

PROCESAMIENTO DE SEÑALES DE RADAR EN PRESENCIA DE CLUTTER DINÁMICO

Tesis de Doctorado

Juan Pablo PASCUAL

Presentada ante la Facultad de Ingeniería de la
Universidad Nacional de La Plata



como requisito para la obtención del grado académico de
DOCTOR EN INGENIERÍA

Dirección de tesis:

Director: Dr. Nicolás von Ellenrieder

Co-director: Dr. Carlos Horacio Muravchik

Jurado de tesis:

Dr. Giorgio M. Caranti

Dr. Isidoro Vaquila

Dr. Juan E. Cousseau

Fecha de la defensa oral y pública:

19 de diciembre de 2014

Instituto de Investigaciones en Electrónica, Control y Procesamiento de Señales-LEICI

Dto. de Electrotecnia, Facultad de Ingeniería, UNLP

Resumen

El radar es un sistema de sensado remoto que utiliza técnicas de procesamiento estadístico de señales para obtener información de la señal recibida. Los radares convencionales son sistemas activos que operan transmitiendo energía en forma de ondas electromagnéticas y recibiendo las señales reflejadas por el entorno y el objeto iluminado.

Una de las dificultades a tener en cuenta en los sistemas de radar es que la señal de interés suele encontrarse oscurecida por las reflexiones producidas por el ambiente, fenómeno al que se denomina *clutter*. Generalmente, y dependiendo de la aplicación, el clutter es considerado una fuente de interferencia y perturbaciones cuyos efectos se deben eliminar o reducir. Por lo tanto, en vista de su naturaleza aleatoria, es importante el desarrollo de métodos estadísticos de procesamiento de señales para poder detectar objetivos y estimar sus propiedades en situaciones de clutter intenso y dinámico. Para obtener algoritmos eficientes, es fundamental utilizar modelos realistas de las señales recibidas por el radar. Estos modelos deben enfatizar las diferencias entre el objeto de interés y el clutter. De esta forma, los métodos de procesamiento de señal son usados para separar el objetivo del clutter y reducir el efecto degradante de este último.

En esta tesis se aborda el problema de detección en presencia de clutter dinámico para aplicaciones de radar. En especial se desarrollan modelos que contemplan las variaciones del escenario y utilizan la historia del clutter para mejorar su caracterización en el instante actual y las predicciones a tiempo futuro. La primera alternativa considera el clutter como una serie temporal que presenta heteroscedasticidad condicional autorregresiva generalizada, utilizando los denominados procesos GARCH. Este tipo de procesos poseen la característica de ser impulsivos, pero presentan la desventaja de que no cuentan con una expresión explícita para su función densidad de probabilidad. Por este motivo, se analizan alternativas para estimar sus parámetros y determinar la calidad de la estimación. Asimismo, se adaptan los test de hipótesis usuales para deducir un esquema de detección basado en el modelo GARCH.

Con el fin de incorporar información de múltiples pulsos en los instantes de decisión, se extiende el modelo anterior combinando un proceso GARCH en dos dimensiones (GARCH-2D) con un proceso autorregresivo (AR) y se deriva el detector correspondiente para este modelo de clutter. La parte GARCH-2D del modelo preserva la propiedad impulsiva de los procesos GARCH y la AR en las innovaciones permite modelar la correlación pulso a pulso que existe en los datos.

En ambos casos se deducen expresiones para las probabilidades de falsa alarma y, dada su complejidad matemática, la probabilidad de detección se evalúa por medio de simulaciones numéricas. Además, se analiza la sensibilidad del desempeño de los detectores ante errores en la estimación de sus parámetros. A pesar de que no resultan de tasa de falsa alarma constante, muestran un comportamiento robusto en situaciones prácticas. Por último, el desempeño de los detectores propuestos es comparado con algoritmos de detección existentes en la literatura utilizando mediciones reales de clutter marítimo. Los resultados muestran que presentan un mejor desempeño respecto de los demás detectores, es decir, una probabilidad de detección mayor para una tasa de falsa alarma menor, independientemente de la relación señal a clutter.

Finalmente se estudia el problema de estimación secuencial de los parámetros de los procesos GARCH. Si bien de los análisis de sensibilidad se concluye que en los detectores propuestos no es necesaria una actualización frecuente de los mismos, su estimación es la etapa de mayor costo computacional en los esquemas de detección propuestos. Siguiendo el enfoque de estimación Bayesiano se deduce un estimador lineal de mínimo error cuadrático medio para la varianza condicional de los procesos GARCH, que es el parámetro del cual depende el estadístico de los detectores desarrollados. La deducción del algoritmo es análoga a la del filtro de Kalman, pero en este caso las matrices del sistema son aleatorias.

Abstract

Radar is a remote sensing system that uses techniques of statistical signal processing to obtain information from the received signal. The most common radars are active systems that radiate an electromagnetic wave into space and receive the echo signals reflected from a target and its environment.

One of the difficulties to consider in radar systems is that the signal of interest is obscured by reflections from the environment, a phenomenon called *clutter*. In general, and depending on the application, clutter is considered an interference source whose effects should be eliminated or mitigated. Due to its random nature it is important to develop statistical signal processing methods for target detection and for estimation of its properties in intense and dynamic clutter situations. Realistic modeling of the received signals is essential to obtain efficient algorithms. These models should emphasize the differences between the target and the clutter. Thus, signal processing methods are employed to detect target echoes in presence of clutter.

In this thesis we address the detection problem in the presence of dynamic clutter for radar applications. Specifically, we develop models which consider the environment changes and use the process history to improve their characterization at current time and for future predictions. First, we propose modeling the clutter as a time series, using Generalized Autoregressive Conditional Heteroscedastic (GARCH) processes. One of their main characteristic of these models is their heavy tailed probability density function, but they have the disadvantage that an explicit expression for their probability density function does not exist. For this reason, we analyze alternatives to estimate their parameters and to determine the estimates quality. Moreover, we adapt a conventional hypothesis test to derive a detection scheme based on the GARCH model.

In addition, we extend the previous ideas by proposing a second clutter model that combines a GARCH-2D process with an AR process to use the information of multiple radar pulses at each decision instant. The GARCH-2D part of the model preserves the impulsivity property of the GARCH processes, i.e., it preserves the heavy tailed pdf, and the AR model in the innovations lets us model pulsewise correlation. In both cases, we carry out a theoretical analysis to determine its false alarm probability and we compute by means of numerical simulations the detection probability to evaluate the detectors performance. We also analyze how the error in the estimation of their parameters affect their performance. Although the resulting detectors are not of constant false alarm rate, we show that they are very robust with respect to the pro-

cess coefficients in practical situations. We compare the performance of the proposed detectors with the performance of other detection algorithms in the literature using the sea radar data measurements. The results show that our detectors outperform the others detectors, i.e. our method results in a higher detection probability for a lower false alarm rate, independently of the value of signal to clutter ratio.

Finally, we study the sequential estimation problem of the GARCH processes parameters. While from the sensitivity analysis we conclude the coefficients estimates of the proposed detectors do not need to be updated often, the estimation of the GARCH process parameters is the stage of greatest computational cost of the proposed detection schemes. Since the statistics of the proposed detectors depend on the conditional variance of the GARCH process, we derive a linear mean square error estimator for it using a Bayesian estimation approach. The derivation parallels the Kalman filter, but in this case the system matrices are random matrices.

Agradecimientos

A Euge por acompañarme durante este largo proceso, por ayudarme y animarme, por su paciencia para que terminara el trabajo, permitiendo que la tesis durma entre nosotros en más de una oportunidad.

Quiero agradecer a mis padres, por su esfuerzo y por brindarme todo para mi realización personal; y también a mis hermanos por su apoyo inestimable en cualquier tipo de circunstancia.

A mi director Nicolás von Ellenrieder, por la confianza que depositó en mí, su comprensión y por la libertad que me dio para trabajar en la tesis. Quiero agradecer también a mi co-director Carlos Muravchik, a quien debo la mayor parte de mi formación en el área del procesamiento de señales. La guía de ambos y en especial su paciencia, teniendo en cuenta mi temperamento, fueron una parte fundamental para la conclusión de este trabajo.

Otros se han mostrado muy generosos conmigo y con esta tesis. Quiero agradecer a Agustín que en muchas oportunidades me ofreció voluntariamente su tiempo para ayudarme en la investigación necesaria para llevar adelante este trabajo y ha sido un sostén anímico durante esta etapa. Quiero expresar mi gratitud también a Martín que me asistió de modo incondicional a la hora de corregir y revisar cada uno de los trabajos que dieron lugar a esta tesis.

A todos y cada uno de los integrantes del grupo de Procesamiento Estadístico de Señales del LEICI, por tantos intercambios de ideas y sugerencias, en especial a Jorge, Javier S., Mariano, Juan Ignacio y Javier G.

A mis compañeros de cátedra en las que me desempeñé durante estos años, y en especial a Adrián y Marite quienes no solo me han enseñado desde lo profesional sino también a través de sus valores.

Quiero agradecer también a Fabricio, quien me prestó su oído y no dudo en instarme reiteradamente para reincorporarme al ámbito académico.

A todos los integrantes del LEICI por generar el ambiente de trabajo propicio para llevar adelante la tesis y por los debates con múltiples aristas, porque esa diversidad de pensamiento ha formado parte de mi crecimiento.

A la Universidad Nacional de La Plata y el Consejo Nacional de Investigaciones Científicas y Técnicas, instituciones que financiaron la realización de esta tesis. Al LEICI - Instituto de

Investigaciones en Electrónica, Control y Procesamiento de Señales por proporcionar el lugar y el material de trabajo.

Índice general

| | |
|---|-----------|
| 1. Introducción | 1 |
| 1.1. Resultados preexistentes | 4 |
| 1.2. Contribuciones originales | 7 |
| 1.3. Organización de la tesis | 8 |
| 2. Consideraciones generales | 11 |
| 2.1. Introducción | 11 |
| 2.2. Radar IPIX | 12 |
| 2.2.1. Muestreo y cuantización de las señales | 13 |
| 2.2.2. Datos utilizados | 15 |
| 2.3. Modelos existentes | 16 |
| 2.3.1. Modelo Rayleigh | 17 |
| 2.3.2. Modelo Weibull | 18 |
| 2.3.3. Modelo K | 18 |
| 2.3.4. Ajuste a los datos reales | 19 |
| 2.4. Detección | 22 |
| 2.4.1. Criterio de Neyman-Pearson | 22 |
| 2.4.2. Test de relación de verosimilitud generalizado | 23 |
| 2.4.3. Enfoque Bayesiano | 23 |
| 2.5. Conclusiones | 24 |
| 3. Detección en rango | 25 |
| 3.1. Introducción | 25 |
| 3.2. Procesos GARCH | 27 |
| 3.2.1. Estructura | 28 |
| 3.2.2. Estimación de cuasi-máxima verosimilitud | 28 |
| 3.2.3. Propiedades asintóticas del QMLE | 30 |
| 3.3. Aproximación de la Cota de Crámer-Rao | 31 |
| 3.3.1. Distribución de máxima entropía | 31 |
| 3.3.2. Cota de Cramér-Rao | 32 |
| 3.3.3. Algoritmo numérico | 33 |

