

# Ensayos sobre la Distribución $\mathcal{G}_0$ *Single-Look*: Propiedades, Simulación, Estimación y Reducción de Ruido Speckle

Tesis presentada para optar al título de Doctora en Ingeniería  
mención Imágenes y Señales

**Débora Chan**

Dirección de tesis: Dr Alejandro C. Frery

Codirección de tesis: Dra. María Juliana Gambini

Buenos Aires, 2020



## **Dedicatoria**

*A Maia, Hernán y Gabriela por multiplicar exponencialmente el amor.*

*A mis padres que me dieron la fuerza para creer que todo es posible.*

*A mis maestros, compañeros y amigos por iluminar siempre mi camino.*

*A la UTN-FRBA por ofrecerme la oportunidad para crecer.*

*A la vida por su infinita generosidad. . .*



# Resumen

Las imágenes de Radar de Apertura Sintética (*Synthetic Aperture Radar* – SAR), son fundamentales para la comprensión y estudio del medio ambiente, ya que proveen información que no es posible adquirir por ningún otro instrumento de teledetección.

Sin embargo, estas imágenes se encuentran corrompidas por un ruido muy particular y característico denominado ruido *speckle*. Este ruido dificulta su procesamiento automático así como su análisis e interpretación.

En esta tesis se presenta una nueva estrategia para filtrar el ruido característico de datos provenientes de imágenes SAR basada en diferencias de entropías.

Se inicia el trabajo con una exposición de las características de estas imágenes y describe la distribución  $\mathcal{G}_I^0$  que modela este tipo de datos y ha demostrado ser flexible para diferentes grados de textura del objetivo. Todo el trabajo se focaliza sobre el caso *Single Look* y sus particularidades.

Luego se presenta un estudio comparativo, llevado a cabo en diversas plataformas, de estrategias de generación de datos para experimentos Monte Carlo esenciales para garantizar la calidad de los experimentos de simulación.

Posteriormente se presenta una evaluación de las distintas estrategias de estimación para los parámetros de la distribución  $\mathcal{G}_I^0$  considerando casos puros y casos contaminados.

A continuación se deduce la distribución asintótica de las entropías de Shannon y Rényi para el caso *Single Look* de la distribución  $\mathcal{G}_I^0$  y se presenta una aplicación a la eliminación del ruido *speckle* característico de estas imágenes. La propuesta es la construcción de un filtrado no local basado en un test de diferencia de entropías. Se aplica esta estrategia sobre imágenes sintéticas y reales.

Finalmente, se evalúa la calidad de esta estrategia y se la compara con otros métodos de eliminación del ruido *speckle*.

**Palabras clave:** Radar de Apertura Sintética, Inferencia, Ruido *Speckle*, Entropía.

## Abstract

The Synthetic Aperture Radar – SAR images provide information that cannot be acquired by other remote sensing instruments, for this reason they became a fundamental tool for the understanding and study of the environment characteristics.

Nevertheless, a very particular noise called *speckle* corrupts those images. This noise difficults their automatic processing as well as their analysis and interpretation.

In this thesis we present a new denoising strategy for SAR images data based on differences of entropies. The work begins with an brief exposition of these images characteristics and describes the  $\mathcal{G}_I^T$  family of distributions for intensity data that has been shown to be flexible for different degrees of texture of the target. All the work focuses on the Single Look case and its particularities.

Then, we analyze techniques for data generation, in several computational environments, for Monte Carlo experiments. We present a comparison among strategies for parameter estimation for the single look case of the  $\mathcal{G}_I^0$  distribution for contaminated and non contaminated data.

We obtain the asymptotic distribution of Shannon and Rényi entropies for the Single Look case of the  $\mathcal{G}_I^0$  distribution, and we present an application to denoise these images. The proposal is based on a non-local filter and entropy difference tests. This strategy is applied to simulated and real images.

Finally, we assess the properties of this filter and we compare them with a classical denoising method.

**Key words:** Synthetic Aperture Radar, Inference, *Speckle* Noise, Entropy.



# Índice general

<b>I</b>	<b>Introducción y Presentación del Modelo Distribucional</b>	<b>19</b>
<b>1.</b>	<b>Imágenes SAR</b>	<b>31</b>
1.1.	Aspectos físicos y tecnológicos . . . . .	31
1.2.	Geometría del SAR . . . . .	35
1.3.	Ruido Speckle . . . . .	39
1.4.	Conclusiones . . . . .	41
<b>2.</b>	<b>La distribución <math>\mathcal{G}_I^0</math> <i>Single Look</i></b>	<b>43</b>
2.1.	Modelo Multiplicativo . . . . .	43
2.1.1.	Supuesto Distribucional para el Ruido Speckle . . . . .	44
2.1.2.	Supuesto Distribucional para la Retrodispersión . . . . .	46
2.1.3.	Supuesto Distribucional para el Retorno . . . . .	48
2.1.3.1.	Propiedades de la distribución $\mathcal{G}_I^0$ . . . . .	50
2.1.3.1.1.	Comportamiento de la cola de la distribución . . . . .	50
2.1.3.2.	Caso <i>Single Look</i> . . . . .	52
2.2.	Conclusiones . . . . .	54

<b>II Simulación Estimaciones y Entropías</b>	<b>55</b>
<b>3. Simulaciones</b>	<b>57</b>
3.1. Las rutinas de generación . . . . .	58
3.2. Experimentos . . . . .	61
3.3. Los códigos . . . . .	62
3.4. Costo Computacional . . . . .	65
3.4.1. Momentos . . . . .	66
3.5. Bondad de Ajuste . . . . .	69
3.6. Distribución muestral de los valores extremos . . . . .	70
3.6.1. Distribución del máximo muestral . . . . .	70
3.6.2. Distribución del mínimo muestral . . . . .	72
3.7. Intervalos de Confianza para Cuantiles . . . . .	73
3.8. Discusión y Conclusiones . . . . .	74
<b>4. Estimación de parámetros de la distribución <math>\mathcal{G}_I^0</math>- <i>Single Look</i></b>	<b>83</b>
4.1. Contaminaciones . . . . .	86
4.2. Los Estimadores . . . . .	88
4.2.1. Máxima Verosimilitud (MV) . . . . .	90
4.2.2. Máxima Verosimilitud Penalizado (MVP) . . . . .	91
4.2.3. Momentos (Mom) . . . . .	92
4.2.4. Momentos Ponderados con la Probabilidad (MPP) . . . . .	93
4.2.5. Verosimilitud-Momentos (VM) . . . . .	93
4.2.6. Medianas (Med) . . . . .	94
4.2.6.1. Funciones de Influencia . . . . .	94
4.2.6.2. El estimador . . . . .	95
4.2.7. Mínima Divergencia de Densidad (mDD) . . . . .	97
4.2.8. Máxima Bondad de Ajuste (MBA) con el estadístico de Anderson Darling Cola Derecha (ADD) . . . . .	97

4.3.	El Umbral . . . . .	98
4.4.	Resultados del Estudio de Simulación . . . . .	100
4.4.1.	Datos Puros . . . . .	100
4.4.2.	Datos Contaminados . . . . .	102
4.4.3.	Datos SAR Reales . . . . .	103
4.5.	Discusión y Conclusiones . . . . .	106
<b>5.</b>	<b>Entropía de la distribución <math>\mathcal{G}_I^0</math></b>	<b>119</b>
5.1.	Entropías $h-\phi$ . . . . .	119
5.1.1.	Entropía de Shannon para la distribución $\mathcal{G}_I^0$ Single Look	120
5.1.2.	Entropía de Rényi para la distribución $\mathcal{G}_I^0$ Single Look . .	121
5.2.	Distribución asintótica de las Entropías $h-\phi$ . . . . .	122
5.3.	Test de la distancia cuadrática entre $h-\phi$ entropías . . . . .	125
5.4.	Evaluación de la Prueba . . . . .	126
5.4.1.	El diseño de la comparación . . . . .	126
5.4.2.	La Comparación . . . . .	128
5.5.	Conclusiones . . . . .	131
<b>III</b>	<b>Propuesta para la eliminación del Ruido <i>Speckle</i></b>	<b>135</b>
<b>6.</b>	<b>Filtrado No local basado en Diferencia de Entropías</b>	<b>137</b>
6.1.	Algoritmo propuesto para la reducción del Ruido Speckle . . .	137
6.2.	Resultados . . . . .	138
6.3.	Evaluación del Filtro basado en Entropías . . . . .	143
6.4.	Los recursos disponibles para la evaluación . . . . .	144
6.5.	Calidad del Filtrado Propuesto . . . . .	148
6.5.1.	Preservación de la Media . . . . .	148
6.5.2.	Imagen Sintética . . . . .	149
6.5.3.	Comparación de desempeño sobre Imágenes Sintéticas .	151

6.5.4. Aplicación sobre Imagen Real . . . . .	156
6.5.5. Imagen de Munich . . . . .	156
6.6. Observaciones y Comentarios . . . . .	157
6.7. Conclusiones . . . . .	160

**IV Consideraciones Finales 169**

**7. Discusión y Conclusiones 171**

# Índice de figuras

1.1. Imagen captada por el sensor E-SAR en los alrededores de Munich, Alemania . . . . .	34
1.2. Geometría del Radar . . . . .	36
1.3. Proceso Multilook . . . . .	38
1.4. Suma de retornos individuales de un pixel . . . . .	40
2.1. Gráficas de densidades de probabilidad de la distribución $\Gamma(L, L)$ . . . . .	45
2.2. Densidades de probabilidad de la distribución $\Gamma^{-1}(\alpha, \gamma)$ . . . . .	47
2.3. Densidades de probabilidad de la distribución $\mathcal{G}_I^0$ . . . . .	49
2.4. Diferencia entre Media y Mediana de la distribución $\mathcal{G}_I^0$ . . . . .	53
3.1. Tiempo de procesamiento de cada generador de acuerdo con el lenguaje de programación. . . . .	65
3.2. Errores de la media por Rutina de Generación . . . . .	68
3.3. Errores de la media por lenguaje de programación . . . . .	68
3.4. Errores de la Varianza por Rutina . . . . .	69
3.5. Errores de la Varianza por Lenguaje de Programación . . . . .	69
3.6. Porcentaje de cobertura del $p_{75}$ . . . . .	74
3.7. Porcentaje de cobertura del $p_{90}$ . . . . .	74

4.1. Efecto double bounce . . . . .	86
4.2. Ventana Deslizante sobre dos regiones diferentes . . . . .	87
4.3. Propiedades Estimadores: 'Norm.Asint' denota normalidad asintótica, 'Efic.Norm.Asint' eficiencia y normalidad asintótica, 'Costo Comp.' costo computacional y 'sol expl.' solución explícita. . . . .	99
4.4. Tasa de Convergencia por textura . . . . .	109
4.5. Sesgo por en función del tamaño muestral y la textura . . . . .	110
4.6. ECM por textura y tamaño muestral . . . . .	111
4.7. Tiempo de Procesamiento por Estimador . . . . .	112
4.8. Función Empírica de Influencia Estilizada – FEIS, para $\alpha = -5, \gamma = 100$ . . . . .	113
4.9. Imagen E-SAR de una sola vista utilizada para estimar el parámetro de textura(4.10(a)). Diez regiones centradas en el reflector de esquina de diferente tamaño seleccionadas para las estimaciones(4.10(b)). . . . .	114
4.10. Estimaciones del parámetro de textura usando las muestras demarcadas en la Figura 4.10(b). . . . .	115
4.11. Clasificación de imagen por SVM con y sin corner reflector. . . . .	116
4.12. Precisión de la clasificación para cada método, con y sin corner reflector. . . . .	117
4.13. Segmentación usando la estimación del parámetro $\alpha$ y el método $k$ -means. . . . .	118
5.1. Gráfico de la entropía de Shannon en función de $\alpha$ y $\gamma$ . . . . .	121
5.2. Gráfico de la entropía de Rényi, con $\beta = 0,5$ en función de $\alpha$ y $\gamma$ . . . . .	122
5.3. Nivel de Significación alcanzado por Tamaño y por Prueba . . . . .	128
5.4. Características de los tests en función del Tamaño Muestral . . . . .	130
6.1. La ventana deslizante y los parches deslizantes interiores a la misma . . . . .	140

6.2. Resultado de aplicar el método de filtrado basado en la entropía de Shannon . . . . .	141
6.3. Resultado de la aplicación del filtro basado en la entropía de Shannon. . . . .	142
6.4. Análisis de Preservación de la Media . . . . .	149
6.5. Imagen Sintética con dos bandas verticales . . . . .	149
6.6. Imagen Sintética filtrada con Filtro de Distancia de Entropías de la Imagen 6.5 . . . . .	150
6.7. Imagen Sintética cociente del Filtro de Distancia de Entropías correspondiente a la Imagen Filtrada 6.6 . . . . .	150
6.8. Imagen Sintética filtrada con Filtro de SRAD . . . . .	151
6.9. Imagen Sintética filtrada con Filtro de Lee Mejorado . . . . .	151
6.10. Imagen Sintética filtrada con Filtro FANS . . . . .	152
6.11. Imagen Sintética cociente con Filtro de SRAD . . . . .	152
6.12. Imagen Sintética cociente con Filtro de Lee Mejorado . . . . .	162
6.13. Imagen Sintética cociente con Filtro FANS . . . . .	163
6.14. Imagen Real Single Look de la ciudad de Munich . . . . .	163
6.15. Imagen Real de Munich filtrada con el Filtro Propuesto . . . . .	164
6.16. Imagen Real de Munich Cociente con el Filtro basado en Entropía	164
6.17. Imagen Real de Munich filtrada con el Filtro SRAD . . . . .	165
6.18. Imagen Real de Munich Cociente con el Filtro SRAD . . . . .	165
6.19. Imagen Real de Munich filtrada con el Filtro de Lee Mejorado .	166
6.20. Imagen Real de Munich Cociente con el Filtro de Lee Mejorado	166
6.21. Imagen Real de Munich filtrada con el Filtro FANS . . . . .	167
6.22. Imagen Real de Munich Cociente con el Filtro FANS . . . . .	167



# Índice de tablas

3.1. Clasificación de las Rutinas . . . . .	60
3.2. Resultados de los tests KS y AD . . . . .	71
3.3. $p$ -valores del test KS para máximos muestrales . . . . .	76
3.4. $p$ -valores del test AD para máximos muestrales . . . . .	77
3.5. $p$ -valores del test KS para mínimos muestrales . . . . .	78
3.6. $p$ -valores del test AD para mínimos muestrales . . . . .	79
3.7. Cobertura del intervalo del $q_3$ . . . . .	80
3.8. Cobertura del intervalo del $p_{90}$ . . . . .	81
3.9. Comparaciones post-hoc por lenguaje ( $p_{75}$ ) . . . . .	82
4.1. Lista de Abreviaturas . . . . .	89
4.2. Estimaciones del parámetro de textura para las regiones de la Figura 4.10(b). . . . .	105
5.1. Entropía de Shannon . . . . .	132
5.2. Resultados de cada Test . . . . .	133
5.3. Comparación de Características por tipo de Test y Tamaño mues- tral . . . . .	133
6.1. Estimaciones del ENL para las imágenes de la Figure 6.2. . . . .	141

6.2. Estimaciones del número equivalente de looks para las imágenes de la Figura 6.3. . . . .	142
6.3. Evaluación Filtro Basado en Distancias de Entropías . . . . .	153
6.4. Evaluación Filtro de Lee Mejorado . . . . .	154
6.5. Evaluación Filtro SRAD . . . . .	154
6.6. Evaluación Filtro FANS . . . . .	155
6.7. Comparación Resultados Filtrado Imagen Sintética . . . . .	155
6.8. Evaluación Filtro de entropías para imagen de Munich con ventanas de $15 \times 15$ y $13 \times 13$ . . . . .	157
6.9. Evaluación Filtro SRAD sobre imagen de Munich . . . . .	158
6.10. Evaluación Filtro Lee-M sobre imagen de Munich . . . . .	158
6.11. Evaluación Filtro FANS sobre imagen de Munich . . . . .	159
6.12. Comparación Resultados Filtrado Imagen de Munich . . . . .	159

Parte I

# Introducción y Presentación del Modelo Distribucional



# Introducción

Los sistemas de RaDAR (*Radio Detection And Ranging*: detección y medición de distancias por radio) miden la respuesta del terreno a la radiación electromagnética emitida en forma de pulsos, mediante la detección de la posición y de la distancia de un objeto.

El valor de esta respuesta es almacenado para su posterior procesamiento, y los radares imageadores la utilizan para formar una imagen de la zona de interés. Estas imágenes resultan de gran utilidad ya que permiten obtener información sobre recursos naturales, como también permiten monitorear el medio-ambiente y detectar efectos de la acción del hombre, tales como deforestación, cultivos y embalses.

Diversos factores influyen directamente en el retorno detectado por el sensor como, por ejemplo, la rugosidad y la geometría del terreno, pendientes y orientaciones del ángulo de incidencia, así como la frecuencia y la polarización de operación del sistema.

Un sistema de radar de apertura sintética (designado por su acrónimo inglés *SAR – Synthetic Aperture Radar*) procesa, mediante algoritmos, la información capturada por la antena del radar y logra el mismo rendimiento que se obtendría mediante una antena mucho más grande que la que tiene en realidad, por eso se llaman de apertura sintética [77] y, con ello, se logran resoluciones espaciales mucho más finas que las permitidas por antenas físicamente

viables.

Las interferencias propias de las ondas reflejadas por muchos dispersores elementales durante el proceso de adquisición de imágenes SAR dan origen a un ruido multiplicativo y no gaussiano que las caracteriza y se conoce como ruido *speckle* [83].

El *speckle* es un patrón que, aunque determinístico, se puede modelar y tratar más adecuadamente como estocástico y, por esta razón se lo suele denominar *ruido*. Este ruido, inherente al proceso de captura de la imagen, no es ni gaussiano ni aditivo y, por lo tanto, es diferente al ruido presente en imágenes ópticas.

La presencia de este ruido corrompe la imagen y compromete el contraste entre áreas diferentes, con lo cual dificulta todas las operaciones de procesamiento y análisis de los datos como, por ejemplo, la detección de bordes, la identificación de cambios, la clasificación, y la segmentación.

El modelado estadístico ocupa por lo tanto un rol esencial en el proceso de interpretación y análisis de imágenes SAR, dado que conduce a una comprensión más profunda y guía a los investigadores al filtrado del ruido *speckle*, segmentación, detección y reconocimiento de objetos.

Los procedimientos de modelado de imágenes pueden subdividirse en dos grandes categorías: paramétricos y no paramétricos. Usualmente los parámetros de estas distribuciones son desconocidos y deben estimarse de acuerdo con los datos provenientes de una imagen real.

Los métodos de filtrado de imágenes o detección de bordes utilizan frecuentemente máscaras deslizantes de tamaño  $3 \times 3$ ,  $5 \times 5$ ,  $7 \times 7$ ,  $9 \times 9$  u  $11 \times 11$  que constituyen muestras aleatorias cuyos parámetros es de interés estimar correctamente.

Debido a ello, encontrar buenos estimadores de los parámetros de la distribución subyacente es una necesidad, especialmente para el caso de muestras pequeñas y medianas.

---

Los procedimientos paramétricos comprenden, generalmente la estimación de los parámetros de acuerdo con el modelo seleccionado para describir la variable de interés así como la evaluación del grado de adecuación de la distribución resultante a los datos reales.

Entre los métodos tradicionales más usados para la estimación de los parámetros se deben destacar el método de momentos (MoM) [104], el de cuantiles [109], el de máxima verosimilitud (ML) [30] y el de máxima verosimilitud con corrección de sesgo basado en la propuesta teórica de Cox y Snell [21] y el basado en bootstrap [42].

Sin embargo, propuestas de estimación más recientes consideran el uso de la transformada de Mellin [81] introduciendo los conceptos de funciones características de segundo tipo, los momentos característicos de segundo tipo o log-momentos y los cumulantes característicos de segundo tipo o log-cumulantes [2, 33] definiendo un nuevo enfoque metodológico.

En el objeto de interés de esta tesis, la distribución subyacente de estas imágenes veremos que es un caso especial de la distribución de Pareto Generalizada Tipo II que ha sido caracterizada como una distribución de colas pesadas [95].

Para esta distribución se han propuesto diversos estimadores que utilizan el índice de su cola [1], entre ellas una de las más populares es el estimador de Hill [54] basado en la verosimilitud condicional del parámetro que describe la cola. Ha sido también muy utilizada la estimación de la pendiente del histograma graficado en escala logarítmica, que a pesar de su sencillez, introduce un sesgo sistemático, y el estimador de máxima verosimilitud [80].

En el contexto de las imágenes SAR, la eficiencia de los estimadores debe ser analizada en función del tamaño muestral y de los valores de los parámetros de brillo y de textura así como por su robustez frente a ciertos esquemas de contaminación.

Es importante también destacar que muchos de estos análisis se realizan

sobre imágenes sintéticas, generadas mediante experimentos de Montecarlo. Esto ha motivado el análisis de la calidad de diversas alternativas de generación de estos datos en diferentes plataformas, con el objeto de que la evaluación de los estimadores se realice bajo la garantía de calidad de los datos provistos por el experimento.

Se exponen en la tesis los resultados y conclusiones relativos a la evaluación y comparación de estos algoritmos de generación de datos utilizando criterios de bondad de ajuste, fidelidad del comportamiento de momentos y precisión de la representación de las colas.

Las imágenes SAR pueden ser univariadas o multivariadas. Las imágenes SAR multivariadas se forman con diferentes estados de polarización de las señales electromagnéticas emitidas y recibidas. En este trabajo, trataremos con imágenes de una sola polarización, es decir monopolarizadas, esto es, univariadas.

Las técnicas de supresión de ruido pueden agruparse en dos grandes categorías.

En la primera categoría que se denomina procesamiento multi-look, la mejora se obtiene promediando varias imágenes obtenidas de una porción del ancho de banda espectral de acimut [89]. De esta forma se logra reducir la varianza del ruido. El interés del presente trabajo se centra, sin embargo, en imágenes que no han sido promediadas, conocidas como *Single Look*.

La segunda categoría, donde se ubica esta tesis, agrupa las técnicas de filtrado del ruido *speckle*. El objetivo de un filtro es suprimir el ruido pero conservar los detalles esenciales del escenario, tales como la textura, pequeños objetos y los bordes [45].

Los primeros intentos fueron realizados con filtros heurísticos, como el de la media. Sin embargo éstos no resultaron efectivos dada la naturaleza multiplicativa del ruido que corrompe estas imágenes. Una alternativa popular al filtro de la media, es el filtro de la mediana, debido a su simplicidad y su

---

capacidad para preservar bordes [65].

Los clásicos filtros lineales, como el basado en máscaras de convolución con pesos proporcionales a una función gaussiana, poseen propiedades deseables tales como la invarianza por desplazamiento y la sencillez de aplicación y resultan óptimos en caso de ruidos aditivos y normales. Sin embargo, en imágenes SAR su desempeño es deficiente dada su tendencia al borroneo de bordes, especialmente en píxeles afectados por altos niveles de ruido.

Cabe destacar que estas propuestas locales sólo aprovechan la información disponible en los píxeles vecinos para restaurar la información de un píxel. El ruido *speckle* requiere algoritmos más sofisticados para su corrección.

El mayor inconveniente de este enfoque radica en el fuerte supuesto de que la vecindad del píxel provee información suficiente respecto de los parámetros de la distribución. Luego se propuso [118] aprovechar la información disponible en píxeles más distantes.

Para superar las dificultades que presenta el filtro de medias, se han propuesto filtros adaptativos locales, que utilizan la media y la varianza de las ventanas de escaneo. Este procedimiento se basa en el deslizamiento de dos ventanas, una pequeña que se utiliza para construir la máscara de convolución, y una mayor que se utiliza como marco de referencia para el conjunto de ventanas deslizantes interiores a ella. Estos filtros han sido adaptados para ruidos aditivos, multiplicativos o combinaciones de ambos [64].

Son consideraciones de importancia el tamaño y forma de la ventana deslizante elegidos. Una ventana demasiado grande (en comparación con la escala de interés, producirá un exceso de suavizado y los bordes se verán borrosos. Una ventana demasiado pequeña disminuirá la capacidad de suavizado del filtro y dejará *speckle*. En cuanto a la forma de la ventana, el cuadrado típico producirá redondeo en las esquinas que no estén orientadas a  $90^\circ$ .

El filtro Lee es un filtro adaptativo diseñado especialmente para eliminar el ruido *speckle* preservando bordes y características puntuales de las imágenes

SAR, tales como los reflectores de esquina (*corner reflectors*). Este filtro minimiza el error cuadrático medio, y utiliza el número estimado de looks (ENL) como insumo de entrada.

Una generalización del filtro de estadísticas locales (*Local Statistics – LS*) que propone un algoritmo diferente para calcular la varianza local, fue propuesto [63] logrando una mejora en la preservación de los bordes.

Otra propuesta interesante es el Filtro Sigma para imágenes SAR [66] basado en la distribución de probabilidades gaussiana. Filtra el ruido de la imagen promediando solo esos píxeles dentro del rango de dos sigma del píxel central de cada ventana deslizante. Los píxeles fuera del rango de dos desvíos son ignorados en la construcción de la máscara, pues se consideran atípicos. En consecuencia, se conservan las características de alto contraste.

Una alternativa al filtro LS, es el filtro de Wiener adaptativo de dominio espacial [38] cuya respuesta impulsiva se elige con el criterio del mínimo error cuadrático medio (MMSE). Este error se define como la diferencia entre la salida del filtro y la respuesta deseada que es conceptual. Las propiedades estadísticas de la respuesta deseada así como su relación con la señal de entrada del filtro son asumidas por el diseñador.

Otros filtros se basan en la estimación de los parámetros de la distribución subyacente. Los filtros MAP (máximo a posteriori) han sido utilizados para eliminar el ruido en imágenes SAR de 1 look en [62], modelando la distribución a priori del backscatter con distribución gaussiana y con otras distribuciones a priori como las distribuciones beta y gama [70, 99].

Otros algoritmos adaptativos se fundamentan en un modelo autorregresivo como el filtro de Kalman [6] o el criterio de error cuadrático medio, o la doble ventana medio recortada o el filtro EPOS basado en una ventana adaptativa [101, 67].

Un enfoque diferente es el basado en el espacio de frecuencias, éstos utilizan transformaciones como Fourier, Wavelet y Curvelet [84].

---

Enfoques alternativos basados en características geométricas y morfología matemática también fueron presentados [24, 97].

Como una alternativa robusta se han propuesto filtros basados en estadísticos de orden, denominados filtros L, con funciones de influencia acotadas.

Asimismo se han aplicado inferencias bayesianas de primero y segundo orden [44] y se han combinado con una sigma preselección para evitar los inconvenientes de sesgo que presentan los filtros no locales [123].

Los filtros Stack han mostrado que preservan los bordes y los detalles [27]. Los filtros Stack adaptativos proponen algoritmos mejorados [69] o más rápidos [68].

En [122] con el objetivo de preservar bordes en imágenes SAR fue propuesto un algoritmo que aplica análisis de conglomerados. Cabe destacar que para muchos de ellos la estimación de parámetros de la distribución subyacente es un insumo fundamental.

Las técnicas de eliminación de ruido basadas en enfoques no locales tienen buen rendimiento en general, sin embargo comportan una notable complejidad computacional que los hace poco adecuados para algunas aplicaciones que requieren resultados inmediatos. El algoritmo SAR-BM3D propone un tamaño variable del área de búsqueda impulsado por el nivel de actividad de cada parche y un enfoque probabilístico que aprovecha las características del ruido *speckle* para acelerar la coincidencia de bloques [103].

Recientemente el filtro FANS (Fast Adaptive Nonlocal SAR despeckling) basado en el algoritmo SAR-BM3D combina un excelente rendimiento y una baja complejidad [22].

En años recientes también se aprovecharon las distancias estocásticas para la reducción del ruido y se han desarrollado algoritmos basados en la descomposición de umbrales y el ordenamiento [105].

Recientemente se añadieron propuestas denominadas híbridas que integran diversos filtros, entre ellos cabe mencionar el filtro de la media, moda

y mediana (MMM) [102] y el filtro de difusión anisotrópica [86] que ha demostrado ser superador en la preservación de bordes.

Los filtros de Lee y Kuan tienen supuestos distribucionales diferentes pero la misma base; forman una imagen salida como combinación lineal de la intensidad del pixel central y la intensidad promedio de la ventana, logrando un equilibrio entre el promedio simple (en regiones homogéneas) y el filtro identidad (preservando bordes y características puntuales). Equilibrio que depende del coeficiente de variación de intensidad de la ventana deslizante.

Por su parte el SRAD (*speckle reduction anisotropic diffusion*) resulta una combinación entre el clásico filtro de difusión anisotrópica y el filtro adaptativo de Lee [119]. Mientras los filtros de Lee y de Frost utilizan el coeficiente de variación, el SRAD utiliza el coeficiente de variación instantáneo que resulta ser función del gradiente y el laplaciano locales. El SRAD se deduce de una ecuación diferencial parcial, es adaptativo y no utiliza umbralizaciones. En [119] los autores muestran que el SRAD se vincula directamente con los filtros de Lee (Kuan) y Frost, siendo además una extensión del filtro de difusión anisotrópica original.

Frente a esta diversidad de propuestas se generó la necesidad de definir criterios a fin de evaluar la calidad de un filtro en cuanto a la reducción del ruido *speckle* de la imagen, la preservación de los bordes, líneas delgadas y objetos [113].

Un criterio universal para la conservación de la textura es difícil de definir dado que un criterio (como ECM) favorecerá filtros con alto rendimiento en reducción de ruido *speckle* como el filtro de la media de  $7 \times 7$ , sin embargo estos filtros destruyen la información respecto de la textura.

Entre los criterios de evaluación de la calidad del filtrado podemos mencionar el recíproco del cuadrado del coeficiente de variación [83], la estimación de la media y el desvío del ruido *speckle*, la relación señal ruido, el factor  $\beta$ , definido como el coeficiente de variación de la imagen filtrada [52], la estimación

---

del número equivalente de looks y el error cuadrático medio normalizado [99]. Recientemente se han propuesto coeficientes que no requieren de una imagen de referencia y se basan en las características de la imagen cociente entre la original y la filtrada *speckle* [47].

Deben considerarse, entre los criterios de evaluación, la complejidad computacional y la simplicidad de implementación. Por último la adecuación de un filtro puede depender de la aplicación de interés y eventualmente una combinación sucesiva de dos filtros puede recomendarse.

Durante el desarrollo de este trabajo de tesis se han realizado publicaciones científicas que dan cuenta de los sucesivos avances. Inicialmente se exploró la teoría de valores extremos en el contexto de los estimadores de cola para la distribución  $\mathcal{G}_I^0$  y se la aplicó a la construcción de mapas de textura [14]. Debido a algunas dificultades que estas estimaciones generaban surgió la necesidad de evaluar la calidad de los métodos de generación de números aleatorios con esta distribución y se realizó un análisis minucioso primero sobre el lenguaje  $\mathbb{R}$  [16] y posteriormente se extendió esta evaluación en cuatro entornos computacionales [17]. Los resultados de estos análisis fueron posteriormente aplicados a la estimación del parámetro de textura para la distribución  $\mathcal{G}_I^0$  considerando el caso de datos puros y también el de datos contaminados con reflectores de esquina y presentando aplicaciones a la clasificación y segmentación de imágenes sintéticas y reales [18]. Por último se dedujeron las expresiones de las entropías de Shannon y Rényi para la distribución  $\mathcal{G}_I^0$  así como el estudio de su distribución asintótica y con base en el test de distancia cuadrática entre entropías se propuso un algoritmo de reducción de ruido *speckle* para imágenes SAR probando su desempeño en imágenes reales y sintéticas en base a criterios novedosos y comparando éste con el de filtros reconocidos en la literatura de estas imágenes de teledetección [15].



# Capítulo 1

## Imágenes SAR

### 1.1. Aspectos físicos y tecnológicos

El sensado remoto constituye una fuente de información cuyo desarrollo data de finales del siglo XIX y cuyo auge de desarrollo se inició en el siglo XX. Su evolución ha sido notable y se ha aplicado en diversas áreas de estudio, fundamentalmente en estudios geográficos, dado que provee datos que permiten evaluar de forma eficiente y con bajos costos áreas de interés.

Los sistemas de RaDAR (*Radio Detection And Ranging*: detección y medición de distancias por radio) son instrumentos de sensado remoto que, a través de la emisión de ondas electromagnéticas y la recepción de la energía reflejada, detectan la presencia y la ubicación de objetos.

El radar, ya sea transportado en plataforma aérea o espacial, observa la superficie objetivo lateralmente, por lo cual se los denomina instrumentos de “visión lateral”. De este modo los retornos provenientes de puntos del terreno con diferente distancia en rango llegan en tiempos distintos. De esta forma resulta posible distinguir entre ellos y calcular la distancia del radar al objetivo.

Este instrumento utiliza el efecto Doppler que se produce por el desplazamiento relativo entre el radar y el blanco para lograr la resolución espacial en la dirección del movimiento de la plataforma donde se transporta no dependa

de la distancia a la zona sensada sino solamente de la longitud de su antena.

Cuando la radiación emitida incide sobre la superficie objetivo, se inicia un proceso de dispersión de la radiación en todas direcciones. La antena capta posteriormente una fracción de la energía dispersada que se denomina retrodispersión que resulta distinta según la naturaleza del objetivo.

La emisión de señales de microondas es a intervalos regulares sobre una zona de interés, una parte de esta energía es retrodispersada y detecta la intensidad y la distancia del objetivo a través del retardo en el tiempo del retorno de la señal emitida.

El valor de esta respuesta es almacenado para su posterior procesamiento y, en el caso de radares imageadores, se utiliza para formar una imagen de la zona de interés.

Un sistema de radar de apertura sintética (de ahora en más designado por su acrónimo inglés *SAR – Synthetic Aperture Radar*) procesa, mediante algoritmos, la información capturada por su antena. Además de sistemas SAR satelitales y aerotransportados, hay dispositivos montados sobre rieles [76].

Mediante este procesamiento se logra el mismo rendimiento que se obtendría mediante una antena mucho mayor que la real, de ahí la denominación de apertura sintética [77]. Las resoluciones espaciales obtenidas resultan mucho más finas que las permitidas por antenas físicamente viables.

Las imágenes SAR, tanto monopolarizadas como polarimétricas, constituyen una herramienta de suma importancia en el estudio, la comprensión y el control del medioambiente y la acción del hombre, dado que a partir de ellas se obtienen valiosas informaciones que ningún otro tipo de imágenes provee.

Este tipo de sensado remoto ha sido utilizado en múltiples áreas del conocimiento, tales como planeamiento urbano [32], estimación de la densidad de construcciones [116], monitoreo de cultivos [7], estimación de daños provocados por desastres naturales tales como derrames de petróleo [110], Tsunamis [19, 100], terremotos [26], actividad volcánica, acumulación de nieve y

---

derrumbes [117], o detección temprana de incendios forestales [11].

Los llamados radares monopolares utilizan una única polarización de emisión y se detecta una sola de las componente de la radiación retrodispersada. El término polarización alude a la orientación del campo dieléctrico de la radiación emitida y recibida. De modo tal que es posible tener radares monopolares HH, HV, VH o VV, representando la primera letra la polarización de la radiación emitida y la segunda la componente de polarización detectada.

Se dice que la polarización es horizontal cuando el vector del campo eléctrico está orientado horizontalmente respecto del plano de La Tierra y diremos que es vertical cuando el campo eléctrico tiene orientación perpendicular al plano de la tierra.

Cuando la antena emisora es la misma que la receptora se trata de un radar mono-estático y son los más comunes y las componentes de polarización cruzada  $S_{HV}$  y  $S_{VH}$  son iguales. Cuando por el contrario las antenas emisora y receptora son diferentes se trata de un radar bi-estático.

En Argentina se desarrollan trabajos con diversos objetivos tales como monitoreo de embalses, evaluación de recursos hídricos superficiales, prevención de inundaciones, climatología, topografía del fondo oceánico, detección de humedales, derrames de petróleo y control de deforestación entre otras aplicaciones.

Entre las ventajas de esta modalidad de teledetección, podemos mencionar que los radares poseen un sistema de iluminación propio que permite la adquisición de imágenes independientemente de otras fuentes.

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 información de áreas bajo estudio que no son visibles para el espectro óptico.

A pesar de lograrse imágenes de alta resolución, el modo en que se generan estas imágenes origina problemas tanto para su visualización como para

su análisis, lo que ha motivado el desarrollo de muchos algoritmos para su comprensión e interpretación.

Entre las desventajas podemos mencionar que son difíciles de analizar e interpretar debido a la presencia del *ruido speckle* [5].

La presencia de este ruido corrompe la imagen y compromete el contraste entre áreas diferentes; dificultando todas las operaciones de procesamiento y análisis de los datos como, por ejemplo, la detección de bordes, la identificación de cambios, la clasificación, y la segmentación.

La Figura 1.1 de la página página 34 muestra una imagen SAR en que predominan áreas urbanas (más claras), y de pastos o cultivos (más oscuras).

**Figura 1.1:** Imagen captada por el sensor E-SAR en los alrededores de Munich, Alemania



Con el objetivo de disminuir los efectos de este ruido, se utilizan filtros como así también una técnica llamada procesamiento multilook.

Esta técnica se aplica a menudo durante el proceso de formación de la imagen y consiste en generar, de forma estadísticamente independiente, varias vistas o *looks* a partir del mismo conjunto de pulsos crudos durante el proceso de generación de la imagen. La imagen final se obtiene promediando estas vistas.

---

En el caso de las imágenes *Single Look*, o de una única vista, la relación señal-ruido es baja, lo cual dificulta sensiblemente su análisis.

Tanto el procesamiento multilook como la aplicación de filtros reducen la presencia del ruido *speckle* pero sacrifican resolución espacial; estableciendo de este modo un compromiso entre la calidad visual y la presencia de ruido.

## 1.2. Geometría del SAR

En la Figura 1.2 de la página 36 pueden observarse la geometría del radar, la altura sobre la tierra del desplazamiento  $h$ , la velocidad del desplazamiento  $v$ , el ángulo de incidencia  $\theta$ , la dirección de avance denominada 'azimuth', su perpendicular llamada 'rango' y la perpendicular a la tierra denominada 'nadir', el ancho de la zona iluminada se lo denomina 'swath'. Se designan con  $L_X$  y  $L_Y$  respectivamente a las longitudes de la antena en la dirección de la altura y del azimuth.

La radiación electromagnética emitida se concentra fundamentalmente en el cono central que en un instante dado 'ilumina' al terreno y que se ve representado en una elipse gris. Durante su desplazamiento, el radar ilumina una franja del terreno objetivo paralela a su trayectoria (azimut).

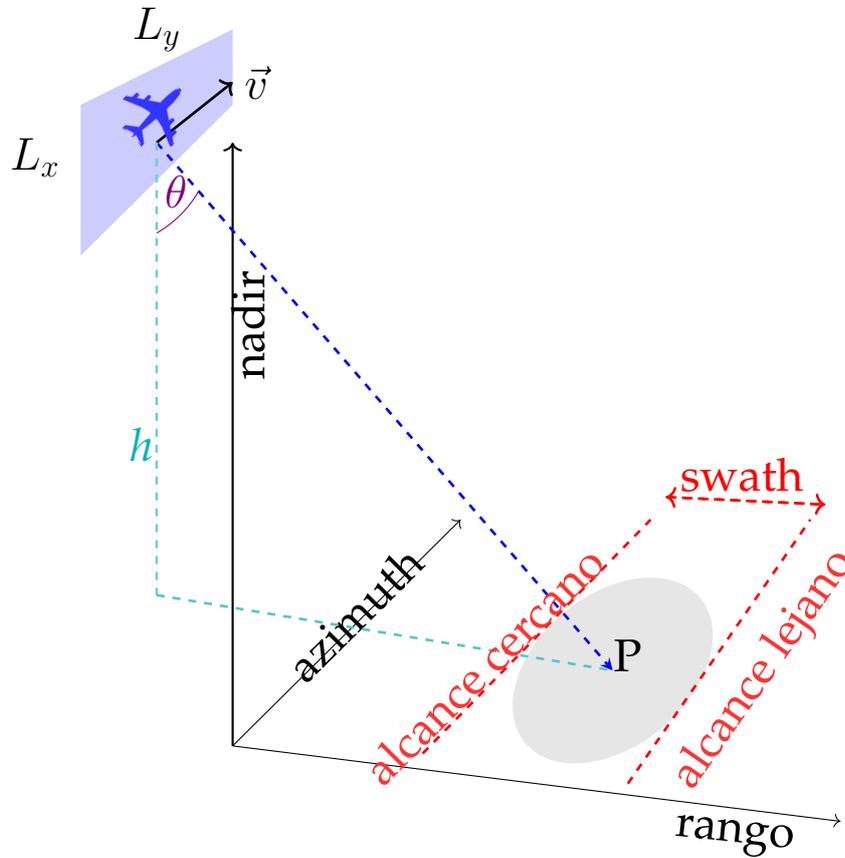
La ecuación del radar [107] es la relación de vínculo fundamental entre las características del sensor, el objetivo y la señal recibida. Su expresión es:

$$P_r = \frac{P_t G^2 \lambda^2 \sigma}{(4\pi)^3 R^4}, \quad (1.1)$$

donde  $P_r$  la potencia retrodispersada,  $P_t$  la potencia emitida por el radar,  $G$  la ganancia de la antena del sensor,  $\lambda$  la longitud de onda,  $R$  la distancia entre el objetivo y el instrumento, y  $\sigma$  la sección eficaz de la retrodispersión.

