Least Square Cross Validation (LSCV) para $L = 3$ y $\alpha = -5$	91
6.1. Distancias Hellinger, Renyi y Triangular, $\alpha_0 = -2, -3, -4, -5, -6, -7$ y $L = 1$	102
6.2. Distancias Hellinger, Renyi y Triangular, $\alpha_0 = -2, -3, -4, -5, -6, -7$ y $L = 3$	103
6.3. Distancias Hellinger, Renyi y Triangular, $\alpha_0 = -2, -3, -4, -5, -6, -7$ y $L = 8$	104
6.4. $\widehat{\text{Sesgo}}$ y $\widehat{\text{ECM}}$ de $\hat{\alpha}$, $L=1$	107
6.5. $\widehat{\text{Sesgo}}$ y $\widehat{\text{ECM}}$ de $\hat{\alpha}$, $L=3$	108
6.6. $\widehat{\text{Sesgo}}$ y $\widehat{\text{ECM}}$ de $\hat{\alpha}$, $L=8$	109
6.7. Resultado de estimar el parámetro α para cada pixel, usando ventanas deslizantes de tamaño 7×7	110
6.8. Resultado de estimar el parámetro α para cada pixel, usando ventanas deslizantes de tamaño 3×3	110
6.9. Densidad de $\mathcal{G}_I^0(\alpha, \gamma^*, L)$	111
6.10. $\widehat{\text{Sesgo}}$ y $\widehat{\text{ECM}}$ para datos sin contaminar, $L = 1$	113
6.11. $\widehat{\text{Sesgo}}$ y $\widehat{\text{ECM}}$ para datos sin contaminar, $L = 3$	113
6.12. $\widehat{\text{Sesgo}}$ y $\widehat{\text{ECM}}$ para datos sin contaminar, $L = 8$	114
6.13. Datos contaminados: Caso 1, $\epsilon = 0.01$ y $L = 1$	117
6.14. Datos contaminados: Caso 1, $\epsilon = 0.01$ y $L = 3$	118
6.15. Datos contaminados: Caso 1, $\epsilon = 0.01$ y $L = 8$	119
6.16. Datos contaminados: Caso 2, $\epsilon = 0.001$ y $L = 1$	119
6.17. Datos contaminados: Caso 2, $\epsilon = 0.001$ y $L = 3$	120
6.18. Datos contaminados: Caso 2, $\epsilon = 0.001$ y $L = 8$	121
6.19. Datos contaminados: Caso 3, $\epsilon = 0.005$ y $L = 1$	121
6.20. Datos contaminados: Caso 3, $\epsilon = 0.005$ y $L = 3$	122
6.21. Datos contaminados: Caso 3, $\epsilon = 0.005$ y $L = 8$	122

6.22. Image real E-SAR junto con la regiones usadas para estimar el parámetro α	123
6.23. Imagen real SAR de 1-look con un corner reflector.	124
6.24. Muestras de varios tamaños en una imagen real SAR de 1-look con un corner reflector.	125
6.25. Sesgo y ECM estimados para datos sin contaminar y $L = 3$	130
6.26. Sesgo y ECM estimados para datos sin contaminar, $L = 3$ y $\alpha = -1.5, -3$	131
6.27. Sesgo y ECM estimados para datos sin contaminar y $L = 8$	131
6.28. Sesgo y ECM estimados para datos sin contaminar, $L = 8$ y $\alpha = -1.5, -3$	132
6.29. Sesgo y ECM estimados para datos contaminados, Caso 1, $\epsilon = 0.05$ y $L = 3$	133
6.30. Sesgo y ECM estimados para datos contaminados, Caso 1, $\epsilon = 0.05$, $\alpha = -1.5, -3$ y $L = 3$	134
6.31. Sesgo y ECM estimados para datos contaminados, Caso 1, $\epsilon = 0.05$ y $L = 8$	135
6.32. Sesgo y ECM estimados para datos contaminados, Caso 1, $\epsilon = 0.05$, $\alpha = -1.5, -3$ y $L = 8$	135
6.33. FIES para $\hat{\alpha}_{MV}$, $\hat{\alpha}_{\Gamma}$, $\hat{\alpha}_{LN}$, $\hat{\alpha}_{LC}$ para $L = 3$, $n = 25$ y $\alpha = -1.5, -3$	136
6.34. FIES para $\hat{\alpha}_{MV}$, $\hat{\alpha}_{\Gamma}$, $\hat{\alpha}_{LN}$, $\hat{\alpha}_{LC}$ para $L = 3$, $n = 25$ y $\alpha = -5, -8$	136
6.35. FIES para $\hat{\alpha}_{MV}$, $\hat{\alpha}_{\Gamma}$, $\hat{\alpha}_{LN}$, $\hat{\alpha}_{LC}$ para $L = 8$, $n = 25$ y $\alpha = -1.5, -3$	137
6.36. FIES para $\hat{\alpha}_{MV}$, $\hat{\alpha}_{\Gamma}$, $\hat{\alpha}_{LN}$, $\hat{\alpha}_{LC}$ para $L = 8$, $n = 25$ y $\alpha = -5, -8$	137
6.37. Muestras para estimación de ENL en una imagen Esar multilook.	138
6.38. Tres muestras utilizadas para comparar $\hat{\alpha}_{MV}$, $\hat{\alpha}_{\Gamma}$, $\hat{\alpha}_{LN}$ y $\hat{\alpha}_{LC}$	139
6.39. Cinco muestras utilizadas para comparar $\hat{\alpha}_{MV}$, $\hat{\alpha}_{\Gamma}$, $\hat{\alpha}_{LN}$ y $\hat{\alpha}_{LC}$	141
6.40. $\hat{\alpha}_{MV}$, $\hat{\alpha}_{\Gamma}$, $\hat{\alpha}_{LN}$ y $\hat{\alpha}_{LC}$ para cinco muestras variando el tamaño de las ventanas de la imagen 6.39.	142

Indice de tablas

3.1. Descripción de algunos sensores SAR junto con sus características. Adaptado de Moreira et al. [68].	30
3.2. Bandas SAR, datos adaptados de [87]	33
6.1. Tiempos medios para datos simulados sin contaminación, $L = 1$, $n = 81$.	114
6.2. Porcentaje de casos de no convergencia para los estimadores de Momentos y Logcumulante en el Caso 1, $\alpha_2 = -15$, $\epsilon = 0.01$. . .	118
6.3. $\hat{\alpha}$ para las muestras de la figura 6.22.	120
6.4. $\hat{\alpha}$ y tiempos de procesos para las muestras presentadas en la figu- ra 6.24	121
6.5. MISE y porcentaje de casos de no convergencia $L = 3$	127
6.6. Sesgo y ECM para $L = 3$	128
6.7. Porcentaje de casos de no convergencia $L=3$	129
6.8. Tiempos medios para datos simulados sin contaminación, $L = 3$, $n = 81$.	129
6.9. Valores de $\hat{\alpha}_{MV}$, $\hat{\alpha}_T$, $\hat{\alpha}_{LN}$ y $\hat{\alpha}_{LC}$ para las tres muestras de la ima- gen 6.38.	140
6.10. Valores de $\hat{\alpha}_{MV}$, $\hat{\alpha}_T$, $\hat{\alpha}_{LN}$ y $\hat{\alpha}_{LC}$ para la muestra amarilla sin el primer elemento.	140
6.11. $\hat{\alpha}$ para cinco muestras variando el tamaño de las ventanas de ima- gen 6.39.	142

Capítulo 1

Resumen

Las imágenes obtenidas con dispositivos que poseen iluminación coherente como lo son el ultrasonido, laser o radar de apertura sintética, son afectadas por la presencia de un ruido que es inherente al proceso de captura de la imagen, llamado ruido speckle. Este ruido se aparta de las hipótesis clásicas ya que no es gaussiano, no es aditivo y es difícil de eliminar.

Durante la última década se ha dedicado especial atención al modelado de datos que provienen de imágenes de radares de apertura sintética (SAR). El uso de modelos estadísticos para modelar estos datos es una herramienta fundamental para explicar las características estadísticas de los mismos y, en este sentido, el modelo \mathcal{G}^0 es una buena elección porque bajo él se pueden caracterizar regiones con diferente grado de textura y brillo a través de sus parámetros. Luego, la estimación de parámetros cumple un rol fundamental en el análisis de imágenes SAR.

La familia de distribuciones \mathcal{G}^0 tienen soporte positivo y se indexan por tres parámetros: textura, escala y número de looks. Se han propuesto muchas estrategias para estimar el parámetro de textura, entre ellas, el método de Máxima Verosimilitud, Momento, métodos robustos y Logcumulantes.

Esta tesis propone y analiza una nueva estrategia para la estimación del parámetro de textura del modelo \mathcal{G}_0 para datos de intensidad, por medio de la

minimización de distancias estocásticas entre la función de densidad teórica \mathcal{G}_I^0 y una estimación no paramétrica de la función de densidad subyacente que proviene de los datos observados. La propuesta es estimar dicha función de densidad utilizando núcleos asimétricos ya que la distribución tiene soporte positivo. Se comparará el desempeño de estos estimadores, en términos de sesgo, error cuadrático medio, con los obtenidos por el métodos de Momentos, Máxima Verosimilitud y Logcumulantes. Además se estudiará el comportamiento de estos estimadores y sus propiedades, como así también la robustez de los mismos bajo diferentes escenarios de contaminación.

Capítulo 2

Introducción

La teledetección constituye una herramienta de gran utilidad para el desarrollo de sistemas de prevención, seguimiento y evaluación de superficie terrestre. Se define como la técnica que permite adquirir imágenes de la superficie terrestre desde sensores instalados en plataformas espaciales o aereotransportadas. Este proceso se basa en la detección y medición de radiación electromagnética emitida ya sea de manera natural o a partir de una fuente artificial emisora. La energía es captada por sensores y permite identificar objetivos por sus propiedades físicas y sustancias que lo componen. Estos sensores emiten radiación electromagnética comprendida en varias zonas del espectro electromagnético. La parte del espectro electromagnético usada en teledetección se extiende aproximadamente desde $10^{-1}\mu m$ hasta $10^9\mu m$, donde μm se llama al micrómetro que es la millonésima parte del metro.

Esta técnica de teledetección comprende dos procesos. Uno de estos procesos es la adquisición de información de la superficie terrestre captando la radiación electromagnética emitida o reflejada por ésta. El otro es el almacenamiento de la información obtenida para su posterior procesamiento en virtud de interpretarla y utilizarla.

El sensor remoto recibe la radiación que proviene de los objetos situados sobre la superficie terrestre. En función de la forma en la que se refleje esta radiación

podremos clasificar los distintos tipos de terrenos.

Los sistemas de teledetección se pueden dividir principalmente en dos grupos: los teledetectores activos y los pasivos. Los sensores activos tienen su propia fuente de energía para iluminar los objetos que observan. Ellos emiten radiación en la dirección del objetivo a ser investigado, detecta y mide la radiación que es reflejada o retrodispersada desde el objetivo. Los sensores pasivos, por otra parte, detectan la radiación electromagnética emitida o reflejada de fuentes naturales, no necesitan una fuente de energía externa. La luz del sol reflejada es la fuente más común de la radiación medida por los sensores pasivos. Un ejemplo de sensor pasivo es el espectrómetro óptico o espectroscopio, que es un instrumento que sirve para medir las propiedades de la luz en una determinada porción del espectro electromagnético. El Radar es un ejemplo de sensor activo.

Los sistemas de Radar (radio detection and ranging: detección y medición de distancias por radio) son instrumentos que, a través de ondas electromagnéticas, detectan un objeto e indican su distancia y posición. Estos instrumentos miden la respuesta del terreno a la radiación electromagnética emitida en forma de pulsos, el valor de esta respuesta es almacenada para su posterior procesamiento y se utiliza para formar una imagen de la zona de interés. Un Radar de Apertura Sintética (synthetic aperture radar: SAR) es un tipo de sistema que consiste en procesar, mediante algoritmos, la información capturada por la antena del radar. Mediante este proceso se consigue el mismo rendimiento que se obtendría si se utilizara una antena mucho más grande que la que tiene en realidad, por eso se llaman de apertura sintética.

Los dispositivos de captura de imágenes que emplean iluminación coherente, como sucede en las imágenes de ultrasonido B, laser y de radar de apertura sintética – SAR (Synthetic Aperture Radar) entre otros, introducen un ruido que es propio del sistema de captura de la imagen. Este ruido, llamado speckle, no es gaussiano ni aditivo y, por lo tanto, diferente al ruido que se observa en imágenes ópticas. Existen diferentes técnicas para disminuir la presencia de este

ruido, pero como contrapartida la imagen pierde resolución.

La utilización de modelos estadísticos ha sido una herramienta fundamental para analizar e interpretar datos con ruido speckle. Se han presentado varias distribuciones para modelar este tipo de imágenes. Las distribuciones Lognormal y Weibull fueron introducidas para caracterizar datos de alta resolución [73], las distribuciones K y Weibull también fueron estudiadas en [72], en [96] los autores presentaron la distribución de Fisher para modelar varios tipos de áreas. Li et al. [63] propusieron una distribución Gamma generalizada para modelar imágenes SAR que tiene las distribuciones Weibull, K y Fisher, entre otras, como casos particulares. Cabe señalar que la distribución \mathcal{G}^0 está relacionada con la distribución de Fisher [66].

El modelo propuesto para datos provenientes de un sistema de iluminación por radiación coherente, como son los datos SAR, es un modelo multiplicativo que considera que el valor observado en cada celda de la imagen es una variable aleatoria Z que resulta del producto de dos variables aleatorias independientes: una correspondiente al backscatter o retrodispersión X (que es lo que observaríamos sin la presencia del ruido speckle) y la otra correspondiente al ruido speckle Y (que es inherente a todo sistema de captura de imágenes con iluminación coherente) [73].

En los últimos años, se ha utilizado exitosamente el modelo \mathcal{G}^0 : \mathcal{G}_A^0 para datos de amplitud y \mathcal{G}_I^0 para datos de intensidad, porque tiene la capacidad de discriminar áreas muy texturadas o extremadamente texturadas mejor que otros modelos. Este modelo fue propuesto por Frery et al. [36] y está caracterizado por tres parámetros; el parámetro α que explica la textura, el parámetro γ que informa sobre el brillo de la imagen y el número de looks L que está relacionado con la relación señal-ruido. Debido a esta interpretabilidad, es crucial obtener estimaciones de calidad de dichos parámetros, en particular para α .

Varios autores han estudiado el problema de la estimación de parámetros para la familia \mathcal{G}^0 . Freitas et al. [34] usaron el primer y segundo momento para estimar

los parámetros de textura y escala en el caso de datos de intensidad. Vasconcellos et al. [97] cuantifican el error en la estimación del parámetro de textura para la distribución \mathcal{G}_A^0 y proponen una técnica analítica para mejorar la estimación a través de una corrección de segundo orden en el sesgo de la estimación del parámetro de textura. Li et al [63] obtuvieron una expresión para los estimadores de la distribución Gamma generalizada basada en cumulants de segundo tipo y una aproximación de segundo orden de la función Polygamma. Cribari-Neto et al. [24] implementaron técnicas de remuestreo para mejorar la estimación de parámetros para el modelo \mathcal{G}_I^0 .

Nicolas y Anfinsen [71] propusieron el método basado en Logmomentos y Logcumulantes para estimar los parámetros de una distribución. Estos métodos dependen de la relación entre los momentos de la distribución y su transformada de Mellin. La misma idea fue aplicada por Khan y Guida [58] para estimar los parámetros del modelo \mathcal{G} multivariado en el caso de datos SAR polarimétricos y también fue utilizada por Tison et al. [96] para la estimación del modelo \mathcal{G} para datos de amplitud.

Unas de las propiedades deseables para un estimador es su robustez, esto es, su capacidad para dar buenas estimaciones aún en presencia de datos atípicos. Bustos et al. [16] y Allende et al. [2] propusieron M y AM estimadores para mejorar el comportamiento del estimador de Máxima Verosimilitud bajo contaminación. Ellos mostraron que su propuesta, si bien supera la performance del estimador de Máxima Verosimilitud, presenta problemas numéricos especialmente para el caso de muestras de pequeño tamaño.

Por otro lado, la teoría de la información ha sido aplicada a los métodos de estadística y probabilidades con éxito [64]. Shannon [91] definió la información $I(X, Y)$ entre las variables aleatorias X e Y como una divergencia calculada entre sus densidades de probabilidad. Estas divergencias fueron ampliamente estudiadas por Kullback y Leibler [60] y por Rényi [82] entre otros y es una medida de la distancia entre dos funciones de densidad. Este tipo de divergencias poseen

múltiples aplicaciones en procesamiento de señales e imágenes [6], análisis de imágenes médicas [98], clasificación de texturas [78], restauración de imágenes [11] e incluso en detección automática de regiones con diferente grado de rugosidad en imágenes SAR [38, 92].

Los estimadores de mínima distancia (MDE) son una alternativa con buenas propiedades para el problema de la estimación de parámetros. Surgen de la idea de encontrar un estimador que sea el valor que minimiza la medida de la distancia entre las funciones de distribución empírica y teórica. Wolfowitz [99, 101] estudió esta clase de estimadores y demostró que, bajo condiciones generales, estos estimadores son fuertemente consistentes. Boos [9] estudió la distancia ponderada de Cramer-von Mises entre la función de distribución empírica y el modelo verdadero. El autor mostró que estos estimadores son consistentes y asintóticamente eficientes bajo cierta función de peso. Hettmansperger et al. [46] estudiaron estimadores de mínima distancia no pesados aplicados a un modelo de posición-escala a partir de la distancia de Cramer-von Mises. El autor demostró que estos estimadores son asintóticamente normales y tienen buenas propiedades de eficiencia y robustez. Beran [7] propone un MDE estimador utilizando la distancia de Hellinger entre un modelo teórico y un estimador de densidad no paramétrico utilizando núcleos simétricos y mostró que este estimador es asintóticamente eficiente bajo ciertas familias paramétricas de densidades. Parr y Schucany [75] demostraron que, considerando ciertas condiciones, el estimador MDE entre la función de distribución empírica y la función de distribución teórica es fuertemente consistente.

Cao et al. [18] propuso minimizar una distancia entre la función de densidad teórica y un estimador de la función de densidad subyacente utilizando el estimador de núcleo simétrico clásico para obtener estimadores de distancia mínima. Demostraron, siguiendo [75], la consistencia fuerte de estos estimadores bajo ciertas consideraciones, y también estudiaron su normalidad asintótica para el caso de la métrica L^2 .

Dentro de los estimadores no paramétricos de la función de densidad subya-

cente se encuentran los estimadores de kernel clásico, con kernel simétrico [93] que son populares en la estimación de la función de densidad. Sin embargo, si la función de densidad a estimar tiene soporte acotado, estos estimadores pueden dar estimaciones sesgadas en los bordes del soporte porque asignan probabilidad positiva fuera del soporte de la función. Una alternativa para mejorar esto es utilizar kernels asimétricos, Chen [22, 23] presenta núcleos Beta y Gamma, Scaillet [88] introduce núcleos Inverso Gaussianos (IG) y Recíproco Inverso Gaussiano (RIG), Bouezmarni et al. [10] demuestran propiedades teóricas de los núcleos Gamma, IG y RIG. En [54] los autores proponen los núcleos Birnbaum-Saunders (BS) y Lognormal (LN). Es interesante señalar que estos estimadores varían su forma de acuerdo con la observación, una característica que permite obtener diferentes grados de suavizamiento sin incurrir en los problemas antes mencionados [88].

En esta tesis se propone un nuevo estimador de los parámetros de la distribución \mathcal{G}_I^0 definido como el punto del espacio paramétrico que minimiza la distancia estocástica que existe entre la función de densidad teórica \mathcal{G}_I^0 y una estimación no paramétrica de la función de densidad subyacente utilizando núcleos asimétricos. Se muestra que la distancia triangular es una buena elección para tratar este problema y se prueba la consistencia y la convergencia en casi todo punto de este estimador. Asimismo, a través de simulaciones Monte Carlo, se analiza:

- El desempeño de este estimador para pequeñas muestras comparándolo, en términos de sesgo y error cuadrático medio y tasa de convergencia, con los estimadores presentes en la literatura.
- La robustez del estimador propuesto bajo diferentes esquemas de contaminación.

Esta tesis está organizada de la siguiente manera:

- En el capítulo 3 presentamos conceptos de teledetección y los radares de apertura sintética dando una breve reseña de principales sensores SAR.

Presentamos también la generación de imágenes SAR y la formación del ruido speckle.

- En el capítulo 4 se introduce el modelo multiplicativo base de la metodología adoptada para explicar el comportamiento de las imágenes SAR. Presentamos las distribuciones de cada una de las variables que forman parte del modelo multiplicativo y, finalmente, presentamos la familia de distribuciones \mathcal{G}_I objeto de estudio en esta tesis.
- En el capítulo 5 se describen dos metodologías de estimación de la función de densidad, paramétrica y no paramétrica. Damos una descripción de cada una de estas metodologías y presentamos el principal aporte de esta tesis, que es la propuesta de un nuevo método de estimación para los parámetros del modelo \mathcal{G}_I^0 .
- En el capítulo 6 se presenta el estudio realizado, a través de simulaciones Monte Carlo, del estimador propuesto. Se muestran los resultados obtenidos de estas simulaciones con datos contaminados y sin contaminar, y la aplicación de estos métodos en imágenes reales.
- En el capítulo 7 se presentan resultados de convergencia del estimador propuesto.
- En el capítulo 8 se presentan las conclusiones finales de este trabajo.

Capítulo 3

Radar de Apertura Sintética

Los sistemas de Radar (radio detection and ranging: detección y medición de distancias por radio) son instrumentos que, a través de ondas electromagnéticas, detectan un objeto e indican su distancia y posición. Estos instrumentos miden la respuesta del blanco a la radiación electromagnética emitida en forma de pulsos, el valor de esta respuesta es almacenada para su posterior procesamiento y se utiliza para formar una imagen de la zona de interés. A diferencia de los sensores ópticos e infrarrojos que son inherentemente pasivos, el radar es un sensor activo, proporciona su propia iluminación en forma de microondas. Las microondas son ondas electromagnéticas (EM) que se encuentran en aproximadamente la zona de 1 a 300 GHz del EM, es decir, longitudes de onda entre 1 centímetros a 1 milímetro.

Muchos de los radares montados sobre plataformas móviles son de vista lateral (SLAR: Side Looking Airbone Radar), dentro de éstos podemos encontrar los radares de apertura real (RAR: Real Aperture Radar) y los radares de apertura sintética (SAR: Synthetic Aperture Radar).

¿Por qué vista lateral? Porque los retornos provenientes de distintos puntos del terreno llegarán en tiempos diferentes, ya que difieren en su distancia al radar. En cambio, en los radares de vista vertical, se pueden recibir señales provenientes de puntos equidistantes. Esto llevaría a ambigüedades ya que el tiempo de arribo

de la señal sería el mismo para este tipo de sensores.

¿Por qué SAR y no RAR? La resolución en la dirección del desplazamiento (Δx) es la mínima distancia que tienen que tener dos puntos diferentes en el terreno para que el sensor sea capaz de distinguirlos. Los sensores RAR tienen una baja resolución espacial, ésta se define como $\Delta x \approx \frac{\lambda \cdot r}{L}$ [87], donde λ es la longitud de onda de la señal emitida, r la altura entre el sensor y el terreno y L la longitud de la antena. Esto se muestra en la figura 3.1.

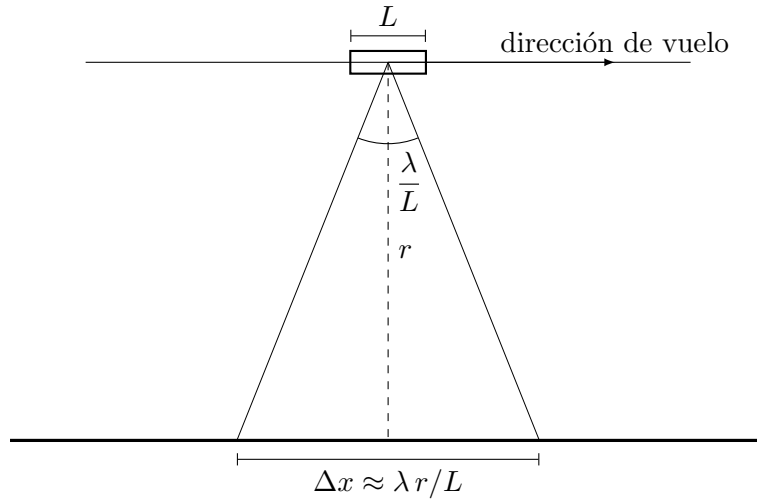


Figura 3.1: Resolución de un radar.

De esta manera para disminuir el valor de Δx , es decir, aumentar la resolución es necesario disminuir r , λ o aumentar L . Los sensores ópticos son los que emiten radiación electromagnética con una longitud de onda menor a la correspondiente a microondas. La ventaja de esto es aumentar la resolución a expensas de las condiciones climáticas entre otros factores.

La tecnología SAR permite simular una antena más larga aprovechando que estos sensores son aerotransportados. Cuando un blanco se mueve con respecto a la fuente de emisión de ondas se observa un desplazamiento en frecuencia entre la señal emitida y la recibida. Esta diferencia en frecuencia se conoce como efecto Doppler y afecta la frecuencia observada cuando hay un movimiento relativo entre el objeto y el sensor. La tecnología SAR hace uso del efecto Doppler para

generar una apertura virtual más grande que la real, por eso se llama de apertura sintética.

Este tipo de imágenes son de gran utilidad ya que permiten obtener información sobre recursos naturales, como así también, permiten detectar efectos de la acción del hombre tales como deforestación, cultivos y embalses. Además presentan ventajas y desventajas sobre las imágenes obtenidas por sensores ópticos, entre las que se encuentran

Ventajas:

- El radar posee un sistema de iluminación propio que permite la adquisición de imágenes tanto de día como de noche.
- Los sistemas SAR emiten radiación electromagnética a frecuencias que permiten atravesar las nubes sin pérdida en la calidad de la imagen obtenida, como así también atravesar zonas de forestación obteniendo mayor información de la zona bajo estudio.

Desventajas:

- Las imágenes SAR tienen la desventaja de poseer un ruido que es inherente al proceso de captura de la imagen, ya que el tipo de iluminación que se utiliza para formar la misma es de tipo coherente [42]. Este ruido, llamado speckle, es multiplicativo, no gaussiano y, por lo tanto, diferente al ruido que se observa en imágenes ópticas. La presencia de este ruido hace que en este tipo de imágenes se observen, en algunos casos, un granulado o falta de contraste.

En este capítulo presentamos una breve descripción del funcionamiento de un radar de apertura sintética y de la forma en que se adquieren este tipo de imágenes, ya que éstas son el objeto de estudio en esta tesis.

3.1. Sensores Remotos SAR

Acá poner algo de historia de los sensores, e ir directamente a los más importantes actualmente. Hay que dar detalles de aquellos de los cuales usamos imágenes.

Un radar de apertura sintética SAR es un sensor activo que ilumina la escena emitiendo señales en la región de microondas del espectro electromagnético.

Antes del desarrollo de las imágenes de radar, las imágenes de alta resolución provenían de sensores pasivos que son sensibles tanto a la radiación solar como a la térmica emitida por la superficie de la tierra. Por lo tanto estos sensores necesitan de una fuente de iluminación externa para poder captar el objetivo de interés.

Los sensores SAR poseen una tecnología diferente para obtener información de la superficie terrestre. Estos sensores son activos, es decir, poseen un sistema de iluminación propia que les permite obtener imágenes en forma independiente de la luz solar. Esto aumenta la capacidad de obtener imágenes ya que pueden operar en forma continua durante el día y la noche.

Dado que estos sistemas emiten pulsos en el espectro de las microondas, ni las nubes, ni la niebla, ni las precipitaciones impiden la adquisición de imágenes. Esto hace que los sensores SAR puedan operar en forma independiente de las condiciones climáticas. Por lo tanto estos instrumentos son capaces de obtener información de la tierra en forma continua lo que permite observar fenómenos como corrientes oceánicas, detectar cambios en la vegetación o movimientos del hielo.

De acuerdo a [68], durante los años 50 y 60 la investigación militar utilizó ampliamente los sistemas SAR con el propósito del reconocimiento del terreno. Sin embargo, a partir de los años 70 y 80 se desarrollaron varios sensores SAR aerotransportados para aplicaciones civiles con el objetivo final de recuperar parámetros geo/biofísicos y detectar cambios en la superficie terrestre.

El primer satélite civil que tuvo a bordo un radar SAR fue lanzado en 1978 [61] a fines de junio por el Jet Propulsion Laboratory (JPL) de la NASA. Este satélite fue diseñado para teledetección de los océanos de la tierra, registrando información de los vientos de la superficie marina y la temperatura, altura de las olas, características hielo marino entre otros parámetros de interés. Debido a un problema eléctrico, este satélite finalizó sus operaciones a mediados de octubre del mismo año.

La Comisión Nacional de Actividades Espaciales (CONAE) y el INVAP de Argentina, junto con la Agencia Espacial Italiana (ASI) integran el Sistema Ítalo Argentino de Satélites para la Gestión de Emergencias (SIASGE) [50]. De acuerdo a [51] el objetivo de la misión SAOCOM es la puesta en órbita de dos constelaciones: SAOCOM1 y SAOCOM2. Cada constelación se compone de dos satélites llamados A y B. El satélite SAOCOM1A fue puesto en órbita el 7 de octubre de 2018 en Vandenberg, California. En [51] se puede ver un video de su lanzamiento.

Entre los objetivos de estos satélites se encuentra la medición de la humedad del suelo, detectar derrames de hidrocarburos en el mar y el seguimiento de la cobertura de agua en las inundaciones. Están equipados con tecnología SAR polarimétrica y operan en banda L. La banda L permite penetrar a través de la superficie hasta 2 m de profundidad dependiendo del tipo de suelo.

En la tabla 3.1, y de acuerdo a la información obtenida de [68], se muestran los principales sensores SAR junto con el período, banda y frecuencia de operación y las instituciones/países que lo desarrollaron.

Sensor	Operación	Banda y Polarización	Institución/País
Seasat	1978	L (HH)	NASA/JPL, USA
ERS -1/2	1991–2000/1995–2011	C (VV)	ESA, Europe
J-ERS -1	1992–1998	L (HH)	JAXA, Japan
Radarsat-1	1995–2013	C (HH)	CSA, Canada
SRTM	Lanzado en 2000	C (HH+VV) y X (VV)	NASA/JPL, USA, DLR, Germany, ASI, Italy
ENVISAT / ASAR	2002–2012	C (dual)	ESA, Europe
ALOS /PalSAR	2006–2011	L (quad)	JAXA, Japan
TerraSAR -X/ TanDEM -X	Lanzado en 2007 Lanzado en 2010	X (quad)	SDLR/Astrium,Germany
Radarsat-2	Lanzado en 2007	C (quad)	CSA, Canada
COSMO -SkyMed-1/4	Lanzado en 2007	X (dual)	ASI/MiD, Italy
RISAT -1	Lanzado en 2012	C (quad)	ISRO, India
HJ-1C	Lanzado en 2012	S (VV)	CRESDA/CAST / NRSCC, China
Kompsat-5	Lanzado en 2013	X (dual)	KARI, Korea
PAZ	Lanzado en 2018	X (quad)	CDTI, Spain
ALOS -2	Lanzado en 2014	L (quad)	JAXA, Japan
Sentinel-1a/1b	Lanzado en 2014	C (dual)	ESA , Europe
Radarsat	Lanzado en 2019	C (quad)	CSA, Canada
Constellation-1/2/3	Lanzado en 2018	L (quad)	CONAE, Argentina

Tabla 3.1: Descripción de algunos sensores SAR junto con sus características. Adaptado de Moreira et al. [68].

3.2. Sistema de iluminación coherente

Un haz de luz esta formado por ondas electromagnéticas cuya longitud de onda está comprendida entre 400 y 800 nanómetros.

La superposición de dos o más ondas producen diferentes tipo de sistemas de iluminación, esto se muestra la figura 3.2.

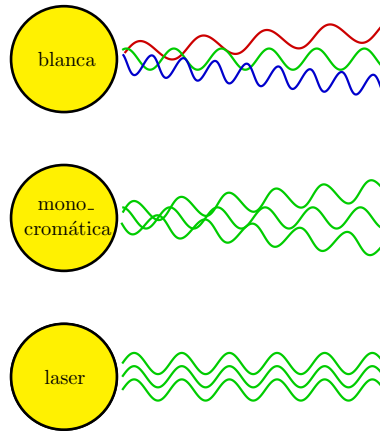


Figura 3.2: Diferentes tipos de sistemas de iluminación.

Cuando se suman dos o más señales se superponen de manera constructiva o destructiva, produciendo de esta forma intensidades máximas o mínimas. Una luz con estas características se denomina incoherente como lo es la luz blanca.

En cambio, con una fuente de luz coherente en frecuencia, todas las ondas emitidas tienen la misma longitud de onda. A este tipo de señal se llama luz monocromática. Si además estas señales tienen una diferencia de fase constante en el tiempo serán coherentes en fase.

Las señales emitidas por un laser o por un radar de apertura sintética son coherentes en fase, en frecuencia, y además los máximos de la señal se producen al mismo tiempo. Esto se conoce como sistema de iluminación coherente.

3.3. Espectro Electromagnético

Las ondas electromagnéticas se modelan con una sinusoidal donde la longitud de onda λ es la distancia entre los máximos de la senoide. La frecuencia f de la onda es el número de máximos por unidad de tiempo. Estas medidas están relacionadas por la ecuación $c = \lambda \cdot f$ donde c es la velocidad de la luz. De esta ecuación vemos que λ y f están inversamente relacionadas dado que c es un valor constante.

Como dijimos anteriormente el radar SAR emite señales en la región de micro-

ondas del espectro electromagnético (EM) y en diferentes bandas. La figura 3.3 muestra cómo está formado el EM de acuerdo a la longitud de onda de la señal.

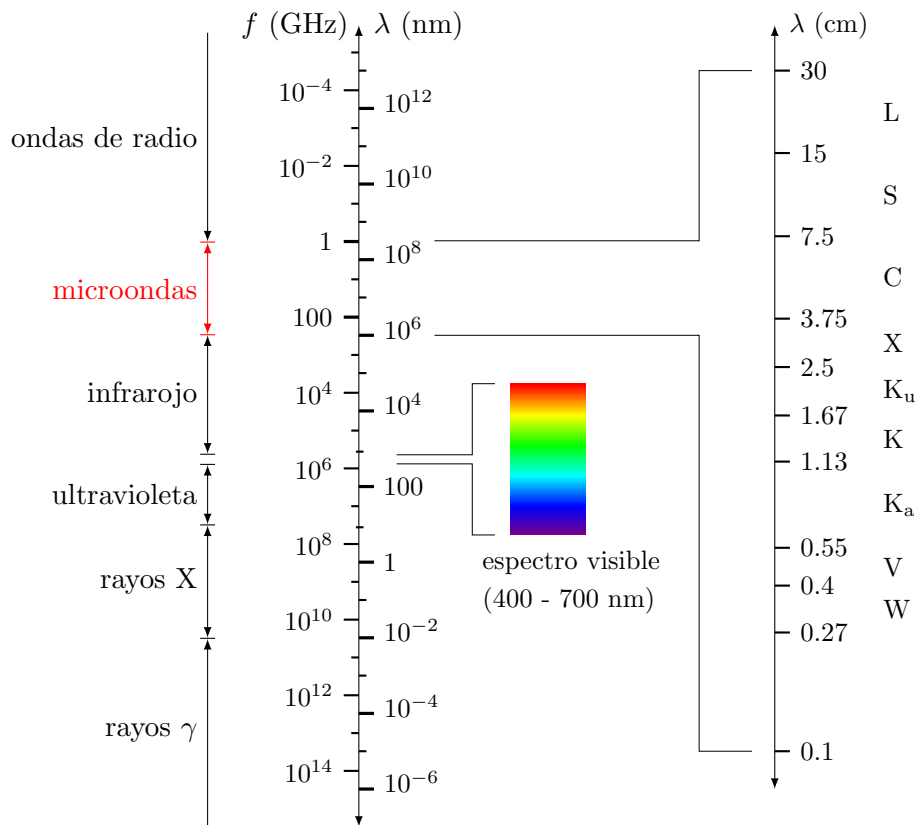


Figura 3.3: Espectro electromagnético.

Las microondas, que son parte del espectro electromagnético, tienen longitudes de onda considerablemente más grande que la luz visible. Un aspecto a tener en cuenta al momento de elegir una longitud de onda es la penetración. Cuanto más grande es la longitud de onda (frecuencias más chicas), más fuerte es la penetración en la vegetación y el suelo. La tabla 3.2 se presentan las bandas más usadas en teledetección por microondas.

Banda	Frecuencia	Longitud de onda
X-band	12.5 a 8 GHz	2.4 a 3.75 cm
C-band	8 a 4 GHz	3.75 a 7.5 cm
S-band	4 a 2 GHz	7.5 a 15 cm
L-band	2 a 1 GHz	15 a 30 cm
P-band	0.999 a 0.2998 GHz	30 a 100 cm

Tabla 3.2: Bandas SAR, datos adaptados de [87]

Cada banda es utilizada en diferente tipos de aplicaciones:

X-band: se utiliza para tener imágenes de alta resolución.

C-band: atraviesa las nubes y la lluvia.

S-band: útil para medir niveles de precipitación.

L-band: útil para aplicaciones en agricultura y mediciones de humedad del suelo.

P-band: adecuada para imágenes con importante penetración en la vegetación.

3.4. Polarización

Las ondas electromagnéticas son ondas transversales, es decir, la dirección del vector campo eléctrico \vec{E} es perpendicular a la dirección de propagación de la onda. Un aspecto a tener en cuenta es la polarización de una onda electromagnética, que se define como la orientación del vector campo eléctrico \vec{E} de la radiación transmitida o recibida. La luz correspondiente al espectro visible por lo general no está polarizada, todas las direcciones de \vec{E} son igualmente probables en el plano perpendicular a la dirección de propagación de la onda. Si la señal se descompone en dos ondas de igual amplitud pero con una diferencia de fase de 90° , entonces se dice que la luz está polarizada circularmente porque \vec{E} gira describiendo un

círculo en el plano ortogonal a la dirección de propagación, a medida que la señal avanza. Si las dos ondas tienen diferente amplitud y están desfasadas entre sí 90° , o si el desfase es distinto de 90° , la luz se dice que está polarizada elípticamente. Cuando el vector \vec{E} se mantiene paralelo a una dirección fija la polarización se llama lineal. La figura 3.4 muestra luz circular o linealmente polarizada.

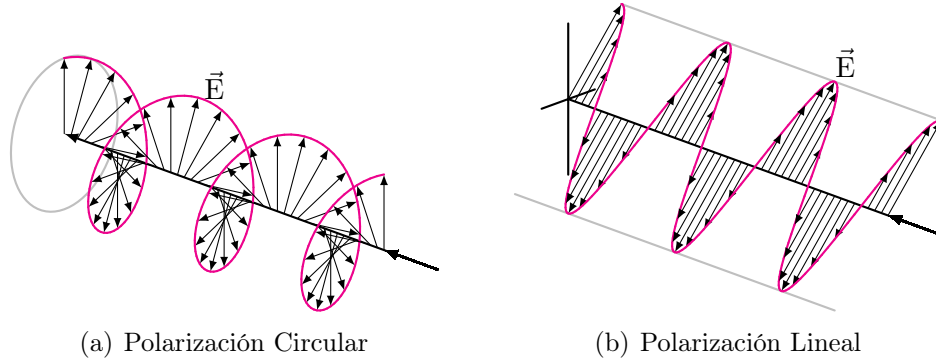


Figura 3.4: Diferentes tipos de polarizaciones.

Los radares de teledetección generalmente están diseñados para transmitir radiación vertical u horizontalmente polarizada.

La antena del radar puede recibir radiación polarizada vertical u horizontalmente y, a veces, ambas. Los planos de polarización transmitida y recibida se designan con las letras H para horizontal y V para vertical. Por lo tanto, la polarización de una imagen de radar puede ser HH, para transmisión horizontal, recepción horizontal, VV para transmisión vertical, recepción vertical, HV para transmisión horizontal, recepción vertical y viceversa (VH).

En el caso donde la polarización de la radiación recibida coincide con la polarización de la radiación transmitida, se conoce como imagen polarizada de manera similar. En cambio, cuando la polarización de la radiación recibida es ortogonal a la polarización de la radiación transmitida, se dice que la polarización es cruzada.

Los sistemas SAR monopolarizados operan con una única polarización de emisión y detecta una sola de las componentes de la radiación recibida. Generalmente usan polarización similar, es decir igual polarización al emitir y recibir, porque las señales de polarización cruzada son demasiado débiles para producir una buena

imagen. De esta forma podemos tener radares monopolares o monopolarimétricos HH, VV, HV o VH donde la primera letra indica la polarización de la radiación emitida y la segunda indica la componente de la radiación detectada. Los primeros sensores SAR, como por ejemplo, RADARSAT-1, que ha sido sustituido por RADARSAT-2, adquieren imágenes únicamente en la banda C y en la combinación HH.

Si un radar tiene la capacidad de captar distintas polarizaciones se lo identificará como polarimétrico. En este tipo de radares se emite una señal electromagnética con polarización horizontal y vertical y se detectan ambas polarizaciones en la radiación recibida. Es decir, operan en las cuatro combinaciones citadas al mismo tiempo generando, en principio, imágenes de cuatro bandas. En esta tesis trabajaremos con imágenes monopolarimétricas.

3.5. Geometría SAR

El radar de apertura sintética SAR es un sensor activo, de vista lateral, que emite energía en un período corto de tiempo y en el intervalo de frecuencias de microondas recibiendo la señal reflejada por el blanco iluminado. Está montado sobre una plataforma móvil (aeronave o satélite) que transmite y recibe las señales con la misma antena, la cual tiene su eje longitudinal paralelo a la trayectoria de vuelo. Debido a la gran velocidad del desplazamiento de la plataforma y aprovechando el efecto Doppler que tiene lugar como consecuencia de dicho desplazamiento, la antena del dispositivo SAR se convierte en una antena virtual de mayor tamaño.

El radar se desplace a una velocidad \vec{v} y a una altura h portando una antena que emite radiación electromagnética e ilumina la superficie del terreno. En la figura 3.5 se muestra un esquema del movimiento del radar, la región sombreada es la zona iluminada por el radar. Esta zona permanece en el haz de la antena durante unos instantes y es observado por el radar desde numerosos puntos a lo

largo de la trayectoria de la plataforma.

Vamos a dar algunas definiciones sobre la geometría SAR:

- Azimut: es la dirección paralela a la trayectoria de vuelo de la plataforma.
- Rango: es la dirección perpendicular a la trayectoria de vuelo de la aeronave, es decir, perpendicular al azimut.
- Nadir: es el punto de la superficie terrestre justo debajo del rada.
- Swath: es el ancho de la zona iluminada.
- Slant Range: es la distancia oblicua entre el objetivo y la antena del radar.

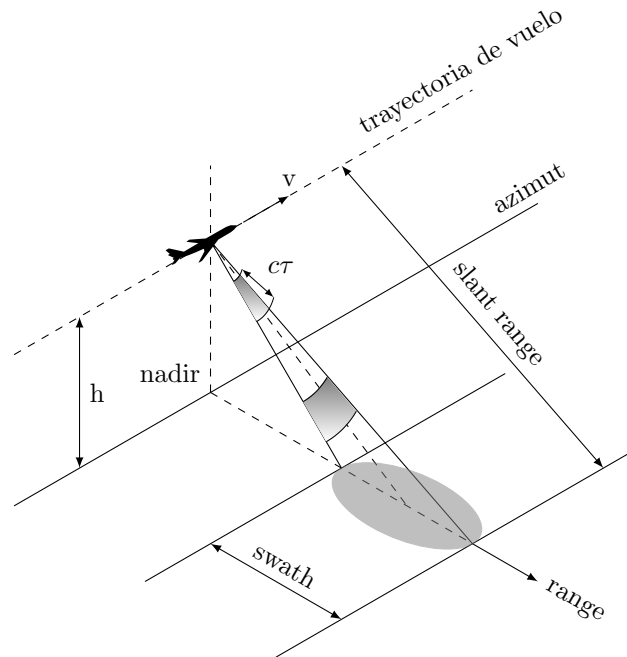


Figura 3.5: Geometría del Radar de Apertura Sintética SAR.