| | | |
|-----------|---|-----------|
| 3.3.4. | Aproximación alternativa | 34 |
| 3.3.5. | Resultados | 34 |
| 3.4. | Modelado del clutter | 38 |
| 3.4.1. | Método alternativo para la estimación de los coeficientes | 38 |
| 3.4.2. | Error de estimación | 39 |
| 3.4.3. | Ajuste a los datos reales | 41 |
| 3.5. | Esquema de detección | 43 |
| 3.5.1. | Probabilidad de detección | 46 |
| 3.5.2. | Análisis de sensibilidad | 49 |
| 3.5.3. | Resultados | 50 |
| 3.6. | Conclusiones | 55 |
| 4. | Detección en rango y tiempo | 57 |
| 4.1. | Introducción | 57 |
| 4.2. | Modelado del clutter | 59 |
| 4.2.1. | Procesos AR-GARCH-2D | 59 |
| 4.2.2. | Estimación de parámetros | 61 |
| 4.3. | Esquema de detección | 62 |
| 4.4. | Simulaciones numéricas | 66 |
| 4.4.1. | Estimación | 66 |
| 4.4.2. | Probabilidad de detección | 69 |
| 4.4.3. | Análisis de sensibilidad | 70 |
| 4.4.4. | Comparación de desempeño | 71 |
| 4.5. | Análisis de desempeño con mediciones reales | 73 |
| 4.5.1. | Ajuste del modelo a los datos de clutter | 74 |
| 4.5.2. | Comparación de desempeño | 76 |
| 4.6. | Discusión | 77 |
| 4.7. | Conclusiones | 79 |
| 5. | Estimación secuencial de parámetros de procesos GARCH | 83 |
| 5.1. | Introducción | 83 |
| 5.2. | Estimador LMMSE | 84 |
| 5.2.1. | Formulación del problema | 84 |
| 5.2.2. | Deducción del estimador | 86 |
| 5.2.3. | Simulaciones numéricas | 92 |
| 5.3. | Estimador LMMSE extendido | 93 |
| 5.3.1. | Formulación del problema | 93 |
| 5.3.2. | Deducción del estimador | 96 |
| 5.3.3. | Simulaciones numéricas | 98 |
| 5.4. | Conclusiones | 102 |

| | |
|--|------------|
| 6. Conclusiones | 103 |
| A. Gradientes de los funcionales utilizados | 109 |
| A.1. Gradiente de $\mathcal{L}_2(\boldsymbol{\theta})$ | 109 |
| A.2. Gradiente de $\mathcal{L}_3(\boldsymbol{\theta})$ | 110 |
| A.3. Gradiente de $\mathcal{L}_4(\boldsymbol{\theta})$ | 110 |
| B. Condición de varianza finita para proceso AR-GARCH-2D | 113 |
| C. Modelo Autorregresivo | 117 |
| C.1. Estimación de parámetros | 117 |

Índice de figuras

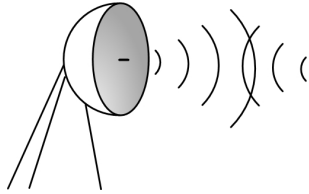
| | |
|---|----|
| 2.1. Diagrama en bloques de un radar convencional con receptor superheterodino. . . | 13 |
| 2.2. Estrategia de almacenamiento de los datos. | 14 |
| 2.3. Intensidad de las mediciones correspondientes al conjunto de datos <i>stare0</i> | 15 |
| 2.4. Intensidad de las mediciones correspondientes al conjunto de datos <i>stare6</i> | 15 |
| 2.5. Correlación estimada del conjunto de datos <i>stare0</i> , sin considerar las celdas de rango correspondientes al objetivo. | 17 |
| 2.6. Correlación estimada del conjunto de datos <i>stare6</i> | 17 |
| 2.7. Funciones densidad de probabilidad resultantes del ajuste e histograma correspondiente al conjunto de datos <i>stare0</i> para la polarización VV. | 21 |
| 2.8. Funciones densidad de probabilidad resultantes del ajuste e histograma correspondiente al conjunto de datos <i>stare6</i> para la polarización VV. | 21 |
| 3.1. Realizaciones de dos casos particulares de proceso GARCH y sus varianzas condicionales. | 29 |
| 3.2. Histograma del proceso ARCH de primer orden y su aproximación de la fdp a través de la distribución de máxima entropía. | 36 |
| 3.3. CRB para la estimación de los coeficientes del proceso ARCH de primer orden. | 36 |
| 3.4. Modelado del clutter a través de procesos GARCH. | 38 |
| 3.5. Estimados de los coeficientes de un proceso GARCH(0,1) en función del número de muestras. | 40 |
| 3.6. Histogramas de la magnitud del clutter del conjunto de datos <i>stare0</i> para la polarización VV, de la magnitud del proceso GARCH(0,1) y funciones densidad de probabilidad resultantes del ajuste. | 43 |
| 3.7. Desempeño del detector GARCH y comparación con el detector Gaussiano, utilizando datos sintéticos. | 48 |
| 3.8. Efecto del error en la estimación de los coeficientes sobre el desempeño del detector GARCH. | 49 |
| 3.9. Mapa de detección en rango y tiempo con objetivo sintético para el conjunto de datos <i>stare0</i> , polarización VV. La ubicación del objetivo es la celda 27, correspondiente a 2391 m. | 52 |

| | |
|---|-----|
| 3.10. Mapa de detección en rango y tiempo con objetivo real para el conjunto de datos <i>stare0</i> , polarización VV. La ubicación del objetivo es 2691 m. | 52 |
| 3.11. Curvas ROC empíricas para el conjunto de datos <i>stare0</i> , polarización VV. | 54 |
| 4.1. Modelado del clutter a través de procesos AR-GARCH-2D. | 62 |
| 4.2. Estimados de los coeficientes del proceso AR-GARCH-2D simulado en función del número de muestras. | 68 |
| 4.3. Estimados del coeficiente k , del proceso AR-GARCH-2D simulado, en función del número de muestras. | 69 |
| 4.4. Desempeño del detector AR-GARCH-2D utilizando datos sintéticos. | 70 |
| 4.5. Efecto del error en la estimación de los coeficientes sobre el desempeño del detector AR-GARCH-2D. | 71 |
| 4.6. Probabilidad de falsa alarma en función del umbral de los detectores, utilizando datos de clutter sintéticos. | 73 |
| 4.7. Probabilidad de detección de los detectores en función de la SCR. | 73 |
| 4.8. Probabilidad de falsa alarma en función del umbral de los detectores AR-GARCH-2D, utilizando datos de clutter del radar IPIX. | 75 |
| 4.9. Desempeño de los detectores AR-GARCH-2D, utilizando datos de clutter del radar IPIX y objetivo sintético. | 76 |
| 4.10. Probabilidad de falsa alarma en función del umbral, utilizando datos de clutter del radar IPIX. | 77 |
| 4.11. Comparación de desempeño entre los detectores, utilizando datos de clutter del radar IPIX y objetivo sintético. | 77 |
| 4.12. Histogramas de la magnitud del clutter del conjunto de datos <i>stare6</i> para la polarización VV y de la magnitud del proceso AR-GARCH-2D. | 78 |
| 4.13. Correlaciones estimadas correspondientes al conjunto de datos <i>stare6</i> para la polarización VV y al proceso AR-GARCH-2D. | 78 |
| 5.1. Estimado de la varianza condicional a partir de datos sin ruido y estado inicial desconocido. | 92 |
| 5.2. Estimado LMMSE a partir de datos con ruido y estado inicial desconocido. | 93 |
| 5.3. Error en la estimación LMMSE a partir de datos con ruido y estado inicial desconocido. | 94 |
| 5.4. RMSE en la estimación LMMSE de la varianza condicional. | 94 |
| 5.5. Resultados de la implementación del estimador LMMSE extendido empleando datos sintéticos. | 100 |
| 5.6. Resultados de la implementación del estimador LMMSE extendido con variaciones en el parámetro k | 101 |

Lista de abreviaturas

- ALQ: *Adaptive Linear-Quadratic* - adaptivo lineal-cuadrático, 4.
- AR: *Autoregressive* - autorregresivo, 3.
- ARCH: *Autoregressive Conditional Heteroscedastic* - heteroscedasticidad condicional autorregresiva, 5.
- ARGLR: *Autoregressive Generalized Likelihood Ratio* - relación de verosimilitud autorregresiva generalizada, 4.
- ARMA: *Autoregressive Moving Average* - autorregresivo de media móvil, 4.
- BIC: *Bayesian Information Criterion* - criterio de información Bayesiano, 42.
- CFAR: *Constant False Alarm Rate* - tasa de falsa alarma constante, 6.
- CRB: *Cramér-Rao bound* - cota de Cramér-Rao, 26.
- CW: *Continuous-Wave* - onda continua, 11.
- fdp: función de densidad probabilidad, 3.
- FRP: Frecuencia de Repetición de Pulso, 13.
- GARCH: *Generalized Autoregressive Conditional Heteroscedastic* - heteroscedasticidad condicional autorregresiva generalizada, 3.
- GC-pdf: *Generalized Compound probability density function* - función densidad de probabilidad compuesta generalizada, 5.
- GLRT: *Generalized Likelihood Ratio Test* - test de relación de verosimilitud generalizado, 4.
- IF: *Intermediate Frequency* - frecuencia intermedia, 13.
- iid: independiente e idénticamente distribuido, 28.
- IRP: Intervalo de Repetición de Pulso, 13.
- LMMSE: *Linear Minimum Mean Square Error* - lineal de error cuadrático medio mínimo, 4.
- LNA: *Low Noise Amplifier* - amplificador de bajo ruido, 13.
- MLE: *Maximum Likelihood Estimator* - estimador de máxima verosimilitud, 19.
- MSE: *Mean Square Error* - error cuadrático medio, 86.
- NSCME: *Normalized Sample Covariance Matrix Estimate* - estimado muestral normalizado de la matriz de covarianza, 72.
- QMLE: *Quasi-Maximum Likelihood Estimator* - estimador de cuasi-máxima verosimilitud, 26.
- RMSE: *Root Mean Square Error* - raíz cuadrada del error cuadrático medio, 37.
- ROC: *Receiver Operating Characteristic* - característica operativa del receptor, 48.
- SCR: *Signal to Clutter Ratio* - relación señal a clutter, 46.

SQP: *Sequential Quadratic Programming* - programación cuadrática secuencial, 40.



Capítulo 1

Introducción

En general los murciélagos producen repugnancia o temor entre los seres humanos. Los motivos de este comportamiento pueden justificarse en su aspecto físico, en sus hábitos nocturnos, en que algunas de sus especies se nutren de sangre o bien en que son uno de los íconos del cine de terror. Sin embargo, la forma de orientarse que utilizan la mayor parte de las especies de murciélagos, la ecolocación, podría ser la fuente de inspiración en la naturaleza de los sistemas de radar. La ecolocación, es un sistema de percepción que consiste en la emisión de sonidos que rebotan al encontrar un obstáculo (Neuweiler, 2000). Los ecos producidos a su retorno se transmiten al cerebro a través del sistema nervioso auditivo y les ayuda a orientarse. Logran así, por ejemplo saber la distancia hasta el objeto, o los objetos, donde se produce el rebote, midiendo el tiempo de retardo entre la señal que ha emitido y la que ha recibido.

Los radares convencionales son sistemas activos que operan transmitiendo energía en forma de ondas electromagnéticas y recibiendo las señales reflejadas por la región y el objeto iluminado. En el receptor, se utilizan técnicas de procesamiento estadístico de señales para obtener información de interés de la señal recibida (Skolnik, 2001).