El factor de retrodispersión  $\sigma$  depende de las condiciones dieléctricas de la superficie objetivo y también de la textura. Es sabido que mayor textura ocasiona mayor intensidad de retorno y mayor ángulo de incidencia  $\theta$  menor

Figura 1.2: Geometría del Radar



pulso de retorno.

El coeficiente de retrodispersión  $\sigma$  se define como el cociente entre la energía total  $E_1$ , es decir si el sensor hubiera retrodispersado la totalidad en forma isotrópica y la energía efectivamente recibida por el sensor  $E_2$ :

$$\sigma = \frac{E_1}{E_2}. \quad (1.2)$$

El coeficiente de retrodispersión, la rugosidad y la geometría del terreno, humedad, pendientes y orientaciones del ángulo de incidencia de la señal, distancia entre sensor y objetivo son factores que influyen directamente en la señal de retorno detectada por el sensor, así como la frecuencia y la polarización de operación del sistema.

---

La radiación electromagnética está conformada por el vector del campo eléctrico  $\vec{E}$  y el vector del campo magnético  $\vec{H}$  que son ortogonales entre sí y respecto de la dirección de propagación. La radiación emitida es linealmente polarizada y el vector campo eléctrico incidente tendrá una dirección horizontal o vertical respecto del plano de la tierra que determina su polarización.

Si el campo eléctrico incidente sobre un objeto en el centro de la zona iluminada por el radar, simbolizado con el punto P de la Figura 1.2 está dado por valores complejos que definen al campo eléctrico en las direcciones horizontal y vertical.

Simbólicamente:

$$\vec{E}^i = \begin{pmatrix} E_v^i \\ E_h^i \end{pmatrix}. \quad (1.3)$$

El campo eléctrico dispersado se vincula de la siguiente forma con el campo eléctrico incidente:

$$\vec{E}^s = E_v^s \check{v} + E_h^s \check{h} = \frac{e^{ik_0 r}}{r} \begin{bmatrix} S_{vv} & S_{vh} \\ S_{hv} & S_{hh} \end{bmatrix} \begin{pmatrix} E_v^i \\ E_h^i \end{pmatrix}, \quad (1.4)$$

donde  $\check{v}$  y  $\check{h}$  son los versores de las direcciones vertical y la horizontal, respectivamente,  $r$  es la distancia entre el radar y el punto de incidencia,  $k_0 = 2\pi/\lambda$  y  $\lambda$  es la longitud de la onda emitida,  $S$  es la matriz de dispersión que caracteriza la respuesta de P en dirección al radar:

$$S = \begin{bmatrix} S_{vv} & S_{vh} \\ S_{hv} & S_{hh} \end{bmatrix}.$$

La retrodispersión es inherente al blanco y es lo que constituye nuestro objetivo de interpretación. En las áreas urbanas la retrodispersión tiene alta variabilidad, es decir que los niveles de gris se dispersan dentro de un amplio intervalo. Se dice que estas zonas son extremadamente heterogéneas o muy rugosas, por el contrario zonas de pasturas tienen una retrodispersión con poca

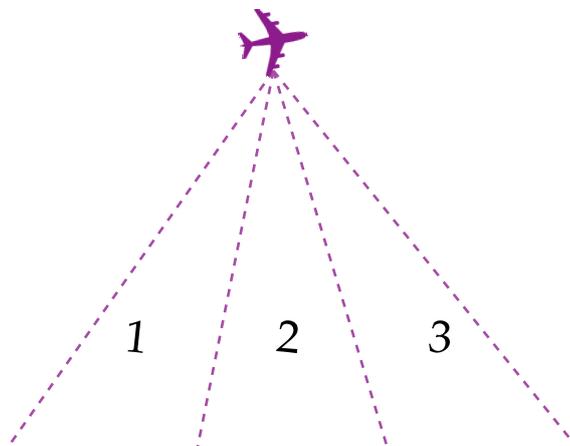
variabilidad es decir niveles de grises concentrados en un intervalo pequeño y se denominan zonas homogéneas o con poca rugosidad.

El demodulado del retorno genera dos magnitudes numéricas que son la parte real y la imaginaria de números complejos que constituyen los datos o pulsos crudos. La señal recibida por el radar es un número complejo, y a partir de esta señal se forma la imagen. Si consideramos el módulo de la señal recibida decimos que los datos están en formato de amplitud. En cambio, si utilizamos el cuadrado del módulo de dicha señal, decimos que los datos están dados en formato de intensidad. Sin pérdida de generalidad en el tratamiento, en este trabajo se considerarán datos dados en este segundo formato.

En caso se obtenerse  $L$  vistas o 'looks' se tendrán  $L$  valores complejos que podemos designar con  $z_{(1)}, z_{(2)}, \dots, z_{(L)}$  cada uno de los cuales corresponde a una vista de la imagen.

Cada vista proviene de dividir el haz del radar en varios subhaces que son más angostos. Cada subhaz proporciona un 'look' o 'vista' independiente al área iluminada. Cada look contiene *speckle*, al promediarlos se reduce el *speckle* de la imagen final. En la Figura 1.3 se han representado 3 haces.

**Figura 1.3:** Proceso Multilook



---

Luego las intensidades observadas podemos designarlas con:

$$z_{(1)}^2, z_{(2)}^2, \dots, z_{(L)}^2.$$

Si se desea tener información en formato intensidad, la intensidad resultante es el promedio de las intensidades de cada una de las vistas, simbólicamente:

$$z_I^{(L)} = \frac{z_{(1)}^2 + z_{(2)}^2 + \dots + z_{(L)}^2}{L}$$

Si por el contrario se desea trabajar con datos en formato amplitud, la conversión es  $z_A^{(L)} = \sqrt{z_I^{(L)}}$ . Destaquemos que el subíndice  $I$  indica intensidad, mientras que el subíndice  $A$  se refiere amplitud.

Tanto en el primer caso como en el segundo, la imagen resultante recibe el nombre de 'multilook' con  $L$  looks.

Si bien teóricamente  $L$  es un valor entero positivo, en las situaciones prácticas, cuando se lo estima puede ser un número real positivo. Existen diversas propuestas para su estimación en caso de ser desconocido [25].

### 1.3. Ruido Speckle

El *speckle* es un patrón que, aunque determinístico, se puede modelar y tratar más eficazmente como estocástico y, por ello, se lo denomina *ruido*. Este ruido no es ni gaussiano ni aditivo y, por lo tanto, es diferente al ruido presente en imágenes ópticas.

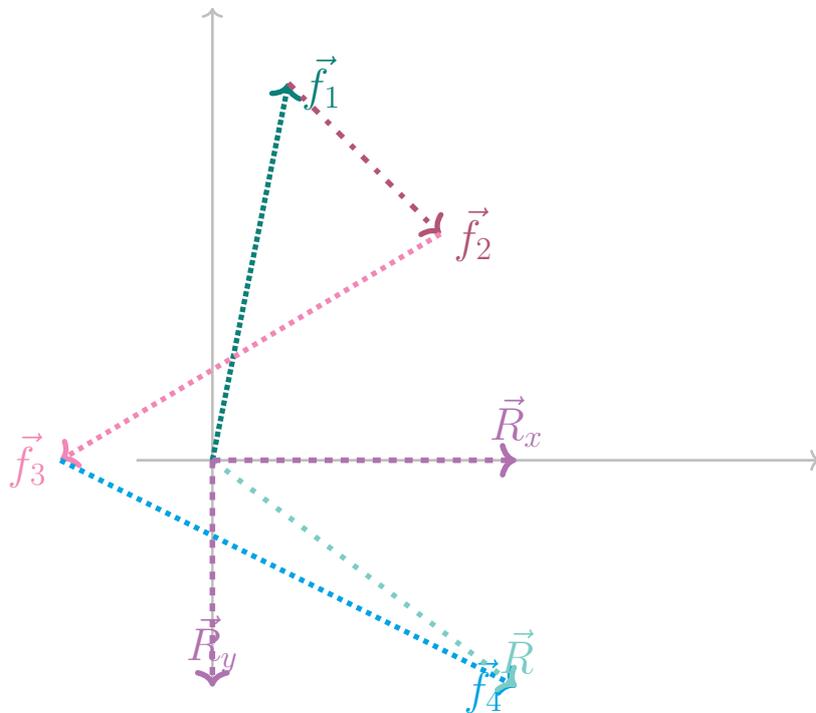
Este ruido es propio del proceso de captura de los sistemas de iluminación coherente tales como imágenes ecográficas, de sonar y de laser. Dentro de cada pixel, existen muchos elementos dispersores; la suma en amplitud y fase de la respuesta de cada uno de ellos determina el retorno registrado en cada pixel.

El ruido *speckle* resulta de la adición coherente de los retornos individuales de los elementos presentes en cada pixel o celda de resolución. En un pixel

donde se encuentra por ejemplo un cultivo, los retornos del cultivo del suelo y algún otro objeto presente como un alambrado se suman con su correspondiente fase y el valor de intensidad de gris del pixel será el cuadrado del módulo de la suma resultante.

En la Figura 1.4 se aprecian los fasores designados con  $f_i$  y su suma designada con  $R$ , así como las componentes de la suma en la dirección de los ejes cartesianos del plano complejo, designadas con  $R_x$  y  $R_y$ .

Figura 1.4: Suma de retornos individuales de un pixel



El *speckle* es siempre un efecto no deseado, debido a que modifica la textura de la imagen generando un aspecto granular aleatorio similar al efecto del ruido sal y pimienta sobre la textura de la imagen. Esto deviene en una percepción visual de incertidumbre radiométrica y reduce la capacidad de distinguir bordes, clasificar, segmentar y detectar objetos en la escena de interés.

---

## 1.4. Conclusiones

En este capítulo se han presentado algunos de los contextos de aplicación de las imágenes SAR que motivan el desarrollo de algoritmos así como su evaluación para mejorar la interpretación de las mismas. Se explica someramente el proceso de adquisición de estas imágenes detallando ventajas y desventajas del mismo puntualizando especialmente en el ruido *speckle*.



## Capítulo 2

# La distribución $\mathcal{G}_I^0$ *Single Look*

### 2.1. Modelo Multiplicativo

Según, entre otras, las referencias [41, 115] el modelo más exitoso propuesto para datos provenientes de un sistema de iluminación por radiación coherente, como son los datos SAR, es el modelo multiplicativo. 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 denotada con  $X$ , que corresponde a lo que se observaría sin la presencia del ruido *speckle*, y la otra correspondiente al ruido *speckle* denotada con  $Y$ .

El tono de gris de cada pixel, es el resultado de la suma de los retornos de cada uno de los elementos con su respectiva fase y trabajaremos con el cuadrado de ese número complejo que es nuestro dato de intensidad  $Z$ .

El modelo multiplicativo considerado es el siguiente:

$$Z = XY, \tag{2.1}$$

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

El análisis de imágenes de esta naturaleza demanda información respecto de un subconjunto de pixeles de la misma. Este análisis se basa generalmente

en propiedades estadísticas de esta muestra de  $n$  celdas de resolución de la imagen, lo cual nos conduce a las propiedades estadísticas de estos valores. La primera propiedad relevante es la distribución a la cual pertenecen estos valores observados.

### 2.1.1. Supuesto Distribucional para el Ruido Speckle

La componente de *speckle* asociada a pixel sigue una distribución  $\Gamma(L, L)$  donde  $L$  es el número de looks de la imagen. Luego la función de densidad de probabilidad del ruido *speckle* de un pixel de una imagen de  $L$  looks es:

$$Y \sim \Gamma(L, L) \iff f(y) = \frac{L^L}{\Gamma(L)} y^{L-1} e^{-yL} \mathbb{1}_{\mathbb{R}^+}(y), \quad (2.2)$$

siendo  $\mathbb{1}_A$  la función indicadora del conjunto  $A$  y  $\Gamma(u) = \int_0^\infty t^{u-1} e^{-t} dt$ . La forma de la función de densidad de probabilidad que caracteriza esta distribución varía notablemente en función del número de looks  $L$ , como puede apreciarse en la Figura 2.1. Nótese que tenemos una distribución exponencial cuando  $L = 1$ .

La expresión de los momentos de esta distribución es:

$$E(Y^r) = \frac{\Gamma(L+r)}{\Gamma(L)}. \quad (2.3)$$

De la ecuación (2.3) se deducen las expresiones de la varianza y de los coeficientes de asimetría de Fisher y de curtosis de la distribución:

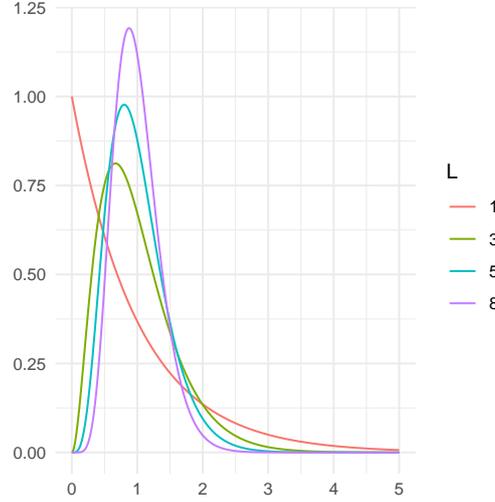
$$v_F(y) = E[(Y - E(Y))^2] = E(Y^2) - E(Y)^2 = L \quad (2.4)$$

$$sk_F(Y) = \frac{E[(Y - E(Y))^3]}{v(Y)^{3/2}} = \frac{2}{\sqrt{L}}, \quad y \quad (2.5)$$

$$Ku(Y) = \frac{E[(Y - E(Y))^4]}{v(Y)} - 3 = \frac{6}{L}, \quad (2.6)$$

De las ecuaciones (2.5) y (2.6) se deduce que ambas características decrecen a medida que el número de looks crece.

**Figura 2.1:** Gráficas de densidades de probabilidad de la distribución  $\Gamma(L, L)$



En esta tesis suponemos conocido el número de looks, más precisamente trabajamos con el caso de una sola vista, denominado *Single Look* (SL), la distribución de  $Y$  resulta:

$$Y \sim \Gamma(1, 1) \iff f(y) = e^{-y} \mathbb{1}_{\mathbb{R}^+}(y) \quad (2.7)$$

donde  $\mathbb{1}_{\mathbb{R}^+}$  es la función indicadora del conjunto de números reales positivos y la expresión 2.7 que resulta la distribución exponencial con parámetro  $\lambda = 1$ .

En este caso:

$$E(Y^r) = \Gamma(1 + r), \quad (2.8)$$

$$sk_F(y) = 2, \quad (2.9)$$

$$Ku(Y) = 6. \quad (2.10)$$

La asimetría y la curtosis están fijas dado el número de looks, que es dato.

### 2.1.2. Supuesto Distribucional para la Retrodispersión

Un modelo general para describir la distribución de la retrodispersión  $X$  es la distribución Gaussiana Inversa Generalizada (GIG), que tiene como casos particulares a la distribución Gamma, la Recíproca Gamma (RG) y la Gaussiana Inversa (GI) [56].

La expresión de la densidad de la GIG está dada por:

$$f_{\text{GIG}}(x) = \frac{(\lambda/\gamma)^{\alpha/2}}{2K_{\alpha}(\sqrt{\lambda\gamma})} x^{\alpha-1} e^{-\frac{1}{2}(\lambda x + \frac{\gamma}{x})} \mathbb{1}_{\mathbb{R}^+}(x), \quad (2.11)$$

cuyo espacio paramétrico está definido de la siguiente manera:

$$\begin{cases} \gamma > 0 \text{ y } \lambda \geq 0 \text{ si } \alpha < 0, \\ \gamma > 0 \text{ y } \lambda > 0 \text{ si } \alpha = 0, \\ \gamma \geq 0 \text{ y } \lambda > 0 \text{ si } \alpha > 0, \end{cases} \quad (2.12)$$

donde  $K_{\alpha}$  es la función de Bessel modificada de tercera especie y orden  $\alpha$  cuya expresión está dada por:

$$K_{\alpha}(\sqrt{\lambda\gamma}) = \left(\frac{\lambda}{\gamma}\right)^{\alpha/2} \frac{1}{2} \int_{\mathbb{R}^+} x^{\alpha-1} e^{-\frac{1}{2}(\lambda x + \frac{\gamma}{x})} dx. \quad (2.13)$$

Es importante considerar las siguientes dos propiedades de la función de Bessel modificada de tercera especie:

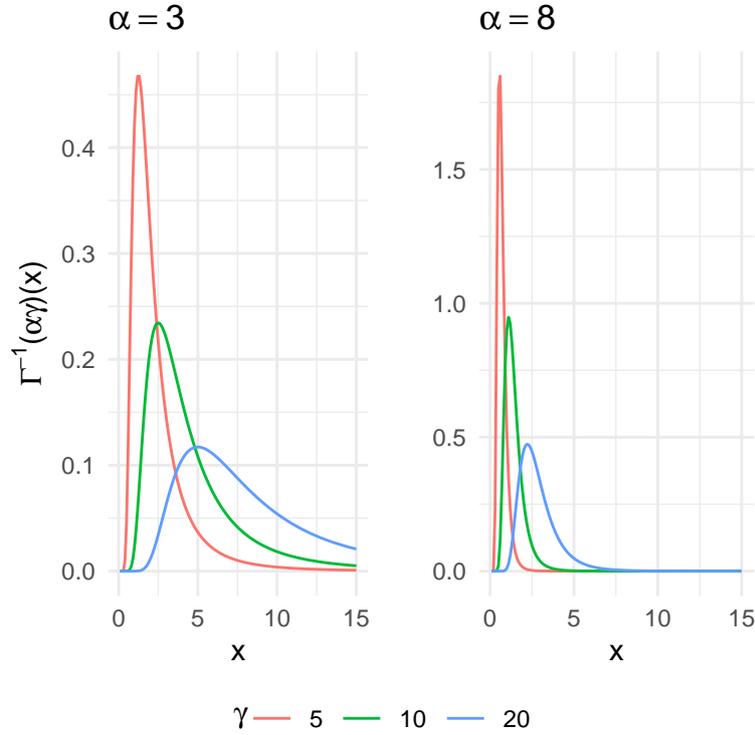
- $K_{-\alpha}(x) = K_{\alpha}(x)$
- $K_{\alpha}(x) \approx 2^{|\alpha|-1} \Gamma(|\alpha|) x^{-|\alpha|}$  para valores de  $x$  cercanos a 0.

Si consideramos el caso en que  $\lambda \rightarrow 0$ , con  $-\alpha, \gamma > 0$ , la densidad de  $X$  se reduce a:

$$f_{\Gamma^{-1}}(x) = \frac{2^{\alpha}}{\gamma^{\alpha} \Gamma(-\alpha)} x^{\alpha-1} e^{-\frac{\gamma}{2x}} \mathbb{1}_{\mathbb{R}^+}(x), \quad (2.14)$$

que corresponde a una distribución Recíproca Gamma (RG), denotada por  $\Gamma^{-1}(\alpha, \gamma)$ . La Figura 2.2 muestra gráficas de densidades que caracterizan esta distribución, donde se aprecian las variaciones para distintos valores de los parámetros  $\alpha$  y  $\gamma$ .

**Figura 2.2:** Densidades de probabilidad de la distribución  $\Gamma^{-1}(\alpha, \gamma)$



Para esta distribución la expresión de los momentos de orden  $r$  es:

$$E(X^r) = \left(\frac{\gamma}{2}\right)^r \frac{\Gamma(-\alpha - r)}{\Gamma(-\alpha)}, \quad (2.15)$$

con  $r > -\alpha$ .

A partir de la ecuación (2.15) las expresiones de la asimetría de Fisher y la curtosis de la distribución resultan:

$$As_F(Z) = \frac{4\sqrt{-\alpha - 2}}{-\alpha + 1}, \quad (2.16)$$

donde  $-\alpha > 3$ , y

$$Kw_F(Z) = \frac{-3(\alpha + 22)}{(\alpha + 3)(\alpha + 4)}, \quad (2.17)$$

donde  $-\alpha > 4$ .

La asimetría y la curtosis decrecen a medida que  $\alpha$  crece en valor absoluto.

### 2.1.3. Supuesto Distribucional para el Retorno

Supuestas las distribuciones del *speckle* y de la retrodispersión o *backscatter*, la distribución de la variable aleatoria que describe el retorno, es decir los datos observados, queda completamente determinada.

En el caso que  $Y \sim \Gamma^{-1}(\alpha, \gamma)$ , el retorno  $Z$  del modelo multiplicativo obedece una distribución  $\mathcal{G}_I^0$  [36].

Gao [41] revisa diversos modelos propuestos para datos de imágenes de radar de apertura sintética y señala a la distribución  $\mathcal{G}^0$  un gran avance dada su tratabilidad y expresividad:

*After 1997, papers on SAR statistical modeling have appeared in renowned journals almost every year. The most attractive achievement among them is the statistical modeling on extremely heterogeneous region of SAR images proposed by Frery, who works in Brazil and has introduced the original idea that for the purpose of statistical modeling, SAR images can be divided into homogeneous regions, heterogeneous regions and extremely heterogeneous regions, according to their contents.*

Su función de densidad está dada por:

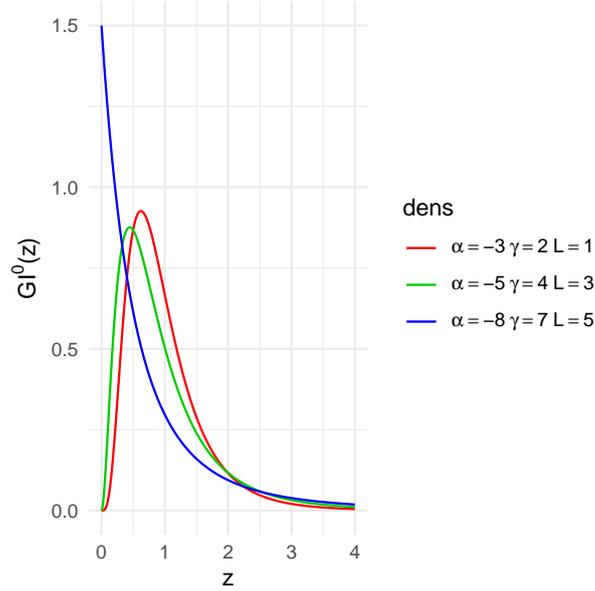
$$f_{\mathcal{G}_I^0}(z) = \frac{L^L \Gamma(L - \alpha)}{\gamma^\alpha \Gamma(-\alpha) \Gamma(L)} \cdot \frac{z^{L-1}}{(\gamma + zL)^{L-\alpha}} \mathbb{1}_{\mathbb{R}^+}(z), \quad (2.18)$$

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

Es importante señalar que los dos primeros parámetros de la distribución  $\alpha$  y  $\gamma$  varían de una a otra región de la imagen, mientras que el tercer parámetro,  $L$ , puede considerarse el mismo para toda la imagen.

El parámetro  $\alpha$  de la distribución  $\mathcal{G}_I^0$  caracteriza el grado de textura del blanco, y permite describir regiones con muy variados niveles de textura [108]. Para valores de  $\alpha$  cercanos a cero (típicamente en el intervalo  $(-3, 0)$ ), la zona de la imagen corresponde a una región extremadamente texturada, en la cual

**Figura 2.3:** Densidades de probabilidad de la distribución  $\mathcal{G}_I^0$



los niveles de gris de la imagen están dispersos en un intervalo grande de valores, como es el caso de las zonas urbanas. A medida que el valor del parámetro  $\alpha$  disminuye, corresponde a zonas con menos textura, los niveles de gris son parecidos entre sí, como son las regiones de forestación (usualmente  $(-6, -3]$ ) y pastura (típicamente en  $(-\infty, -6)$ ). Por otro lado, el parámetro  $\gamma$ , denominado parámetro de escala, posee una interpretación en términos del brillo. Cuanto mayor es su valor, mayor nivel de intensidad posee la imagen en esa región. Por último el parámetro  $L$ , que puede suponerse conocido, está vinculado a la relación señal-ruido. La Figura 2.3 muestra la forma de la distribución  $\mathcal{G}_I^0$  para distintos valores de los parámetros.

La ecuación de los momentos de orden  $r$  es:

$$E(Z^r) = \left(\frac{\gamma}{L}\right)^r \frac{\Gamma(-\alpha - r)}{\Gamma(-\alpha)} \cdot \frac{\Gamma(L + r)}{\Gamma(L)}. \quad (2.19)$$

Estos momentos son finitos cuando  $-\alpha > r$ .

### 2.1.3.1. Propiedades de la distribución $\mathcal{G}_I^0$

En esta sección presentaremos y demostraremos un conjunto características especiales y propiedades importantes de la distribución  $\mathcal{G}_I^0$  y particularizaremos algunas de ellas para el caso *single look*.

**2.1.3.1.1. Comportamiento de la cola de la distribución** En [95] se realiza una descripción minuciosa del comportamiento de las colas de las distribuciones. Utilizaremos algunos de estos conceptos para caracterizar las colas de la distribución  $\mathcal{G}_I^0$ .

**Definición 2.1.** Diremos que una familia de distribuciones  $\mathcal{F}$  es de variación lenta en el infinito cuando se verifica que para cualquier función  $\ell: (0, \infty) \rightarrow \mathbb{R}$  de esta familia y  $\forall t > 0$

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

**Definición 2.2.** Diremos que una familia de distribuciones  $\mathcal{F}$  es de colas pesadas, cuando para cualquier función  $f$  de la familia y para algún valor real  $\beta > 0$  se cumple que:

$$f(x) = \ell(x)x^{-\beta}.$$

Siendo  $\ell(x)$  una función de variación lenta en el infinito,  $\beta$  recibe el nombre de índice de la cola.

Otro enfoque posibles para caracterizar las colas es considerar una distribución de referencia. En esta línea se han propuesto como referencias la curva normal, tan difundida en el ámbito de la estimación estadística, y la curva exponencial.

Utilizando a la curva gaussiana como distribución de referencia diremos que

**Definición 2.3.** Diremos que una familia de distribuciones  $\mathcal{F}$  tiene colas pesa-

---

das cuando para cualquier función  $f$  de la familia se cumple que:

$$\lim_{x \rightarrow \infty} \frac{f(x)}{\phi(x)} = 0.$$

siendo  $\phi$  la curva de densidad correspondiente a la distribución normal estándar.

Si, en cambio, utilizamos la curva exponencial como distribución de referencia diremos que  $f$  tiene colas pesadas cuando para cualquier valor de  $\lambda > 0$  se cumple que:

$$\lim_{x \rightarrow \infty} \frac{e^{-\lambda x}}{f(x)} = 0.$$

Se ha demostrado que [39] la distribución  $\mathcal{G}_I^0$  tiene colas pesadas. Utilizando la distribución gaussiana como referencia:

$$\lim_{z \rightarrow \infty} \frac{f_{\mathcal{G}_I^0}(z)}{\phi(z)} = \lim_{z \rightarrow \infty} \frac{L^L \Gamma(L - \alpha) \sqrt{2\pi} \sigma z^{L-1} e^{-\frac{1}{2} z^2}}{\gamma^\alpha \Gamma(-\alpha) \Gamma(L) (\gamma + zL)^{L-\alpha}} = 0.$$

También puede probarse utilizando a la distribución exponencial de referencia:

$$\lim_{z \rightarrow \infty} \frac{e^{-\lambda z}}{f_{\mathcal{G}_I^0}(z)} = \lim_{z \rightarrow \infty} \frac{\gamma^\alpha \Gamma(-\alpha) \Gamma(L) (\gamma + zL)^{L-\alpha} e^{-\lambda z}}{L^L \Gamma(L - \alpha) z^{L-1}} = 0.$$

Dado que su comportamiento asintótico es comparable al de  $f(z) = z^{\alpha-1}$ , si consideramos  $\ell(z) = f(z)z^{-\alpha+1}$  se deduce sencillamente que el índice de cola de esta distribución es  $-\alpha + 1$ .

**Definición 2.4.** Diremos que una familia de distribuciones  $\mathcal{F}$  es absolutamente propensa a outliers, si dada una función  $f$  perteneciente a dicha familiar y muestra ordenada de esta distribución  $Z_{1:n} \leq Z_{2:n} \leq \dots \leq Z_{n:n}$  existen dos constantes positivas  $\epsilon$  y  $\delta$  y un entero  $n_0$  tales que:

$$P(Z_{m:n} - Z_{m-1:n} > \epsilon) \geq \delta \quad \forall m \geq n_0.$$

Una condición suficiente para que una distribución sea absolutamente propensa a outliers es que existan constantes  $\epsilon$  y  $\delta$  tales que para algún  $z_0$  se cumple que:

$$\frac{f(z + \epsilon)}{f(z)} \geq \delta, \quad \text{para cualquier } z \geq z_0.$$

### 2.1.3.2. Caso *Single Look*

Existen muchas aplicaciones como detección de estructuras urbanas, detección de mapas de carreteras, o detección de barcos, entre otras, para las cuales las imágenes de radar de apertura sintética (SAR) de alta resolución (HR) son requeridas. Para estas aplicaciones, las imágenes *Single Look* con la máxima resolución espacial posible resultan esenciales y las imágenes *multilook* no son una buena opción. En [58] los autores estudiaron la distribución polarimétrica  $\mathcal{G}$  que cuando se restringe su dominio paramétrico, se reduce a  $\mathcal{G}_p^0$  o a  $\mathcal{K}_p$

Es nuestro interés en este trabajo el caso monopolarizado más ruidoso correspondiente a una sola vista, es decir  $L = 1$ . Este caso es conocido como *single look* y la expresión de su función de densidad de probabilidad (2.20) es

$$f_Z(z) = \frac{-\alpha}{\gamma} \left( \frac{z}{\gamma} + 1 \right)^{\alpha-1} \mathbb{1}_{\mathbb{R}^+}(z). \quad (2.20)$$

y de su función de distribución acumulada (2.21) es

$$F_Z(z) = 1 - \left( \frac{z}{\gamma} + 1 \right)^{\alpha}, \quad \text{para } z \geq 0. \quad (2.21)$$

En el caso *Single Look*, la distribución  $\mathcal{G}_I^0$  adopta la forma de una distribución de Pareto generalizada de tipo II [14]

$$f_{GP_{II}}(z) = \frac{\beta}{\sigma} \left( 1 + \frac{z - \mu}{\sigma} \right)^{-\beta-1} \mathbb{1}_{\mathbb{R}^+}(z),$$

con  $\mu = 0$ ,  $\sigma = \gamma$  y  $\beta = -\alpha$ .

La ecuación de los momentos de orden  $r$  es:

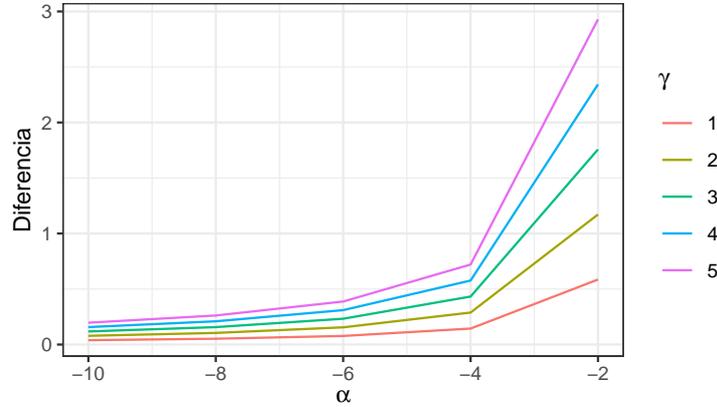
$$E(Z^r) = \gamma^r \frac{\Gamma(-\alpha - r)\Gamma(1 + r)}{\Gamma(-\alpha)}, \quad (2.22)$$

que quedan definidos y son finitos cuando  $-\alpha > r$

Para esta distribución, la expresión de su valor esperado es:

$$E(Z) = -\frac{\gamma}{\alpha + 1}. \quad (2.23)$$

**Figura 2.4:** Diferencia entre Media y Mediana de la distribución  $\mathcal{G}_I^0$



La expresión de la mediana de la distribución, que se deduce de la función de distribución (2.21) es  $\tilde{Z} = \gamma(2^{-1/\alpha} - 1)$ . Si bien la diferencia entre la media y la mediana, que es una medida de la asimetría de la distribución, es en todos los casos positiva, se aprecia que aumenta cuando  $\gamma$  aumenta mientras se mantiene fijo el valor de  $\alpha$  y disminuye cuando el valor de  $\alpha$  decrece y se mantiene fijo el valor de  $\gamma$ , como puede apreciarse en la Figura 2.4.

De la ecuación (2.22) se deduce la expresión de la varianza de la distribución:

$$V_F(Z) = \frac{\gamma^2 \alpha}{(\alpha + 1)^2 (\alpha + 2)} \quad (2.24)$$

que está definida cuando  $-\alpha > 2$ . De la ecuación (2.22) se deduce la expresión del coeficiente de asimetría de Fisher, que es otra manera de cuantificar esta característica de la distribución:

$$As_F(Z) = \frac{\alpha^{1/2} (2\alpha^2 + 22\alpha + 24)}{\alpha^{3/2}}. \quad (2.25)$$

En esta última expresión se aprecia nuevamente que la asimetría decrece cuando el valor absoluto del parámetro de textura crece.

El coeficiente de curtosis es:

$$Ku(Z) = \frac{3(\alpha + 2)}{\alpha^2} (\alpha^3 - 17\alpha^2 - 36\alpha - 24). \quad (2.26)$$

**Definición 2.5.** Cuando la probabilidad de ocurrencia de un valor particular varía en forma inversamente proporcional a la potencia de ese valor, se dice que la distribución sigue una ley de potencia o ley de Zipf.

La forma de la densidad de estas distribuciones es  $P(x) \propto Cx^{-\beta}$ .

Claramente, por lo expuesto, la distribución  $\mathcal{G}_I^0(\alpha, \gamma, 1)$  es una ley de potencia. Una estrategia común para confirmar que un conjunto de datos proviene de una distribución o ley de potencia es representar un histograma en escalas logarítmicas y verificar que tiene estructura lineal.

Para el caso *Single Look* la expresión del logaritmo de la función de densidad presentada en la Ecuación (2.20) es

$$\log f_Z(z) = \log(-\alpha) - \alpha \log \gamma + (\alpha - 1) \log(z + \gamma), \quad (2.27)$$

que resulta una función lineal en el argumento  $\log(z + \gamma)$ .

## 2.2. Conclusiones

En este capítulo se explica el modelo multiplicativo y en este contexto se presentan las distribuciones estadísticas utilizadas para modelar el retorno, el ruido *speckle* y la retrodispersión. Con el objetivo de caracterizar la distribución  $\mathcal{G}_I^0$ , se definen y demuestran algunas propiedades que la distinguen y se brinda un análisis minucioso de éstas para el caso *Single Look*.

El reciente trabajo de Yue et al. [120] relaciona las distribuciones marginales de los datos SAR con la distribución y la estructura de correlación de los retrodispersores elementales.

ParteII

# Simulación Estimaciones y Entropías



## Capítulo 3

# Simulaciones

Para evaluar el rendimiento de algoritmos para la interpretación y el análisis automáticos de imágenes es importante contar con datos que reproduzcan fielmente la información de imágenes SAR y usualmente se realizan simulaciones de Montecarlo que permiten realizar inferencias en una amplia variedad de situaciones.

Los datos simulados son el insumo básico para el diseño de imágenes sintéticas que permiten comparar el desempeño de distintas herramientas para el procesamiento de imágenes SAR [16].

Gambini et al. [40] examinaron la precisión de los procedimientos de detección de bordes usando datos simulados. Otros autores han utilizado datos SAR simulados para evaluar técnicas de detección de bordes por medio de test no paramétricos [43], distancias estocásticas [79] y distancias geodésicas [78]. También han sido esenciales para evaluar el desempeño de filtros [22, 105].

En este capítulo se relata un estudio realizado con el fin de evaluar y comparar cuatro maneras alternativas de generar datos con distribución  $\mathcal{G}_I^0$  [17] en una diversidad de entornos de uso frecuente.

En esta comparación se utilizaron diversos criterios para estimar la calidad de los datos generados en términos de la fidelidad de los mismos con la distribución así como el costo computacional de cada una de estas alternativas. El experimento fue realizado sobre cuatro lenguajes de programación: R, Matlab,

Julia y Ox.

La experiencia señaló que si bien no hay una recomendación válida sobre los cuatro entornos, sí se pueden realizar consideraciones específicas para cada una de ellas.

Se han desarrollado algoritmos de simulación de imágenes SAR que aprovechan propiedades físicas de los datos SAR [59], algunos basados en modelos fractales [49] y otros en la teoría de difracción geométrica [31]. El esquema utilizado en este trabajo involucra la generación de números pseudoaleatorios.

Devroye [28] discute el desempeño de la generación de números pseudoaleatorios para distintas familias de probabilidad en términos de algoritmos, sencillez del código, portabilidad y tiempos de ejecución. En este trabajo se supone que las rutinas subyacentes para el muestreo están libres de errores, y que se dispone de un generador de números pseudoaleatorio uniforme perfecto.

La evaluación de los datos con distribución  $\mathcal{G}_l^0$  desde esta perspectiva y con los criterios que estableceremos es original y constituye uno de los aportes de este trabajo.

### 3.1. Las rutinas de generación

Utilizando equivalencias distribucionales presentamos cuatro rutinas para la generación de datos con distribución  $\mathcal{G}_l^0$ , las primeras dos basadas en el modelo multiplicativo, la tercera en el método de inversión y la última, específica para el caso Single Look, en la equivalencia con la Pareto Generalizada Tipo II, expuesta en el capítulo anterior.

Evaluaremos estas alternativas tomando en cuenta la distribución del máximo, la bondad de ajuste, la precisión de los momentos y cuantiles. Esta evaluación se realiza usando cuatro lenguajes de programación (Matlab, R, Ox y Julia), destacando similitudes y diferencias entre ellas.

Mejail et al. [75] demostraron que la familia  $\mathcal{G}^0$  es una reparametrización de

---

la distribución de Fisher-Snedekor  $\mathcal{F}$ . La relación entre estas dos distribuciones puede expresarse de la siguiente manera:

$$F_{\mathcal{G}_I^0(\alpha, \gamma, L)}(x) = \Upsilon_{2L, -2\alpha}(-\alpha x / \gamma), \quad (3.1)$$

donde  $\Upsilon_{s,t}$  es la función de distribución de una variable de Fisher-Snedekor  $\mathcal{F}_{s,t}$  con  $s$  y  $t$  grados de libertad. Esta relación puede utilizarse para calcular los cuantiles de la distribución  $\mathcal{G}_I^0(\alpha, \gamma, L)$ .

La variable aleatoria  $Z \sim \mathcal{G}_I^0(\alpha, \gamma, L)$  puede definirse como el producto de dos variables aleatorias independientes. Las siguientes relaciones distribucionales fundamentan las cuatro alternativas utilizadas para la generación de estos datos.