La antena del radar tiene su eje longitudinal paralelo a la trayectoria de vuelo y emite pulsos de energía electromagnética perpendicular a dicha trayectoria, como se ilustra en la figura 3.5. Los pulsos electromagnéticos iluminan un área en el terreno, la antena recibe la energía reflejada en diferentes momentos, dependiendo de la distancia que existe entre el objeto y la antena. Con el procesamiento de

la señal reflejada se genera la imagen. A medida que el radar se desplaza emite pulsos de radiación electromagnética y recibe la señal correspondiente a la energía reflejada (retrodispersada) por la superficie sensada. En el momento en el que este pulso incide sobre el blanco la radiación emitida sufre un proceso de dispersión como se muestra en la figura 3.6 dependiendo del tipo de textura que tiene el blanco iluminado.

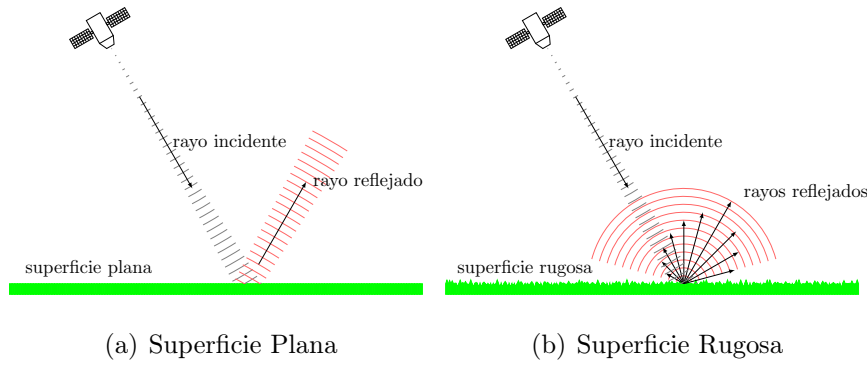


Figura 3.6: Dispersión de la señal emitida por un radar SAR.

3.5.1. Resolución

La resolución de un radar es la capacidad que tiene para distinguir la posición entre objetos que están muy cercanos en rango o en azimut. Entonces, la resolución se divide en dos categorías: resolución en rango y resolución en azimut.

El grado de resolución en rango depende de la longitud temporal del pulso transmitido. Si el radar emite pulsos de longitud temporal τ la resolución en slant range queda determinada por $\Delta sl = \frac{c\tau}{2}$ donde c es la velocidad de la luz [68, 87]. Entonces, señales con pulsos de longitud temporal pequeña mejoran la resolución en rango.

En los radares de apertura real RAR la resolución en azimut se define como $\Delta x_{\text{rar}} = R \frac{\lambda}{L}$ donde R es la distancia del objeto al radar, L es la longitud de la antena y λ es el ancho de banda del pulso emitido por el sensor. Esto tiene el inconveniente de que la resolución depende de la distancia al sensor. Cuanto

mayor es esta distancia mayor es el valor de Δx_{rar} y, por lo tanto, menor la resolución. Con el fin de mejorar dicha resolución se debería disminuir el ancho de banda λ o aumentar la longitud de la antena L . Si se reduce λ dejamos de estar en el espectro de las microondas. Por otro lado la antena no puede ser excesivamente grande por razones de ingeniería estructural.

De acuerdo a [68] este problema se superó al usar un radar coherente junto con el efecto Doppler. Cuando un emisor de pulsos se mueve respecto de un objetivo en la tierra se observa lo que se conoce como efecto Doppler. Cuando dos puntos sobre el terreno están levemente separados en la dirección azimut, tienen ángulos ligeramente diferentes de la antena con respecto a la línea de vuelo. Debido a esto, las velocidades con la que los objetivos se aproximan al sensor son ligeramente diferentes. Por lo tanto la señal recibida de cada punto tendrá su frecuencia desplazada una cantidad diferente de la original.

La idea básica del SAR es iluminar a un punto de la tierra no sólo con un único pulso sino con una secuencia de pulsos. En la figura 3.7 se esquematiza la forma de generar una apertura sintética. Cuando el objetivo A es iluminado por el radar es cuando se comienza a generar dicha apertura. A medida que la plataforma continúa moviéndose hacia adelante, se registra la retrodispersión que generó cada pulso emitido por el radar, durante todo el tiempo que el objetivo estuvo iluminado. El punto en el que el objetivo deja de estar iluminado por el radar, determina la longitud B de la antena simulada.

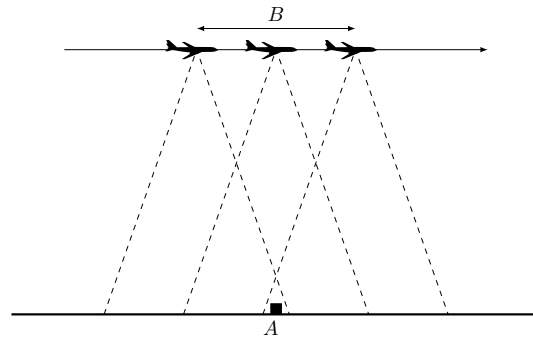


Figura 3.7: Apertura Sintética en un SAR.

Esta idea de generar una apertura sintética llevó a mejorar la resolución en azimuth. La resolución resultante en azimuth se vuelve igual a la mitad de la longitud de la antena, es decir, $\Delta x_{\text{sar}} = \frac{L}{2}$. Observemos que esta resolución es independiente de la longitud de onda y de la distancia al sensor. Además, si se achica la longitud de la antena mejora la resolución.

3.5.2. Ecuación del radar

La ecuación del radar representa la relación que existe entre la energía transmitida y la energía recibida por la antena del radar. Esta relación se define como

$$P_r = \frac{P_e G^2 \lambda^2 \sigma}{(4\pi)^3 R^4} \quad (3.5.1)$$

donde

- P_r =potencia recibida.
- P_e =potencia emitida.
- G =ganancia de la antena (constante adimensional que indica la relación que existe entre la potencia radiada en una dirección y a una distancia, y la potencia que radiaría una antena isotrópica con la misma potencia emitida y a la misma distancia. Una antena isotrópica es una antena ideal que emite la misma intensidad de radiación en todas las direcciones del espacio. λ =longitud de onda de la señal.
- R =distancia entre el sensor y el terreno.
- σ = coeficiente de retrodispersión.

Entre los parámetros que intervienen en la ecuación del radar se encuentra el coeficiente de retrodispersión σ . Este parámetro es de interés porque su valor depende de la textura del terreno el cual influye en la intensidad de la señal del retorno.

3.6. Formación del ruido Speckle

El láser HeNe es un tipo de fuente que utiliza como medio activo una mezcla gaseosa de helio y neón. De acuerdo a [41, 80] Ali Javan, un físico que trabajaba para Laboratorios Bell Telephone, fue el responsable del funcionamiento del primer láser HeNe en 1960. Este láser mostró un fenómeno inesperado: los objetos vistos en luz altamente coherente adquieren una apariencia granular peculiar que no tiene una relación obvia con las propiedades macroscópicas del objeto iluminado, sino que presentan un patrón irregular.

Esta apariencia granular se conoce como *ruido speckle*. Las imágenes adquiridas por sistemas de iluminación coherente como son las imágenes SAR, las que utilizan laser o ultrasonido entre otros, se ven afectadas por la presencia del ruido speckle. Este tipo de ruido es inherente al proceso de captura de la imagen, no es aditivo ni gaussiano y suele dificultar el análisis e interpretación de las imágenes porque pueden degradarlas significativamente.

En la teledetección por microondas la celda de resolución es de unos metros mientras que la longitud de onda es alrededor de centímetros (banda X o C) o decenas de centímetros (banda S, L y P). Dentro de una celda de resolución existen elementos que dispersan la señal que incide en la celda.

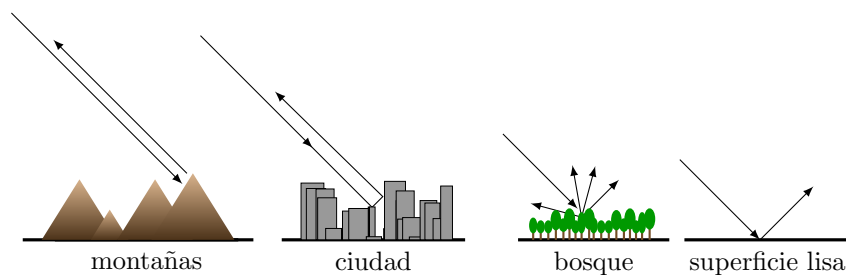


Figura 3.8: Radiación reflejada por distintos objetos en la superficie terrestre.

Cuando el radar ilumina una superficie la señal que retorna está compuesta por la suma de los pulsos reflejados por cada uno de los dispersores elementales que se encuentran dentro de la celda, esto se muestra en la figura 3.8. El tipo de

terreno dependerá de la forma en que esta señal se refleje.

Estos dispersores elementales se distribuyen en forma aleatoria y, por lo tanto, la distancia entre ellos y al radar también es aleatoria [62]. En el momento de la reflexión, la señal que retorna ya no es coherentes en fase pero si lo es en frecuencia. Esto se debe principalmente a la textura de la superficie. La figura 3.9 muestra las señales incidentes y reflejadas en una superficie rugosa. La línea sólida negra corresponde a la primera y la línea punteada roja corresponde a la segunda. La distancia d da información sobre la rugosidad del objeto, la figura 3.9(b) representa una superficie más rugosa que 3.9(a). Se puede observar que las ondas recibidas son coherentes en frecuencia, pero no necesariamente en fase. Cuanto mayor es el tamaño de d mayor será la diferencia de fase de las ondas reflejadas por los dispersores.

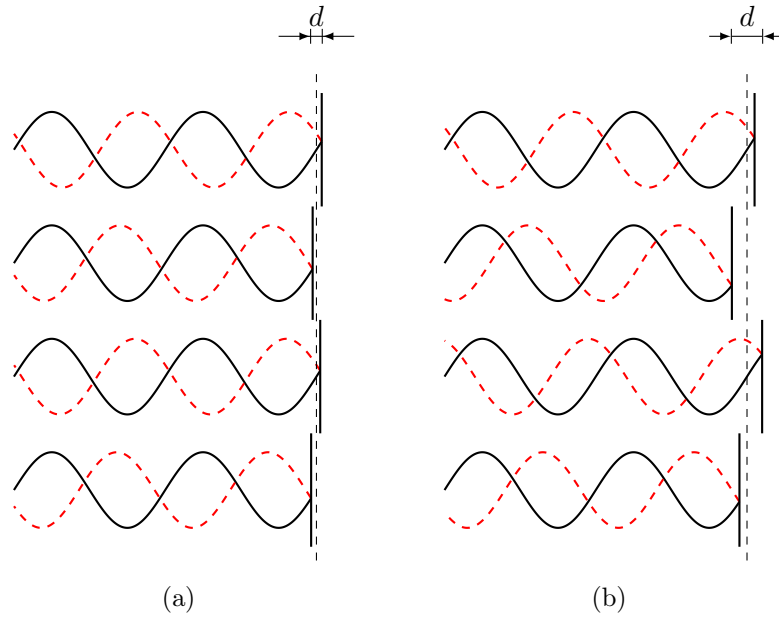


Figura 3.9: Coherencia de la señal reflejada.

Entonces, las ondas electromagnéticas que provienen de la misma celda de resolución se suman de manera coherente y dan como resultado fuertes fluctuaciones en la retrodispersión de una celda de resolución a otra [73]. Por lo tanto el valor de gris en un pixel se obtiene como superposición coherente de las seña-

les incidentes que, al estar desfasadas, pueden interferir de manera constructiva o destructiva, lo que produce que las señales recibidas sean más fuertes o más débiles [45]. Estas interferencias causan variaciones en la magnitud de la señal y cambios aleatorios de intensidad de píxel a píxel. Estos cambios en la intensidad se traducen en la imagen como diferentes niveles de gris. Esto explica el patrón granular que se observa en las imágenes SAR y que produce negros, correspondientes a interferencias destructivas, y puntos brillantes debido a las constructivas [102]. Este patrón angular es lo que se conoce como *Ruido Speckle*.

La presencia de este ruido dificulta la interpretación de las imágenes y puede conducir a interpretaciones y detecciones de bordes erróneas entre otros inconvenientes, debido a la degradación de la calidad de los píxeles.

Un enfoque común para la presencia de este ruido durante el procesamiento de la señal cruda captada por el radar es promediar varias estimaciones independientes de la señal reflejada. Dado que el radar SAR genera una apertura sintética la longitud de esta antena virtual se puede subdividir en N segmentos independientes, conocidos también como *looks* [62]. Cada uno de estos segmentos registra la energía retrodispersada por el mismo blanco y se procesa en forma independientemente para formar una imagen SAR. Estas N imágenes se promedian entre sí para formar una sola imagen SAR de N *looks*. Este procesamiento reduce el efecto del ruido speckle pero, al reducir la longitud de la antena en segmentos, se pierde resolución en azimut. La utilidad de este procedimiento se manifiesta en la mejora de la relación señal-ruido. Cuantas más subdivisiones se hagan, es decir, cuanto más *looks* formen la imagen menos resolución tendrá pero también será menos ruidosa.

3.7. Doble Bounce o Doble Rebote

El efecto Doble Rebote es una característica importante que tienen las imágenes de alta resolución como lo son las imágenes SAR. Este efecto resulta en

un fuerte mecanismo de dispersión de la señal emitida por el radar. Aparece por ejemplo, cuando la señal emitida rebota en la calle, luego en los edificios (llamado un doble rebote) como se ve en la figura 3.10(b), o cuando rebota en un edificio y en sus alrededores (corner reflector). Este efecto se observa en las imágenes SAR como puntos muy brillantes, en cambio, las autopistas aparecen oscuras al ser superficies planas. En la figura 3.10 se puede ver un esquema de diferentes reflexiones de una señal emitida por el radar.

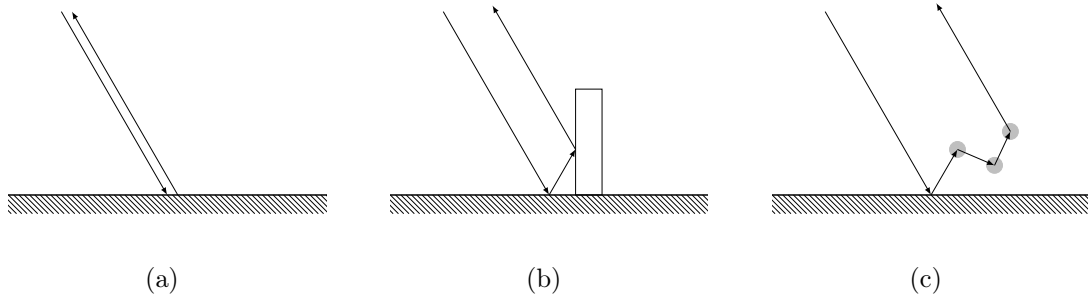


Figura 3.10: Diferentes reflexiones de una señal emitida.

Debido a la aleatoriedad y a la fuerte dispersión que puede tener la señal retrodispersada es necesario contar con modelos estadísticos para poder obtener información de las características del suelo.

3.8. Retorno Complejo

Supongamos que iluminamos un área a través de un haz de energía electromagnética con n dispersores elementales como se mostró en la figura 3.6. Como presentan Frery y Wu [35], cada dispersor j retornará una fracción de la energía incidente A_j con fase ϕ_j . El retorno total S es una señal compleja definida por:

$$S = \sum_{j=1}^N A_j e^{i\phi_j} = \underbrace{\sum_{j=1}^N A_j \cos \phi_j}_{\text{Re}(Z)} + i \underbrace{\sum_{j=1}^N A_j \sin \phi_j}_{\text{Im}(Z)}, \quad (3.8.1)$$

donde i es la unidad imaginaria y $A_j e^{i\phi_j}$ representa el retorno del j -ésimo dispersor. La ecuación (3.8.1) describe la suma coherente de las contribuciones del retorno total de cada uno de los dispersores. En general, en vez de tratar directamente con el retorno complejo se trabajan con los datos dados en formato amplitud $A = |S|$ o en formato intensidad $I = |S|^2$, donde $I = A^2$. En esta tesis se trabajará con este último formato.

De acuerdo a [79] si no hay un retrodispersor que domine, la forma de extraer información del retorno de los retrodispersores es aproximándose al problema desde un punto de vista estadístico. Este enfoque se verá en el capítulo 4.

3.9. Conclusiones

En este capítulo presentamos los conceptos básicos de teledetección, iluminación coherente y polarización. Estos conceptos son de utilidad porque son el pilar de la formación de imágenes SAR. Se muestran las ventajas y desventajas que poseen estos sistemas frente a aquellos que adquieren imágenes ópticas. Entre las ventajas se encuentran la capacidad que tienen los sensores SAR de adquirir imágenes independientemente de las condiciones climáticas y de la luz, ya que ellos poseen un sistema de iluminación propia. Entre las desventajas se encuentra el ruido speckle que está presente en este tipo de imágenes porque es propio a la forma de adquirir las mismas.

Asimismo se muestra que el aprovechamiento del efecto Doppler contribuye a generar la apertura sintética del radar SAR. Las bondades de esta apertura sintética se pone de manifiesto en la obtención de imágenes con mejor resolución.

Finalmente se presenta la aleatoriedad de la formación del ruido speckle, lo que pone de manifiesto la necesidad de encontrar modelos estadísticos que permitan analizar e interpretar este tipo de imágenes.

Capítulo 4

Modelo \mathcal{G}_0

Las imágenes adquiridas por radares de apertura sintética SAR han encontrado muchas aplicaciones debido a sus ventajas respecto de las imágenes que provienen de sensores ópticos. La independencia de las condiciones climáticas y de la iluminación junto con la capacidad que tienen estos sensores de penetrar la nubes, las copas de los árboles e incluso el suelo, han contribuido para que este tipo de imágenes resulten cada vez más populares.

Sin embargo, las imágenes SAR tienen la desventaja de ser difíciles de interpretar y procesar debido a la presencia del ruido speckle. Como hemos visto en capítulos anteriores, este ruido es inherente al proceso de captura de la imagen y se hace presente debido a la superposición coherente de las ondas reflejadas por muchos dispersores. Esto causa una variación en la intensidad pixel a pixel que se manifiesta en la imagen como un patrón granular.

Debido a la aleatoriedad y a la fuerte dispersión que puede tener la señal retrodispersada, es necesario contar con modelos estadísticos que contribuyan a una mejor extracción de información a través del desarrollo de filtros, detección de bordes, clasificación de áreas entre otras metodologías.

El modelo multiplicativo es muy apropiado para explicar las características estadísticas de la imagen de un objeto cuando es iluminado por radiación coherente, como lo son las imágenes SAR. Este modelo considera que el valor observado

en cada celda de la imagen es una variable aleatoria Z que resulta del producto de dos variables aleatorias independientes: una correspondiente al backscatter o retrodispersión X (que es lo que observaríamos sin la presencia del ruido speckle) y la otra correspondiente al ruido speckle Y (que es inherente a todo sistema de captura de imágenes con iluminación coherente).

Se han utilizado algunas distribuciones estadísticas para modelar los datos SAR, entre ellas se encuentra la distribución Lognormal, la Weibull y la distribución K [62].

Un modelo común para datos que provienen de zonas homogéneas y con mucha presencia de ruido speckle son la distribución exponencial y la Rayleigh [36]. En cambio las distribuciones K [53] han sido muy utilizada en datos SAR especialmente cuando los datos provienen de áreas texturadas.

Frery et al. [36, 37] presentan la familia de distribuciones \mathcal{G} para datos dados tanto en formato amplitud como en intensidad. Es un modelo más general porque incluye a las distribuciones clásicas, como K, ya que son casos particulares de esta nueva familia de distribuciones.

Como caso particular de esta clase de distribuciones está la familia \mathcal{G}^0 que tiene tantos parámetros como la distribución K y es capaz de modelar zonas extremadamente rugosas, como las zonas urbanas, mejor que esta última distribución [67].

La familia de distribuciones \mathcal{G}^0 están indexadas por tres parámetros: α que está relacionado con la textura del terreno, el parámetro γ considerado un parámetro de escala e informa sobre el brillo de la imagen y el parámetro L que es el número de looks. Este último parámetro da información sobre la relación señal-ruido de la imagen, cuanto mayor es el número de looks menor presencia de ruido speckle se encuentra en la imagen, a expensas de una pérdida en la resolución.

En este capítulo se presenta el modelo multiplicativo junto con la distribución \mathcal{G}^0 y la relación con otras distribuciones. Asimismo se presentan resultados teóricos producto de esta tesis, donde se muestra la capacidad que tiene esta dis-

tribución de generar valores atípicos, y resultados de convergencia uniforme de la distribución \mathcal{G}^0 cuando el parámetro de textura $\alpha \rightarrow -1$ y cuando $\alpha \rightarrow -\infty$.

4.0.1. Modelo Multiplicativo

Una aproximación para comprender la distribución de los datos SAR es proponer modelos que ajusten razonablemente bien a ellos. Algunas distribuciones conocidas como la Exponential, Gamma [35], Lognormal [73, 103], Rayleigh, Weibull [62, 103] entre otras, fueron propuestas como "modelos empíricos" para describir a los datos SAR. La pertinencia de estas distribuciones depende del formato de los datos (amplitud, intensidad, complejo, etc.), el número de looks y el grado de homogeneidad de la zona bajo estudio.

Lee y Pottier [62] explican la naturaleza multiplicativa del ruido speckle utilizando el modelo Rayleigh para 1-look con datos en formato amplitud en zonas homogéneas. Ellos realizaron un scatter plot del desvío standard muestral respecto de la media muestral en muchas zonas homogéneas en una imagen SAR. Para el caso de n muestras provenientes de una imagen de 1-look utilizan en modelo Rayleigh. En el caso de muestras de 4-looks modelan con la distribución χ con $2n$ grados de libertad.

Si se supone un backscatter constante C en zonas homogéneas, el retorno queda modelado por $Z = C \cdot Y$ bajo el modelo multiplicativo. Entonces, calculando la esperanza y el desvío de Z se obtiene

$$E(Z) = C E(Y) \quad (4.0.1)$$

$$SD(Z) = C \cdot SD(Y) \quad (4.0.2)$$

Despejando C de (4.0.1) y reemplazando en (4.0.2) se obtiene que

$$SD(Z) = \frac{SD(Y)}{E(Y)} E(Z) \quad (4.0.3)$$

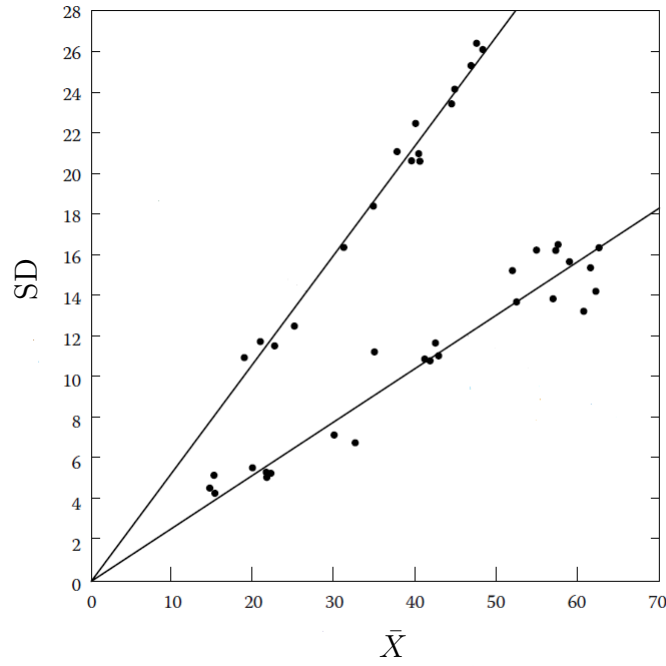


Figura 4.1: Modelo Multiplicativo, fuente Lee et.al [62].

En la figura 4.0.1 tomada de Lee y Pottier [62] se muestra el scatter plot basado en datos tomados por el radar SIR-B (Spaceborne Imaging Radar), un sensor SAR que opera en banda C y X . Se tomaron 16 y 25 muestras de zonas homogéneas para el caso de 1-look y 4-looks respectivamente. En cada muestra se calculó \bar{X} la media muestral y SD el desvío muestral y se grafica SD vs. \bar{X} .

La naturaleza multiplicativa del ruido speckle se manifiesta en la recta que pasa por el origen producto del ajuste lineal que se realizó al conjunto de datos. De acuerdo a Lee y Pottier [62] los valores de las pendientes para 1-look y 4-look son 0.54 y 0.26 respectivamente, que son razonablemente cercanos a los valores teóricos 0.5227 y 0.261 respectivamente.

Por este motivo en esta tesis se utilizará el modelo multiplicativo para modelar datos SAR. Este enfoque se basa en la física de la formación de la imagen donde toma en cuenta la iluminación coherente propia de este tipo de sensores y propone aprovechar la naturaleza multiplicativa del ruido speckle.

El modelo multiplicativo considerado es el siguiente:

$$Z = X \cdot Y$$

donde X e Y corresponden al backscatter o retrodispersión e Y corresponde al ruido speckle, respectivamente.

En esta tesis se consideran imágenes monopolarimétricas, una banda y una polarización en formato intensidad. Daremos modelos para el speckle, la retrodispersión y finalmente para el retorno bajo el modelo multiplicativo.

4.0.2. Modelo para el Ruido Speckle

Como mencionamos en 3.8.1 en esta tesis se modelarán los datos SAR en formato intensidad. De acuerdo a la ecuación (3.8.1) el retorno S es un número complejo. Entonces $S = S_R + i S_{Im}$ donde i es la unidad imaginaria, S_R y S_{Im} son la parte real e imaginaria de S respectivamente. Entonces $Z = S_R^2 + S_{Im}^2$ donde Z es la intensidad del retorno.

Frery y Wu [35] indican que, bajo las condiciones:

- Se observan retrodispersores (backscatters) cuyo tamaño es menor que la longitud de onda.
- Las amplitudes A_j de los dispersores elementales son independientes y positivas.
- Ninguno de estos backscatters domina el retorno.
- Las fases ϕ_i son eventos independientes e igualmente distribuidas con distribución uniforme en $[-\pi, \pi]$.

se puede probar que $S_R \sim N(0, \sigma^2/2)$ y $S_{Im} \sim N(0, \sigma^2/2)$, entonces

$$Z \sim Exp(1/\sigma^2) \equiv \Gamma(1, 1/\sigma^2), \quad (4.0.4)$$

donde σ^2 es la intensidad media del retorno.

Cabe señalar que en esta tesis se van a considerar las siguientes parametrizaciones para las distribuciones $Exp(\lambda)$ y $\Gamma(\alpha, \beta)$. Sean $T \sim Exp(\lambda)$ y $W \sim \Gamma(\alpha, \beta)$ entonces las funciones de densidad f_T y f_W están dadas por

$$f_T(t) = \lambda e^{-\lambda t} \quad (4.0.5)$$

$$f_W(t) = \frac{\beta^\alpha}{\Gamma(\alpha)} t^{\alpha-1} e^{-\beta t} \quad (4.0.6)$$

respectivamente.

Entonces, si modelamos al backscatter X como la intensidad media del retorno σ^2 , el modelo para Z para 1-look será $Z = \sigma^2 Y$ donde $Y \sim \Gamma(1, 1)$ es el modelo para el ruido speckle.

Como se vio en 3.6 una técnica muy utilizada para disminuir el efecto del ruido speckle en la imagen SAR y, por lo tanto, mejorar la relación señal-ruido es el proceso multilook. Como se indicó en 3.6 este proceso consiste en promediar L vistas independientes de retornos de 1-look. Entonces la intensidad del retorno final es el promedio de las intensidades de los retornos Z_l de cada una de estas vistas. Es decir

$$Z = \frac{1}{L} \sum_{l=1}^L Z_l. \quad (4.0.7)$$

Entonces, por 4.0.4 y utilizando propiedades de la distribución $\Gamma(\alpha, \beta)$, Z queda distribuido como $Z \sim \Gamma(L, L/\sigma^2)$. Entonces, si consideramos a la retrodispersión constante, el retorno Z para el caso de L -looks queda modelado como

$$Z = \sigma^2 Y \text{ con } Y \sim \Gamma(L, L). \quad (4.0.8)$$

La pregunta que nos podemos hacer es ¿tiene sentido considerar un retorno constante? Y, obviamente, la respuesta es no. Dependiendo cómo se modele a X surgen distintas familias de distribuciones para modelar el retorno.

Si $X \sim \Gamma(\alpha, \lambda)$ entonces $Z \sim \mathcal{K}(\alpha, \lambda, L)$. Esta familia de distribuciones fueron propuestas por [53], ha sido un modelo ampliamente utilizado pero no modela en forma adecuada datos que provienen de zonas extremadamente heterogéneas como son las áreas urbanas como lo indicaron Frery et al. en [36].

Frery et al [37] propusieron un nuevo modelo para el backscatter, dando origen a una nueva familia de distribuciones que modelan muy bien datos provenientes de zonas extremadamente heterogéneas mejor que la familia \mathcal{K} . Esta familia de distribuciones se llama \mathcal{G} y fueron ampliamente aplicadas tanto para imágenes monopolarimétricas[40, 69], como en imágenes polarimétricas [34, 74].

4.0.3. Modelo para el Backscatter

El grado de homogeneidad de un área de la imagen está modelado por el backscatter. Por este motivo es importante encontrar un modelo para la retrodispersión que sea lo más general posible, para que pueda explicar los diferentes grados de textura que tiene una imagen.

Frery et al. [37] consideraron un modelo general para la retrodispersión que tiene como casos especiales modelos más simples que permiten explicar las diferentes texturas que puede tener un determinado área. En general, estas texturas se indican como: áreas homogéneas (pastura), heterogéneas (bosque) y extremadamente heterogéneas (áreas urbanas). Estos modelos simples tienen la ventaja de tener menos parámetros para estimar.

El modelo que proponen en [37] para el backscatter para datos de intensidad es una distribución Gaussiana Inversa Generalizada con parámetros α , γ y λ . En este caso decimos que el backscatter $X \sim N^{-1}(\alpha, \gamma, \lambda)$ cuya función de densidad viene dada por

$$f_X(x) = \frac{\left(\frac{\lambda}{\gamma}\right)^{\frac{\alpha}{2}}}{2K_\alpha(2\sqrt{\lambda\gamma})} x^{\alpha-1} e^{-\frac{\gamma}{x} - \lambda x} \text{ para } x > 0, \quad (4.0.9)$$

donde K_α indica la función de Bessel modificada de tercera especie y de orden α .

El espacio paramétrico está dado por

$$\begin{cases} \gamma > 0, \lambda \geq 0, & \text{si } \alpha < 0 \\ \gamma > 0, \lambda > 0, & \text{si } \alpha = 0 \\ \gamma \geq 0, \lambda > 0, & \text{si } \alpha > 0. \end{cases}$$

Este modelo general presenta algunas ventajas porque tiene como casos particulares a algunas distribuciones que son muy utilizadas para modelar el retorno para datos en formato intensidad. Estos casos son:

- Una constante β que se utiliza para modelar áreas homogéneas dando lugar a un retorno que es puro speckle. Este caso lo vimos anteriormente y da lugar a un retorno distribuido con una distribución Gamma.
- La distribución Gamma, dando lugar a la familia de distribuciones \mathcal{K} .
- La distribución Recíproca de una Gamma (Γ^{-1}) dando lugar a la familia de distribuciones \mathcal{G} para el retorno, que se verán en la próxima sección.

En la figura 4.2 y de acuerdo a [37] se esquematiza el comportamiento asintótico de la distribución $N^{-1}(\alpha, \gamma, \lambda)$ en algunos casos límites, D indica convergencia en distribución y p convergencia en probabilidad.

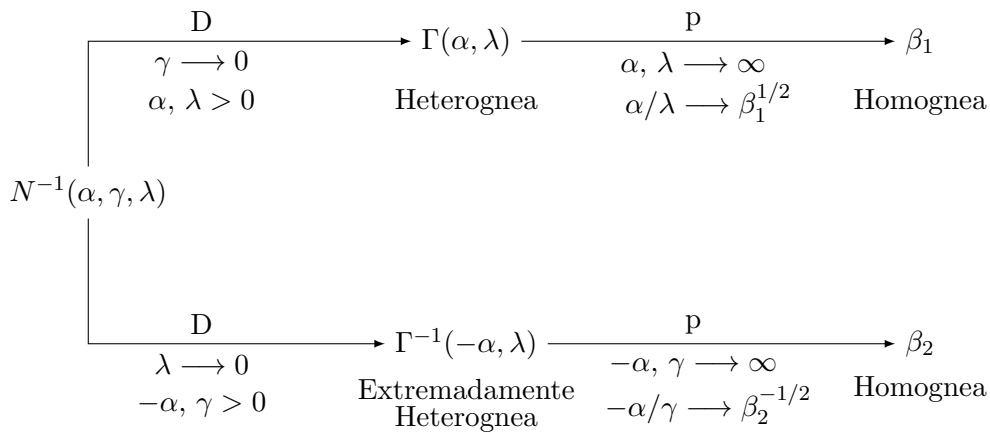


Figura 4.2: Comportamiento asintótico da la distribución $N^{-1}(\alpha, \gamma, \lambda)$.

Este comportamiento asintótico muestra que tanto backscatters homogéneos como heterogéneos y extremadamente heterogéneos pueden tratarse a partir de la distribución $N^{-1}(\alpha, \gamma, \lambda)$.

4.0.4. Modelo para el Retorno

De acuerdo a lo visto en las subsecciones 4.0.2 y 4.0.3 surge un modelo para la intensidad del retorno para datos multilook a partir del modelo multiplicativo 4.0.1

$$Z = X \cdot Y$$

donde X e Y son variables aleatorias que corresponden al backscatter y al ruido speckle, respectivamente.

Si se considera que $X \sim \mathcal{N}^{-1}(\alpha, \lambda, \gamma)$, $Y \sim \Gamma(L, L)$, entonces Z obedece una distribución \mathcal{G}_I [37], que es el modelo \mathcal{G} para datos de intensidad. En [36] se presenta el modelo \mathcal{G} para datos de amplitud a partir de un análisis similar.

Entonces, el retorno $Z \sim \mathcal{G}_I$ si su función de densidad está dada por

$$f_{Z_I}(z) = \frac{z^{L-1} L^L \sqrt{(\frac{\lambda}{\gamma})^\alpha}}{\Gamma(L) \mathcal{K}_\alpha(2\sqrt{\lambda\gamma})} \left(\frac{\gamma + Lz}{\lambda} \right) \mathcal{K}_{\alpha-L}(\sqrt{\lambda(\gamma + Lz)}) \quad (4.0.10)$$

donde $L \geq 1$ es el número de looks y el espacio paramétrico para α , γ y λ es el mismo que en 4.0.3.

Cabe señalar que el número de looks L debería ser, en principio, un número natural por la forma en que fué definido. En general este parámetro es considerado conocido porque viene con la información de la imagen, pero puede ser que no lo sea. En este caso se estima a partir de los datos de la imagen con muestras extraídas de zonas homogéneas y se llama número equivalente de looks (ENL). De acuerdo a [4], $\text{ENL} = 1/\widehat{\text{CV}}^2$ donde $\widehat{\text{CV}} = \widehat{\sigma}/\widehat{\mu}$, con $\widehat{\sigma}$ es el desvío estandar muestral y $\widehat{\mu}$ la media muestral.

Haciendo un análisis similiar al que se presentó en la figura 4.2 y que se muestra

en la figura 4.3 se obtienen casos particulares de la distribución \mathcal{G}_I , dependiendo del comportamiento asintótico de sus parámetros.

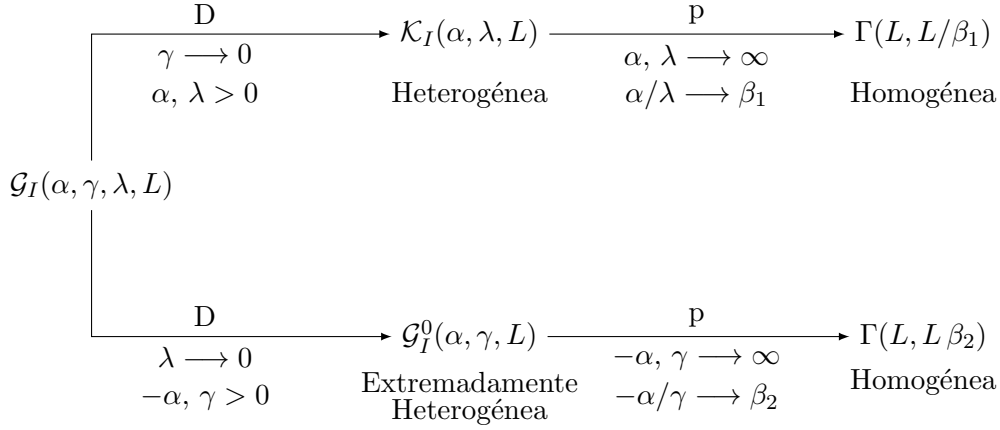


Figura 4.3: Comportamiento asintótico de la distribución $\mathcal{G}_I(\alpha, \gamma, \lambda, L)$.

Se puede observar que, cuando

- $\gamma \rightarrow 0$ y α, λ son positivos, $\mathcal{G}_I \xrightarrow{D} \mathcal{K}(\alpha, \lambda, L)$.
- $\lambda \rightarrow 0$ y α, γ son positivos, $\mathcal{G}_I \xrightarrow{D} \mathcal{G}_I^0(\alpha, \gamma)$.

Entonces, bajo el modelo multiplicativo, tenemos los siguientes modelos para el retorno que son casos particulares de la distribución \mathcal{G}_I . Si el ruido speckle $Y \sim \Gamma(L, L)$ y considerando diferentes modelos para el backscatter, tenemos las siguientes distribuciones para la intensidad del retorno Z .

- Si $X \sim \Gamma(\alpha, \lambda)$, entonces $Z \sim \mathcal{K}(\alpha, \lambda, L)$ con $\gamma, \lambda > 0$ y $L \leq 1$.
- Si $X \sim \Gamma^{-1}(-\alpha, \gamma)$, entonces $Z \sim \mathcal{G}_I^0(\alpha, \gamma, L)$ con $-\alpha, \gamma > 0$ y $L \leq 1$.
- Si $X \sim IG(\gamma, \lambda)$, entonces $Z \sim \mathcal{G}^H(\gamma, \lambda, L)$ con $\gamma, \lambda > 0$ y $L \leq 1$.

donde las distribuciones Γ^{-1} e IG están definidas por

Definición 4.1. Una variable aleatoria X tiene una distribución Inversa Gamma de parámetros $\lambda, \gamma > 0$ y se indicará $\Gamma^{-1}(\lambda, \gamma)$ si su función de densidad viene

dada por

$$f_X(x) = \frac{\beta^\lambda}{\Gamma(\lambda)} \frac{1}{x^{\lambda+1}} e^{-\beta \frac{1}{x}}, \text{ para } x > 0. \quad (4.0.11)$$

Definición 4.2. Una variable aleatoria X tiene una distribución Inversa Gaussiana de parámetros $\mu, \lambda > 0$ y se indicará $IG(\lambda, \gamma)$ si su función de densidad viene dada por

$$f_X(x) = \sqrt{\frac{\lambda}{2\pi x^3}} e^{-\frac{\lambda(x-\mu)^2}{2\mu^2 x}}, \text{ para } x > 0. \quad (4.0.12)$$

La distribución IG es un caso particular de la distribución Gaussiana Inversa Generalizada al considerar $\alpha = -1/2$ en la fórmula 4.0.9. Estas distribución han sido ampliamente estudiada por Jorgensen [56]. La familia de distribuciones $\mathcal{G}^H(\gamma, \lambda, L)$ para el retorno fueron aplicadas a datos SAR en [14, 52].

La familia de distribuciones \mathcal{G}_I modelan mejor datos que provienen de áreas heterogéneas y extremadamente heterogéneas que la familia \mathcal{K} [36, 66]. Mejail et al. [66] mostraron que, con una elección adecuada de los parámetros, la distribución \mathcal{G}_I puede aproximar muy bien a la distribución \mathcal{K} . Por estas razones, el modelo \mathcal{G}_I se llama Modelo Universal para datos SAR [35] y es el modelo utilizado en esta tesis.

4.1. Modelo \mathcal{G}_I^0

Diremos que si $Z = X \cdot Y$ con $X \sim \Gamma^{-1}(-\alpha, \gamma)$ e $Y \sim \Gamma(L, L)$ independiente, el retorno $Z \sim \mathcal{G}_I^0(\alpha, \gamma, L)$. Su función de densidad está dada por

$$f_{\mathcal{G}_I^0}(z) = \frac{L^L \Gamma(L - \alpha)}{\gamma^\alpha \Gamma(-\alpha) \Gamma(L)} \cdot \frac{z^{L-1}}{(\gamma + zL)^{L-\alpha}}, \quad (4.1.1)$$

donde $-\alpha, \gamma, z > 0$ y $L \geq 1$.

Los momentos de orden r están dados por

$$E(Z^r) = \left(\frac{\gamma}{L}\right)^r \frac{\Gamma(-\alpha - r)}{\Gamma(-\alpha)} \frac{\Gamma(L + r)}{\Gamma(L)}, \quad (4.1.2)$$

y son finitos si $\alpha < -r$.

En la figura 4.4 se muestra el gráfico de la densidad \mathcal{G}_I^0 para $L = 3$ y tres diferentes valores de α , en escala lineal 4.4(a) y en escala semilogarítmica 4.4(b). Se puede observar en 4.4(b) la diferencia que existe en el comportamiento de las colas de la distribución, estudiaremos este comportamiento en 4.1.1. Se puede observar que cuanto mayor es el parámetro de textura más pesadas son las colas de la distribución.

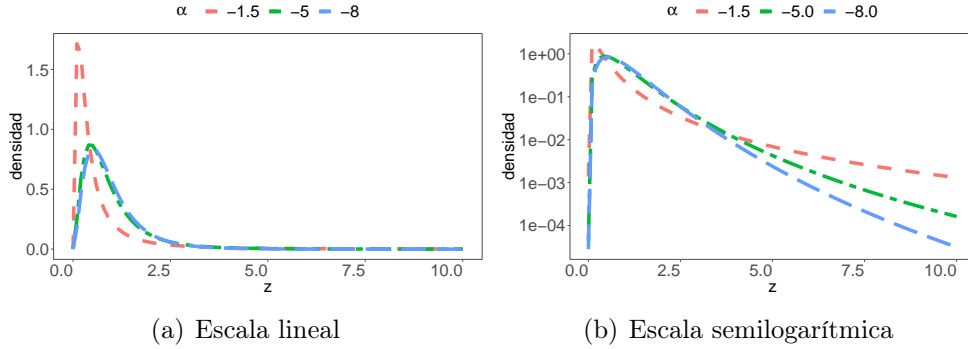


Figura 4.4: Densidad \mathcal{G}_I^0 para $L = 3$ y $\alpha = -1.5, -5, -8$.

En la figura 4.5 se muestra el gráfico de la densidad \mathcal{G}_I^0 para $L = 1, 3, 8$ y $\alpha = -5$. Se puede observar el efecto multilook, a medida que el número de looks aumenta, las colas son menos pesadas.

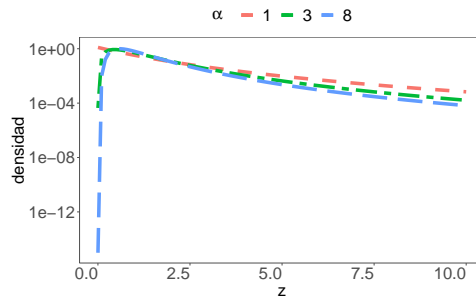


Figura 4.5: Densidad \mathcal{G}_I^0 para $L = 1, 3, 8$ y $\alpha = -5$.

Bajo este modelo se pueden caracterizar regiones con diferentes grado de textura a través de los parámetros de la distribución \mathcal{G}_I^0 . Para valores de α cercanos a cero (típicamente en el intervalo $(-3, 0)$), la zona de la imagen corresponde a una región muy texturada, como es el caso de las zonas urbanas en las imágenes SAR. A medida que el valor del parámetro α disminuye, corresponde a zonas con cada vez menos textura, como son las regiones de forestación (usualmente $(-6, -3]$) y pastura (en $(-\infty, -6)$). Por otro lado, el parámetro γ (llamado parámetro de escala) posee una interpretación en términos del brillo. Cuanto mayor es su valor, mayor intensidad posee la imagen en esa región. Por estas razones, la estimación precisa de los parámetros es de suma importancia en el análisis de imágenes con ruido speckle.