La historia del radar comienza con la teoría moderna del electromagnetismo. Sin embargo, como ha ocurrido con gran parte de los avances tecnológicos los desarrollos de los sistemas radar fueron impulsados por necesidades militares. La aparición de aviones bombarderos pesados hacia fines de 1930 fue uno de los hechos que dio lugar a la utilización del radar en operaciones militares (Skolnik, 2001). El objetivo era contar con alertas tempranas a grandes distancias cuando un bombardero se aproximaba. Durante la segunda guerra mundial la tecnología de radar creció rápidamente, con grandes desarrollos independientes en los distintos países invo-

lucrados (Skolnik, 2001). En la actualidad, el uso del radar con fines militares sigue siendo una de sus aplicaciones dominantes, utilizado en vigilancia de fronteras, navegación o guiado de misiles, en tierra, aire o agua. Sin embargo, su utilización se ha ampliado en numerosos campos de aplicación. Su uso más común es el radar para control de tráfico, cuya función es la de medir la velocidad de automoviles en rutas y autopistas con el objetivo de controlar que se respeten los límites existentes. Otro ejemplo son los radares meteorológicos utilizados para monitorear y predecir el clima o en estudios atmosféricos. Una aplicación habitual del radar, que afecta a un gran número de personas, son los sistemas de control de tráfico aéreo, donde su función es guiar aviones comerciales en sus trayectos y en la vecindad de los aeropuertos. A bordo de aviones los radares son usados para determinar altitud y evitar tormentas severas. Asimismo, son utilizados por barcos para evitar colisiones y para la detección de boyas. También, radares montados en vehículos espaciales son una importante herramienta utilizada en el mapeo de la topología de la tierra y de características medio ambientales, como condiciones de los océanos o de los hielos continentales, condiciones forestales, el uso de la tierra o contaminación (Richards, 2005). Finalmente, y de manera aún experimental, los radares se están utilizando para ordenar automáticamente el tránsito de vehículos en autopistas.

Una de las dificultades a tener en cuenta en los sistemas de radar es que la señal de interés suele encontrarse obscurecida por las reflexiones producidas por el ambiente, fenómeno al que se denomina *clutter* (Haykin, 1985). Los ejemplos más comunes de clutter son las reflexiones producidas por superficies de tierra (u objetos sobre el nivel del terreno), superficies marinas, y ciertas condiciones atmosféricas (lluvia, granizo y nieve). Generalmente, y dependiendo de la aplicación, el clutter es considerado una fuente de interferencia y perturbaciones cuyos efectos se deben eliminar o reducir. Por lo tanto, en vista de su naturaleza aleatoria, es importante el desarrollo de métodos estadísticos de procesamiento de señales para poder detectar objetivos y estimar sus propiedades en situaciones de clutter intenso y dinámico. Para obtener algoritmos eficientes, es fundamental utilizar modelos realistas de las señales recibidas por el radar. Estos modelos deben enfatizar las diferencias entre el objeto de interés y el clutter. De esta forma, los métodos de procesamiento de señal son usados para separar la señal proveniente de reflexiones en el objetivo del clutter y reducir el efecto degradante de este último.

Para representar los datos en problemas de radar los métodos convencionales utilizan distintas distribuciones estadísticas que se ajustan a los datos reales de clutter. Dependiendo de la aplicación y el escenario de operación, las distribuciones más usadas son la lognormal, Weibull y K (Goldstein, 1973; Sekine *et al.*, 1979; Shnidman, 1999; Sangston *et al.*, 2012), y en ambientes con ruido impulsivo distribuciones alfa-estables o con colas pesadas (Zha y Qiu., 2006). Los métodos mencionados para caracterizar el clutter son independientes del tiempo. No obstante, es importante destacar que en muchas aplicaciones el escenario puede cambiar. Por ejemplo, el cambio en las condiciones climáticas modifica el clutter marítimo; o en radares montados en vehículos aéreos, el escenario cambia a medida que la aeronave sobrevuela el territorio.

En particular los radares que operan en ambientes marítimos poseen una seria limitación

en su desempeño debido a las reflexiones no deseadas (Gini y Greco, 2001; Jakeman y Pusey, 1976). El clutter en estos casos es impulsivo y los impulsos en muchos casos son procesados por el radar como objetivos, incrementando la probabilidad de falsa alarma. Por lo tanto uno de los problemas de estudio en los sistemas de radar es el modelado del clutter con modelos que no sean Gaussianos para el diseño de detectores óptimos o sub-óptimos.

En esta tesis se desarrollan modelos dinámicos de clutter que contemplan las variaciones del escenario y utilizan la historia del clutter para mejorar su caracterización en el instante actual y las predicciones a tiempo futuro. En primer lugar se estudia el modelado del clutter a través de series temporales que presentan heteroscedasticidad condicional autorregresiva (GARCH por sus siglas en inglés). Estos procesos son muy utilizados en econometría para modelar diversos índices financieros (Engle, 1982; Bollerslev, 1986). Presentan como características las de ser impulsivos y de agrupamiento de su volatilidad, es decir grandes cambios tienden a ser seguidos de grandes cambios y pequeños cambios tienden a ser seguidos de pequeños cambios. En contraposición, poseen la desventaja de que no cuentan con una expresión explícita para su función densidad de probabilidad (fdp). Por este motivo, en esta tesis se analizan alternativas para estimar sus parámetros y determinar la calidad de la estimación, esto último a partir de una aproximación de la cota de Cramér-Rao y a través de simulaciones numéricas. También se incluye el ajuste del modelo a mediciones reales de clutter marítimo. El análisis realizado confirma que los procesos GARCH permiten una adecuada representación de las estadística de los datos.

En segundo lugar se propone un modelo para el clutter que resulta de la combinación de un proceso GARCH en dos dimensiones (GARCH-2D) con un proceso autorregresivo (AR). La parte GARCH-2D del modelo preserva la propiedad impulsiva de los procesos GARCH y la AR en las innovaciones permite modelar la correlación. Como en el tópico anterior se desarrolla un algoritmo de estimación para sus parámetros y se utiliza para realizar un ajuste a mediciones de clutter marítimo.

Se estudia también el problema de detección asociado a cada uno de los modelos descriptos. Con este objetivo, se adaptan los test de hipótesis usuales para sortear la dificultad impuesta por la falta de expresiones explícitas para las fdp's de los modelos propuestos. El enfoque del detector resultante a partir del segundo modelo es diferente del que se obtiene basado en un modelo GARCH para el clutter. Debido a que los procesos AR-GARCH-2D modelan el clutter en dos dimensiones su detector asociado toma cada decisión usando información de múltiples pulsos. En cambio, en el caso del detector GARCH el clutter es modelado considerando las realizaciones en rango, es decir, se cuenta con una única realización del proceso por pulso transmitido.

Para ambos detectores se deducen expresiones para sus probabilidades de falsa alarma y, dada su complejidad matemática, la probabilidad de detección se evalúa por medio de simulaciones numéricas. Asimismo, se analiza la sensibilidad del desempeño de los detectores ante errores en la estimación de sus parámetros. A pesar de que no resultan de tasa de falsa alarma constante, muestran un comportamiento robusto en situaciones prácticas.

Además, se evalúa el desempeño de los detectores obtenidos utilizando las mediciones de clutter marítimo. Al mismo tiempo, se comparan con detectores del mismo tipo en cada caso. El GARCH se contrasta con un detector Gaussiano (Richards, 2005) y con uno Weibull (Ravid y Levanon, 1992), mientras que el AR-GARCH-2D se compara con el test de relación de verosimilitud generalizado (GLRT por sus siglas en inglés) (Kelly, 1986), el detector adaptivo lineal-cuadrático (ALQ por sus siglas en inglés) (Gini y Greco, 1999) y el test de relación de verosimilitud autorregresiva generalizada (ARGLR por sus siglas en inglés) (Sheikhi *et al.*, 1998).

Las características de los procesos del tipo GARCH surgen del hecho de que su varianza instantánea se modela como un proceso autorregresivo de media móvil (ARMA por sus siglas en inglés) y es justamente de este parámetro del cual depende el umbral en los esquemas de detección correspondientes a estos modelos. Por este motivo, finalmente se estudia el problema de estimación secuencial de los parámetros de los procesos GARCH, con el objetivo de contar con un mecanismo que permita actualizar la varianza en cada instante de decisión.

Suponiendo que los coeficientes del proceso son conocidos, las ecuaciones que describen el modelo pueden escribirse como un sistema lineal de tiempo discreto. En estas condiciones, siguiendo el enfoque de estimación Bayesiano se deduce un estimador lineal de mínimo error cuadrático medio (LMMSE por sus siglas en inglés) para la varianza condicional de los procesos GARCH, que es el parámetro del cual depende el estadístico de los detectores desarrollados. La deducción del algoritmo es análoga a la del filtro de Kalman, pero en este caso las matrices del sistema son aleatorias (Koning, 1984). En una segunda etapa se considera que los parámetros del proceso son desconocidos, lo que lleva a un sistema de estados no lineal y como consecuencia a un estimador LMMSE extendido, con un enfoque similar al procedimiento seguido para obtener el filtro Kalman extendido (Simon, 2006).

Desde un punto de vista general los análisis del clutter presentados en la tesis están motivados por su relevancia en el diseño exitoso de sistemas de radar. En particular el desarrollo de algoritmos de detección se lleva a cabo por medio de modelos que describen adecuadamente el clutter, pero principalmente, que resultan significativos al momento de predecir sus desempeños.

1.1. Resultados preexistentes

En radares de baja resolución, durante muchos años el clutter fue modelado de manera aceptable, excepto para ángulos bajos de incidencia, por medio de la distribución Gaussiana para la amplitud de las mediciones o, equivalentemente, la distribución Rayleigh para su intensidad (Marcum, 1960; Swerling, 1960). Sin embargo en radares modernos de mejor resolución se ha observado que la estadística del clutter se aparta de la del modelo Gaussiano. En Haykin *et al.* (1991) y en Farina *et al.* (1997) se realizan análisis detallados de datos experimentales recogidos por el radar IPIX de la Universidad de McMaster desde un acantilado en Dartmouth, Nueva Escocia. Ambos trabajos se concentran en el ajuste de diferentes distribuciones a los datos y se

concluye que éstos se apartan de la distribución Gaussiana, indicando que la K y la Log-normal son las que describen mejor su intensidad, dentro del conjunto de distribuciones consideradas. Por otro lado, en Anastassopoulos *et al.* (1999) se analizan datos marítimos provenientes de un radar de apertura sintética (SAR por sus siglas en inglés) en las bandas X y C, arribando a la conclusión de que la función densidad de probabilidad compuesta generalizada (GC-pdf por sus siglas en inglés) es la más conveniente para representar las mediciones de SAR. En los radares de alta resolución las celdas de resolución en rango son pequeñas dando como resultado densidades de clutter con colas pesadas que no pueden representarse adecuadamente a través de la distribución gaussiana.

Estos estudios dieron lugar a la modelización del clutter utilizando una amplia variedad de distribuciones. Dependiendo de la aplicación, las más usadas son la Weibull, la Log-normal, la K (Shnidman, 1999; Sangston y Gerlach, 1994), también se ha empleado la distribución Pareto (Weinberg, 2011), las GC-pdf's (Shang y Song, 2011; Sangston *et al.*, 2012; Cui *et al.*, 2012) o bien se ha caracterizado como un proceso aleatorio esféricamente invariante (Conte y Longo, 1987). Algunas de estas distribuciones logran un buen ajuste a las mediciones del clutter, sin embargo es importante destacar dos cuestiones. La primera es que son independientes del tiempo y en muchos casos el entorno del radar puede cambiar abruptamente (Li *et al.*, 2009) llevando a una degradación en el desempeño del radar en un escenario real. En segundo lugar, en general no es posible obtener un esquema de detección asociado a estas distribuciones, llevando a soluciones sub-óptimas o *ad hoc* que luego son evaluadas con datos sintéticos del modelo propuesto.