- Si  $X \sim \Gamma(\alpha, \gamma)$  entonces  $\frac{1}{X} \sim \Gamma^{-1}(-\alpha, \gamma^{-1})$ , donde  $\Gamma^{-1}$  denota la distribución Gama Inversa.
- La distribución  $\chi^2$  es un caso particular de la distribución  $\Gamma$ , es decir que si  $X \sim \chi_n^2$  entonces  $X \sim \Gamma(n/2, 2)$ .
- Además si  $X \sim \Gamma(m, n)$  entonces  $aX \sim \Gamma(m, n/a)$  para  $a > 0$ .

Basados en estas relaciones, se proponen las siguientes rutinas para la generación aleatoria de datos con distribución  $\mathcal{G}_I^0(\alpha, \gamma, 1)$ :

**1. Rutina  $\Gamma$  :**

$$Z = Y \frac{1}{X}$$

siendo  $X$  e  $Y$  variables aleatorias independientes, tales que  $X \sim \Gamma(-\alpha, \gamma)$  o bien  $\frac{1}{X} \sim \Gamma^{-1}(-\alpha, \gamma^{-1})$  e  $Y \sim \Gamma(1, 1)$ .

**2. Rutina  $\chi^2$**

$$Z = \gamma X \frac{1}{Y},$$

siendo  $X$  e  $Y$  variables aleatorias independientes tales que  $X \sim \chi_2^2$  e  $Y \sim \chi_{-2\alpha}^2$ .

La construcción de la tercera rutina se basa en el hecho de que la variable aleatoria Fisher-Snedecor  $\mathcal{F}$  se define como el cociente de dos variables aleatorias independientes con distribuciones  $\chi^2$  cada una de las cuales dividida por sus grados de libertad:

**3. Rutina  $F$**

$$Z = -\frac{\gamma}{\alpha}F,$$

siendo  $F$  una variable aleatoria con distribución  $\mathcal{F}_{2,-2\alpha}$ .

La cuarta rutina se basa en la relación de la distribución  $\mathcal{G}_I^0(\alpha, \gamma, 1)$  con la distribución de Pareto Generalizada de Tipo II.

**4. Rutina  $P$**

$$Z \sim P_{(II)}(0, \gamma, -\alpha),$$

donde  $P_{(II)}$  alude a la distribución de Pareto Generalizada de Tipo II.

En todos los casos,  $-\alpha, \gamma > 0$ . La Tabla 3.1 muestra la clasificación de los métodos de generación de los números aleatorios.

**Tabla 3.1:** Clasificación de las Rutinas

Modelo Multiplicativo	Método de Inversión
$\Gamma$	$F$
$\chi^2$	$P$

La tercera y cuarta rutinas  $F$  y  $P$  pueden expresarse en función de una única variable aleatoria, es decir que requieren de un solo número aleatorio mientras que las dos primeras  $\Gamma$  y  $\chi^2$  requieren el llamado de dos variables aleatorias.

---

## 3.2. Experimentos

Dado que el parámetro  $\alpha$  de la distribución  $\mathcal{G}_I^0$  puede ser interpretado como la textura de la región correspondiente de la imagen, como se detalló en la Sección 2.1.3, y dado que estamos interesados en estudiar todo tipo de áreas, consideraremos valores representativos de cada textura y diferentes niveles de brillo, definido por los valores de  $\gamma$  para cada experimento.

Los lenguajes de programación utilizados para cada uno de los experimentos, que son utilizados en la mayoría de las aplicaciones de sensado remoto y procesamiento de imágenes:

- **R** [92], versión 3.3.1, es un entorno de software (*open source*) para computación estadística y procesamiento de imágenes, entre otras aplicaciones. Es posible ejecutar rutinas de este entorno sobre varias plataformas y provee herramientas estadísticas y gráficas.
- **MatLab** [73], version R2017a, es un entorno computacional numérico. Tiene su propio lenguaje de programación y una amplia librería de herramientas de procesamiento de imágenes. Este entorno permite una variedad de operaciones matriciales y la creación de interfaces de usuario.
- **Ox** [29], version 7.00, es un lenguaje de programación orientado a objetos con una extensa librería de herramientas estadísticas que también está disponible en varios entornos de programación.
- **Julia** [8], versión 0.5.0, es un lenguaje de programación dinámico de alto rendimiento para computación científica. Proporciona un compilador de ejecución paralela distribuida con buena precisión numérica y librerías avanzadas con funciones matemáticas.

Todos estos entornos de programación, excepto el segundo, son abiertos.

### 3.3. Los códigos

En esta sección incluimos los códigos utilizados para las rutinas en cada uno de los lenguajes de programación. El código 3.1 corresponde a R, el 3.2 a Matlab, el 3.3 a Ox, y el 3.4 a Julia. Denotamos con  $\alpha$  y  $\gamma$  a los parámetros de textura y de escala respectivamente, y con  $n$  al tamaño muestral, mientras  $gen$  indica la rutina de generación utilizada.

Siguiendo un criterio de generalidad, siempre que sea posible dejamos libre el parámetro  $L$  para el número de looks.

**Rutina 3.1:** Generación de muestras  $\mathcal{G}_I^0$  en R

```
library(stats4)
library(rmutil)
rGI0= function(alpha ,gam,n,L = 1,gen){
  switch(gen,
    "Pareto" = rpareto(n,-alpha ,gam) ,
    "Gamma" = rgamma(n,L,L)/rgamma(n,-alpha ,gam) ,
    "Chi2" = gam*rchisq(n,2)/rchisq(n,-2*alpha) ,
    "F-Snedecor" = (-1)*gam*rf(n,2,-2*alpha)/ alpha)
}
```

**Rutina 3.2:** Generación de muestras  $\mathcal{G}_I^0$  en Matlab

```
function data = rGI0(alpha ,gam,n,gen , varargin)
  L = 1;
  fprintf('Total_number_of_inputs_=_%d\n' ,nargin);
  if nargin == 4
    gen = 'Pareto';
  end
  switch gen
    case 'Gamma'
      data=gamrnd(L,1/L,n,1) ./gamrnd(-alpha ,1/gam,n,1);
    case 'Chi2'
      data=gam*chi2rnd(2 ,n,1) ./chi2rnd(-2*alpha ,n,1);
```

```

    case 'F-Snedecor'
        data=(-1)*gam*frnd(2,-2*alpha,n,1)./alpha;
    case 'Pareto'
        data=gprnd(-1/alpha,-gam/alpha,0,n,1);
end

```

### Rutina 3.3: Generación de muestras $\mathcal{G}_I^0$ en Ox

```

#include <oxstd.h>
#include <oxprob.h>

rGI0(const alpha, const gam, const n, const gen)
{ decl data, i, XG, YG, XC, YC, XF, XU, DG, L = 1;
  switch (gen)
  {
    case "Gamma":
        YG = rangamma(n,1,L,L);
        XG = rangamma(n,1,-alpha,gam);
        data = YG./XG;
        break;
    case "Chi2":
        XC=ranchi(n,1,2);
        YC=ranchi(n,1,-2*alpha);
        data = gam * XC./YC;
        break;
    case "F-Snedecor":
        XF=ranf(n,1,2,-2*alpha);
        data = -(gam/alpha) * XF;
        break;
    case "Pareto":
        XU=ranu(n,1);
        data=gam*((1-XU).^(1/alpha)-1);
        break;
  }
  return data; }

```

**Rutina 3.4:** Generación de muestras  $\mathcal{G}_I^0$  en Julia

```

Pkg.add("Distributions")
Pkg.add("Switch")

using Distributions
using Switch

function rGI0(alpha ,gam,n,gen)
    data=ones(n);
    L = 1;
    @switch gen begin
        @case "Gamma"
            fg=Gamma(L,1/L)
            dg=Gamma(-alpha,1/gam)
            YG=rand(fg,n)
            XG=rand(dg,n)
            data=YG./XG

            break
        @case "Chi2"
            fc=Chisq(2)
            dc=Chisq(2*alpha)
            XC=rand(fc,n)
            YC=rand(dc,n)
            data=gam.*XC./YC

            break
        @case "F-Snedecor"
            ff=FDist(2,2*alpha)
            XF=rand(ff,n)
            data=gam.*XF./alpha

            break
        @case "Pareto"
            XU=rand(n)
            data=gam.*((1-XU).^(-1/alpha)-1)

    end
    return data
end

```

end

### 3.4. Costo Computacional

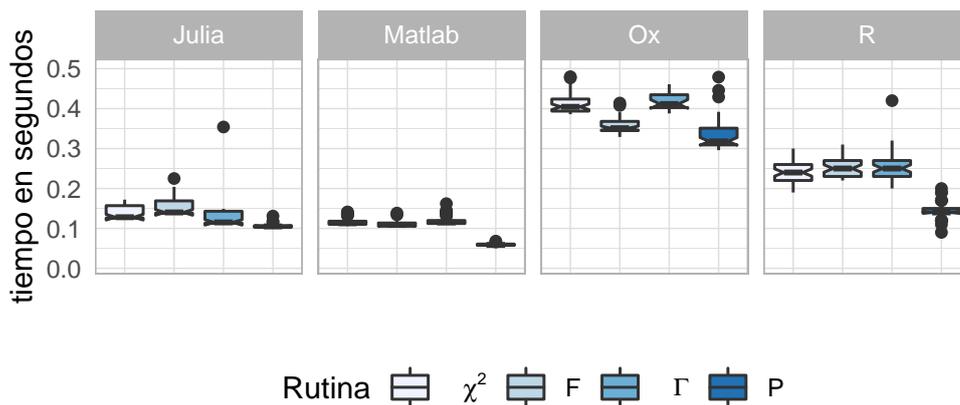
Una característica importante que debe analizarse en este contexto es el tiempo que consume cada una de las rutinas de generación de datos.

Se desarrolló la siguiente experiencia utilizando un procesador Intel© Core™, i7-6700K CPU 3.4 GHz, 16 SIGB RAM, con sistema operativo de 64 bit.

Se consideraron las combinaciones de parámetros  $(\alpha, \gamma) \in \{-8, -7, -6, -5, -4, -3, -2\} \times \{0, 1, 1, 5, 10, 25, 50, 100\}$ , en cada uno de los lenguajes de programación detallados en la Sección 3.2. Se generaron muestras de tamaño  $10^6$  con cada una de las rutinas detalladas para cada combinación de parámetros.

En la Figura 3.1, página 65 se grafica el tiempo consumido para generar estas muestras con estos parámetros, particionando la representación por 'generador y por lenguaje de programación.

**Figura 3.1:** Tiempo de procesamiento de cada generador de acuerdo con el lenguaje de programación.



Con el objetivo de comparar el costo computacional de cada una de las rutinas en los diferentes lenguajes de programación considerados se replicó

esta experiencia dos veces y se aplicó el test de Friedman [37] que corresponde a un análisis de la varianza en bloques, utilizando en el primer caso como bloques a los lenguajes y en el segundo caso como bloques a las rutinas.

Se observa que en todas las plataformas la Rutina  $P$  consume un tiempo computacional notablemente menor que las restantes.

En las pruebas post-hoc del test surge que las diferencias entre los distintos generadores no alcanzan significación estadística ( $p$ -valor = 0,065), mientras que las diferencias entre lenguajes sí ( $p$ -valor = 0,007) y la mayor diferencia fue registrada a posteriori entre los tiempos consumidos por Ox y por Matlab.

### 3.4.1. Momentos

Se generaron 1000 muestras aleatorias de tamaño 500, con los siguientes parámetros:  $(\alpha, \gamma) \in \{-8, -5, -3\} \times \{0,1, 1, 10, 100, 1000\}$ .

Para cada muestra se calcularon la media y la varianza muestrales. Se desconsideró en esta sección el valor de  $\alpha = -1,5$ , puesto que el momento de segundo orden no está definido en dicho caso.

Utilizando la ecuación de los momentos (2.19) se calculó el verdadero valor de media y varianza de la correspondiente distribución  $\mathcal{G}_I^0(\alpha, \gamma, 1)$  de cada muestra generada.

Para evaluar la precisión de los momentos muestrales ( $\hat{\theta}$ ) respecto de los verdaderos valores ( $\theta$ ) se utilizaron las siguientes medidas:

- **Error Cuadrático Medio (ECM)**

$$\widehat{\text{ECM}}(\hat{\theta}) = \frac{1}{n} \sum_{i=1}^n (\hat{\theta}_i - \theta)^2.$$

- **Error Absoluto Medio (EAM)**

$$\widehat{\text{EAM}}(\hat{\theta}) = \frac{1}{n} \sum_{i=1}^n |\hat{\theta}_i - \theta|.$$

---

- **Máximo Error Absoluto (MxEA)**

$$\widehat{\text{MxEA}}(\hat{\theta}) = \max_{1 \leq i \leq n} \{|\hat{\theta}_i - \theta|\}.$$

donde  $n$  es el tamaño muestral. El EAM fue utilizado en múltiples aplicaciones de procesamiento de imágenes dado que sus requerimientos de memoria son sensiblemente menores que los del ECM [96].

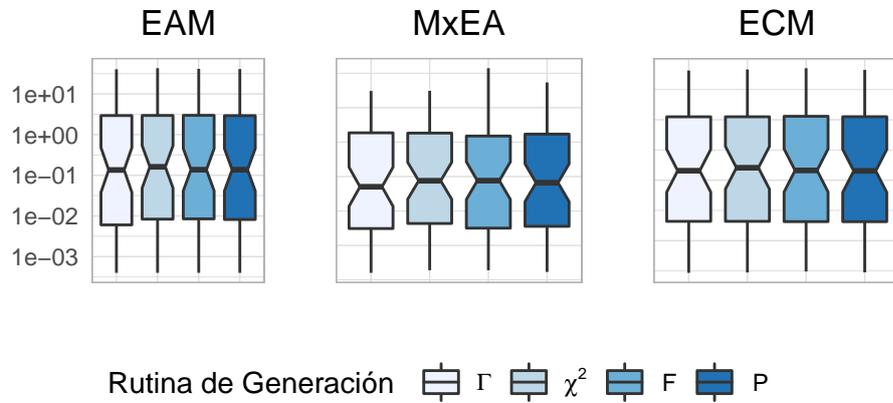
Con el objeto de comparar estas medidas por rutina de generación y por lenguaje se condujeron pruebas de análisis de la varianza no paramétricas de Kruskal-Wallis (KW) [61] y sus respectivas comparaciones post-hoc en caso de resultar significativas las diferencias de posiciones centrales.

Las Figuras 3.2 y 3.3 de la página 68) exhiben los boxplots de las medidas de error para la media en escala logarítmica, de acuerdo con la rutina de generación y con el lenguaje de programación respectivamente. Estas pruebas señalan que Julia es diferente de los restantes lenguajes de programación en el EAM y también se encontraron diferencias significativas entre Ox y R, sin embargo no se hallaron diferencias significativas en la bondad e los momentos entre Matlab y R.

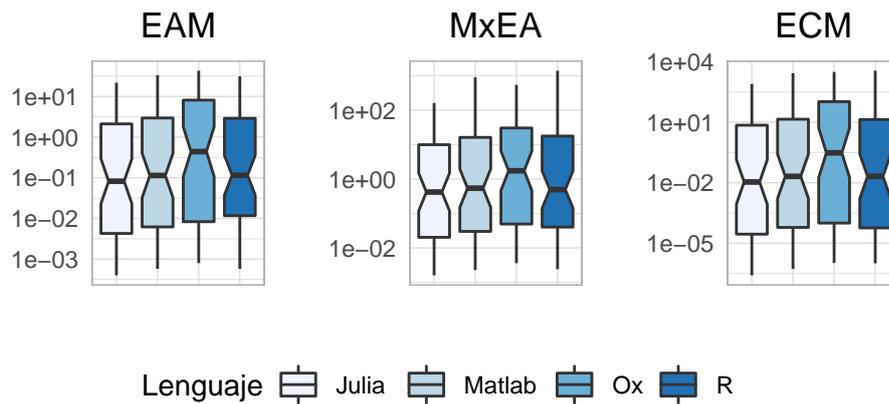
Respecto del MxEA, donde también se hallaron diferencias significativas, se puede comentar que Ox es el que alcanza menor precisión en la estimación del momento de primer orden, mientras que Julia resultó el más exacto y entre estos dos extremos se hallan sin diferencias entre si Matlab y R. Para el ECM las observaciones son similares.

Las Figuras 3.4 y 3.5 páginas 69, 69 muestran los boxplots de las medidas de error para la varianza en escala logarítmica, de acuerdo con la rutina de generación y con el lenguaje de programación respectivamente. En este caso ECM, EAM y MxEA presentan diferencias estadísticamente significativas de acuerdo con el lenguaje de programación, siendo en todos los casos el de peor desempeño el Ox. Estos boxplots tienen incorporada una muesca y un intervalo de confianza centrado en la mediana y cuyos extremos son el valor mediano

**Figura 3.2:** Errores de la media por Rutina de Generación



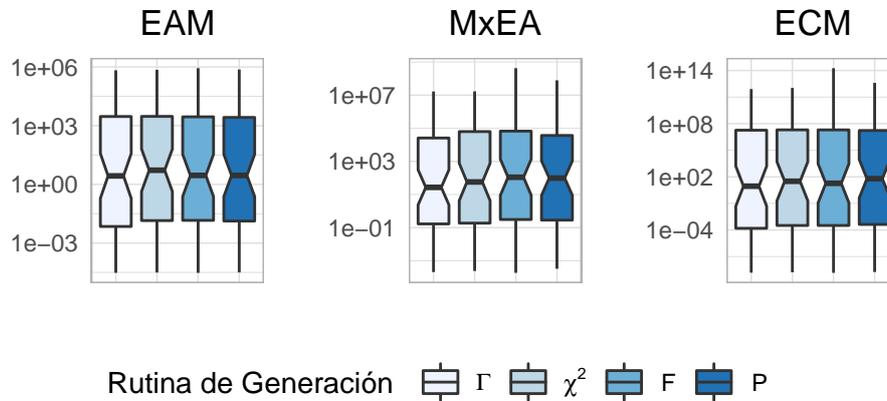
**Figura 3.3:** Errores de la media por lenguaje de programación



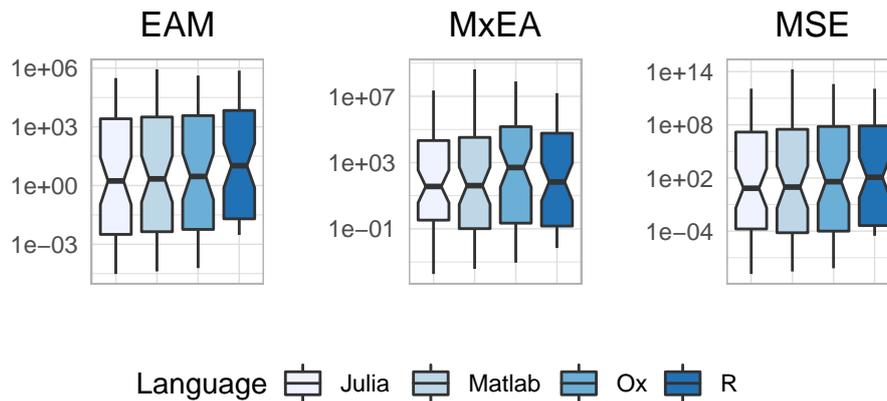
$\pm 1,58\text{IQR}$  donde IQR es la distancia entre el primero y el tercer cuartil de la distribución empírica. Cuando las muescas de dos diagramas de caja no se superponen, se considera una fuerte evidencia de que las medianas difieren de ambas distribuciones.

Para el caso del EAM se aprecia la presencia de dos subgrupos: Ox y R por un lado, y Julia y Matlab por el otro.

**Figura 3.4:** Errores de la Varianza por Rutina



**Figura 3.5:** Errores de la Varianza por Lenguaje de Programación



### 3.5. Bondad de Ajuste

En esta sección analizamos la bondad de ajuste de los datos generados a su distribución teórica utilizando valores representativos del parámetro  $\alpha$  y un valor del parámetro  $\gamma$  ligado al primero por la condición de esperanza unitaria; es decir  $\gamma^* = -\alpha - 1$ .

En todos los entornos de programación y con cada una de las cuatro rutinas propuestas se generaron 1000 muestras aleatorias con 1000 observaciones cada una de ellas considerando los valores del parámetro  $\alpha \in \{-8, -5, -3, -1,5\}$  y

su correspondiente  $\gamma^*$ .

Se han aplicado el clásico test de Kolmogorov-Smirnov [91] y el test de Anderson Darling [93] para evaluar bondad de ajuste a la distribución original. El primero de estos tests está basado en el comportamiento de la máxima diferencia registrada entre las funciones de distribución empírica y teórica (Ec. (2.18)). Una de las ventajas de esta prueba es que su estadístico es independiente de la distribución de los datos cuando los verdaderos parámetros de la distribución son conocidos, que es nuestro caso. Sin embargo, una desventaja es que esta prueba resulta menos sensible cuando los apartamientos respecto de la distribución teórica se hallan en las colas de la distribución. Debido a ello se le han hecho propuestas de varias correcciones, una de ellas es la segunda prueba que otorga mayor peso a las colas.

El nivel de significación utilizado en todas las pruebas fue  $\eta = 0,05$ .

La Tabla 3.2 muestra los  $p$ -valores promedio obtenidos en cada una de estas dos pruebas en cada uno de los lenguajes de programación, considerando todas las rutinas. Los niveles medio de rechazo tanto de la prueba de KS como del test de AD para los datos generados no tiene diferencias estadísticamente significativas respecto del nivel teórico establecido, ni por Rutina ni por Lenguaje de Programación.

## 3.6. Distribución muestral de los valores extremos

### 3.6.1. Distribución del máximo muestral

Consideremos una muestra  $Z_1, \dots, Z_n$  de variables aleatorias independientes con idéntica distribución  $\mathcal{G}_I^0(\alpha, \gamma, 1)$  cuyo ordenamiento en forma no decreciente es  $Z_{1:n} \leq \dots \leq Z_{n:n}$ .

**Tabla 3.2:** Resultados de los tests KS y AD

Test	Lenguaje	$\Gamma$	$\chi^2$	$F$	$P$
KS	R	0,0498	0,0495	0,0485	0,0470
	Matlab	0,0420	0,0420	0,0420	0,0420
	Ox	0,0410	0,0520	0,0440	0,0523
	Julia	0,0490	0,0409	0,0310	0,0530
AD	R	0,0493	0,0498	0,0490	0,0488
	Matlab	0,0430	0,0430	0,0430	0,0430
	Ox	0,0350	0,0523	0,0440	0,0578
	Julia	0,0510	0,0470	0,0510	0,0570

Definamos a partir de la muestra de tamaño  $n$  dos nuevas variables aleatorias el mínimo  $U$  y el máximo  $V$  muestrales.

$$U_n = \min_{1 \leq i \leq n} \{Z_1, Z_2, \dots, Z_n\} \quad (3.2)$$

y

$$V_n = \max_{1 \leq i \leq n} \{Z_1, Z_2, \dots, Z_n\} \quad (3.3)$$

Las funciones de densidad de probabilidad de estas nuevas variables aleatorias son:

$$f_{U_n}(u) = -n(1 - F_Z(u))^{n-1} f_Z(u) = -\frac{n\alpha}{\gamma} \left(1 + \frac{u}{\gamma}\right)^{n\alpha-1}, \quad (3.4)$$

y

$$f_{V_n}(v) = n(F_Z(v))^{n-1} f_Z(v) = -\frac{n\alpha}{\gamma} \left(1 + \frac{v}{\gamma}\right)^{\alpha-1} \left[1 - \left(1 + \frac{v}{\gamma}\right)^\alpha\right]^{n-1}, \quad (3.5)$$

donde  $F_Z$  and  $f_Z$  son las funciones de distribución y densidad respectivamente de una variable  $Z \sim \mathcal{G}_I^0(\alpha, \gamma, 1)$  respectivamente.

Las colas de una distribución son importantes en el análisis de distribuciones de valores extremos, como es la  $\mathcal{G}_I^0$ .

Dada la propensión a valores extremos altos y la variación lenta en el infinito de la distribución  $\mathcal{G}_I^0$ , es pertinente analizar el comportamiento de los valores extremos generados por cada una de las rutinas.

Con este fin se generaron 1000 muestras de máximos muestrales correspondientes a muestras de cada uno de los siguientes tamaños  $n \in \{50, 100, 500, 1000\}$ . Luego se evaluó la bondad de ajuste de estas muestras de máximos a sus distribuciones teóricas correspondientes mediante los tests de KS y AD. Las Tablas 3.3 y 3.4 muestran los  $p$ -valores muestrales medios para un nivel de significación del  $\eta = 0,05$ .

La rutina  $F$  alcanza el menor nivel de significación en todos los casos tanto para el test KS como para el test AD, para la distribución de los máximos muestrales. Sin embargo no se hallaron diferencias estadísticamente significativas para los niveles medios de rechazo ni por rutina ni por lenguaje de programación.

### 3.6.2. Distribución del mínimo muestral

La presencia de outliers bajos perjudica el comportamiento del estimador de máxima verosimilitud, dado que este se basa en la suma de los logaritmos de las observaciones. En [3], los autores muestran que los outliers altos, cercanos al modelo no generan tantos inconvenientes como los outliers bajos para lograr una estimación de los parámetros, especialmente en regiones extremadamente heterogéneas como las urbanas.

Esta observación ha motivado el estudio del comportamiento del mínimo muestral para esta distribución.

---

Con este fin, se generaron 1000 muestras de mínimos muestrales correspondientes a muestras de cada uno de los siguientes tamaños  $n \in \{50, 100, 500, 1000\}$ . Luego se evaluó la bondad de ajuste de estas muestras de mínimos a sus distribuciones teóricas correspondientes mediante los tests de KS y AD. Hemos considerado  $n \in \{50, 100, 500, 1000\}$  y generamos 1000 muestras de tamaños 1000 de mínimos muestrales de muestras de tamaño  $na$ .

En las Tablas 3.5 y 3.6 se presentan los  $p$ -valores de las pruebas de bondad de ajuste para los mínimos muestrales con un nivel de significación del  $\eta = 0,05$ . Como en los casos anteriores se compararon los niveles de rechazo y no se hallaron diferencias estadísticamente significativas ni por lenguaje de programación ni por rutina.

En el caso de la distribución del mínimo muestral, la rutina  $F$  presenta mejor comportamiento para ambos tests puesto que sus niveles de significación empíricos son más cercanos a los teóricos correspondientes.

### 3.7. Intervalos de Confianza para Cuantiles

Para cada una de las muestras cuyo detalle de generación fue referido en la Sección Sec. 3.6.1, registramos el tercer cuartil y el percentil 90 así como los intervalos de confianza de nivel 95% para ellos, basados en el método de Bootstrap [111].

Las Tablas 3.7 y 3.8 exhiben el porcentaje de cobertura de estos intervalos sobre el verdadero valor del cuantil correspondiente.

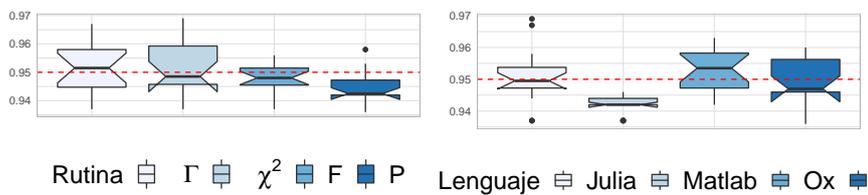
La primera observación interesante que cabe destacar es que el porcentaje de cobertura no depende del valor del parámetro de textura.

La interacción entre la rutina y el lenguaje de programación no resultó significativa ni para el  $p_{75}$  ni para el  $p_{90}$ . El test KW señala sin embargo diferencias significativas en el porcentaje de cobertura por lenguaje de programación. Las comparaciones a posteriori indican que las coberturas de  $\chi^2$  y  $P$  son estadís-

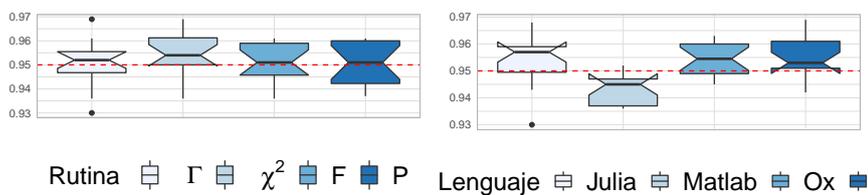
ticamente distintas para el  $p_{75}$ . Estas rutinas se distinguen en la cantidad de variables aleatorias generadas, mientras que el generador  $P$  realiza una única generación y es una aplicación del método de inversión, el generador  $\chi^2$  requiere de la generación de dos números aleatorios, es probable que de este modo pierda precisión.

En las gráficas 3.6 and 3.7 de la página 74 se pueden apreciar los porcentajes de cobertura de los intervalos de confianza de los dos percentiles y con una línea roja punteada se señala el nivel de confianza establecido en la construcción de los intervalos.

**Figura 3.6:** Porcentaje de cobertura del  $p_{75}$



**Figura 3.7:** Porcentaje de cobertura del  $p_{90}$



### 3.8. Discusión y Conclusiones

Este capítulo está dedicado a la generación de datos con distribución  $\mathcal{G}_I^0$ , las rutinas  $P$  y  $F$  tiene códigos de programación más sencillos que las otras, puesto que se basan en el método de inversión. La última rutina sólo es aplicable en el caso Single Look.

---

Si focalizamos en el tiempo de procesamiento, claramente la rutina  $P$  es la mejor ya que aventaja a las restantes en el costo computacional sobre todos los entornos de programación. Si comparamos entre los lenguajes, la diferencia es marcada entre Ox y Matlab.

Si en cambio nos preocupa la bondad de ajuste tanto de los datos crudos como de los extremos, no se aprecia ventaja en ningún lenguaje ni rutina.

Respecto de la precisión de los momentos de primer y segundo orden, podemos establecer dos grupos: por un lado Ox y R, y por el otro Julia y Matlab siendo el segundo grupo el que logra mayor fidelidad de los momentos teóricos en los momentos empíricos de los datos generados.

Por último si enfocamos en las coberturas de los intervalos correspondientes a las rutinas  $\chi^2$  y  $P$  son significativamente diferentes para el tercer cuartil pero no hay diferencias para el  $P_{90}$ . Sin embargo en ambos casos hay diferencias entre lenguajes de programación y en este caso el desempeño de Ox, R y Julia es superior al de Matlab.

Los resultados de la comparación del desempeño de cada una de estas cuatro rutinas alternativas para generar datos aleatorios con distribución  $\mathcal{G}_I^0$  se comunicaron inicialmente sobre la plataforma R en [16] y luego sobre se añadió el análisis sobre Matlab, Julia y Ox en [17].

**Tabla 3.3:**  $p$ -valores del test KS para máximos muestrales

KS test		Tamaño muestral				Promedio
Leng.	Rutina	50	100	500	1000	
R	$\Gamma$	0,049	0,046	0,050	0,051	0,049
	$\chi^2$	0,048	0,047	0,051	0,056	0,050
	$F$	0,053	0,075	0,042	0,046	0,054
	$P$	0,041	0,049	0,053	0,047	0,048
Matlab	$\Gamma$	0,053	0,050	0,043	0,050	0,049
	$\chi^2$	0,048	0,049	0,046	0,044	0,047
	$F$	0,052	0,046	0,046	0,047	0,048
	$P$	0,052	0,057	0,053	0,048	0,053
Ox	$\Gamma$	0,054	0,045	0,051	0,049	0,050
	$\chi^2$	0,052	0,041	0,052	0,052	0,051
	$F$	0,048	0,043	0,051	0,052	0,048
	$P$	0,049	0,050	0,047	0,056	0,051
Julia	$\Gamma$	0,048	0,043	0,051	0,052	0,048
	$\chi^2$	0,053	0,047	0,047	0,047	0,048
	$F$	0,045	0,047	0,049	0,047	0,047
	$P$	0,054	0,058	0,070	0,052	0,059

**Tabla 3.4:**  $p$ -valores del test AD para máximos muestrales

AD test		Tamaño Muestral				Promedio
Leng.	Rutina	50	100	500	1000	
R	$\Gamma$	0,049	0,044	0,054	0,056	0,051
	$\chi^2$	0,049	0,052	0,047	0,057	0,051
	$F$	0,053	0,065	0,047	0,047	0,053
	$P$	0,038	0,048	0,059	0,048	0,048
Matlab	$\Gamma$	0,051	0,048	0,047	0,050	0,049
	$\chi^2$	0,047	0,051	0,051	0,047	0,049
	$F$	0,052	0,045	0,047	0,049	0,048
	$P$	0,051	0,051	0,049	0,055	0,052
Ox	$\Gamma$	0,053	0,045	0,047	0,048	0,048
	$\chi^2$	0,054	0,044	0,055	0,054	0,052
	$F$	0,048	0,046	0,052	0,055	0,050
	$P$	0,053	0,052	0,051	0,053	0,052
Julia	$\Gamma$	0,053	0,054	0,054	0,043	0,051
	$\chi^2$	0,050	0,048	0,045	0,038	0,045
	$F$	0,046	0,047	0,052	0,046	0,048
	$P$	0,056	0,060	0,057	0,045	0,055

Tabla 3.5:  $p$ -valores del test KS para mínimos muestrales

KS test		Tamaño muestral				Promedio
Leng.	Rutina	50	100	500	1000	
R	$\Gamma$	0,050	0,049	0,046	0,047	0,048
	$\chi^2$	0,050	0,049	0,046	0,047	0,048
	$F$	0,049	0,049	0,043	0,047	0,047
	$P$	0,055	0,044	0,050	0,050	0,050
Matlab	$\Gamma$	0,048	0,048	0,052	0,047	0,049
	$\chi^2$	0,055	0,042	0,045	0,048	0,048
	$F$	0,044	0,043	0,051	0,041	0,045
	$P$	0,045	0,050	0,055	0,042	0,048
Ox	$\Gamma$	0,092	0,088	0,096	0,088	0,091
	$\chi^2$	0,065	0,061	0,065	0,074	0,066
	$F$	0,071	0,066	0,076	0,068	0,071
	$P$	0,054	0,067	0,064	0,085	0,068
Julia	$\Gamma$	0,055	0,048	0,053	0,056	0,053
	$\chi^2$	0,050	0,046	0,051	0,049	0,049
	$F$	0,046	0,048	0,049	0,051	0,048
	$P$	0,049	0,047	0,050	0,046	0,048

**Tabla 3.6:**  $p$ -valores del test AD para mínimos muestrales

AD test		Tamaño muestral				Promedio
Leng.	Rutina	50	100	500	1000	
R	$\Gamma$	0,047	0,050	0,050	0,051	0,050
	$\chi^2$	0,047	0,050	0,050	0,051	0,050
	$F$	0,049	0,051	0,047	0,049	0,049
	$P$	0,055	0,047	0,052	0,050	0,051
Matlab	$\Gamma$	0,049	0,046	0,052	0,048	0,049
	$\chi^2$	0,053	0,043	0,050	0,055	0,050
	$F$	0,048	0,049	0,050	0,047	0,048
	$P$	0,047	0,050	0,055	0,044	0,049
Ox	$\Gamma$	0,086	0,077	0,086	0,077	0,081
	$\chi^2$	0,059	0,057	0,058	0,070	0,061
	$F$	0,073	0,062	0,073	0,063	0,068
	$P$	0,051	0,064	0,067	0,082	0,066
Julia	$\Gamma$	0,051	0,048	0,053	0,054	0,051
	$\chi^2$	0,053	0,049	0,049	0,047	0,049
	$F$	0,051	0,049	0,050	0,054	0,051
	$P$	0,049	0,049	0,055	0,046	0,050

Tabla 3.7: Cobertura del intervalo del  $q_3$

Leng.	Rutina	$\alpha$				Promedio
		-8	-5	-3	-1,5	
R	$\Gamma$	0,960	0,957	0,946	0,955	0,9545
	$\chi^2$	0,960	0,959	0,946	0,960	0,9470
	$F$	0,946	0,947	0,948	0,956	0,9493
	$P$	0,947	0,936	0,936	0,936	0,9388
Matlab	$\Gamma$	0,937	0,942	0,944	0,946	0,9423
	$\chi^2$	0,937	0,942	0,944	0,946	0,9423
	$F$	0,937	0,942	0,944	0,946	0,9423
	$P$	0,942	0,942	0,942	0,942	0,9420
Ox	$\Gamma$	0,955	0,958	0,945	0,961	0,9548
	$\chi^2$	0,961	0,948	0,959	0,963	0,9578
	$F$	0,953	0,951	0,954	0,950	0,9520
	$P$	0,942	0,958	0,944	0,943	0,9468
Julia	$\Gamma$	0,948	0,937	0,958	0,967	0,9525
	$\chi^2$	0,969	0,956	0,949	0,945	0,9548
	$F$	0,950	0,948	0,953	0,937	0,9470
	$P$	0,948	0,944	0,951	0,953	0,9490

**Tabla 3.8:** Cobertura del intervalo del  $p_{90}$

Leng.	Rutina	$\alpha$				Promedio
		-8	-5	-3	-1,5	
R	$\Gamma$	0,969	0,951	0,952	0,946	0,955
	$\chi^2$	0,969	0,942	0,952	0,951	0,954
	$F$	0,954	0,960	0,950	0,961	0,956
	$P$	0,952	0,961	0,961	0,961	0,959
Matlab	$\Gamma$	0,947	0,954	0,952	0,936	0,945
	$\chi^2$	0,947	0,945	0,952	0,936	0,945
	$F$	0,947	0,945	0,952	0,936	0,945
	$P$	0,937	0,937	0,937	0,937	0,937
Ox	$\Gamma$	0,949	0,954	0,961	0,957	0,955
	$\chi^2$	0,953	0,959	0,961	0,963	0,959
	$F$	0,945	0,952	0,946	0,946	0,947
	$P$	0,955	0,960	0,960	0,949	0,956
Julia	$\Gamma$	0,930	0,953	0,959	0,955	0,949
	$\chi^2$	0,962	0,968	0,959	0,955	0,961
	$F$	0,960	0,943	0,959	0,959	0,955
	$P$	0,950	0,944	0,948	0,959	0,950

**Tabla 3.9:** Comparaciones post-hoc por lenguaje ( $p_{75}$ )

	Diferencia Observada	Diferencia Crítica	Significación
Julia-Matlab	22,44	17,37	<b>V</b>
Julia-Ox	4,56	17,37	<b>F</b>
Julia-R	3,88	17,37	<b>F</b>
Matlab-Ox	27	17,37	<b>V</b>
Matlab-R	18,57	17,37	<b>V</b>
Ox-R	8,44	17,37	<b>F</b>

## Capítulo 4

# Estimación de parámetros de la distribución $\mathcal{G}_I^0$ - *Single Look*

La estimación de parámetros en el estudio de datos provenientes de imágenes de radar de apertura sintética constituyen un insumo fundamental por su utilidad en los métodos de segmentación, clasificación y filtrado de las mismas al tiempo que posibilitan una mayor comprensión de las mismas.

Muchos de los métodos de filtrado de imágenes o detección de bordes utilizan máscaras deslizantes de tamaños  $3 \times 3$ ,  $5 \times 5$ ,  $7 \times 7$ ,  $9 \times 9$  u  $11 \times 11$ . Dado que algunas de estas muestras son pequeñas y otras contaminadas es importante encontrar estimadores para los parámetros de la distribución  $\mathcal{G}_I^0$  que se desempeñen eficazmente frente a estas limitaciones de las muestras.

Entre las propiedades deseables de un buen estimador cabe destacar: el insesgamiento, la consistencia, la eficiencia, la normalidad asintótica, la simplicidad en su implementación, el bajo costo computacional y la resistencia a diversos tipos de contaminación de los datos, es decir la robustez.

El procedimiento paramétrico consiste, generalmente, en estimar los parámetros del modelo elegido para describir la variable de interés y evaluar el grado de adecuación de la distribución resultante a los datos reales. Para ello se consideran el sesgo, la varianza y el error cuadrático medio de cada estimador o bien se aplican otras medidas de bondad de ajuste disponibles.

Entre los métodos clásicos más populares para la estimación de los parámetros de la distribución  $\mathcal{G}_I^0$  se puede mencionar al método de analogía que incluye como casos particulares al método de momentos (MoM) [83, 94, 104] y al de logmomentos [81], al de cuantiles [74], y al de máxima verosimilitud (ML) [30, 33].