4.1.1. Resultados Teóricos

En esta sección se darán nuevos resultados teóricos que muestran que la distribución \mathcal{G}_I^0 es una distribución de colas pesadas [40]. Esto explica en parte los problemas que se enfrentan al buscar estimadores de los parámetros de esta distribución. También se presentan nuevos resultados sobre el comportamiento asintótico de la distribución \mathcal{G}_I^0 respecto de sus parámetros, que se necesitarán para demostrar la consistencia del estimador del parámetro de textura α propuesto en esta tesis.

\mathcal{G}_I^0 distribución de colas pesadas

Siguiendo a [27, 43, 83] daremos las siguientes definiciones.

Definición 4.3. La función $\ell: \mathbb{R} \rightarrow \mathbb{R}$ es de variación lenta en el infinito si para todo $t > 0$ vale que

$$\lim_{x \rightarrow +\infty} \frac{\ell(tx)}{\ell(x)} = 1.$$

Definición 4.4. Una función de densidad de probabilidad $f(x)$ tiene colas pesa-

das si para cualquier $\eta > 0$ vale que

$$f(x) = \ell(x)x^{-\eta},$$

donde ℓ es una función de variación lenta en el infinito y η es el índice de la cola.

Cuanto menor es el índice de la cola, mayor propensa es la distribución a producir observaciones extremas.

Proposición 4.1.1. La distribución \mathcal{G}_I^0 tiene colas pesadas con índice de la cola igual a $1 - \alpha$.

Demostración. Si definimos $\ell(x) = f_{\mathcal{G}_I^0}(x)x^{-\alpha+1}$ tenemos que

$$\begin{aligned} \lim_{x \rightarrow +\infty} \frac{\ell(tx)}{\ell(x)} &= \lim_{x \rightarrow +\infty} \frac{(tx)^{L-1} (Ltx + \gamma)^{\alpha-L} (tx)^{1-\alpha}}{x^{L-1} (Lx + \gamma)^{\alpha-L} x^{1-\alpha}} \\ &= \lim_{x \rightarrow +\infty} t^{L-\alpha} \left(\frac{Lx + \gamma}{Ltx + \gamma} \right)^{L-\alpha} = 1. \end{aligned}$$

Esto vale para todo $t > 0$, por lo tanto ℓ es de variación lenta en el infinito. \square

Se puede ver que el índice de la cola es una función decreciente de α , entonces la distribución \mathcal{G}_I^0 es más propensa a producir valores extremos cuando el parámetro de textura es más grande. Esto va de acuerdo a lo mostrado en la figura 4.5.

Comportamiento asintótico de la distribución \mathcal{G}_I^0 .

Estos resultados se probarán considerando el caso $E(X) = 1$ que nos da una relación entre los parámetros de la distribución \mathcal{G}_I^0 : $\gamma^* = -\alpha - 1$. Como el parámetro γ es considerado un parámetro de escala, con esta relación logramos que los resultados para diferentes áreas resulten comparables.

En este caso se puede escribir a $f_{\mathcal{G}_I^0}(z)$ en términos de γ . De esta forma, la

función de distribución del modelo \mathcal{G}_I^0 bajo esta condición queda:

$$f_{\mathcal{G}_I^0}(z) = \frac{\Gamma(L + \gamma + 1)}{\Gamma(\gamma + 1) \Gamma(L)} \left(\frac{L}{\gamma}\right)^L \frac{z^{L-1}}{\left(1 + \frac{Lz}{\gamma}\right)^{L+\gamma+1}}. \quad (4.1.3)$$

Vamos a utilizar esta forma de escribir a $f_{\mathcal{G}_I^0}(z)$ en las proposiciones 4.1.2 y 4.1.3.

Proposición 4.1.2. Para todo intervalo compacto $[z_1, z_2] \subset (0, \infty)$, $f_{\mathcal{G}_I^0}(z)$ converge uniformemente a $\Gamma(L, L)$ en $[z_1, z_2]$ si $\alpha \rightarrow -\infty$, donde $\Gamma(L, L)$ es la función de densidad del modelo Γ con parámetros (L, L) .

Demostración. Si consideramos $f_{\mathcal{G}_I^0}(z)$ escrita como en (4.1.3) estudiar la convergencia de $f_{\mathcal{G}_I^0}(z)$ cuando $\alpha \rightarrow -\infty$ es equivalente a estudiar el $\lim_{\gamma \rightarrow +\infty} f_{\mathcal{G}_I^0}(z)$.

Si llamamos:

$$\begin{aligned} a(\gamma) &= \frac{\Gamma(\gamma + 1) e^\gamma}{\sqrt{2\pi\gamma^{\gamma+1/2}}}, \\ b(L, \gamma, z) &= L\gamma^{L-1} \left(\frac{\frac{Lz}{\gamma}}{1 + \frac{Lz}{\gamma}}\right)^{L-1} \quad \text{y} \\ c(L, \gamma, z) &= \left(1 + \frac{Lz}{\gamma}\right)^{-2-\gamma}, \end{aligned}$$

entonces

$$f_{\mathcal{G}_I^0}(z) = \frac{a(L + \gamma)}{a(\gamma) \Gamma(L)} e^L \left(1 + \frac{L}{\gamma}\right)^{-L-\gamma-1/2} b(L, \gamma, z) c(L, \gamma, z). \quad (4.1.4)$$

De la fórmula de Stirling se obtiene

$$\lim_{\gamma \rightarrow +\infty} a(\gamma) = \lim_{\gamma \rightarrow +\infty} \frac{\Gamma(\gamma + 1) e^\gamma}{\sqrt{2\pi\gamma^{\gamma+1/2}}} = 1,$$

entonces

$$1 = \lim_{\gamma \rightarrow +\infty} \frac{a(L + \gamma)}{a(\gamma)} = \lim_{\gamma \rightarrow +\infty} \frac{\Gamma(L + \gamma + 1)}{\Gamma(\gamma + 1)} e^L \gamma^{-L} \left(1 + \frac{L}{\gamma}\right)^{-L-\gamma-1/2}.$$

Por otro lado, se verifica

$$\begin{aligned} \lim_{\gamma \rightarrow +\infty} b(L, \gamma, z) &= \lim_{\gamma \rightarrow +\infty} L \gamma^{L-1} \left(\frac{\frac{Lz}{\gamma}}{1 + \frac{Lz}{\gamma}} \right)^{L-1} \\ &= \lim_{\gamma \rightarrow +\infty} \frac{L^L z^{L-1}}{\left(1 + \frac{Lz}{\gamma}\right)^{L-1}} = L^L z^{L-1}, \end{aligned} \quad (4.1.5)$$

$$\lim_{\gamma \rightarrow +\infty} c(L, \gamma, z) = \lim_{\gamma \rightarrow +\infty} \left(1 + \frac{Lz}{\gamma}\right)^{-2-\gamma} = e^{-Lz}, \quad (4.1.6)$$

donde la convergencia en (4.1.5) y (4.1.6) es uniforme en $[z_1, z_2]$.

Entonces se obtiene que $\lim_{\alpha \rightarrow -\infty} f_{\mathcal{G}_I^0}(z) = L^L / \Gamma(L) e^{-Lz} z^{L-1}$ uniformemente en $[z_1, z_2]$, donde $f_{\mathcal{G}_I^0}(z)$ es la función de densidad dada en (4.1.4). \square

La figura 4.6 muestra la convergencia uniforme de $f_{\mathcal{G}_I^0}(z)$ a $\Gamma(3, 3)$, cuando $\alpha \rightarrow -\infty$ y $L = 3$. El área gris es una banda de ancho $\epsilon = 0.1$ alrededor de la función de densidad $\Gamma(3, 3)$ (línea sólida). La línea punteada y la línea discontinua representan $f_{\mathcal{G}_I^0}(z)$ para $\alpha = -20$ y $\alpha = -8$ respectivamente. Se puede ver que, cuando α disminuye, $f_{\mathcal{G}_I^0}(z)$ cae dentro del área gris. Mas aún, para este valor de ϵ , $f_{\mathcal{G}_I^0}(-20)$ ya cae dentro de esa banda.

Proposición 4.1.3. Para todo intervalo compacto $[z_1, z_2] \subset (0, \infty)$, $f_{\mathcal{G}_I^0}(z)$ converge uniformemente a 0 en $[z_1, z_2]$, si $\alpha \rightarrow -1^-$.

Demostración. En forma equivalente a la proposición 4.1.2 y usando la relación entre α y γ , estudiar la convergencia de $f_{\mathcal{G}_I^0}(z)$ cuando $\alpha \rightarrow -1^-$ es equivalente a estudiar ese límite cuando $\gamma \rightarrow 0^+$. De (4.1.3):

$$\begin{aligned} f_{\mathcal{G}_I^0}(z) &= \frac{\Gamma(L + \gamma + 1)}{\Gamma(\gamma + 1) \Gamma(L)} L^L \frac{z^{L-1}}{(\gamma + LZ)^{L+\gamma+1}} \gamma^{\gamma+1} \\ &\leq \frac{\Gamma(L + \gamma + 1)}{\Gamma(\gamma + 1) \Gamma(L)} L^L \frac{z_2^{L-1}}{(\gamma + Lz_1)^{L+\gamma+1}} \gamma^{\gamma+1} \end{aligned}$$

Dado que Γ es una función continua, para $z \in [z_1, z_2]$, $\lim_{\gamma \rightarrow 0^+} f_{\mathcal{G}_I^0}(z) = 0$ uniformemente en z , con lo que se obtiene el resultado. \square

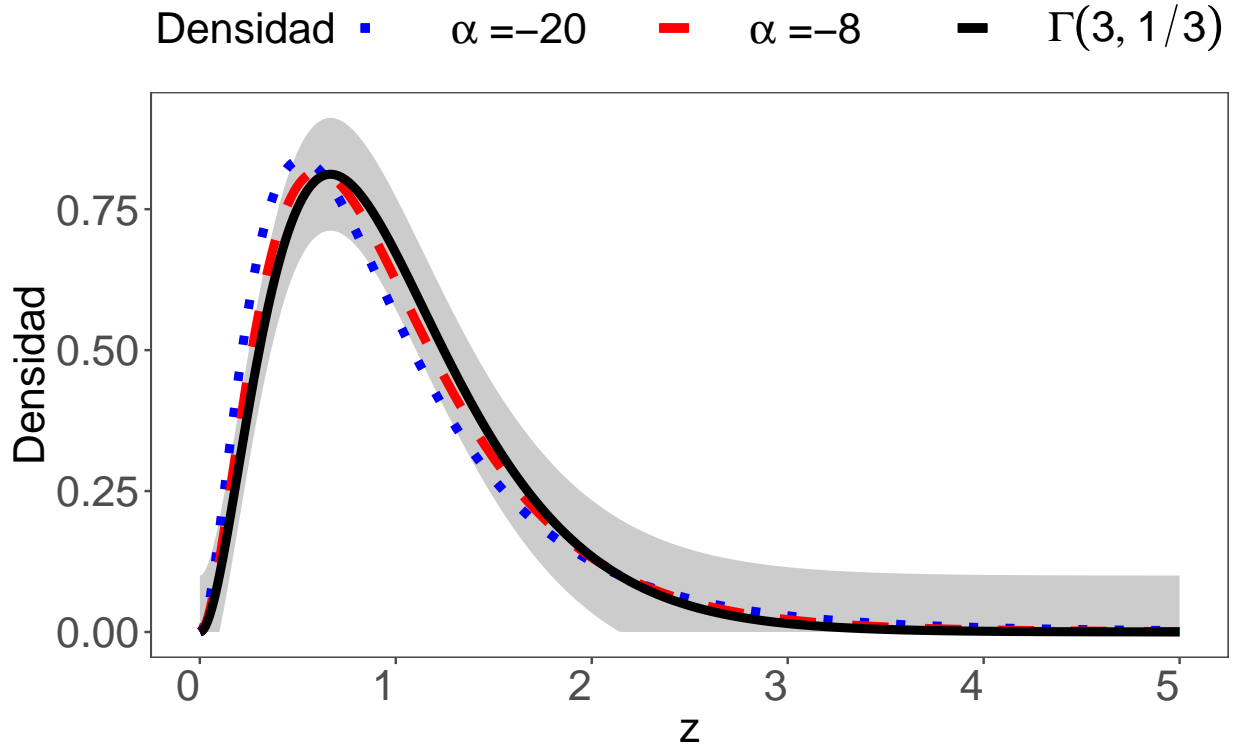


Figura 4.6: Convergencia uniforme de $f_{\mathcal{G}_l^0}(z)$ a $\Gamma(3, 3)$ cuando $\gamma = \gamma^*$, $\alpha \rightarrow -\infty$, $L = 3$ y $\epsilon = 0.1$.

La figura 4.7 muestra la convergencia uniforme de $f_{\mathcal{G}_l^0}(z)$ cuando $\alpha \rightarrow -1$ y $L = 3$. La banda gris es de ancho $\epsilon = 0.2$ alrededor del eje z . Se graficaron tres diferentes densidades con líneas punteada, discontinua y sólida, las cuales representan $f_{\mathcal{G}_l^0}(z)$ para $\alpha = -1.01$, -1.2 y -1.5 respectivamente. Se puede ver que cuando el valor de α se acerca a -1 , $f_{\mathcal{G}_l^0}(z)$ se aplanan a 0, excepto para $z = 0$. Para $\epsilon = 0.2$ y cualquier intervalo $[z_1, z_2]$ considerado, $f_{\mathcal{G}_l^0}(z)$ cae dentro de esta banda para α suficientemente cercano a -1 . Por ejemplo, para el intervalo $[1.5, 2]$ y para este valor de ϵ , $f_{\mathcal{G}_l^0}(z)$ entra en la banda gris desde $\alpha = -1.5$. Si el intervalo es $[0.5, 1]$ esto sucede a partir de $\alpha = -1.01$.

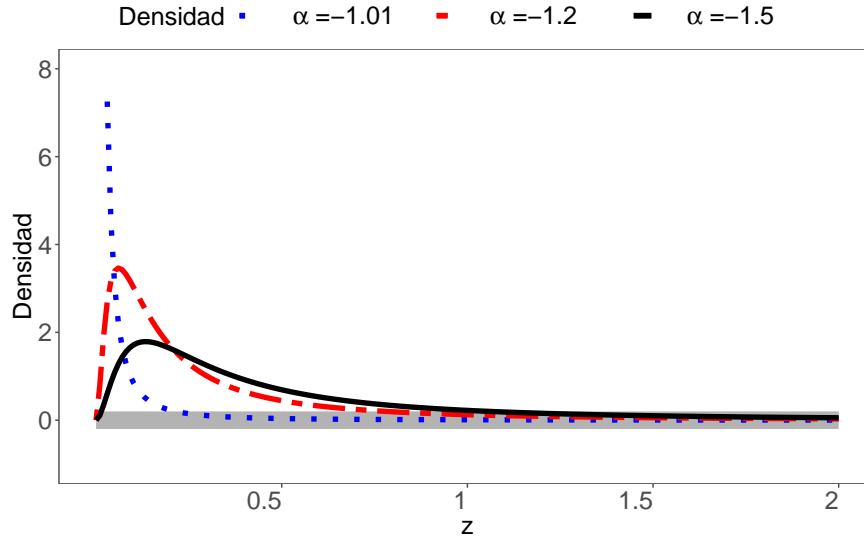


Figura 4.7: Convergencia uniforme de $f_{\mathcal{G}_0^0}(z)$ a 0 cuando $\gamma = \gamma^*$, $\alpha \rightarrow -1$, $L = 3$ y $\epsilon = 0.2$.

4.2. Conclusiones

En este capítulo presentamos varias distribuciones que se utilizan para modelar datos provenientes de imágenes SAR:

- Los llamados modelos empíricos que son distribuciones que se proponen para ajustar datos SAR, como la distribución Lognormal, Weibull entre otras.
- El modelo multiplicativo, que establece que el retorno Z se modela como el producto de dos variables aleatorias: X para el backscatter e Y para el ruido speckle

Mostramos un abordaje donde se explica la naturaleza multiplicativa del ruido speckle. A partir de esto se dedujo una distribución para el ruido speckle en el caso de datos en formato intensidad y monopolarizados, proponiendo $Y \sim \Gamma(L, L)$ para el modelo multilook, donde L es el número de looks. En cuanto a la retrodispersión X se propone modelarla con una distribución Gaussiana Inversa Generalizada $N^{-1}(\alpha, \gamma, \lambda)$. Para valores particulares de los parámetros de la distribución

$N^{-1}(\alpha, \gamma, \lambda)$, se obtienen las distribuciones $\Gamma(\alpha, \lambda)$, $\Gamma^{-1}(\alpha, \gamma)$ e $IG(\gamma, \lambda)$. Bajo el modelo multiplicativo estas distribuciones para la retrodispersión dan origen a las distribuciones \mathcal{K} , \mathcal{G}_0 y \mathcal{G}^H para el retorno Z , respectivamente.

En esta tesis utilizamos el modelo \mathcal{G}_I^0 para datos de intensidad. Presentamos la función de distribución y demostramos que es una distribución que tiene colas pesadas.

Asimismo demostramos resultado sobre la convergencia uniforme de la distribución \mathcal{G}_I^0 cuando el parámetro de textura se acerca a -1 y a $-\infty$ para el caso donde $E(Z) = 1$.

Capítulo 5

Metodología

Es sabido que tener información sobre la función de densidad o distribución de una variable aleatoria permite tener una completa descripción de la misma. Por este motivo es un problema fundamental de la Estadística obtener, a partir de la información proporcionada por una muestra, buenas estimaciones de las funciones de densidad de una variable o vector aleatorio.

En este sentido existen dos enfoques para abordar este problema.

- Un enfoque paramétrico: donde se considera que la función de densidad teórica pertenece a una familia de funciones de densidad $f_X(\theta)$ conocida, indexada por el vector de parámetros θ . Bajo esta suposición estimar la función de densidad teórica se reduce a estimar el valor de los parámetros del modelo a partir de la información proporcionada por la muestra. Los métodos clásicos de estimación paramétrica son: Método de los Momentos y Método de Máxima Verosimilitud. En los últimos años el Método de Logcumulantes, que se basa en la transformada de Mellin, ha ganado interés por su buena performance.
- Un enfoque no paramétrico: donde no se hace ninguna suposición inicial sobre la familia de densidades a las que pertenece la función de densidad teórica, sino que trata de estimarla teniendo como única información los

datos muestrales, imponiendo solamente las condiciones necesarias para que dicha estimación sea una función de densidad.

El principal aporte de esta tesis es la propuesta de un nuevo método de estimación para los parámetros del modelo \mathcal{G}_I^0 con buenas propiedades que serán estudiadas a través del sesgo, del error cuadrático medio y de su capacidad para resistir presencia de datos atípicos, aún en el caso de muestras de pequeño y moderado tamaño. Este estimador surge de la minimización de distancias estocásticas entre la función de densidad teórica y una estimación no paramétrica de la función densidad subyacente utilizando núcleos asimétricos.

En este capítulo describimos dos metodologías de estimación de la función de densidad: paramétrica y no paramétrica. Presentamos los estimadores de momento, de Máxima Verosimilitud y Logcumulantes del parámetro de textura de la distribución \mathcal{G}_I^0 , que son los estimadores que utilizaremos para evaluar la performance del propuesto en esta tesis. También presentamos los conceptos de estimación no paramétrica de la función de densidad, núcleos asimétricos, distancias estocásticas y estimadores de mínima distancia estocástica porque son el corazón del estimador planteado. Asimismo hacemos una discusión de diferentes distancias estocásticas estudiadas y diferentes núcleos utilizados para estimar la función de densidad subyacente. Finalmente nuestra propuesta para estimar el parámetro de textura del modelo \mathcal{G}_I^0 que es el argumento que minimiza una distancia estocástica entre la función de densidad teórica del modelo teórico y una estimación no paramétrica de la función de densidad subyacente.

5.1. Estimación Paramétrica

Esta metodología de estimación propone, a partir del conocimiento de la familia de densidades a la que pertenece la función de densidad teórica, estimar los parámetros que caracterizan dicha función de densidad a partir de datos muestrales. A continuación presentaremos métodos clásicos de estimación paramétrica.

5.1.1. Método de los Momentos

El Método de los Momentos (MOM) fue introducido por Karl Pearson en el año 1894 y se basa en la idea intuitiva de pensar que, si los datos muestrales provienen de una determinada función de densidad entonces los momentos muestrales deberían dar buenas estimaciones de los correspondientes momentos poblacionales.

Definición 5.1. Si X_1, \dots, X_n es una muestra de variables aleatorias independientes e idénticamente distribuidas con función de densidad $f_{X;\theta}$, donde $\theta \in \mathbb{R}$ es el parámetro a estimar, y $g(x)$ una función de \mathbb{R} en \mathbb{R} , el Método de los Momentos propone como estimador $\hat{\theta}_{mom}$ de θ aquel valor del parámetro que es solución de la ecuación $\frac{1}{n} \sum_{i=1}^n g(X_i) = E_{\theta}(g(X_i))$. Si en cambio $X_i \sim f_{X;\theta}$ con $\theta \in \Theta \subset \mathbb{R}^k$, entonces $\hat{\theta}_{mom}$ será el valor θ que verifica el sistema de ecuaciones $\frac{1}{n} \sum_{i=1}^n g_s(X_i) = E_{\theta}(g_s(X_i))$ donde $s = 1, \dots, k$ y $g_s(x)$ son k funciones de \mathbb{R} en \mathbb{R} .

Obsevemos que si $g_s(x) = x^s$ la definición es consistente con la idea intuitiva mencionada anteriormente.

El estimador por Momentos Fraccionales ha sido estudiado en los últimos tiempos [36, 39, 57]. En esta tesis se utilizarán los momentos de orden 1 y de orden $\frac{1}{2}$ ya que estos momentos existen para valores de $\alpha < -1$.

Las ecuaciones de momento de orden 1 y de orden $\frac{1}{2}$ para estimar los parámetros de la distribución \mathcal{G}_I^0 son:

$$\begin{aligned} \overline{X} &= \frac{\gamma}{-\alpha - 1}, \\ \frac{\sum_{i=1}^n X_i^{1/2}}{n} &= \frac{\gamma^{1/2}}{L^{1/2}} \frac{\Gamma(-\alpha - 1/2)}{\Gamma(-\alpha)} \frac{\Gamma(L + 1/2)}{\Gamma(L)}. \end{aligned} \quad (5.1.1)$$

5.1.2. Método de Máxima Verosimilitud

El Método de Máxima Verosimilitud fue popularizado por R. Fisher entre 1912 y 1922. Sin embargo este método había sido utilizado antes por Gauss, Laplace entre otros investigadores. La idea del método consiste en, dada una muestra, estimar los parámetros de la distribución que maximice la probabilidad de que haya salido esa muestra. Esto requiere mirar a la función de densidad conjunta como función de los parámetros, a esta función se llama función de verosimilitud.

Definición 5.2. Si X_1, \dots, X_n es una muestra de variables aleatorias independientes e idénticamente distribuidas con función de densidad $f_{X;\theta}$ con $\theta \in \Theta \subset \mathbb{R}^k$, diremos que $\hat{\theta}_{MV}$ es el estimador de Máxima Verosimilitud de θ si verifica que

$$\hat{\theta}_{MV} = \arg \max_{\theta \in \Theta} L(\theta \mid \vec{X} = \vec{x}), \quad (5.1.2)$$

siendo $L(\theta \mid \vec{X} = \vec{x})$ la función de verosimilitud, que es la función de densidad conjunta del vector aleatorio \vec{X} mirada como función del vector de parámetros θ , sabiendo que salió la muestra \vec{x} .

Este estimador tiene muy buenas propiedades asintóticas, además de ser consistente es asintóticamente normal y eficiente bajo condiciones generales.

El parámetro γ en la distribución \mathcal{G}_I^0 es proporcional al brillo y es un parámetro de escala. Con el fin de reducir nuestro análisis a un solo parámetro, y con el propósito de hacer los resultados comparables, basamos nuestro análisis en la condición $E(Z) = 1$ que vincula la textura con el brillo. Esta condición nos da una relación entre γ y α que es la siguiente:

$$\gamma = -\alpha - 1. \quad (5.1.3)$$

A este valor de γ lo llamaremos γ^* .

Entonces, si Z_1, \dots, Z_n una muestra de variables aleatorias independientes e idénticamente distribuidas donde $Z_i \sim f_{\mathcal{G}_I^0(\alpha, \gamma)}(z)$, el logaritmo de la función de

verosimilitud bajo la condición (5.1.3) es:

$$\begin{aligned} \log(L(\alpha) \mid \vec{Z} = \vec{z}) &= n(L \log L + \log \Gamma(L - \alpha) - \log \Gamma(-\alpha) - \alpha \log(-\alpha - 1)) \\ &\quad + (L - 1) \sum_{i=1}^n \log(z_i) - (L - \alpha) \sum_{i=1}^n \log(-\alpha - 1 + z_i L). \end{aligned} \quad (5.1.4)$$

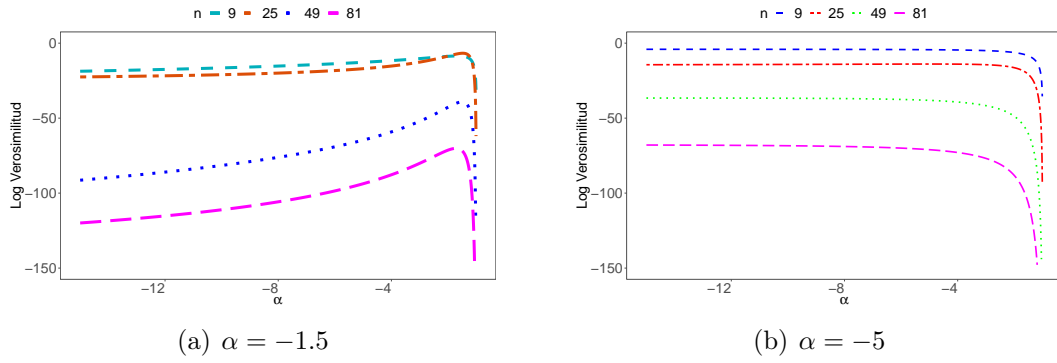
Por lo tanto, el estimador de Máxima Verosimilitud (MV) del parámetro de textura es:

$$\hat{\alpha}_{\text{MV}} = \arg \max_{\alpha} \log(L(\alpha) \mid \vec{Z} = \vec{z}). \quad (5.1.5)$$

Algunos autores evaluaron la performance para el estimador de Máxima Verosimilitud del parámetro de textura del modelo \mathcal{G}_I^0 . Vasconcellos et al. [97] cuantificaron el sesgo de este estimador y propusieron una corrección de segundo orden de acuerdo a [26]. Cribari-Neto et al. [24] desarrollaron una corrección del sesgo de $\hat{\alpha}_{\text{MV}}$ del modelo \mathcal{G}^0 para datos de amplitud usando técnicas de remuestreo a expensas de un alto costo computacional.

En la figura 5.1 se muestra el gráfico del logaritmo de la función de verosimilitud en función de α , para distintos tamaños de muestra n y para dos valores del parámetro de textura: $\alpha = -1.5$ y $\alpha = -5$ para el caso de $L = 3$. Se puede observar que, para zonas extremadamente texturadas la función de verosimilitud muestra claramente un máximo. Sin embargo, para el caso de zonas texturadas esta función es bastante plana, incluso podría ser monótona dependiendo de la muestra. Esto puede originar problemas numéricos al momento de encontrar el máximo.

También se podría encontrar $\hat{\alpha}_{\text{MV}}$ como la solución de la siguiente ecuación no lineal.

Figura 5.1: Función de verosimilitud para $L = 3$

$$\begin{aligned}
& \Psi^0(-\hat{\alpha}_{MV}) - \Psi^0(L - \hat{\alpha}_{MV}) - \log(-1 - \hat{\alpha}_{MV}) + \\
& \frac{\hat{\alpha}_{MV}}{-1 - \hat{\alpha}_{MV}} + \frac{1}{n} \sum_{i=1}^n \log(-1 - \hat{\alpha}_{MV} + Lz_i) - \\
& \frac{\hat{\alpha}_{MV} - L}{n} \sum_{i=1}^n \frac{1}{-1 - \hat{\alpha}_{MV} + Lz_i} = 0, \tag{5.1.6}
\end{aligned}$$

donde $\Psi^0(\cdot)$ es la función digamma. Este sistema no tiene una solución analítica cerrada, por lo tanto es necesario la utilización de rutinas numéricas para encontrar dicha solución que puede incluso no existir.

5.1.3. Logmomentos y Logcumulantes

Es conocida la relación que existe entre los momentos y la función característica de una variable aleatoria con la transformada de Fourier de su función de densidad.

Definición 5.3. Si X es una variable aleatoria continua con función de densidad $f_X(x)$, la función característica $\phi_X : \mathbb{R} \rightarrow \mathbb{C}$ se define como

$$\phi_X(t) = E(e^{itX}) = \int_{-\infty}^{\infty} e^{itx} f_X(x) dx = \mathcal{F}\{f_X\}(t), \tag{5.1.7}$$

siendo $\mathcal{F}\{f_X\}$ la transformada de Fourier de f_X .

Cuando los momentos de orden n de una variable aleatoria existen, es decir, cuando $E(X^k) < \infty$, se puede apelar a la relación (5.2.30) para calcular dichos momentos. Entonces

$$\phi_X^{(k)}(0) = i^k E(X^k), \quad (5.1.8)$$

con $\phi_X^{(k)}$ la derivada de orden k de ϕ_X evaluada en $t = 0$. Luego, los momentos de orden k se pueden definir como

$$m_k = E(X^k) = i^{-k} \phi_X^{(k)}(t) \big|_{t=0}. \quad (5.1.9)$$

Si llamamos $\psi_X = \log(\phi_X)$, los cumulantes de orden k se definen como

$$\kappa_k = i^{-k} \psi_X^{(k)} \big|_{t=0}. \quad (5.1.10)$$

Nicolas et al. [71] proponen utilizar la transformada de Mellin en lugar de la transformada de Fourier en (5.2.30), para el caso donde la función de densidad f_X de la variable aleatoria X tiene soporte positivo. De esta manera surgen nuevas características para la variable aleatoria que son los Logmomentos (MomL) y los Logcumulantes (LC). Los autores proponen llamar a esta metodología Estadísticas de Segundo Tipo. Entonces, análogamente a la metodología descrita en 5.2.30, definen:

- Primera función característica de segundo tipo.

$$\Phi_X(s) = \int_0^\infty x^{s-1} f_X(x) dx = \mathcal{M}\{f_X\}(s). \quad (5.1.11)$$

- Momentos de segundo tipo.

$$\tilde{m}_k = \Phi_X^{(k)}(s) \Big|_{s=1}. \quad (5.1.12)$$

Una propiedad interesante que cumplen las derivadas de orden k de la transformada de Mellin de una función $h(x)$ es $\mathcal{M}\{\ln(x)^k h(x)\}(s) = \mathcal{M}^{(k)}\{h(x)\}(s)$.

Entonces, aplicando esta propiedad a (5.1.11) y (5.1.12), y considerando $h = f_X$ se obtiene:

$$\begin{aligned} \tilde{m}_k &= \Phi_X^{(k)}(s) \Big|_{s=1} = \mathcal{M}^{(k)}\{f(x)\}(s) = \mathcal{M}\{f(x) \log(x)^k\}(s) \Big|_{s=1} \\ &= \int_0^\infty \log(x)^k f_X(x) dx. \end{aligned} \quad (5.1.13)$$

Esta última ecuación sugiere llamar a los momentos de segundo tipo como Logmomentos (MomL).

- Segunda función característica de segundo tipo.

Manteniendo la analogía con las estadísticas de primer tipo, Nicolas et al. [71] definen la *Segunda función característica de segundo tipo* como

$$\Psi_X(s) = \log(\Phi)(s). \quad (5.1.14)$$

- Cumulantes de segundo tipo.

Las derivadas de orden k de la segunda función característica de segundo tipo, evaluada en $s = 1$, se definen como los cumulantes de segundo tipo. Entonces,

$$\tilde{\kappa}_k = \Psi_X^{(k)} \Big|_{s=1}. \quad (5.1.15)$$

Como en el caso de MomL, los cumulantes de segundo tipo se llamarán

Logcumulantes (LC).

Para el caso donde $f_X(u) = \mathcal{G}_I^0(u)$ se obtiene:

$$\Phi_x(s) = \frac{\left(\frac{L}{\gamma}\right)^{1-s} \Gamma(-1 + L + s) \Gamma(1 - s - \alpha)}{\Gamma(L) \Gamma(-\alpha)} \quad (5.1.16)$$

$$\begin{aligned} \tilde{m}_1 &= \left. \frac{d\Phi_x(s)}{ds} \right|_{s=1} \\ &= -\log\left(\frac{L}{\gamma}\right) + \Psi^0(L) - \Psi^0(-\alpha). \end{aligned} \quad (5.1.17)$$

Usando los desarrollos presentados en Tison et al. [96] tenemos que $\tilde{k}_1 = \tilde{m}_1$. Por lo tanto

$$\tilde{k}_1 = -\log\left(\frac{L}{\gamma}\right) + \Psi^0(L) - \Psi^0(-\alpha). \quad (5.1.18)$$

Además, si Z_1, \dots, Z_n es una muestra de variables aleatorias independientes e idénticamente distribuidas donde $Z_i \sim \mathcal{G}_I^0$, usando (5.1.13) tenemos que LC es

$$\hat{\tilde{k}}_1 = \frac{1}{n} \sum_{i=1}^k \log z_i. \quad (5.1.19)$$

Asumiendo que vale la relación (5.1.3), $\gamma^* = -\alpha - 1$, el estimador Logcumulantes del parámetro de textura α de la distribución \mathcal{G}_I^0 , que lo llamaremos $\hat{\alpha}_{LC}$, es la solución de la ecuación

$$\hat{\tilde{k}}_1 = -\log\left(\frac{L}{-\alpha - 1}\right) + \Psi^0(L) - \Psi^0(-\alpha), \quad (5.1.20)$$

es decir, la solución de la ecuación

$$\frac{1}{n} \sum_{i=1}^n \log(z_i) = -\log\left(\frac{L}{-\alpha - 1}\right) + \Psi^0(L) + \Psi^0(-\alpha). \quad (5.1.21)$$

5.2. Estimación No Paramétrica

Como mencionamos anteriormente, otra posibilidad para estimar los parámetros de un modelo es la estimación no paramétrica. Esta metodología de estimación no determina a priori ningún modelo para la distribución de la variable aleatoria de interés, y propone estimadores de la función de densidad sin más límites que los necesarios para que estos estimadores cumplan con las condiciones de ser una función de densidad.

5.2.1. Distancias Estocásticas

Al momento de proponer un modelo estadístico es importante contar con una medida apropiada que indique la diferencia entre el modelo teórico y los datos muestrales. Si nos basamos en la información muestral existen medidas que vinculan la función de distribución empírica y el modelo teórico, y otras que miden la discrepancia entre una estimación no paramétrica de la función de densidad subyacente y la función de densidad que proviene del modelo utilizado. En esta tesis se trabajará con esta última propuesta.

Es conocido el aporte que realizó Claude E. Shannon a la creación de lo que se conoce como *Teoría de la Información*. En su trabajo [91] propone una nueva manera de medir la transmisión de información a través de un canal, pensando a la información como un concepto estadístico. Propone como medida de información o de desorden a lo que se conoce como entropía de Shannon. En el caso discreto y considerando X una variable aleatoria que toma n valores con probabilidades p_1, \dots, p_n , entonces la entropía de Shannon se define como $H(X) = -\sum_{i=1}^n p_i \log(p_i)$. Para el caso continuo se define $H(X) = \int_{-\infty}^{\infty} p_X(x) \log(p_X(x))$.

Kullback y Leibler (KL) [60] estaban interesados en proponer una medida que discrimine dos poblaciones, considerando una cantidad que involucre medidas de información. Su interés era definir una magnitud que indique la diferencia entre dos funciones de distribución pensando en una generalización de la entropía de

Shannon. Si $p(x)$ y $q(x)$ son dos funciones de probabilidad puntual con el mismo soporte Ω , la KL divergencia se define como $d_{\text{KL}}(p\|q) = -\sum_{x \in \Omega} p(x) \log \frac{q(x)}{p(x)} = \sum_{x \in \Omega} p(x) \log \frac{p(x)}{q(x)}$.

En el caso continuo si $P(x)$ y $Q(x)$ son funciones de distribución de una misma variable aleatoria X continua, la divergencia KL se define como

$$d_{\text{KL}}(P\|Q) = \int_{-\infty}^{+\infty} p(x) \log \frac{p(x)}{q(x)} dx, \quad (5.2.1)$$

donde $p(x)$ y $q(x)$ representan las funciones de densidad de P y Q respectivamente.

Se puede pensar que si $p(x)$ es el modelo verdadero, la KL divergencia mide cuánto me alejo del modelo verdadero, si utilizo el modelo q en vez del modelo p . Hay que notar que si $p(x) = q(x)$ la distancia $\text{KL} = 0$. Esta medida de “distancia” no es simétrica ni cumple con la desigualdad triangular, en cambio sí cumple que $d_{\text{KL}}(p\|q) \geq 0$ y $d_{\text{KL}}(p\|q) = 0 \iff p = q$.

Muchas de las medidas utilizadas que se basan en determinar diferencias entre una estimación de la función de densidad subyacente y el modelo teórico no son métricas, donde por métrica entendemos a aquellas funciones $D : X \times X \rightarrow \mathbb{R}^+$ que cumplen:

1. $D(x, y) \geq 0 \forall x, y \in X$.
2. $D(x, y) = 0 \iff x = y$.
3. $D(x, y) = D(y, x) \forall x, y \in X$.
4. $D(x, z) \leq D(x, y) + D(y, z) \forall x, y, z \in X$.

Es decir, son medidas que no cumplen la propiedad 3 ni 4, pero sí cumplen las propiedades 1 y 2. Si la medida de discrepancia cumple solamente las propiedades 1 y 2 la llamaremos divergencia. Si además cumple la propiedad 3 de simetría la llamaremos distancia estocástica.

Salicrú et al. [86] propusieron una familia de medidas de divergencia, llamadas (h, ϕ) -divergencias que incluyen a la KL divergencia, a las ϕ -divergencias presentadas por Csiszar [25] y a las generalizaciones de las J -divergencias y R -divergencias que fueron definidas por Taneja en [94] entre otras. Algunas de las medidas que se trabajan en esta tesis son ϕ divergencias y otras (h, ϕ) -divergencias. Por eso daremos la definición de ambas.

Definición 5.4 (ϕ -divergencias). Sean $F(x)$ y $G(x)$ dos funciones de distribución definidas sobre el mismo conjunto $S \subset \mathbb{R}$, ambas absolutamente continuas con respecto a la medida Lebesgue, es decir existen $f(x)$ y $g(x)$ funciones de densidad correspondientes a F y G respectivamente. Además $F(x)$ absolutamente continua con respecto a $G(x)$, entonces para toda función convexa $\phi : [0, +\infty) \rightarrow \mathbb{R}$ tal que $\phi(1) = 0$ la ϕ -divergencia entre f y g se define como

$$d_\phi(f, g) = \int_S \phi\left(\frac{f}{g}\right) g dx = E_g\left(\phi\left(\frac{f}{g}\right)\right). \quad (5.2.2)$$

Cabe realizar algunas observaciones:

- Por la desigualdad de Jensen y por ser f una función de densidad se verifica que:

$$\int_S \phi\left(\frac{f(x)}{g(x)}\right) g(x) dx \geq \phi\left(\int_S \frac{f(x)}{g(x)} g(x) dx\right) = \phi(1). \quad (5.2.3)$$

Entonces, la condición $\phi(1) = 0$ garantiza que la medida de divergencia cumpla que $d_\phi(f, g) \geq 0$. Además, si $f = g$ entonces $d_\phi(f, f) = 0$.

- La condición que $F(x)$ sea absolutamente continua con respecto a $G(x)$ es equivalente a pedir que, en casi todo punto, $\text{Sop}(g) \subset \text{Sop}(f)$. En general, en la literatura se pide que ambas funciones de densidad tengan un soporte común.

Ejemplo 5.2.1. Ejemplos de ϕ -divergencias. En estos ejemplos llamaremos $S = \text{Sop}(f) = \text{Sop}(g)$.

- Kullback-Leibler definida en 5.2.1 donde $\phi(x) = x \log(x)$.
- Hellinger

$$d_H(f, g) = \frac{1}{2} \int_S \left(\sqrt{f(x)} - \sqrt{g(x)} \right)^2 dx, \quad (5.2.4)$$

donde $\phi(x) = (\sqrt{x} - 1)^2$.

- Triangular

$$d_T(f, g) = \int_S \frac{(f(x) - g(x))^2}{f(x) + g(x)} dx, \quad (5.2.5)$$

donde $\phi(x) = \frac{(x - 1)^2}{x + 1}$.

Definición 5.5. (h, ϕ) -divergencias.

Si se mantienen las hipótesis de la definición (5.4) y $h : (0, +\infty) \rightarrow [0, +\infty)$ es una función estrictamente creciente que verifica que $h(0) = 0$ entonces se define la (h, ϕ) -divergencia como

$$d_\phi^h(f, g) = h \left(\int_S \phi \left(\frac{f(x)}{g(x)} \right) g(x) dx \right) = h \left(E_g \left(\phi \left(\frac{f}{g} \right) \right) \right). \quad (5.2.6)$$

Ejemplo 5.2.2. Ejemplos de (h, ϕ) -divergencia son:

- Bhattacharyya

$$d_B = -\log \int_S \sqrt{f_v f_w}, \quad (5.2.7)$$

donde $\phi(x) = -\sqrt{x} + \frac{x+1}{d}$ y $h(y) = -\log(-y + 1)$ con $0 \leq y < 1$.

- Rényi de orden $\beta \in (0, 1)$

$$d_{\text{R}}^{\beta}(V, W) = \frac{1}{2(\beta - 1)} \log \int_{-\infty}^{\infty} (f_{\text{v}}^{\beta} f_{\text{w}}^{1-\beta}) + \log \int_{-\infty}^{\infty} (f_{\text{v}}^{1-\beta} f_{\text{w}}^{\beta}), \quad (5.2.8)$$

donde

$$\phi(x) = \frac{x^{\beta} - \beta(x - 1) - 1}{\beta - 1} \text{ con } 0 < \beta < 1, \text{ y}$$

$$h(y) = \frac{1}{\beta - 1} \log((\beta - 1)y + 1) \text{ con } 0 \leq y < \frac{1}{1 - \beta}.$$

Se han encontrado muchas aplicaciones de estas medidas de divergencia en señales y procesamiento de imágenes [6], en el análisis de imágenes médicas [98], detección automática de regiones en imágenes SAR [65, 70, 92].

Cabe señalar que muchos autores han encontrado cotas superiores e inferiores que vinculan la distancia triangular con otras medidas de divergencia que se encuentran en la clase de las ϕ -divergencias, especialmente para el caso discreto. Particularmente esto es útil ya que contribuye a obtener resultados de convergencia entre distintas divergencias. En Dragomir [30] el autor encuentra cotas superiores e inferiores para las ϕ -divergencias en términos de la KL divergencia. Jain et. al. [59] también encuentra cotas que vinculan a la distancia triangular con las discrepancias Jensen – Shannon, Simétrica Chi-square, Aritmética entre otras. Taneja [95] estudió nuevas cotas entre la discriminación triangular (distancia triangular para el caso discreto) y la media armónica y la simétrica Chi-square divergencia.

En Casseti et al. [20] se estudiaron estas divergencias o distancias estocásticas en el contexto de proponer un nuevo estimador para α , el parámetro de textura para el modelo \mathcal{G}_I^0 , a través de la minimización de la distancia estocástica entre el modelo teórico y una estimación no paramétrica de la función de densidad subyacente. En ese trabajo se analizaron las distancias estocásticas de Hellinger, Bhattacharyya, Rényi and Triangular y se compararon con el estimador de Máxima Verosimilitud de α mostrando evidencia que la distancia triangular era una

buena elección para este problema. Por este motivo nos focalizaremos en estudiar la distancia triangular en particular. En lo siguiente definiremos la distancia triangular generalizada y veremos que es una métrica en el sentido que cumple con las condiciones (5.2.1). Probaremos algunos resultados que nos serán útiles al momento de estudiar la convergencia del estimador propuesto en esta tesis.