En esta tesis se propone la utilización de series temporales que presentan heteroscedasticidad condicional autorregresiva (ARCH por sus siglas en inglés). Los procesos ARCH fueron propuestos por Engle (1982) y han sido ampliamente utilizados en econometría para modelar diversos índices financieros (Ding *et al.*, 1993; Gouriéroux, 1997). Se caracterizan por ser condicionalmente Gaussianos, con varianza condicional variante en el tiempo. Posteriormente Bollerslev (1986) introdujo una extensión de los procesos ARCH a los procesos GARCH, de manera similar a la extensión de un proceso AR, en series temporales, al proceso ARMA general. Debido a su popularidad en el campo de la econometría este tipo de procesos han sido extensamente estudiados. Se han propuesto métodos para la estimación de sus parámetros (Berkes *et al.*, 2003; Francq y Zakoian, 2004; Straumann, 2005; Francq y Zakoian, 2007) y se han analizado las propiedades de los dichos estimadores (Weiss, 1986; Lee y Hansen, 1994; Berkes y Horváth, 2003, 2004).

Recientemente los procesos GARCH también fueron utilizados en la modelización del ruido en aplicaciones de sonar (Amiri *et al.*, 2007). Asimismo, en Noiboar y Cohen (2007) se propuso una extensión multidimensional de estas series temporales y se empleó en el modelado de coeficientes de la transformada wavelet para la detección de anomalías en aplicaciones de sonar. Más tarde, se realizó un análisis análogo para la supresión de speckle en procesamiento de imágenes SAR utilizando procesos GARCH-2D (Amirmazlaghani *et al.*, 2009).

Por otra parte existen numerosos trabajos donde se estudian algoritmos de detección adaptati-

va para sistemas de radar. Este problema cobra importancia debido a que los radares modernos operan en ambientes no-homogéneos y no-estacionarios. Naturalmente, a causa de su tratabilidad matemática los primeros resultados encontrados en la literatura suponen que el clutter es Gaussiano (Reed *et al.*, 1974; Kelly, 1986; Bose y Steinhardt, 1995; Sheikhi *et al.*, 1998; Kraut y Scharf, 1999). En Reed *et al.* (1974) se analiza un método para la detección de una señal conocida en presencia de interferencia Gaussiana con matriz de covarianza desconocida. Se utilizan dos conjuntos de datos de entrada, los cuales se denominan primario y secundario. La posibilidad de que exista señal presente solo se contempla para los datos primarios, mientras que se supone que los datos secundarios solo se componen de muestras de clutter, independientes e idénticamente distribuidas de las componentes de clutter de los datos primarios. En este algoritmo los datos de entrada secundarios se utilizan para estimar la matriz de covarianza del clutter. Luego, los datos primarios son pasados por el filtro adaptado con ruido coloreado y la salida es comparada con un umbral. Sin embargo, no se da una regla para determinar este último, cuyo valor es el que controla la probabilidad de falsa alarma.

En Kelly (1986) se reconsidera el problema como un ejercicio de test de hipótesis con una generalización del modelo de señal, se supone que no es completamente conocida. En el método de Reed *et al.* (1974) se reemplaza por un test de verosimilitud generalizado donde se incluyen tanto los datos de entrada primarios como los secundarios. La probabilidad de falsa alarma de la regla de decisión resultante posee la propiedad de no depender de la matriz de covarianza del clutter. En el caso del detector propuesto en Sheikhi *et al.* (1998) el planteo es análogo, pero la interferencia se modela por medio de un proceso autorregresivo con el objetivo de imponerle una estructura a su matriz de covarianza. A diferencia de los esquemas anteriores en este caso el detector se ajusta a si mismo utilizando únicamente el conjunto de datos primario.

Con el sustento del análisis de datos experimentales, el problema de detección adaptiva con clutter no-Gaussiano también ha sido investigado (Conte *et al.*, 1996; Richmond, 1996; Gini *et al.*, 1998; Gini y Greco, 1999; Sangston *et al.*, 1999). En el caso estudiado por Richmond (1996) la interferencia se modela como Gaussiana-compuesta y se supone que la textura de los datos de entrada primarios y secundarios está completamente correlacionada. Por otro lado, Gini y Greco (1999) proponen una solución sub-óptima para la detección de objetivos con modelo Swerling I inmersos en clutter con distribución Gaussiana-compuesta, en particular con distribución K. Este detector, al que se denomina *adaptivo lineal-cuadrático*, es de falsa alarma constante (CFAR por sus siglas en inglés) respecto de la fdp de la amplitud de la señal y si bien es bastante insensible a la estructura de la matriz de covarianza del clutter, no resulta CFAR respecto de este parámetro.

Otros trabajos estudian comparativamente el desempeño de estos detectores utilizando mediciones reales de radar (Gini *et al.*, 2000; Maio *et al.*, 2005). De estos resultados es interesante destacar que aún cuando la estadística de los datos de clutter no se ajusta a la distribución Gaussiana, el desempeño de detectores asociados a ésta es comparable, y en algunos casos mejor, al de esquemas de detección diseñados para modelos que la representan mejor. La causa de este comportamiento es que la complejidad matemática de los modelos no-Gaussianos lleva

a soluciones *ad hoc* o sub-óptimas.

Finalmente, los sistemas de tiempo discreto con parámetros aleatorios aparecen en aplicaciones tales como control digital de procesos químicos (Yaz y Skelton, 1994), en seguimiento de objetivos múltiples (Luo *et al.*, 2012) o en sistemas de control a través de redes de datos (Sun y Ma, 2014). Koning (1984) aborda el problema de estimación lineal óptima de estados en el caso general de un sistema lineal de tiempo discreto con parámetros estocásticos, los cuales son estadísticamente independientes con respecto al tiempo. El estimador se deduce transformando el sistema en un sistema con parámetros determinísticos y ruidos de estado y observación aditivos y dependientes de los estados. Estos resultados han sido utilizados en aplicaciones de sensado múltiple (Luo *et al.*, 2008) y en algoritmos de asociación de datos en problemas seguimiento de objetivos múltiples (Luo *et al.*, 2012).

Por su parte, en los sistemas de control en red los retardos aleatorios, la pérdida de paquetes y los devanecimientos de los datos son los fenómenos que llevan al modelado estocástico del sistema. El problema de estimación lineal óptima para este tipo de problemas ha sido estudiado en Sun *et al.* (2008); Moayedi *et al.* (2010); Ma y Sun (2011).

1.2. Contribuciones originales

Las principales contribuciones de esta tesis están relacionadas con la modelización estadística del clutter y la detección en aplicaciones de radar convencional.

Se propuso como modelo dinámico del clutter la utilización de series temporales que presentan heterocedasticidad condicional autorregresiva generalizada (GARCH). Se adecuó el método de estimación comúnmente utilizado para los parámetros de estos procesos aleatorios, a situaciones donde el número de muestras disponible no es suficiente para obtener un error razonable. Esto permitió ajustar el modelo a mediciones reales de clutter marítimo. En base al modelo GARCH para el clutter se dedujo un esquema de detección adaptando los test de hipótesis usuales, debido a que no se cuenta con una expresión explícita para su función densidad de probabilidad. Se evaluó el desempeño del detector obtenido y se comparó con el de otros detectores existentes en la literatura. Los resultados se presentaron en Pascual *et al.* (2011) y fueron publicados en Pascual *et al.* (2013b).

Otro aporte realizado es la evaluación de desempeño en la estimación de los coeficientes de un proceso GARCH, utilizando una aproximación de la cota de Cramér-Rao a partir de la densidad de probabilidad de máxima entropía que se obtiene como solución al problema de momentos (Pascual *et al.*, 2013a).

Con el objetivo de incluir información de múltiples pulsos en los instantes de decisión, se extendió el modelo mencionado combinando un proceso GARCH de dos dimensiones con uno autorregresivo, dando lugar a lo que se denominó procesos AR-GARCH-2D. Se analizaron condiciones para su estabilidad y se dedujo un procedimiento de estimación de sus parámetros, análogo al de los procesos GARCH. Se ajustó el modelo a datos reales de clutter marítimo. Se debe tener en cuenta que los procesos AR-GARCH-2D representan una familia de modelos, lo

que hizo necesario utilizar un criterio de selección de modelos para elegir aquel que presenta el compromiso óptimo entre complejidad y calidad del ajuste. Además, se obtuvo un detector asociado a este modelo, se evaluó su desempeño y se comparó con el de otros detectores existentes. Los resultados que se muestran en la tesis fueron publicados en Pascual *et al.* (2014c).

Los estadísticos de los detectores propuestos dependen de la varianza condicional de los procesos, la cual varía temporalmente. Por este motivo, se estudió el problema de su estimación de forma secuencial. Bajo la hipótesis de que los coeficientes del proceso son conocidos, el modelo puede escribirse como un sistema de estados lineal de tiempo discreto con matrices aleatorias. Siguiendo el enfoque de estimación Bayesiano se derivó un estimador lineal de mínimo error cuadrático medio para la varianza condicional de los procesos GARCH. Los resultados obtenidos fueron presentados en Pascual *et al.* (2014a).

Por otro lado, se generalizaron estas ideas considerando el caso en que los coeficientes son desconocidos. Haciendo un paralelismo con el filtro de Kalman, se aumenta el vector de estados con los parámetros desconocidos. En estas condiciones el sistema de estados resultante es no lineal dando lugar a un estimador LMMSE para el vector de estados análogo al filtro de Kalman extendido. Respecto a este tema, los resultados presentados en esta tesis aún no han sido publicados.

1.3. Organización de la tesis

El capítulo 2 presenta una introducción donde se describen las mediciones de radar que se utilizarán en el desarrollo de la tesis. Se exponen los modelos matemáticos usualmente empleados para representar el clutter, los que se ajustan a datos reales y se realiza un análisis comparativo entre ellos. Además, se detallan las herramientas existentes para la deducción de esquemas de detección que permitirán desarrollar los algoritmos propuestos a lo largo del trabajo.

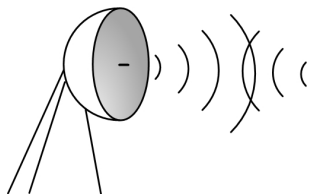
Del capítulo 3 en adelante se presentan las diferentes contribuciones originales de la tesis en el ámbito de la detección de radar en presencia de clutter dinámico.

En el tercer capítulo se propone la modelización en rango del clutter utilizando los procesos GARCH. Se describe la estructura de este tipo de procesos, se detalla un método de estimación de sus parámetros y se caracterizan sus estimados. Debido a que no se cuenta con una expresión explícita para su fdp, se da una aproximación de la cota de Cramér-Rao para la varianza de los estimadores de sus coeficientes, que resulta de considerar la distribución de máxima entropía dado un conjunto de momentos conocidos. Además, se realiza el ajuste del modelo a las mediciones a través de un método alternativo propuesto, a causa de que el número de muestras por pulso no es suficiente para alcanzar un error de estimación razonable en una situación práctica. Finalmente, se deduce un esquema de detección basado en el modelo GARCH para el clutter y se analiza su desempeño tanto de manera teórica como experimental por medio de los datos reales.