Los trabajos [35, 74] estudian el comportamiento del estimador de máxima verosimilitud (ML) que evidencia un excelente comportamiento en términos de consistencia y normalidad asintótica; sin embargo, presenta serios problemas de convergencia y sesgo. Entre las propuestas de mejora de estos métodos podemos citar al de máxima verosimilitud con corrección de sesgo basado en la propuesta teórica de Cox y Snell [21, 109] y el basado en bootstrap [23, 34].

Entre las alternativas robustas, los M-estimadores resultan confiables, cuando la imagen presenta contaminación por la presencia de 'corner reflectors', que son dispositivos esenciales para la calibración de los datos y dan lugar a retornos mucho más altos en un píxel que en el resto de la imagen.

En [72] se verifica que los GM-estimadores funcionan tan bien como el estimador de máxima verosimilitud para datos no contaminados y superan a estos en presencia de outliers. Se han propuesto también los AM estimadores [3] es decir M-estimadores con funciones de influencia asimétricas; estos mostraron mejor comportamiento que el estimador de máxima verosimilitud y que los M-estimadores para ciertos esquemas de contaminación.

Entre las estimaciones no paramétricas se encuentra la estimación por núcleos asimétricos minimizando las distancias estocásticas entre la distribución empírica y la distribución teórica [13, 39].

Recientemente, la metodología basada en logmomentos y logcumulantes (MoLC) [2, 4] ha sido agregada como un nuevo enfoque. Nicolas [81] propone sustituir la transformada de Fourier, que no ha mostrado buenos resultados con funciones de soporte sobre los reales positivos, por la transformada de Mellin. Por analogía con las funciones características, los cumulantes y log-

---

cumulantes, el autor introduce los conceptos de funciones características de segundo tipo, los logmomentos y los logcumulantes.

Dado que el comportamiento asintótico de la distribución  $\mathcal{G}_I^0$  es comparable con una distribución de potencia con índice de cola  $1 - \alpha$ , inicialmente también fue muy utilizada la estimación de la pendiente del histograma graficado en escala logarítmica que, a pesar de su sencillez, introduce un sesgo sistemático [80]. También se han propuesto estimadores basados en el índice de la cola de la distribución de Pareto [1], entre ellas una de las más populares es el estimador de Hill [54] definido en función de la función de verosimilitud condicional.

Múltiples correcciones han sido propuestas para este estimador, algunas basadas en la mejora del gráfico de Hill [60], otras basadas en la reducción del sesgo del mismo [10] y algunas enfocadas en el comportamiento del estimador en algunas situaciones específicas [82].

La estimación del parámetro de textura para la distribución  $\mathcal{G}_I^0$  es básica para el caso de los filtros estadísticos locales, sobre todo para muestras chicas y/o contaminadas. Una de nuestras principales motivaciones es evaluar las técnicas de estimación para tales situaciones.

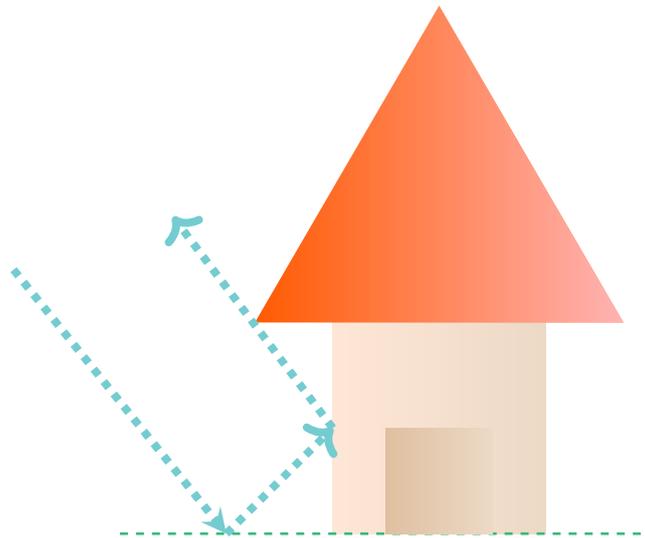
En este capítulo proponemos evaluar el comportamiento de varios estimadores disponibles para el parámetro de textura de la distribución unidimensional  $\mathcal{G}_I^0$  en el caso más ruidoso, es decir *single look* ( $L = 1$ ), cuya precisión y exactitud son esenciales para la interpretación de estas imágenes.

Con este objetivo utilizamos criterios como la convergencia, el sesgo, el ECM, la robustez y el costo computacional, considerando muestras de diferentes tamaños, considerando muestras con datos puros y muestras con datos contaminados.

## 4.1. Contaminaciones

Algunos efectos como el double bounce o doble rebote de los pulsos entrantes, son característicos de este tipo de imágenes. Este efecto ocurre, por ejemplo, cuando el pulso rebota en la calle y luego en un edificio y vuelve al radar (ver Figura 4.1). Los objetos aparecen muy brillantes (blancos) en las imágenes de radar [83].

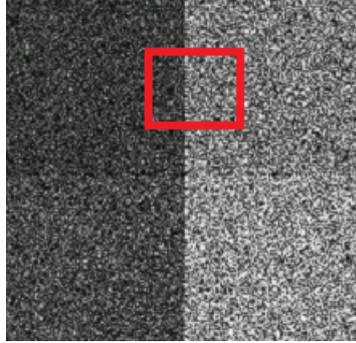
**Figura 4.1:** Efecto double bounce



Como hemos destacado previamente, la robustez, es decir la habilidad de producir estimaciones aceptables aún cuando una pequeña proporción de los datos no proviene del modelo que se ha supuesto, es una propiedad deseable para cualquier buen estimador. Al utilizar ventanas deslizantes para la estimación de los parámetros es posible que en una de estas ventanas encontremos datos con texturas diferentes, especialmente en el borde de dos regiones, como ilustra la Figura 4.2.

A continuación se describen los esquemas de contaminación más frecuentes en datos SAR. Con el propósito de formalizar los tres casos, consideremos dos variables aleatorias con distribución  $W \sim \mathcal{G}_I^0(\alpha_1, \gamma_1, 1)$  y  $U \sim \mathcal{G}_I^0(\alpha_2, \gamma_2, 1)$ ,

**Figura 4.2:** Ventana Deslizante sobre dos regiones diferentes



donde  $\gamma_1, \gamma_2, -\alpha_1, -\alpha_2 > 0$  y  $C \in \mathbb{R}_+$  un valor grande .

Nuestra muestra es  $\{z_1, \dots, z_n\}$  donde cada  $Z_I$  tiene la forma:

- Caso 1: Cierta proporción de datos corresponden a cada una de dos distribuciones de la familia  $\mathcal{G}_I^0$  single look pero con distintos parámetros. Se define  $Z = BU + (1 - B)W$ , con función de distribución

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

donde  $\mathcal{F}_{\mathcal{G}_I^0(\alpha, \gamma, 1)}$  indica la función de distribución acumulada de una variable aleatoria  $\mathcal{G}_I^0(\alpha, \gamma, 1)$ .

- Caso 2: Una pequeña proporción de los datos son outliers altos,  $Z = BC + (1 - B)W$ , con  $C = 10^k$   $k \in \mathbb{N}$ , donde  $B \sim Be(\epsilon)$  que indica la proporción de datos contaminados con la forma de outliers.
- Caso 3: Una pequeña proporción de los datos pertenece a la misma distribución pero con el parámetro de escala mucho mayor,  $Z = BU + (1 - B)W$  con  $\gamma_2 = 10^k \gamma_1$  y  $\alpha_1 = \alpha_2$

Todos estos casos constituyen apartamientos del modelo que establece para el retorno una distribución  $\mathcal{G}_I^0(\alpha, \gamma^*, 1)$ . El primero de los casos asume que una pequeña porción  $\epsilon$  tienen la misma distribución pero distintos valores paramétricos. El segundo asume una pequeña porción de outliers altos. El tercer caso

constituye un caso especial del primero con la particularidad de que el valor de  $\gamma$  es el único diferente y en cierto sentido. La contaminación de interés es la presencia de unos pocos valores muy alejados del centro del conjunto de observaciones resultado del fenómeno de *double bounce* o de un *corner reflector* que corresponde al segundo caso de contaminación listado.

Si bien la distribución  $\mathcal{G}_I^0$  es propensa a outliers como detallamos en la Sección 2.1, vamos a suponer como *datos puros* una colección de variables aleatorias independientes idénticamente distribuidas (v.a.i.i.d).

## 4.2. Los Estimadores

La estimación del parámetro de textura  $\alpha$  de la distribución  $\mathcal{G}_I^0$  es fundamental para la interpretación de la imagen tanto en filtros no locales como en segmentación.

En [39] se analiza un método no paramétrico basado en la minimización de la distancia Triangular entre la distribución  $\mathcal{G}_I^0$  y la función de distribución empírica estimada mediante núcleos asimétricos. Esta propuesta si bien es robusta, tiene un alto costo computacional y no se aplica al caso *Single Look*.

Por este motivo, en este trabajo se elige este caso para analizar, que se corresponde con el más ruidoso.

Dada la flexibilidad de la distribución Pareto Generalizada para modelar fenómenos en diversas áreas del conocimiento, esta ley ha sido intensivamente estudiada. Se han presentado un conjunto importante de estimadores cuyas propiedades se han analizado, aunque nunca se ha hecho en el contexto de las imágenes SAR.

Evaluamos el estimador de Máxima Verosimilitud (MV), Máxima Verosimilitud Penalizado (MVP), Verosimilitud-Momentos (VM), Momentos Ponderados con Probabilidad (MPP), Máxima Bondad de Ajuste (MBA) con diferentes estadísticos y mínima Divergencia de Densidad (mDD). En [88] se realizan

ajustes del perfil de la función de verosimilitud a fin de obtener estimaciones más precisas del parámetro de forma de la distribución GPD, focalizando en el estimador de máxima verosimilitud y el test de cociente de verosimilitudes. En esta sección vamos a detallar los métodos de estimación considerados conjuntamente con algunas de sus propiedades destacadas. En la estimación se utilizó la función `fitGPD` del paquete POT (Generalized Pareto Distribution and Peaks Over Threshold) de R [92]. Este paquete utiliza la optimización basada en el algoritmo del gradiente conjugado, Nelder-Mead y quasi-Newton.

Todas las simulaciones se realizaron utilizando el lenguaje computacional R versión 3.3 [92], en una computadora con procesador Intel© Core™, i7-4790K CPU 4 GHz, 16 GB RAM, con sistema operativo 64 bit.

En la Tabla 4.1 se presentan los métodos de estimación revisados en este capítulo, conjuntamente con sus abreviaturas.

**Tabla 4.1:** Lista de Abreviaturas

Abreviatura	Estimador	Notación
MV	Máxima Verosimilitud	$\hat{\alpha}_{MV}$
MVP	Máxima Verosimilitud Penalizado	$\hat{\alpha}_{MVP}$
Mom	Momentos	$\hat{\alpha}_{Mom}$
MPP	Momentos Ponderados con la Probabilidad	$\hat{\alpha}_{MPP}$
VM	Verosimilitud Momentos	$\hat{\alpha}_{VM}$
Med	Medianas	$\hat{\alpha}_{Med}$
mDD	Mínima divergencia de densidad	$\hat{\alpha}_{mDD}$
ADD	Máxima Bondad de Ajuste(MBA) con estadístico de Anderson Darling (cola derecha)	$\hat{\alpha}_{ADD}$

El código utilizado para esta implementación se consigna a continuación:

**Rutina 4.1:** Métodos para el parámetro de forma de la GPD

```

#samp es la muestra aleatoria simulada

#sh.meth es el metodo de estimacion elegido
#Metodos para el parametro de forma de la GPD

#alpha.meth es la estimacion lograda para el metodo meth
# para el parametro de textura de la distribucion GIO
#"method" es el metodo de estimacion elegido

sh.meth=fitgpd(samp, threshold=0,
               est='method')$param[[2]]
#esta es la estimacion del parametro de escala de la GPD

alpha.meth=min(-0.1, max(-20,-1/sh.meth))
#alpha.meth = -20 or alpha.meth = -0.1
#controla la convergencia del metodo de estimacion

#el usuario debe reemplazar "method" por
#"mle": Metodo de Maxima Verosimilitud
#"mple": Metodo de Maxima Verosimilitud Penalizado
#"mom": Metodo de los Momentos
#"pwmb": Momentos Ponderados con la Probabilidad
#method
#"lme": Metodo de Verosimilitud-Momentos
#"med": Metodo de las Medianas
#"mdpd": Metodo de la minima divergencia de la distribucion
#"mgf" (con el estadistico stat='ADR'): Maxima Bondad de ajuste
#con el estadistico Anderson Darling (cola derecha).

```

#### 4.2.1. Máxima Verosimilitud (MV)

Dada una muestra aleatoria  $z = (z_1, \dots, z_n)$  de variables aleatorias independientes con idéntica distribución  $\mathcal{G}_I^0(\alpha, \gamma, 1)$  con  $(\alpha, \gamma) \in \Theta = \mathbb{R}_- \times \mathbb{R}_+$ , el

---

estimador de máxima verosimilitud  $(\hat{\alpha}_{MV}, \hat{\gamma}_{MV})$  satisface la siguiente condición

$$(\hat{\alpha}, \hat{\gamma}) = \arg \max_{(\alpha, \gamma) \in \Theta} \mathcal{L}(\alpha, \gamma, \mathbf{1}, \mathbf{z}) = \arg \max_{(\alpha, \gamma) \in \Theta} \prod_{i=1}^n f_{G_T^0(\alpha, \gamma, 1)} z_i,$$

siendo  $\mathcal{L}$  la función de verosimilitud. La maximización de la función de logverosimilitud nos conduce a las estimaciones de  $\hat{\alpha}$  y  $\hat{\gamma}$  que satisfacen simultáneamente

$$\frac{1}{\hat{\alpha}_{MV}} - \ln \hat{\gamma}_{MV} + \frac{1}{n} \sum_{i=1}^n \ln(z_i + \hat{\gamma}_{MV}) = 0 \text{ y} \quad (4.1)$$

$$\frac{\hat{\alpha}_{MV}}{\hat{\gamma}_{MV}} + \frac{1}{n} (1 - \hat{\alpha}_{MV}) \sum_{i=1}^n (\hat{\gamma}_{MV} + z_i)^{-1} = 0. \quad (4.2)$$

La estimación de máxima verosimilitud también puede obtenerse maximizando la función de log-verosimilitud reducida.

El estimador de máxima verosimilitud ha sido ampliamente estudiado debido a sus propiedades asintóticas [48], sin embargo presenta problemas de sesgo y de robustez especialmente en muestras pequeñas. En [35] se propone una técnica para corregir el sesgo en pequeñas muestras. En [112] se desarrolló un estimador robusto basado en la asignación de ponderadores aleatorios. Los autores evaluaron este estimador bajo distintos esquemas de contaminación y concluyeron que cuanto mayor es el número de looks, menor es la proporción de casos de no convergencia. Esto justifica el análisis particular del caso Single Look. Grimshaw en [48] estudió las propiedades del estimador de máxima verosimilitud para la distribución de Pareto Generalizada. Mostró su consistencia asintótica, su eficiencia y normalidad, pero cabe destacar que no tiene solución explícita y presenta problemas de convergencia y robustez, especialmente en muestras pequeñas.

#### 4.2.2. Máxima Verosimilitud Penalizado (MVP)

Coles y Dixon [20], con el propósito de lograr convergencia en pequeñas muestras, propusieron minimizar la función de log-verosimilitud negativa y

penalizada.

Este estimador se obtiene

$$(\hat{\alpha}_{\text{MVP}}, \hat{\gamma}_{\text{MVP}}) = \arg \min_{(\alpha, \gamma)} \{-\ln L(\alpha, \gamma) + \ln P(\alpha)\},$$

donde  $P(\alpha)$  es la función de penalización. Los autores analizan varias funciones de penalización y proponen:

$$P(\alpha) = \begin{cases} e^{-\lambda \left(\frac{-1}{1+\alpha}\right)^\nu} & \text{si } \alpha < -1 \\ 0 & \text{si } \alpha \geq -1, \end{cases}$$

siendo  $\lambda$  y  $\nu$  valores no negativos, y sugieren el uso de  $\nu = \lambda = 1$ .

Para muestras de tamaño grande, el MVP hereda propiedades optimales del estimador de MV, pero logra sortear las limitaciones de este último para muestras pequeñas.

### 4.2.3. Momentos (Mom)

De la expresión de los momentos de orden  $r$  dada en la Ecuación 2.22 para la distribución  $\mathcal{G}_I^0(\alpha, \gamma, 1)$ .

Usando primero y segundo orden, el estimador de Momentos resulta:

$$\hat{\alpha}_{\text{Mom}} = \frac{2\hat{\sigma}^2}{\bar{z}^2 - \hat{\sigma}^2} \quad \text{y} \quad \hat{\gamma}_{\text{Mom}} = \frac{\bar{z}(\bar{z}^2 + \hat{\sigma}^2)}{\bar{z}^2 - \hat{\sigma}^2},$$

donde  $\bar{z}$  y  $\hat{\sigma}^2$  denotan la media y la varianza muestral respectivamente.

El estimador (Mom) posee propiedades asintóticas optimales tales como normalidad y eficiencia y tiene una expresión explícita para esta función, sin embargo es muy sensible a la presencia de outliers y frecuentemente no converge. No reportaremos resultados experimentales para este estimador, dado su pobre desempeño frente a los restantes.

---

#### 4.2.4. Momentos Ponderados con la Probabilidad (MPP)

Toda distribución de probabilidad puede ser caracterizada por sus momentos, así como también por sus momentos ponderados [55]. Los momentos ponderados constituyen una generalización de los momentos ordinarios, y se definen por:

$$M_{p,r,s} = E [Z^p (F(Z))^r (1 - F(Z))^s],$$

siendo  $F$  la función de distribución de la variable aleatoria  $Z$ , y  $p, r, s \in \mathbb{R}$ .

La expresión de la función de distribución para la ley  $\mathcal{G}_I^0(\alpha, \gamma, 1)$  es

$$F(z) = 1 - \left( \frac{z}{\gamma} + 1 \right)^\alpha.$$

Los momentos ponderados para  $p = 1, r = 0$  y  $s \in \mathbb{R}$  en esta ley tienen la expresión:

$$M_{1,0,s} = E(Z(1 - F_Z(z))^s) = \frac{\gamma}{(s+1)[1 + \alpha(s+1)]}.$$

Los estimadores de ambos parámetros pueden, entonces, calcularse como

$$\begin{aligned} \hat{\alpha}_{\text{MPP}} &= \frac{\widehat{M}_{1,0,0} - 2\widehat{M}_{1,0,1}}{4\widehat{M}_{1,0,1} - \widehat{M}_{1,0,0}}, \\ \hat{\gamma}_{\text{MPP}} &= 2\hat{\alpha}_{\text{MPP}} \frac{\widehat{M}_{1,0,1}\widehat{M}_{1,0,0}}{2\widehat{M}_{1,0,1} - \widehat{M}_{1,0,0}}, \end{aligned}$$

donde  $\widehat{M}_{1,0,0} = n^{-1} \sum_{i=1}^n z_i$  y  $\widehat{M}_{1,0,1} = n^{-1} \sum_{i=1}^n \frac{n-i}{n-1} z_i$ .

Hosking y Wallis [55] analizaron las propiedades de los estimadores MV, MOM y MPP específicamente para la distribución de Pareto Generalizada.

Mostraron también que MPP es una alternativa al estimador de momentos ordinario con ventaja para pequeñas muestras pero con baja eficiencia asintótica.

#### 4.2.5. Verosimilitud-Momentos (VM)

Zhang [121] propone el estimador de Verosimilitud-Momentos (VM) combinando Mom y MV. La solución para esta combinación de ecuaciones siempre

existe, es eficiente y asintóticamente normal.

La expresión del momento de orden  $r$  de  $(Z + \gamma)$  es

$$E(Z + \gamma)^r = \alpha \gamma^r (r + \alpha)^{-1}, \quad r < -\alpha. \quad (4.3)$$

El estimador VM  $(\hat{\alpha}_{VM}, \hat{\gamma}_{VM})$  se define como la solución del sistema de ecuaciones (4.4):

$$\begin{aligned} \frac{1}{n} \sum_{i=1}^n \left(1 + \frac{Z_i}{\hat{\gamma}_{MV}}\right)^p - (1 - r)^{-1} &= 0, \\ \hat{\alpha}_{VM} &= \frac{-n}{\sum_{i=1}^n \left(1 + \frac{Z_i}{\hat{\gamma}}\right)}, \end{aligned} \quad (4.4)$$

siendo  $p = \frac{rn}{\sum_{i=1}^n \left(1 + \frac{Z_i}{\gamma}\right)}$ ,  $r < -\alpha$ .

Cabe destacar que VM es más eficiente que MPP y que Mom, el valor de su eficiencia asintótica es cercano a 1. Además si se estima el parámetro de escala y luego con ese valor se estima el parámetro de forma, repitiendo ese procedimiento varias veces, en caso de existir el estimador de máxima verosimilitud, el algoritmo del VM y MV convergen al mismo valor.

#### 4.2.6. Medianas (Med)

Cuando se pretende estudiar el comportamiento asintótico de un estimador a medida que su tamaño muestral tiende a infinito, es usual referirse a la consistencia, la distribución asintótica, el punto de ruptura y la función de influencia. En particular estas dos últimas características aluden a la robustez del estimador.

##### 4.2.6.1. Funciones de Influencia

Dada una muestra aleatoria  $\mathbf{z} = \{z_1, z_2, \dots, z_n\}$  constituida por  $n$  variables aleatorias independientes idénticamente distribuidas y un estimador  $\hat{\theta}_n = \hat{\theta}_n(\mathbf{z})$ .

---

**Definición 4.1.** Definimos el valor asintótico del estimador en  $F$  como:

$$\widehat{\theta}_n \xrightarrow{p} \widehat{\theta}_\infty(F) \quad (4.5)$$

Indicando  $\xrightarrow{p}$  la convergencia en probabilidad.

El interés se centra en el análisis del comportamiento de  $\widehat{\theta}_\infty(F)$  cuando  $F$  pertenece a un 'entorno de contaminación de  $F$ ', situación muy usual en el contexto de las imágenes SAR, como fue descrito en la sección 4.1.

**Definición 4.2.** Definimos entorno de contaminación de la distribución  $F$  a la familia de funciones:

$$\mathcal{F}, \epsilon = \{(1 - \epsilon)F + \epsilon G : G \in \mathcal{G}\} \quad (4.6)$$

La función de influencia, también llamada *curva de influencia* de un estimador, introducida por Hampel en [51], es una aproximación del comportamiento de  $\widehat{\theta}_\infty(F)$  contiene una pequeña porción  $\epsilon$  de datos atípicos idénticos.

**Definición 4.3.** Definimos curva o función de influencia de un estimador  $\widehat{\theta}$  para una función  $F$ :

$$IF(z_0, F) = \lim_{\epsilon \rightarrow 0^+} \frac{\widehat{\theta}_\infty [(1 - \epsilon)F + \epsilon\delta_{z_0}] - \theta_\infty(F)}{\epsilon}, \quad (4.7)$$

$$IF(z_0, F) = \frac{\partial \widehat{\theta}_\infty [(1 - \epsilon)F + \epsilon\delta_{z_0}]}{\partial \epsilon}, \quad (4.8)$$

donde  $\delta_{z_0}$  es la función masa puntual de  $z_0$ . Cuando  $\epsilon$  es pequeño el sesgo del estimador  $\widehat{\theta}_\infty [(1 - \epsilon)F + \epsilon\delta_{z_0}]$  se puede aproximar por la expresión  $\epsilon IF_{\widehat{\theta}}(z_0, F)$ .

#### 4.2.6.2. El estimador

Peng and Welsh [85] proponen el estimador de medianas Median (Med) como una alternativa robusta, dado que su función de influencia es acotada. Este estimador es asintóticamente normal, sin embargo su desempeño no es

destacable. Por esta razón sólo se recomienda su uso en caso que sea notoria la presencia de outliers [57].

Designemos con  $\theta = (\alpha, \gamma)$  al vector de parámetros de interés. La función de scores se define como [85]:

$$\psi_{\theta}(z) = \frac{\partial \ln f_Z(z; \theta)}{\partial \theta}, \quad (4.9)$$

que, en nuestro caso, la expresión del vector de scores es:

$$\begin{aligned} \psi_{(\alpha, \gamma)}(z) &= (\psi_{(\alpha, \gamma)}^1(z), \psi_{(\alpha, \gamma)}^2(z)) \\ &= \left( \frac{1}{\alpha} - \ln \gamma + \ln(z + \gamma), -\frac{\alpha}{\gamma} + \frac{\alpha - 1}{z + \gamma} \right). \end{aligned} \quad (4.10)$$

Este método iguala la mediana muestral de  $\{\psi_{(\alpha, \gamma)}(z_i)\}$ ,  $i = 1, \dots, n$  con la mediana poblacional  $Q_{1/2}$ . Para hallar la mediana poblacional, calculamos la función de densidad correspondiente a las variables aleatorias auxiliares  $W_1 = \ln(Z + \gamma)$  y  $W_2 = (Z + \gamma)^{-1}$ , que son

$$f_{W_1}(w_1) = -\alpha \gamma^{-\alpha} e^{\alpha w_1} \mathbb{1}_{[\ln \gamma, +\infty)}(w_1), \quad \text{y} \quad (4.11)$$

$$f_{W_2}(w_2) = -\alpha \gamma^{-\alpha} w_2^{-\alpha-1} \mathbb{1}_{(0, \gamma^{-1})}(w_2). \quad (4.12)$$

Dado que cada una de las funciones componentes del vector de scores es una función monótona, la mediana de la función de scores resulta la especialización de esa función en la mediana muestral.

Por lo tanto, el estimador de la mediana puede ser obtenido como la solución del siguiente sistema de ecuaciones:

$$\begin{aligned} \psi_{(\hat{\alpha}_{\text{Med}}, \hat{\gamma}_{\text{Med}})}^1(q_{1/2}(z)) &= \ln \hat{\gamma}_{\text{Med}} - \frac{\ln 2}{\hat{\alpha}_{\text{Med}}}, \\ \psi_{(\hat{\alpha}_{\text{Med}}, \hat{\gamma}_{\text{Med}})}^2(q_{1/2}(z)) &= \frac{2^{1/\hat{\alpha}_{\text{Med}}}}{\hat{\gamma}_{\text{Med}}}. \end{aligned}$$

donde  $q_{1/2}$  denota la mediana muestral. No reportaremos resultados experimentales para este estimador, dado su pobre desempeño frente a los restantes.

---

#### 4.2.7. Mínima Divergencia de Densidad (mDD)

Este estimador constituye otra alternativa robusta dado que su función de influencia también está acotada. Está indexado por una constante positiva  $\omega$  que controla el compromiso entre eficiencia y robustez.

Cuando  $\omega$  crece, aumenta la robustez pero baja la eficiencia. En el caso particular que  $\omega = 0$  el estimador mDD maximiza la logverosimilitud y, por lo tanto, es el estimador MV. Es decir que el estimador MV puede pensarse como un caso particular del mDD.

Juarez y Schucany [57] probaron que el MV es más eficiente para datos no contaminados, pero el comportamiento de mDD es superior bajo ciertos esquemas de contaminación.

Este estimador, propuesto por Juárez y Schucany [57], se define como

$$(\hat{\alpha}_{\text{MPD}}, \hat{\gamma}_{\text{MPD}}) = \arg \min_{(\alpha, \gamma) \in \Theta} H_{\omega}(\theta, z_1, \dots, z_n),$$

donde

$$H_{\omega}(\theta) = \int_{\mathbb{R}} f^{1+\omega}(z, \theta) dz - \left(1 + \frac{1}{\omega}\right) \frac{1}{n} \sum_{i=1}^n f^{\omega}(z_i, \theta),$$

siendo  $f$  la función de densidad de probabilidad dada en la Ecuación (2.18) cuando  $\theta = (\alpha, \gamma)$  pertenece al espacio paramétrico  $\Theta$ .

En este trabajo utilizamos el valor sugerido en [57],  $\omega = 0,1$ .

#### 4.2.8. Máxima Bondad de Ajuste (MBA) con el estadístico de Anderson Darling Cola Derecha (ADD)

Luceño [71] propuso esta metodología de estimación y mostró su consistencia y eficiencia asintótica, así como su superioridad para el caso de datos provenientes de distribuciones con colas pesadas.

Los estimadores de máxima bondad de ajuste, también conocidos como de mínima distancia, pueden obtenerse minimizando alguno de los estadísticos

de prueba correspondientes a alguno de los test de bondad de ajuste basados en la distribución empírica.

Estos estadísticos cuantifican la distancia entre la función de distribución teórica  $F_Z(z)$  y la función de distribución empírica  $\widehat{F}_Z(z_1, \dots, z_n)$  basada en la muestra  $\mathbf{z} = (z_1, z_2, \dots, z_n)$ .

Entre los estadísticos de bondad de ajuste disponibles seleccionamos el de Anderson Darling con peso en la cola Derecha ADD y reportamos resultados experimentales para éste. La decisión se basó en que su comportamiento resultó superior en términos de convergencia, sesgo y error cuadrático medio al de Kolmogorov-Smirnov (KS), Cramer Von Mises (CM), Anderson Darling (AD), AD cola Izquierda (ADI), AD de segundo orden cola Derecha (AD2D), AD de segundo orden cola izquierda (AD2I) y AD de segundo orden (AD2).

La expresión del estadístico ADD es:

$$(\widehat{\alpha}_{\text{ADD}}, \widehat{\gamma}_{\text{ADD}}) = \arg \min_{(\alpha, \gamma) \in \Theta} G_{\text{ADD}}(\alpha, \gamma, z_1, \dots, z_n),$$

donde  $\Theta$  designa al espacio paramétrico y

$$G_{\text{ADD}}(\alpha, \gamma, z_1, \dots, z_n) = n \int_{\mathbb{R}} \frac{(F_Z(z) - \widehat{F}_Z(z_1, \dots, z_n))^2}{1 - F_Z(z)} dF_Z(z).$$

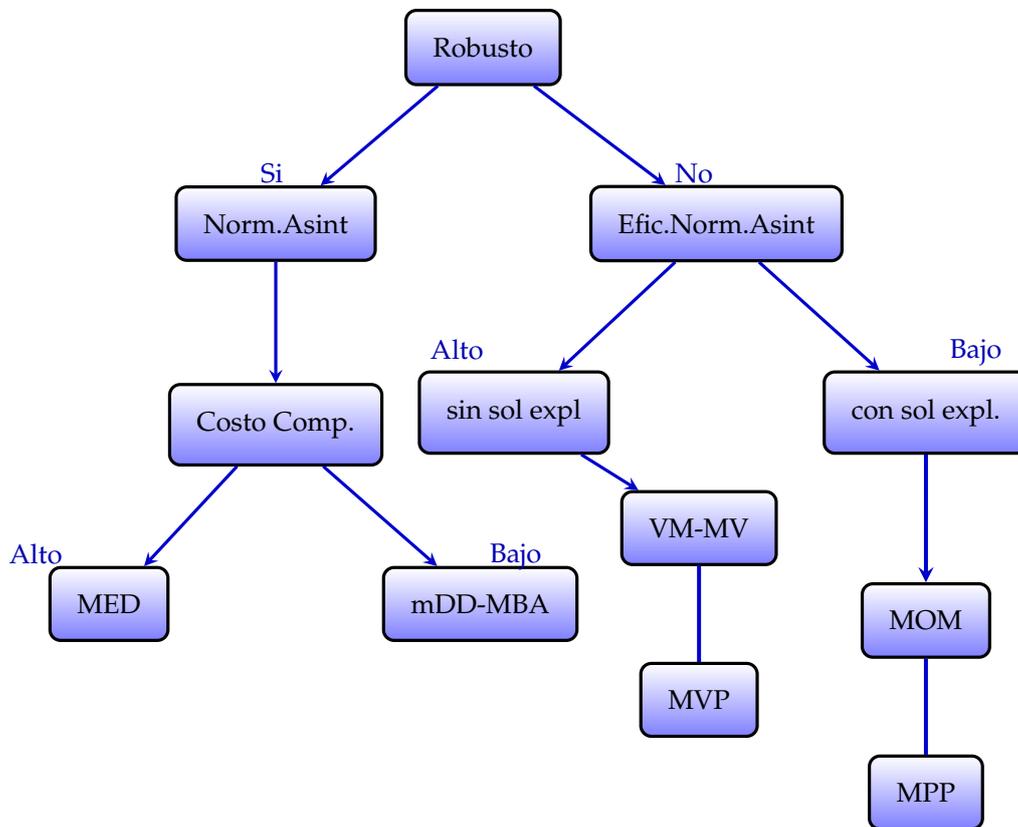
La Figura 4.3 de la página 99 resume las principales características de los métodos de estimación propuestos.

### 4.3. El Umbral

La distribución de Pareto Generalizada es una ley límite de las excedencias sobre un umbral [87]. De este modo, la selección del umbral es crucial para una estimación precisa.

Se evaluaron las estimaciones con cada uno de los estimadores presentados con los siguientes cinco métodos de selección de umbral.

**Figura 4.3:** Propiedades Estimadores: 'Norm.Asint' denota normalidad asintótica, 'Efic.Norm.Asint' eficiencia y normalidad asintótica, 'Costo Comp.' costo computacional y 'sol expl.' solución explícita.



- $u_0$  considera en la estimación a todos los valores muestrales.
- $u_{q_{10}}$  que usa los valores mayores, seleccionando 90 % de la muestra.
- $u_{q_{20}}$  que usa los valores mayores, seleccionando 80 % de la muestra.
- $u_{\text{Hill}}$  basado en el gráfico de Hill.

El gráfico de Hill representa las estimaciones realizadas con el estimador de Hill considerando proporciones decrecientes de la cola de la distribución. Se elige como umbral el primer cuantil empírico de la distribución a partir del cual se estabiliza el valor de la estimación.

- e)  $u_{AD}$  es un umbral seleccionado en forma automática basado en los  $p$ -valores de la prueba de AD de bondad de ajuste.

Con el objeto de decidir cuál es el umbral más adecuado para cada uno de los estimadores, se generaron 1000 muestras de tamaños 25, 49 y 81, para todas las combinaciones de parámetros  $\alpha \in \{-8, -5, -2\}$  y  $\gamma \in \{0, 1; 1; 10\}$ .

La conclusión de este análisis es que  $u_{q_{10}}$  para el tamaño muestral  $n = 49$ , es el mejor para los estimadores MV y mDD. En todos los restantes casos,  $u_0$  es el mejor.

## 4.4. Resultados del Estudio de Simulación

Si bien hemos presentado un conjunto amplio de estimadores, sólo trabajaremos en esta sección con aquellos que mostraron mejor desempeño en las experiencias preliminares.

Presentamos los resultados correspondientes a la comparación de los siguientes métodos de estimación MBA (ADD), MV, VM, mDD, MVP y MPP utilizando como criterios la tasa de convergencia, el costo computacional, el sesgo de la estimación y el error cuadrático medio (ECM) estimado. Esto ha sido replicado para datos sin contaminación y para datos contaminados.

### 4.4.1. Datos Puros

Se generaron 1000 muestras para cada uno de los tamaños  $\{25, 49, 81, 121, 500\}$  con datos distribuidos con  $\mathcal{G}_I^0(\alpha, \gamma, 1)$  combinando los siguientes valores de parámetros  $\alpha \in \{-8, -5, -2\}$  y  $\gamma \in \{0, 1, 1, 10, 100\}$ . Las muestras fueron obtenidas aplicando los resultados presentados en la Ref. [17]. Se consideran muestras pequeñas a aquellas cuyo tamaño es inferior a 30 y grandes a las restantes.

La Figura 4.4 página 109 muestra la tasa de convergencia para distintos grados de textura y para valores del parámetro de escala en  $\gamma = \{1, 10\}$ . Pa-

---

ra muestras pequeñas, la mejor tasa de convergencia se observó para ADD, sin embargo la tasa de convergencia para tamaños muestrales mayores resulta inaceptable.

Excepto el MVP, los restantes estimadores considerados alcanzan un buen nivel de convergencia en muestras grandes y razonablemente aceptable en muestras pequeñas. El estimador VM es el que logra mejor desempeño para valores bajos del parámetro de textura.

La Figura 4.5 de la página 110 muestra el comportamiento de los estimadores en cuanto al sesgo de estimación para diferentes grados de textura.

Se aprecia en esta figura que todos los métodos subestiman el parámetro de textura para valores bajos ( $\alpha \leq -5$ ) de textura y que lo sobreestiman para valores altos ( $\alpha > -5$ ). Un caso problemático de esta estimación lo constituyen los valores muy bajos de textura, como por ejemplo  $\alpha = -8$ . En este caso el mejor estimador, en el sentido de menor sesgo es el MVP, aún en el caso de pequeñas muestras.

MPP mejora su rendimiento a medida que disminuye el valor de la textura.

VM se destaca en algunos casos, mientras que todos los estimadores logran una precisión aceptable para muestras de tamaños superiores a 121.

La Figura 4.6 de la página 111 ilustra el comportamiento del ECM. En ella se aprecia que todos los estimadores tienen una conducta similar en este sentido. La falta de monotonía de algunas curvas es debida a la aleatoriedad y a la característica de colas pesadas del modelo.

La Figura 4.7 de la página 112 muestra el tiempo en milisegundos consumido por cada uno de los métodos de estimación para generar 1000 muestras de tamaño 500 para todas las combinaciones de parámetros detalladas.

La comparación a posteriori de Tukey HSD correspondientes a un análisis de la varianza señalan que el MPP es mejor que los restantes en términos del costo computacional.

#### 4.4.2. Datos Contaminados

La presencia de outliers puede causar severos sesgos en la estimación de los parámetros, por ello en la evaluación de un método de estimación es necesario probar su desempeño frente a ciertos esquemas de contaminación. En esta tesis analizamos el comportamiento de los estimadores de interés frente a situaciones de contaminación típicas de datos provenientes de imágenes SAR.

Consideremos una variable aleatoria Bernoulli  $B$  con probabilidad de éxito  $0 < \epsilon \ll 1$ .

El conjunto  $\mathcal{C}_{\mathcal{F}} = \{(1 - \epsilon)\mathcal{F} + \epsilon G : G \in \mathcal{G}\}$  se conoce como entorno de contaminación de la familia  $\mathcal{F}$  donde  $\mathcal{G}$  puede ser otra familia de distribuciones o bien una constante.

Focalizamos nuestro interés en una variable aleatoria  $Z$  contaminada de la siguiente manera:  $Z = BC + (1 - B)W$ , siendo  $C \in \mathbb{R}_+$  y  $W \sim \mathcal{G}_I^0(\alpha, \gamma, 1)$ .

La función empírica de influencia de un estimador (FEI) o curva de influencia empírica de un estimador  $T_n$  basado en una muestra aleatoria  $Z_1, Z_2, \dots, Z_{n-1}$  se define como:

$$\text{FEI}(z) = T_n(Z_1, Z_2, \dots, Z_{n-1}, z), \text{ con } z \in \mathbb{R}^+. \quad (4.13)$$

La idea consiste en evaluar la robustez de un estimador, midiendo el impacto sobre éste de la presencia de outliers. Nuestro modelo de outliers consiste en observar una pequeña porción de los datos  $\epsilon$  que no tiene la distribución general de la mayoría de los datos, sino que toman un valor  $C$  generalmente muy alto.

Una manera cualitativa de describir la robustez de los estimadores es por medio de la función de influencia empírica (EIF) [50]. La FEI (función empírica de influencia) muestra lo que sucede con un estimador  $\hat{\theta}$  cuando una observación  $z$  que se mueve sobre el soporte de la distribución es añadida a la muestra sobre la que se realiza la estimación.

---

Para evitar la influencia de la aleatoriedad de la muestra  $Z_1, Z_2, \dots, Z_{n-1}$ , característica del muestreo, se construye una 'muestra artificial' con los  $n - 1$  cuantiles de la distribución [3], es decir

$$Z_i = F_{\mathcal{G}_I^0(\alpha, \gamma, 1)}^{-1} \left( \frac{i - 1/3}{n + 1/3} \right), \text{ con } i \in \{1, 2, \dots, n - 1\}. \quad (4.14)$$

Esta curva es conocida como la función empírica de influencia estilizada (FEIS) y resulta un caso especial del entorno de contaminación.

En el caso de la FEIS se considera  $\epsilon = 1/n$  y  $C = z$  es una constante que recorre el soporte de la distribución.