5.2.2. Distancia triangular generalizada

Sea $\delta_s : \mathbb{R}_+ \times \mathbb{R}_+ \rightarrow \mathbb{R}_+$ la función definida por

$$\delta_s(a, b) = \begin{cases} \frac{|a - b|}{(a + b)^{1-1/s}} & \text{si } (a, b) \neq (0, 0) \\ 0 & \text{si } (a, b) = (0, 0) \end{cases}$$

Lema 5.2.1. Para $s \geq 1$, δ_s define una distancia en \mathbb{R}_+ y para $a, b \in \mathbb{R}_+$, se verifica $\delta_s(a, b) \leq |a - b|^{1/s}$.

Demostración. Claramente $\delta_s(a, b) = \delta_s(b, a) \geq 0$ y $\delta_s(a, b) = 0$ si y solo si $a = b$. Para mostrar que δ_s define una distancia en \mathbb{R}_+ basta mostrar que se cumple la desigualdad triangular

$$\delta_s(a, c) \leq \delta_s(a, b) + \delta_s(b, c).$$

El resultado es obvio si dos de los tres números son iguales, por lo tanto podemos suponerlos diferentes. Vamos a considerar los siguientes casos:

- I) Si $b \leq a, c$, entonces $(a + b)^{1-1/s}, (b + c)^{1-1/s} \leq (a + c)^{1-1/s}$. Como se verifica $|a - c| \leq |a - b| + |b - c|$, tenemos

$$\frac{|a - c|}{(a + c)^{1-1/s}} \leq \frac{|a - b|}{(a + b)^{1-1/s}} + \frac{|b - c|}{(c + b)^{1-1/s}}$$

II) Tomemos el caso $a, c \leq b$. Si $a = 0$ (el caso $c = 0$ es similar), entonces

$$\delta_s(a, c) = |c|^{1/s} \leq |b|^{1/s} + \delta_s(b, c) = \delta_s(a, b) + \delta_s(b, c).$$

Supongamos $a, c > 0$, tenemos que $b^{-1} \leq a^{-1}, c^{-1}$ y por el caso anterior

$$\delta_s(a^{-1}, c^{-1}) \leq \delta_s(a^{-1}, b^{-1}) + \delta_s(b^{-1}, c^{-1}),$$

usando que $\delta_s(a, b) = (ab)^{1/s} \delta_s(a^{-1}, b^{-1})$

$$\begin{aligned} \delta_s(a, c) &= (ac)^{1/s} \delta_s(a^{-1}, c^{-1}) \\ &\leq (ac)^{1/s} \delta_s(a^{-1}, b^{-1}) + (ac)^{1/s} \delta_s(b^{-1}, c^{-1}) \\ &\leq (ab)^{1/s} \delta_s(a^{-1}, b^{-1}) + (bc)^{1/s} \delta_s(b^{-1}, c^{-1}) \\ &\leq \delta_s(a, b) + \delta_s(b, c). \end{aligned}$$

III) Para $a \leq b \leq c$, usamos que $a+c = b+(1-t)a+tc$ donde $t = (c-b)/(c-a)$ y que la función $g(z) = (b+z)^{-1+1/s}$ es convexa, para obtener

$$(a+c)^{-1+1/s} \leq \frac{b-a}{c-a} (a+b)^{-1+1/s} + \frac{c-b}{c-a} (b+c)^{-1+1/s},$$

lo que prueba la desigualdad.

□

Definición 5.6. Distancia triangular generalizada.

Dadas f, g dos densidades de probabilidad, definimos

$$\begin{aligned} d_{T,s}(f, g) &= \int_{\Lambda} \frac{|f(x) - g(x)|^s}{(f(x) + g(x))^{s-1}} \\ &= \int_{\Lambda} \phi_s\left(\frac{f}{g}\right) g, \end{aligned} \tag{5.2.9}$$

donde $\Lambda = \text{Sop}(f) \cup \text{Sop}(g)$ y $\phi_s(x) = \frac{|x-1|^s}{(x+1)^{s-1}}$, lo que muestra que $d_{T,s}(f, g)$

es una ϕ - divergencia.

Observemos que

$$\delta_s^s(f, g) = \frac{|f - g|^s}{(f + g)^{s-1}} = \frac{|f - g|^{s-1}}{(f + g)^{s-1}} |f - g| \leq |f - g|, \quad (5.2.10)$$

lo que muestra que $d_{T,s}(f, g)$ está definido si $f - g$ es integrable.

Proposición 5.2.1. Para $s \geq 1$ el funcional $D_{T,s} = d_{T,s}^{1/s}$ define una distancia en el espacio de las funciones no negativas integrables.

Demostración. Sean f, g, h funciones no negativas integrables, por el lema anterior tenemos (eliminamos la variable x para hacerlo más legible)

$$0 \leq \delta_s(f, h) \leq \delta_s(f, g) + \delta_s(h, g),$$

por lo tanto

$$D_{T,s}(f, h) = \left(\int \delta_s^s(f, h) \right)^{1/s} \leq \left(\int (\delta_s(f, g) + \delta_s(h, g))^s \right)^{1/s}.$$

Usando la desigualdad de Minkowski en L^s , obtenemos

$$\left(\int (\delta_s(f, g) + \delta_s(h, g))^s \right)^{1/s} \leq D_{T,s}(f, g) + D_{T,s}(g, h),$$

lo que prueba la proposición. □

Remark 5.2.1. Cabe señalar que si $h_s(y) = y^{1/s}$ y ϕ como la función definida en 5.2.9, entonces $d_{T,s}^{h_s}$ es una (h, ϕ) - divergencia.

5.2.3. Núcleos

Estimar una función de densidad desconocida a partir de datos muestrales es un problema clásico en estadística. Si se conoce la familia de distribuciones de

donde proviene la muestra, se puede dar una estimación de los parámetros de la función de densidad subyacente usando la metodología presentada en 5.1.

Otra forma de abordar el problema de la estimación de la función de densidad subyacente f es a través de una estimación no paramétrica de la misma. El más sencillo y más conocido de estos estimadores es el histograma. Supongamos que x_1, \dots, x_{n+1} son los datos muestrales donde $x_1 < x_2 < \dots < x_{n+1}$, sea $B_i = [x_i, x_{i+1})$ con $i = 1, \dots, n$. Entonces el estimador \hat{f} de f se define como

$$\hat{f}_b(x) = \frac{1}{nb} \sum_{i=1}^n I_{B_i}(x_i) = \frac{N_i}{nb} \text{ para } x \in B_i, \quad (5.2.11)$$

donde b es la longitud del intervalo B_i que suponemos la misma, N_i la cantidad de observaciones muestrales que caen dentro del intervalo B_i e $I_{B_i}(x)$ es la función indicadora definida en B_i . En el caso donde los intervalos tienen distinta longitud se deberá dividir por la longitud de cada uno dentro de la sumatoria.

Este estimador es muy utilizado especialmente cuando se quiere hacer un estudio exploratorio de la distribución de los datos, pero provee una estimación de la función de densidad que no es continua y, por lo tanto, no derivable. Estas propiedades son deseables, principalmente cuando se trabaja con variables aleatorias continuas.

Núcleos simétricos

Rosenblatt [84] y Parzen [77] avanzaron en la búsqueda de un estimador para f y propusieron los estimadores tipo núcleo, con núcleo simétrico (KS). Si consideramos X_1, \dots, X_n una muestra de variables aleatorias independientes e idénticamente distribuidas este estimador se define como

$$\hat{f}_{sn}(x) = \frac{1}{nb} \sum_{i=1}^n K\left(\frac{x - X_i}{b}\right), \quad (5.2.12)$$

donde b es el parámetro de suavizamiento llamado también bandwidth o ancho de banda. $K(x)$ es una función llamada núcleo o kernel que satisface ciertas propiedades, entre ellas la simetría y $\int_{-\infty}^{\infty} K(x)dx = 1$ entre otras. Estos estimadores fueron muy estudiados por diversos autores, en Silverman [93] y Scott [90] se puede ver un exhaustivo estudio de ellos y sus propiedades. En particular, se puede mostrar que estos estimadores son consistentes en cada x si se cumple que

$$\lim_{n \rightarrow \infty} nb = \infty. \quad (5.2.13)$$

Asimismo se pueden obtener expresiones para el sesgo y la varianza, estas vienen dadas por

$$B(\hat{f}(x)) = \frac{b^2 f''(x) k_2}{2} + O(b^4) \quad (5.2.14)$$

$$Var(\hat{f}(x)) \approx \frac{1}{nb} \int_{-\infty}^{\infty} K^2(t) dt \quad (5.2.15)$$

$$(5.2.16)$$

respectivamente, donde $k_2 = \int_{-\infty}^{\infty} x^2 K(t) dt$. Esto muestra el compromiso que existe entre el sesgo y la varianza, una rápida convergencia a 0 del parámetro b genera una disminución del sesgo pero un aumento importante en la varianza. Por eso es necesaria la condición (5.2.13), el ancho de banda debe converger a cero a una velocidad más lenta que n^{-1} .

Medidas de error

Dado un problema de estimación es necesario disponer de criterios que permitan comparar entre varios estimadores con el objetivo de elegir el estimador óptimo.

En el caso de estimación paramétrica donde se supone que X es una variable aleatoria con función de densidad f_{θ} y se quiere estimar el parámetro θ , un criterio ampliamente utilizado es elegir el estimador que tenga el menor error cuadrático

medio (MSE), definido como:

$$\text{MSE}(\hat{\theta}) = \text{Var}(\hat{\theta}) + B^2(\hat{\theta}). \quad (5.2.17)$$

En cambio, en el enfoque de la estimación no paramétrica donde se obtiene una estimación y representación completa de la función de densidad, es necesario contar con medidas de error que midan el comportamiento del estimador de la función de densidad subyacente en forma global. Existen muchas medidas de error, dentro de las más utilizadas se encuentra el error cuadrático integrado (ISE) que se define, para un ancho de banda b dado, como:

$$\text{ISE}(\hat{f}_b(X_1, \dots, X_n)) = \int_{\mathbb{R}} (\hat{f}_b(x, X_1, \dots, X_n) - f(x))^2 dx. \quad (5.2.18)$$

El ISE es una variable aleatoria que depende del núcleo utilizado, del tamaño muestral, de la verdadera función de densidad y del estimador. Pero también depende de una realización particular de la muestra. Por lo tanto, una medida que contemple un comportamiento promedio del estimador sobre diversas realizaciones muestrales es considerar un promedio del ISE sobre dichas realizaciones. Es decir, considerar

$$\text{MISE}(\hat{f}_b(X_1, \dots, X_n)) = E\left[\int_{\mathbb{R}} (\hat{f}_b(x, X_1, \dots, X_n) - f(x))^2 dx\right]. \quad (5.2.19)$$

Núcleos Asimétricos

En general los estimadores \hat{f}_{KS} tienen un buen comportamiento cuando la función de densidad está definida en toda la recta real. Pero cuando la función de densidad tiene soporte acotado o semiacotado inferiormente este estimador presenta sesgo especialmente cerca del borde del soporte. Esto se debe a que estos estimadores asignan probabilidad fuera del soporte de la distribución.

Diferentes alternativas se han propuesto para resolver el problema del sesgo en el borde utilizando núcleos simétricos, alguna de ellas pueden dar estimaciones

negativas de la función de densidad. Entre estas alternativas se encuentran los métodos de reflexión, de renormalización y los boundary kernels entre otras. Una descripción de estos métodos se pueden encontrar en [55].

Otra forma de encarar el problema del sesgo en el borde es la utilización de núcleos asimétricos (A). Brown and Chen [12, 22] propusieron los Beta kernel, Chen [23] los Gamma kernels (GA), Scaillet [88] el Inverso Gaussiano (IG) y Recíproco Gaussiano Inverso (RIG), Jin et. al. [54] el Lognormal kernel (LN) y Birnbaum-Saunders. Todos estos núcleos tienen soporte en $[0, +\infty)$, no asignan peso fuera del soporte de la distribución y alcanzan una tasa de convergencia en el sentido del error cuadrático medio integrado de $n^{-4/5}$. Además, de acuerdo a [88] los núcleos GA, IG y RIG son libres de sesgo en el borde en el sentido que el sesgo es un $O(b)$, mientras que el sesgo correspondiente al núcleo LN es un $O(b^2)$ [29].

Observemos que la expresión 5.2.12 correspondiente al estimador \hat{f}_{s_n} se puede escribir como

$$\hat{f}_{s_n} = \frac{1}{n} \sum_{i=1}^n K_{(x,b)}(X_i), \quad (5.2.20)$$

donde $K_{(x,b)}(\cdot) = \frac{1}{b} K\{(\cdot - x)/h\}$. Como indica Hirukawa en [47] esta expresión se puede generalizar considerando el caso donde el núcleo $K_{(x,b)}$ es una función asimétrica y su forma puede variar de acuerdo a la posición de x . Ya en 1986 Silverman mencionó en [93] la posibilidad de utilizar las funciones de densidad Γ y Lognormal como núcleos cuando la función de densidad a estimar es sesgada a derecha y tiene soporte positivo.

Entonces, los núcleos asimétricos se definen como:

Definición 5.7. Sea X_1, \dots, X_n una muestra aleatoria donde $X_i \sim f_\theta$ con f_θ la función de densidad teórica, el estimador \hat{f}_{A_n} de f_θ utilizando núcleos asimétricos se define como

$$\hat{f}_{A_n}(x) = \frac{1}{n} \sum_{i=1}^n K_{s(x,b)}(X_i), \quad (5.2.21)$$

donde b es el ancho de banda, $s(x, b)$ es el vector de parámetros que depende del núcleo usado, y K el núcleo asimétrico.

Es necesario que el núcleo asimétrico:

- a) Sea una función de densidad con soporte en el intervalo $[0, 1]$, $(0, +\infty)$ o en el intervalo $[0, +\infty)$.
- b) Los parámetros de posición y escala sean funciones del punto x donde se realiza la estimación y del ancho de banda b .

La propiedad a) garantiza que la densidad estimada sea siempre positiva. La propiedad b) indica que la forma del núcleo asimétrico cambia de acuerdo al punto donde se realiza la estimación. Esta flexibilidad permite una mejor aproximación en las colas de la distribución.

Como se indica en [29, 47] los núcleos asimétricos presentan el inconveniente que, $\int_{-\infty}^{\infty} \hat{f} dx$ no siempre es igual a 1. Una forma de resolver esto es normalizando por la constante adecuada.

Dentro de los núcleos utilizados se encuentran los propuestos en [10] y en [29].

$$K_{\Gamma^1_{(\frac{x}{b}+1, b)}}(t) = \frac{1}{\Gamma(\frac{x}{b}+1)b^{\frac{x}{b}+1}} t^{-x/b} \exp\{-t/b\}, \quad (5.2.22a)$$

$$K_{\Gamma^2_{(\rho_b(x), b)}}(t) \text{ donde } \rho_b(x) = \begin{cases} x/b & 2b \leq x \\ \frac{1}{4}(x/b)^2 + 1 & x \in [0, 2b), \end{cases} \quad (5.2.22b)$$

$$K_{\text{IG}_{(x, \frac{1}{b})}}(t) = \frac{1}{\sqrt{2\pi b t^3}} \exp\left\{-\frac{1}{2bx}\left(\frac{t}{x} + \frac{x}{t} - 2\right)\right\}, \quad (5.2.22c)$$

$$K_{\text{RIG}_{(\frac{1}{x-b}, \frac{1}{b})}}(t) = \frac{1}{\sqrt{2\pi b t}} \exp\left\{-\frac{x-b}{2b}\left(\frac{t}{x-b} - 2 + \frac{x-b}{t}\right)\right\}, \quad (5.2.22d)$$

$$K_{\text{LN}_{(\log(x)+b^2, b)}}(t) = \frac{1}{t\sqrt{2\pi b}} \exp\left\{-\frac{(\log t - \log x - b^2)^2}{2b^2}\right\}, \quad (5.2.22e)$$

para $t, x, b > 0$.

Se consideraron estos núcleos porque el modelo \mathcal{G}^0 propone a la distribución Γ^{-1} para modelar el backscatter. El modelo \mathcal{G}^H proviene de considerar la distribución IG para modelar el ruido speckle. Además, de acuerdo a lo visto en el capítulo 4.1 la distribución Lognormal fue introducida en [73] y fue propuesta como modelo empírico para describir datos SAR.

~~Se hizo una primera selección de los núcleos estudiando el MISE muestral para distintas combinaciones de los parámetros y tamaños muestrales. Se consideraron valores de $\alpha = \{-1.5, -3, -5, -8\}$, $\gamma = -\alpha - 1$ que es el valor de γ para que $E(X) = 1$, y valores de $n = \{9, 25, 49, 81, 121\}$. Estos son los valores de los parámetros que se analizarán en esta tesis. Se utilizó el método LSCV para encontrar el ancho de banda b . Dado que los valores del MISE son similares para los núcleos Γ^1 y Γ^2 , y para los núcleos IG y RIG respectivamente, se eligieron los núcleos Γ^1 e IG junto con el núcleo LN para abordar el problema de estimación. A partir de ahora el núcleo Γ^1 se llamará Γ .~~

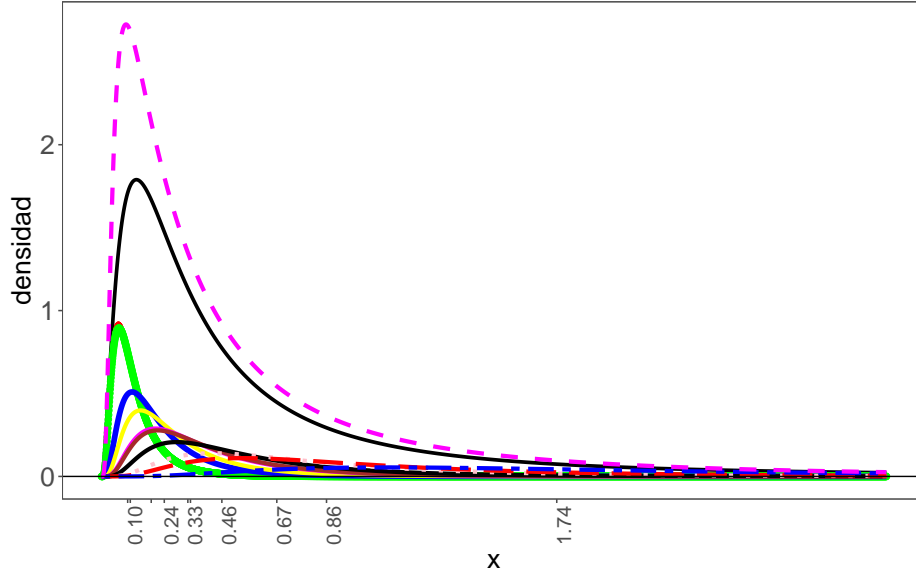


Figura 5.2: Estimación de \mathcal{G}_L^0 para $L = 3$, $\alpha = -1.5$, $\gamma = 0.5$ y $n = 10$

En la figura 5.2 se muestra un ejemplo de estimación de la función de densidad \mathcal{G}_L^0 para el caso de $\alpha = -1.5$, $\gamma = 0.5$, $L = 3$ y $n = 10$, utilizando núcleo LN

donde se utilizó un ancho de banda que se obtiene de aplicar el método LSCV explicado en la subsección 5.2.3. La línea sólida es la función de densidad teórica, la línea discontinua es la estimación \hat{f} obtenida, el resto son los distintos núcleos. Para cada x el valor de $\hat{f}(x)$ se obtiene como promedio de los distintos núcleos en ese x . Se eligió un tamaño de muestra chico para ejemplificar la metodología de estimación. Vale la pena observar que el parámetro de posición depende del valor de x donde se realiza la estimación. Por eso estos núcleos cambian de forma dependiendo del valor x .

En la figura 5.3 se muestra cómo estiman los núcleos K_Γ , K_{LN} y K_{IG} al modelo \mathcal{G}_I^0 para un tamaño de muestra $n = 25$, $L = 8$ utilizando también el método LSCV para elegir el ancho de banda. Se eligieron estos núcleos porque son los que finalmente evaluaremos. Esta elección se justificará en el capítulo 6.

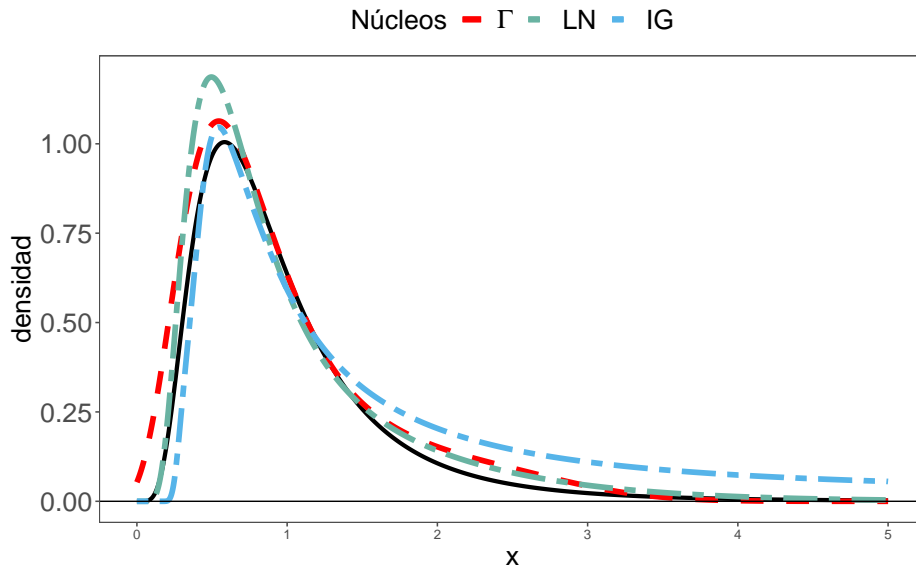


Figura 5.3: Estimación de \mathcal{G}_I^0 para $L = 8$, $\alpha = -5$, $\gamma^* = 4$ y $n = 25$ utilizando núcleos K_Γ , K_{LN} y K_{IG} .

Propiedades globales

Para poder estudiar propiedades globales de estos estimadores como el MISE, se asume que:

a) f es dos veces diferenciable para todos los núcleos considerados.

b) $\int_0^\infty f'^2(x)dx$ y que $\int_0^\infty (xf''(x))^2 dx$ son finitas para el caso de núcleo Γ .

c) $\int_0^\infty (x^3 f''(x))^2 dx$ es finita para el caso de los núcleos IG.

d) $\int_0^\infty \frac{f'(x)}{x} dx$ y que $\int_0^\infty (xf'(x) + x^2 f''(x))^2 dx$ para el caso de núcleo LN.

Es necesario aclarar que la condición b) se cumple para valores del número de looks L mayores a $3/2$.

Llamaremos b^* al valor del ancho de banda que minimiza el MISE, y MISE^* al MISE cuando $b=b^*$. Expresiones para b^* y MISE^* para los distintos núcleos son:

■ Núcleo Gamma

$$b_\Gamma^* = \frac{\left[\frac{1}{2\sqrt{\pi}} \int_0^\infty x^{-1/2} f(x) dx \right]^{2/5}}{4^{2/5} \left[\int_0^\infty \left\{ x f'(x) + \frac{1}{2} x f''(x) \right\}^2 dx \right]^{2/5}} n^{-2/5}, \quad (5.2.23)$$

$$\begin{aligned} \text{MISE}^*(\hat{f}_\Gamma) &= \frac{5}{4^{4/5}} \left[\frac{1}{2\sqrt{\pi}} \int_0^\infty x^{-1/2} f(x) dx \right]^{4/5} \\ &\times \left[\int_0^\infty \left\{ x f'(x) + \frac{1}{2} x f''(x) \right\}^2 dx \right]^{1/5} n^{-4/5}. \end{aligned} \quad (5.2.24)$$

■ Núcleo Inverso Gaussiano

$$b_{\text{IG}}^* = \frac{\left[\frac{1}{2\sqrt{\pi}} \int_0^\infty x^{-3/2} f(x) dx \right]^{2/5}}{\left[\int_0^\infty (x^3 f''(x))^2 dx \right]^{2/5}} n^{-2/5}, \quad (5.2.25)$$

$$\begin{aligned} \text{MISE}_{\text{IG}}^* &= \frac{5}{4} \left[\frac{1}{2\sqrt{\pi}} \int_0^\infty x^{-3/2} f(x) dx \right]^{4/5} \\ &\times \left[\int_0^\infty \{ x^3 f''(x) \}^2 dx \right]^{1/5} n^{-4/5}. \end{aligned} \quad (5.2.26)$$

■ Núcleo Lognormal

$$b_{LN}^* = \left(\frac{\int_0^\infty \frac{f(x)}{x}}{n \int_0^\infty x^2 (3f'(x) + xf''(x))^2} \right)^{1/5}, \quad (5.2.27)$$

$$\begin{aligned} \text{MISE}_{LN}^* &= \frac{3}{4} \left[\int_0^\infty \frac{f(x)}{x} \right]^{4/5} \\ &\times \left[\int_0^\infty x^2 (3f'(x) + xf''(x))^2 \right]^{1/5} n^{-4/5}. \end{aligned} \quad (5.2.28)$$

Por lo tanto, \hat{f}_Γ , \hat{f}_{IG} y \hat{f}_{LN} alcanzan una tasa de convergencia en términos del mise del orden de $n^{-4/5}$, comparabale con la que alcanzan otros núcleos simétricos.

Ancho de banda

La elección del ancho de banda b es un punto crucial en estimación no paramétrica de la función de densidad. Si b es muy pequeño la función de densidad estimada puede ser muy ruidosa, presentando varios picos. Si en cambio, el valor de b es demasiado grande, entonces la estimación puede pasar por alto características clave porque la variabilidad disminuye.

En esta tesis se estudiarán tres estrategias para determinar el ancho de banda b .

- El b_{opt} que es el valor de b que minimiza el MISE. Expresiones de b_{opt} se darán en la siguiente subsección.
- Un valor específico de b encontrado empíricamente.
- El método Least Squared Cross Validation (LSCV) ampliamente utilizado en la literatura.

El método LSCV consiste en encontrar el valor de b que minimiza el ISE.

Observemos que, dada una muestra,

$$\begin{aligned} \text{ISE}(b) &= \int_{\mathbb{R}} (\hat{f}_b - f)^2 dx \\ &= \int_{\mathbb{R}} \hat{f}_b^2 dx - 2 \int_{\mathbb{R}} \hat{f}_b f dx + \int_{\mathbb{R}} f^2 dx. \end{aligned} \quad (5.2.29)$$

El último término de la ecuación (5.2.29) no depende de b . Entonces

$$\min_b \text{ISE}(b) = \min_b \phi(\hat{f}_b) \text{ donde ,}$$

donde

$$\phi(\hat{f}) = \int_{\mathbb{R}} \hat{f}^2 dx - 2 \int_{\mathbb{R}} \hat{f} f dx. \quad (5.2.30)$$

Notemos que el segundo término de la ecuación (5.2.30) se puede ver como $E_X(\hat{f}_b(x))$ con $X \sim f$ desconocida. Por lo tanto se puede estimar $E_X(\hat{f}_b(x))$ por medio de $\hat{E}_X(\hat{f}_b(x)) = \frac{1}{n} \sum_{i=1}^n \hat{f}_b(X_i)$. Observemos que la observación X_i donde se evalúa \hat{f} es utilizada también para encontrar \hat{f}_b . Una forma de asegurar independencia entre X_i y \hat{f}_b es considerar el leave-one out estimador:

$$\hat{E}_X(\hat{f}_b(x)) = \frac{1}{n} \sum_{i=1}^n \hat{f}_{b,-i}(X_i), \quad (5.2.31)$$

donde es el estimador de $\hat{f}_{b,-i}$ sin considerar la i^{th} observación. De esta forma nos aseguramos que las observaciones utilizadas para calcular $\hat{f}(\cdot)$ son independientes de X_i .

El valor de b_{LSCV} óptimo que propone este método es aquel valor que minimiza la función de cross validación

$$\text{LSCV}(b) = \int_{\mathbb{R}} \hat{f}^2(x) dx - \frac{2}{n} \sum_{i=1}^n \hat{f}_{-i}(X_i). \quad (5.2.32)$$

Hall and Marron [44] muestran que, para el caso de núcleos simétricos, esta fun-

ción puede tener mínimos locales. Entonces el b_{LSCV} será el más grande de los mínimos locales porque tiene un mejor comportamiento empírico que el mínimo global.

En la figura 5.4 se puede ver la función de cross validación LSCV en función del ancho de banda, usando núcleo Γ , para $L = 3$, $\alpha = \{-3, -5\}$ y $n = \{9, 25, 49, 81, 121\}$. Se puede ver que, en todos los casos, esta función presenta un único mínimo.

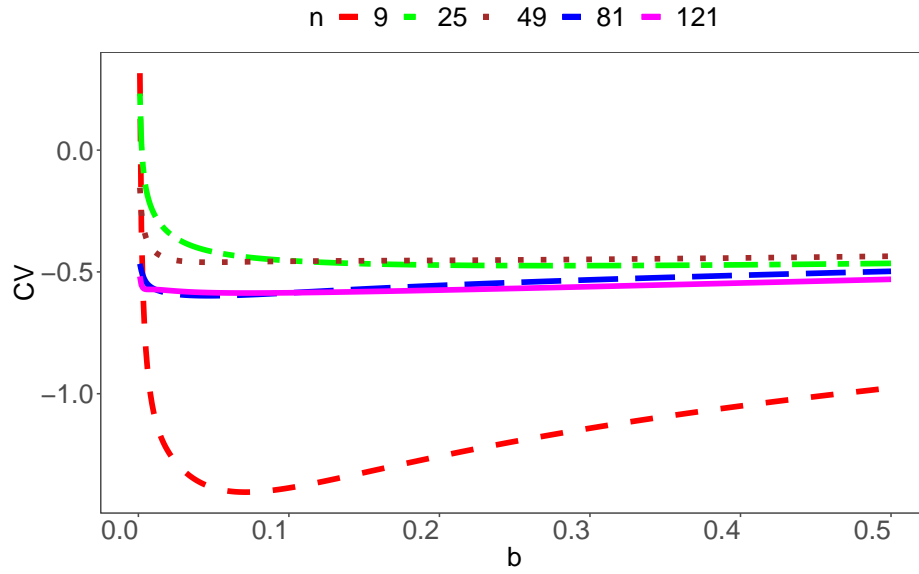


Figura 5.4: Least Square Cross Validation (LSCV) para $L = 3$ y $\alpha = -5$

Convergencia

Es interesante contar con resultados de convergencia de un estimador para medir la calidad del mismo. En [29] el autor presenta propiedades asintóticas del estimador \hat{f}_{A_n} de f , donde \hat{f}_{A_n} es el estimador de núcleos asimétricos utilizado en esta tesis. El autor demuestra resultados de consistencia y de normalidad asintótica, como así también demuestra propiedades de convergencia globales. Algunos de estos resultados son necesarios para probar la consistencia del estimador del parámetro de textura de la distribución \mathcal{G}_I^0 que se propone en esta tesis y que presentaremos en la sección 5.2.4.

■ **Convergencia fuerte de \widehat{f}_{A_n} .**

Teorema 5.2.1. Sea $f \in C^2(\text{Sop}(f))$ la densidad a estimar, donde $\text{Sop}(f)$ es el soporte de f , y \widehat{f}_{A_n} su estimador de núcleo. Para cada punto x fijo en $\text{Sop}(f)$ tenemos que

$$\widehat{f}_{A_n}(x) \xrightarrow{c.s.} f(x) \text{ cuando } n \rightarrow \infty. \quad (5.2.33)$$

Si se considera un conjunto de condiciones más fuertes sobre el ancho de banda se obtienen resultados sobre la convergencia en L_1 sobre conjuntos compactos.

■ **Convergencia en L_1 de \widehat{f}_{A_n} .**

Teorema 5.2.2. Sea f una función de densidad continua y acotada en $\text{Sop}(f)$, sea \widehat{f}_{A_n} el estimador de f utilizando núcleos definidos en [29] dentro de los cuales se encuentran Γ , IG y LN . Supongamos que existe un número real $r_1 = r_1(K) > 0$ que depende del núcleo elegido tal que, para cada $x \in \text{Sop}(f)$

$$b_n^{r_1} \int_0^\infty |dK_{x,b_n}(s)| \leq c_1(x), \quad (5.2.34)$$

donde $c_1(x)$ es una función integrable en cualquier compacto que contenga a $x \in \text{Sop}(f)$. Entonces si $\lim_{n \rightarrow \infty} b_n = 0$ y

a) $\lim_{n \rightarrow \infty} n b_n^{2r_1} = +\infty$, entonces

$$\int_0^{+\infty} |\widehat{f}_{A_n}(x) - f(x)| dx \xrightarrow{p} 0 \text{ cuando } n \rightarrow \infty. \quad (5.2.35)$$

b) $\lim_{n \rightarrow \infty} \frac{nb^{2r_1}}{\ln n} = +\infty$, entonces

$$\int_0^{+\infty} |\hat{f}_{A_n}(x) - f(x)| dx \xrightarrow{c.s.} 0 \text{ cuando } n \rightarrow \infty, \quad (5.2.36)$$

con $r_1 = 1$ para los núcleos Γ , $r_1 = 5/2$ para el núcleo IG y $r_1 = 2$ para el núcleo LN.

El resultado 5.2.36 es esencial para demostrar la consistencia fuerte del estimador de mínima distancia para el parámetro de textura α de la distribución \mathcal{G}_I^0 como veremos en sección 5.2.4.

5.2.4. Estimadores de Mínima Distancia

Los estimadores de mínima distancia (MDE) es una metodología no paramétrica que permite encontrar nuevos estimadores de los parámetros de una distribución. Su propósito es encontrar los valores de los parámetros que hacen que el modelo teórico esté lo más cerca posible de la información que provee la muestra. En este sentido hay dos elementos involucrados: uno es la medida de distancia que permite cuantificar la proximidad, el otro es la información que proviene de la muestra.

Teniendo en cuenta ambos elementos, la forma habitual de definir el estimador MDE es

$$\hat{\theta}_n = \underset{\theta \in \Omega}{\operatorname{argmin}} D(F_\theta, F_n), \quad (5.2.37)$$

donde F_n es la función de distribución empírica, F_θ es la función de distribución teórica, Ω es el espacio paramétrico y D es una medida de distancia entre las funciones de distribución que, usualmente, es una métrica.

Mucho se ha investigado respecto de estos estimadores. Ya en la década del 50 Wolfowitz [99, 101] estudió esta clase de estimadores y demostró que la sucesión

de estimadores $\hat{\theta}_n$ definida en 5.2.37 es fuertemente consistente bajo ciertas condiciones generales, cuando la distancia D es la distancia supremo entre la función de distribución empírica y la distribución teórica. Boos en [9] estudió esta clase de estimadores donde utiliza una medida de distancia ponderada por ciertas funciones de peso. Para determinadas elecciones de estas funciones de peso se obtiene la distancia de Cramer-von Mises, de Anderson-Darling entre otras. El autor demostró que estos estimadores son consistentes y asintóticamente eficientes bajo ciertas condiciones para las funciones de peso.

Hettmansperger et al. [46] estudiaron los MDE estimadores utilizando la distancia de Cramer-von Mises. Los autores demostraron que estos estimadores son asintóticamente normales y tienen buenas propiedades de eficiencia y buenas propiedades de robustez. Parr and Schucany [76] probaron que, bajo ciertas condiciones el estimador de mínima distancia entre la función de distribución empírica y el modelo teórico es fuertemente consistente bajo una medida de discrepancia general poniendo algunas condiciones sobre esta medida. Muchas de estas medidas son utilizadas como estadísticos para test de bondad de ajuste basados en la función de distribución empírica. También mencionan que los MDE estimadores comparten la propiedad de invarianza que tienen los MLE estimadores, en el sentido que $\widehat{g(\theta)}_{\text{MDE}} = g(\widehat{\theta}_{\text{MDE}})$. Posteriormente, Parr and Schucany [75] generalizan el resultado de consistencia dando condiciones para asegurar la unicidad del mínimo del MDE estimador.

Dado que, en general, F_θ es una distribución absolutamente continua, se pueden pensar a los estimadores MDE en términos de la función de densidad. Pensando en este sentido estos estimadores se pueden definir, de una forma más general, como

$$\hat{\theta}_n = \operatorname{argmin}_{\theta \in \Omega} d(f_\theta, \hat{f}_n), \quad (5.2.38)$$

donde d es una medida de discrepancia, discimilarity o distancia estocástica, entre la función de densidad teórica f_θ y un estimador no paramétrico \hat{f}_n de la

densidad teórica.

¿Por qué proponer un nuevo estimador si se conocen las buenas propiedades asintóticas que tiene el MLE estimador? Justamente, un buen estimador debe tener un buen comportamiento si los datos provienen del modelo asumido, pero además no debe sufrir grandes desvíos si los datos sufren perturbaciones que hacen que el verdadero modelo no sea completamente válido.

Beran [7] señala que, en general y haciendo referencia al trabajo de Huber [49], el MLE estimador no es estable bajo pequeñas perturbaciones del modelo asumido. Por este motivo el autor propone el estimador MDE utilizando la distancia de Hellinger entre \hat{f}_n y la función de densidad teórica, estimando \hat{f}_n con un estimador de núcleos clásico. Estudió propiedades asintóticas y el desempeño de estos estimadores frente a pequeños desvíos del modelo asumido, mostrando que su propuesta es estable frente a estos desvíos. Eslinger et al. [32] implementó y realizó un estudio empírico del estimador propuesto por Beran para estudiar la performance de este estimador para pequeñas muestras en términos de eficiencia, distribución muestral y robustez. El autor compara este estimador con otros dentro de los cuales se encuentra el MLE estimador.

Cabe señalar que uno de los puntos de interés en esta tesis es encontrar un estimador del parámetro de textura de la distribución \mathcal{G}_I^0 que tenga un buen comportamiento para pequeños tamaños de muestra. Este aspecto es interesante de tener en cuenta porque mucho de los métodos de filtrado de imágenes o detección de bordes utilizan máscaras deslizantes para estimar parámetros, las cuales suelen ser de tamaño 3×3 , 5×5 , 7×7 , 9×9 y 11×11 . En este sentido es conocido que el MLE estimador puede no presentar un buen comportamiento cuando el tamaño de muestra es pequeño. Por este motivo es necesario encontrar estimadores que tengan un comportamiento asintótico similar al MLE, que mejore la performance del MLE para pequeños tamaño de muestras y para desvíos del modelo teórico.

Por lo anteriormente expuesto en esta tesis se propone un estimador del tipo MDE para estimar el parámetro de textura de la distribución \mathcal{G}_I^0 donde se utilizan

distancias estocásticas para medir la discrepancia entre el modelo teórico y una estimación no paramétrica de la función de densidad subyacente.

Las distancias estocásticas que serán estudiadas son:

- Distancia de Hellinger $d_H(f_V, f_W) = 1 - \int_{-\infty}^{\infty} \sqrt{f_V f_W}$.
- Distancia de Bhattacharyya $d_B(f_V, f_W) = -\log \int_{-\infty}^{\infty} \sqrt{f_V f_W}$.
- Distancia Triangular $d_T(f_V, f_W) = \int_{-\infty}^{\infty} \frac{(f_V - f_W)^2}{f_V + f_W}$.

Observar que, de acuerdo a la definición dada en 5.6, $d_T(f_V, f_W) = d_{T,2}(f_V, f_W)$.

- Distancia de Rényi con parámetro $\beta \in (0, 1)$

$$d_R^\beta(f_V, f_W) = \frac{1}{2(\beta - 1)} \log \int_{-\infty}^{\infty} (f_V^\beta f_W^{1-\beta}) + \log \int_{-\infty}^{\infty} (f_V^{1-\beta} f_W^\beta).$$

Dado que las distancias de Hellinger y Bhattacharyya verifican $d_B = -\log(1 - d_H)$, y la función logaritmo es creciente resulta que $\arg \min_\alpha d_B(\alpha) = \arg \min_\alpha d_H(\alpha)$. Por lo tanto estas distancias poseen el mismo mínimo y por eso utilizaremos solamente la distancia de Hellinger que está más estudiada en la literatura.

5.3. Conclusiones

En este capítulo se introducen los elementos metodológicos que se utilizan en esta tesis. Se presentan los conceptos de estimación paramétrica y no paramétrica de los parámetros de una función de densidad. En este sentido se describen los estimadores paramétricos existentes en la literatura y se introduce la estimación por núcleos, tanto simétricos como asimétricos, de la función de densidad subyacente. Se presentan las (h, ϕ) divergencias como medida de discrepancia entre dos funciones de densidad y, en particular, se estudia la distancia triangular. Se introducen los estimadores de mínima distancia como metodología no paramétrica para estimar los parámetros de una distribución. Finalmente se presenta el

aporte principal de esta tesis que es la propuesta de un nuevo estimador para el parámetro de textura de la distribución \mathcal{G}_I^0 basada en la minimización de una distancia estocástica entre el modelo teórico y una estimación no paramétrica de la función de densidad de probabilidad subyacente utilizando núcleos asimétricos.

Capítulo 6

Simulaciones y aplicación a una imagen real

En este capítulo se analizan varias estrategias para la estimación del parámetro de textura de la distribución \mathcal{G}_I^0 . Se ha demostrado que esta distribución es capaz de caracterizar un gran número de objetivos en las imágenes SAR, en particular, las monopolarizadas que son las estudiadas en esta tesis. La distribución está caracterizada por tres parámetros: el número de looks L que se puede estimar en toda la imagen, un parámetro de escala γ , y un parámetro de textura α . Este último está estrechamente relacionado con el número de retrodispersores elementales para cada píxel, esto es una de las razones por las que recibe atención en la literatura.

Otras de las razones es que el parámetro α de la distribución \mathcal{G}_I^0 puede interpretarse en términos de la textura de la imagen. Valores cercanos a 0, típicamente $\alpha \in [-3, 0)$ sugieren zonas extremadamente texturadas como zonas urbanas. Conforme el valor de α decrece disminuye la heterogeneidad de la imagen. Regiones con moderada textura como son los bosques, están representadas por valores de $\alpha \in [-6, -3)$, zonas de pastura son compatibles con valores de $\alpha \in (-\infty, -6)$. Esta es otra de las razones por las cuales es muy importante obtener estimaciones precisas de este parámetro.

Entre las técnicas de estimación disponibles, el método de los momentos y el método de Máxima Verosimilitud son las técnicas clásicas de estimación paramétrica. En [8] los autores proponen un método basado en momentos fraccionarios de bajo orden para estimar la matriz de covarianza en una imagen SAR polarimétrica bajo una distribución alfa estable. En [24, 97] los autores utilizan técnicas de remuestreo para mejorar el rendimiento de estos estimadores en términos de sesgo y error cuadrático medio, a expensas de un alto costo computacional.

Más recientemente [5, 15, 58] los estimadores basados en la transformada de Mellin, como lo son los Logcumulantes y Logmomentos estimadores, han sido implementados con éxito. Tison et al. [96] mostraron que los estimadores basados en Logcumulantes tienen una mejor performance respecto del MV estimador para el modelo \mathcal{G}^0 para datos de amplitud. Sin embargo estos estimadores no han sido evaluados para el caso de datos de intensidad ni bajo la presencia de contaminación.

Una propiedad que es deseable que un estimador posea es ser resistente a la presencia de datos atípicos. Bustos et al. [16] y Allende et al. [2] estudiaron la performance del MV estimador bajo el modelo \mathcal{G}^0 para datos de amplitud, mostrando que hay una pérdida de robustez ante la presencia de outliers moderados en áreas extremadamente texturadas, mientras que para zonas texturadas la pérdida de robustez ocurre en presencia de outliers severos. Asimismo esta falta de robustez es más pronunciada en el caso de muestras de pequeño tamaño. Ellos propusieron como alternativa los M y AM estimadores que tienen un buen comportamiento bajo contaminación pero presentan problemas numéricos especialmente para tamaños de muestras chicos.

La idea principal de este trabajo es presentar un nuevo método de estimación para el parámetro de textura del modelo \mathcal{G}_I^0 que tenga buenas propiedades midiéndolas en término de su sesgo, su error cuadrático medio, su capacidad para resistir diferentes niveles de contaminación y de bajo costo computacional. Estas propiedades serán estudiadas incluso para muestras de tamaño pequeño y

moderado.

Para lograr esta tarea en esta tesis se propone como estimador del parámetro de textura al valor del argumento que minimiza la distancia estocástica entre el modelo teórico y una estimación no paramétrica de la función de densidad subyacente, es decir,

$$\hat{\theta}_n = \operatorname{argmin}_{\theta \in \Theta} d(f_\theta, \hat{f}_n) \quad (6.0.1)$$

donde Θ es el espacio paramétrico y d es una medida de discrepancia, discimilarity o distancia estocástica, entre la función de densidad teórica f_θ y un estimador no paramétrico \hat{f}_n de la función de densidad teórica.