En el capítulo 4 se ataca el problema de detección en rango y tiempo. Se presenta un modelo que permite modelar el clutter en dos dimensiones combinando un proceso GARCH-2D con un proceso autorregresivo. Se describe un procedimiento de estimación de sus coeficientes, a través del cual se realiza un ajuste a las mediciones de radar. Se deduce un nuevo esquema de detección asociado al modelo AR-GARCH-2D y se analiza su desempeño de forma análoga a lo realizado a con el detector GARCH en el capítulo 3. Utilizando un criterio de selección de modelos se elige el proceso que mejor representa a los datos. Por último, se compara el desempeño del detector asociado al modelo elegido con el de algoritmos de detección existentes en la literatura.

En el quinto capítulo se estudia el problema de estimación secuencial de los parámetros de un proceso GARCH. Este problema se divide en dos etapas. En la primera se supone que los coeficientes son conocidos y se deduce un estimador LMMSE para la varianza condicional. Como se verá, si bien esta hipótesis no es realista en una situación práctica, sienta las bases para la segunda etapa. En esta última se considera que los coeficientes son desconocidos y se agregan al vector de estados. En consecuencia el sistema resultante pierde la linealidad. En estas condiciones, se deduce un nuevo estimador LMMSE para la estimación conjunta de los coeficientes y de la varianza condicional. En ambos casos se presentan ejemplos numéricos donde se verifica el comportamiento esperado mediante simulaciones.

Por último en el capítulo 6 se exponen las conclusiones generales de la tesis, así como posibles líneas de investigación a seguir.



Capítulo 2

Consideraciones generales

2.1. Introducción

El radar es un sistema de sensado remoto que utiliza técnicas de procesamiento estadístico de señales para obtener información de interés de la señal recibida. Los radares convencionales son sistemas activos que operan transmitiendo energía en forma de ondas electromagnéticas hacia una región específica del espacio en busca de objetos a los que se denominan objetivos. Parte de esta energía es reflejada por el objetivo y su entorno en la región iluminada y retorna al radar. Las señales reflejadas son recogidas por la antena receptora y enviadas al receptor del radar, donde son procesadas para detectar la presencia del objetivo y extraer su información, como ubicación, velocidad, etc.

Existen numerosos tipos de radar y su clasificación varía de acuerdo con características tales como forma de onda que utilizan, banda de frecuencias en la que operan, la ubicación del transmisor en relación al receptor y su funcionalidad, entre otras (Mahafza y Elsherbeni, 2004). De acuerdo con la forma de onda se dividen en radares de onda continua, CW (por sus siglas en inglés), los cuales transmiten continuamente o en radares pulsados cuando la forma de onda que transmiten es pulsada. Por otro lado, en los casos en los que el transmisor del radar se encuentra en la misma ubicación que el receptor se dice que el radar es monoestático. En cambio cuando el transmisor se encuentra alejado del receptor el radar se denomina biestático. Existen también sistemas de radar más generales compuestos de uno o más transmisores y más de un receptor distribuidos en un cierto radio conformando una red, a este tipo de configuraciones se les denomina sistemas de radar multiestáticos (Peebles, 1998). Respecto de su funcionalidad

los radares se clasifican en numerosas categorías, aquí se mencionan solo aquellas que fueron descritas en la introducción del capítulo 1: militares, de control de velocidad, de control de tráfico aéreo, marítimos, meteorológicos, espaciales.

En la sección 2.2 se exhiben las características del radar IPIX, del cual provienen las mediciones utilizadas a lo largo de la tesis para validar los modelos de clutter propuestos y evaluar el desempeño de los algoritmos de detección desarrollados. Asimismo se detalla toda la información inherente a los conjuntos de datos utilizados.

En las aplicaciones de radar las reflexiones no deseadas producidas en el entorno del objetivo se denominan clutter. Debido a su naturaleza aleatoria por lo general, para modelizar los datos de clutter los métodos convencionales utilizan distintas distribuciones (Shnidman, 1999). En la sección 2.3 se presentan los modelos más utilizados en la literatura basados en las distribuciones Rayleigh, Weibull y K. Se describen las técnicas de estimación de sus parámetros por medio del método de máxima verosimilitud o por medio del método de momentos. Finalmente, se ajustan a los conjuntos de datos del radar IPIX y se realiza una comparación de los resultados.

Además del adecuado modelado del clutter, uno de los objetivos de este trabajo consiste en el desarrollo de algoritmos de detección en base a los modelos propuestos. En la sección 2.4 se hace una breve descripción de las técnicas de test de hipótesis que serán luego utilizadas a lo largo de la tesis. En relación al enfoque clásico se presenta el criterio de Neyman-Pearson para establecer la regla de decisión y se describe el test de relación de verosimilitud generalizado como su extensión cuando existen parámetros del problema que son desconocidos y deben estimarse. En contraposición, también se describe el enfoque Bayesiano para establecer un criterio de decisión cuando existen parámetros desconocidos, pero que son considerados variables aleatorias.

2.2. Radar IPIX

Para el desarrollo de esta tesis se utilizaron datos medidos por el radar IPIX, diseñado y construido con fines científicos en la Universidad McMaster, Canadá (Haykin *et al.*, 1991). El radar IPIX es pulsado, monoestático (Peebles, 1998), y utiliza una única antena para desempeñar las funciones de transmisión y recepción. Además, es coherente, opera en la banda X a la frecuencia de 9.4 GHz, posee la versatilidad de transmitir y recibir con dos polarizaciones lineales, horizontal (H) y vertical (V), conmutables pulso a pulso, utiliza múltiples frecuencias y también puede transmitir formas de ondas arbitrarias en un ancho de banda de 25 MHz (Haykin *et al.*, 1991).

En la figura 2.1 se muestra un diagrama en bloques de un radar pulsado monoestático con receptor superterodino. El generador de onda produce el pulso con la forma de onda que se desea transmitir. El transmisor (denotado como Tx en la figura) es el encargado de modular una portadora de radio frecuencia con esta onda y de amplificarla a los niveles de potencia indicados para ser transmitida. La señal de salida del transmisor es dirigida hacia la antena a través del duplexor.

La señal recibida, producto de las reflexiones en el objetivo y su entorno, también es dirigida

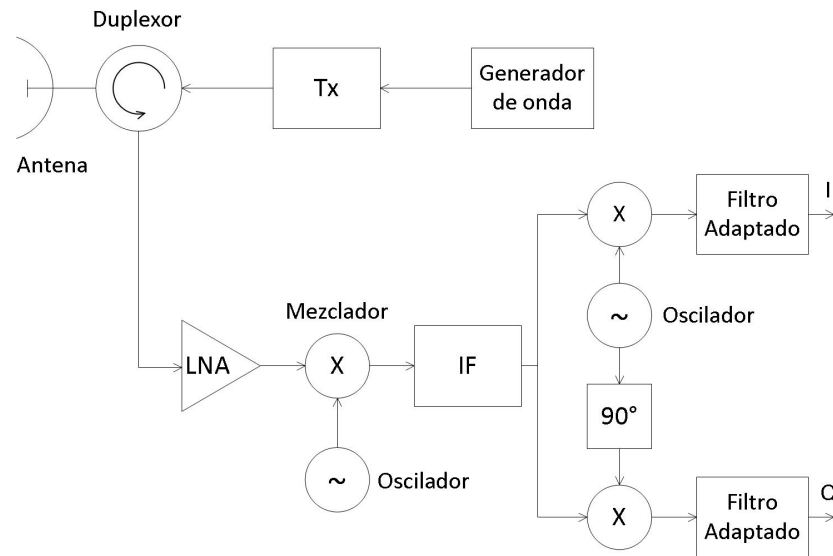


Figura 2.1: Diagrama en bloques de un radar convencional con receptor superheterodino.

por el duplexor, pero en este caso hacía el receptor. La primera etapa suele constar de un amplificador de bajo ruido (LNA por sus siglas en inglés), seguida por una o más etapas de frecuencia intermedia (IF por sus siglas en inglés). Finalmente, la señal es llevada a banda base utilizando un demodulador en cuadratura, es decir se la divide en dos canales uno en fase (denotado como I en la figura) y otro en cuadratura (denotado como Q en la figura), para extraer la información tanto de la amplitud como de la fase de la señal. Con el fin de maximizar la relación señal a ruido en los instantes en que se muestreará la señal, el filtro pasabajos utilizado en cada rama del demodulador es el filtro adaptado a la señal recibida (Richards, 2005).

El radar IPIX fue construido con el objetivo de estudiar técnicas de detección e identificación de objetivos pequeños en el océano. La base de esta investigación es el estudio de la naturaleza del clutter marítimo y el comportamiento correspondiente de los objetivos de interés en diversas condiciones del mar. Con este fin en noviembre de 1993 se creó una base de datos con el radar IPIX iluminando al océano Atlántico en la costa este de Canadá, desde un acantilado en Dartmouth, Nueva Escocia, bajo diversas condiciones de clima y estado del mar (Bakker y Currie, 2001).

2.2.1. Muestreo y cuantización de las señales

Existen diferentes intervalos de muestreo independientes para los radares pulsados (Richards, 2005). Aquí se identifican tres que corresponden a los utilizados en los conjuntos de datos que se emplearán, los cuales se ilustran en la figura 2.2.

El radar emite una serie periódica de pulsos, este período se denomina intervalo de repetición de pulso (IRP), y su inversa es la frecuencia de repetición de pulso (FRP). En general la FRP toma valores en el rango de frecuencias que va de cientos de Hz a cientos de kHz, mientras que la duración de los pulsos va de 100 ns a 100 μ s. En el caso del radar IPIX el valor máximo de

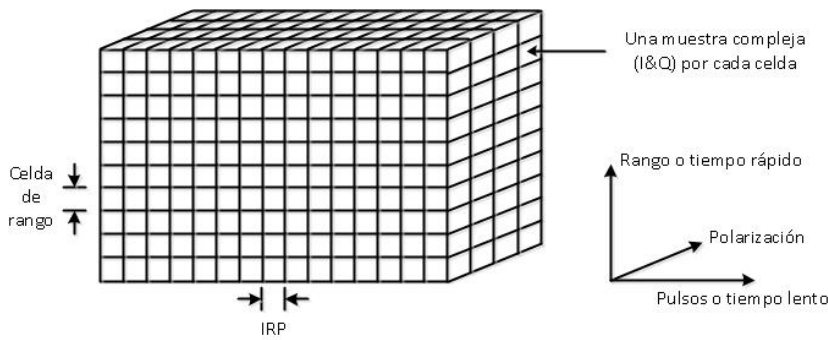


Figura 2.2: Estrategia de almacenamiento de los datos.

la FRP es de 20 kHz y la duración de los pulsos puede variarse de 20 a 5000 ns (Drosopoulos, 1994).

En una porción del intervalo de tiempo entre pulsos, la señal recibida en cada polarización es muestreada a la salida del filtro adaptado a una tasa de muestreo alta, típicamente en el orden de algunos MHz, siendo como máximo de 50 MHz para el IPIX (Drosopoulos, 1994). Cada una de estas muestras tiene su correspondencia con la distancia que viajó la onda transmitida hasta el objetivo donde se produjo su reflexión, debido a que en el tiempo transcurrido hasta que la señal retorna al radar la onda recorre dos veces esta distancia a la velocidad de la luz. Por este motivo a estas muestras se las denomina celdas de rango.