Se construyó una curva de influencia empírica estilizada considerando muestras de tamaños  $n \in \{25, 49, 81, 121\}$ , para cada uno de los estimadores comparados y para todas las combinaciones de parámetros estudiadas.

Dado que el valor esperado para la función no contaminada es 25 el eje de abscisas recorre los valores [25; 1000]. En este caso tomamos  $\epsilon = (n + 1)^{-1}$ , siendo  $n$  el tamaño original de la muestra de cuantiles.

La Figura 4.8 de la página 113 muestra las curvas empíricas de influencia de los estimadores considerados para  $\alpha = -5$  y  $\gamma = 100$ .

La curva de influencia de un estimador robusto no debería alejarse notablemente del valor del parámetro que desea estimar. En la Figura 4.8 se aprecia que el estimador mDD presenta un buen desempeño frente a contaminaciones con valores altos, sin embargo en casi todos los casos es destacable la superioridad del comportamiento del MPP, debido a los leves apartamientos del verdadero valor del parámetro a estimar así como en la estabilidad de sus estimaciones.

### 4.4.3. Datos SAR Reales

Realizamos estimaciones sobre la imagen real que se muestra en la Figura 4.10(a). Esta es una imagen E-SAR de una sola vista banda-L y polarización

HH contaminada con un reflector de esquina (*corner reflector*). La Figura 4.10(b) de la página 114 muestra las subregiones utilizadas para la estimación del parámetro de textura.

Las estimaciones se presentan en la Figura 4.10 de la página (114) donde la línea horizontal negra punteada indica el valor considerado verdadero para el parámetro de textura que se tomó como el valor límite del estimador MV dadas sus propiedades asintóticas. El valor considerado verdadero es la estimación límite de MV sin contaminación.

Los valores pueden apreciarse en la Tabla 4.2.

Tanto en la Tabla 4.2 como en la Figura 4.10 de la página 115 se puede apreciar que tanto el MPP como el MVP son los que producen estimaciones más precisas en presencia de un corner reflector, especialmente en un caso de interés que son las pequeñas muestras de imágenes muy ruidosas.

El estimador MV también proporciona estimaciones precisas, para muestras grandes. En muestras de tamaño superior a 600, todas las estimaciones son cercanas al valor real, aunque el mDD resulta ser el más sesgado.

Todas las estimaciones tienden al valor verdadero hipotético, pero a velocidades y con direcciones muy disímiles; algunas lo subestiman (ADD y VM), mientras que otras lo sobreestiman (MV, mDD, MPP, MPP).

Además, cabe destacar que la tasa de convergencia y el comportamiento es bastante diferente en el caso de muestras pequeñas. La siguiente aplicación pone de relieve el impacto de tal diferencia en un proceso de clasificación.

El siguiente experimento fue proyectado con el objeto de ilustrar la importancia de disponer de diferentes estrategias de estimación:

1. Se construye una imagen SAR simulada utilizando datos con distribución  $\mathcal{G}_I^0(-5, 1, 1)$  y con distribución  $\mathcal{G}_I^0(-2, 1, 1)$  para la región central y la región borde respectivamente, como puede apreciarse en la Figura 4.12(a).
2. La imagen de las estimaciones se construye aplicando el método de MPP,

**Tabla 4.2:** Estimaciones del parámetro de textura para las regiones de la Figura 4.10(b).

$n$	ADD	VM	mDD	MV	MVP	MPP
30	-1,067	-1,572	-1,130	-1,366	-2,288	-2,141
88	-1,630	-1,496	-0,440	-0,504	-1,216	-1,308
154	-1,552	-1,945	-0,457	-0,488	-1,142	-1,196
266	-1,703	-1,588	-0,555	-0,546	-1,114	-1,127
368	-1,601	-1,350	-0,612	-0,594	-1,107	-1,109
532	-1,782	-1,536	-0,770	-0,702	-1,108	-1,094
744	-1,761	-1,588	-0,866	-0,782	-1,109	-1,097
1044	-1,617	-1,499	-0,920	-0,845	-1,110	-1,106
1386	-1,636	-1,409	-1,037	-0,933	-1,121	-1,114
1692	-1,537	-1,589	-1,127	-1,055	-1,138	-1,124

realizando luego sobre estas estimaciones del parámetro de textura una clasificación utilizando el método supervisado de la máquina de soporte vectorial (SVM) con núcleos lineales. La Figura 4.12(b) muestra el resultado de estos dos procedimientos combinados.

3. Se agrega un outlier a la imagen original y se vuelve a realizar el procedimiento descrito en el inciso anterior. Los resultados de la aplicación de ambos procedimientos en presencia del outlier se muestran en la Figura 4.12(c).
4. Clasificamos la misma imagen con el outlier pero utilizando el método de estimación mDD. Luego aplicamos el método de clasificación supervisada SVM y los resultados de estos procedimientos se presentan en la Figura 4.12(d).

Se puede apreciar que la eficacia del método de estimación de parámetros

depende en gran medida de la estructura de la imagen. Se observa que en ausencia del reflector de esquina, el método MPP proporciona buenos resultados de clasificación, pero cuando el reflector de esquina está presente, es recomendable aplicar otro método. La precisión de la clasificación para cada caso se graficó en la Figura 4.12 de la página 117.

La Figura 4.12 muestra la precisión de la clasificación realizada con cada método de estimación, con y sin reflector de esquina.

A pesar de que los otros estimadores superan al MPD para los datos no contaminados, en el caso de contaminación con reflector de esquina, el método MPD opera con mayor precisión como puede apreciarse en la Figura 4.11 de la página 116.

La Figura 4.13 de la página 118 muestra el resultado de la aplicación de una segmentación de dos etapas a la imagen de la Figura 4.10(a). En la primera etapa se realiza una estimación del parámetro de textura con los métodos VM y MV utilizando ventanas de  $21 \times 21$  píxeles, en la segunda etapa se realiza una clasificación no supervisada mediante la metodología no jerárquica de  $k$ -medias considerando el número de clases  $k = 2, 3$ . Se puede apreciar que la segmentación obtenida usando el estimador VM coincide con lo observado en la Figura 4.10(a) de la página (114). Esto es debido a la robustez del procedimiento frente a la presencia de un reflector de esquina.

## 4.5. Discusión y Conclusiones

En este capítulo, hemos comparado seis métodos de estimación para el parámetro de textura de la distribución  $\mathcal{G}_I^0$  en contexto de datos puros y de datos contaminados.

Para los datos puros se aprecia que MVP presenta problemas de convergencia para muestras chicas en cualquiera de las combinaciones de valores de parámetros consideradas. ADD tiene un comportamiento errático y VM es el

---

que presenta mejor tasa de convergencia en muestras chicas.

ADD tiene la menor tasa de convergencia para muestras grandes pero alta tasa en el caso de muestras chicas. El sesgo y el ECM son aceptables para muestras de tamaños superiores a 121 en el caso de datos no contaminados.

La comparación entre los estimadores fue realizada teniendo en cuenta criterios de convergencia, sesgo, error cuadrático medio, costo computaciones y propiedades asintóticas deseable.

Las observaciones realizadas para datos no contaminados, en general no son válidas para el caso de contaminaciones y específicamente con corner reflector.

Para los casos de contaminación es aconsejable construir la curva de influencia empírica y realizar observaciones sobre la misma.

La función de influencia pone en evidencia que las contaminaciones con valores bajos del parámetro  $C$  pueden causar errores groseros en las estimaciones.

La curva de influencia empírica estilizada muestra que el mDD logra estimaciones más precisas para todos los tamaños muestrales cuando un único píxel está contaminado.

Los métodos de MVP y VM son superiores a los restantes frente a este tipo de contaminación, excepto por mDD que tiene muy buen desempeño especialmente para outliers altos.

ADD produce estimaciones precisas para muestras de tamaños superiores a 100 en presencia de outliers.

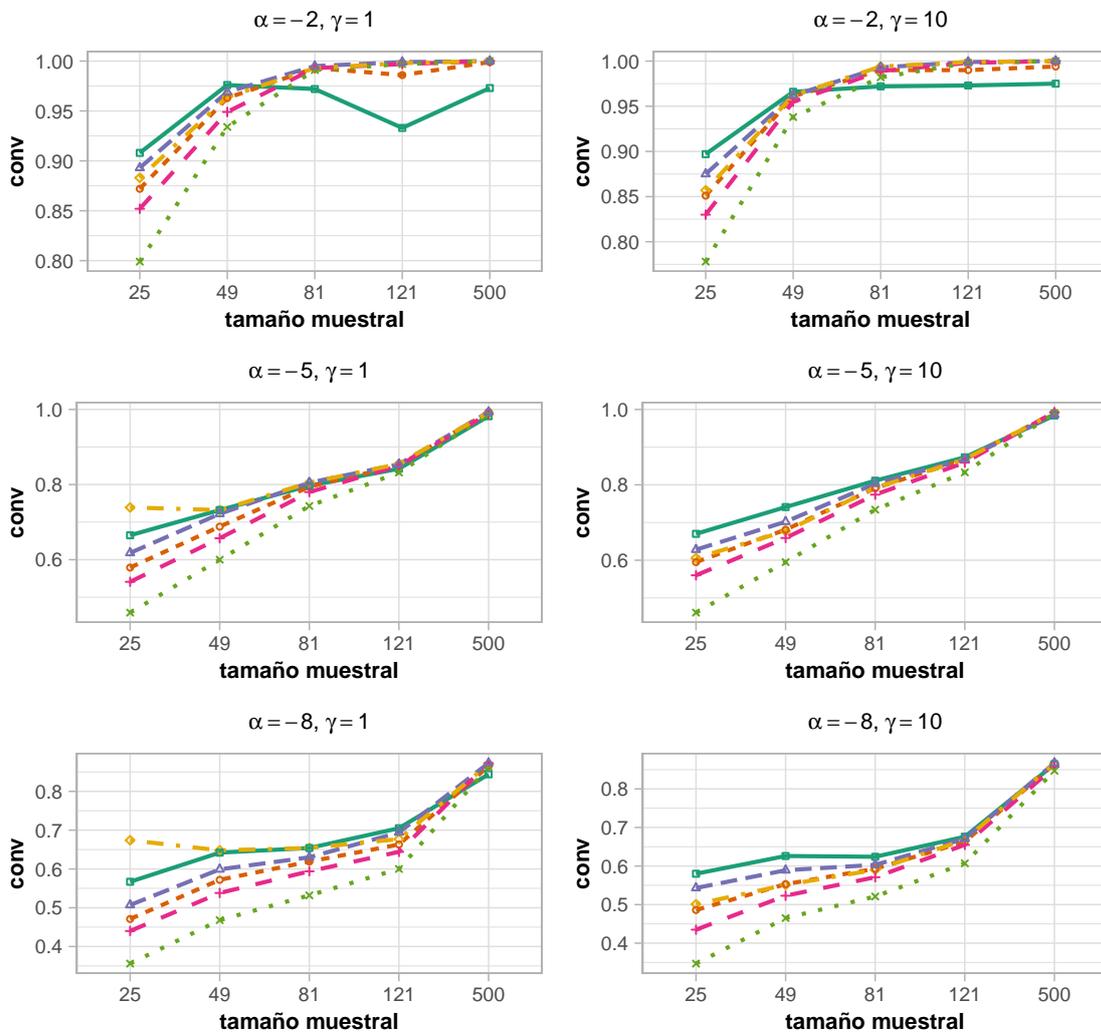
MV y mDD convergen en el caso de muestras grandes, sin embargo presentan sesgos significativos para el caso de muestras chicas. En este último caso el mejor desempeño se ha registrado para MPP, MVP y VM. Cabe señalar también que MPP es el método que consume el menor tiempo computacional mientras que ADD es el método que evidencia mayor costo computacional.

Teniendo en cuenta todas estas consideraciones, se recomienda el uso de

mDD o MPP para este tipo de datos.

Los resultados de esta comparación de estimadores en casos de datos  $\mathcal{G}_I^0$  considerando datos puros y también datos contaminados con reflectores de esquina, así como las aplicaciones a segmentación de imágenes se han publicado en [18].

Figura 4.4: Tasa de Convergencia por textura



Estimador    —■— ADD    -▲- MPP    ···· MVP  
                  -○- mDD    -+ - MV    -◇- VM

Figura 4.5: Sesgo por en función del tamaño muestral y la textura

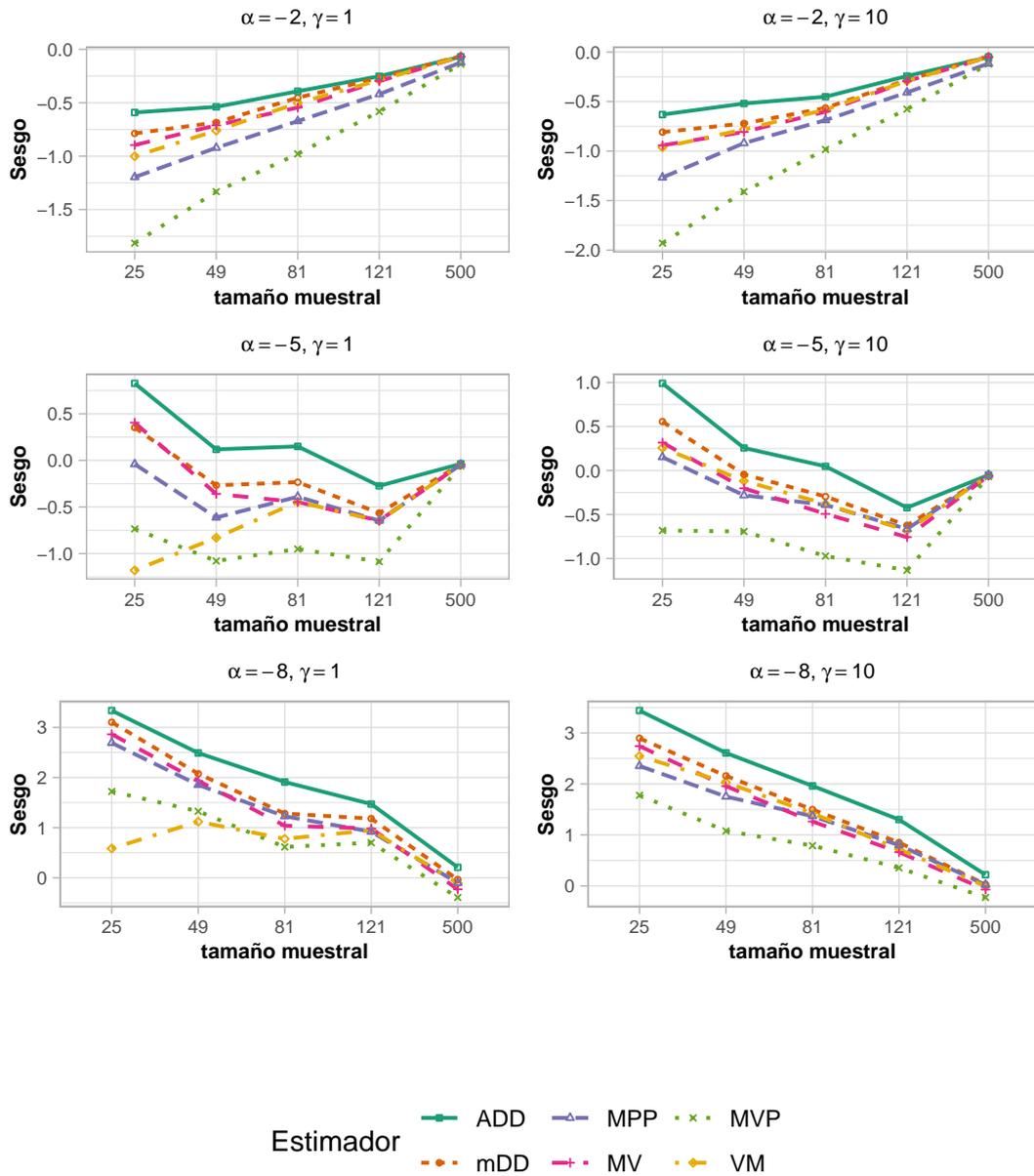
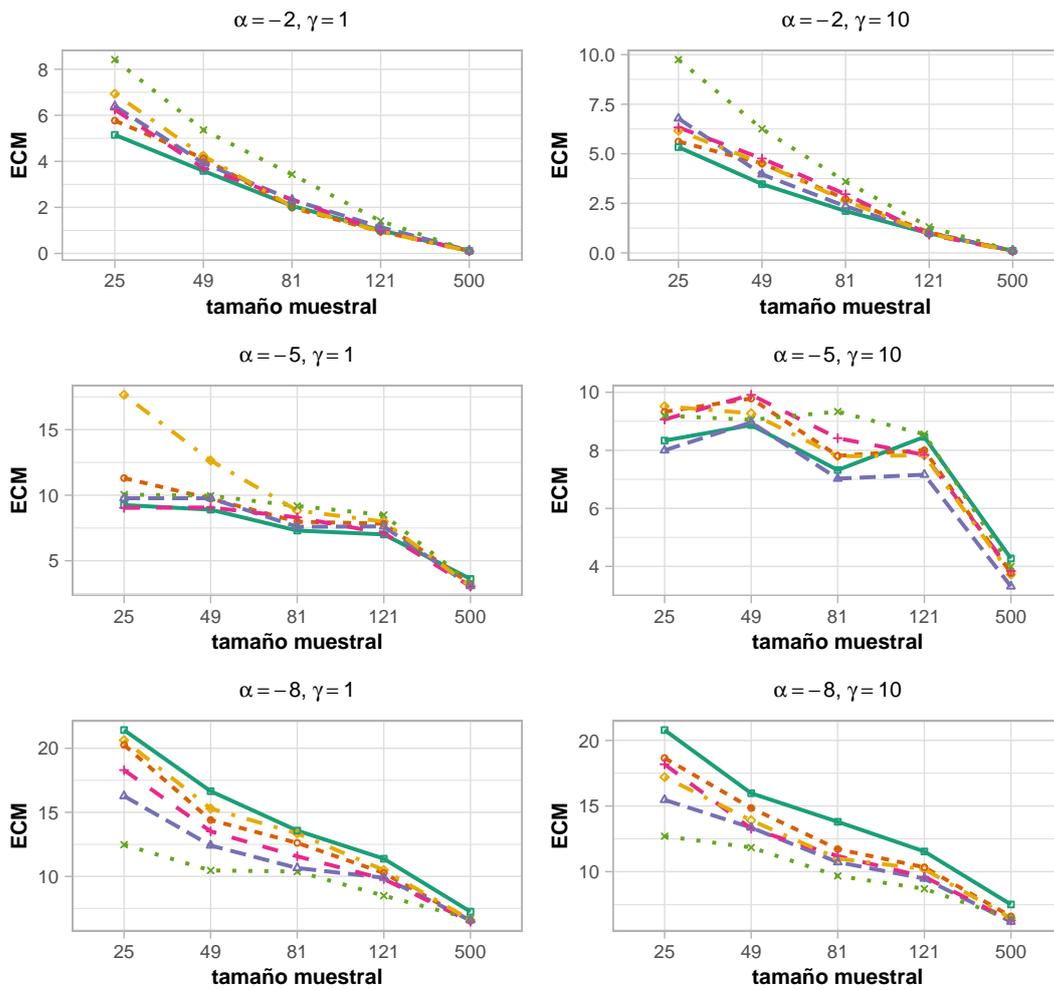
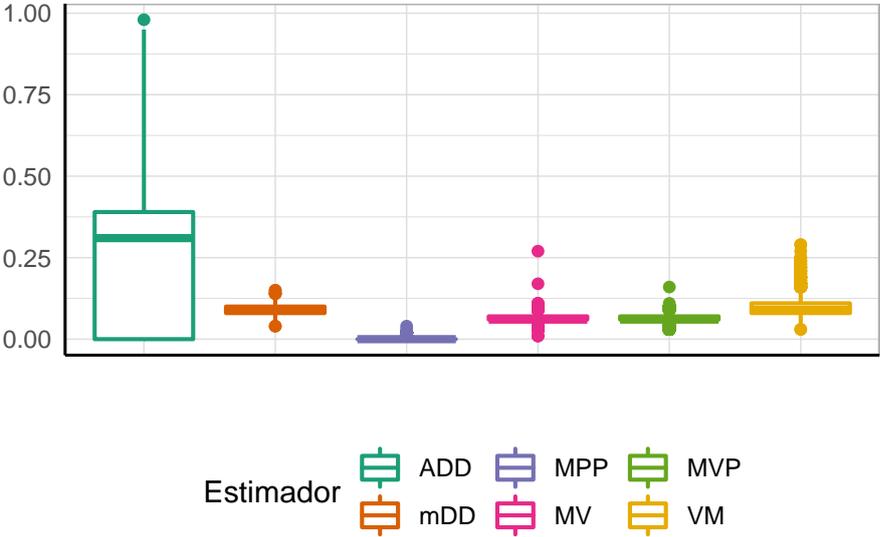


Figura 4.6: ECM por textura y tamaño muestral

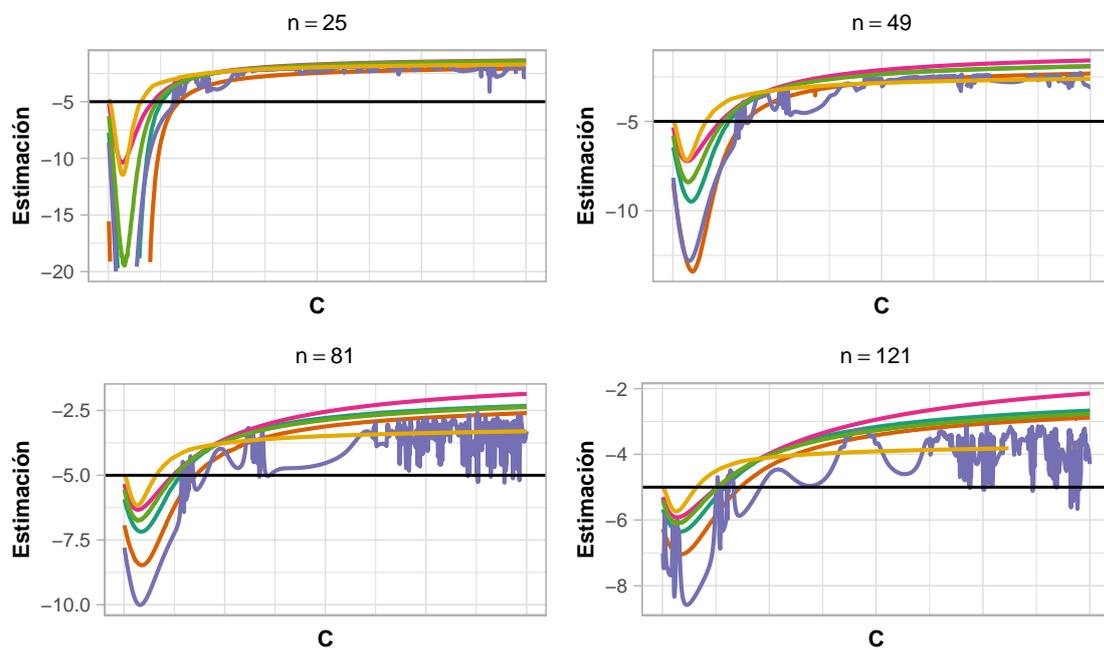


Estimador    —■— ADD    -▲- MPP    ···· MVP  
                  -○- mDD    -+·- MV    -◇- VM

Figura 4.7: Tiempo de Procesamiento por Estimador

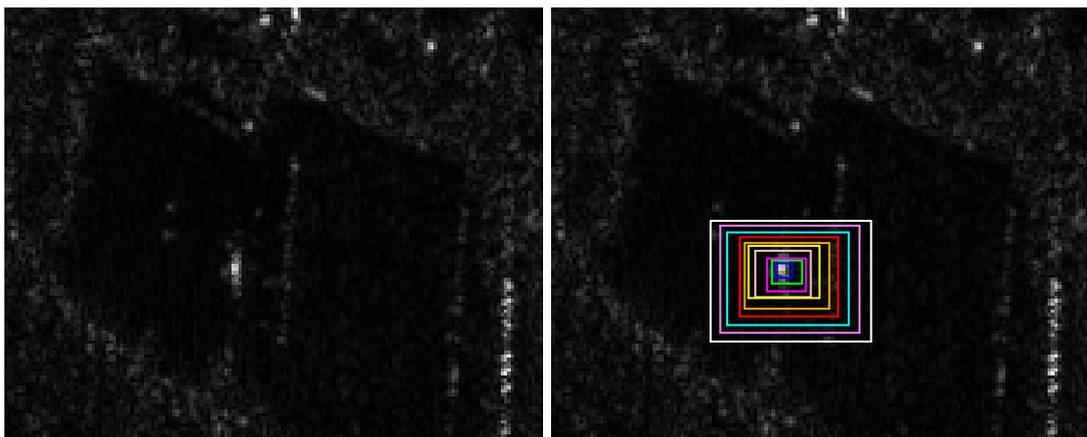


**Figura 4.8:** Función Empírica de Influencia Estilizada – FEIS, para  $\alpha = -5, \gamma = 100$ .



Estimador    — MV    — mDD    — VM  
                  — MVP    — MPP    — ADD

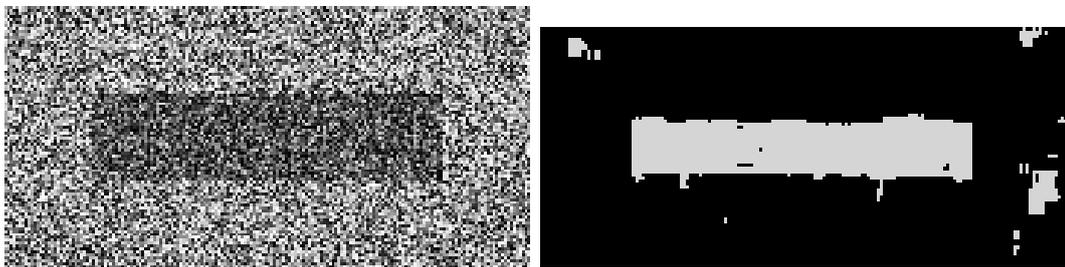
**Figura 4.9:** Imagen E-SAR de una sola vista utilizada para estimar el parámetro de textura(4.10(a)). Diez regiones centradas en el reflector de esquina de diferente tamaño seleccionadas para las estimaciones(4.10(b)).



(a) Imagen E-SAR de una vista contaminada con un reflector de esquina. (b) Regiones de interés utilizadas para calcular las estimaciones del parámetro de textura.

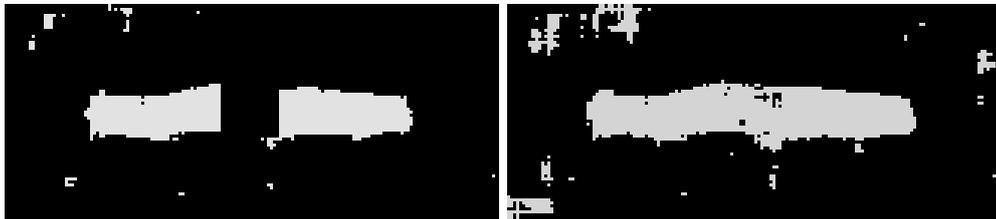


**Figura 4.11:** Clasificación de imagen por SVM con y sin corner reflector.



(a) Imagen Sintética con dos regiones.

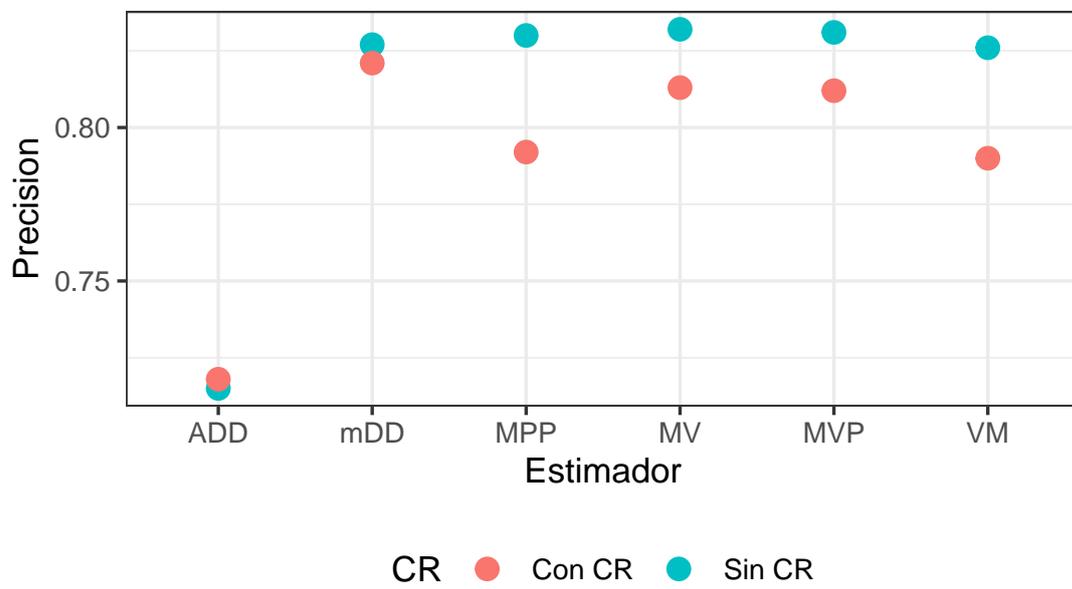
(b) Clasificación con SVM usando estimador MPP, precisión: 0,82.



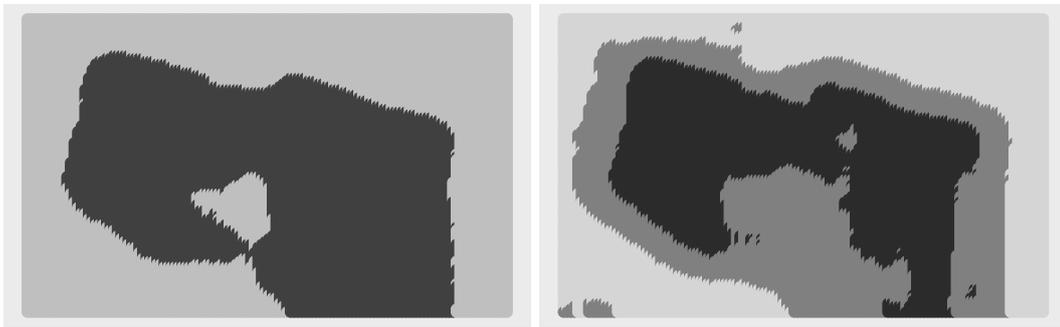
(c) Clasificación con SVM usando estimador MPP con corner reflector, precisión: 0,77.

(d) Clasificación con SVM usando estimador MPD con corner reflector, precisión: 0,81.

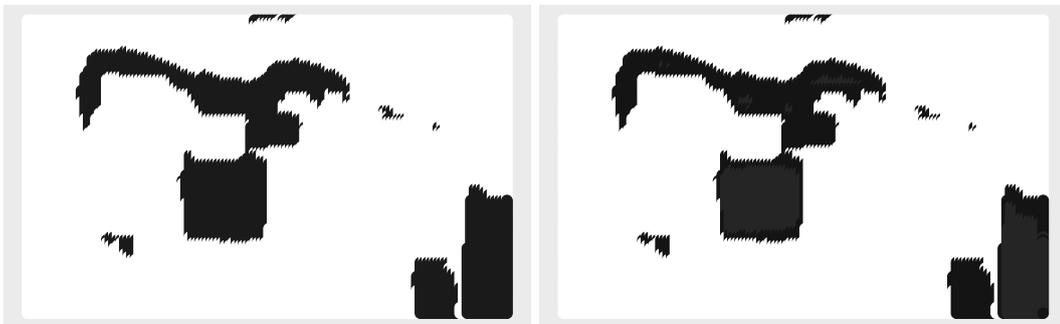
**Figura 4.12:** Precisión de la clasificación para cada método, con y sin corner reflector.



**Figura 4.13:** Segmentación usando la estimación del parámetro  $\alpha$  y el método  $k$ -means.



(a)  $k$  - means clasificación en dos regiones con el método VM. (b)  $k$  - means clasificación en tres regiones con el método VM.



(c)  $k$  - means clasificación en dos regiones con el método MV.. (d)  $k$  - means clasificación en tres regiones con el método MV..

## Capítulo 5

# Entropía de la distribución $\mathcal{G}_I^0$

### 5.1. Entropías $h$ - $\phi$

El concepto entropía se utiliza en diversos campos del conocimiento entre ellos en teoría de la información, dado que los conceptos de información y entropía están íntimamente relacionados. En todos los casos, la entropía cuantifica la cantidad de información de una fuente. Debido a ello es concebida como una 'medida del desorden' o de la 'particularidad de ciertos patrones'. Asimismo puede ser considerada como una medida de la incertidumbre. La relación entre frecuencia e información es que a menor probabilidad de ocurrencia, mayor información. A pesar de que la entropía fue definida originalmente para el caso discreto, es posible extender esta media para variables continuas. En esos casos se la suele llamar 'entropía diferencial'.

**Definición 5.1.** Dada una variable aleatoria continua  $Z$  soporte en  $\Omega$  y vector paramétrico  $\theta \in \Theta \subseteq \mathbb{R}^p$ , siendo su función de densidad de probabilidad  $f_Z(z | \theta)$  se define la entropía  $H_\phi^h(f)$  mediante la siguiente ecuación:

$$H_\phi^h(f) = h \left[ \int_{\Omega} \phi(f(x/\theta)) dx \right], \quad (5.1)$$

siendo  $\phi : [0, +\infty) \rightarrow \mathbb{R}$  una función cóncava y  $h : \mathbb{R} \rightarrow \mathbb{R}$  una función monótona creciente o bien  $\phi : [0, +\infty) \rightarrow \mathbb{R}$  una función convexa y  $h : \mathbb{R} \rightarrow \mathbb{R}$  una función decreciente [98].

Si bien se han propuesto varias alternativas para las funciones  $h$  y  $\phi$  en este trabajo utilizamos las entropías de Shannon y de Rényi que son casos particulares de las entropías  $h$ - $\phi$ . Las funciones  $h$  y  $\phi$  correspondientes a esos dos casos son:

- Entropía de Shannon:  $h(x) = x$  y  $\phi(x) = -x \ln(x)$ .
- Entropía Rényi:  $h(x) = \frac{1}{1-\beta} \ln x$  y  $\phi(x) = x^\beta$ ,  $\beta \in (0, 1)$ .

### 5.1.1. Entropía de Shannon para la distribución $\mathcal{G}_I^0$ Single Look

En esta sección calculamos la entropía de Shannon para la distribución  $\mathcal{G}_I^0(\alpha, \gamma, 1)$ .

La entropía de Shannon de una distribución con densidad de probabilidad  $f$  y parámetros  $\theta = (\theta_1, \dots, \theta_n)$  está dada por la siguiente ecuación:

$$H_S(f, \theta) = - \int_{\Omega} f(z, \theta) \ln [f(z, \theta)] dz. \quad (5.2)$$

siendo  $\Omega$  el soporte de la distribución de probabilidad  $f$ .

Entonces para nuestro caso de interés, considerando  $\theta = (\alpha, \gamma)$  y  $L = 1$ :

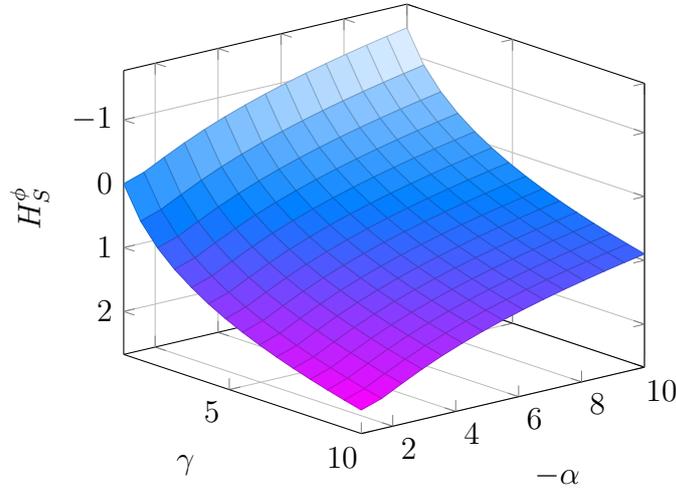
$$H_S(f_{\mathcal{G}_I^0}(z/\alpha, \gamma, 1)) = - \int_0^{+\infty} -\frac{\alpha}{\gamma} \left(1 + \frac{z}{\gamma}\right)^{\alpha-1} \ln \left(-\frac{\alpha}{\gamma} \left(1 + \frac{z}{\gamma}\right)^{\alpha-1}\right) dz \quad (5.3)$$

$$= \frac{\alpha-1}{\alpha} - \ln \left(-\frac{\alpha}{\gamma}\right). \quad (5.4)$$

La Figura 5.1 muestra un gráfico de la entropía de Shannon para un amplio rango de combinaciones de parámetros  $\alpha$  y  $\gamma$  de la distribución  $\mathcal{G}_I^0$  con  $L = 1$ . Estas combinaciones de valores de los parámetros cubren el producto cartesiano de los intervalos  $(\alpha, -\gamma) \in [-10, \dots, 1] \times [-10, \dots, -1]$ .

En la tabla 5.1, se presentan los valores de la entropía de Shannon para combinaciones de los parámetros  $\alpha$  y  $\gamma$  dentro del rango especificado. Puede observarse que el valor de la entropía, para un valor de  $\gamma$  fijo, es una función creciente en  $\alpha$ , mientras que para  $\alpha$  fijo, resulta creciente en  $\gamma$ .

**Figura 5.1:** Gráfico de la entropía de Shannon en función de  $\alpha$  y  $\gamma$ .



### 5.1.2. Entropía de Rényi para la distribución $\mathcal{G}_I^0$ Single Look

**Definición 5.2.** La entropía de Rényi depende de un parámetro  $\beta$ , para una función de densidad de probabilidad  $f$  con soporte  $\Omega$  está dada por:

$$H_R^\beta(f, \theta) = \frac{1}{1-\beta} \ln \int_{\Omega} [f(x)]^\beta d\mu(x), \quad \beta \in (0, 1). \quad (5.5)$$

Cuando  $\beta \rightarrow 1$  la entropía de Rényi tiende a la entropía de Shannon.

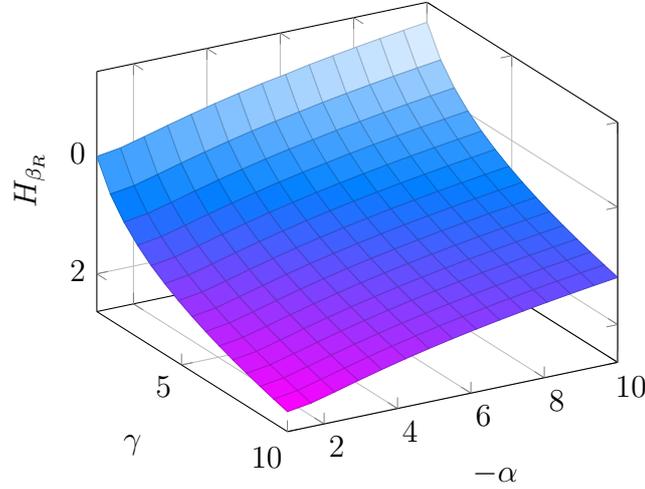
La expresión de la entropía de Rényi para para el caso Single Look de la distribución  $\mathcal{G}_I^0(\alpha, \gamma, 1)$  es:

$$H_R^\beta(\mathcal{G}_I^0(z/\alpha, \gamma, 1)) = \frac{\beta}{1-\beta} \ln \left( -\frac{\alpha}{\gamma} \right) + \frac{1}{1-\beta} \ln \left( -\frac{\gamma}{\beta(\alpha-1)+1} \right) \quad (5.6)$$

siendo  $\frac{1}{2} < \beta < 1$  y  $\alpha < -1$  para que quede definida la entropía de Rényi.

La Figura 5.2 muestra la variación de la entropía de Rényi para algunas combinaciones de valores de los parámetros  $\alpha$  y  $\gamma$  de la distribución  $\mathcal{G}_I^0$  con  $L = 1$ , con  $\beta = 0,5$ . Estas combinaciones de valores de los parámetros cubren el producto cartesiano de los intervalos  $(\alpha, -\gamma) \in [-10, \dots, 1] \times [-10, \dots, -1]$ . Se aprecia en la Figura 5.2 que para valores fijos de  $\alpha$  la entropía de Rényi es una función creciente en  $\gamma$  y para valores fijos de  $\gamma$  es creciente como función de  $\alpha$ .