Esta propuesta está basada en dos pilares: la forma de estimar \hat{f}_n y la distancia estocástica elegida. En este capítulo daremos argumentos que sugieren que los núcleos asimétricos ofrecen una buena alternativa para estimar la función de densidad subyacente en forma no paramétrica, en particular los núcleos Gamma y Lognormal son una buena elección. También proponemos el método para elegir el ancho de banda, como así también mostraremos que la distancia triangular es una buena medida de discrepancia para este caso.

6.1. Generando muestras \mathcal{G}_I^0

Para generar muestras que provengan de una distribución \mathcal{G}_I^0 vamos a explotar la naturaleza multiplicativa del modelo que la genera.

De acuerdo a 4.1 la distribución \mathcal{G}_I^0 es el producto de las distribuciones $\Gamma^{-1}(-\alpha, \gamma)$ y $\Gamma(L, L)$ donde $-\alpha, \gamma > 0$ y $L \geq 1$. Además, se sabe que si $X \sim \Gamma(a, \beta)$, entonces $Y = 1/X \sim \Gamma^{-1}(a, \beta)$.

Luego, para generar una muestra Z_1, \dots, Z_n de variables aleatorias independientes e idénticamente distribuidas con $Z_i \sim \mathcal{G}_I^0(\alpha, \gamma, L)$ con $i = 1, \dots, n$ hacemos lo siguiente:

- Fijamos $a = -\alpha, \gamma$ ambos positivos y $L \geq 1$.

- Generamos una muestra de tamaño n de una variable aleatoria $W \sim \Gamma(a, \gamma)$.
- Definimos $X = 1/W$, entonces $X \sim \Gamma^{-1}(a, \beta)$.
- Generamos una muestra de tamaño n de una variable aleatoria $Y \sim \Gamma(L, L)$.
- Obtenemos $Z \sim \mathcal{G}_I^0(\alpha, \gamma, L)$ definiendo $Z = X \cdot Y$.

6.2. Eligiendo distancias estocásticas

El primer objetivo de este trabajo fue elegir una distancia entre las posibles medidas de discimilarity que ofrece la literatura, para luego elegir el estimador del parámetro de textura a través de la minimización de dicha distancia entre la función de densidad teórica del modelo \mathcal{G}_I^0 y una estimación no paramétrica de la función de densidad subyacente. Como se mencionó en el capítulo 5, se consideraron las siguientes distancias estocásticas:

- Distancia de Hellinger $d_H(f_v, f_w) = 1 - \int_{-\infty}^{\infty} \sqrt{f_v f_w}$.
- Distancia de Bhattacharyya $d_B(f_v, f_w) = -\log \int_{-\infty}^{\infty} \sqrt{f_v f_w}$.
- Distancia Triangular $d_T(f_v, f_w) = \int_{-\infty}^{\infty} \frac{(f_v - f_w)^2}{f_v + f_w}$.
- Distancia de Rényi con parámetro $\beta \in (0, 1)$

$$d_R^\beta(f_v, f_w) = \frac{1}{2(\beta - 1)} \log \int_{-\infty}^{\infty} (f_v^\beta f_w^{1-\beta}) + \log \int_{-\infty}^{\infty} (f_v^{1-\beta} f_w^\beta).$$

donde V y W son dos variables aleatorias definidas sobre el mismo espacio de probabilidad cuyas funciones de densidad son f_v y f_w . Recordemos que por lo visto en la sección 5.2.4 las distancias de Hellinger y Bhattacharyya poseen el mismo mínimo, por este motivo solamente se analizará la distancia de Hellinger.

Para cumplir con el objetivo inicialmente planteado comenzamos calculando las curvas de distancias entre las funciones de densidad de la distribución

$\mathcal{G}_I^0(\alpha_0, 1 - \alpha_0, L)$ y $\mathcal{G}_I^0(\alpha, 1 - \alpha, L)$, $\alpha < -1$, para distintos valores de α_0 y de L . Recordemos que vamos a considerar que el parámetro de escala cumpla la condición $\gamma^* = -\alpha - 1$ para que la $E(Z) = 1$ donde $Z \sim \mathcal{G}_I^0(\alpha, \gamma, L)$.

Las figuras 6.1, 6.2 y 6.3 muestran los gráficos de dichas curvas, para los casos donde el número de looks $L = 1, 3, 8$, el valor de $\alpha_0 = -2, -3, -4, -5, -6, -7$ y las distancias de Hellinger, Rényi y Triangular.

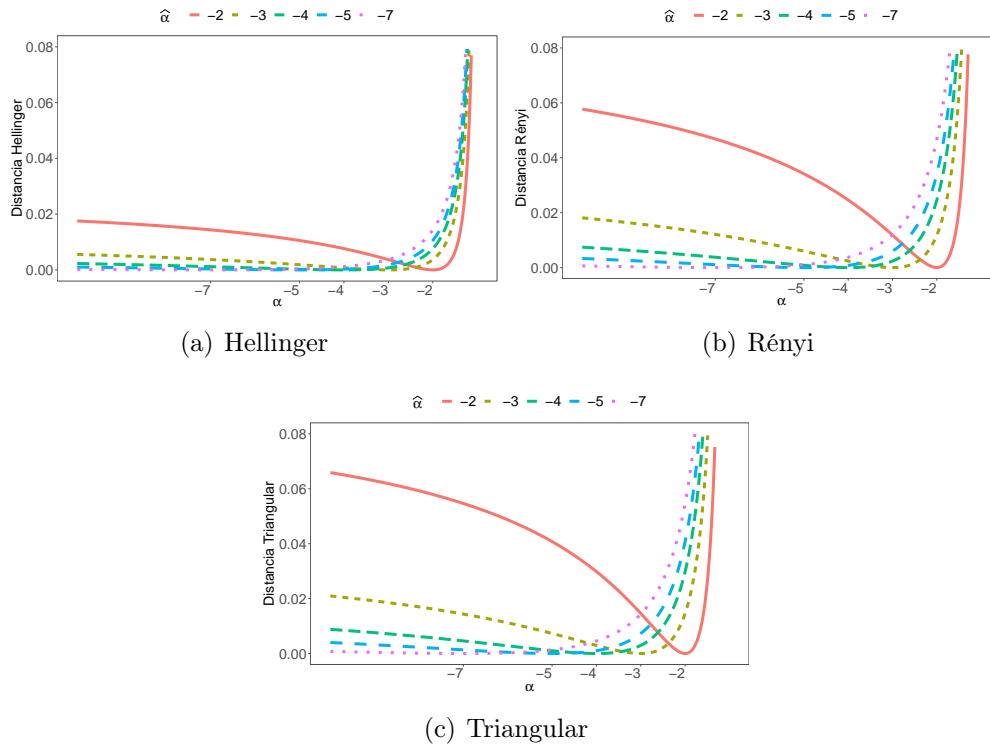


Figura 6.1: Distancias Hellinger, Renyi y Triangular, $\alpha_0 = -2, -3, -4, -5, -6, -7$ y $L = 1$.

En estos gráficos se puede observar que, cuanto menor es el valor de α_0 la curva se hace más plana. Por este motivo los métodos de optimización resultan más inestables y el mínimo es más difícil de encontrar en forma precisa. Se puede observar también que, a medida que aumenta el valor de L , esta situación se revierte y es posible encontrar el mínimo de una forma más eficiente. Recordemos que el número de looks representa la relación señal ruido, por lo tanto, cuanto mayor es el número de looks menor ruido presenta la imagen siendo $L = 1$ el caso

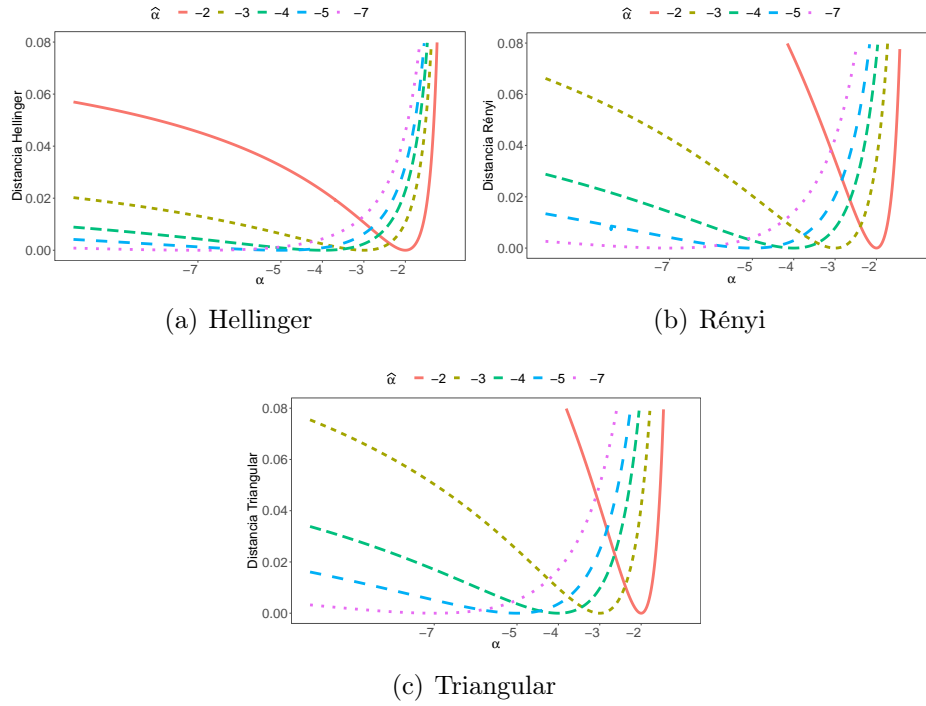


Figura 6.2: Distancias Hellinger, Renyi y Triangular, $\alpha_0 = -2, -3, -4, -5, -6, -7$ y $L = 3$.

de mayor ruido. Se puede observar también que la distancia de Hellinger es la que resulta más plana en un entorno del mínimo, y que las distancias de Rényi y Triangular tienen un comportamiento similar.

Siguiendo en la búsqueda de la mejor distancia se realizaron simulaciones Monte Carlo para elegir la distancia estocástica que mejor performance tiene a la hora de estimar el parámetro de textura de la distribución \mathcal{G}_I^0 . En Cassetti et al. [19] se evaluaron dichas distancias pero ahora estimando el parámetro de textura y comparando la calidad de estos estimadores con el MV estimador en términos del sesgo y del error cuadrático medio.

Se realizaron 1000 replicaciones del experimento que consiste en generar muestras Z_1, Z_2, \dots, Z_n de variables aleatorias independientes e idénticamente distribuidas donde $Z_i \sim f_{\mathcal{G}_I^0(\alpha, \gamma^*, L)}$ donde el espacio paramétrico está formado por:

- Textura: $\alpha = \{-1.5, -3, -5, -8\}$. Con estos valores se describen regiones extramadamente texturadas ($\alpha = -1.5$), texturadas ($\alpha = \{-3, -5\}$) y

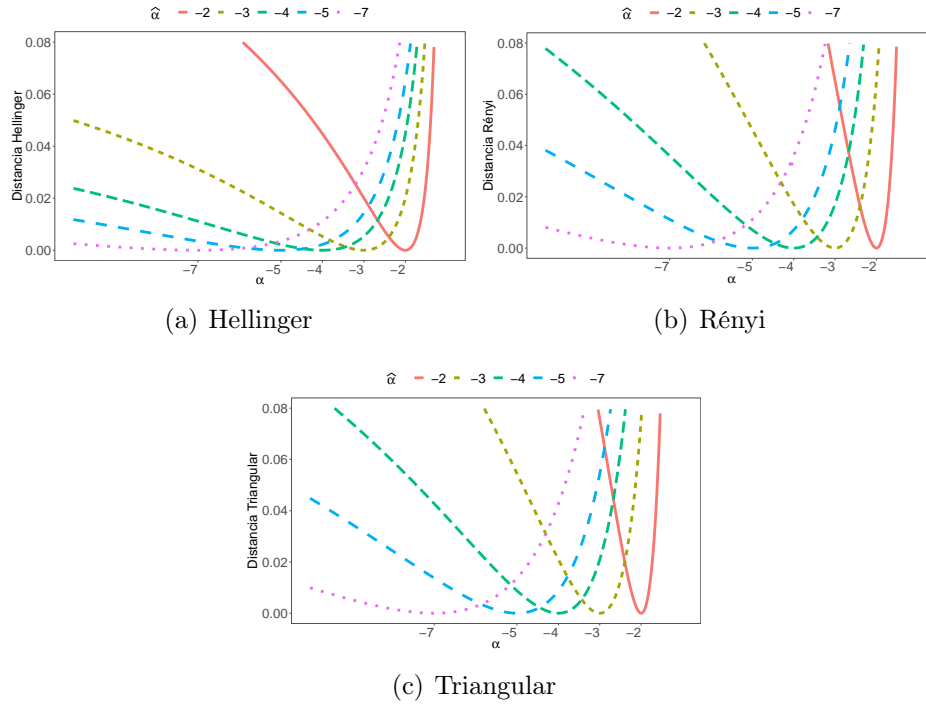


Figura 6.3: Distancias Hellinger, Renyi y Triangular, $\alpha_0 = -2, -3, -4, -5, -6, -7$ y $L = 8$.

homogéneas ($\alpha = -8$).

- Looks: $L = \{1, 3, 8\}$, para modelar varios niveles de procesamiento.
- Tamaño de muestra: $n = \{9, 25, 49, 81, 121, 1000\}$.

Como se mencionó anteriormente la elección de estos tamaños de muestra se basa en que muchos de los métodos de filtrado de imágenes o detección de bordes utilizan ventanas deslizantes para estimar los parámetros, las cuales suelen ser de tamaño 3×3 , 5×5 , 7×7 , 9×9 y 11×11 . Se consideró también el tamaño de la muestra $n = 1000$ para poder estudiar el comportamiento del estimador cuando el tamaño de muestra es grande. Se sigue considerando el parámetro de escala γ^* para que los resultados sean comparables. Además esta relación nos brinda la posibilidad de llevar un problema de estimación de dos parámetros a estimar un sólo parámetro.

En cada replicación:

- se estima la función de densidad subyacente \hat{f} utilizando histogramas junto con el método de Freedman Diaconis para elegir el ancho de banda b dado por $b = 2 \text{ IQR}(z_1, \dots, z_n) n^{-1/3}$ donde IQR es el rango intercuartil.
- se calcula $\hat{\alpha} = \underset{\alpha}{\operatorname{argmin}} d_D(f_{\mathcal{G}_I^0(z, \alpha, 1-\alpha, L)}, \hat{f})$ donde D es alguna de las distancias definidas en 6.2.

De esta manera se obtuvieron 1000 valores estimados de α , $\{\hat{\alpha}_1, \dots, \hat{\alpha}_{1000}\}$. Con estos valores se estimó la esperanza, el sesgo y el error cuadrático medio que están definidos por:

- $\bar{\alpha} = (1000)^{-1} \sum_{i=1}^{1000} \hat{\alpha}_i$
- $\hat{B}(\hat{\alpha}) = \bar{\alpha} - \alpha$
- $\widehat{\text{ECM}} = (1000)^{-1} \sum_{i=1}^{1000} (\hat{\alpha}_i - \alpha)^2$

Las figuras 6.4(a), 6.5(a) y 6.6(a) muestran el sesgo del estimador propuesto para el parámetro de textura ($\hat{\alpha}$) utilizando las distancias de Hellinger, Rényi, Triangular y también Máxima Verosimilitud, para distintos tamaños de muestras n y número de looks L . La línea recta azul representa el eje x .

En las figuras 6.4(b), 6.5(b) y 6.6(b) se muestra el error cuadrático medio (ECM) estimado para las mismas combinaciones de parámetros. Los resultados se muestran en escala semilogarítmica para su mejor comprensión visual. Cabe señalar que la minimización se realizó recorriendo un rango de valores de alfa variando entre -10 y -1 con paso 0.1 .

Se grafica también el intervalo de confianza de nivel aproximado 95 % para cada una de las estimaciones realizadas. Se utilizó el método del percentil que está en descripto [13]. Este método consiste en utilizar como límites superior e inferior del intervalo de confianza a los percentiles $(\alpha/2)100\%$ y $(1 - \alpha/2)100\%$ de la distribución de $\hat{\alpha}$ obtenida con las estimaciones en cada una de las muestras generadas. Este método fue evaluado por [13] en el caso donde la distribución subyacente es una exponencial donde falla la simetría de la distribución. El autor

muestra que este intervalo tiene un mejor desempeño que el intervalo de confianza utilizando el método de aproximación normal. El método de percentiles se aplica en cada uno de los intervalos construidos en esta tesis.

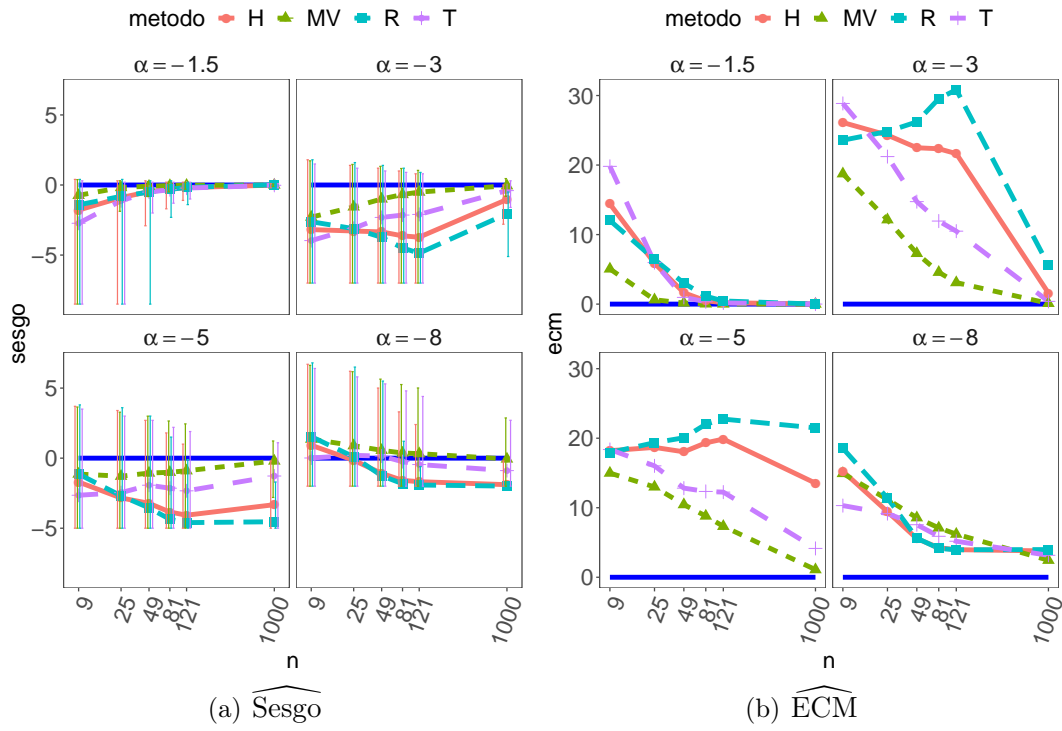
Se puede observar que, en la mayoría de los casos estudiados, los estimadores del parámetro α utilizando distancia triangular y el método de Máxima Verosimilitud ($\hat{\alpha}_{DT}$ y $\hat{\alpha}_{MV}$ respectivamente) tienen un comportamiento similar en término de sesgo y ECM. Mientras que el estimador utilizando las distancias de Hellinger y Renyi ($\hat{\alpha}_{DH}$ y $\hat{\alpha}_{DR}$ respectivamente) tienen una performance parecida.

En el gráfico de $L = 1$ se observa que las distancias de Hellinger y Rényi muestran mayor ECM para tamaños de muestra mayores que 9, salvo para zonas homogéneas donde el ECM es similar en todas las distancias estudiadas. Asimismo estas distancias muestran un mayor sesgo para muestras grandes.

Para el caso de $L = 3$ y $L = 8$ vemos que, en algunos casos el $\hat{\alpha}_{DT}$ mejora la performance del $\hat{\alpha}_{MV}$. Se puede observar que las otras distancias presentan un sesgo y un ECM mayor para zonas texturadas y homogéneas. Asimismo, en estos gráficos se observa que los métodos de minimización de distancias poseen menor ECM para valores grandes del número de looks.

En Casseti et al. [20] se aplicó la metodología anterior a una imagen SAR real. En este caso se utilizó una imagen E-SAR [48] de un look de los alrededores de Munich, banda L, polarización HH, con $L = 1$ en formato intensidad. E-SAR es un sistema formado por sensores y algoritmos de procesamiento de señales desarrollado, entre otras instituciones, por el centro aeroespacial Deutsches Zentrum für Luft- und Raumfahrt German (DLR), un centro de investigación cuya área principal es la teledetección por microondas. Este centro utiliza, mayoritariamente, sistemas de radar de apertura sintética (SAR) tanto en plataformas espaciales como aerotransportadas. El compromiso de la institución en los proyectos internacionales ERS-1 y SIR-100 / X-SAR dio origen a los conocidos sistemas aerotransportados SAR y E-SAR.

En las figuras 6.7(a) y 6.7(b) se muestran dos imágenes E-SAR, de un look,

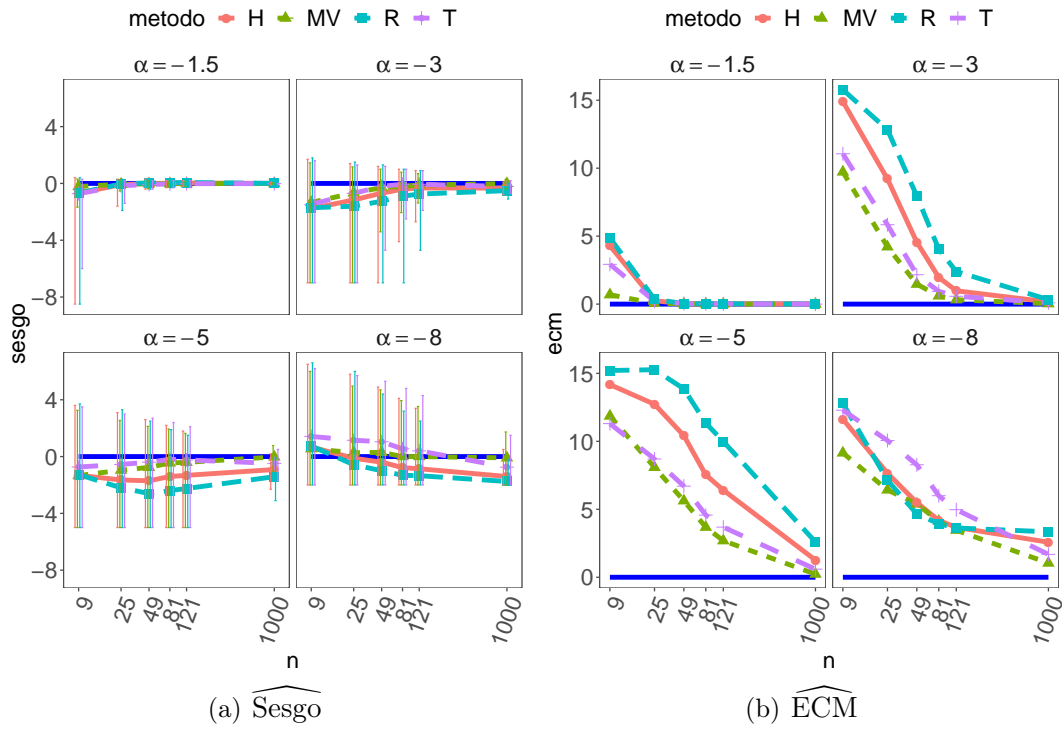
Figura 6.4: $\widehat{\text{Sesgo}}$ y $\widehat{\text{ECM}}$ de $\hat{\alpha}$, $L=1$.

que se utilizaron para testear la performance de las distancias estudiadas y el MV estimador. En ambas figuras se aplicó el método para cada pixel, usando ventanas deslizantes de tamaño 7×7 y 3×3 respectivamente. Se puede observar que los resultados son prometedores ya que se analizó el caso de mayor ruido donde todos los métodos presentaron un comportamiento similar.

Por lo anteriormente expuesto, especialmente por el resultado de las simulaciones, elegimos la distancia Triangular para continuar con el estudio.

6.3. Incorporando núcleos y evaluando robustez

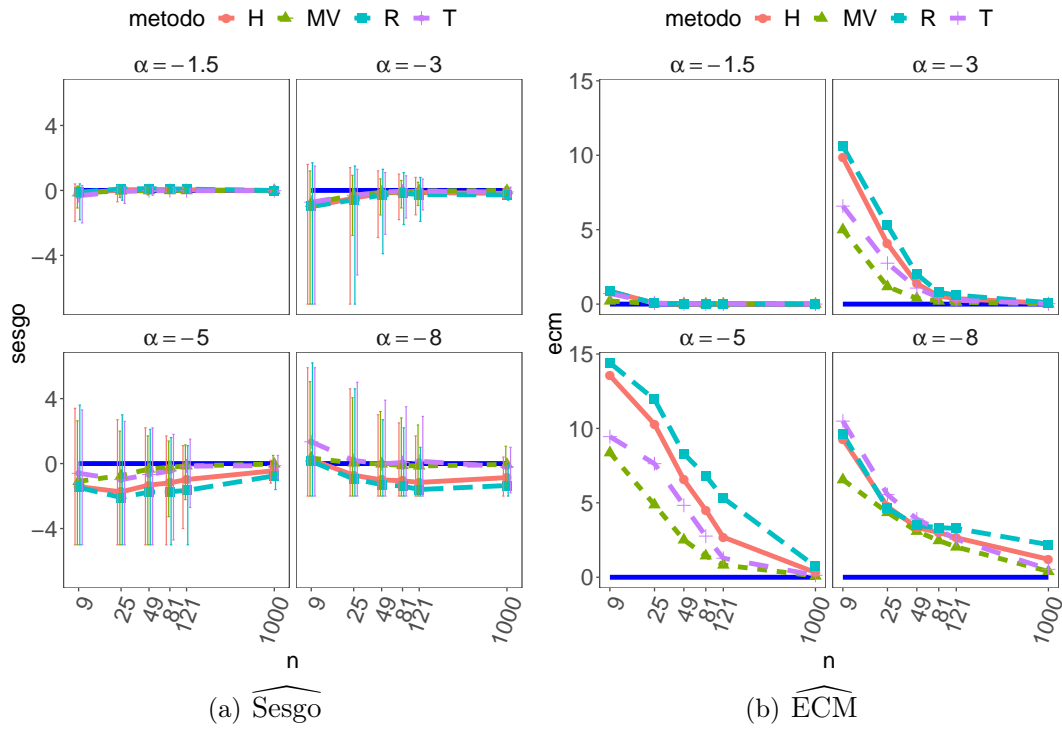
Como se mencionó en la sección 6.2, en Casseti et al. [20] se compararon estimadores basados en la distancia de Hellinger, Bhattacharyya, Renyi y Triangular con el MV estimador. Se mostró que la distancia Triangular es la mejor opción para definir el MDE estimador. En ese trabajo se utilizaron histogramas para

Figura 6.5: $\widehat{\text{Sesgo}}$ y $\widehat{\text{ECM}}$ de $\hat{\alpha}$, $L=3$.

estimar la función de densidad subyacente f_s . En [40] presentamos mejoras con respecto a los resultados obtenidos en [20] ya que:

- Se evaluó el impacto de la contaminación en la estimación de α .
- Se emplearon núcleos asimétricos en lugar de histogramas para estimar f_s .
- Se compararon los estimadores basados en la distancia Triangular ($\hat{\alpha}_{\text{DT}}$) con el MV estimador ($\hat{\alpha}_{\text{MV}}$), Momentos Fraccionales ($\hat{\alpha}_{\text{Mom12}}$) y Logcumulantes ($\hat{\alpha}_{\text{LC}}$).

Los estimadores por momentos fraccionales han sido usados por [36, 39] en el contexto de estimar el parámetro α para datos de amplitud. Para el caso de datos de intensidad, usando la relación $\gamma^* = -\alpha - 1$ en 5.1.1 se tiene que, $\hat{\alpha}_{\text{Mom12}}$ es el valor de α que resuelve la ecuación:

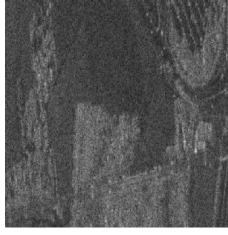
Figura 6.6: $\widehat{\text{Sesgo}}$ y $\widehat{\text{ECM}}$ de $\hat{\alpha}$, $L=8$.

$$\frac{1}{n} \sum_{i=1}^n \sqrt{Z_i} - \sqrt{\frac{-\hat{\alpha}_{\text{Mom12}} - 1}{L} \frac{\Gamma(-\hat{\alpha}_{\text{Mom12}} - \frac{1}{2})}{\Gamma(-\hat{\alpha}_{\text{Mom12}})} \frac{\Gamma(L + \frac{1}{2})}{\Gamma(L)}} = 0. \quad (6.3.1)$$

donde Z_1, \dots, Z_n es una muestra de variables aleatorias i.i.d proveniente del modelo \mathcal{G}_I^0 .

En el capítulo 5 se indicó que $\hat{\alpha}_{\text{MV}}$ y $\hat{\alpha}_{\text{LC}}$ son soluciones de las ecuaciones 5.1.6 y 5.1.21 respectivamente.

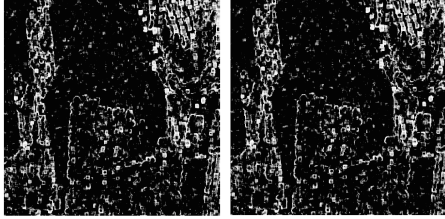
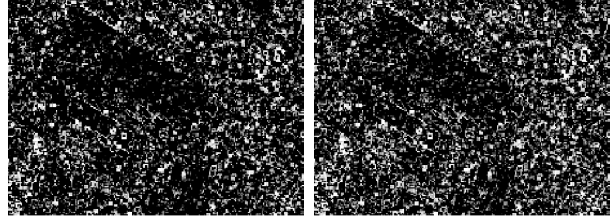
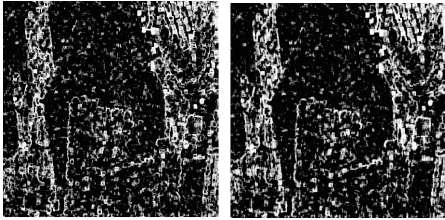
Como se indicó en 5.2.3, dentro de los núcleos más estudiados se encuentran los núcleos IG , Γ y Lognormal. Estas distribuciones están presentes en el modelado de datos SAR, ya sea porque son propuestas para describir al backscatter a partir del modelo multiplicativo, o bien porque fueron propuestas como modelo empírico. En el trabajo [40] estudiamos la performance del núcleo IG para estimar la función de densidad subyacente (\hat{f}_{IG}) en el contexto del MDE estimador y proponemos



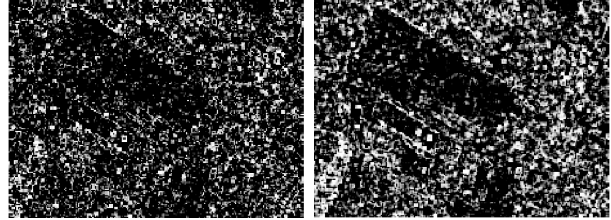
(a) Imagen E-SAR, L=1.



(a) Imagen E-SAR, L=1.

(b) Hellinger. (c) Rényi, $\beta = 0.8$.(b) Hellinger. (c) Rényi, $\beta = 0.8$.

(d) Triangular. (e) MV.



(d) Triangular. (e) MV.

Figura 6.7: Resultado de estimar el parámetro α para cada pixel, usando ventanas deslizantes de tamaño 7×7 .

Figura 6.8: Resultado de estimar el parámetro α para cada pixel, usando ventanas deslizantes de tamaño 3×3 .

como estimador del parámetro de textura del modelo \mathcal{G}_I^0 a:

$$\hat{\alpha}_{IG} = \arg \min_{-20 \leq \alpha \leq -1} d_T(f_{\mathcal{G}^0}(\alpha, \gamma^*, L), \hat{f}_{IG}(z)), \quad (6.3.2)$$

donde d_T es la distancia triangular definida en 5.2.5 y $\gamma^* = -\alpha - 1$.

Como es sabido, uno de los puntos importantes en estimación no paramétrica de la función de densidad es el ancho de banda. En esta oportunidad y, a través de estudios empíricos, se eligió un ancho de banda fijo $b = \frac{n^{-1/2}}{5}$.

Ninguno de los problemas planteados para encontrar $\hat{\alpha}$ tiene una solución

explícita, por lo tanto apelamos a procedimientos numéricos acordes a cada estimador, para encontrar dichas soluciones.

Respecto del estimador $\hat{\alpha}_{MV}$ utilizamos la función *mle* del software R-project que tiene incorporado el algoritmo L-BFGS-B [17] para encontrar el óptimo de una función de una variable. Este algoritmo es un método de optimización tipo cuasi newton que admite restricciones de borde y es ampliamente utilizado en computación gráfica y computación científica [33]. Este algoritmo aproxima al método BFGS (Broyden, Fletcher, Goldfarb y Shanno) ya que no necesita calcular la matriz Hessiana sino que utiliza una aproximación de ella que se actualiza en cada iteración, usando solamente la función y su gradiente. Respecto del estimador $\hat{\alpha}_{LC}$ y $\hat{\alpha}_{Mom12}$ utilizamos la función *uniroot* del mismo software que implementa el método de bisección para encontrar el cero de una función.

En la figura 6.9 se muestra el gráfico de la densidad del modelo \mathcal{G}_I^0 para valores de $\alpha = \{-20, -25\}$, $\gamma = \gamma^*$, $L = \{3, 8\}$. Se puede observar que estas densidades son muy similares ya que la distribución \mathcal{G}_I^0 converge en distribución a una $\Gamma(L, L)$ [37] cuando $\alpha \rightarrow -\infty$ y $\gamma = \gamma^*$. Por este motivo se plantea que α tome valores en el intervalo $[-20, -1]$ en cada método numérico empleado.

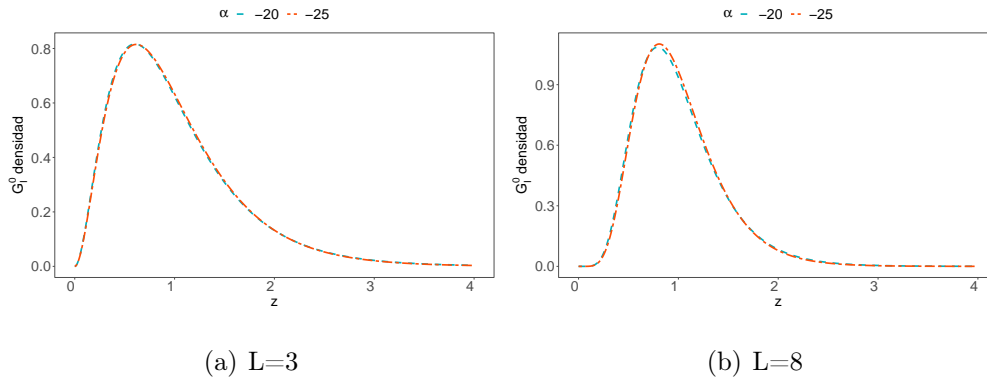


Figura 6.9: Densidad de $\mathcal{G}_I^0(\alpha, \gamma^*, L)$

Se realizó un experimento Monte Carlo para evaluar la performance de cada uno de los estimadores propuestos. El espacio paramétrico consiste en una grilla formada por:

- tres valores de textura $\alpha = \{-1.5, -3, -5\}$. Los cuales representan áreas texturadas y extremadamente texturadas.
- tres niveles de procesamiento señal-ruido $L = \{1, 3, 8\}$.
- tamaños de muestra $n = \{9, 25, 49, 81, 121, 1000\}$ compatibles con tamaños de ventana 3, 5, 7, 9.

Se generaron 1000 muestras para cada punto del espacio paramétrico, y se obtuvieron $\{\hat{\alpha}_1, \dots, \hat{\alpha}_{1000}\}$ estimaciones para cada una de estas combinaciones. De esta forma se estimaron, para cada uno de los estimadores estudiados

- la media $\bar{\hat{\alpha}} = (1000)^{-1} \sum_{i=1}^{1000} \hat{\alpha}_i$,
- el sesgo $\hat{B}(\hat{\alpha}) = \bar{\hat{\alpha}} - \alpha$.
- el error cuadrático medio $\widehat{\text{ECM}} = (1000)^{-1} \sum_{i=1}^{1000} (\hat{\alpha}_i - \alpha)^2$.

En las siguientes figuras “MV”, “DT”, “Mom12” and “LC” denotarán al estimador basado en el método de Máxima Verosimilitud, distancia Triangular con núcleo IG, $\frac{1}{2}$ -momento y Logcumulante, respectivamente. En el eje de las abscisas se grafica el tamaño muestral, el cual se presenta en escala semilogarítmica.

Las figuras 6.10(a), 6.11(a) y 6.12(a) muestran el sesgo de $\hat{\alpha}$ para datos sin contaminar y para diferentes valores de n y L , la línea azul representa el eje x . Se puede observar que $\hat{\alpha}_{\text{MV}}$ y $\hat{\alpha}_{\text{DT}}$ tienen un comportamiento similar y un sesgo medio cercano a cero en la mayoría de los casos estudiados. Es notable el sesgo que presentan $\hat{\alpha}_{\text{LC}}$ y $\hat{\alpha}_{\text{Mom12}}$ para $\alpha = -5$. Se puede observar también que $\hat{\alpha}_{\text{MV}}$ tiende a subestimar al verdadero valor de α , y que ningún método de estimación tiene el menor error cuadrático medio en todos los casos. El estimador $\hat{\alpha}_{\text{Mom12}}$ es el que peor se comporta para zonas texturadas.

Todos los métodos presentan alta variabilidad para el caso de $L = 1$, esta variabilidad disminuye a medida que aumenta el número de looks, incluso la longitud de los intervalos de confianza van disminuyendo a medida que aumenta el tamaño de muestra.

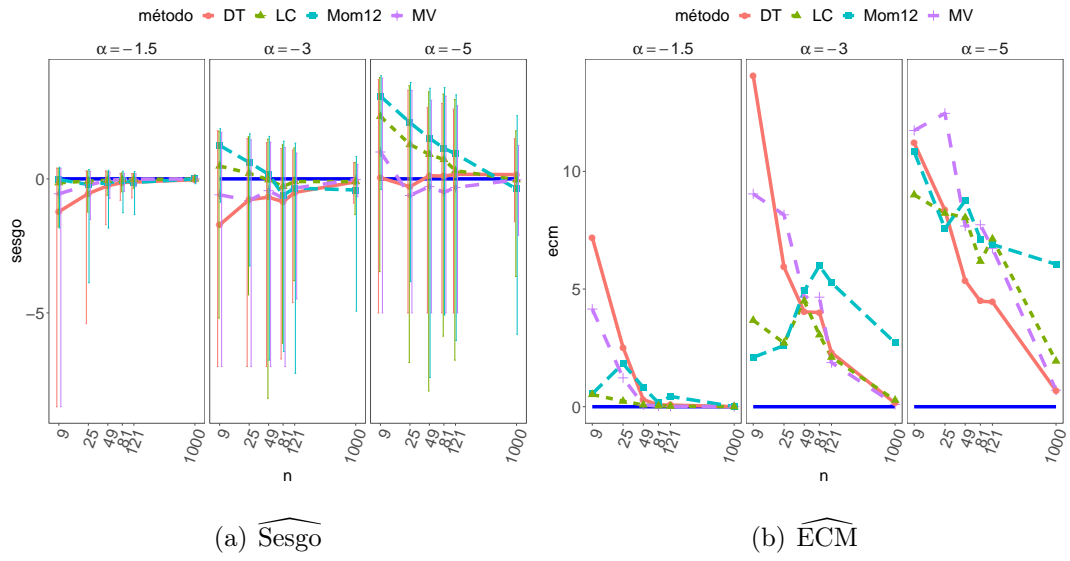


Figura 6.10: $\widehat{\text{Sesgo}}$ y $\widehat{\text{ECM}}$ para datos sin contaminar, $L = 1$.

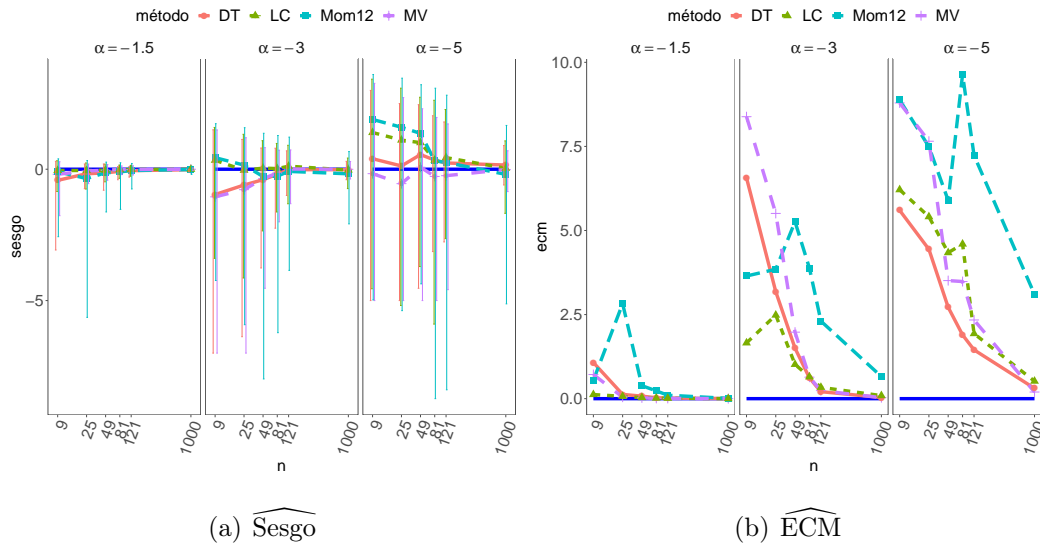


Figura 6.11: $\widehat{\text{Sesgo}}$ y $\widehat{\text{ECM}}$ para datos sin contaminar, $L = 3$.

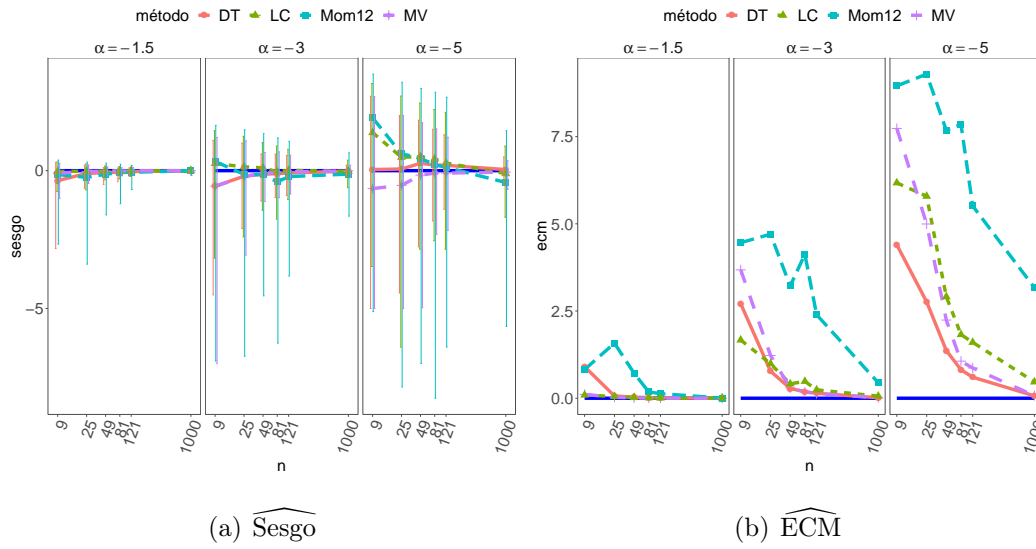


Figura 6.12: $\widehat{\text{Sesgo}}$ y $\widehat{\text{ECM}}$ para datos sin contaminar, $L = 8$.

Se calculó el tiempo medio de procesamiento, medido en segundos, para cada método y cada caso estudiado. Como ejemplo se presenta, en la tabla 6.1, el tiempo medio para $L = 1$ y $n = 81$. Se puede observar que nuestra propuesta tiene un mayor costo computacional debido a la integración numérica presente en la definición del estimador. Los otros casos son consistentes con los datos de esta tabla. La plataforma informática que se utilizó para realizar los procesos fue Intel(R) Core i7, con 8GB de memoria y 64 bits Windows 7.