Se puede pensar entonces que el conjunto total de muestras son almacenadas en un arreglo de tres dimensiones, donde se tiene una matriz por cada polarización. Las columnas las conforman las celdas de rango para un dado pulso y las filas corresponden a muestras de la misma celda de rango para los distintos pulsos. Debido a que el radar puede transmitir y recibir las señales en dos polarizaciones lineales, en el caso más general son 4 las matrices que conforman este arreglo tridimensional. De aquí en adelante, se denotará como XY a la matriz correspondiente a las muestras de la señal transmitida con polarización X y recibida con polarización Y, donde X e Y pueden ser H (horizontal) o V (vertical) según corresponda.

Esta estrategia de almacenamiento facilitará la comprensión de las operaciones en el procesamiento de las muestras a lo largo del trabajo. En general se denomina dimensión en rango, o tiempo rápido, a la dimensión vertical y dimensión en tiempo, o tiempo lento, a la dimensión horizontal debido a la gran diferencia entre el intervalo de muestreo entre celdas de rango y el IRP (Richards, 2005).

Las muestras de este arreglo son complejas, donde la parte real representa la componente en fase y la parte imaginaria la componente en cuadratura de la señal en banda base, respectivamente. En el caso del IPIX cada una de estas componentes es cuantizada en 8 bits (Drosopoulos, 1994).

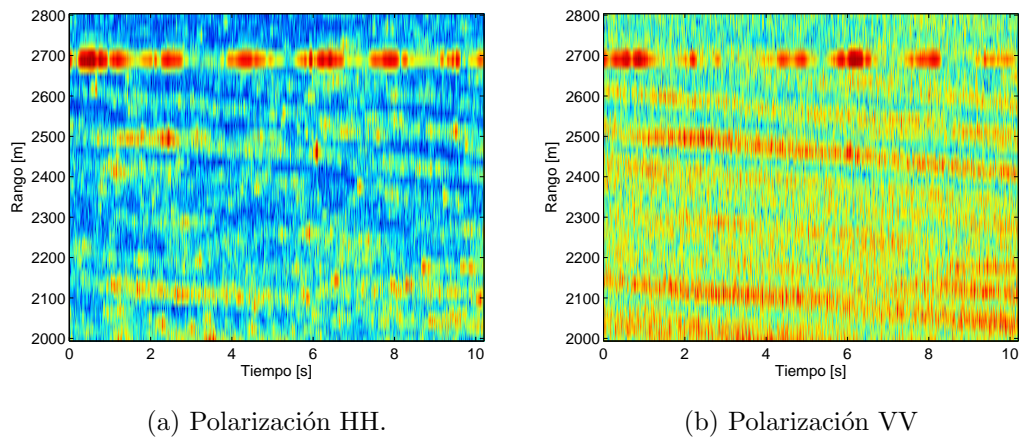


Figura 2.3: Intensidad de las mediciones correspondientes al conjunto de datos *stare0*.

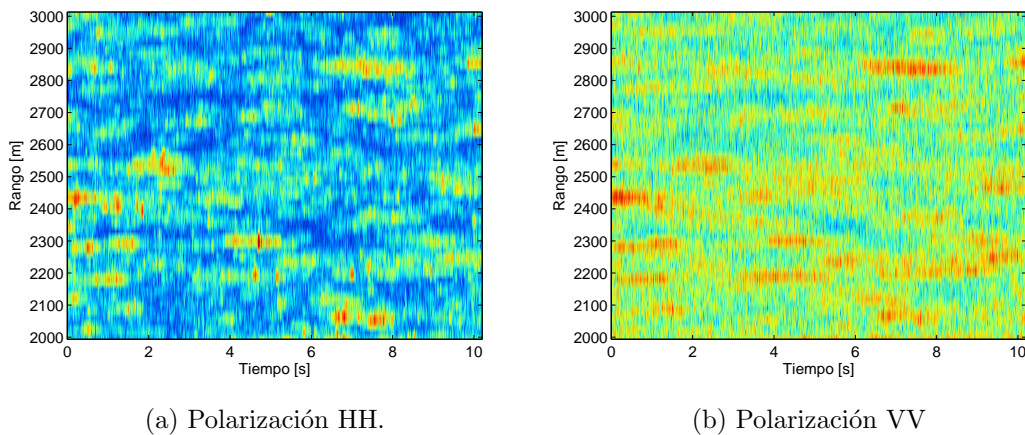


Figura 2.4: Intensidad de las mediciones correspondientes al conjunto de datos *stare6*.

2.2.2. Datos utilizados

A continuación se resumen los parámetros de las mediciones utilizadas a lo largo de la tesis para validar los resultados. El primer conjunto de datos denominado *stare0* corresponde a mediciones realizadas el día 11 de noviembre de 1993 a las 16:11 hs. En la figura 2.3 se muestran los gráficos de intensidad de las mediciones para las polarización HH y VV. Estos datos corresponden a reflexiones de un pequeño objetivo inmerso en clutter inhomogéneo. El objetivo consiste de una pelota de playa envuelta en papel de aluminio. Su ubicación nominal es 2691 m. Los demás conjuntos de datos utilizados fueron registrados el día 10 de noviembre de 1993 a las 00:34:24, a las 00:36:51 y a las 00:39:10 hs. y corresponden a los conjuntos de datos denominados *stare6*, *stare7* y *stare8* respectivamente. En la figura 2.4 se muestran los gráficos de intensidad de las mediciones de los datos *stare6* para las polarización HH y VV. Estos datos corresponden a reflexiones de clutter marítimo inhomogéneo sin objetivo presente. En todos los casos los datos fueron registrados con el radar IPIX a 30 m de altura iluminando una región fija del mar. En la tabla 2.1 se detallan los parámetros técnicos del radar y las condiciones del mar en el momento que se realizaron las mediciones (Bakker y Currie, 2001).

| Conjunto de datos | <i>stare0</i> | <i>stare6</i> | <i>stare7</i> | <i>stare8</i> |
|--------------------------------|---------------|---------------|---------------|---------------|
| Período de las olas | 8 s | 4 s | 4 s | 4 s |
| Altura de las olas | 1 m | 1.2 m | 1.2 m | 1.2 m |
| PRF | 200 Hz | 500 Hz | 500 Hz | 500 Hz |
| Número de pulsos | 2048 | 8192 | 8192 | 8192 |
| Duración del pulso | 200 ns | 200 ns | 200 ns | 200 ns |
| Frecuencia | 9.39 GHz | 9.39 GHz | 9.39 GHz | 9.39 GHz |
| Resolución en rango | 30 m | 30 m | 30 m | 30 m |
| Intervalo de muestreo en rango | 15 m | 15 m | 15 m | 15 m |
| Número de celdas de rango | 54 | 68 | 68 | 68 |

Tabla 2.1: Parámetros de los datos utilizados.

En las figuras 2.5 y 2.6 se presentan estimaciones de las correlaciones de los datos en tiempo, o tiempo lento de acuerdo con la definición de la sección 2.2.1, para las polarizaciones HH y VV de los conjuntos de datos *stare0* y *stare6*, respectivamente. Las mismas fueron estimadas como el promedio de las correlaciones muestrales obtenidas para cada celda de rango. En el caso del conjunto de datos *stare0* se consideró hasta la celda de rango 45, excluyendo de este modo las celdas de rango correspondientes a reflexiones en el objetivo. Como se puede observar en ambos conjuntos de datos, luego de 25 ms para la polarización VV y 20 ms para la polarización HH la función de correlación es aproximadamente cero, es decir que las muestras de dos pulsos separados en un tiempo mayor al mencionado no están correlacionadas. La diferencia en la cantidad de muestras, o pulsos, dentro de los intervalos de tiempo mencionados entre las correlaciones estimadas para los distintos conjuntos de datos, se debe a que no es la misma la FRP utilizada en cada caso. Por otro lado, haciendo un análisis análogo en rango se obtiene que la correlación de los datos en esta dimensión es de dos a tres celdas. Esto es consecuencia de que en ambos casos las muestras fueron tomadas cada 15 m cuando la resolución del radar IPIX es de 30 m.

2.3. Modelos existentes

El procesamiento estadístico de las señales recibidas por el radar permite detectar la presencia de objetivos y extraer su información. En este sentido para obtener algoritmos eficientes, es fundamental utilizar modelos realistas de las señales recibidas. Por este motivo, el modelado del clutter ha sido ampliamente estudiado en la literatura. Naturalmente, debido a su simplicidad en el tratamiento matemático el modelo Gaussiano para los datos, o Rayleigh para su intensidad, es la primera opción (Marcum, 1960; Swerling, 1960). Sin embargo, el análisis de mediciones muestra que en casos como los de clutter marítimo o de clutter en radares de alta resolución, la estadística de los datos se aparta del modelo Rayleigh. En estudios más recientes se ha encontrado que existe una familia de distribuciones que logran un buen ajust-

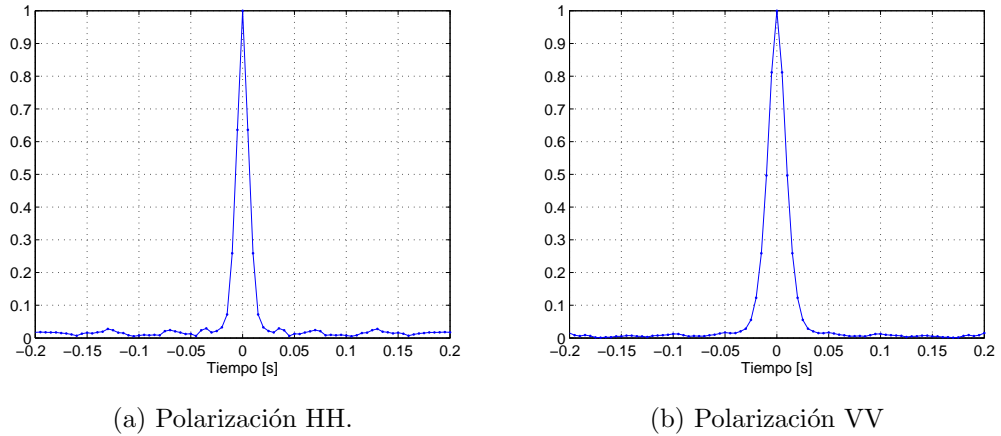


Figura 2.5: Correlación estimada del conjunto de datos *stare0*, sin considerar las celdas de rango correspondientes al objetivo.

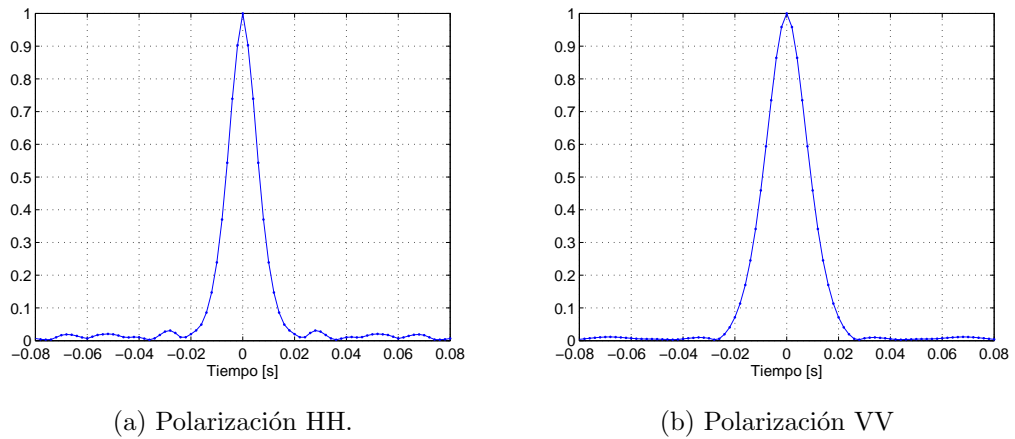


Figura 2.6: Correlación estimada del conjunto de datos *stare6*.

te de los datos en distintas condiciones entre las que se incluyen la Weibull y los modelos Gaussianos-compuestos (Farina *et al.*, 1997; Anastassopoulos *et al.*, 1999). Para la aplicación efectiva de los modelos teóricos es necesario validarlos por medio de ajustes a los datos reales en distintas condiciones ambientales. A continuación se realiza el ajuste de los datos del radar IPIX descritos anteriormente a las distribuciones Rayleigh, Weibull y K .