**Figura 5.2:** Gráfico de la entropía de Rényi, con  $\beta = 0,5$  en función de  $\alpha$  y  $\gamma$



## 5.2. Distribución asintótica de las Entropías $h$ - $\phi$

Para las entropías de Shannon y Rényi, vale el siguiente resultado [98].

**Lema 1.** Dada una muestra aleatoria simple  $z = \{Z_1, Z_2, \dots, Z_n\}$  tamaño  $N$  es decir, variables aleatorias independientes con idéntica distribución  $f_\theta(Z)$ , donde  $\theta = (\theta_1, \theta_2, \dots, \theta_p)$  es el vector de parámetros y  $\hat{\theta} = (\hat{\theta}_1, \hat{\theta}_2, \dots, \hat{\theta}_p)$  el correspondiente estimador de máxima verosimilitud, entonces:

$$\sqrt{N} \left( H_h^\phi(\hat{\theta}) - H_h^\phi(\theta) \right) \xrightarrow[N \rightarrow \infty]{D} \mathcal{N}(0, \sigma_H^2), \quad (5.7)$$

siendo  $\sigma_H^2(\theta) = \delta^T K(\theta)^{-1} \delta$ , donde  $K(\theta)$  es la matriz de información de Fisher de  $f_\theta(Z)$  y  $\delta$  es el vector de scores de la entropía dado por:

$$\delta = [\delta_1, \delta_2, \dots, \delta_p] \quad \text{con} \quad \delta_i = \frac{\partial H_h^\phi(\theta)}{\partial \theta_i}. \quad (5.8)$$

La distribución asintótica de las entropías de Shannon y Rényi para la distribución  $\mathcal{G}_I^0$ , depende de la matriz de información de Fisher de la distribución, por ello debemos hallar su expresión.

---

**Definición 5.3.** La matriz de información de Fisher para una función de densidad de probabilidad  $f$  se define como:

$$K(\theta) = E \left( -\frac{\partial^2 \ln f(Z, \theta)}{\partial \theta^2} \right). \quad (5.9)$$

Considerando la Eq. (2.20), se tiene que

$$\ln(f(z, \alpha, \gamma)) = \ln(-\alpha) - \ln \gamma + (\alpha - 1) \ln \left( 1 + \frac{z}{\gamma} \right), \quad (5.10)$$

las derivadas de segundo orden de la expresión (5.10) son:

$$\frac{\partial^2 \ln f}{\partial \alpha^2} = -\frac{1}{\alpha^2}, \quad (5.11)$$

$$\frac{\partial^2 \ln f}{\partial \alpha \partial \gamma} = -\frac{Z}{\gamma^2 + Z\gamma}, \quad (5.12)$$

$$\frac{\partial^2 \ln f}{\partial \gamma^2} = \frac{1}{\gamma^2} + \frac{\alpha - 1}{\gamma^2} \left[ -\frac{Z^2}{(\gamma + Z)^2} + \frac{2Z}{Z + \gamma} \right]. \quad (5.13)$$

luego la expresión de la matriz de Información para la distribución  $\mathcal{G}_I^0$  es:

$$K(\theta) = E \left( \begin{array}{cc} \frac{1}{\alpha^2} & \frac{Z}{\gamma(Z + \gamma)} \\ \frac{Z}{\gamma(Z + \gamma)} & -\frac{1}{\gamma^2} - \frac{\alpha - 1}{\gamma^2} \left[ -\frac{Z^2}{(\gamma + z)^2} + \frac{2Z}{Z + \gamma} \right] \end{array} \right) \quad (5.14)$$

$$K(\theta) = \left( \begin{array}{cc} \frac{1}{\alpha^2} & \frac{1}{\gamma(1 - \alpha)} \\ \frac{1}{\gamma(1 - \alpha)} & \frac{1}{\gamma^2(\alpha - 2)} \end{array} \right). \quad (5.15)$$

De acuerdo a la Ecuación 5.8, en el caso de la entropía de Shannon para la distribución  $\mathcal{G}_I^0(\alpha, \gamma, 1)$  el vector de scores tiene la forma

$$\delta_S^T = [\delta_1, \delta_2] = \left[ \frac{\partial H_S(\alpha, \gamma)}{\partial \alpha}, \frac{\partial H_S(\alpha, \gamma)}{\partial \gamma} \right], \quad (5.16)$$

y por lo tanto

$$\delta_S^T = \left[ \frac{1}{\alpha^2} - \frac{1}{\alpha}, \frac{1}{\gamma} \right]. \quad (5.17)$$

Entonces, la forma cuadrática que define la varianza asintótica de la entropía de Shannon es:

$$\sigma_{H_s}^2 = \left( \frac{1}{\alpha^2} - \frac{1}{\alpha}; \frac{1}{\gamma} \right) \begin{pmatrix} \frac{1}{\alpha^2} & \frac{1}{\gamma(1-\alpha)} \\ \frac{1}{\gamma(1-\alpha)} & \frac{1}{\gamma^2(\alpha-2)} \end{pmatrix}^{-1} \begin{pmatrix} \frac{1}{\alpha^2} - \frac{1}{\alpha} \\ \frac{1}{\gamma} \end{pmatrix}. \quad (5.18)$$

La expresión de la inversa de la matriz de información de Fisher de la  $\mathcal{G}_I^0$  es:

$$[K(\theta)]^{-1} = \begin{pmatrix} (\alpha-1)^2\alpha^2 & \alpha\gamma(\alpha-1)(\alpha-2) \\ \alpha\gamma(\alpha-1)(\alpha-2) & \gamma^2(\alpha-1)^2(\alpha-2)/\alpha \end{pmatrix}. \quad (5.19)$$

Luego, la expresión de la varianza asintótica de la entropía de Shannon resulta

$$\sigma_{H_s}^2 = \frac{(\alpha-1)^2 [(-2 + 4\alpha^3 - \gamma^2 - \alpha^2(10 + \gamma^2) + \alpha(5 + 2\gamma^2))]}{\alpha^2 [(\alpha(5 + 2\gamma^2) - 2 - \gamma^2 - \alpha^2(2 + \gamma^2))]} \quad (5.20)$$

La forma cuadrática que define la varianza asintótica de la entropía de Rényi para esta distribución es:

$$\delta_R^T = \left[ \frac{\partial H_R(\alpha, \gamma)}{\partial \alpha}, \frac{\partial H_R(\alpha, \gamma)}{\partial \gamma} \right] = \left[ \frac{\beta}{1-\beta} \left( \frac{1}{\alpha} - \frac{1}{\beta(\alpha-1)+1} \right), \frac{1}{\gamma} \right]. \quad (5.21)$$

La varianza asintótica de la entropía de Rényi de esta distribución, considerando los desarrollos previos resulta:

$$\sigma_{H_R^\beta}^2 = \begin{pmatrix} \frac{\beta}{1-\beta}\Omega \\ \frac{1}{\gamma} \end{pmatrix} \begin{pmatrix} \frac{1}{\alpha^2} & \frac{1}{\gamma(\alpha-1)} \\ \frac{1}{\gamma(\alpha-1)} & \frac{1}{\gamma^2 - (\alpha-2)} \end{pmatrix}^{-1} \begin{pmatrix} \frac{\beta}{1-\beta}\Omega \\ \frac{1}{\gamma} \end{pmatrix}, \quad (5.22)$$

siendo  $\Omega = \left( \frac{1}{\alpha} - \frac{1}{\beta(\alpha-1)+1} \right)$ .

Operando:

$$\sigma_{H_R^\beta}^2 = \frac{(\alpha-1)^2(-2 + 4\beta + 4\alpha^3\beta^2 - \beta^2(4 + \gamma^2))}{(1 + (\alpha-1)\beta)^2(\alpha(5 + 2\gamma^2) - \alpha^2(2 + \gamma^2) - \gamma^2 - 2)} + \frac{\alpha(1 - 10\beta + 2\beta^2(7 + \gamma^2 - \alpha^2\beta(-4 + \beta(14 + \gamma^2))))}{(1 + (\alpha-1)\beta)^2(\alpha(5 + 2\gamma^2) - \alpha^2(2 + \gamma^2) - \gamma^2 - 2)}.$$

---

### 5.3. Test de la distancia cuadrática entre $h$ - $\phi$ entropías

Dadas dos muestras  $\mathbf{x} = (x_1, \dots, x_{N_1})$  e  $\mathbf{y} = (y_1, \dots, y_{N_2})$  con distribución  $\mathcal{G}_I^0$ , con parámetros  $\theta_1 = (\alpha_1, \gamma_1)$  y  $\theta_2 = (\alpha_2, \gamma_2)$ , respectivamente. Queremos testear si ambas muestras pertenecen a la misma distribución, o lo que es lo mismo si  $\theta_1 = \theta_2$ .

El planteo de las hipótesis correspondientes a esta prueba es:

$$H_0 : H_{\mathcal{M}}(\theta_1) = H_{\mathcal{M}}(\theta_2) = v$$

versus

$$H_1 : H_{\mathcal{M}}(\theta_1) \neq H_{\mathcal{M}}(\theta_2), \mathcal{M} \in \{S, R\}$$

donde  $S$  y  $R$  indican las entropías de Shannon y de Rényi, respectivamente.

Sea  $\hat{\theta}_i$  el estimador de máxima verosimilitud de  $\theta_i$  basado en una muestra de tamaño  $N_i$ . Siendo  $r$  la cantidad de poblaciones comparadas, con  $r \geq 2$  del Lema 1, sabemos que:

$$\sum_{i=1}^r \frac{N_i (H_{\mathcal{M}}(\hat{\theta}_i) - \bar{v})^2}{\sigma_{\mathcal{M}}^2(\hat{\theta}_i)} \xrightarrow[N \rightarrow \infty]{D} \chi_{r-1}^2 \quad (5.23)$$

siendo

$$\bar{v} = \left[ \sum_{i=1}^r \frac{N_i}{\sigma_{\mathcal{M}}^2(\hat{\theta}_i)} \right]^{-1} \sum_{i=1}^r \frac{N_i H_{\mathcal{M}}(\hat{\theta}_i)}{\sigma_{\mathcal{M}}^2(\hat{\theta}_i)} \quad (5.24)$$

Para nuestro caso de interés, es decir  $r = 2$  y  $N_1 = N_2 = N$ , el estadístico del test es:

$$S_{\mathcal{M}}(\theta_1, \theta_2) = N \frac{(H_{\mathcal{M}}(\hat{\theta}_1) - H_{\mathcal{M}}(\hat{\theta}_2))^2}{\sigma_{\mathcal{M}}^2(\hat{\theta}_1) + \sigma_{\mathcal{M}}^2(\hat{\theta}_2)} \quad (5.25)$$

Para muestras suficientemente grandes, si  $S_{\mathcal{M}}(\theta_1, \theta_2) = s$ , la hipótesis nula puede ser rechazada con nivel de significación  $\eta$  si:

$$P(\chi_{r-1}^2 > s) \leq \eta \quad (5.26)$$

donde  $P(\chi_{r-1}^2 > s)$  representa el  $p$ -valor para el estadístico  $S_{\mathcal{M}}$ .

## 5.4. Evaluación de la Prueba

La calidad y tamaño de la muestra tienen un rol fundamental, al momento de medir la cantidad de evidencia que existe en contra o a favor de alguna hipótesis. Pero sin dudas, otro elemento que influye es el tipo de test que se utiliza. En esta sección comparamos la prueba basada en diferencias de entropías propuesta con el test Z clásico de diferencia de medias de dos poblaciones independientes basado en el Teorema Central del Límite. Es importante destacar que las hipótesis de estos tres tests no son idénticas. Los test presentados comparan la entropía de ambas muestras que es función de sus parámetros. Esta función es diferente para el caso de las entropías de Shannon (5.4) y de Renyi (5.6). Por su parte, el test Z compara los valores medios poblacionales de los cuales provienen estas dos muestras que se trata como hemos visto de otra función de sus parámetros 2.23.

Para realizar esta comparación presentaremos un estudio de simulación.

En la comparación de estas pruebas vamos a considerar el tamaño de la muestra, y la variación de cada uno de los dos parámetros así como la variación conjunta de ambos.

### 5.4.1. El diseño de la comparación

Para comparar el desempeño de estas pruebas se realiza un experimento de simulación. Se consideraron las combinaciones  $(\alpha, \gamma) \in \{(-2, 1), (-5, 4), (-8, 7)\}$  y los tamaños muestrales  $n \in \{81, 121, 289, 500, 1000\}$ . Puesto que se trata de test asintóticos debemos distinguir entre su comportamiento en pequeñas muestras y su comportamiento cuando el tamaño muestral es grande.

Para cada una de estas combinaciones de parámetros y tamaños muestrales se construyó una muestra teórica con los  $n$  cuantiles de la distribución que se considera válida en la hipótesis nula, a esta muestra teórica nos referiremos en

---

adelante con 'ventana central' (VC).

Luego se construyeron muestras de datos con distribución  $\mathcal{G}_I^0$  con las siguientes características.

Cada base tiene 4000 muestras del mismo tamaño:

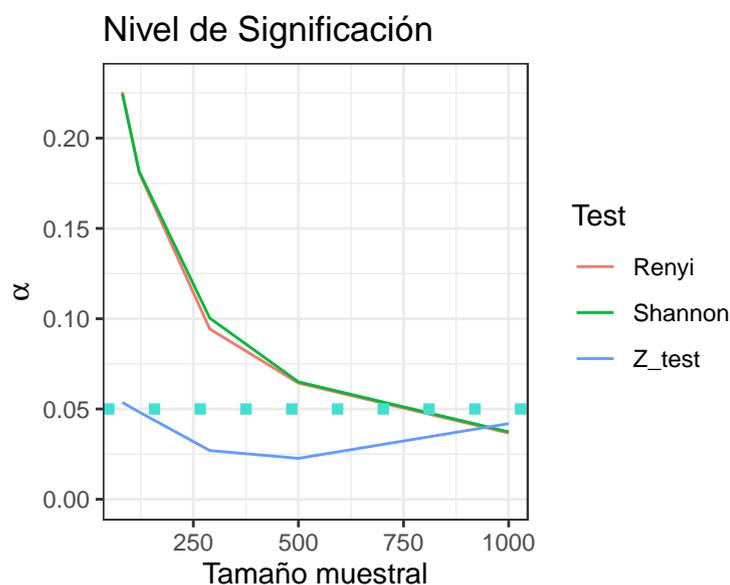
1. Las primeras 1000 muestras con idénticos parámetros que la VC, es decir  $\mathcal{G}_I^0(\alpha_C, \gamma_C, 1)$ .
2. Las siguientes 1000 muestras tienen el mismo valor de  $\alpha_{VC}$  pero varía secuencialmente el valor de  $\gamma$  alejándose del de la VC en pasos de 0,2 unidades en sentido creciente, es decir  $\mathcal{G}_I^0(\alpha_C, \gamma_C + 0,2 * k, 1)$  con  $1 \leq k \leq 1000$ .
3. Las siguientes 1000 muestras tienen el mismo valor de  $\gamma_{VC}$  pero varía secuencialmente el valor de  $\alpha$  alejándose del de la ventana central en pasos de 0,2 unidades en sentido decreciente es decir  $\mathcal{G}_I^0(\alpha_C - 0,2k, \gamma_C, 1)$  con  $1 \leq k \leq 1000$ .
4. Las últimas 1000 muestras tienen valores de  $\alpha$  y  $\gamma$  diferentes a los de la ventana central que se apartan de ella como se indicó en los ítems anteriores, pero en este último caso los dos parámetros varían simultáneamente, es decir  $\mathcal{G}_I^0(\alpha_C - 0,2k, \gamma_C + 0,2k, 1)$  con  $1 \leq k \leq 1000$ .

El objetivo de este experimento es evaluar la fortaleza de estos tests en distintos tipos de casos donde se presenta un apartamiento respecto del supuesto establecido en la hipótesis de nulidad.

Para cada una de estas 4000 muestras se estiman los parámetros de la distribución por máxima verosimilitud. Con estas estimaciones se calculan las varianzas asintóticas y los estadísticos de los tres test que se desean comparar.

Para cada una de estos contrastes se guardan los valores  $p$  correspondientes. Considerando estos valores y un nivel de significación del 0,05 se cuentan

**Figura 5.3:** Nivel de Significación alcanzado por Tamaño y por Prueba



los rechazos originados por cada una de las pruebas en las muestras generadas.

### 5.4.2. La Comparación

En la Figura 5.3 se representa el nivel de significación alcanzado, con línea punteada se marca el nivel de significación asintótico elegido para cada uno de los contrastes. En esta figura se aprecia que los tres test convergen hacia el nivel asintótico establecido 0,05.

Con estos  $p$ -valores y sabiendo cuáles casos son iguales y distintos se analizan las proporciones de falsos positivos, falsos negativos, de verdaderos positivos y verdaderos negativos.

Con estos valores se estiman las siguientes medidas para cada una de las muestras construidas.

En este experimento, para cada combinación de valores de parámetros y tamaño muestral,  $n_1 = 1000$ ,  $n_2 = 3000$  y  $n = 4000$ .

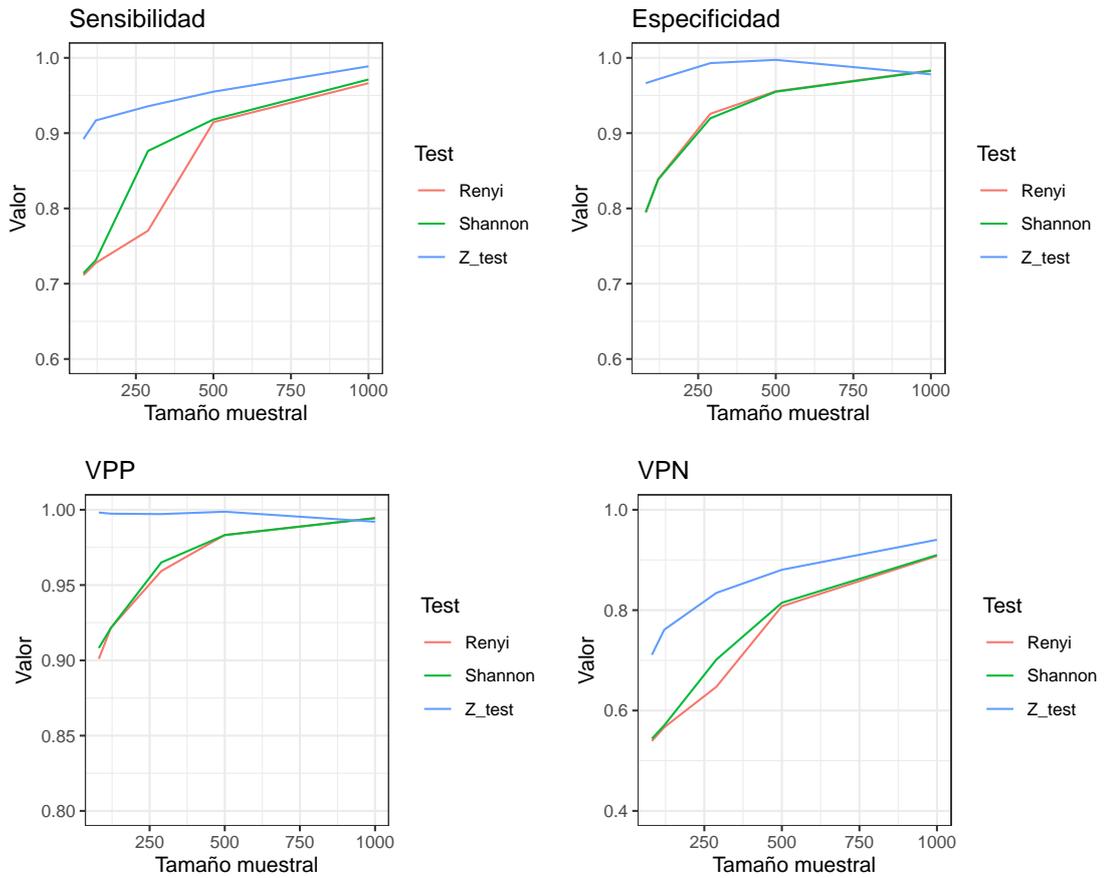
- 
- **Sensibilidad (SENS)** proporción de rechazos del test cuando efectivamente debe rechazar ( $SENS = P(\text{rech}H_0/H_0 \text{ es F}) = \frac{c}{3000}$ ).
  - **Especificidad (ESP)** proporción de no rechazos del test cuando efectivamente no debe rechazar ( $ESP = P(\text{no rech}H_0/H_0 \text{ es V}) = \frac{a}{1000}$ ).
  - **Valor Predictivo Positivo (VPP)** proporción de muestras que efectivamente son distintas a la ventana central en cuanto a las funciones de sus parámetros que el test decide correctamente rechazar ( $VPP = P(H_0 \text{ sea F} / \text{rech}H_0) = \frac{c}{a+c}$ ).
  - **Valor Predictivo Negativo (VPN)** proporción de muestras que no son distintas a la ventana central en cuanto a las funciones de sus parámetros que el test decide correctamente no rechazar ( $VPN = P(H_0 \text{ sea V} / \text{no rech}H_0) = \frac{b}{b+d}$ ).

Luego para cada tamaño muestral se promedian la sensibilidad, especificidad, valor predictivo positivo, valor predictivo negativo y nivel de significación logrado para todas las combinaciones de parámetros consideradas. Dichos valores medios en función del tamaño muestral se representan en la Figura 5.4.

En los gráficos de la Figura 5.4 se aprecia que el test Z alcanza valores superiores a los tests basados en entropías para muestras pequeñas, que los tres test alcanzan valores altos de estas características para muestras de tamaño superior a 500, que los test basados en entropías son muy similares en cuanto a estas propiedades y que la característica menos lograda por las tres pruebas es el VPN.

Luego para identificar en cuáles de estas características o métricas establecidas se diferencian entre sí los tests comparados, se conduce un análisis de la varianza no paramétrico para comparar su desempeño en términos de estas cuatro características.

Figura 5.4: Características de los tests en función del Tamaño Muestral



---

Esta comparación se hace por test y por tamaño muestral y los  $p$  valores de estos análisis se consignan en la Tabla 5.3.

De estos resultados se infiere que la especificidad, la sensibilidad y el VPN alcanzados por los tres tests es similar. Sin embargo, no resulta así con el VPP, donde la diferencia resulta estadísticamente significativa.

También surge de estos valores que la sensibilidad y el VPN varían significativamente por el tamaño muestral considerado.

## 5.5. Conclusiones

Si bien hemos notado que el test  $Z$  clásico es superior a los test basados en diferencia de entropías en cuanto a alguna de las características estudiadas, vamos a desarrollar un filtro basado en diferencia de entropías y no en el test  $Z$ . Esta elección se motiva por el hecho de que la entropía tiene valor *per se*, ya que es un indicador de la complejidad o desorden del proceso subyacente a los datos. Otro motivo es que sabemos que los datos no siguen una distribución normal.

La entropía ha sido utilizada para detectar cambios, clasificar, segmentar y umbralizar imágenes con muy buenos resultados, dado que cuantifica la cantidad de información de la imagen, mientras que la prueba  $Z$ , basada en las medias no considera la cantidad de información sino solamente su valor medio.

**Tabla 5.1:** Entropía de Shannon

$\downarrow \alpha / \gamma \rightarrow$	0,1	1	10	100	1000
-10	-3,505	-1,203	1,100	3,403	5,705
-9,5	-3,449	-1,146	1,157	3,459	5,762
-9	-3,389	-1,086	1,216	3,519	5,822
-8,5	-3,325	-1,022	1,280	3,583	5,885
-8	-3,257	-0,954	1,348	3,651	5,953
-7,5	-3,184	-0,882	1,421	3,724	6,026
-7	-3,106	-0,803	1,500	3,802	6,10
-6,5	-3,021	-0,718	1,585	3,887	6,190
-6	-2,928	-0,625	1,677	3,980	6,283
-5,5	-2,826	-0,523	1,780	4,082	6,385
-5	-2,712	-0,409	1,893	4,196	6,498
-4,5	-2,584	-0,282	2,021	4,323	6,626
-4	-2,439	-0,136	2,166	4,469	6,771
-3,5	-2,270	0,033	2,336	4,638	6,941
-3	-2,068	0,235	2,537	4,840	7,142
-2,5	-1,819	0,484	2,786	5,089	7,391
-2	-1,496	0,807	3,109	5,412	7,715
-1,5	-1,041	1,261	3,564	5,866	8,169
-1	-0,303	2,000	4,303	6,605	8,908

---

**Tabla 5.2:** Resultados de cada Test

<b>La distribución</b>	<b>Rechazo <math>H_0</math></b>	<b>No rechazo <math>H_0</math></b>	
<b>Coincide con la de la VC</b>	$a$	$b$	$n_1$
<b>No coincide con la de la VC</b>	$c$	$d$	$n_2$
	$a + c$	$b + d$	$n$

**Tabla 5.3:** Comparación de Características por tipo de Test y Tamaño muestral

	<b>test</b>	<b>tamaño</b>
<b>Sensibilidad</b>	0,2101	<b>0,0342</b>
<b>Especificidad</b>	0,6522	0,1824
<b>VPP</b>	<b>0,0189</b>	0,6808
<b>VPN</b>	0,2645	<b>0,0271</b>



Parte III

## Propuesta para la eliminación del Ruido *Speckle*



## Capítulo 6

# Filtrado No local basado en Diferencia de Entropías

### 6.1. Algoritmo propuesto para la reducción del Ruido Speckle

En esta sección vamos a explicar el método para la reducción del ruido *speckle* basado en la comparación de las entropías de dos muestras provenientes de la distribución  $\mathcal{G}_I^0$ .

El modelo considera dos muestras de igual tamaño  $\mathbf{X} = (X_1, X_2, \dots, X_n)$  e  $\mathbf{Y} = (Y_1, Y_2, \dots, Y_n)$  bajo el supuesto de que  $X \sim \mathcal{G}_I^0(\alpha_1, \gamma_1, 1)$  e  $Y \sim \mathcal{G}_I^0(\alpha_2, \gamma_2, 1)$  y las hipótesis que se contrastan son:

$$H_0 : (\alpha_1, \gamma_1) = (\alpha_2, \gamma_2)$$

versus

$$H_1 : (\alpha_1, \gamma_1) \neq (\alpha_2, \gamma_2)$$

Se trata de una prueba de nivel aproximado basada en la distribución asintótica del estadístico  $S_{\mathcal{M}}$ ,  $\mathcal{M} \in \{S, R\}$  que ha sido descrita en el capítulo anterior.

El filtro se ha construido de la siguiente forma. Se considera una secuencia de ventanas deslizantes a través de toda la imagen,  $\mathbf{W}_i$ ,  $i = 1, \dots, n$  de tamaño  $t \times t$ . Para cada ventana  $\mathbf{W}_i$ , consideramos una ventana central dentro de la misma,  $W_i^c$ , de tamaño  $k \times k$ , donde  $k < t$  y una serie de sub-ventanas de tamaño  $k \times k$ ,  $W_i^j$ ,  $j = 1 \dots, n$  centradas en cada pixel de  $\mathbf{W}_i$ , como se muestra en la Figura 6.1. Para cada  $i \in \{1, \dots, n\}$ , la muestra  $W_i^c$  corresponde a un entorno del píxel central que denotamos  $x_i^c$ . Los píxeles de la muestra  $W_i^c$  se comparan con los píxeles de las ventanas deslizantes  $W_i^j$  utilizando el estadístico  $S_M$ ,  $M \in \{S, R\}$  y se calculan los  $p$ -valores,  $p_i^j$ , correspondientes. Utilizando una función suavizante y un nivel de significancia  $\eta$ , se construye la máscara de convolución adaptativa [15].

En este trabajo utilizamos como familia de funciones suavizantes a la logística centrada en  $2 * \eta$ :

$$F_\eta(x) = \frac{1}{1 + e^{-2\beta*(x-2\eta)}}, \beta = K * \eta, K \in \mathbb{N}$$

El Algoritmo 1 muestra los pasos del método.

## 6.2. Resultados

Para estudiar la calidad del suavizado, calculamos inicialmente el número equivalente de looks o vistas y el promedio correspondiente a toda la imagen.

Un buen método de suavizado debería incrementar el número equivalente de looks y no cambiar el valor medio de la imagen original completa en la imagen resultante.

Considerando que la imagen filtrada es como una estimación de la imagen original, es deseable que  $\hat{Z}$  se aproxime al valor real de  $Z$ , por eso es deseable que el filtro preserve la media.

Dado que una imagen con mayor número de looks es menos ruidosa, esta estimación suele utilizarse como una medida de la reducción del ruido [83].

---

**Input:** I imagen original de tamaño  $n \times n$ ,  $t, k, t > k$  tamaños de las ventanas grande y pequeña, respectivamente.

**Output:** Imagen suavizada.

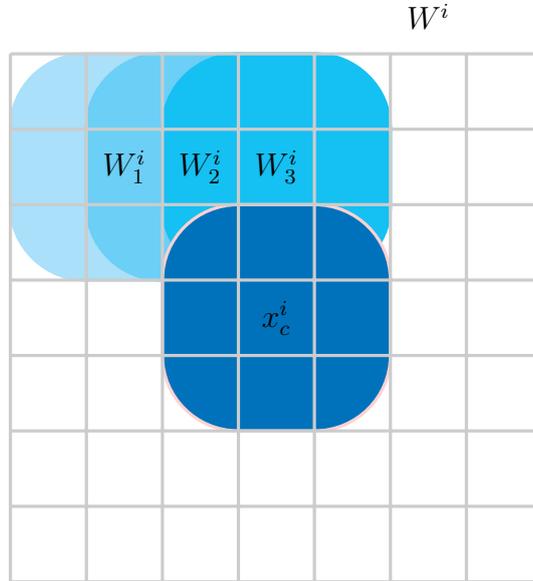
**for**  $i = 1$  **to**  $n$  **do**

- 1- Considere una ventana deslizante  $\mathbf{W}_i$  de tamaño  $t \times t$ .
- 2- Sea  $x_c^i$  el píxel central  $\mathbf{W}_i$ .
- 3- Considere un entorno de tamaño  $k \times k$  del píxel central  $x_c^i$ , denominado  $W_i^c$ .
- 4- Calcule los estimadores de máxima verosimilitud  $(\hat{\alpha}_i^c, \hat{\gamma}_i^c)$ , la entropía y la varianza asintótica de la misma para la muestra formada por los píxeles del entorno  $W_i^c$ .
- 5- Considere una secuencia de ventanas de tamaño  $k \times k$  que designamos con  $\{W_i^j\}$ ,  $j = 1, \dots, t^2$  que son solapadas e interiores a la ventana mayor  $\mathbf{W}_i$ , correspondientes al entorno de cada píxel  $x_j \in \mathbf{W}_i$ ,  $j = 1, \dots, t^2$ , como se muestra en la Figura 6.1.
- 6- Calcule los estimadores  $(\hat{\alpha}_i^j, \hat{\gamma}_i^j)$ , de la entropía y su varianza asintótica para la muestra  $W_i^j$ ,  $j = 1, \dots, t^2$ .
- 7- Calcule el estadístico  $S_i^j$ , usando la Eq. (5.25) y el  $p$ -valor  $p_i^j$ ,  $j = 1, \dots, t^2$  correspondiente usando la distribución asintótica  $\chi_1^2$ .
- 8- Calcule  $x_i^{\text{smooth}} = \sum_{j=1}^{t^2} F(p_i^j) x_i^j$ , donde  $x_i^j$  es el  $j$ -ésimo píxel de  $\mathbf{W}_i$  y  $F : \mathbb{R} \rightarrow [0, 1]$ .

**return**  $x^{\text{smooth}}$

**Algoritmo 1:** Algoritmo de Eliminación de Ruido Speckle

**Figura 6.1:** La ventana deslizante y los parches deslizantes interiores a la misma

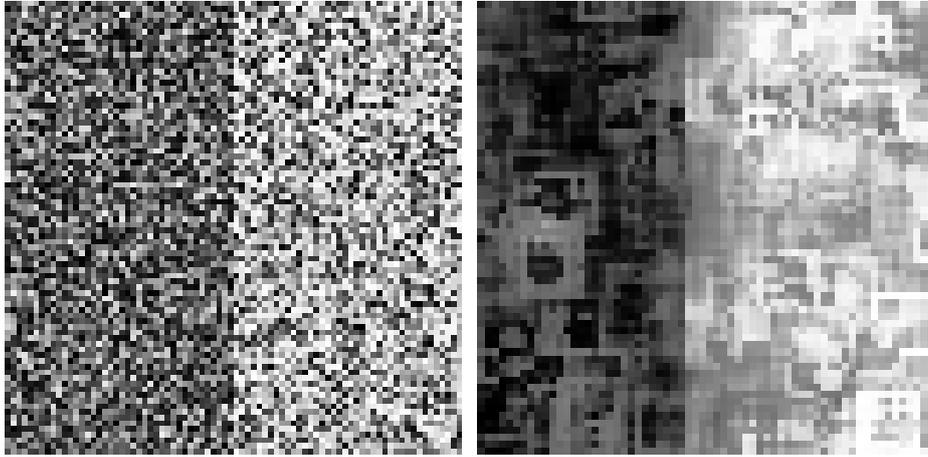


En el próximo capítulo realizaremos un análisis más exhaustivo de la calidad del filtrado y lo compararemos con otros filtros disponibles que han sido extensamente evaluados y aplicados en la comunidad SAR.

La Figura 6.2 muestra el resultado de aplicar el método a una imagen sintética generada bajo la distribución  $\mathcal{G}_I^0$ . La Figura 6.3(a) muestra la imagen de tamaño  $80 \times 80$  sintética original, que fue generada con datos  $\mathcal{G}_I^0(-4, 3, 1)$  la banda vertical de la izquierda y con datos  $\mathcal{G}_I^0(-1, 5, 0, 5, 1)$  la banda vertical de la derecha. Figure 6.3(b) muestra la imagen filtrada con ventanas deslizantes de tamaño  $7 \times 7$ , dentro de una ventana marco de  $15 \times 15$  también deslizante, aplicando el estadístico basado en la entropía de Shannon, con función suavizante logística de parámetros  $\eta = 0,25$  y  $K = 40$ .

La Tabla 6.1 presenta el número equivalente de looks de la imagen original y el promedio para la imagen original y para la imagen suavizada. Se puede apreciar un incremento en el valor del número de looks ENL mientras que el promedio general de la imagen no cambia.

**Figura 6.2:** Resultado de aplicar el método de filtrado basado en la entropía de Shannon



(a) Imagen Original.

(b) Imagen Filtrada  $\mathbf{W} = 15 \times 15$ ,  $W_i^c = 7 \times 7$ ,  $\eta = 0,25$  y  $K = 40$

**Tabla 6.1:** Estimaciones del ENL para las imágenes de la Figure 6.2.

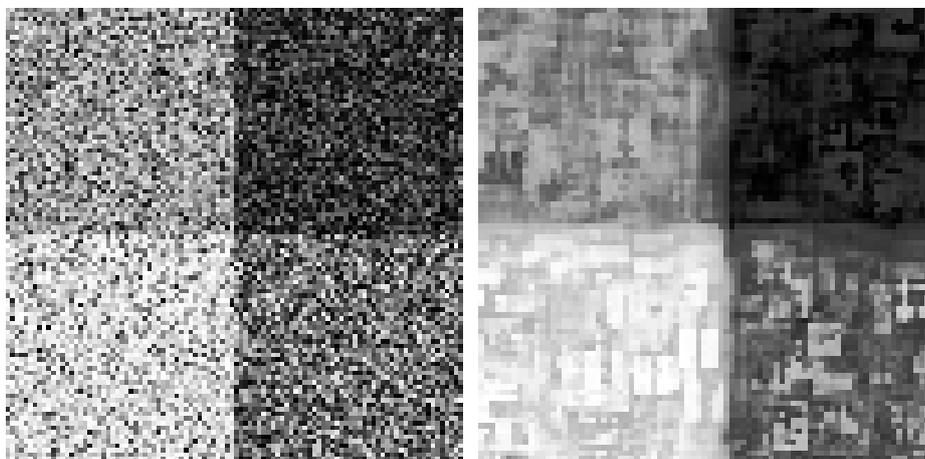
Figura	ENL	Promedio
6.3(a)	0,603	1,1965
6.3(b)	3,27	1,1978

La Figura 6.3 (página 142) muestra el resultado de aplicar el método a una imagen sintética generada utilizando la distribución  $\mathcal{G}_I^0$ .

La Figura 6.4(a) (página 142) muestra los datos originales, las cuatro regiones de la imagen se han generado utilizando los valores  $\mathcal{G}_I^0(-5, 10, 1)$ ,  $\mathcal{G}_I^0(-5, 1, 1)$ ,  $\mathcal{G}_I^0(-1,5, 10, 1)$  y  $\mathcal{G}_I^0(-1,5, 1, 1)$  de la distribución. La Figura 6.4(b) muestra la imagen filtrada con pequeñas ventanas de tamaño  $9 \times 9$ , dentro de una ventana marco de tamaño  $15 \times 15$  aplicando el algoritmo basado en la entropía de Shannon, con función suavizante logística de parámetros  $\eta = 0,25$  y  $K = 40$ .

La Tabla 6.2 muestra el número equivalente de Looks (ENL) y los prome-

**Figura 6.3:** Resultado de la aplicación del filtro basado en la entropía de Shannon.



(a) Imagen Original.

(b) Imagen Filtrada  $W = 15 \times 15$ ,  $W_i^c = 9 \times 9$ ,  $\eta = 0,25$  y  $K = 40$ .

dios de la imagen original y de la filtrada.

Puede apreciarse que el filtro preserva el valor medio y al mismo tiempo logra un incremento en el ENL, también para este segundo experimento con imagen sintética.

**Tabla 6.2:** Estimaciones del número equivalente de looks para las imágenes de la Figura 6.3.

Figura	ENL	Promedio
6.3(a)	0,0624	5,2338
6.3(b)	0,402	5,1767

---

### 6.3. Evaluación del Filtro basado en Entropías

En esta sección se propone una evaluación de la calidad del filtrado del algoritmo propuesto para imágenes SAR, utilizando un criterio comparativo con otros filtros disponibles y criterios extensamente validados.

Un filtro de calidad para imágenes SAR elimina el ruido *speckle* pero preserva las características fundamentales de la escena, tales como la textura, los objetos y los bordes. Un indicador de importancia de la calidad del filtrado es la preservación de la media local en los datos transformados.

Muchas técnicas de filtrado así como indicadores de calidad de estos procedimientos han sido propuestas en la literatura, sin embargo, evaluar objetivamente la superioridad de un filtro sobre otros sigue considerándose un asunto no trivial. Existe, por otro lado, un consenso respecto de que una medida universal de la calidad de un filtrado no está aún disponible para imágenes SAR.

Una cantidad importante de indicadores de calidad de filtrado han sido propuestos en la literatura, entre ellos cabe mencionar: la razón pico entre señal y ruido (PSNR), la correlación  $\beta$  [2], la cifra de mérito de Pratt (PFoM) [90], y el índice de similitud estructural (SSIM) [114], que han sido utilizados con frecuencia en literatura de la comunidad SAR.

Existe un consenso acerca de la obligatoriedad de evaluar la preservación del valor medio y la reducción de la varianza en un área homogénea seleccionada por el usuario, como orientativas de la calidad del filtrado.

Dada la simplicidad de su aplicación, muy frecuentemente los investigadores utilizan el número equivalente de looks (ENL) [83], que debe ser tan alto como sea posible en la imagen filtrada. Un valor alto de ENL indica mayor rechazo de *speckle* y por ende una capacidad mejorada de distinguir los diferentes niveles de textura de la imagen. Sin embargo, es sabido que valores altos de ENL se obtienen fácilmente debido a un sobrefiltrado de la imagen, dando

a la imagen filtrada un aspecto degradado indeseable.