MV	DT	Mom $\frac{1}{2}$	LC
0.003	2.223	0.0001	0.003

Tabla 6.1: Tiempos medios para datos simulados sin contaminación, $L = 1$, $n = 81$.

Es muy importante contar con estimadores que sean resistentes a la presencia de datos contaminados, es decir, que sean capaces de producir buenas estimaciones incluso cuando una proporción de los datos no proviene del modelo supuesto como verdadero. Esta situación es de particular importancia en el caso de muestras de pequeño tamaño, por ejemplo, cuando se utilizan filtros que emplean estimadores basados, por lo general, en pequeñas muestras ya que recorren la imagen

a través de ventanas deslizantes de tamaño 3×3 , 5×5 o, por ejemplo, 7×7 . Estas muestras pueden tener datos de zonas con diferentes grado de textura, por ejemplo, en el borde entre diferentes regiones. Esta capacidad de producir buenas estimaciones cuando las observaciones no provienen exactamente del modelo asumido se llama Robustez.

Una de las fuentes de contaminación en las imágenes SAR es el fenómeno de Double Bounce, donde algunos píxeles tienen un alto valor de retorno. La presencia de tales valores atípicos puede provocar grandes errores en la estimación. La descripción de este fenómeno se hizo en 3.10.

Con el fin de evaluar la robustez de los estimadores, proponemos tres escenarios capaces de describir los desvíos del modelo teórico. Para cada uno de estos escenarios generamos muestras contaminadas donde $0 < \epsilon \ll 1$ es la proporción de contaminación. Sea B una variable aleatoria Bernoulli con probabilidad p de ocurrencia de la contaminación. Sea $C \in \mathbb{R}_+$ un valor grande. Entonces

- Caso 1: Sean W y U variables aleatorias tales que $W \sim \mathcal{G}_I^0(\alpha_1, \gamma_1^*, L)$, y $U \sim \mathcal{G}_I^0(\alpha_2, \gamma_2^*, L)$. Definimos $Z = BU + (1 - B)W$, entonces generamos $\{z_1, \dots, z_n\}$ variables aleatorias independientes, idénticamente distribuidas con función de distribución acumulada dada por:

$$(1 - \epsilon)\mathcal{F}_{\mathcal{G}_I^0(\alpha_1, \gamma_1^*, L)}(z) + \epsilon\mathcal{F}_{\mathcal{G}_I^0(\alpha_2, \gamma_2^*, L)}(z),$$

donde $\mathcal{F}_{\mathcal{G}_I^0(\alpha, \gamma, L)}$ es la función de distribución acumulada bajo el modelo $\mathcal{G}_I^0(\alpha, \gamma, L)$.

- Case 2: Consideramos $W \sim \mathcal{G}_I^0(\alpha_1, \gamma_1^*, L)$ y el retorno definido como $Z = BC + (1 - B)W$.
- Case 3: Consideramos $W \sim \mathcal{G}_I^0(\alpha, \gamma^*, L)$ y $U \sim \mathcal{G}_I^0(\alpha, 10^k \gamma^*, L)$ con $k \in \mathbb{N}$. El retorno Z se define como el Caso 1 y la función de distribución acumulada

está dada por:

$$(1 - \epsilon)\mathcal{F}_{\mathcal{G}_I^0(\alpha, \gamma^*, L)}(z) + \epsilon\mathcal{F}_{\mathcal{G}_I^0(\alpha, 10^k \gamma^*, L)}(z).$$

Todos estos modelos consideran desvíos de la hipótesis del modelo teórico. El primer tipo de contaminación asume que, con probabilidad ϵ , los datos pueden provenir de una distribución perteneciente a la familia de distribuciones \mathcal{G}_I^0 pero con otros parámetros. El segundo tipo de contaminación modela, con probabilidad ϵ , un retorno con un valor grande y fijo, digamos $C = 100$. El tercer tipo de contaminación es un caso particular del primero, donde la contaminación asume que los datos provienen de una distribución cuyo factor de escala es k órdenes de magnitud mayor que el factor de escala correspondiente al modelo teórico. Analizamos estos tres casos de la contaminación en la evaluación de cada estimador.

Las figuras 6.13, 6.14 y 6.15 muestran el sesgo y el ECM estimado ($\widehat{\text{ECM}}$) para el Caso 1 de contaminación con $\alpha_2 = -15$, $\epsilon = 0.01$, y variando n y L . Este tipo de contaminación introduce, con probabilidad $\epsilon = 0.01$, observaciones casi sin textura en la muestra bajo análisis. Como se espera, la influencia de tal perturbación es más notable en aquellas situaciones donde el modelo subyacente está más alejado de la contaminación. Es decir, para valores grandes de α . Esto se ve con claridad en las figuras 6.13(b), 6.14(b) y 6.15(b) donde el $\widehat{\text{ECM}}$ de $\hat{\alpha}_{\text{MV}}$, $\hat{\alpha}_{\text{Mom12}}$ y $\hat{\alpha}_{\text{LC}}$ son mayores que el error cuadrático medio de $\hat{\alpha}_{\text{DT}}$ para $L = 3, 8$, pero esto no es tan claro para el caso de $L = 1$. Lo que si se puede decir que $\hat{\alpha}_{\text{DT}}$ es competitivo respecto de los otros estimadores para $\alpha = -3, -5$.

Las figuras 6.16, 6.17 y 6.18 presentan $\hat{\alpha}$ y $\widehat{\text{ECM}}$ para el Caso 2 de contaminación con $\epsilon = 0.001$ y $C = 100$. Este tipo de contaminación inyecta un valor constante, en este caso 100, con probabilidad $\epsilon = 0.001$. Como estamos considerando muestras con media unitaria, este valor de C respresenta un valor grande de contaminación. En este caso $\hat{\alpha}_{\text{DT}}$ es, en media, más cercano al verdadero valor

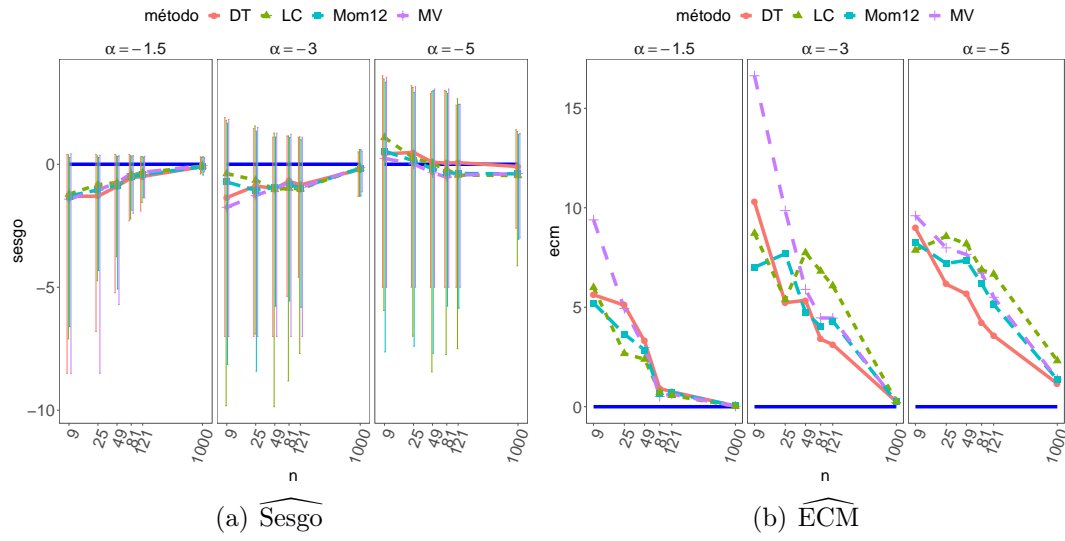


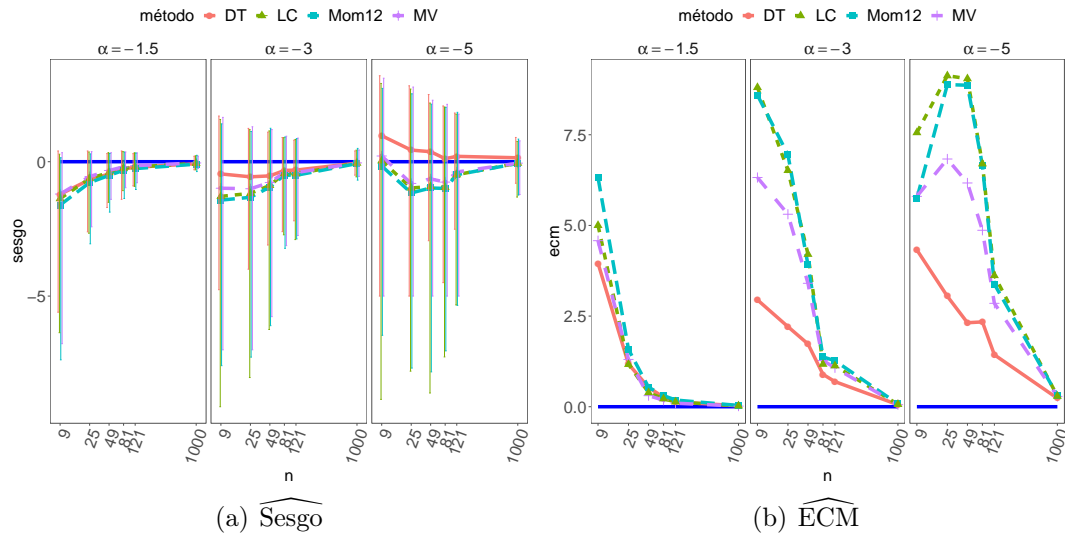
Figura 6.13: Datos contaminados: Caso 1, $\epsilon = 0.01$ y $L = 1$.

que los otros métodos, y el valor de \widehat{ECM} es menor.

Las figuras 6.19, 6.20 y 6.21 muestran $\widehat{\alpha}$ y \widehat{ECM} para el Caso 3 de contaminación con $\epsilon = 0.005$ y $k = 2$. En este caso se contamina la muestra proveniente del modelo verdadero con un observaciones que, con probabilidad $\epsilon = 0.005$, provienen de una distribución \mathcal{G}_I^0 con un factor de escala cien veces más grande que el correspondiente al modelo verdadero. El comportamiento de estos estimadores sigue el mismo patrón para $L = 3$ y 8, $\widehat{\alpha}_{DT}$ produce estimaciones más cercanas al verdadero valor con un ECM más chico. No hay un buen estimador para el caso $L = 1$ en este caso de contaminación.

Como se mencionó anteriormente, los algoritmos implementados para los casos de los estimadores Mom12 y Logcumulante no convergen en algunos casos. En el experimento Monte Carlo, si Mom12 o Logcumulante no convergen, no consideramos esta iteración al momento de estimar, por lo que la cantidad de elementos para calcular $\widehat{\alpha}$, el sesgo y el error cuadrático medio es un poco menor que 1000.

La tabla 6.2 informa, a modo de ejemplo, el porcentaje de casos que fueron removidos para el Caso 1 y $\alpha_2 = -15$, $\epsilon = 0.01$. Estos resultados son consistentes con los otros escenarios de contaminación, y sugieren que estos métodos son progresivamente más propensos a fallar en casos con mayor presencia de ruido

Figura 6.14: Datos contaminados: Caso 1, $\epsilon = 0.01$ y $L = 3$.

speckle.

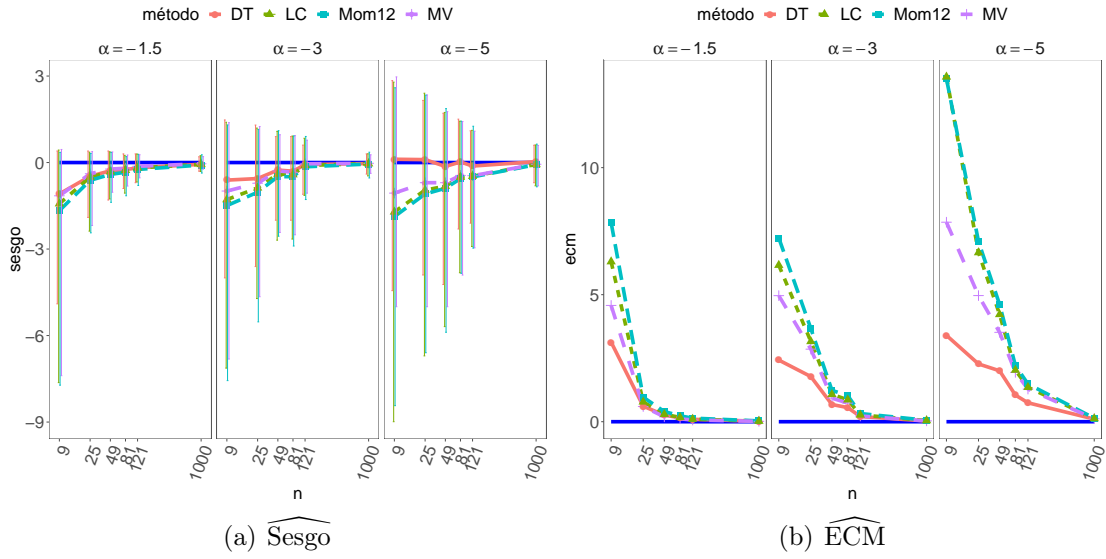
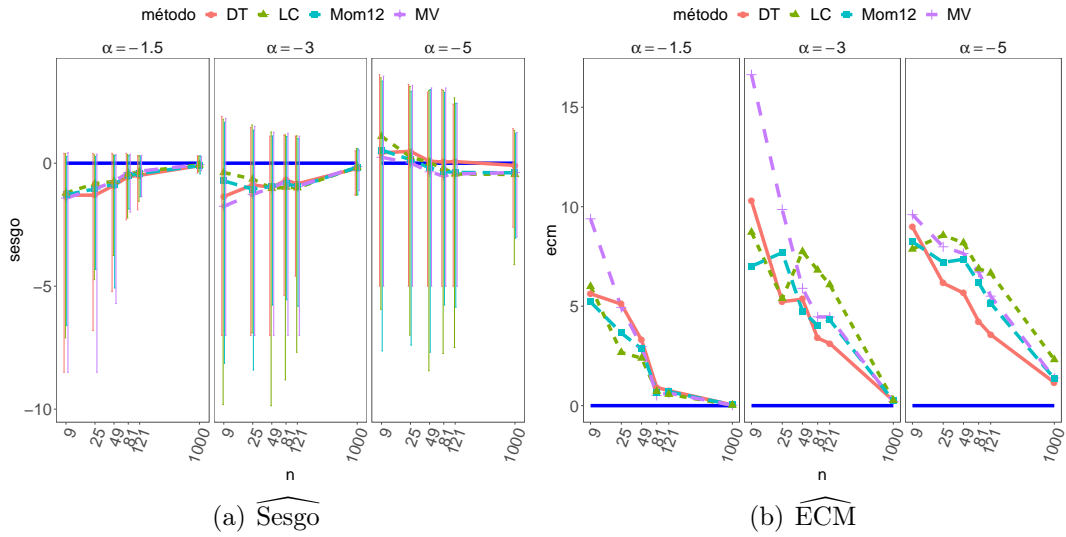
L	Mom12	LC
1	22.87	21.56
3	11.71	11.87
8	5.81	6.04

Tabla 6.2: Porcentaje de casos de no convergencia para los estimadores de Momentos y Logcumulante en el Caso 1, $\alpha_2 = -15$, $\epsilon = 0.01$

Hemos aplicado estos métodos a una imagen real, la misma imagen utilizada en Cassetti et al. [20], una imagen E-SAR [48] de un look de los alrededores de Munich, banda L, polarización HH en formato intensidad. La figura 6.22 muestra las regiones usadas para estimar el parámetro de textura. Esta imagen tiene 300×250 pixels y comprende principalmente dos áreas de cultivo diferentes.

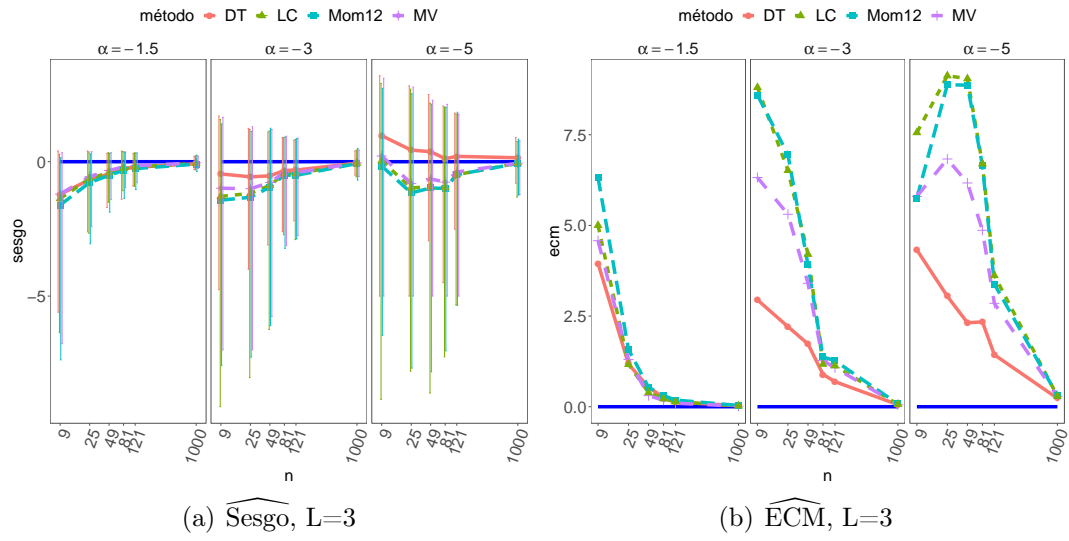
La tabla 6.3 muestra los resultados de las estimaciones del parámetro α para cada región rectangular junto con los tiempos de procesamiento, donde NA significa que el algoritmo correspondiente no convergió. El resto de las estimaciones arrojan valores compatibles con el mismo tipo de zona en cada una de las muestras analizadas.

La figura 6.23 muestra una imagen con la presencia de un corner reflector,

Figura 6.15: Datos contaminados: Caso 1, $\epsilon = 0.01$ y $L = 8$.Figura 6.16: Datos contaminados: Caso 2, $\epsilon = 0.001$ y $L = 1$.

mientras que la figura 6.24 muestra las regiones usadas para estimar el parámetro de textura para esta situación. La tabla 6.4 presenta los valores de $\hat{\alpha}$ para cada área rectangular en la imagen de la figura 6.24.

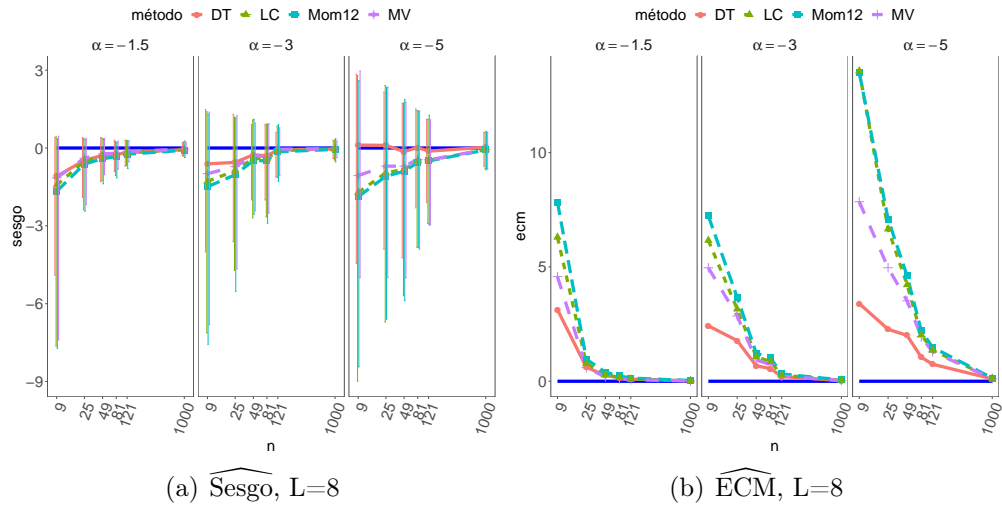
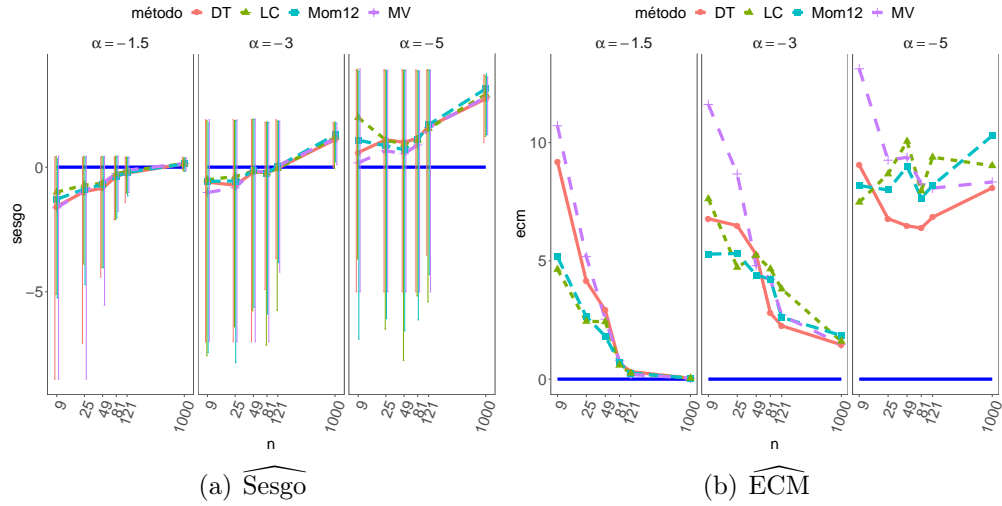
Los estimadores MV, Mom12 y Logcumulantes son incapaces de producir una estimación en pequeñas muestras. Se puede observar que tanto $\hat{\alpha}_{\text{Mom12}}$ como $\hat{\alpha}_{\text{LC}}$ requieren al menos un tamaño de muestra de 90 observaciones para que se pueda

Figura 6.17: Datos contaminados: Caso 2, $\epsilon = 0.001$ y $L = 3$.

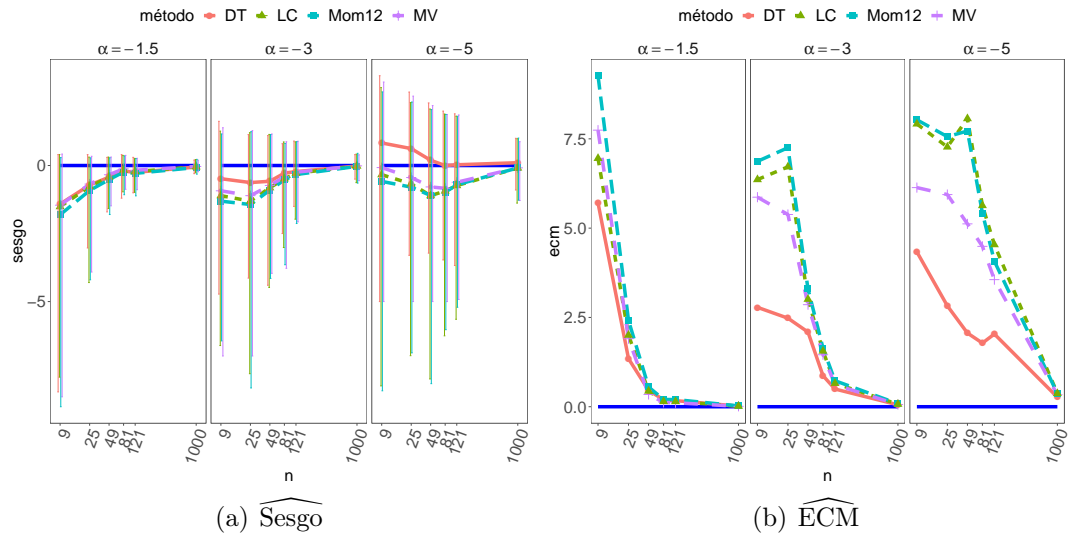
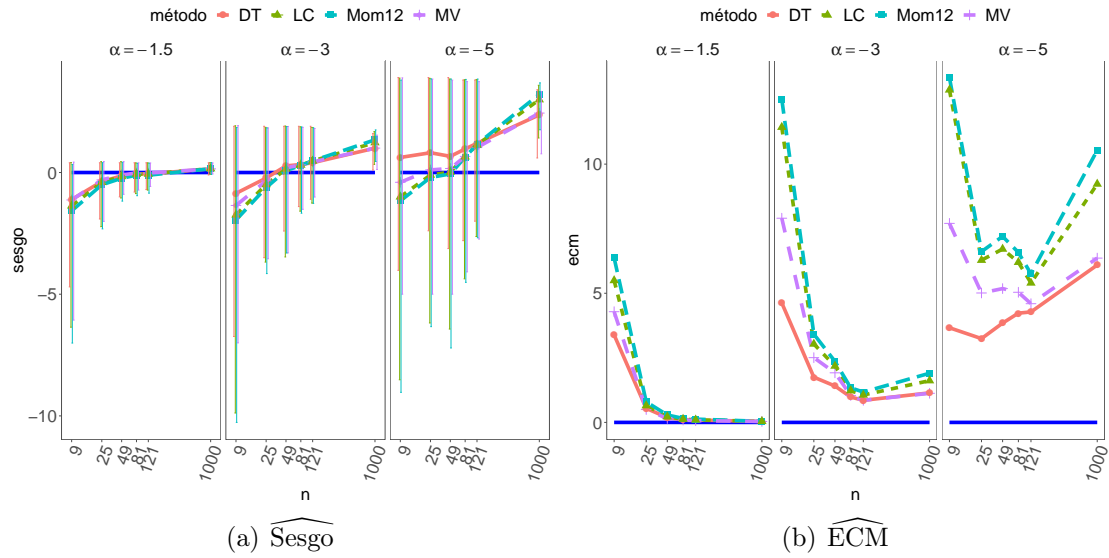
Color	n	$\hat{\alpha}_{MV}$	$\hat{\alpha}_{DT}$	$\hat{\alpha}_{Mom12}$	$\hat{\alpha}_{LC}$	Tiempo MV	Tiempo DT	Tiempo Mom12	Tiempo LC
Magenta	100	-1.9	-2.7	-1.9	-1.7	0.03	5.85	0.03	0.02
Amarillo	90	-6.2	-5.1	-6.6	-6.8	0.00	5.16	0.00	0.00
Rojo	64	-1.8	-1.9	-1.9	-1.8	0.00	4.17	0.00	0.00
Verde	48	-2.5	-2.5	-2.9	-3.1	0.00	3.31	0.00	0.00
Azul	25	-4.9	-3.0	NA	NA	0.00	2.08	0.00	0.00

Tabla 6.3: $\hat{\alpha}$ para las muestras de la figura 6.22.

dar una estimación para α . El estimador basado en la distancia triangular produce valores plausibles bajo contaminación incluso con muestras muy pequeñas.

Figura 6.18: Datos contaminados: Caso 2, $\epsilon = 0.001$ y $L = 8$.Figura 6.19: Datos contaminados: Caso 3, $\epsilon = 0.005$ y $L = 1$.Tabla 6.4: $\hat{\alpha}$ y tiempos de procesos para las muestras presentadas en la figura 6.24

Color	n	$\hat{\alpha}_{\text{MV}}$	$\hat{\alpha}_{\text{T}}$	$\hat{\alpha}_{\text{Mom12}}$	$\hat{\alpha}_{\text{LC}}$	Tiempo MV	Tiempo DT	Tiempo Mom12	Tiempo LC
Magenta	15	-20.0	-4.1	NA	NA	0.03	1.95	0.03	0.03
Verde	42	-9.2	-5.0	NA	NA	0.00	4.04	0.00	0.00
Azul	90	-3.5	-2.7	-4.7	-14.2	0.02	4.85	0.00	0.00
Amarillo	156	-2.2	-1.8	-2.6	-3.4	0.01	8.35	0.00	0.00
Rojo	225	-1.9	-1.7	-2.1	-2.5	0.02	10.97	0.00	0.00

Figura 6.20: Datos contaminados: Caso 3, $\epsilon = 0.005$ y $L = 3$.Figura 6.21: Datos contaminados: Caso 3, $\epsilon = 0.005$ y $L = 8$.

Estos resultados nos llevan a concluir que la distribución \mathcal{G}_I^0 es una buena opción para modelar adecuadamente datos provenientes de imágenes SAR con datos de intensidad ante la presencia de contaminación. Asimismo podemos concluir que el estimador propuesto es una buena elección para estimar el parámetro de textura de la distribución \mathcal{G}_I^0 , ya que es competitivo frente al MV estimador y mejora a los otros dos estimadores estudiados especialmente para el caso de

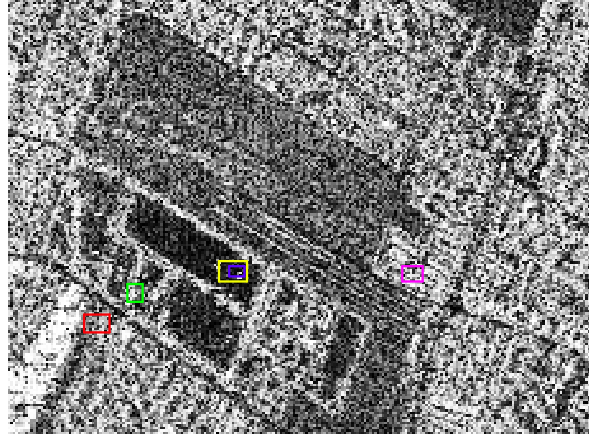


Figura 6.22: Image real E-SAR junto con la regiones usadas para estimar el parámetro α .

muestras de pequeño tamaño.

6.4. Mejorando la propuesta

Para el caso $L = 1$ la distribución $\mathcal{G}_I^0(\alpha, \gamma)$ es una distribución de Pareto Generalizada de parámetros $\mathcal{PG}(\mu, \sigma, \beta)$, donde $\mu = 0$, $\sigma = \gamma$ y $\beta = -\alpha$. El parámetro de textura de la distribución \mathcal{G} está relacionado con el índice de la cola. En el trabajo que presentamos en [21] empleamos teoría de valores extremos para proponer un estimador para este parámetro. Por este motivo en esta tesis continuamos el análisis para el caso multilook.

En [40] presentamos el MDE estimador utilizando distancia triangular y encontrando el mínimo recorriendo valores en una grilla, estimando \hat{f}_{AK} con núcleo IG y proponiendo un valor de ancho de banda b encontrado empíricamente.

En esta sección proponemos mejoras:

- Consideramos otros núcleos.
- Utilizamos el método LSCV para encontrar el ancho de banda que se pre-

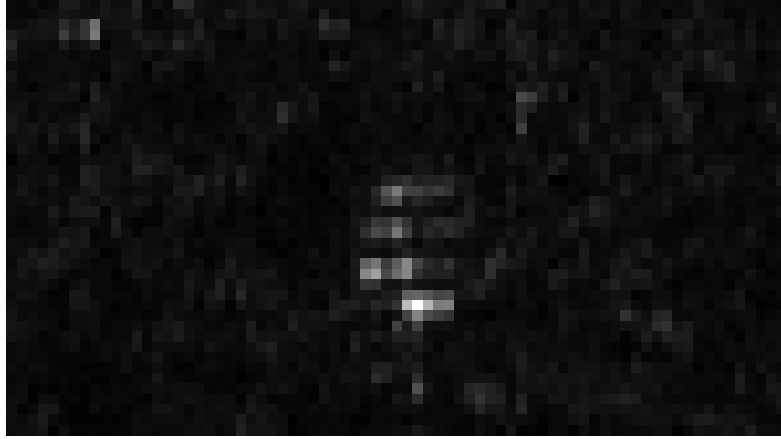


Figura 6.23: Imagen real SAR de 1-look con un corner reflector.

sentó en la subsección 5.2.32.

- Utilizamos el algoritmo L-BFGS-B 6.3 para encontrar el estimador $\hat{\alpha}_{DT}$ minimizando la distancia definida en 6.3.2.
- Contamos casos de no convergencia de los algoritmos numéricos empleados, entendiendo que un algoritmo no converge si devuelve como valor estimado del parámetro a alguno de los extremos del intervalo de búsqueda.

La pregunta es ¿por qué el núcleo IG y no otro? ¿Cómo elegir el mejor núcleo? Para responder a estas preguntas estudiamos otros núcleos. En 5.2.3 se mencionó que, dentro de los núcleos más estudiados, se encuentran los núcleos IG , Γ^1 , Γ^2 y LN . En [40] evaluamos la performance del núcleo IG , en esta sección vamos a evaluar los otros dos núcleos y compararemos su performance con el núcleo IG .

En una primera instancia estimamos el MISE (5.2.19) (\widehat{MISE}) para cada núcleo estudiado, para distintas combinaciones de los parámetros, tamaños muestrales y número de looks.

Realizamos 500 replicaciones, en cada replicación generamos una muestra pro-

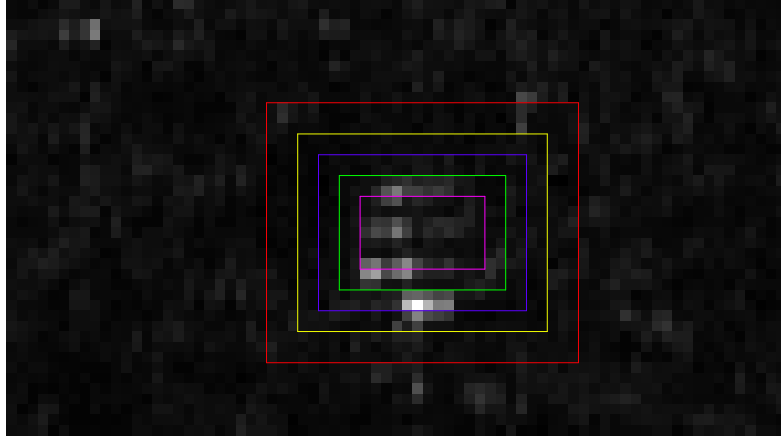


Figura 6.24: Muestras de varios tamaños en una imagen real SAR de 1-look con un corner reflector.

veniente de una distribución $\mathcal{G}_I^0(\alpha, \gamma^*)$ y calculamos

$$\widehat{\text{MISE}} = \frac{1}{500} \sum_{i=1}^{500} \int_0^{+\infty} (\hat{f}_{b,K}^i(x) - f(x))^2 dx, \quad (6.4.1)$$

donde $\hat{f}_{b,K}^i(x)$ es el estimador de la función de densidad subyacente en la replicación i para el núcleo K , que depende del ancho de banda b .

En la tabla 6.5 se muestran los valores de $\widehat{\text{MISE}}$ para $\alpha = \{-1.5, -3, -5, -8\}$, $L = 3$, $\gamma^* = -\alpha - 1$, $n = \{9, 25, 49, 81, 121\}$ y $K = \{\Gamma^1, \Gamma^2, \text{LN}, \text{IG}, \text{IGJstar}\}$, donde Γ^1 , Γ^2 , LN, IG, IGJstar corresponden a los núcleos Γ^1 , Γ^2 , Lognormal e Inverso Gaussiano respectivamente. La columna IGJstar corresponde al núcleo estudiado en [40] donde se consideró un ancho de banda fijo encontrado empíricamente. En esta oportunidad se aplicó el método LSCV para encontrar el ancho de banda en cada replicación y para cada combinación de los parámetros en el resto de los núcleos estudiados.

Se puede observar que:

- Los valores del MISE son muy similares para los núcleos Γ^1 y Γ^2 , por lo

que se eligió el núcleo Γ^1 ya que presenta menor MISE en la mayoría de los casos planteados. A partir de ahora el núcleo Γ^1 se llamará Γ .

- Los valores del $\widehat{\text{MISE}}$ correspondientes a los núcleos IG y con IGJstar son de varios órdenes de magnitud mayor a los otros núcleos.

Luego evaluamos la cantidad de casos de no convergencia de cada método. La tabla 6.5 muestra el porcentaje de casos donde el valor de $\hat{\alpha} = -20$, no se observaron situaciones donde $\hat{\alpha} = -1$. En negrita están marcados los casos donde se presentan el mayor porcentaje. Se puede observar que, en general, el núcleo IG es el que presenta mayor porcentaje de ocurrencias de no convergencia.

Además estudiamos el sesgo y el ECM para el caso donde el número de looks es $L = 3$, la tabla 6.6 muestra los resultados obtenidos. En negrita están marcados los casos donde el sesgo y el ECM son menores con IG e IGJstar respecto de los núcleos Γ y LN. Se puede observar que la performance de estos dos últimos núcleos mejora al núcleo IG (en sus dos versiones) en la mayoría de los casos analizados. Sin embargo IG muestra una menor variabilidad salvo para zonas extremadamente texturadas, lo que no sucede con IGJstar.

Entonces, priorizando la calidad del ajuste, el porcentaje de casos de convergencia y el sesgo continuamos el análisis con los núcleos Γ y LN.

6.5. Evaluación de la performance entre los distintos métodos de estimación

En esta sección presentamos los resultados empíricos que obtuvimos sobre la calidad de los estimadores para el parámetro de textura del modelo \mathcal{G}_I^0 . Como dijimos en la sección anterior elegimos los métodos MV (Máxima Verosimilitud), Γ (núcleo Gamma), LN (núcleo Lognormal) y LC (Logcumulante). El primero y último estimador son los más usados en la literatura, los otros son los propuestos en esta tesis.

α	n	MISE					Porcentaje de casos de no convergencia			
		Γ^1	Γ^2	LN	IG	IGJstar	Γ	LN	IG	IGJstar
-1.5	9	0.407	0.411	0.811	6.119	41.330	0	0.4	0.4	0
	25	0.123	0.145	0.184	2.464	17.024	0	0	0	0
	49	0.082	0.100	0.538	1.157	9.671	0	0	0	0
	81	0.064	0.079	0.085	0.753	6.126	0	0	0	0
	121	0.064	0.076	0.081	0.528	4.000	0	0	0	0
-3	9	0.253	0.273	0.563	12.473	50.020	5.2	7.2	8	6.8
	25	0.080	0.093	0.106	3.217	25.401	0.2	1	0.6	0.6
	49	0.044	0.053	0.071	0.686	16.401	0	0.4	0	0
	81	0.030	0.033	0.034	0.255	11.577	0	0	0	0
	121	0.022	0.022	0.029	0.166	8.985	0	0	0	0
-5	9	0.238	0.266	0.432	15.315	63.384	9.6	13.2	17	13.2
	25	0.066	0.085	0.087	4.647	32.939	3.4	5	2	1.6
	49	0.042	0.048	0.064	0.677	24.405	1.4	1.2	1	0.4
	81	0.027	0.027	0.033	0.285	18.906	0.4	0.8	0.2	0
	121	0.019	0.016	0.025	0.179	15.920	0	0.2	0	0
-8	9	0.237	0.263	0.383	18.106	73.081	16.6	19	26.2	18
	25	0.072	0.088	0.112	5.134	40.479	81.81	11	8.8	4.2
	49	0.041	0.044	0.048	0.930	30.231	5	5.4	4.2	1.2
	81	0.026	0.024	0.031	0.329	25.712	4.4	4	2	0.2
	121	0.017	0.014	0.020	0.214	21.193	1.8	2	0.8	0.4

Tabla 6.5: MISE y porcentaje de casos de no convergencia $L = 3$.

A través de un experimento de Monte Carlo evaluamos la calidad de los estimadores estudiando:

- el porcentaje de casos de no convergencia y el costo computacional.
- el sesgo y el ECM para diferentes escenarios.
- la robustez de estos estimadores por medio del Caso 1 de contaminación y de la función de influencia empírica que será definida en la sección 6.6.

El experimento Monte Carlo que se implementó es como el descrito en 6.3, en esta oportunidad se realizaron 500 replicaciones independientes para cada uno de los valores de los parámetros: $\alpha \in \{-1.5, -3, -5, -8\}$, $L \in \{1, 3, 8\}$, $n \in \{9, 25, 49, 81, 121, 500\}$, valores similares a los estudiados en [40].

α	n	Sesgo				ECM			
		Γ	LN	IG	IGJstar	Γ	LN	IG	IGJstar
-1.5	9	-0.183	-0.117	-0.84	-0.525	0.407	0.417	1.432	1.348
	25	-0.058	-0.007	-0.538	-0.182	0.07	0.049	0.496	0.129
	49	-0.035	0.002	-0.404	-0.105	0.034	0.019	0.217	0.040
	81	-0.036	0.005	-0.348	-0.072	0.024	0.011	0.158	0.021
	121	-0.041	0.014	-0.295	-0.046	0.024	0.008	0.108	0.011
-3	9	-0.652	-0.552	-1.659	-1.584	5.873	8.945	10.65	11.965
	25	-0.338	-0.066	-0.582	-0.684	3.542	2.352	1.843	3.475
	49	-0.014	0.015	-0.215	-0.319	0.928	1.34	0.549	1.145
	81	0.079	0.091	-0.03	-0.165	0.397	0.441	0.212	0.490
	121	0.117	0.104	0.025	-0.108	0.239	0.285	0.294	0.287
-5	9	0.084	0.67	-0.645	-0.743	11.348	10.726	9.943	13.158
	25	0.225	0.632	0.092	-0.531	6.506	6.648	4.603	8.359
	49	0.362	0.339	0.604	-0.099	3.989	5.787	2.338	4.688
	81	0.318	0.437	0.771	-0.101	3.045	2.761	1.852	3.623
	121	0.36	0.397	0.896	0.108	1.95	2.041	1.454	1.505
-8	9	2.346	2.952	1.14	1.382	18.944	21.516	16.879	19.118
	25	1.787	2.209	1.96	0.93	13.309	16.633	11.463	13.551
	49	1.41	1.735	2.442	1.073	13.02	12.658	10.667	10.100
	81	1.26	1.462	2.498	0.779	9.727	9.893	10.468	10.278
	121	1.062	1.235	2.689	0.915	7.425	8.555	8.983	5.975

Tabla 6.6: Sesgo y ECM para $L = 3$.

Porcentaje de casos de no convergencia y tiempos de proceso

En la tabla 6.7 se muestra el porcentaje de casos de no convergencia para $L = 3$, el caso $L = 8$ arrojan valores similares. Se puede observar que los métodos MV y LC son los que mayor porcentaje de no convergencia presentan, seguidos por el método LN y por último el método Γ .

En la tabla 6.8 se muestran los tiempos medios, medidos en segundos, para el caso de datos sin contaminación, $L = 3$ y $n = 81$. En esta oportunidad la plataforma informática utilizada para realizar los procesos fue Intel(R) Core i7, con 16GB de memoria, 64 bits Windows 7. Si bien se duplicó la memoria del procesador, no cambió el tiempo de proceso de los estimadores MV y LC. En cambio el tiempo de proceso de los MDE estimadores se redujo un 18%. Esta

Tabla 6.7: Porcentaje de casos de no convergencia $L=3$.

α	n	MV	Γ	LN	LC
-1.5	9	0	0	0.4	2.8
	25	0	0	0	0.2
-3	9	13	5.2	7.2	28.4
	25	1	0.2	1	11.4
	49	0.2	0	0.4	3.8
	81	0	0	0	2.4
	121	0	0	0	0.2
-5	9	26.8	9.6	13.2	35.2
	25	10	3.4	5	28.6
	49	3.4	1.4	1.2	18.6
	81	0.2	0.4	0.8	15.8
	121	0.4	0	0.2	9.6
	500	0	0	0	0.6
-8	9	39.6	16.6	19	44.2
	25	28.6	9	11	36.4
	49	18.4	5	5.4	31.6
	81	12	4.4	4	27.2
	121	5.8	1.8	2	24.6
	500	0	0	0.2	9

reducción se puede adjudicar al cambio de equipo informático como a la forma de encontrar la minimización.

MV	Γ	LN	LC
0.003	1.87	1.87	0.003

Tabla 6.8: Tiempos medios para datos simulados sin contaminación, $L = 3$, $n = 81$.

Evaluación de los métodos de estimación para datos sin contaminar

En la figura 6.25 se muestra el sesgo y el ECM estimados para datos sin contaminar y $L=3$. Las iteraciones consideradas son aquellas en donde todos los métodos convergen. Se puede observar que a medida que el valor del α disminuye, tanto $\widehat{\text{Sesgo}}$ como $\widehat{\text{ECM}}$ van aumentando su valor en la mayoría de los métodos.