Suponiendo que \mathbf{v} es un vector de $N \times 1$ que contiene las muestras de clutter medidas y denotando v_i como el i -ésimo elemento de \mathbf{v} , se define A como la variable aleatoria que describe la intensidad de v_i , es decir $A = |v_i|$.

2.3.1. Modelo Rayleigh

La distribución Rayleigh para la intensidad de un conjunto de datos surge cuando se supone que los datos tienen distribución Gaussiana. Para el caso en que $v_i \sim \mathcal{CN}(0, \sigma^2)$ la fdp de A es (Papoulis, 1991)

$$f_A(a) = \frac{2a}{\sigma^2} e^{-a^2/\sigma^2} u(a), \quad (2.1)$$

donde σ se conoce como parámetro de escala de la distribución Rayleigh y $u(a)$ es la función escalón unitario. Los momentos de esta distribución son (Papoulis, 1991)

$$\mathbb{E}\{A^n\} = \int_{-\infty}^{\infty} a^n f_A(a) da = \sigma^n \Gamma(n/2 + 1), \quad (2.2)$$

donde $\Gamma(t) = \int_0^{\infty} x^{t-1} e^{-x} dx$ es la función Gamma (Abramowitz y Stegun, 1972).

2.3.2. Modelo Weibull

La distribución Weibull es una generalización de la distribución Rayleigh cuya fdp está dada por (Shnidman, 1999)

$$f_A(a) = \frac{c}{b} \left(\frac{a}{b}\right)^{c-1} e^{-(a/b)^c} u(a), \quad (2.3)$$

donde b es el parámetro de escala y c es el parámetro de forma de la distribución Weibull. Notar que cuando $c = 2$, esta distribución coincide con la distribución Rayleigh. Los momentos de la distribución Weibull son (Gini *et al.*, 2000)

$$\mathbb{E}\{A^n\} = \int_{-\infty}^{\infty} a^n f_A(a) da = b^n \Gamma(n/c + 1). \quad (2.4)$$

2.3.3. Modelo K

Los modelos Gaussianos-compuestos surgen de considerar cada muestra de clutter como el producto de dos variables aleatorias independientes

$$v_i = \sqrt{\tau} z_i, \quad (2.5)$$

donde $z_i \sim \mathcal{CN}(0, 1)$ se denomina speckle y la variable aleatoria τ se llama textura y representa la potencia local de esa muestra.

Dependiendo de la distribución que se le asigne a τ se tiene un modelo Gaussiano-compuesto u otro. En particular, cuando τ tiene distribución Gamma, se obtiene lo que se conoce como modelo o distribución K . En este caso, la fdp de τ está dada por

$$f_{\tau}(\tau) = \frac{\nu}{\mu \Gamma(\nu)} \left(\frac{\nu \tau}{\mu}\right)^{\nu-1} e^{-\nu \tau / \mu} u(\tau). \quad (2.6)$$

Notar que $v_i | \tau \sim \mathcal{CN}(0, \tau)$, lo que implica que la densidad condicional $f_{A|\tau}(a|\tau)$ es la fdp de una distribución Rayleigh con parámetro de escala τ . Entonces la fdp de A se obtiene como $\mathbb{E}_{\tau}\{f_{A|\tau}(a|\tau)\}$, resultando (Gini *et al.*, 2000)

$$f_A(a) = \frac{\sqrt{4\nu/\mu}}{2^{\nu-1} \Gamma(\nu)} \left(a \sqrt{\frac{4\nu}{\mu}}\right)^{\nu} K_{\nu-1} \left(a \sqrt{\frac{4\nu}{\mu}}\right) u(\tau), \quad (2.7)$$

donde μ es el parámetro de escala, ν es el parámetro de forma de la distribución K y $K_\nu(x)$ es la función de Bessel modificada de tercer tipo de orden ν (Lebedev, 1965). Los momentos de la distribución K son (Gini *et al.*, 2000)

$$\mathbb{E}\{A^n\} = \int_{-\infty}^{\infty} a^n f_A(a) da = \frac{\Gamma(n/2 + 1)\Gamma(n/2 + \nu)}{\Gamma(\nu)} \left(\frac{\mu}{\nu}\right)^{n/2}. \quad (2.8)$$

2.3.4. Ajuste a los datos reales

Cuando se habla de ajuste se hace referencia a la estimación de los parámetros de la densidad correspondiente utilizando como observaciones mediciones reales de clutter. El método más popular para realizar la estimación de los parámetros es el método de máxima verosimilitud. Dada la función de densidad de probabilidad conjunta $f(\mathbf{x}; \boldsymbol{\theta})$ parametrizada por el vector de parámetros $\boldsymbol{\theta}$, a la que se denomina función de verosimilitud, se define el estimador de máxima verosimilitud $\hat{\boldsymbol{\theta}}$ (MLE por sus siglas en inglés) como el valor de $\boldsymbol{\theta}$ que maximiza $f(\mathbf{x}; \boldsymbol{\theta})$ (Kay, 1993). En general, cuando existe y se puede hallar, el MLE posee asintóticamente, propiedades que son deseables para un estimador. El MLE es asintóticamente insesgado y eficiente, alcanza la cota de Crámer-Rao, y tiene distribución asintótica Gaussiana. Se puede decir que es asintóticamente óptimo (Kay, 1993).

Considerando que se miden N intensidades independientes $x_i = |v_i|$ con $i = 1, \dots, N$, las mismas pueden agruparse en un vector de observaciones de $N \times 1$ cuyas entradas están idénticamente distribuidas con fdp $f_A(x_i; \boldsymbol{\theta})$. Luego, la función de verosimilitud de los datos es

$$f(\mathbf{x}; \boldsymbol{\theta}) = \prod_{i=1}^N f_A(x_i; \boldsymbol{\theta}), \quad (2.9)$$

donde $f_A(x_i; \boldsymbol{\theta})$ está dada por (2.1), (2.3) o (2.7) para la distribuciones Rayleigh, Weibull o K respectivamente. En el caso de la distribución Rayleigh el vector de parámetros se reduce a un escalar y está dado por $\theta = \sigma^2$ y para las distribuciones Weibull y K el vector de parámetros es $\boldsymbol{\theta} = [b \ c]^T$ y $\boldsymbol{\theta} = [\mu \ \nu]^T$ respectivamente.

Para la distribución Rayleigh el MLE de σ^2 tiene una expresión cerrada dada por

$$\hat{\sigma}^2 = \frac{1}{N} \sum_{i=1}^N |v_i|^2. \quad (2.10)$$

En el caso de la distribución Weibull al hallar el gradiente de la función de verosimilitud e igualarlo a cero resulta el siguiente sistema de ecuaciones (Ravid y Levanon, 1992)

$$\frac{1}{\hat{c}} = \frac{\sum_{i=1}^N \ln(|v_i|) |v_i|^{\hat{c}}}{\sum_{i=1}^N |v_i|^{\hat{c}}} - \frac{1}{N} \sum_{i=1}^N \ln(|v_i|) \quad (2.11)$$

$$\hat{b}^{\hat{c}} = \frac{1}{N} \sum_{i=1}^N |v_i|^{\hat{c}}. \quad (2.12)$$

Para obtener \hat{c} es necesario resolver numéricamente (2.11), empleando, por ejemplo, el método de Newton-Raphson (Burden y Faires, 2011). Finalmente, reemplazando su valor en (2.12) se obtiene \hat{b} .

| Distribución | Parámetro | Estimados | |
|--------------|-----------|-------------------------|-------------------------|
| | | <i>stare0</i> , pol. VV | <i>stare6</i> , pol. VV |
| Rayleigh | σ | 0.135 | 0.070 |
| Weibull | b | 0.119 | 0.062 |
| | c | 1.433 | 1.445 |
| K | μ | 0.018 | 0.005 |
| | ν | 1.152 | 1.140 |

 Tabla 2.2: Estimados de los parámetros de las distribuciones Rayleigh, Weibull y K .

En la estimación de los parámetros de la distribución K la resolución del sistema de ecuaciones no lineal resultante es considerablemente más compleja que este último caso. Por este motivo habitualmente se recurre a un método alternativo para obtener $\hat{\mu}$ y $\hat{\nu}$, conocido como método de los momentos. Este método consiste simplemente en igualar los momentos muestrales de los datos a los momentos teóricos de la distribución en cuestión. Aunque este estimador no posee condiciones de optimabilidad es simple de determinar y es útil cuando el conjunto de datos es suficientemente grande, debido a que por lo general es un estimador consistente (Kay, 1993).

Suponiendo que A es una variable aleatoria con distribución K , entonces a partir de (2.8) con $n = 2$ se obtiene que $\mathbb{E}\{A^2\} = \mu$. Por lo tanto, utilizando el método de momentos el estimador de μ es

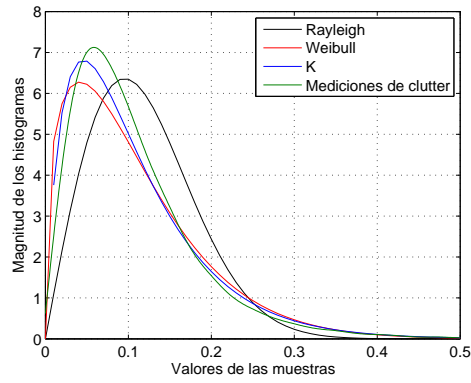
$$\hat{\mu} = \frac{1}{N} \sum_{i=1}^N |v_i|^2. \quad (2.13)$$

La ecuación restante, que permite obtener $\hat{\nu}$, surge de hallar la media utilizando nuevamente (2.8), con $n = 1$, e igualar el resultado a la media muestral de los datos, de esta manera la misma resulta

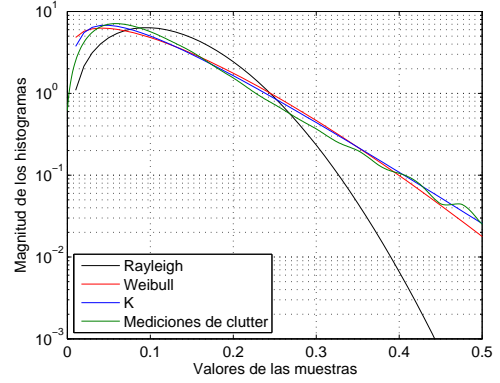
$$\sqrt{\frac{\pi \hat{\mu}}{4 \hat{\nu}}} \frac{\Gamma(1/2 + \hat{\nu})}{\Gamma(\hat{\nu})} = \frac{1}{N} \sum_{i=1}^N |v_i|. \quad (2.14)$$

La ecuación (2.14) también debe resolverse numéricamente debido a que no tiene solución analítica.