## 6.4. Los recursos disponibles para la evaluación

Dada una región homogénea de una imagen, con un valor medio  $\bar{x}$  y una desviación standard  $s$ , se estima:

$$\text{ENL} = \frac{\mu^2}{\sigma^2} \quad (6.1)$$

Análisis recientes utilizan la relación de cociente punto a punto entre la imagen original y la imagen filtrada [9, 46]. La idea es inspeccionar el patrón de regularidad de la denominada imagen cociente  $H$ , definida como:

$$H(x, y) = \frac{U(x, y)}{V(x, y)}, \quad (6.2)$$

siendo  $U$  la imagen SAR original y  $V$  su versión filtrada. En una operación de filtro ideal, la imagen  $U$  debería ser ruido *speckle* puro. Por lo cual si de la inspección de esta imagen surge la presencia de alguna estructura geométrica relacionada con la imagen original implicaría una deficiencia en el procedimiento de filtrado.

Un filtrado en áreas donde el *speckle* es dominante, la imagen cociente debería mostrar solo ruido y el ENL estimado debería ser cercano al ENL de la imagen SAR original. De acuerdo con el modelo multiplicativo, la media del ruido *speckle* es 1. Por lo tanto la media de la imagen cociente debería ser cercana a 1.

Varios enfoques han sido considerados para cuantificar la presencia de contenido geométrico en la imagen cociente. Entre ellos el análisis estadístico utilizando bondad de ajuste a una distribución de probabilidad o el análisis del histograma sin que se hayan alcanzado resultados convincentes.

También se trabajó con detectores de bordes como el de Canny [12] y el de falsa alarma constante [106] sobre las imágenes filtradas. Debido a la naturaleza multiplicativa del *speckle*, el tratamiento diferencial como el de Canny no

---

resulta adecuado para las imágenes SAR, mientras que el análisis de la imagen cociente surge como el tratamiento natural.

Bajo el supuesto de que una imagen bien filtrada debería dejar como diferencia entre imagen original y filtrada solamente ruido sin contenido geométrico, en [45] se propone un nuevo estimador, denominado estimador de relación  $\alpha\beta$ .

Este estimador considera por un lado el valor medio del *speckle* y el ENL medido sobre un área homogénea dentro de la imagen cociente así como el contenido geométrico restante de la imagen cociente cuantificado mediante el  $\beta$ ratio. La relación  $\alpha\beta$  se define:

$$\alpha\beta = \alpha |\delta_{\text{ENL}}| + (1 - \alpha) |\delta_{\mu}| + \beta_H. \quad (6.3)$$

Siendo  $\alpha \in [0, 1]$ ,  $\delta_{\text{ENL}} = \text{ENL}_U - \text{ENL}_H$ ,  $\delta_{\mu} = 1 - \mu_H$ . Es importante destacar que tanto  $\mu_H$  como  $\text{ENL}_H$  se miden sobre zonas homogéneas de la imagen.

El ENL de la imagen original en general es conocido, pero si fuera desconocido se estima también sobre una zona homogénea. Por su parte  $\beta_H$  es el estimador de borde estándar [2].

El valor de  $\beta$  mide la correlación entre los bordes de la imagen original y la filtrada, siendo su valor ideal en la preservación de bordes el 1, pero pudiendo variar en general entre 0 y 1.

En la ecuación (6.3) el estimador  $\beta_H$  estima la correlación entre los bordes de la imagen original y los de la imagen cociente, por lo cual su valor ideal sería 0.

El valor de  $\alpha$  opera como una suma convexa donde se ponderan los términos de la diferencia en ENL y la distancia a ruido puro de la imagen cociente. Si bien el valor de este parámetro es elegido por el usuario, para poder comparar eficazmente varios métodos de filtrado, el valor de  $\alpha$  debe ser el mismo para todos.

Por todo lo considerado, el mejor método de filtrado será el que logre el

menor valor del coeficiente  $\alpha\beta$ .

En [47] los autores definen un nuevo indicador de la calidad de filtrado que no requiere ninguna referencia básica y resulta compatible con las medidas clásicas de preservación del valor medio, índice  $\beta$  de correlación de bordes, ENL, PSNR y MSSIM.

Esta medida utiliza solamente la imagen original y la imagen filtrada y la estimación del ENL se mide la desviación del filtro ideal como una combinación de desviaciones de propiedades marginales ideales con una medida de la estructura restante en la imagen cociente.

Esta medida de calidad del filtrado se basa en descriptores de primer y segundo orden. La componente de primer orden es una medida estadística del ruido *speckle* remanente, esta componente comprende dos términos: uno relativo a la conservación de la media y el otro a la preservación del ENL, mientras que la componente de segundo orden se vincula con el contenido geométrico de la imagen cociente.

El residuo de primer orden resulta del promedio de dos residuos. Considerando  $n$  áreas homogéneas que designamos con  $A_i$  con  $1 \leq i \leq n$  se define sobre cada una de ellas, el residuo del ENL como el valor absoluto de los residuos debidos a las desviaciones respecto del ENL ideal que es el de la imagen original:

$$r_{\widehat{\text{ENL}}}(i) = \frac{|\widehat{\text{ENL}}_U(i) - \widehat{\text{ENL}}_H(i)|}{\widehat{\text{ENL}}_U(i)}. \quad (6.4)$$

El residuo de media como el valor absoluto de las desviaciones respecto de la media ideal de la imagen cociente que es 1:

$$r_{\widehat{\mu}}(i) = |1 - \widehat{\mu}_H(i)|, \quad (6.5)$$

y luego, promediando estos dos residuos, tenemos el indicador de primer orden

$$r_{\widehat{\text{ENL}}, \widehat{\mu}} = \frac{1}{2n} \sum_{i=1}^n (r_{\widehat{\text{ENL}}}(i) + r_{\widehat{\mu}}(i)). \quad (6.6)$$

---

En una operación de filtrado ideal debería ocurrir que el promedio de ambos residuos sea nulo, es decir:  $r_{\widehat{\text{ENL}}, \hat{\mu}} = 0$ .

Haralick [53] mostró la necesidad de un modelo estadístico para describir un conjunto de características significativas de la textura de una imagen. Para definir estas características se debe disponer de la matriz de niveles de gris y para considerar las relaciones espaciales, de la matriz de coocurrencia de niveles de gris de segundo orden.

Esta matriz describe la frecuencia con la que ese nivel de gris ocurre en una relación espacial específica con otro valor de gris dentro de una ventana determinada. Si se han considerado  $k$  niveles de gris, esta matriz será de orden  $k \times k$ . Haralick definió 14 características que brindan información relativa a la naturaleza de la textura de una imagen a partir de la matriz de coocurrencia de niveles de gris (MCNG).

El contenido geométrico de la imagen restante se estima mediante el momento de la diferencia inversa también conocido como homogeneidad de la MCNG.

Valores bajos de este indicador se asocian con baja variación de textura, la medida de homogeneidad  $h$  de segundo orden propuesta es:

$$h = \sum_i \sum_j \frac{1}{1 + (i - j)^2} P(i, j), \quad (6.7)$$

siendo  $P(i, j)$  el elemento de posición  $(i, j)$  de la matriz de co-ocurrencia.

Este coeficiente de homogeneidad  $h$  se calcula sobre todas las posiciones y bajo la hipótesis de que si la imagen cociente no tiene contenido geométrico, el valor de  $h$  es invariante por permutaciones, se compara entonces el valor de  $h$  en la imagen cociente  $h_0$  con el de sus  $g$  permutaciones. Llamando  $h_0$  a la media de homogeneidad de la imagen cociente original y  $h_m$  con  $1 \leq m \leq g$  la media de sus  $g$  permutaciones, se define:

$$\delta_h = 100 \left| \frac{h_0 - \bar{h}_g}{h_0} \right|, \quad (6.8)$$

siendo  $\bar{h}_g$  el promedio de  $h_i$  sobre las  $g \geq 1$  permutaciones  $I_i$  de la imagen cociente original  $I_0$ .

El valor absoluto de la ecuación 6.8 indica la variación relativa de  $h_0$  en porcentaje como una medida del apartamiento de la hipótesis de nulidad que establece que la homogeneidad es la misma para cada permutación de la imagen cociente con resultados comparables con  $r_{\widehat{\text{ENL}}}$ , luego de experimentar con imágenes sintéticas y reales. Finalmente el estimador propuesto en [47] combina la medida que cuantifica la estructura geométrica remanente en la imagen cociente y las desviaciones respecto de las propiedades estadísticas de ella respecto de las ideales.

Cuanto mayor es  $\delta_h$  mayor es la cantidad de estructura geométrica remanente en la imagen cociente.

Este valor 100 es la propuesta de los autores para producir resultados consistentes.

$$\mathcal{M}_0 = r_{\widehat{\text{ENL}}, \hat{\mu}} + \delta_h. \quad (6.9)$$

Un proceso de filtrado es ideal cuando  $\mathcal{M}_0 = 0$ , cuanto mayor sea su valor más alejado del ideal estará el proceso de filtrado.

## 6.5. Calidad del Filtrado Propuesto

### 6.5.1. Preservación de la Media

Como hemos destacado una cualidad importante para cualquier filtro es preservar el valor de la media local. Para analizar si el filtro propuesto satisface este requerimiento se utilizó la imagen sintética de la Figura 6.3.

Se eligió una línea al azar y se representaron los valores de la imagen original, los valores de la imagen filtrada con tamaños de ventanas  $w$  y se señalaron con líneas horizontales los valores esperados teóricos de cada subregión.

La Figura 6.4 muestra los resultados. En la Figura 6.4 se aprecia que la

**Figura 6.4:** Análisis de Preservación de la Media

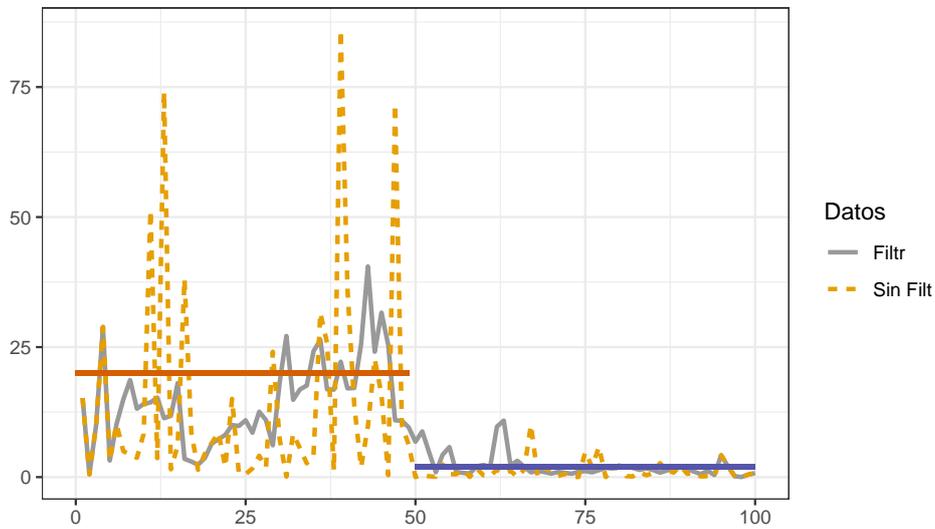


imagen filtrada se aproxima al valor medio de la región mientras que la imagen original tiene importantes variaciones respecto de este valor.

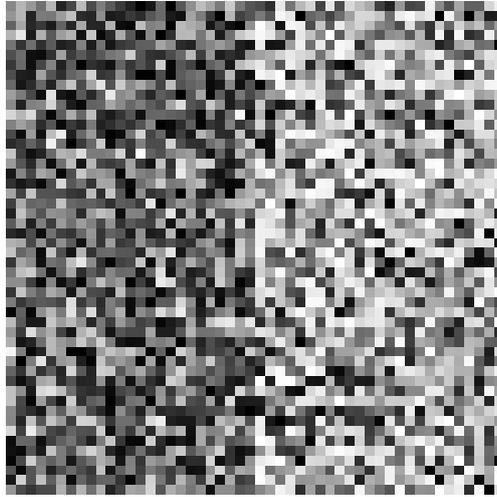
### 6.5.2. Imagen Sintética

Con el objetivo de evaluar el desempeño del filtro propuesto, se compara su desempeño con otros filtros clásicos de amplia difusión. En este apartado se generó una imagen sintética de tamaño  $100 \times 100$  con dos bandas verticales utilizando los valores  $\mathcal{G}_I^0(-1,5, 1, 1)$  para la región izquierda y  $\mathcal{G}_I^0(-4, 1, 1)$  para la región derecha, como se muestra en la Figura 6.5.

Se aplicaron sobre esta imagen el filtro basado en distancias entre entropías descrito en el Capítulo 6 utilizando diferentes tamaños de ventanas centrales y ventanas deslizantes dentro de ellas. En la Figura 6.6 de la página 150 se presenta el resultado del filtrado con una ventana marco de  $11 \times 11$  y ventanas deslizantes dentro de ella de  $7 \times 7$ .

La Imagen Sintética Cociente correspondiente al filtro basado en distancia de entropías se presenta en la Figura 6.7 de la página 150.

**Figura 6.5:** Imagen Sintética con dos bandas verticales



La evaluación comparativa del desempeño del filtro propuesto con los filtros de Lee-M( Filtro de Lee mejorado), SRAD y FANS presentados en el Capítulo I se detalla en las próximas secciones de este Capítulo.

### **6.5.3. Comparación de desempeño sobre Imágenes Sintéticas**

Se realizó el filtrado de la imagen sintética con cada uno de los siguientes cuatro filtros: el propuesto, SRAD, Lee Mejorado y FANS. Luego, para cada caso se halló la imagen cociente.

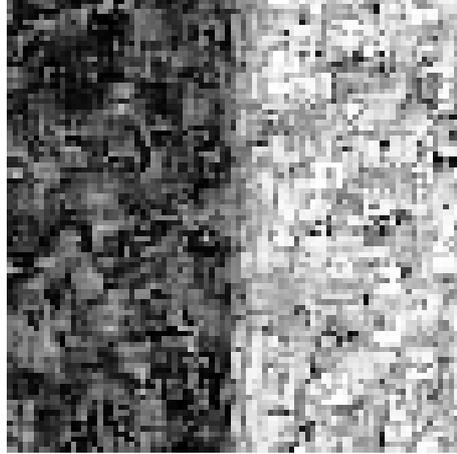
La imagen sintética filtrada con los filtros de Lee-M, SRAD y FANS se presenta en las figuras 6.9, 6.8 y 6.10 respectivamente.

La imagen sintética cociente correspondiente a los filtros de Lee-M, SRAD y FANS se presentan en las figuras 6.12, 6.11 y 6.13 respectivamente en la página 162.

Se consideraron sobre la imagen sintética original y sobre la imagen cociente correspondiente a cada uno de los cuatro filtros,  $n = 10$  muestras de zonas homogéneas, sobre las cuales se calcularon las medias, los desvíos y se estimó el ENL. Con los promedios de estas estimaciones se calcularon los valores de

---

**Figura 6.6:** Imagen Sintética filtrada con Filtro de Distancia de Entropías de la Imagen 6.5



$r_{\hat{\mu}}, r_{\widehat{\text{ENL}}}, r_{\hat{\mu}, \widehat{\text{ENL}}}$ . En las Tabla 6.3 de la página 153 se presentan los valores correspondientes al Filtro propuesto. En las Tabla 6.4 de la página 154 se presentan los valores correspondientes al Filtro de Lee Mejorado. En las Tabla 6.5 de la página 154 se presentan los valores correspondientes al Filtro SRAD. En las Tabla 6.6 de la página 155 se presentan los valores correspondientes al Filtro FANS.

Para cada imagen cociente se calculó el coeficientes de homogeneidad de Haralick  $h_0$  y se realizaron  $g = 10$  permutaciones de los valores  $I_m$  con  $1 \leq m \leq 10$  a las cuales también se les calculó el coeficiente de homogeneidad y se estimó para cada filtrado de estas imágenes. Con el promedio de estos coeficientes  $h_g$  se estimó el valor  $\delta_h$  para cada procedimiento de filtrado. Los resultados se presentan en la Tabla 6.7 de la página 155.

#### **6.5.4. Aplicación sobre Imagen Real**

La Figura 6.14 de la página 163 se muestra una imagen real Single Look, obtenida por el sensor E-SAR correspondiente a los alrededores de la ciudad de Munich, Alemania. En esta imagen se aprecian zonas oscuras correspondien-

**Tabla 6.3:** Evaluación Filtro Basado en Distancias de Entropías

	$\hat{\mu}_U$	$\hat{\sigma}_U$	$\widehat{\text{ENL}}_U$	$\hat{\mu}_H$	$\hat{\sigma}_H$	$\widehat{\text{ENL}}_H$	$r_{\widehat{\text{ENL}}}$	$r_{\hat{\mu}}$
<i>ROI</i> <sub>1</sub>	0.2541	0.2622	0.9391	1.1951	1.3070	0.8361	0.1232	0.1951
<i>ROI</i> <sub>2</sub>	0.2549	0.3129	0.6639	1.2456	1.4866	0.7020	0.0543	0.2456
<i>ROI</i> <sub>3</sub>	0.2890	0.4440	0.4236	1.3232	2.1122	0.3925	0.0794	0.3232
<i>ROI</i> <sub>4</sub>	0.3698	0.4655	0.6311	1.2388	1.8775	0.4354	0.4496	0.2388
<i>ROI</i> <sub>5</sub>	0.3132	0.3937	0.6330	1.2701	2.1392	0.3525	0.7957	0.2701
<i>ROI</i> <sub>6</sub>	0.3953	0.5054	0.6117	1.4770	2.4503	0.3634	0.6835	0.4770
<i>ROI</i> <sub>7</sub>	1.7412	4.2528	0.1676	1.7573	3.3308	0.2783	0.3978	0.7573
<i>ROI</i> <sub>8</sub>	1.4716	3.4909	0.1777	1.9863	4.8208	0.1698	0.0468	0.9863
<i>ROI</i> <sub>9</sub>	1.5497	1.9587	0.6259	1.3589	1.7656	0.5923	0.0567	0.3589
<i>ROI</i> <sub>10</sub>	1.7620	3.4169	0.2659	1.8073	4.2301	0.1825	0.4569	0.8073

**Tabla 6.4:** Evaluación Filtro de Lee Mejorado

	$\hat{\mu}_U$	$\hat{\sigma}_U$	$\widehat{\text{ENL}}_U$	$\hat{\mu}_H$	$\hat{\sigma}_H$	$\widehat{\text{ENL}}_H$	$r_{\widehat{\text{ENL}}}$	$r_{\hat{\mu}}$
<i>ROI</i> <sub>1</sub>	0.2541	0.2622	0.9391	0.9627	1.0070	0.9141	0.0273	0.0373
<i>ROI</i> <sub>2</sub>	0.2549	0.3129	0.6639	0.9863	1.1672	0.7141	0.0702	0.0137
<i>ROI</i> <sub>3</sub>	0.2890	0.4440	0.4236	0.9438	1.2491	0.5710	0.2580	0.0562
<i>ROI</i> <sub>4</sub>	0.3698	0.4655	0.6311	1.0236	1.1825	0.7493	0.1578	0.0236
<i>ROI</i> <sub>5</sub>	0.3132	0.3937	0.6330	0.9584	1.1499	0.6947	0.0887	0.0416
<i>ROI</i> <sub>6</sub>	0.3953	0.5054	0.6117	1.0109	1.2799	0.6238	0.0194	0.0109
<i>ROI</i> <sub>7</sub>	1.7412	4.2528	0.1676	1.0289	2.1367	0.2319	0.2771	0.0289
<i>ROI</i> <sub>8</sub>	1.4716	3.4909	0.1777	0.8511	1.8884	0.2031	0.1252	0.1489
<i>ROI</i> <sub>9</sub>	1.5497	1.9587	0.6259	0.9055	1.0962	0.6824	0.0827	0.0945
<i>ROI</i> <sub>10</sub>	1.7620	3.4169	0.2659	1.0442	1.8279	0.3263	0.1850	0.0442

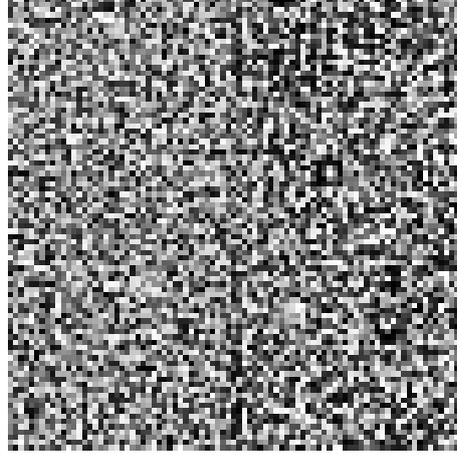
**Tabla 6.5:** Evaluación Filtro SRAD

	$\hat{\mu}_U$	$\hat{\sigma}_U$	$\widehat{\text{ENL}}_U$	$\hat{\mu}_H$	$\hat{\sigma}_H$	$\widehat{\text{ENL}}_H$	$r_{\widehat{\text{ENL}}}$	$r_{\hat{\mu}}$
<i>ROI</i> <sub>1</sub>	0.2541	0.2622	0.9391	0.4801	0.5027	0.9120	0.0297	0.5199
<i>ROI</i> <sub>2</sub>	0.2549	0.3129	0.6639	0.4748	0.5827	0.6641	0.0003	0.5252
<i>ROI</i> <sub>3</sub>	0.2890	0.4440	0.4236	0.5693	0.9728	0.3425	0.2369	0.4307
<i>ROI</i> <sub>4</sub>	0.3698	0.4655	0.6311	0.7902	1.0566	0.5593	0.1284	0.2098
<i>ROI</i> <sub>5</sub>	0.3132	0.3937	0.6330	0.7135	0.9010	0.6271	0.0094	0.2865
<i>ROI</i> <sub>6</sub>	0.3953	0.5054	0.6117	0.8317	1.0657	0.6091	0.0043	0.1683
<i>ROI</i> <sub>7</sub>	1.7412	4.2528	0.1676	1.3212	2.5614	0.2660	0.3699	0.3212
<i>ROI</i> <sub>8</sub>	1.4716	3.4909	0.1777	1.7200	4.0427	0.1810	0.0183	0.7200
<i>ROI</i> <sub>9</sub>	1.5497	1.9587	0.6259	1.7058	3.1821	0.2873	1.1783	0.7058
<i>ROI</i> <sub>10</sub>	1.7620	3.4169	0.2659	1.9368	4.2206	0.2106	0.2628	0.9368

**Tabla 6.6:** Evaluación Filtro FANS

	$\hat{\mu}_U$	$\hat{\sigma}_U$	$\widehat{\text{ENL}}_U$	$\hat{\mu}_H$	$\hat{\sigma}_H$	$\widehat{\text{ENL}}_H$	$r_{\widehat{\text{ENL}}}$	$r_{\hat{\mu}}$
<i>ROI</i> <sub>1</sub>	0.2541	0.2622	0.9391	0.7062	0.3705	3.6325	0.7415	0.2938
<i>ROI</i> <sub>1</sub>	0.2549	0.3129	0.6639	0.6771	0.3794	3.1853	0.7916	0.3229
<i>ROI</i> <sub>1</sub>	0.2890	0.4440	0.4236	0.7101	0.3748	3.5891	0.8820	0.2899
<i>ROI</i> <sub>1</sub>	0.3698	0.4655	0.6311	0.7803	0.3934	3.9352	0.8396	0.2197
<i>ROI</i> <sub>1</sub>	0.3132	0.3937	0.6330	0.7106	0.3935	3.2608	0.8059	0.2894
<i>ROI</i> <sub>1</sub>	0.3953	0.5054	0.6117	0.7628	0.4045	3.5571	0.8280	0.2372
<i>ROI</i> <sub>1</sub>	1.7412	4.2528	0.1676	0.6673	0.4704	2.0127	0.9167	0.3327
<i>ROI</i> <sub>1</sub>	1.4716	3.4909	0.1777	0.6672	0.3949	2.8551	0.9378	0.3328
<i>ROI</i> <sub>1</sub>	1.5497	1.9587	0.6259	0.7367	0.4675	2.4833	0.7479	0.2633
<i>ROI</i> <sub>1</sub>	1.7620	3.4169	0.2659	0.6918	0.4660	2.2035	0.8793	0.3082

**Figura 6.7:** Imagen Sintética cociente del Filtro de Distancia de Entropías correspondiente a la Imagen Filtrada 6.6



**Tabla 6.7:** Comparación Resultados Filtrado Imagen Sintética

	$r_{\widehat{\mu}, \widehat{\text{ENL}}}$	$h_0$	$\bar{h}_g$	$\delta_h$	$\mathcal{M}_0$
<b>SRAD</b>	0.3531	0.9994	0.9994	0.0048	0.3579
<b>Entropía</b>	0.3902	0.9702	0.9727	0.2596	0.6497
<b>Lee M</b>	0.0896	0.7837	0.7950	1.4436	1.5332
<b>FANS</b>	0.5630	0.1995	0.1927	3.4085	3.9715

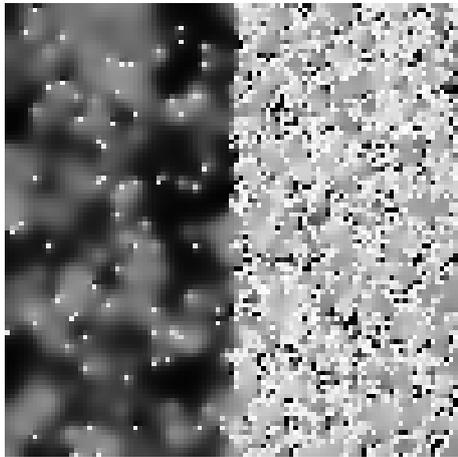
tes a campos de cultivos y áreas claras o brillantes correspondientes a zonas urbanas. En la Figura 6.15 de la página 164 se muestra la imagen de Munich con el filtro propuesto con ventanas de  $27 \times 27$  y de  $9 \times 9$  respectivamente para la central y las deslizantes dentro de la central.

En la Figura 6.16 de la página 164 se muestra la imagen cociente correspondiente al filtro basado en entropías de Munich.

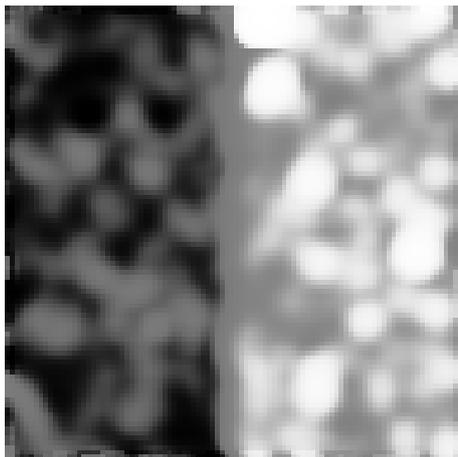
En la Figura 6.16 de la página 164 se muestra la imagen cociente de Munich correspondiente al filtro propuesto.

---

**Figura 6.8:** Imagen Sintética filtrada con Filtro de SRAD



**Figura 6.9:** Imagen Sintética filtrada con Filtro de Lee Mejorado

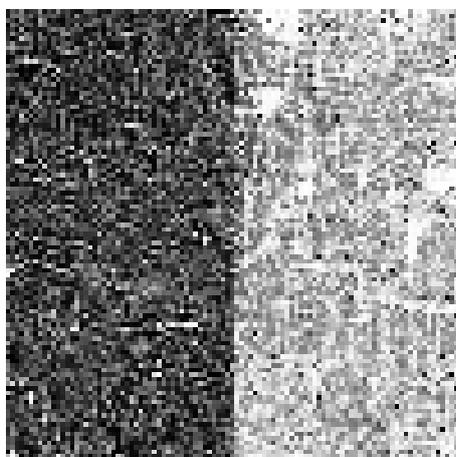


### 6.5.5. Imagen de Munich

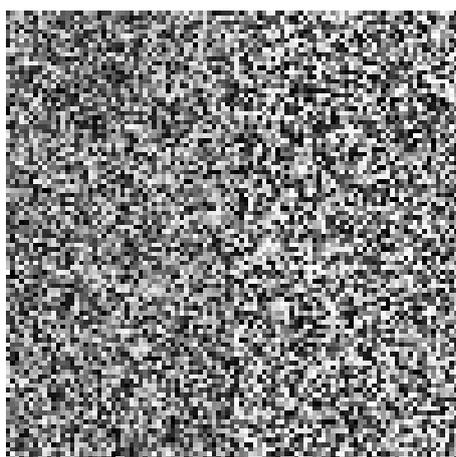
En la Figura 6.17 de la página 165 se muestra la imagen de Munich filtrada con el filtro SRAD. En la Figura 6.18 de la página 165 se muestra la imagen cociente de Munich correspondiente al filtro SRAD. En la Figura 6.19 de la página 166 se muestra la imagen de Munich filtrada con el filtro de Lee Mejorado.

En la Figura 6.20 de la página 166 se muestra la imagen cociente de Munich correspondiente al filtro de Lee mejorado. En la Figura 6.21 de la página 167 se

**Figura 6.10:** Imagen Sintética filtrada con Filtro FANS



**Figura 6.11:** Imagen Sintética cociente con Filtro de SRAD



muestra la imagen de Munich filtrada con el filtro FANS.

En la Figura 6.22 de la página 167 se muestra la imagen cociente de Munich correspondiente al filtro FANS.

Las Tablas 6.8, 6.9, 6.10, 6.8, de la página 159 resumen los coeficientes de preservación de la media y el ENL de la imagen de Munich correspondientes al filtro propuesto, al filtro SRAD, al filtro LEE-M y al filtro FANS respectivamente. En cada una de estas imágenes se han considerado las mismas  $n = 10$  regiones homogéneas para realizar la evaluación.

Finalmente en la Tabla 6.12 de la página 159 presenta la comparación de

**Tabla 6.8:** Evaluación Filtro de entropías para imagen de Munich con ventanas de  $15 \times 15$  y  $13 \times 13$

	$\hat{\mu}_U$	$\hat{\sigma}_U$	$\widehat{\text{ENL}}_U$	$\hat{\mu}_H$	$\hat{\sigma}_H$	$\widehat{\text{ENL}}_H$	$r_{\widehat{\text{ENL}}}$	$r_{\hat{\mu}}$
<i>ROI</i> <sub>1</sub>	13.4862	4.0407	11.1396	0.9840	0.2687	13.4153	0.1696	0.0160
<i>ROI</i> <sub>2</sub>	13.3525	3.9950	11.1712	0.9855	0.2720	13.1241	0.1488	0.0145
<i>ROI</i> <sub>3</sub>	16.8347	4.6650	13.0229	0.9947	0.3031	10.7727	0.2089	0.0053
<i>ROI</i> <sub>4</sub>	16.9642	4.6616	13.2434	0.9960	0.2369	17.6780	0.2509	0.0040
<i>ROI</i> <sub>5</sub>	20.9158	8.0521	6.7473	0.9813	0.3221	9.2807	0.2730	0.0187
<i>ROI</i> <sub>6</sub>	15.7209	7.5346	4.3534	0.9891	0.2828	12.2332	0.6441	0.0109
<i>ROI</i> <sub>7</sub>	13.0488	3.9085	11.1457	1.0145	0.2471	16.8493	0.3385	0.0145
<i>ROI</i> <sub>8</sub>	20.5704	7.7688	7.0110	0.9918	0.3610	7.5471	0.0710	0.0082
<i>ROI</i> <sub>9</sub>	13.2419	3.8741	11.6834	1.0000	0.2698	13.7397	0.1497	0.0000
<i>ROI</i> <sub>10</sub>	12.7489	3.3205	14.7414	1.0084	0.2794	13.0257	0.1317	0.0084

**Tabla 6.9:** Evaluación Filtro SRAD sobre imagen de Munich

	$\hat{\mu}_U$	$\hat{\sigma}_U$	$\widehat{\text{ENL}}_U$	$\hat{\mu}_H$	$\hat{\sigma}_H$	$\widehat{\text{ENL}}_H$	$r_{\widehat{\text{ENL}}}$	$r_{\hat{\mu}}$
<i>ROI</i> <sub>1</sub>	13.49	4.04	11.14	13536.66	5857.21	5.34	1.09	13535.66
<i>ROI</i> <sub>2</sub>	13.35	3.99	11.17	13677.68	5339.82	6.56	0.70	13676.68
<i>ROI</i> <sub>3</sub>	16.83	4.67	13.02	15222.19	7370.71	4.27	2.05	15221.19
<i>ROI</i> <sub>4</sub>	16.96	4.66	13.24	14788.13	7278.92	4.13	2.21	14787.13
<i>ROI</i> <sub>5</sub>	20.92	8.05	6.75	15352.31	8656.69	3.15	1.15	15351.31
<i>ROI</i> <sub>6</sub>	15.72	7.53	4.35	14971.25	5792.09	6.68	0.35	14970.25
<i>ROI</i> <sub>7</sub>	13.05	3.91	11.15	13257.44	7291.15	3.31	2.37	13256.44
<i>ROI</i> <sub>8</sub>	20.57	7.77	7.01	14530.81	6433.07	5.10	0.37	14529.81
<i>ROI</i> <sub>9</sub>	13.24	3.87	11.68	13158.65	8262.43	2.54	3.61	13157.65
<i>ROI</i> <sub>10</sub>	12.75	3.32	14.74	11858.75	6417.62	3.41	3.32	11857.75

**Tabla 6.10:** Evaluación Filtro Lee-M sobre imagen de Munich

	$\hat{\mu}_U$	$\hat{\sigma}_U$	$\widehat{\text{ENL}}_U$	$\hat{\mu}_H$	$\hat{\sigma}_H$	$\widehat{\text{ENL}}_H$	$r_{\widehat{\text{ENL}}}$	$r_{\hat{\mu}}$
<i>ROI</i> <sub>1</sub>	13.4862	4.0407	11.1396	0.9973	0.9283	1.1543	8.6503	0.0027
<i>ROI</i> <sub>2</sub>	13.3525	3.9950	11.1712	1.0511	0.9714	1.1709	8.5409	0.0511
<i>ROI</i> <sub>3</sub>	16.8347	4.6650	13.0229	0.9511	0.4936	3.7117	2.5086	0.0489
<i>ROI</i> <sub>4</sub>	16.9642	4.6616	13.2434	1.0049	0.6404	2.4617	4.3797	0.0049
<i>ROI</i> <sub>5</sub>	20.9158	8.0521	6.7473	0.9980	0.7496	1.7726	2.8064	0.0020
<i>ROI</i> <sub>6</sub>	15.7209	7.5346	4.3534	1.0510	0.7594	1.9155	1.2728	0.0510
<i>ROI</i> <sub>7</sub>	13.0488	3.9085	11.1457	0.9108	0.7267	1.5710	6.0947	0.0892
<i>ROI</i> <sub>8</sub>	20.5704	7.7688	7.0110	0.9850	0.6235	2.4954	1.8096	0.0150
<i>ROI</i> <sub>9</sub>	13.2419	3.8741	11.6834	0.9469	0.5643	2.8158	3.1493	0.0531
<i>ROI</i> <sub>10</sub>	12.7489	3.3205	14.7414	0.9735	0.5200	3.5052	3.2056	0.0265

**Tabla 6.11:** Evaluación Filtro FANS sobre imagen de Munich

	$\hat{\mu}_U$	$\hat{\sigma}_U$	$\widehat{\text{ENL}}_U$	$\hat{\mu}_H$	$\hat{\sigma}_H$	$\widehat{\text{ENL}}_H$	$r_{\widehat{\text{ENL}}}$	$r_{\hat{\mu}}$
<i>ROI</i> <sub>1</sub>	13.4862	4.0407	11.1396	0.8053	0.3928	4.2044	1.6495	0.1947
<i>ROI</i> <sub>2</sub>	13.3525	3.9950	11.1712	0.8291	0.4071	4.1482	1.6930	0.1709
<i>ROI</i> <sub>3</sub>	16.8347	4.6650	13.0229	0.8761	0.4239	4.2709	2.0493	0.1239
<i>ROI</i> <sub>4</sub>	16.9642	4.6616	13.2434	0.8597	0.4462	3.7129	2.5668	0.1403
<i>ROI</i> <sub>5</sub>	20.9158	8.0521	6.7473	0.8016	0.4175	3.6863	0.8304	0.1984
<i>ROI</i> <sub>6</sub>	15.7209	7.5346	4.3534	0.8447	0.3956	4.5599	0.0453	0.1553
<i>ROI</i> <sub>7</sub>	13.0488	3.9085	11.1457	0.7912	0.4086	3.7490	1.9730	0.2088
<i>ROI</i> <sub>8</sub>	20.5704	7.7688	7.0110	0.8169	0.3783	4.6640	0.5032	0.1831
<i>ROI</i> <sub>9</sub>	13.2419	3.8741	11.6834	0.8505	0.4083	4.3379	1.6934	0.1495
<i>ROI</i> <sub>10</sub>	12.7489	3.3205	14.7414	0.8985	0.4206	4.5629	2.2307	0.1015

resultados considerando  $g = 10$  permutaciones de cada una de las imágenes cociente correspondientes a los cuatro filtros comparados.

**Tabla 6.12:** Comparación Resultados Filtrado Imagen de Munich

	$r_{\hat{\mu}, \widehat{\text{ENL}}}$	$h_0$	$\bar{h}_g$	$\delta_h$	$\mathcal{M}$
<b>SRAD</b>	18.18	0.24	0.20	18.18	36.36
<b>Entropía</b>	0.12	0.21	0.21	2.48	2.60
<b>Lee-M</b>	2.14	0.36	0.32	12.51	14.64
<b>FANS</b>	0.84	0.18	0.18	2.28	3.13

## 6.6. Observaciones y Comentarios

De la observación de la Tabla 6.7 de la página 155 correspondiente a la evaluación de la imagen sintética, se aprecia que el Filtro de Lee-M es el que preserva mejor el valor de la media local junto con el ENL, sin embargo no logra eliminar bien el contenido geométrico de la imagen. Debido a ello SRAD logra un valor global del índice  $\mathcal{M}_0$  inferior al logrado por Lee-M. Por su parte el filtro propuesto basado en distancias de entropías, con la combinación de tamaños de ventana  $11 \times 11$  y  $7 \times 7$  elimina el contenido geométrico mejor que el filtro de Lee-M pero no supera el desempeño de SRAD.

De la Tabla 6.12 de la página 159 podemos inferir que en el caso de la imagen real correspondiente a la Ciudad de Munich, el filtro propuesto y el FANS compiten por el puesto de mejor desempeño y superan notablemente a los restantes evaluados. Sin embargo para esta combinación de tamaños de ventana  $15 \times 15$  y  $13 \times 13$  el filtro propuesto supera en desempeño al FANS dado que preserva mejor la media local y el ENL, considerando el criterio propuesto.

## 6.7. Conclusiones

Hemos propuesto un nuevo filtro de medias no local para imágenes contaminadas con ruido *speckle* utilizando el test de la distancia cuadrática entre  $h - \phi$  entropías de Shannon y Rényi. Este método de reducción de ruido *speckle* analiza la similaridad entre la ventana central y las deslizantes utilizando un estadístico de contraste basado en la distribución asintótica de las entropías.

Se ha fundamentado por qué un buen filtro preserva el valor medio de los datos transformados y mejora la relación señal-ruido en la imagen filtrada.

Posteriormente se compara el desempeño del filtro propuesto basado en la distribución asintótica de distancias entre entropías  $h - \phi$  con filtros destacados en la literatura SAR como el filtro de Lee Mejorado, el filtro de Difusión Anisotrópica SRAD y el filtro FANS.

Los criterios considerados en esta comparación se explican detalladamente. Estas cuantificaciones de la bondad de la eliminación del ruido *speckle* consideran la preservación de la media, el aumento del estimador ENL, la preservación de bordes y el análisis de la geometría de la imagen cociente.

También se utiliza el índice de homogeneidad definido a partir de la matriz de coocurrencia de niveles de gris de Haralick.