Dado que por cuestiones de escala, no se puede evaluar adecuadamente el $\widehat{\text{Sesgo}}$ y el $\widehat{\text{ECM}}$ para $\alpha = -1.5, -3$ lo graficamos aparte, en la figura 6.26. Se puede observar que el estimador MV y LC por un lado, y Γ y LN por el otro tienen un comportamiento similar tanto en sesgo como en ECM en la mayoría de los casos. Este comportamiento similar también se puede observar en los intervalos de confianza, los métodos MV y LC producen intervalos de longitud mayor que los otros métodos, mostrando que son menos precisos respecto de los MDE estimadores.

Para zonas extremadamente heterogéneas y heterogéneas los MDE estimadores presentan una mejor performance que el resto. En cambio para zonas moderadamente texturadas y homogéneas el comportamiento se revierte.

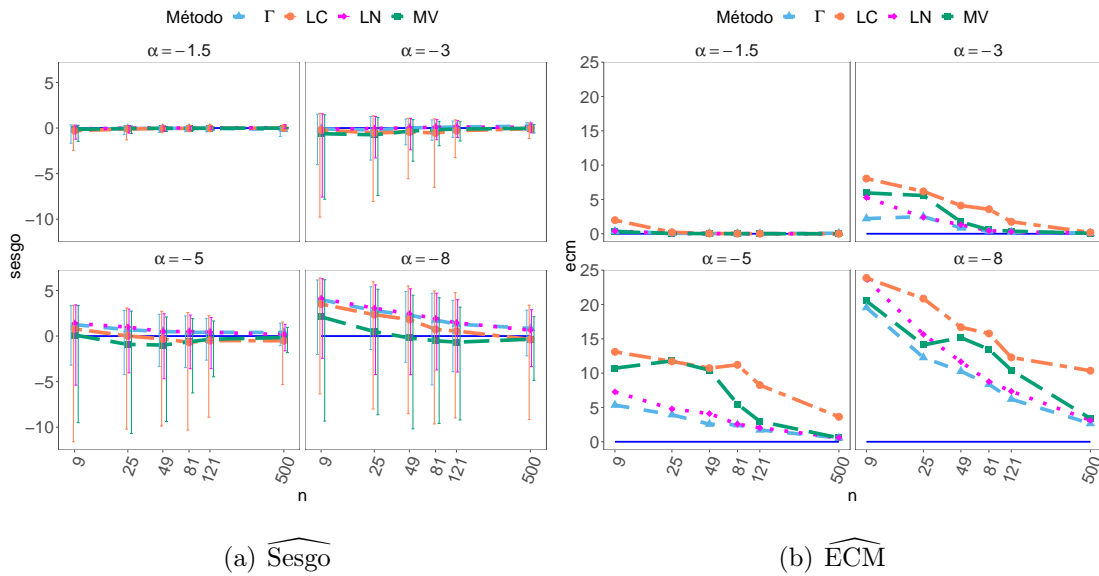


Figura 6.25: Sesgo y ECM estimados para datos sin contaminar y $L = 3$.

Nuevamente, como no se puede visualizar adecuadamente el $\widehat{\text{Sesgo}}$ y el $\widehat{\text{ECM}}$ para $\alpha = -1.5, -5$ hicimos un gráfico mostrando únicamente estos casos.

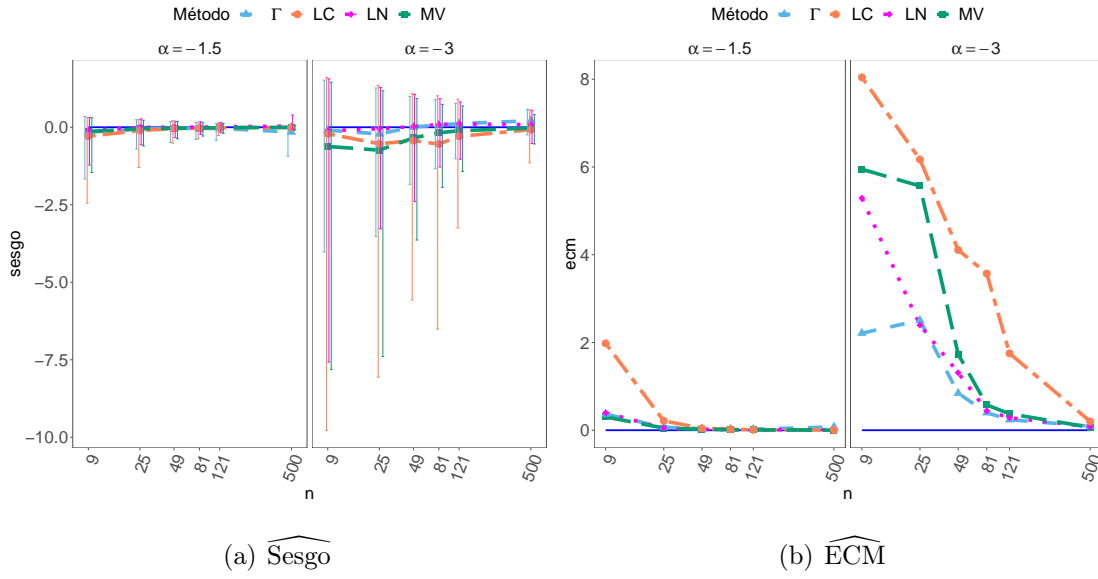


Figura 6.26: Sesgo y ECM estimados para datos sin contaminar, $L = 3$ y $\alpha = -1.5, -3$.

El mismo análisis se presenta en las figuras 6.27 y 6.28 para el caso $L = 8$. En este caso el MV estimador es el que mejor performance muestra para zonas homogéneas, mientras que el LN estimador supera la resto en zonas extremadamente texturadas y con moderada textura.

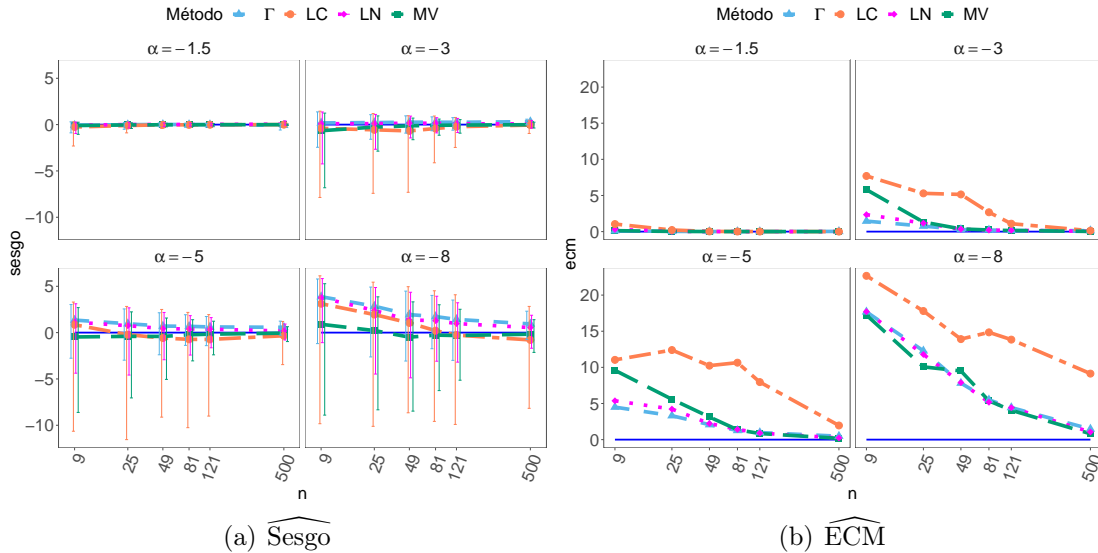


Figura 6.27: Sesgo y ECM estimados para datos sin contaminar y $L = 8$.

De acuerdo al estudio para datos sin contaminación podemos concluir que

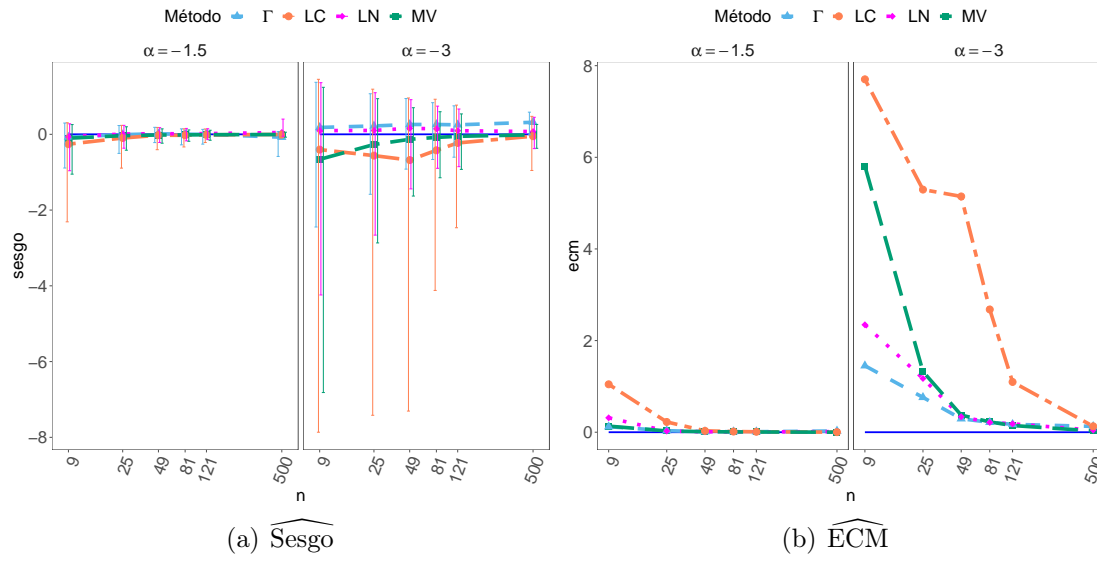


Figura 6.28: Sesgo y ECM estimados para datos sin contaminar, $L = 8$ y $\alpha = -1.5, -3$.

el estimador LN es competitivo frente al resto de los estimadores para datos sin contaminar, incluso en alguno de los casos evaluados presenta una mejor performance que el resto de los métodos estudiados.

Evaluación de los métodos de estimación para datos contaminados

En las figuras 6.29 y 6.30 se muestran los gráficos correspondientes a datos contaminados para el caso $L = 3$, mientras que las figuras 6.31 y 6.32 lo muestran para el caso $L = 8$. En esta instancia se analizó el caso 1 de contaminación, que se presentó en la sección 6.3, con $\epsilon = 0.05$. En ambos niveles de contaminación se ve con mayor claridad cómo se agrupan los estimadores por desempeño, MV y LC por un lado y los MDE por otro lado. Es notable la diferencia, tanto en sesgo como en ECM, a favor de los MDE estimadores para valores grandes de α . En la mayoría de los casos estudiados los MDE estimadores presentan intervalos de menor longitud respecto de los otros estimadores. En zonas moderadamente texturadas todos los métodos son comparables aunque los MDE estimadores presentan un menor ECM. Para zonas homogéneas MV y LC estimadores presentan menor sesgo aunque mayor variabilidad.

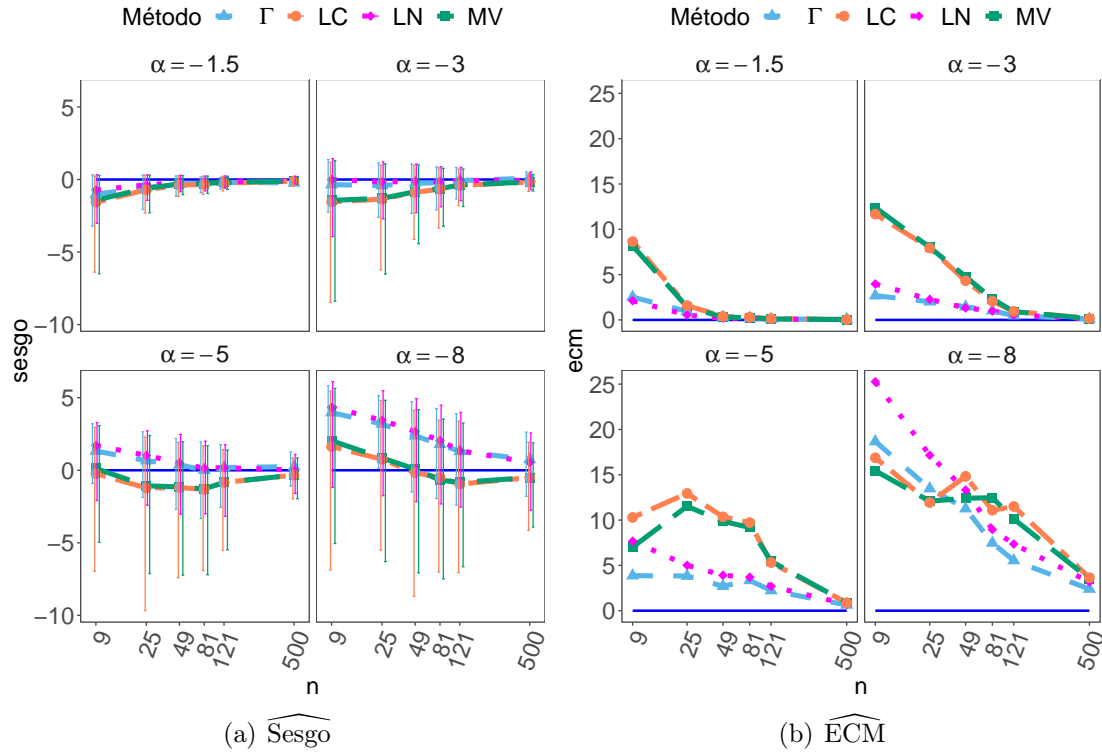


Figura 6.29: Sesgo y ECM estimados para datos contaminados, Caso 1, $\epsilon = 0.05$ y $L = 3$.

De acuerdo a este análisis el LN estimador es el que mejor desempeño presenta, tanto para datos contaminados como sin contaminar, para zonas con textura.

6.6. Función de influencia empírica

Otra forma de evaluar cómo un estimador $T_n(z_1; \dots; z_n)$ se comporta bajo contaminación es fijar $n - 1$ observaciones y permitir que una observación varíe. Esto se conoce como *Función de influencia empírica* (FIE). El problema de esta metodología es que depende de una muestra en particular. Para evitar esto Andrews et al. [3] propusieron generar una muestra formada por los i -ésimos cuantiles de la distribución, F , que se asume como modelo teórico. Estos cuantiles se definen como $z_i = F^{-1}((i - 1/3)/(n + 1/3))$, $1 \leq i \leq n - 1$.

Entonces la FIE se convierte en la *Función de influencia empírica estilizada*

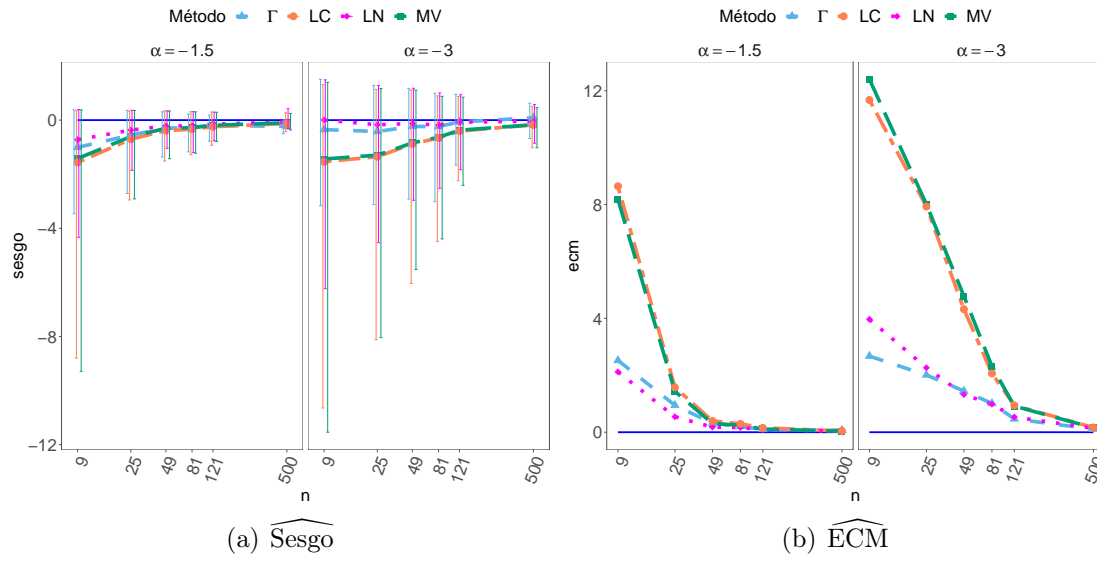


Figura 6.30: Sesgo y ECM estimados para datos contaminados, Caso 1, $\epsilon = 0.05$, $\alpha = -1.5, -3$ y $L = 3$.

(FIES) que fue usada, entre otros, por Rousseeuw y Verboven [85] para estimar en el caso de muestras de pequeño tamaño, y por Allende et al. [2] para los AM estimadores para el parámetro de textura del modelo \mathcal{G}_0 en el caso de datos de amplitud.

Los cuantiles de la distribución \mathcal{G}_I^0 se pueden obtener apelando a la relación que existe con la distribución de Fisher Snedecor.

$$F_{\alpha, \gamma, L}(z) = \Upsilon_{2L, -2\alpha}(-\alpha z / \gamma), \quad (6.6.1)$$

para todo $z > 0$, donde $\Upsilon_{2L, -2\alpha}$ es la función de distribución acumulada de una variable aleatoria con distribución Fisher-Snedecor con $2L$ y -2α grados de libertad.

De esta manera, los cuantiles de la distribución $\mathcal{G}_I^0(\alpha, \gamma, L)$ se pueden obtener a partir de $\Upsilon_{2L, -2\alpha}^{-1}$, que está disponible en la mayoría de los entornos computacionales.

Las figuras 6.33(a), 6.33(b) 6.34(a) y 6.34(b) muestran las FIES para $L = 3$, $n = 25$ y $\alpha = -1.5, -3, -5, -8$ respectivamente. La línea horizontal azul indica

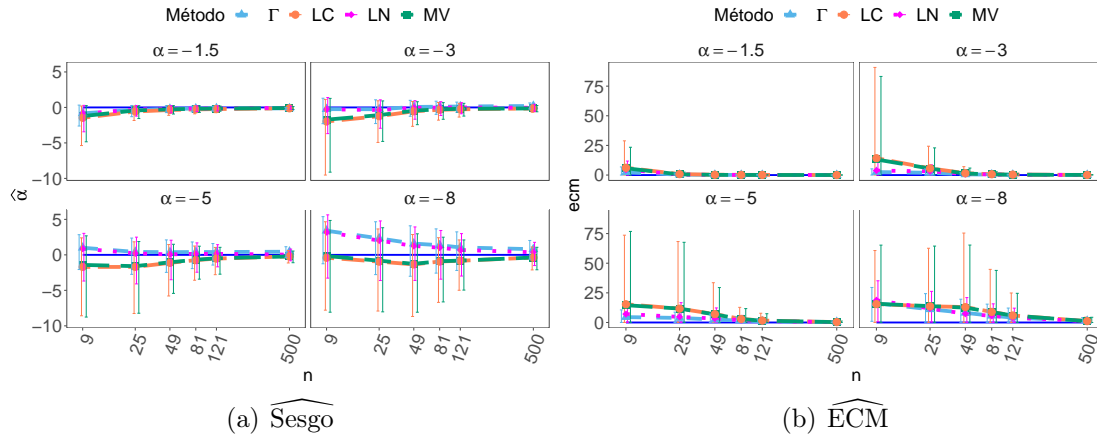


Figura 6.31: Sesgo y ECM estimados para datos contaminados, Caso 1, $\epsilon = 0.05$ y $L = 8$.

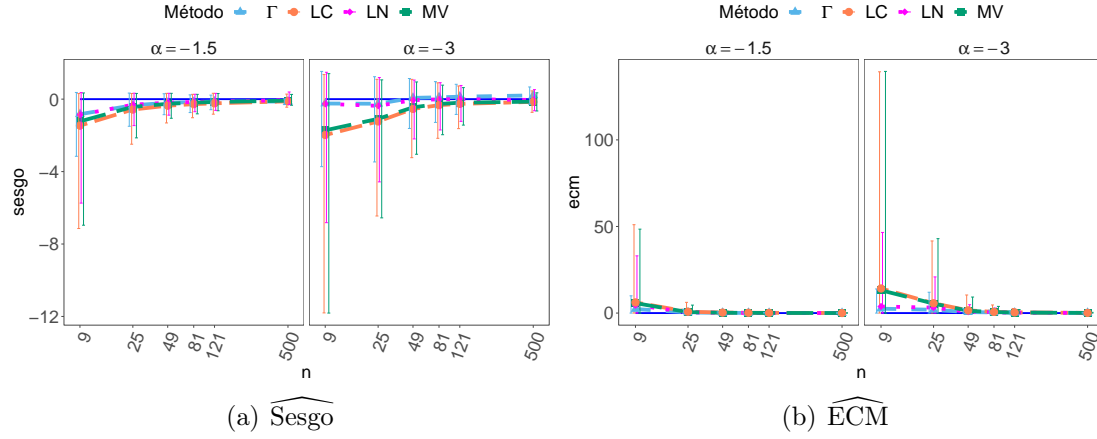


Figura 6.32: Sesgo y ECM estimados para datos contaminados, Caso 1, $\epsilon = 0.05$, $\alpha = -1.5, -3$ y $L = 8$.

el verdadero valor de α , los ticks rojos indican el máximo y mínimo cuantil, es decir, $F_{-5,4,L}^{-1}(2/(3n+1))$ y $F_{-5,4,L}^{-1}((3n-4)/(3n+1))$.

Se puede observar que $\hat{\alpha}_{MV}$, $\hat{\alpha}_{LC}$ y, por otro lado, $\hat{\alpha}_{\Gamma}$ y $\hat{\alpha}_{LN}$ tienen un comportamiento similar para $\alpha = -3, -5, -8$. Para zonas extremadamente texturadas $\hat{\alpha}_{LN}$ es el que mejor performance tiene, mientras que para el resto de las texturas tanto $\hat{\alpha}_{\Gamma}$ como $\hat{\alpha}_{LN}$ tienen una mejor performance respecto de los otros estimadores, para valores de z menores que el último cuantil. Mas aún, $\hat{\alpha}_{MV}$ no converge para zonas moderadamente heterogéneas y homogéneas, mientras que $\hat{\alpha}_{LC}$ no

converge para zonas homogéneas. Todos los estimadores se comportan de manera similar para valores grandes de z . Este comportamiento muestra la sensibilidad y la pérdida de robustez de $\hat{\alpha}_{MV}$ y $\hat{\alpha}_{LC}$ frente a los otros estimadores.

Las figuras 6.35(a), 6.35(b) 6.36(a) y 6.36(b) muestran las FIES para $L = 8$, $n = 25$ y $\alpha = -1.5, -3, -5, -8$ respectivamente. El desempeño de los estimadores para este caso es similar a lo descrito para $L = 3$. Hay que mencionar que todos los métodos convergen para este caso donde hay menor presencia de ruido speckle.

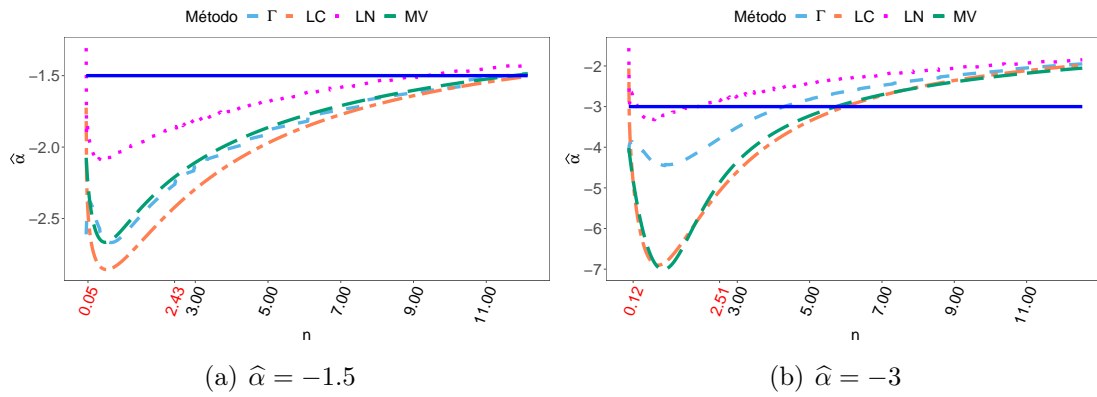


Figura 6.33: FIES para $\hat{\alpha}_{MV}$, $\hat{\alpha}_{\Gamma}$, $\hat{\alpha}_{LN}$, $\hat{\alpha}_{LC}$ para $L = 3$, $n = 25$ y $\alpha = -1.5, -3$.

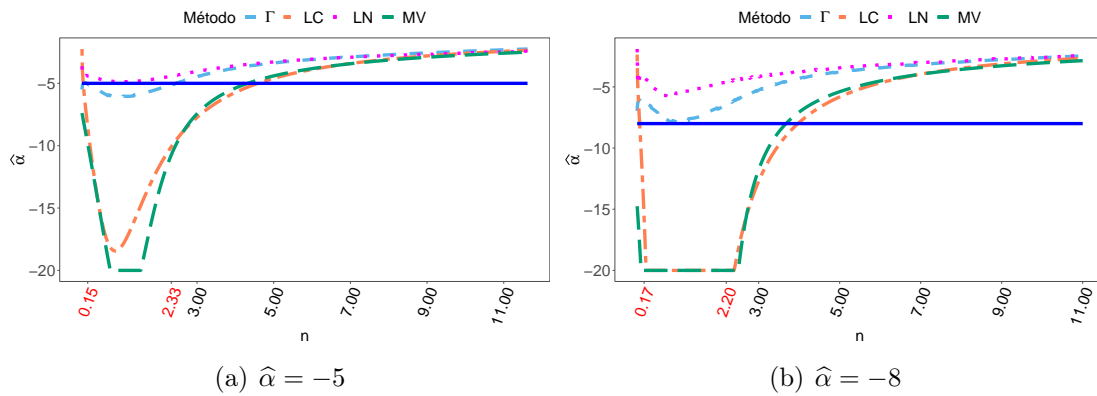


Figura 6.34: FIES para $\hat{\alpha}_{MV}$, $\hat{\alpha}_{\Gamma}$, $\hat{\alpha}_{LN}$, $\hat{\alpha}_{LC}$ para $L = 3$, $n = 25$ y $\alpha = -5, -8$.

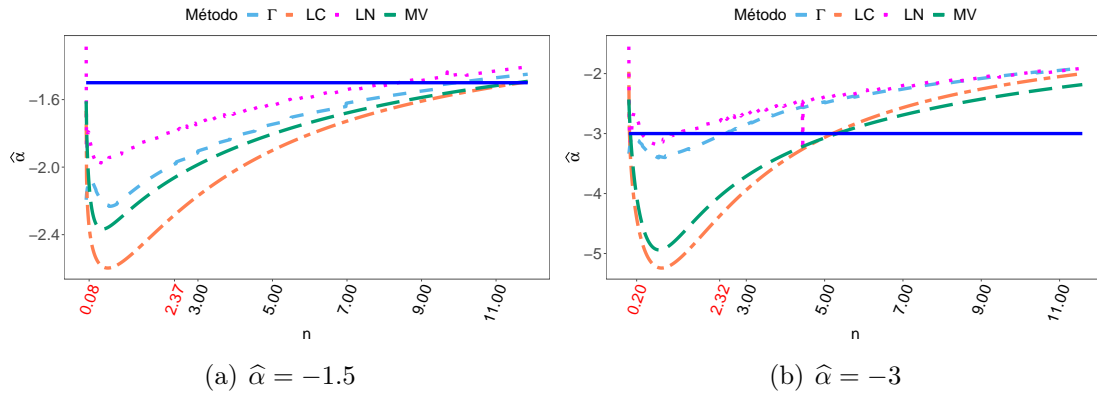


Figura 6.35: FIES para $\hat{\alpha}_{MV}$, $\hat{\alpha}_{\Gamma}$, $\hat{\alpha}_{LN}$, $\hat{\alpha}_{LC}$ para $L = 8$, $n = 25$ y $\alpha = -1.5, -3$.

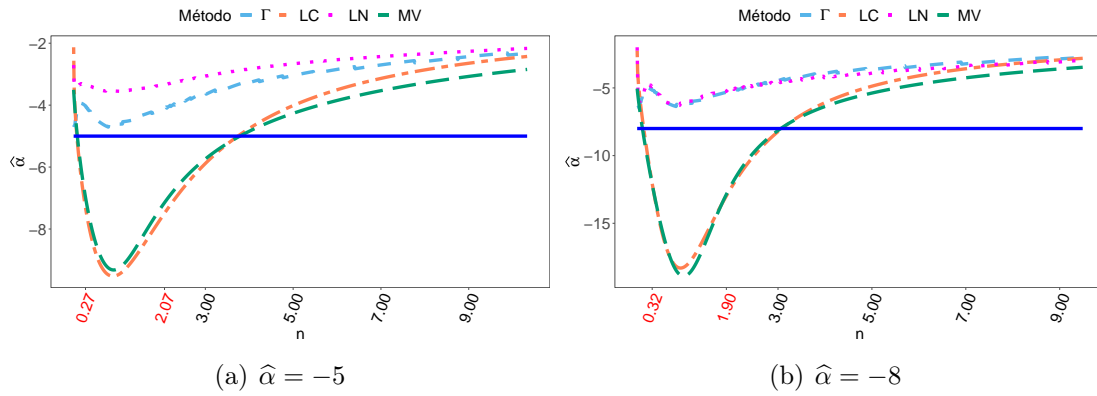


Figura 6.36: FIES para $\hat{\alpha}_{MV}$, $\hat{\alpha}_{\Gamma}$, $\hat{\alpha}_{LN}$, $\hat{\alpha}_{LC}$ para $L = 8$, $n = 25$ y $\alpha = -5, -8$.

6.7. Aplicación a una imagen real

En esta sección evaluaremos a los estimadores $\hat{\alpha}_{MV}$, $\hat{\alpha}_{\Gamma}$, $\hat{\alpha}_{LN}$, $\hat{\alpha}_{LC}$ cuando estimamos en una imagen real. En este caso utilizamos la imagen 6.22. Esta imagen es de un solo look. Para obtener una imagen multilook calculamos nuevos píxeles promediando los valores de intensidad en ventanas deslizantes, no superpuestas, de tamaño 2×2 . Esta técnica se conoce como procesamiento piramidal [1], la imagen resultante es una imagen multilook.

El número de looks se define como $E(I)^2/Var(I)$ donde I es un área homogénea. En general es un parámetro conocido, pero en el caso que no lo sea se puede estimar por el número equivalente de looks definido como $ENL = 1/\widehat{CV(I)}^2$, el

cuadrado del recíproco del coeficiente de variación muestral $\widehat{CV} = \widehat{\sigma}/\widehat{\mu}$, donde $\widehat{\sigma}$ y $\widehat{\mu}$ son el desvío estándar y la media muestral respectivamente.

Para encontrar el ENL elegimos manualmente 4 muestras de zonas homogéneas y calculamos ENL_i con $i = 1, \dots, 4$. El valor final de ENL se obtiene haciendo un promedio ponderado por el tamaño de cada muestra. El valor de ENL obtenido es 3.21. La figura 6.37 muestra la imagen ESAR multilook junto con las muestras utilizadas para obtener ENL .

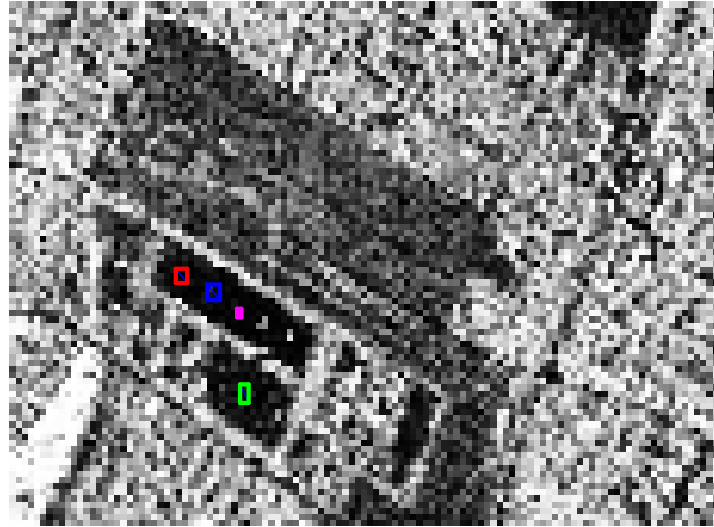


Figura 6.37: Muestras para estimación de ENL en una imagen Esar multilook.

En la figura 6.38 están graficadas tres muestras: amarilla, roja y azul. Las dos primeras de tamaño 16, la última de tamaño 12. La muestra roja es un desplazamiento respecto de la muestra amarilla en una fila de la matriz de datos de la imagen, la muestra azul está contenida dentro de la muestra amarilla. La tabla 6.9 muestra los valores de las estimaciones obtenidas con los diferentes

métodos en cada una de las muestras. Se puede observar que el valor de las estimaciones para los diferentes métodos difieren considerablemente en el caso de la muestra amarilla. Tanto MV como LC arrojan estimaciones de $\hat{\alpha}$ compatibles con una zona homogénea, mientras que $\hat{\alpha}_T$ indica que la zona es heterogénea y $\hat{\alpha}_{LN}$ está en el límite de afirmar si la zona es heterogénea o extremadamente heterogénea.

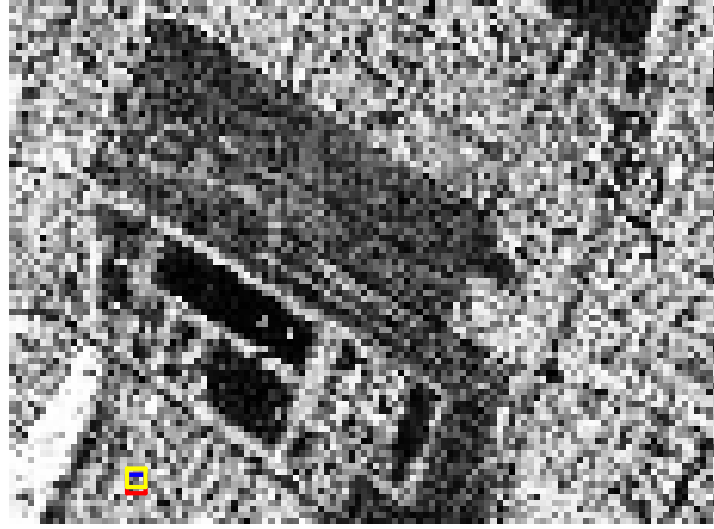


Figura 6.38: Tres muestras utilizadas para comparar $\hat{\alpha}_{MV}$, $\hat{\alpha}_T$, $\hat{\alpha}_{LN}$ y $\hat{\alpha}_{LC}$.

¿Por qué sucede esto? es la pregunta inmediata. Para responder esta pregunta volvimos a estimar en una ventana desplazada (muestra roja) y en una ventana contenida (muestra azul). Las estimaciones se presentan en la tabla 6.9.

Se puede observar que, tanto $\hat{\alpha}_{MV}$ como $\hat{\alpha}_{LC}$ han cambiado sustancialmente el valor de las estimaciones, dando ahora valores compatibles con una zona heterogénea en las muestras Roja y Azul. Si bien $\hat{\alpha}_T$ y $\hat{\alpha}_{LN}$ cambiaron de valor, no se

Tabla 6.9: Valores de $\hat{\alpha}_{MV}$, $\hat{\alpha}_{\Gamma}$, $\hat{\alpha}_{LN}$ y $\hat{\alpha}_{LC}$ para las tres muestras de la imagen 6.38.

Color	n	$\hat{\alpha}_{MV}$	$\hat{\alpha}_{\Gamma}$	$\hat{\alpha}_{LN}$	$\hat{\alpha}_{LC}$
Amarilla	16	-7.212581	-4.343894	-3.234532	-6.738052
Roja	16	-3.036808	-3.420479	-2.123289	-3.265889
Azul	12	-4.611385	-3.851962	-2.350084	-4.516257

modificó el tipo de zona que describen.

Entonces, basándonos en los resultados de las simulaciones obtenidas para las FIES correspondientes a las figuras 6.33(b) y 6.33(b) que muestran la falta de robustez de MV y LC ante una contaminación con un valor bajo, volvimos a estimar quitando la primera observación que corresponde a un valor 86 % más bajo que la media. Las estimaciones obtenidas se presentan en la tabla 6.10

Tabla 6.10: Valores de $\hat{\alpha}_{MV}$, $\hat{\alpha}_{\Gamma}$, $\hat{\alpha}_{LN}$ y $\hat{\alpha}_{LC}$ para la muestra amarilla sin el primer elemento.

n	$\hat{\alpha}_{MV}$	$\hat{\alpha}_{\Gamma}$	$\hat{\alpha}_{LN}$	$\hat{\alpha}_{LC}$
15	-5.733452	-3.970921	-3.128065	-4.802438

Nuevamente se puede observar el cambio significativo en el valor de las estimaciones de los MV y LC estimadores, dando ahora estimaciones compatibles con una zona homogénea. Esto muestra la robustez del estimador de distancias estocásticas frente a los estimadores MV y LC.

En la figura 6.39 se presentan cinco muestras que fueron seleccionadas para realizar otro análisis del desempeño de los métodos de estimación. El tamaño de las muestras es $n = 9, 25, 49, 81, 121$ que corresponden a ventanas cuadradas, una contenida dentro de la otra. En la figura 6.40 se muestra el valor de $\hat{\alpha}$ para cada método y cada muestra en función del tamaño de muestra. Se realizó un bootstrap paramétrico [28] de 2000 replicaciones para hallar un intervalo de confianza de nivel 0.95 para cada método y cada valor de n . Este tipo de bootstrap se utiliza cuando se supone que los datos provienen de un modelo conocido. Este método consiste en:

- Suponer que los datos z_1, \dots, z_n provienen de una distribución paramétrica conocida $F(\theta)$.
- Estimar θ por un estadístico $\hat{\theta}$.
- Generar R muestras bootstrap de $F(\hat{\theta})$.
- Encontrar, para cada muestra bootstrap, $\hat{\theta}^*$.

Luego de generar las muestras bootstrap encontramos el intervalo de confianza por el método del percentil. Este método utiliza los percentiles $\theta_{(\alpha)}^*$ y $\theta_{(1-\alpha)}^*$ de la distribución de $\hat{\theta}$ generada por las muestras bootstrap, donde $\theta_{(\alpha)}^*$ y $\theta_{(1-\alpha)}^*$ son los percentiles muestrales de la distribución de $\hat{\theta}$. De acuerdo a [31], cuando los datos son sesgados, este método es preferible al intervalo normal estándar.

Los valores de las estimaciones obtenidas se presentan en la tabla 6.11.

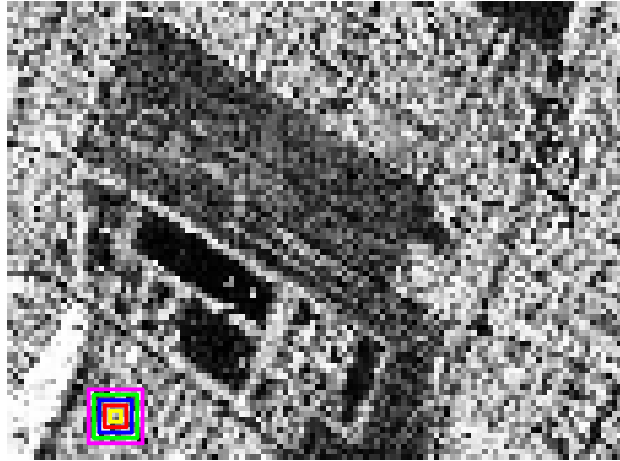


Figura 6.39: Cinco muestras utilizadas para comparar $\hat{\alpha}_{MV}$, $\hat{\alpha}_{\Gamma}$, $\hat{\alpha}_{LN}$ y $\hat{\alpha}_{LC}$.

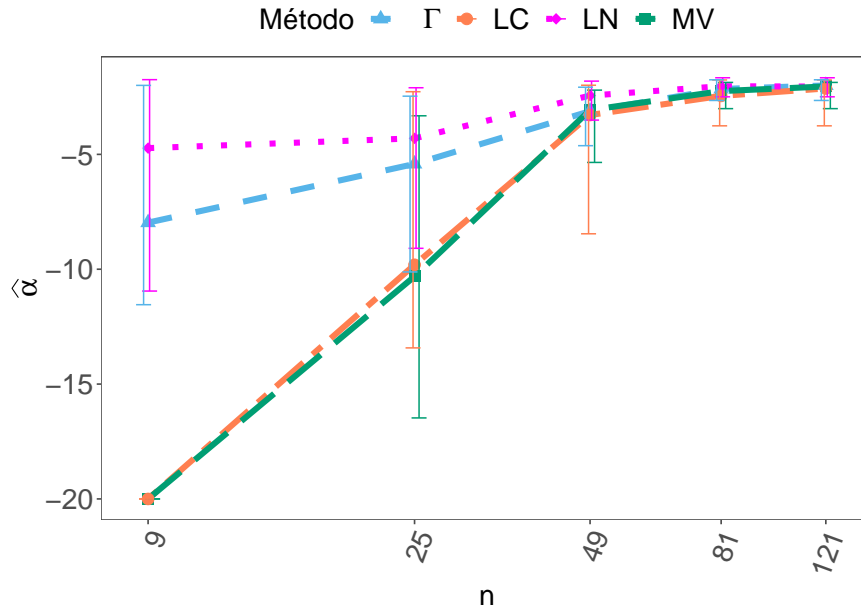


Figura 6.40: $\hat{\alpha}_{MV}$, $\hat{\alpha}_{\Gamma}$, $\hat{\alpha}_{LN}$ y $\hat{\alpha}_{LC}$ para cinco muestras variando el tamaño de las ventanas de la imagen 6.39.

Tabla 6.11: $\hat{\alpha}$ para cinco muestras variando el tamaño de las ventanas de imagen 6.39.

Color	n	$\hat{\alpha}_{MV}$	$\hat{\alpha}_{\Gamma}$	$\hat{\alpha}_{LN}$	$\hat{\alpha}_{LC}$
Amarilla	9	-20	-7.970677	-4.727529	-20
Roja	25	-10.3127	-5.417753	-4.301603	-9.806147
Azul	49	-3.072795	-3.119539	-2.440396	-3.285406
Verde	81	-2.244657	-2.152866	-2.033067	-2.4545
Magenta	121	-2.04354	-1.999166	-2.048823	-2.130614

En este nuevo análisis se pone de manifiesto el comportamiento de los estimadores bajo estudio. Se puede observar que $\hat{\alpha}_{MV}$ y $\hat{\alpha}_{LC}$ tienen un desempeño similar, y por otro lado esto mismo se pone de manifiesto con $\hat{\alpha}_{LN}$ y $\hat{\alpha}_{\Gamma}$. Además, para $n = 9$ no convergen ninguno de los métodos MV y LC. A medida que el tamaño de muestra aumenta las estimaciones se van estabilizando mostrando que el método LN es el más estable mientras que $\hat{\alpha}_{MV}$ y $\hat{\alpha}_{LC}$ son los más inestables, dando malas estimaciones para tamaño de muestras pequeño.

Se puede observar también que los intervalos de confianza correspondientes a los métodos MV y LN son en general más anchos que los de los otros métodos. Todos los intervalos van disminuyendo su longitud a medida que aumenta el tamaño de muestra, es decir, van siendo más precisos. Es importante mencionar que el intervalo de confianza correspondiente al método LN es el que menor longitud presenta en la mayoría de los casos. Cabe señalar que, en el caso donde los métodos MV y LN no convergieron no se da ningún intervalo de confianza y solo se consideraron las muestras donde todos los métodos convergieron.

6.8. Conclusiones

En este capítulo se estudia el desempeño del estimador de mínima distancia propuesto en esta tesis mediante una simulación Monte Carlo, especialmente para muestras de pequeño tamaño. Esta simulación se realizó con el fin de elegir una distancia estocástica entre las estudiadas, seleccionar el núcleo para estimar la función de densidad subyacente y la forma de elegir el ancho de banda.

A través de este estudio se concluyó que la distancia triangular es una buena elección para estimar el parámetro de textura de la distribución \mathcal{G}_I^0 utilizando estimadores de mínima distancia. Se evaluaron tres núcleos para estimar la función de densidad subyacente: IG, Γ y LN en términos de error cuadrático medio integrado, sesgo, ECM y robustez. Se mostró que el el MDE estimador junto con la distancia triangular, el núcleo IG con un ancho de banda elegido empíricamente es competitivo respecto del método de Momentos fraccionales, Máxima Verosimilitud y Logcumulantes en situaciones sin contaminación y mejora el rendimiento de estas otras técnicas en presencia de bajos niveles de contaminación. Sin embargo posee un alto error cuadrático medio integrado comparado con el que presentan los núcleos Γ y Lognormal eligiendo el ancho de banda con el método LSCV. Estos últimos núcleos tienen un mejor desempeño en cantidad de casos de convergencia frente al MV y LC estimadores, compiten con los métodos MV

y LC en términos de sesgo y ECM para datos sin contaminar, incluso mejorando su performance para algunos casos estudiados. Tienen la desventaja de tener un alto costo computacional frente a los otros métodos de estimación.

Respecto del análisis de robustez realizado mediante el estudio de la función de influencia emírica stilizada se mostró que el que mejor desempeño presenta es el núcleo LN. Este comportamiento también se observa cuando se aplican estas técnicas a una imagen real donde se puede observar que el que arrojó estimaciones más estables cuando se varía el tamaño de muestra es el método LN.

Por lo tanto, este análisis nos lleva a concluir que el MDE estimador con la distancia triangular, el núcleo LN y el método LSCV para elegir el ancho de banda es una buena elección para estimar el parámetro de textura de la distribución \mathcal{G}_I^0 para el caso de muestras de pequeño tamaño que es el punto de interés de esta tesis.

Capítulo 7

Resultados Teóricos

Los estimadores de mínima distancia (MDE) surgen como una alternativa más robusta a los estimadores paramétricos clásicos, como Máxima Verosimilitud (MV). La idea de la cual parte resulta muy natural, es encontrar el valor del vector de parámetros que minimiza la distancia entre el modelo teórico y aquel que resulta de la información muestral.

En general, los MDE estimadores se definen midiendo la discrepancia entre funciones de distribución. Wolfowitz [100], Parr and Schucany [76] son algunas de las principales referencias. Como indican Cao et al. [18] en los casos más usuales las funciones de distribución son absolutamente continuas. Ellos aplican los resultados de Parr and Schucany [75] al caso donde se miden distancias entre funciones de densidad. Consideran un estimador por núcleos simétricos para estimar la función de densidad subyacente en vez de utilizar la función de distribución empírica.

En esta capítulo probaremos la convergencia fuerte del estimador propuesto en esta tesis

$$\hat{\alpha}_T = \arg \min_{\alpha \in I} d_T(\mathcal{G}_I^0(\alpha, \gamma^*, L), \hat{f}_n), \quad (7.0.1)$$

donde $\mathcal{G}_I^0(\alpha, \gamma^*, L)$ es la función de densidad del modelo G_I^0 definida en 4.1, $\gamma^* =$

$-\alpha - 1$, $I \subseteq \mathbb{R}$, \hat{f}_n un estimador de la función de densidad subyacente usando núcleos asimétricos como los estudiados en 5.2.3 y d_T es la distancia triangular definida en 6.2.

Consistencia fuerte de $\hat{\alpha}_T$

Para probar la convergencia fuerte del estimador definido en 7.0.1 nos basaremos en el Teorema 1 de Parr et. al [75]. Hemos considerado un espacio métrico general, para incluir funciones integrables que no sean necesariamente funciones de densidad.

Sea \mathcal{F} el conjunto de las funciones no negativas e integrables, D una métrica en \mathcal{F} y $\mathcal{M} = \{f_\theta : \theta \in \Theta\} \subseteq \mathcal{F}$ una familia paramétrica. Supongamos que para $g \in \mathcal{F}$ se cumplen las siguientes condiciones:

C1. existe $\theta_* \in \Theta$ tal que $D(g, f_{\theta_*}) < D(g, f_\theta)$ para $\theta \neq \theta_*$,

C2. si $\{\theta_n\}_{n \in \mathbb{N}} \subseteq \Theta$ verifica $\lim_{n \rightarrow \infty} D(g, f_{\theta_n}) = D(g, f_{\theta_*})$, entonces $\theta_n \rightarrow \theta_*$ cuando $n \rightarrow +\infty$.

Proposición 7.0.1. Asumiendo que se verifican las condiciones C1 y C2, si $\{g_n\}_{n \in \mathbb{N}} \subseteq \mathcal{F}$ es una sucesión convergente a $g \in \mathcal{F}$ y $\{\theta_n\}_{n \in \mathbb{N}} \subseteq \Theta$ que verifica

$$\inf_{\theta \in \Theta} D(g_n, f_\theta) \leq D(g_n, f_{\theta_n}) \leq \inf_{\theta \in \Theta} D(g_n, f_\theta) + k_n, \quad (7.0.2)$$

donde $\{k_n\}_{n \in \mathbb{N}}$ es una sucesión de números positivos que cumplen que $k_n \rightarrow 0$ cuando $n \rightarrow +\infty$, entonces $\theta_n \rightarrow \theta_*$ cuando $n \rightarrow +\infty$.

Demostración. Si definimos $a_n = \inf_{\theta \in \Theta} D(g_n, f_\theta)$, como

$$D(g, f_\theta) - D(g, g_n) \leq D(g_n, f_\theta) \leq D(g, f_\theta) + D(g, g_n),$$

entonces

$$\inf_{\theta \in \Theta} D(g, f_\theta) - D(g, g_n) \leq \inf_{\theta \in \Theta} D(g_n, f_\theta) \leq \inf_{\theta \in \Theta} D(g, f_\theta) + D(g, g_n),$$

de la condición C1 se obtiene $D(g, f_{\theta_*}) - D(g, g_n) \leq a_n \leq D(g, f_{\theta_*}) + D(g, g_n)$. Dado que $g_n \rightarrow g$, tenemos que $a_n \rightarrow D(g, f_{\theta_*})$.

Por la definición de ínfimo y por la condición 7.0.2 se cumple que $a_n \leq D(g_n, f_{\theta_n}) \leq a_n + k_n$, entonces vemos que $0 \leq D(g_n, f_{\theta_n}) - a_n \leq k_n$, y por lo tanto, $D(g_n, f_{\theta_n}) \rightarrow D(g, f_{\theta_*})$.

Entonces

$$\begin{aligned} |D(g, f_{\theta_n}) - D(g, f_{\theta_*})| &\leq |D(g, f_{\theta_n}) - D(g_n, f_{\theta_n})| + |D(g_n, f_{\theta_n}) - D(g, f_{\theta_*})| \\ &\leq D(g, g_n) + |D(g_n, f_{\theta_n}) - D(g, f_{\theta_*})| \rightarrow 0. \end{aligned}$$

Luego, por la condición C2, se deduce $\theta_n \rightarrow \theta_*$. □

Entonces, ¿cómo aplicamos este resultado a nuestro caso? Definimos:

- $\mathcal{F} = \{f \in L^1(0, +\infty) / f \geq 0\}$.
- $D_T(f, g) = d_T^{1/2}(f, g)$ para $f, g \in \mathcal{F}$ que, de acuerdo a la proposición 5.2.1, resulta una distancia en \mathcal{F} .
- $\{\theta_n\}_{n=1}^\infty = \{\hat{\alpha}_n\}_{n=1}^\infty$ los estimadores de mínima distancia definidos, para cada n como

$$\begin{aligned} \hat{\alpha}_n &= \arg \min_{\alpha \in I} D_T(\mathcal{G}_I^0(\alpha, \gamma^*, L), \hat{f}_n) \\ &= \arg \min_{\alpha \in I} d_T(\mathcal{G}_I^0(\alpha, \gamma^*, L), \hat{f}_n), \end{aligned}$$

con $I \subseteq \mathbb{R}$ un intervalo, \hat{f}_n un estimador no paramétrico de la densidad con núcleos asimétricos como los definidos en el capítulo 6. Las igualdades

anteriores valen porque d_T es una transformación monótona creciente de D_T .

Con estas definiciones se cumple lo siguiente.

- La condición C1, dado que D_T es una distancia y $\mathcal{G}_I^0(\alpha, \gamma^*, L)$ un elemento de la familia \mathcal{F} . Esta condición es equivalente a la hipótesis iii) del teorema de Parr et al. [75].
- Por (5.2.10) $d_T(\mathcal{G}_I^0, \hat{f}_n) \leq \| \mathcal{G}_I^0 - \hat{f}_n \|_1$. De acuerdo al teorema 5.2.2 si $\lim_{n \rightarrow \infty} \frac{nb^{2r_1}}{\ln n} = +\infty$, $\| \mathcal{G}_I^0 - \hat{f}_n \|_1 \xrightarrow{cs} 0$ cuando $n \rightarrow +\infty$ con r_1 el establecido en cada caso de acuerdo al núcleo elegido. Por lo tanto, bajo esta condición sobre el ancho de banda $D_T(\mathcal{G}_I^0, \hat{f}_n) \xrightarrow{cs} 0$. Esto corresponde a la hipótesis i) de [75].
- La hipótesis ii) de [75] se verifica ya que D_T es una distancia.

Falta probar la hipótesis iv) de [75] para poder concluir que $\hat{\alpha}_n \xrightarrow{cs} \alpha_*$. Esto lo probaremos a partir de los siguientes resultados.

Sea $f_{\mathcal{G}_I^0(\alpha, \gamma, L)}$ la función de densidad correspondiente al modelo \mathcal{G}_I^0 como la definida en 4.1.1. Recordemos que el espacio paramétrico está formado por los valores de $L \geq 1$, $\gamma > 0$ y $\alpha < -1$ para que exista primer momento finito. A continuación se asume L conocido y $\gamma^* = -\alpha - 1$. De esta manera estamos considerando un modelo reducido a un parámetro, entonces la función de densidad para este caso la llamaremos f_α y queda definida como

$$f_\alpha(z) = \frac{L^L \Gamma(L - \alpha)}{(-\alpha - 1)^\alpha \Gamma(-\alpha) \Gamma(L)} \cdot \frac{z^{L-1}}{((-\alpha - 1) + zL)^{L-\alpha}}. \quad (7.0.3)$$

Proposición 7.0.2. La aplicación del intervalo $(-\infty, -1)$ en $L^1(0, \infty)$ dada por $\alpha \mapsto f_\alpha$ es continua.

Demostración. Por la continuidad de $f_\alpha(z)$ en α , si $\alpha_n \rightarrow \alpha$, entonces $f_{\alpha_n}(z) \rightarrow f_\alpha(z)$ para todo $z > 0$. Siendo f_{α_n} y f_α densidades de probabilidad, por el Teorema de Scheffé [89] obtenemos la convergencia en $L^1(0, \infty)$. \square

Proposición 7.0.3. Sea $\alpha_* \in (-\infty, -1)$ fijo, si la sucesión $\{\alpha_n\}_{n \geq 1} \subset (-\infty, -1)$ verifica $\lim_{n \rightarrow \infty} d_T(f_{\alpha_*}, f_{\alpha_n}) = 0$ siendo d_T la distancia estocástica definida en 5.2.5, entonces $\alpha_n \rightarrow \alpha_*$.

Demostración. Supongamos que $\alpha_n \not\rightarrow \alpha_*$, vamos a probar que existe una sub-sucesión $\{\alpha_{n_k}\}_{k \geq 1}$ tal que $d_T(f_{\alpha_*}, f_{\alpha_{n_k}}) \not\rightarrow 0$, lo que representa una contradicción. Como $\alpha_n \not\rightarrow \alpha_*$, existe una sub-sucesión $\{\alpha_{n_k}\}_{k \geq 1}$ que verifica alguna de las siguientes afirmaciones:

- $\alpha_{n_k} \rightarrow \alpha_0$, donde $\alpha_0 \in (-\infty, -1)$ y $\alpha_0 \neq \alpha_*$.
- $\alpha_{n_k} \rightarrow -1$,
- $\alpha_{n_k} \rightarrow \infty$.

En el primer caso, como consecuencia de la proposición 7.0.2 que muestra la continuidad respecto al parámetro α , se verifica

$$\lim_{k \rightarrow \infty} d_T(f_{\alpha_*}, f_{\alpha_{n_k}}) = d_T(f_{\alpha_*}, f_{\alpha_0}) > 0.$$

En el caso siguiente, por la proposición 4.1.3 tenemos que $f_{\alpha_{n_k}} \rightarrow 0$ uniformemente en $[z_1, z_2]$, por lo tanto

$$\begin{aligned} \liminf_{k \rightarrow \infty} d_T(f_{\alpha_*}, f_{\alpha_{n_k}}) &\geq \liminf_{k \rightarrow \infty} \int_{z_1}^{z_2} \frac{(f_{\alpha_*}(z) - f_{\alpha_{n_k}}(z))^2}{f_{\alpha_*}(z) + f_{\alpha_{n_k}}(z)} dz \\ &= \int_{z_1}^{z_2} f_{\alpha_*}(z) dz > 0. \end{aligned}$$

Por último, siendo que $f_{\alpha_*} \neq f_{\Gamma(L,L)}$ donde $f_{\Gamma(L,L)}$ es la función de densidad del modelo $\Gamma(L, L)$, existe $[z_1, z_2] \subset (0, \infty)$, tal que

$$\int_{z_1}^{z_2} \frac{(f_{\alpha_*}(z) - f_{\Gamma(L,L)}(z))^2}{f_{\alpha_*}(z) + f_{\Gamma(L,L)}(z)} dz > 0,$$

como $f_{\alpha_{n_k}} \rightarrow f_{\Gamma(L,L)}$ uniformemente en $[z_1, z_2]$, el

$$\begin{aligned} \liminf_{k \rightarrow \infty} d_T(f_{\alpha_*}, f_{\alpha_{n_k}}) &\geq \liminf_{k \rightarrow \infty} \int_{z_1}^{z_2} \frac{(f_{\alpha_*}(z) - f_{\alpha_{n_k}}(z))^2}{f_{\alpha_*}(z) + f_{\alpha_{n_k}}(z)} dz \\ &= \int_{z_1}^{z_2} \frac{(f_{\alpha}(z) - f_{\Gamma(L,L)}(z))^2}{f_{\alpha}(z) + f_{\Gamma(L,L)}(z)} dz > 0. \end{aligned}$$

Esto muestra que $d_T(f_{\alpha_*}, f_{\alpha_{n_k}}) \not\rightarrow 0$. La contradicción proviene de suponer que $\alpha_n \not\rightarrow \alpha_*$. \square

Luego, la proposición 7.0.3 prueba que se verifica la condición 2 que corresponde a la hipótesis iv) de [75]. Por lo tanto $\hat{\alpha}_n \xrightarrow{cs} \alpha_*$. Es decir, los estimadores MDE del parámetro de textura del modelo \mathcal{G}_I^0 son fuertemente consistentes.

Capítulo 8

Conclusiones

Esta tesis propone un nuevo estimador para el parámetro de textura de la distribución \mathcal{G}_I^0 basado en la minimización de una distancia estocástica entre el modelo (que depende de dicho parámetro) y una estimación de la función de densidad subyacente utilizando núcleos asimétricos \hat{f} . El objetivo es obtener un estimador que tenga buen desempeño para muestras de pequeño tamaño, sea competitivo con los estimadores de Máxima Verosimilitud (MV) y Logcumulantes (LC) que son los utilizados en la literatura y que tenga una buena performance frente a la presencia de diferentes niveles de contaminación.

En una primera instancia se consideró el núcleo inverso Gaussiano (IG) para estimar \hat{f} con un ancho de banda encontrado empíricamente, y se comparó su desempeño con los estimadores que provienen del método MV, LC y Momento Fraccionario de orden 1/2 en términos de sesgo, error cuadrático medio (ECM), tiempo de procesamiento y bajo contaminación. En esta oportunidad se consideraron tres casos de contaminación y se observó que tanto Momentos Fraccionarios como LC presentaron problemas de convergencia. Además, Momentos Fraccionarios presentó mayor sesgo y ECM en la mayoría de los casos analizados.

Definimos tres modelos de contaminación inspirados en situaciones reales para evaluar el impacto de los valores atípicos en el desempeño de los estimadores. Analizando estos escenarios se observó que

- Caso 1: Independientemente de la intensidad de la contaminación, cuanto mayor es el número de looks menor es la cantidad de casos donde se presentan problemas de convergencia.
- Casos 2 y 3: el porcentaje de situaciones para las cuales no hay convergencia aumenta con el nivel de contaminación y se reduce con el valor de α .

Entonces, en este primer trabajo se observó que, en el caso de un solo look, el estimador propuesto no presenta resultados excelentes. En el caso multilook el nuevo estimador presenta buenas propiedades según lo medido por su sesgo y ECM. Es competitivo con MV y LC estimadores en situaciones sin contaminación, y supera el desempeño de estas métodos incluso en presencia de pequeños niveles de contaminación. Por este motivo se decidió mejorar la propuesta.

Además se investigaron más a fondo las dificultades de estimar el parámetro de textura de la distribución \mathcal{G}_I^0 y se justificaron a través de nuevos resultados teóricos que indican que esta distribución tiene colas pesadas.

En un segundo estudio se comparó la performance del núcleo IG con el LN y el Γ analizando el error cuadrático medio integrado (MISE). Se observó que el núcleo IG es el que mayor MISE presentó frente a los otros núcleos. Por este motivo no se lo consideró en el nuevo análisis realizado.

Se volvió a realizar un análisis similar respecto del primer trabajo considerando en esta oportunidad los estimadores que provienen del método de MV, LC y el que resulta de la minimización de la distancia triangular utilizando los núcleos Γ y LN para estimar \hat{f} . Se fijó un nuevo criterio de convergencia para todos los métodos y se comparó la cantidad de casos de no convergencia. Se estudió el costo computacional y, en esta oportunidad, se incorporó la Función de Influencia Empírica para evaluar la robustez de estos estimadores, como así también el caso 1 de contaminación. Asimismo se utilizó el método LSCV para elegir el ancho de banda que interviene en la estimación de \hat{f} .

En los estudios por Monte Carlo se observó que:

- Por un lado, los métodos MV y LC tienen comportamiento similar entre sí y, por el otro, los estimadores de mínima distancia utilizando núcleos LN y Γ tanto para datos sin contaminar como para datos contaminados. Incluso este comportamiento se presenta al momento de analizar la Función de Influencia Empírica.
- Para datos sin contaminar y con textura, el estimador propuesto es competitivo respecto de los MV y LC estimadores en términos de sesgo y ECM. Sin embargo supera a los estimadores MV y LC en presencia de contaminación, ya sea utilizando núcleo Γ o LN, aunque este último es el que presenta un comportamiento más estable.
- El tiempo computacional de los estimadores de mínima distancia es sensiblemente mayor respecto de los MV y LC estimadores.
- La cantidad de casos de no convergencia que presenta nuestra propuesta es significativamente menor respecto de los MV y LC estimadores.

En la aplicación a una imagen real se observó que el estimador de mínima distancia utilizando núcleo Lognormal es el que mejor desempeño mostró, superando a MV y LC estimadores.

Además, se completó el análisis con la demostración de la consistencia fuerte del estimador de mínima distancia bajo la distancia triangular y *widehatf* por núcleos asimétricos como los trabajados en esta tesis.

Por lo anteriormente expuesto, el nuevo estimador presenta buenas propiedades medidas por su sesgo, error cuadrático medio y cantidad de casos de no convergencia. Es competitivo con los MV estimadores y LC estimadores en situaciones sin contaminación, y supera a estos métodos en presencia de pequeños niveles de contaminación.

Por esta razón nuestra conclusión es que es aconsejable usar $\hat{\alpha}_T$ con núcleo LN, especialmente cuando se usan muestras pequeñas y/o cuando existe la posibilidad de tener datos contaminados. Si bien el costo computacional adicional incurrido

al usar este estimador es mucho mayor que el requerido para calcular $\hat{\alpha}_{\text{ML}}$, sus ventajas son más importantes que este aumento en el tiempo de procesamiento.

Apéndice A

Información computacional

Las simulaciones se realizaron utilizando el lenguaje y el entorno R para computación estadística [81]. La función `adaptIntegrate` del paquete `cubature` se utilizó para realizar la integración numérica requerida para evaluar la distancia triangular, el algoritmo utilizado es una integración multidimensional adaptativa sobre hipercubos. Para encontrar numéricamente $\hat{\alpha}_{LC}$ utilizamos la función “`uniroot`” implementada en R.

Una parte de las simulaciones se realizó en una plataforma compuesta por un Intel(R) Core i7 con 8 GB de memoria y 64 bits Windows 7. Otra parte se hizo en un equipo similar pero con 16 GB de memoria RAM. Los códigos y los datos están disponibles a solicitud del autor correspondiente.

Bibliografía

- [1] E. Adelson, C. Anderson, J. Bergen, P. Burt, and J. Ogden. Pyramid methods in image processing. 1984.
- [2] H. Allende, A. C. Frery, J. Galbiati, and L. Pizarro. M-estimators with asymmetric influence functions: the GA0 distribution case. *Journal of Statistical Computation and Simulation*, 76(11):941–956, 2006. doi: 10.1080/10629360600569154.
- [3] D. Andrews, P. Bickel, F. Hampel, P.J. Huber, W. Rogers, and J. Tuckey. *Robust Estimates of Location: Survey and Advances*. New Jersey: Princeton University Press, 1972.
- [4] S Anfinsen, A Doulgeris, and T Eltoft. Estimation of the equivalent number of looks in polarimetric synthetic aperture radar imagery. *IEEE Transactions on Geoscience and Remote Sensing*, 47(11):3795–3809, 2009.
- [5] S.N. Anfinsen and T. Eltoft. Application of the matrix-variate Mellin transform to analysis of polarimetric radar images. *IEEE Transactions on Geoscience and Remote Sensing*, 49(6):2281–2295, 2011. ISSN 0196-2892. doi: 10.1109/TGRS.2010.2103945.
- [6] S. Aviyente, F. Ahmad, and M. G. Amin. Information theoretic measures for change detection in urban sensing applications. In *IEEE Workshop on Signal Processing Applications for Public Security and Forensics*, pages 1–6, 2007.

- [7] R. Beran. Minimum hellinger distance estimates for parametric models. *Ann. Statist.*, 5(3):445–463, 05 1977. doi: 10.1214/aos/1176343842. URL <https://doi.org/10.1214/aos/1176343842>.
- [8] Y. Bian. Polarimetric sar statistical analysis using alpha-stable distribution and its application in optimal despeckling. *International Journal of Remote Sensing*, 34(19):6796–6836, 2013. doi: 10.1080/01431161.2013.808778. URL <http://dx.doi.org/10.1080/01431161.2013.808778>.
- [9] D. Boos. Stochastic approaches of minimum distance method for region based classification. *Journal of the American Statistical Association*, 76(375):663–670, 1981. doi: 10.1080/01621459.1981.10477701. URL <http://amstat.tandfonline.com/doi/abs/10.1080/01621459.1981.10477701>.
- [10] T. Bouezmarni and O. Scaillet. Consistency of asymmetric kernel density estimators and smoothed histograms with application to income data. *Econometric Theory*, 21(2):390–412, 2005.
- [11] E. Bratsolis and M. Sigelle. Kullback-Leibler divergence and Markov random fields for speckled image restoration. In *Seventh International Symposium on Signal Processing and Its Applications, 2003. Proceedings*, volume 1, pages 425–428, 2003. doi: 10.1109/ISSPA.2003.1224731.
- [12] B. M. Brown and S. X. Chen. Beta-bernstein smoothing for regression curves with compact support. *Scandinavian Journal of Statistics*, 26(1):47–59, 1999. doi: 10.1111/1467-9469.00136. URL <https://onlinelibrary.wiley.com/doi/abs/10.1111/1467-9469.00136>.
- [13] S. Buckland. Monte carlo methods for confidence interval estimation using the bootstrap technique. *Journal of Applied Statistics*, 10(2):194–212, 1983. doi: 10.1080/02664768300000017. URL <https://doi.org/10.1080/02664768300000017>.

- [14] M. Buemi, N. Goussies, J. Jacobo, and M. Mejail. Sar image segmentation using level sets and region competition under the \mathcal{G}^h model. In Eduardo Bayro-Corrochano and Jan-Olof Eklundh, editors, *Progress in Pattern Recognition, Image Analysis, Computer Vision, and Applications*, pages 153–160, Berlin, Heidelberg, 2009. Springer Berlin Heidelberg. ISBN 978-3-642-10268-4.
- [15] F. Bujor, E. Trouve, L. Valet, J-M Nicolas, and J.-P. Rudant. Application of log-cumulants to the detection of spatiotemporal discontinuities in multitemporal SAR images. *IEEE Transactions on Geoscience and Remote Sensing*, 42(10):2073–2084, 2004. ISSN 0196-2892.
- [16] O. H. Bustos, M. M. Lucini, and A. C. Frery. M-estimators of roughness and scale for GA0-modelled SAR imagery. *EURASIP Journal on Applied Signal Processing*, 1(1):105–114, 2002. doi: 10.1155/S1110865702000392.
- [17] R. Byrd, P. Lu, J. Nocedal, and C. Zhu. A limited memory algorithm for bound constrained optimization. *SIAM Journal of Scientific Computing*, 16:1190–1208, 9 1995. ISSN 1064-8275. doi: 10.1137/0916069.
- [18] R. Cao, A. Cuevas, and R. Fraiman. Minimum distance density-based estimation. *Computational Statistics & Data Analysis*, 20(6):611–631, 1995.
- [19] J. Casseti, J. Gambini, and A. Frery. Estimación de parámetros utilizando distancias estocásticas para datos con ruido speckle. In *Argentine Symposium of Technology in JAIIO*, 2013.
- [20] J. Casseti, J. Gambini, and A. C. Frery. Parameter estimation in SAR imagery using stochastic distances. In *Proceedings of The 4th Asia-Pacific Conference on Synthetic Aperture Radar (APSAR)*, Tsukuba, Japan, 2013.
- [21] D. Chan, J. Casseti, and A. C. Frery. Texture parameter estimation in monopolarized sar imagery, for the single look case, using extreme value

- theory. In *2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, pages 3266–3269, July 2016. doi: 10.1109/IGARSS.2016.7729845.
- [22] S. Chen. Beta kernel estimators for density functions. *Computational Statistics and Data Analysis*, 31(2):131–145, 1999.
- [23] S. Chen. Probability density functions estimation using Gamma kernels. *Annals of Institute Statistical Mathematics*, 52(3):471–480, 2000.
- [24] F. Cribari-Neto, A. C. Frery, and M. F. Silva. Improved estimation of clutter properties in speckled imagery. *Computational Statistics and Data Analysis*, 40(4):801–824, 2002. doi: 10.1016/S0167-9473(02)00102-0.
- [25] I. Csiszár. Information-type measures of difference of probability distributions and indirect observations. *Studia Scientiarum Mathematicarum Hungarica*, 2:299–318, 1967.
- [26] E. J. Snell D. R. Cox. A general definition of residuals. *Journal of the Royal Statistical Society. Series B (Methodological)*, 30(2):248–275, 1968. ISSN 00359246.
- [27] J. Danielsson, B. N. Jorgensen, M. Sarma, and C. G. de Vries. Comparing downside risk measures for heavy tailed. *Economics Letters*, 92(2):202–208, 2006.
- [28] A. Davison and D. Hinkley. *Bootstrap Methods and their Application*. Cambridge University Press, 1997.
- [29] Francial Giscard Baudin Libengué Dobélé-Kpoka. *Méthode non-paramétrique des noyaux associés mixtes et applications*. PhD thesis, École Doctorale Carnot-Pasteur – École Doctorale des Sciences et Technologie, June 2013.

- [30] S Dragomir. Upper and lower bounds for csiszar f-divergence in terms of hellinger discrimination and applications. *Nonlinear Analysis Forum*, 01 2002.
- [31] B. Efron and R. Tibshirani. *An Introduction to the Bootstrap*. Number 57 in Monographs on Statistics and Applied Probability. Chapman & Hall/CRC, Boca Raton, Florida, USA, 1993.
- [32] P. Eslinger and W. Woodward. Minimum hellinger distance estimation for normal models. *Journal of Statistical Computation and Simulation*, 39(1-2): 95–114, 1991. doi: 10.1080/00949659108811342.
- [33] Y. Fei, G. Rong, B. Wang, and W. Wang. Parallel l-bfgs-b algorithm on gpu. *Computers & Graphics*, 40:1 – 9, 2014. ISSN 0097-8493. doi: <https://doi.org/10.1016/j.cag.2014.01.002>. URL <http://www.sciencedirect.com/science/article/pii/S0097849314000119>.
- [34] C. Freitas, A. C. Frery, and A. Correia. The polarimetric \mathcal{G} distribution for sar data analysis. *Environmetrics*, 16(1):13–31, 2005. doi: 10.1002/env.658. URL <https://onlinelibrary.wiley.com/doi/abs/10.1002/env.658>.
- [35] A. C. Frery and J. Wu. Operational statistics for SAR imagery, 2019. URL <https://github.com/acfrery/Statistics-SAR-Intensity.git>.
- [36] A. C. Frery, H.J. Müller, C. C. F. Yanasse, and S. J. S. Sant’Anna. A model for extremely heterogeneous clutter. *IEEE Transactions on Geoscience and Remote Sensing*, 35(3):648–659, 1997.
- [37] A. C. Frery, A. H. Correia, C. D. Rennó, C. C. Freitas, J. Jacobo-Berlles, M. E. Mejail, and K. L. P. Vasconcellos. Models for Synthetic Aperture Radar Image Analysis. *Resenhas (IME-USP)*, 4(1):45–77, 1999.
- [38] A. C. Frery, R. J. Cintra, and A. D. C. Nascimento. Entropy-based sta-

- tistical analysis of PolSAR data. *IEEE Transactions on Geoscience and Remote Sensing*, 51(6), in press. doi: 10.1109/TGRS.2012.2222029.
- [39] J. Gambini, M. Mejail, J. Jacobo-Berlles, and A. Frery. Accuracy of local edge detection in speckled imagery. *Statistics & Computing*, 18(1):15–26, 2008.
- [40] J. Gambini, J. Cassetti, M. Lucini, and A. Frery. Parameter estimation in SAR imagery using stochastic distances and asymmetric kernel. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 8:365–375, 2015.
- [41] J. W. Goodman. Statistical properties of laser speckle patterns. In J. C. Dainty, editor, *Laser Speckle and Related Phenomena*, pages 9–75. Springer Berlin Heidelberg, Berlin, Heidelberg, 1975. ISBN 978-3-662-43205-1. doi: 10.1007/978-3-662-43205-1_2. URL https://doi.org/10.1007/978-3-662-43205-1_2.
- [42] J. W. Goodman. *Satatistical Optics, Pure and Applied Optics*. Wiley, New York, USA, 1985.
- [43] R. F. Green. Outlier-prone and outlier-resistant distributions. *American Statistical Association*, 71(354):502–505, 1976. ISSN 0162-1459.
- [44] P Hall and J. S. Marron. Local minima in cross-validation functions. *Journal of the Royal Statistical Society, Series B*, 53:245–252, 1991.
- [45] Iman Heidarpour Shahrezaei and Hyun-cheol Kim. Resolutional analysis of multiplicative high-frequency speckle noise based on sar spatial despeckling filter implementation and selection. *Remote Sensing*, 11(9), 2019. ISSN 2072-4292. doi: 10.3390/rs11091041. URL <https://www.mdpi.com/2072-4292/11/9/1041>.

- [46] T.P. Hettmansperger, I. Hueter, and J. Hüsler. Minimum distance estimators. *Journal of Statistical Planning and Inference*, 41(3):291 – 302, 1994. ISSN 0378-3758. doi: [https://doi.org/10.1016/0378-3758\(94\)90025-6](https://doi.org/10.1016/0378-3758(94)90025-6). URL <http://www.sciencedirect.com/science/article/pii/0378375894900256>.
- [47] M. Hirukawa. *Asymmetric Kernel Smoothing. Theory and Applications in Economics and Finance*. Springer Briefs in Statistics. Springer Singapore, 2018.
- [48] R. Horn. The DLR airborne SAR project E-SAR. In *Geoscience and Remote Sensing Symposium*, volume 3, pages 1624–1628. IEEE Press, 1996. doi: 10.1109/IGARSS.1996.516751.
- [49] P. Huber. The 1972 wald lecture robust statistics: A review. *The Annals of Mathematical Statistics*, 43(4):1041–1067, 08 1972. doi: 10.1214/aoms/1177692459. URL <https://doi.org/10.1214/aoms/1177692459>.
- [50] Invap-Saocom. Misión saocom, . URL [URL{http://www.invap.com.ar/es/espacial-y-gobierno/proyectos-espaciales/satelites-saocom.html/}](http://www.invap.com.ar/es/espacial-y-gobierno/proyectos-espaciales/satelites-saocom.html/). [Web; accedido el 21-10-2019].
- [51] Invap-Saocom. Invap en la misión saocom, . URL [URL{http://saocom.invap.com.ar/}](http://saocom.invap.com.ar/). [Web; accedido el 21-10-2019].
- [52] J. Jacobo Berlles. *Nuevas familias de distribuciones polarimétricas para imágenes SAR*. Tesis, Facultad de Ciencias Exactas y Naturales. Universidad de Buenos Aires, 2005. URL http://hdl.handle.net/20.500.12110/tesis_n3834_JacoboBerlles.
- [53] E. Jakeman and R. J. A. Tough. Generalized K distribution: a statistical model for weak scattering. *Journal of the Optical Society of America A*, 4(9):1764–1772, September 1987.

- [54] X. Jin and J. Kawczak. Birnbaum-Saunders and lognormal kernel estimators for modelling durations in high frequency financial data. *Annals of Economics and Finance*, 4(1):103–124, 2003. URL <http://EconPapers.repec.org/RePEc:cuf:journl:y:2003:v:4:i:1:p:103-124>.
- [55] M. C. Jones. Simple boundary correction for density estimation. *Statistics and Computing*, 3(3):135–146, 1993.
- [56] B. Jorgensen. *Statistical Properties of the Generalized Inverse Gaussian Distribution*, volume 9 of *Lecture Notes in Statistics*. Springer-Verlag New York, 1982. ISBN 978-1-4612-5698-4. doi: 10.1007/978-1-4612-5698-4.
- [57] S. Khan and R. Guida. Single-look polsar statistical analysis using fractional moments of polarimetric whitening filter. In *2013 IEEE International Geoscience and Remote Sensing Symposium - IGARSS*, pages 3187–3190, July 2013. doi: 10.1109/IGARSS.2013.6723504.
- [58] S. Khan and R. Guida. Application of Mellin-Kind statistics to polarimetric \mathcal{G} distribution for SAR data. *IEEE Transactions on Geoscience and Remote Sensing*, 52(6):3513–3528, June 2014. ISSN 0196-2892.
- [59] K.Jain and A. Srivastava. On symmetric information divergence measures of csiszar’s f - divergence class. *Journal of Applied Mathematics, Statistics and Informatics*, 3(1):85–102, 2007.
- [60] S. Kullback and R. Leibler. On information and sufficiency. *Annals of Mathematical Statistics*, 22:79–86, 1951.
- [61] Jet Propulsion Laboratory-Nasa. Mision to earth, seasat. URL [URL{https://www.jpl.nasa.gov/missions/seasat/}](https://www.jpl.nasa.gov/missions/seasat/). [Web; accedido el 21-10-2019].
- [62] Jong-Sen Lee and Eric Pottier. *Polarimetric Radar Imaging : From basics to applications*. CRC Press., Boca Ratón, 02 2009. ISBN 9781420054989. doi: 10.1201/9781420054989.

- [63] 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 Journal of Selected Topics in Signal Processing*, 5(3):386–397, June 2011. ISSN 1932-4553. doi: 10.1109/JSTSP.2011.2138675.
- [64] F. Liese and I. Vajda. On divergences and informations in statistics and information theory. *IEEE Transactions on Information Theory*, 52(10):4394–4412, 2006. ISSN 0018-9448. doi: 10.1109/TIT.2006.881731.
- [65] R. C. P. Marques, F. N. Medeiros, and J. Santos Nobre. SAR image segmentation based on level set approach GA0 model. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 34(10):2046–2057, 2012. ISSN 0162-8828. doi: 10.1109/TPAMI.2011.274.
- [66] M. Mejail, J. C. Jacobo-Berlles, A. C. Frery, and O. H. Bustos. Classification of SAR images using a general and tractable multiplicative model. *International Journal of Remote Sensing*, 24(18):3565–3582, 2003.
- [67] M. E. Mejail. *La Distribución \mathcal{G}_A^0 en el modelado y Análisis de Imágenes SAR*. PhD thesis, Departamento de Computación, Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires, 1999.
- [68] A. Moreira, P. Prats-Iraola, M. Younis, G. Krieger, I. Hajnsek, and K. Papathanassiou. A tutorial on synthetic aperture radar. *IEEE Geoscience and Remote Sensing Magazine*, 1(1):6–43, March 2013. doi: 10.1109/MGRS.2013.2248301.
- [69] A. D. C. Nascimento, R. J. Cintra, and A. C. Frery. Hypothesis testing in speckled data with stochastic distances. *IEEE Transactions on Geoscience and Remote Sensing*, 48(1):373–385, 2010. doi: 10.1109/TGRS.2009.2025498.
- [70] A. D. C. Nascimento, M. M. Horta, A. C. Frery, and R. J. Cintra. Comparing edge detection methods based on stochastic entropies and dis-

- tances for PolSAR imagery. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 7(2):648–663, 2014. URL <http://dx.doi.org/10.1109/JSTARS.2013.2266319>.
- [71] J-M. Nicolas and S. N. Anfinson. Introduction to second kind statistics: Application of log-moments and log-cumulants to the analysis of radar image distributions. *Traitement du Signal*, 19(3):139–167, 2002.
- [72] C. Oliver. Optimum texture estimators for SAR clutter. *Journal of Physics D: Applied Physics*, 26(11):1824, 1993. URL <http://stacks.iop.org/0022-3727/26/i=11/a=002>.
- [73] C. Oliver and S. Quegan. *Understanding Synthetic Aperture Radar Images*. Artech House, Boston, 1998.
- [74] M. G. Palacio, S. B. Ferrero, and A. C. Frery. Revisiting the effect of spatial resolution on information content based on classification results. *International Journal of Remote Sensing*, 40(12):4489–4505, 2019. doi: 10.1080/01431161.2019.1569278. URL <https://doi.org/10.1080/01431161.2019.1569278>.
- [75] W. C. Parr and W. R. Schucany. Minimum distance estimation and components of goodness-of-fit statistics. *Journal of the Royal Statistical Society. Series B (Methodological)*, 44(2):178–189, 1982.
- [76] William C. Parr and William R. Schucany. Minimum distance and robust estimation. *Journal of the American Statistical Association*, 75(371):616–624, 1980. doi: 10.1080/01621459.1980.10477522. URL <http://amstat.tandfonline.com/doi/abs/10.1080/01621459.1980.10477522>.
- [77] E Parzen. On estimation of a probability density function and mode. *Annals of Mathematical Statistics*, 33(3):1065–1076, 1962.

- [78] D. Puig and M.A. Garcia. Pixel classification through divergence-based integration of texture methods with conflict resolution. In *International Conference on Image Processing*, volume 3, pages 1037–1040, 2003. doi: 10.1109/ICIP.2003.1246862.
- [79] S. Quegan, I. Rhodes, and R. Caves. Statistical models for polarimetric-SAR data. In *Geoscience and Remote Sensing Symposium, 1994. IGARSS '94. 'Surface and Atmospheric Remote Sensing: Technologies, Data Analysis and Interpretation'*, *International*, volume 3, pages 1371–1373, 8-12 Aug. 1994. doi: 10.1109/IGARSS.1994.399442.
- [80] E. Quel and C. Rosito. *Introducción al Láser*. Lugar Editorial S. A., Buenos Aires, Argentina, 1994.
- [81] R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2016. URL <https://www.R-project.org/>.
- [82] A. Rényi. On measures of entropy and information. In University of California Press, editor, *Proceedings of the Fourth Berkeley Symposium on Mathematics, Statistics and Probability*, volume 1, pages 547–561, 1961.
- [83] J. Rojo. Heavy-tailed densities. *Wiley Interdisciplinary Reviews Computing Statistics*, 5(10):30–40, 2013.
- [84] M Rosenblatt. Remarks on some nonparametric estimates of a density function. *Annals of Mathematical Statistics*, 27(3):832–837, 1956.
- [85] P. J. Rousseeuw and S. Verboven. Robust estimation in very small samples. *Computational Statistics and Data Analysis*, 40(4):741–758, 2002.
- [86] M. Salicrú, D. Morales, M. L. Menendez, and L. Pardo. On the applications of divergence type measures in testing statistical hypotheses. *Journal of Multivariate Analysis*, 51(2):372–391, 1994.

-
- [87] Sarmap. Synthetic aperture radar and sarscape. the earth observation information gateway, 2009. URL [URL{http://www.sarmap.ch/pdf/SAR-Guidebook.pdf}](http://www.sarmap.ch/pdf/SAR-Guidebook.pdf). [Web; accedido el 21-10-2019].
- [88] O. Scaillet. Density estimation using inverse and reciprocal inverse Gaussian kernels. *Journal of Nonparametric Statistics*, 16(1-2):217–226, 2004.
- [89] H. Scheffe. A useful convergence theorem for probability distributions. *Ann. Math. Statist.*, 18(3):434–438, 09 1947. doi: 10.1214/aoms/1177730390. URL <https://doi.org/10.1214/aoms/1177730390>.
- [90] D.W. Scott. *Multivariate Density Estimation*. John wiley and Sons, New York, first edition, 1992.
- [91] C. E. Shannon. A mathematical theory of communication. *Bell System Technical Journal*, 27:379–423, 1948.
- [92] W. B. Silva, C. C. Freitas, S. J. S. Sant’Anna, and A. C. Frery. Classification of segments in PolSAR imagery by minimum stochastic distances between Wishart distributions. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 6(3):1263–1273, 2013. doi: 10.1109/JSTARS.2013.2248132.
- [93] B. W. Silverman. *Density Estimation for Statistics and Data Analysis*. London: Chapman and Hall, first edition, 1986.
- [94] I. Taneja. On generalized information measures and their applications. In Peter W. Hawkes, editor, *Advances in Electronics and Electron Physics*, volume 76, pages 327 – 413. Academic Press, 1989. doi: [https://doi.org/10.1016/S0065-2539\(08\)60580-6](https://doi.org/10.1016/S0065-2539(08)60580-6).
- [95] I. Taneja. Bounds on triangular discrimination, harmonic mean and symmetric chi-square divergences. *Journal of Concrete and Applicable Mathematics*, 4(1):91–111, 2006.

- [96] C. Tison, J.-M. Nicolas, F. Tupin, and H. Maitre. A new statistical model for Markovian classification of urban areas in high-resolution SAR images. *IEEE Transactions on Geoscience and Remote Sensing*, 42(10):2046–2057, 2004. ISSN 0196-2892.
- [97] K. L. P. Vasconcellos, A. C. Frery, and L. B. Silva. Improving estimation in speckled imagery. *Computational Statistics*, 20(3):503–519, 2005. doi: 10.1007/BF02741311.
- [98] B.C. Vemuri, L. Meizhu, S.-I. Amari, and F. Nielsen. Total Bregman divergence and its applications to DTI analysis. *IEEE Transactions on Medical Imaging*, 30(2):475–483, 2011. ISSN 0278-0062. doi: 10.1109/TMI.2010.2086464.
- [99] J. Wolfowitz. Estimation by the minimum distance method. *Annals of the Institute of Statistical Mathematics*, 5:9–23, 1953. doi: 10.1007/BF02949797.
- [100] J. Wolfowitz. Estimation by the minimum distance method in nonparametric stochastic difference equations. *Ann. Math. Statist.*, 25(2):203–217, 06 1954. doi: 10.1214/aoms/1177728782. URL <https://doi.org/10.1214/aoms/1177728782>.
- [101] J. Wolfowitz. The minimum distance method. *The Annals of Mathematical Statistics*, 28(1):75–88, 03 1957. doi: 10.1214/aoms/1177707038. URL <http://dx.doi.org/10.1214/aoms/1177707038>.
- [102] N. Yahya, NS. Kamel, and AS. Malik. Subspace-based technique for speckle noise reduction in ultrasound images. *Biomedical Engineering Online*, 13(1):6257–6271, 2014. doi: doi:10.1186/1475-925X-13-154.
- [103] C. C. F. Yanasse, A. C. Frery, S. J. S. Sant’Anna, P. F. Hernandez, and L. V. Dutra. Statistical analysis of SAREX data over Tapajós - Brazil. In

M. Wooding and E. Attema, editors, *SAREX-92: South American Radar Experiment*, pages 25–40, Paris, 1993. ESA.