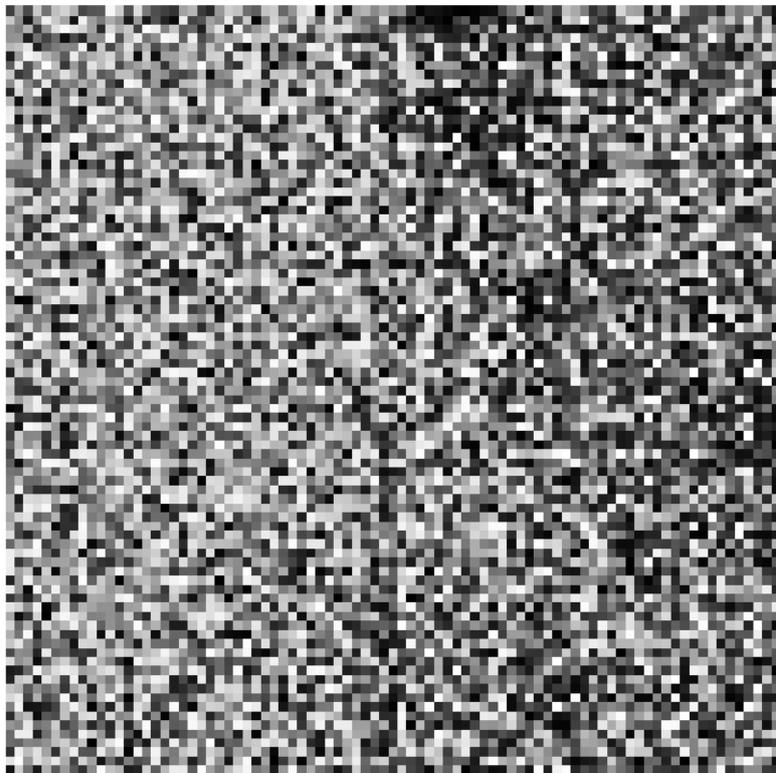
Se realiza esta comparación tanto para imágenes reales como para imágenes sintéticas, y en ambos casos el filtro propuesto mostró un desempeño muy competitivo tanto para las imágenes simuladas como para las reales.

Los resultados, en función de estos los indicadores y medidas considerados son alentadores para el filtro propuesto debido a que, además de satisfacer los criterios clásicos detallados en este capítulo, también logra un desempeño comparativo destacable con los filtros clásicos de la literatura SAR, utilizando en esta comparación criterios integrales novedosos que estudian el contenido geométrico de la imagen cociente. Asimismo se observa en los re-

---

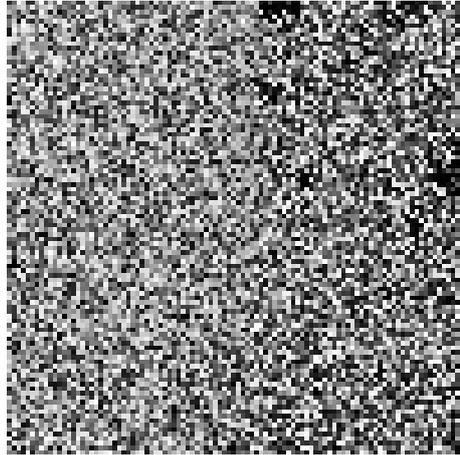
sultados de los dos experimentos que el filtrado basado en distancia cuadrática entre de entropías preserva los bordes y no produce un efecto de borronado sobre los mismos.

**Figura 6.12:** Imagen Sintética cociente con Filtro de Lee Mejorado

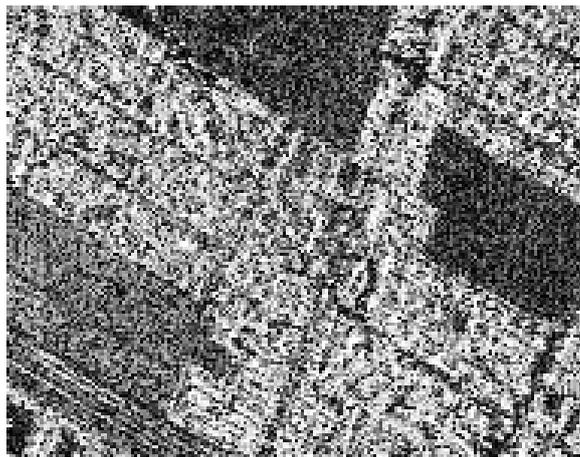


---

**Figura 6.13:** Imagen Sintética cociente con Filtro FANS



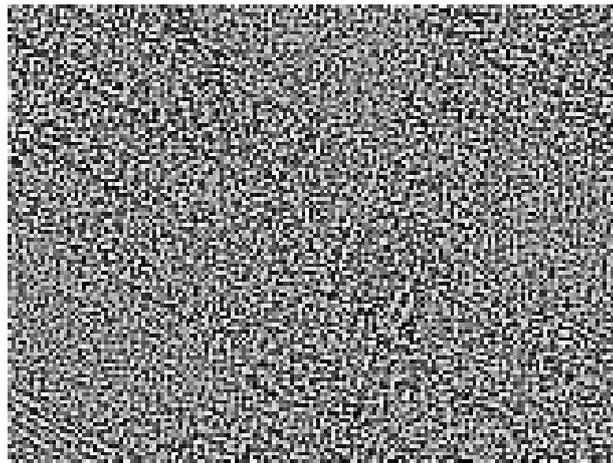
**Figura 6.14:** Imagen Real Single Look de la ciudad de Munich



**Figura 6.15:** Imagen Real de Munich filtrada con el Filtro Propuesto

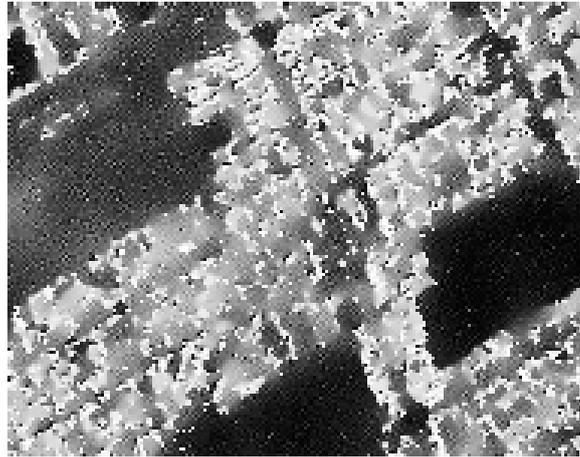


**Figura 6.16:** Imagen Real de Munich Cociente con el Filtro basado en Entropía

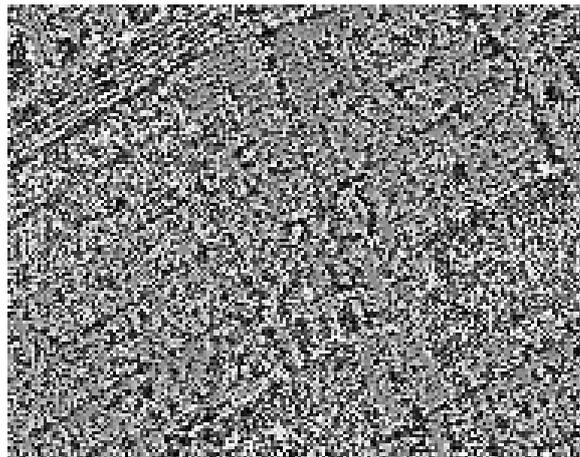


---

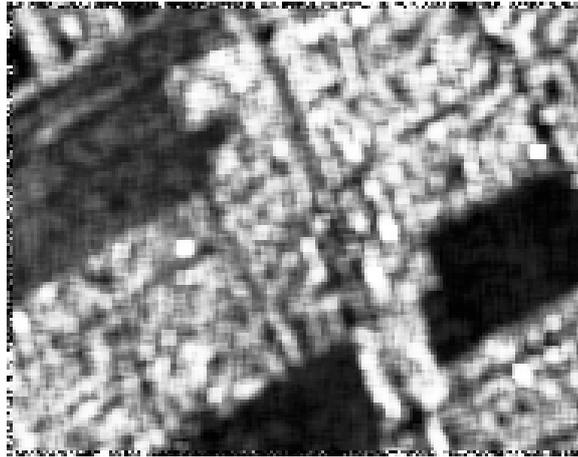
**Figura 6.17:** Imagen Real de Munich filtrada con el Filtro SRAD



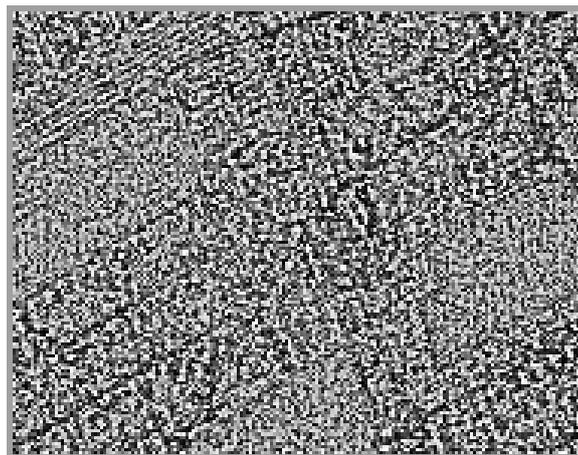
**Figura 6.18:** Imagen Real de Munich Cociente con el Filtro SRAD



**Figura 6.19:** Imagen Real de Munich filtrada con el Filtro de Lee Mejorado

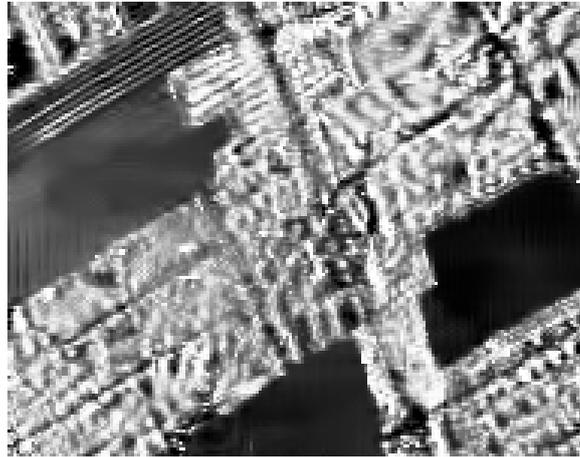


**Figura 6.20:** Imagen Real de Munich Cociente con el Filtro de Lee Mejorado

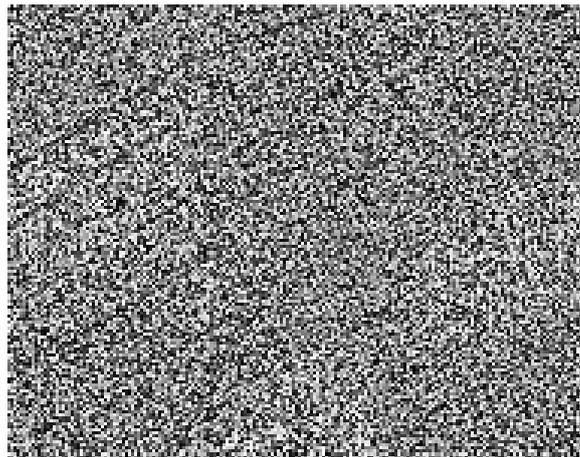


---

**Figura 6.21:** Imagen Real de Munich filtrada con el Filtro FANS



**Figura 6.22:** Imagen Real de Munich Cociente con el Filtro FANS





ParteIV

## Consideraciones Finales



## Capítulo 7

# Discusión y Conclusiones

En este trabajo se destaca la importancia de las imágenes de teledetección correspondientes al radar de apertura sintética. Se detallan sus ventajas y desventajas así como sus vastos campos de aplicación. Se propone el modelo multiplicativo para el análisis del retorno de las mismas y se presentan los modelos distribucionales para el ruido *speckle*, la retrodispersión y el retorno de las mismas.

Luego se estudian en profundidad las propiedades características de la distribución  $\mathcal{G}_I^0$  dado que este modelo ha demostrado ser versátil y eficiente para describir datos correspondientes a imágenes SAR, abarcando zonas rurales y urbanas con diferentes grados de textura.

Se focaliza el análisis en el caso más ruidoso que corresponde a  $L = 1$  y se denomina generalmente Single Look. Se caracteriza este caso aprovechando desarrollos y conceptos específicos de la teoría de valores extremos y las distribuciones de colas pesadas.

Dada la importancia en la interpretación automática de imágenes de contar con datos sintéticos de calidad, se presentan distintas formas de generar estadísticamente números aleatorios con distribución  $\mathcal{G}_I^0$  para el caso particular de Single Look, aprovechando relaciones entre distribuciones de probabilidad.

Se consideran cuatro plataformas diferentes para comparar el desempeño de los distintos modos de generación y se utilizan en esta comparación varios

criterios como la calidad de datos de la cola, las distribuciones del máximo y del mínimo de las distribuciones, la bondad de ajuste y el costo computacional de los procesos. También se estudió la influencia de la elección del umbral en estas estimaciones para cada estimador.

Se realiza este estudio inicialmente sobre la plataforma  $\mathbb{R}$  que se extiende luego a otros tres entornos computacionales: Julia, Ox y Matlab, ampliando los criterios de evaluación para la comparación.

En este análisis, considerando ciertos criterios tales como el ahorro computacional se concluye que hay generadores que aventajan a los otros en cualquiera de las plataformas consideradas, como es el caso del generador de Pareto.

Sin embargo respecto de otros criterios, tales como la bondad de ajuste, no hay generadores que sean preferibles a otros en ninguno de los cuatro entornos computacionales considerados.

Por último respecto de criterios como la precisión de los momentos se prefieren generadores diferentes en cada una de las plataformas. También se esbozaron diferencias entre las plataformas siendo el desempeño de alguna de ellas muy superior al de las otras respecto de ciertas características como la cobertura de los intervalos.

Se concluye esta parte de la investigación con una recomendación de diferencias según la plataforma en la que se decida trabajar con las imágenes.

Se destacan para el caso Single Look algunas propiedades distintivas de la distribución  $\mathcal{G}_I^0$  y se evalúa el comportamiento de un vasto conjunto de estimadores disponibles para distribuciones de valores extremos. Se realiza esta comparación en dos contextos fundamentales como son el de los datos puros y el de los datos contaminados, considerando recursos del área de la robustez.

Se utilizan como criterios de comparación, el sesgo, la varianza, la existencia, la convergencia, la estabilidad, la robustez y el comportamiento asintótico de los estimadores. Basados en estos criterios algunos estimadores son descon-

---

siderados en los últimos análisis.

Dada la presencia de distintos esquemas de contaminación propios de los datos provenientes de imágenes SAR y especialmente del entorno de las muestras correspondientes a las ventanas deslizantes de los procesos de filtrado, se focaliza el análisis de estos estimadores en cuanto a su robustez utilizando criterios propios de este campo y criterios específicos de las imágenes SAR.

Entre los esquemas de contaminación propios de estas imágenes, se estudia el caso de los reflectores de esquina. Se presentan resultados y aplicaciones del desempeño de estos estimadores contextualizados para imágenes sintéticas y para imágenes reales.

Los resultados de estos análisis se aplican a la segmentación de imágenes sintéticas y reales corrompidas con un corner reflector, destacando las bondades en este contexto de los estimadores mDD y mPP.

A continuación se deducen en forma teórica las expresiones correspondientes a las entropías de Shannon y de Rényi para la distribución  $\mathcal{G}_I^0$ . También se desarrolla la expresión de la matriz de Información de Fisher y la del vector de scores.

Se deriva luego la distribución asintótica de las mismas y se propone una prueba basada en la distancia cuadrática entre entropías. Esta prueba se evalúa en su comportamiento y se compara en términos de sensibilidad, especificidad, valor predictivo positivo y valor predictivo negativo con un test clásico.

A continuación se propone un proceso de reducción del ruido *speckle* que aprovecha en la construcción de la máscara de convolución los  $p$ -valores del test de distancia cuadrática entre entropías  $h-\phi$ .

Con el objetivo de ajustar el desempeño de esta máscara se aplica una transformación logística de los  $p$ -valores de la prueba.

Finalmente se analiza el desempeño de este filtro utilizando criterios novedosos tanto en imágenes sintéticas y como en imágenes reales. Se aprovechan en este análisis la homogeneidad de las imágenes definida a partir de matriz

de coocurrencia de niveles de gris de Haralick, la imagen cociente y el análisis de correlación.

Se destaca la ventaja de este procedimiento de eliminación de ruido *speckle* comparando su desempeño con el de filtros bien conocidos y de amplia difusión en la literatura de imágenes SAR como el filtro de Lee Mejorado, el filtro de difusión anisotrópica SRAD y el filtro FANS.

Los resultados obtenidos son muy alentadores. En futuros trabajos, se prevé la extensión de estos resultados para la distribución  $\mathcal{G}_l^0$  en el caso de  $L > 1$ , es decir correspondiente a imágenes multi-look y también para imágenes polarimétricas.

También se estudiará el comportamiento de este filtrado para el caso multilook, lo cual requerirá previamente la deducción de las expresiones de las entropías, los vectores de scores y la matriz de información de Fisher para el caso más general.

En los nuevos estudios se tratarán también los restantes esquemas de contaminación característicos de las imágenes SAR, tanto para el caso single look como para el caso multilook.

## Bibliografía

- [1] I. B. Aban and M. M. Meerschaert. Shifted Hill's estimator for heavy tails. *Communications in Statistics-Simulation and Computation*, 30(4):949–962, 2001.
- [2] A. Achim, E. Kuruoglu, and J. Zerubia. SAR image filtering based on the heavy-tailed rayleigh model. *IEEE Transactions on Image Processing*, 15(9):2686–2693, 2006.
- [3] H. Allende, A. C. Frery, J. Galbiati, and L. Pizarro. M-estimators with asymmetric influence functions: the distribution case. *Journal of Statistical Computation and Simulation*, 76(11):941–956, 2006.
- [4] S. N. Anfinsen and T. Eltoft. Analysis of multilook polarimetric radar data with the matrix-variate Mellin transform. In *Synthetic Aperture Radar (EUSAR), 2010 8th European Conference on*, pages 1–4. VDE, 2010.
- [5] F. Argenti, A. Lapini, T. Bianchi, and L. Alparone. A tutorial on speckle reduction in synthetic aperture radar images. *IEEE Geoscience and Remote Sensing Magazine*, 1(3):6–35, 2013.
- [6] S. Azimi, R. Mahmood, and S. Bannour. Two-dimensional adaptive block Kalman filtering of SAR imagery. *IEEE Transactions on Geoscience and Remote Sensing*, 29(5):742–753, 1991.

- [7] M. Barber, F. Grings, P. Perna, M. Piscitelli, M. Maas, C. Bruscantini, J. Jacobo-Berlles, and H. Karszenbaum. Speckle noise and soil heterogeneities as error sources in a Bayesian soil moisture retrieval scheme for SAR data. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 5(3):942–951, June 2012.
- [8] J. Bezanson, A. Edelman, S. Karpinski, and V. Shah. *Julia: A Fresh Approach to Numerical Computing*. Cornell University Library, 2014.
- [9] A. Buades, B. Coll, and J. M. Morel. A review of image denoising algorithms, with a new one. *Multiscale Modeling & Simulation*, 4(2):490–530, 2005.
- [10] F. Caeiro, M. I. Gomes, and D. Pestana. Direct reduction of bias of the classical Hill estimator. *Revstat*, 3(2):113–136, 2005.
- [11] S. Canale, A. De Santis, D. Iacoviello, F. Pirri, and S. Sagratella. Integrating X-SAR images and anthropic factors for fire susceptibility assessment. In *IEEE International Geoscience and Remote Sensing Symposium*, pages 818–821, 2011.
- [12] J. Canny. A computational approach to edge detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1(6):679–698, 1986.
- [13] J. Casseti, J. Gambini, and A. C. Frery. Parameter estimation in SAR imagery using stochastic distances. In *Conference on Asia-Pacific Synthetic Aperture Radar (APSAR)*, pages 573–576. IEEE, 2013.
- [14] 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 *IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, pages 3266–3269, 2016.

- 
- [15] D. Chan, J. Gambini, and A.C. Frery. Speckle noise reduction in SAR images using Information Theory. In *2020 IEEE Latin American GRSS & ISPRS Remote Sensing Conference (LAGIRS)*, pages 456–461, mar 2020.
- [16] D. Chan, A. Rey, J. Gambini, J. Casseti, and A. C. Frery. Methods and frameworks for sampling  $\mathcal{G}_I^0$  data. In *IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, pages 3003–3006, 2017.
- [17] D. Chan, A. Rey, J. Gambini, and A. C. Frery. Sampling from the  $\mathcal{G}_I^0$  distribution. *Monte Carlo Methods and Applications*, 2018.
- [18] D. Chan, A. Rey, J. Gambini, and A. C. Frery. Low-cost robust estimation for the single-look  $\mathcal{G}_I^0$  model using the pareto distribution. *IEEE Geoscience and Remote Sensing Letters*, 2019.
- [19] S. W. Chen and M. Sato. Tsunami damage investigation of built-up areas using multitemporal spaceborne full polarimetric SAR images. *IEEE Transactions on Geoscience and Remote Sensing*, 51(4):1985–1997, 2013.
- [20] S. G. Coles and M. J. Dixon. Likelihood-based inference for extreme value models. *Extremes*, 2(1):5–23, 1999.
- [21] D. R. Cox and E. J. Snell. A general definition of residuals. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 248–275, 1968.
- [22] D. Cozzolino, S. Parrilli, G. Scarpa, G. Poggi, and L. Verdoliva. Fast adaptive nonlocal SAR despeckling. *IEEE Transactions on Geoscience and Remote Sensing Letters*, 11:524–528, 2014.
- [23] 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.

- [24] T. R. Crimmins. Geometric filter for speckle reduction. *Applied Optics*, 24(10):1438–1443, 1985.
- [25] C. da Costa Freitas, A. H. Correia, A. C. Frery, and S. J. S. Sant Ana. A system for multilook polarimetric SAR image statistical classification. *Anais IX Simposio Brasileiro de Sensoriamento Remoto, Santos, Brasil*, pages 1063–1074, 1998.
- [26] F. Dell’Acqua and P. Gamba. Remote sensing and earthquake damage assessment: Experiences, limits, and perspectives. *Proceedings of the IEEE*, 100(10):2876–2890, 2012.
- [27] D. Dellamonica, P. J. S. Silva, C. Humes, N. S. T. Hirata, and J. Barrera. An exact algorithm for optimal MAE stack filter design. *IEEE Transactions on Image Processing*, 16(2):453–462, 2007.
- [28] L. Devroye. *Non-Uniform Random Variate Generation*. Springer Verlag, 1986.
- [29] J. A. Doornik and M. Ooms. *Introduction to Ox*. Timberlake Consultants Press, London, 2006.
- [30] R. O. Duda, P. E. Hart, and D. G. Stork. *Pattern classification*. John Wiley & Sons, 2012.
- [31] E. M. El-Desouki, K. F. A. Hussien, and H. M. El-Hennawy. Electromagnetic simulation for land imaging using fully polarimetric SAR system. In *34th National Radio Science Conference*, pages 132–141, 2017.
- [32] T. Esch, M. Schmidt, M. Breunig, A. Felbier, H. Taubenböck, W. Heldens, C. Riegler, A. Roth, and S. Dech. Identification and characterization of urban structures using VHR SAR data. In *IEEE International on Geoscience and Remote Sensing Symposium (IGARSS)*, pages 1413–1416, 2011.

- 
- [33] M. Ferreira da Silva, F. Cribari-Neto, and A. C. Frery. Improved likelihood inference for the roughness parameter of the  $\mathcal{G}_A^0$  distribution. *Environmetrics: The Official Journal of the International Environmetrics Society*, 19(4):347–368, 2008.
- [34] D. Firth. Bias reduction of maximum likelihood estimates. *Biometrika*, 80(1):27–38, 1993.
- [35] A. C. Frery, F. Cribari-Neto, and M. O. de Souza. Analysis of minute features in speckled imagery with maximum likelihood estimation. *EURASIP Journal on Advances in Signal Processing*, 2004(16):375370, 2004.
- [36] A. C. Frery, H. Müller, C. Yanasse, and S. Sant’Anna. A model for extremely heterogeneous clutter. *IEEE Transactions on Geoscience and Remote Sensing*, 35(3):648–659, 1997.
- [37] M. Friedman. The use of ranks to avoid the assumption of normality implicit in the analysis of variance. *Journal of the American Statistical Association*, 32(200):675–701, 1937.
- [38] V. S. Frost, J. A. Stiles, K. S. Shanmugan, and J. C. Holtzman. A model for radar images and its application to adaptive digital filtering of multiplicative noise. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, I(2):157–166, 1982.
- [39] J. Gambini, J. Cassetti, M. Lucini, and A. C. Frery. Parameter estimation in SAR imagery using stochastic distances and asymmetric kernels. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 8(1):365–375, 2015.
- [40] J. Gambini, M. Mejail, J. Jacobo Berllés, and A. C. Frery. Accuracy of local edge detection in speckled imagery. *Statistics & Computing*, 18(1):15–26, 2008.

- [41] G. Gao. Statistical modeling of SAR images: A survey. *Sensors*, 10(1):775–795, 2010.
- [42] D. E. Giles, H. Feng, and R. T. Godwin. Bias-corrected maximum likelihood estimation of the parameters of the generalized Pareto distribution. *Communications in Statistics-Theory and Methods*, 45(8):2465–2483, 2016.
- [43] E. Girón, A. C. Frery, and F. Cribari-Neto. Nonparametric edge detection in speckled imagery. *Mathematics and Computers in Simulation*, 82(11):2182–2198, 2012.
- [44] D. Gleich, M. Kseneman, and M. Datcu. Despeckling of Terra SAR-X data using second-generation wavelets. *IEEE Geoscience and Remote Sensing Letters*, 7(1):68–72, 2010.
- [45] L. Gomez, M. E. Buemi, J. C. Jacobo-Berlles, and M. E. Mejail. A new image quality index for objectively evaluating despeckling filtering in SAR images. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 9(3):1297–1307, 2015.
- [46] L. Gomez, C. G. Munteanu, M. E. Buemi, J. C. Jacobo-Berlles, and M. E. Mejail. Supervised constrained optimization of bayesian nonlocal means filter with sigma preselection for despeckling SAR images. *IEEE Transactions on Geoscience and Remote Sensing*, 51(8):4563–4575, 2013.
- [47] L. Gomez, R. Ospina, and A. C. Frery. Unassisted quantitative evaluation of despeckling filters. *Remote Sensing*, 9(4):389, 2017.
- [48] S. D. Grimshaw. Computing maximum likelihood estimates for the generalized Pareto distribution. *Technometrics*, 35(2):185–191, 1993.

- 
- [49] P. Guccione, L. Mascolo, G. Nico, A. Pelusi, and M. Zonno. SAR image simulation of ocean environment and detection of oil slicks. In *10th European Conference on Synthetic Aperture Radar*, pages 1–4, 2014.
- [50] F. R. Hampel, E. M. Ronchetti, P. J. Rousseeuw, and W. A. Stahel. *Robust statistics: the approach based on influence functions*, volume 196. John Wiley & Sons, 2011.
- [51] Frank R Hampel. The influence curve and its role in robust estimation. *Journal of the american statistical association*, 69(346):383–393, 1974.
- [52] X. Hao, S Gao, and X. Gao. A novel multiscale nonlinear thresholding method for ultrasonic speckle suppressing. *IEEE Transactions on Medical Imaging*, 18(9):787–794, 1999.
- [53] R. M. Haralick, K. Shanmugam, and I. Dinstein. Textural features for image classification. *IEEE Transactions on Systems, Man, and Cybernetics*, SMC 3(6):610–621, 1973.
- [54] B. M. Hill. A simple general approach to inference about the tail of a distribution. *The Annals of Statistics*, pages 1163–1174, 1975.
- [55] J. R.M. Hosking and J. R. Wallis. Parameter and quantile estimation for the generalized Pareto distribution. *Technometrics*, 29(3):339–349, 1987.
- [56] J. C. Jacobo Berlles. *Nuevas familias de distribuciones polarimétricas para imágenes SAR*. PhD thesis, Universidad de Buenos Aires Facultad de Ciencias Exactas y Naturales, 2005.
- [57] S. F. Juárez and W. R. Schucany. Robust and efficient estimation for the generalized Pareto distribution. *Extremes*, 7(3):237–251, 2004.

- [58] S. Khan and R. Guida. On Single-Look Multivariate  $\mathcal{G}$  Distribution for PolSAR Data. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 5(4):1149–1163, 2012.
- [59] M. Kiemer and H. Breit. Efficient evaluation of multichannel SAR data recombination filters. *IEEE Transactions on Geoscience and Remote Sensing*, 55:6277–6286, 2017.
- [60] J. H.T. Kim and J. Kim. A parametric alternative to the Hill estimator for heavy-tailed distributions. *Journal of Banking & Finance*, 54:60–71, 2015.
- [61] W.H. Kruskal and W. A. Wallis. Use of ranks in one-criterion variance analysis. *Journal of the American Statistical Association*, 47(260):583–621, 1952.
- [62] D. T. Kuan, A. Sawchuk, T. Strand, and P. Chavel. Adaptive restoration of images with speckle. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 35(3):373–383, 1987.
- [63] D. T. Kuan, A. A. Sawchuk, T. C. Strand, and P. Chavel. Adaptive noise smoothing filter for images with signal-dependent noise. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, I(2):165–177, 1985.
- [64] J. S. Lee. Digital image enhancement and noise filtering by use of local statistics. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 2:165–168, 1980.
- [65] J. S. Lee, L. Jurkevich, P. Dewaele, P.I. Wambacq, and A. Oosterlinck. Speckle filtering of synthetic aperture radar images: A review. *Remote sensing reviews*, 8(4):313–340, 1994.
- [66] J. S. Lee, J. H. Wen, T. L. Ainsworth, K. S. Chen, and A. J. Chen. Improved sigma filter for speckle filtering of SAR imagery. *IEEE Transactions on Geoscience and Remote Sensing*, 47(1):202–213, 2009.

- 
- [67] X. Li, H. Sun, and X Xu. A novel speckle filtering technique for SAR image. In *6th International Conference on Signal Processing, 2002.*, volume 1, pages 788–791. IEEE, 2002.
- [68] J. H. Lin and Y. T. Kim. Fast algorithms for training stack filters. *IEEE Transactions on Signal Processing*, 42(4):772–781, 1994.
- [69] J. H. Lin, T. M. Sellke, and E. J. Coyle. Adaptive stack filtering under the mean absolute error criterion. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 38(6):938–954, 1990.
- [70] A. Lopes, E. Nezry, R. Touzi, and H. Laur. Maximum a posteriori speckle filtering and first order texture models in SAR images. In *Geoscience and Remote Sensing Symposium, 1990. IGARSS'90. Remote Sensing Science for the Nineties'. 10th Annual International*, pages 2409–2412. Ieee, 1990.
- [71] A. Luceño. Fitting the generalized Pareto distribution to data using maximum goodness-of-fit estimators. *Computational Statistics & Data Analysis*, 51(2):904–917, 2006.
- [72] M. M. Lucini, A. C. Frery, and O. H. Bustos. Un método robusto para la clasificación de imágenes de radar SAR. *UNNE Comunicaciones Científicas y Tecnológicas*, 2005.
- [73] MathWorks. *MatLab, The Language of technical Computing*. The MathWorks Inc., 2019.
- [74] M. Mejail, J. J. Berlles, A. C. Frery, and O. Bustos. *Parametric Roughness Estimation in Amplitude: SAR Images Under the Multiplicative Model*. Universidad Nacional de Cordoba Instituto de Matematica, Astronomia y Fisica, 1998.

- [75] 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.
- [76] O. Monserrat, M. Crosetto, and G. Luzi. A review of ground-based SAR interferometry for deformation measurement. *ISPRS Journal of Photogrammetry and Remote Sensing*, 93:40–48, jul 2014.
- [77] A. Moreira, P. Prats-Iraola, M. Younis, G. Krieger, I. Hajnsek, and K. P. Papathanassiou. A tutorial on SyntheticAperture Radar. *IEEE Geoscience and Remote Sensing Magazine*, 1(1):6–43, 2013.
- [78] J. Naranjo-Torres, J. Gambini, and A. C. Frery. The geodesic distance between  $\mathcal{G}_I^0$  models and its application to region discrimination. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 10(3):987–997, 2017.
- [79] A. D. C. Nascimento, M. M. Horta, A. C. Frery, and R. J. Cintra. Comparing edge detection methods based on stochastic entropies and distances for PolSAR imagery. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 7(2):648–663, Feb. 2014.
- [80] M. E. Newman. Power laws, Pareto distributions and Zipf’s law. *Contemporary Physics*, 46(5):323–351, 2005.
- [81] J. M. Nicolas and S. N. Anfinsen. Introduction to second kind statistics: Application of log-moments and log-cumulants to the analysis of radar image distributions. *Trait. Signal*, 19(3):139–167, 2002.
- [82] J. Nuyts. Inference about the tail of a distribution: Improvement on the Hill estimator. *International Journal of Mathematics and Mathematical Sciences*, 2010, 2010.

- 
- [83] C. Oliver and S. Quegan. *Understanding Synthetic Aperture Radar images*. SciTech Publishing, 2004.
- [84] J. A. Parra Plazas, J. Hernández T. Peñaloza, and F. de la Rosa. Estudio de reducción del ruido en imágenes SAR para el análisis de inundaciones. *Ingenium*, 16(32):9–18, 2015.
- [85] L. Peng and A. H. Welsh. Robust estimation of the generalized Pareto distribution. *Extremes*, 4(1):53–65, 2001.
- [86] P. Perona, T. Shiota, and J. Malik. Anisotropic diffusion. In *Geometry-driven diffusion in Computer Vision*, pages 73–92. Springer, 1994.
- [87] J. Pickands. Statistical inference using extreme order statistics. *Annals of Statistics*, pages 119–131, 1975.
- [88] J. F. Pires, A. H. M. A. Cysneiros, and F. Cribari-Neto. Improved inference for the generalized Pareto distribution. *Brazilian Journal of Probability and Statistics*, 32(1):69–85, feb 2018.
- [89] L. J. Porcello, N. G. Massey, R. B. Innes, and J. M. Marks. Speckle reduction in Synthetic-Aperture Radars. *Journal of the Optical Society of America*, 66(11):1305–1311, 1976.
- [90] W. K. Pratt. *Introduction to Digital Image Processing*. CRC Press, 2013.
- [91] Massey J. R. and Frank J. The Kolmogorov-Smirnov test for goodness of fit. *Journal of the American Statistical Association*, 46(253):68–78, 1951.
- [92] R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2020.
- [93] N. M. Razali and Y. B. Wah. Power comparisons of Shapiro-Wilk, Kolmogorov-Smirnov, Lilliefors and Anderson Darling tests. *Journal of Statistical Modeling and Analytics*, 2(1):21–33, 2011.

- [94] R. A. Redner and H. F. Walker. Mixture densities, maximum likelihood and the EM algorithm. *SIAM Review*, 26(2):195–239, 1984.
- [95] J. Rojo. Heavy-tailed densities. *Wiley Interdisciplinary Reviews: Computational Statistics*, 5(1):30–40, 2013.
- [96] R. Roy. Comparison of different techniques to generate normal random variables. *Stochastic Signals and Systems (ECE 330: 541)*, Rutgers, The State University of New Jersey, 2002.
- [97] F. Safa and G. Flouzat. Speckle removal on radar imagery based on mathematical morphology. *Signal Processing*, 16(4):319–333, 1989.
- [98] M. Salicrú, D. Morales, M. L. Menéndez, and L. Pardo. On the applications of divergence type measures in testing statistical hypotheses. *Journal of Multivariate Analysis*, 51(2):372–391, 1994.
- [99] S. J. S. Sant’Anna. Avaliação do desempenho de filtros redutores de speckle em imagens de radar de abertura sintética. Master’s thesis, Instituto Nacional de Pesquisas Espaciais, 1995.
- [100] M. Sato, S. W. Chen, and M. Satake. Polarimetric SAR analysis of tsunami damage following the march 11, 2011 East Japan earthquake. *Proceedings of the IEEE*, 100(10):2861–2875, 2012.
- [101] N. Saxena and N. Rathore. A review on speckle noise filtering techniques for SAR images. *International Journal of Advanced Research in Computer Science and Electronics Engineering (IJARCSEE)*, 2(2):243–247, 2013.
- [102] I. Shanthi and M.L. Valarmathi. Speckle noise suppression of sar image using hybrid order statistics filters. *International Journal of Advanced Engineering Sciences and Technologies*, 5(2):229–235, 2011.

- 
- [103] Francescopaolo Sica, Davide Cozzolino, Xiao Xiang Zhu, Luisa Verdoliva, and Giovanni Poggi. Insar-bm3d: A nonlocal filter for sar interferometric phase restoration. *IEEE Transactions on Geoscience and Remote Sensing*, 56(6):3456–3467, 2018.
- [104] C. Tison, J. M. Nicolas, F. Tupin, and H. Maître. A new statistical model for Markovian classification of urban areas in high-resolution SAR images. *IEEE Transactions on Geoscience and Remote Sensing*, 42(10):2046–2057, 2004.
- [105] L. Torres, S. J. S. Sant’Anna, C. da Costa Freitas, and A. C. Frery. Speckle reduction in polarimetric SAR imagery with stochastic distances and nonlocal means. *Pattern Recognition*, 47(1):141–157, 2014.
- [106] R. Touzi, A. Lopes, and P. Bousquet. A statistical and geometrical edge detector for SAR images. *IEEE Transactions on Geoscience and Remote Sensing*, 26(6):764–773, 1988.
- [107] F. T. Ulaby. Radar remote sensing and surface scattering and emission theory. *Microwave Remote Sensing: Active and Passive*, 1982.
- [108] K. Vasconcellos and A. C. Frey. Maximum likelihood fitting of extremely heterogeneous radar clutter. In *Image Processing Techniques, First Latino-American Seminar on Radar Remote Sensing*, volume 407, page 97, 1997.
- [109] K. L. P. Vasconcellos, A. C. Frery, and L. B. Silva. Improving estimation in speckled imagery. *Computational Statistics*, 20(3):503–519, 2005.
- [110] D. Velotto, S. Lehner, A. Soloviev, and C. Maingot. Analysis of oceanic features from dual-polarization high resolution X-band SAR imagery for oil spill detection purposes. In *IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, pages 2841–2844, 2012.

- [111] J. A. Villaseñor-Alva and E. González-Estrada. A bootstrap goodness of fit test for the Generalized Pareto Distribution. *Computational Statistics & Data Analysis*, 53(11):3835–3841, 2009.
- [112] C. H. Wang, X. B. Wen, and H. X. Xu. A robust estimator of parameters for  $\mathcal{G}_I^0$ - modeled SAR imagery based on random weighting method. *EURASIP Journal on Advances in Signal Processing*, 2017(1):22, 2017.
- [113] Z. Wang, A. C. Bovik, and L. Lu. Why is image quality assessment so difficult? In *ICASSP*, volume 4, pages 3313–3316, 2002.
- [114] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli. Image quality assessment: from error visibility to structural similarity. *IEEE Transactions on Image Processing*, 13(4):600–612, 2004.
- [115] K. D. Ward, S. Watts, and R. J. A. Tough. *Sea clutter: scattering, the K distribution and radar performance*, volume 20. IET, 2006.
- [116] Q. Wu, R. Chen, H. Sun, and Y. Cao. Urban building density detection using high resolution SAR imagery. In *Joint Urban Remote Sensing Event*, pages 45–48, 2011.
- [117] Y. Yamaguchi. Disaster monitoring by fully Polarimetric SAR data acquired with ALOS-PALSAR. *Proceedings of the IEEE*, 100(10):2851–2860, 2012.
- [118] L. P. Yaroslavsky. *Digital picture processing: an introduction*, volume 9. Springer Science & Business Media, 2012.
- [119] Y. Yu and S. T. Acton. Speckle reducing anisotropic diffusion. *IEEE Transactions on Image Processing*, 11(11):1260–1270, 2002.

- 
- [120] D.-X. Yue, F. Xu, A. C. Frery, and Y.-Q. Jin. A generalized Gaussian coherent scatterer model for correlated SAR texture. *IEEE Transactions on Geoscience and Remote Sensing*, 58(4):2947–2964, April 2020.
- [121] J. Zhang. Likelihood moment estimation for the generalized Pareto distribution. *Australian & New Zealand Journal of Statistics*, 49(1):69–77, 2007.
- [122] W. Zhang, F. Liu, L. Jiao, B. Hou, S. Wang, and R. Shang. SAR image despeckling using edge detection and feature clustering in bandelet domain. *IEEE Geoscience and Remote Sensing Letters*, 7(1):131–135, 2010.
- [123] D. Zhong and S. Chang. Long term moving object segmentation and tracking using spatio-temporal consistency. In *International Conference on Image Processing (ICIP)*, 2001.