En la tabla 2.2 se presentan los estimados de los parámetros de las distribuciones descriptas, obtenidos con la polarización VV de los conjuntos de datos *stare0* y *stare6*. En las figuras 2.7 y 2.8 se muestran las fdp's de cada una de estas distribuciones con los valores de los parámetros estimados, junto con los histogramas de los datos correspondientes para la polarización VV de los conjuntos de datos *stare0* y *stare6*, respectivamente, en escala lineal y en escala logarítmica. En estos últimos se observa el parecido entre los histogramas de los datos reales y las densidades de las distribuciones Weibull y K , en particular en lo que respecta a las colas de las fdp's. Por el contrario la distribución Rayleigh es la que peor ajusta los datos, debido a que no puede representar adecuadamente la impulsividad de los mismos.

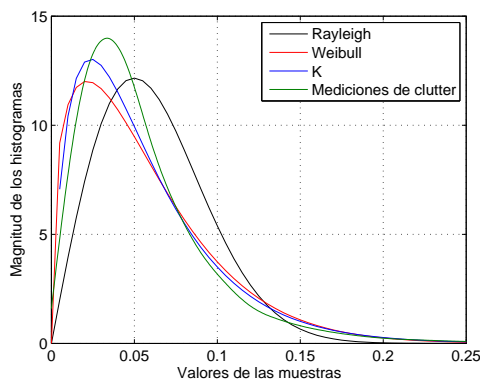


(a) Escala lineal.

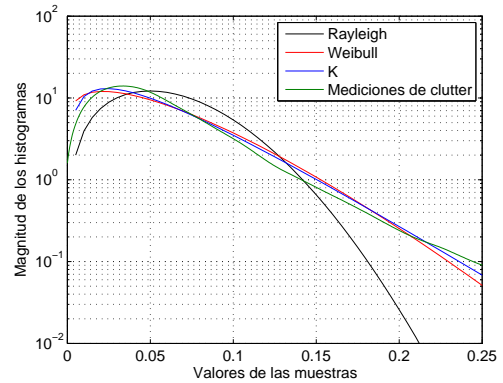


(b) Escala logarítmica.

Figura 2.7: Funciones densidad de probabilidad resultantes del ajuste e histograma correspondiente al conjunto de datos *stare0* para la polarización VV.



(a) Escala lineal.



(b) Escala logarítmica.

Figura 2.8: Funciones densidad de probabilidad resultantes del ajuste e histograma correspondiente al conjunto de datos *stare6* para la polarización VV.

2.4. Detección

Una de las operaciones básicas que desempeña el radar es la detección (Skolnik, 2001). Esto quiere decir, decidir si la señal que arriba es el resultado de la reflexión en un objetivo o si corresponde simplemente a interferencia. Debido a que tanto las reflexiones en los objetivos como el clutter son de naturaleza aleatoria su representación más conveniente es a través de modelos estadísticos de señales (Richards, 2005). En consecuencia, el proceso de detección es un problema de test de hipótesis (Kay, 1998).

Como se mencionó, el radar transmite un pulso y el receptor muestrea la señal reflejada a la salida del filtro adaptado. Denotando \mathbf{y} al vector de muestras complejas a la salida del demodulador en cuadratura, el procedimiento de detección consiste en decidir entre dos hipótesis H_0 y H_1 luego de haberse observado \mathbf{y} para la celda de rango bajo prueba

$$\begin{aligned} H_0 : \mathbf{y} &= \mathbf{v} \\ H_1 : \mathbf{y} &= \mathbf{s} + \mathbf{v}. \end{aligned} \tag{2.15}$$

Bajo la hipótesis nula H_0 se supone que los datos consisten solo de clutter \mathbf{v} , mientras que para la hipótesis H_1 se considera que las mediciones están conformadas por la combinación de clutter y las reflexiones en un objetivo \mathbf{s} .

El análisis comienza con la descripción de las funciones de densidad de probabilidad de las mediciones para cada una de hipótesis. En este caso las fdp's requeridas son la de \mathbf{y} dado que el objetivo no está presente, $f_y(\mathbf{y}|H_0)$, y la de \mathbf{y} dado que el objetivo está presente, $f_y(\mathbf{y}|H_1)$. Gran parte del problema de detección consiste en definir estas densidades en base a las propiedades estadísticas del clutter y de la señal reflejada en el objetivo.

En base a este planteo se pueden definir dos probabilidades que son de interés para evaluar el desempeño de los esquemas de detección. Una de ellas es la probabilidad de detección, P_D , que es la probabilidad de decidir que el objetivo está presente cuando efectivamente lo está. La otra es la probabilidad de falsa alarma, P_{FA} , que es la probabilidad de decidir que el objetivo está presente cuando en realidad no lo está.

2.4.1. Criterio de Neyman-Pearson

Uno de los objetivos en el diseño de un sistema de radar es utilizar adecuadamente la información proveniente de las densidades de los datos para lograr un buen desempeño en el proceso de detección. En tal sentido, el siguiente paso es definir un criterio que establezca la elección óptima entre las dos hipótesis. Por lo general, en aplicaciones de radar, dentro del enfoque clásico de test de hipótesis se utiliza el criterio de Neyman-Pearson (Kay, 1998). En este caso, el algoritmo de detección se diseña de modo de maximizar la probabilidad de detección bajo la restricción de que la probabilidad de falsa alarma no exceda un determinado valor constante. La solución a este problema de optimización da como resultado la regla de decisión que se conoce como test de relación de verosimilitud (Van Trees, 2001)

$$\Lambda(\mathbf{y}) = \frac{f_y(\mathbf{y}|H_1)}{f_y(\mathbf{y}|H_0)} \underset{H_0}{\overset{H_1}{\gtrless}} \eta, \tag{2.16}$$

donde $\Lambda(\mathbf{y})$ es la relación de verosimilitud y η es el umbral de decisión, el cual se obtiene a partir del valor establecido para la probabilidad de falsa alarma (Van Trees, 2001)

$$P_{FA} = \int_{\{\mathbf{y}:\Lambda(\mathbf{y})>\eta\}} f_y(\mathbf{y}|H_0)d\mathbf{y}. \quad (2.17)$$

La regla (2.16) establece que la relación entre las fdp's, cada una evaluada en el valor particular de los datos observados \mathbf{y} , debe compararse con un umbral para tomar una decisión. Si la relación de verosimilitud excede dicho umbral se debe elegir la hipótesis H_1 , es decir declarar que un objetivo está presente. Si por el contrario, no excede el umbral, se debe optar por la hipótesis H_0 y declarar que no existe un objetivo presente.

2.4.2. Test de relación de verosimilitud generalizado

La regla de decisión presentada en la sección anterior supone que las densidades de probabilidad son completamente conocidas. Esto por lo general no es cierto, las fdp's contienen parámetros cuyos valores son desconocidos. Estas situaciones corresponden a las llamadas hipótesis compuestas y dan lugar a tests especiales (Kay, 1998). Uno de los test de hipótesis compuestas más utilizados en la práctica es el test de relación de verosimilitud generalizado (GLRT por sus siglas en inglés). En el GLRT se reemplazan los parámetros desconocidos por sus estimados de máxima verosimilitud. Aunque no existen condiciones de optimalidad asociadas a este test, su uso se debe a que en la práctica funciona bien. En general, el GLRT decide por la hipótesis H_1 cuando (Kay, 1998)

$$\Lambda(\mathbf{y}) = \frac{f_y(\mathbf{y}; \hat{\boldsymbol{\theta}}_1|H_1)}{f_y(\mathbf{y}; \hat{\boldsymbol{\theta}}_0|H_0)} > \eta, \quad (2.18)$$

donde $\hat{\boldsymbol{\theta}}_i$ es el MLE del vector de parámetros $\boldsymbol{\theta}_i$ de la fdp de las mediciones bajo la hipótesis H_i , es decir maximiza $f_y(\mathbf{y}; \boldsymbol{\theta}_i|H_i)$ para $i = 0, 1$.

2.4.3. Enfoque Bayesiano

El GLRT considera que los parámetros desconocidos de las densidades de probabilidad son determinísticos. En cambio, el enfoque Bayesiano considera que los parámetros desconocidos de las fdp's son realizaciones de variables aleatorias a las que se les asigna una fdp a priori. Llamando $f_0(\boldsymbol{\theta}_0)$ y $f_1(\boldsymbol{\theta}_1)$ a las densidades de probabilidad a priori para las hipótesis H_0 y H_1 respectivamente, luego las fdp's de los datos están dadas por

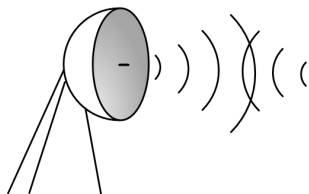
$$f_y(\mathbf{y}|H_i) = \int f_{y|\boldsymbol{\theta}_i}(\mathbf{y}|\boldsymbol{\theta}_i, H_i)f_i(\boldsymbol{\theta}_i)d\boldsymbol{\theta}_i, \quad \text{para } i = 0, 1 \quad (2.19)$$

donde $f_{y|\boldsymbol{\theta}_i}(\mathbf{y}|\boldsymbol{\theta}_i, H_i)$ es la densidad de probabilidad condicional de \mathbf{y} dado $\boldsymbol{\theta}_i$, considerando que H_i es la hipótesis verdadera. Las densidades de probabilidad $f_y(\mathbf{y}|H_0)$ y $f_y(\mathbf{y}|H_1)$ quedan completamente determinadas, es decir no dependen de los parámetros desconocidos. Por lo tanto, con el enfoque Bayesiano el detector óptimo de Neyman-Pearson decide de manera usual, según la regla (2.16).

De acuerdo con el modelo adoptado para los datos resultará conveniente la utilización de uno u otro de los enfoques presentados.

2.5. Conclusiones

En este capítulo se detallaron las características del radar IPIX y los conjuntos de datos que se emplean a lo largo del desarrollo de la tesis para validar los modelos y los métodos de detección propuestos. Se presentaron las distribuciones Rayleigh, Weibull y K utilizadas en la literatura para modelar el clutter, en particular, marítimo. Se describieron los procedimientos de estimación de sus parámetros y se ajustaron a los conjuntos de datos del radar IPIX. De los resultados obtenidos es importante destacar la deficiencia de la fdp de la distribución Rayleigh en el ajuste a las colas pesadas de los histogramas de los datos. Sin embargo, el motivo de la elección de estas distribuciones no se centra solo en la calidad del ajuste a los datos sino que además poseen esquemas de detección asociados, los cuales serán utilizados como referencia para comparación con los métodos de detección propuestos en la tesis. Finalmente se hizo una breve descripción de las técnicas de test de hipótesis que serán retomadas en diferentes instancias del trabajo para deducir los esquemas de detección para los modelos de clutter propuestos.



Capítulo 3

Detección en rango

3.1. Introducción

En las aplicaciones de radar la señal de interés suele encontrarse oscurecida por clutter. Salvo algunas excepciones, el clutter es considerado una fuente de interferencia cuyos efectos se deben eliminar o reducir. En vista de su naturaleza aleatoria, para modelizar los datos de clutter los métodos convencionales utilizan distintas distribuciones. Dependiendo de la aplicación y el escenario de operación las distribuciones más utilizadas son la Gaussiana, Log-normal, Weibull, K (Shnidman, 1999; Sangston y Gerlach, 1994), Pareto (Weinberg, 2011), la función densidad de probabilidad compuesta generalizada (Anastassopoulos *et al.*, 1999; Sangston *et al.*, 2012; Shang y Song, 2011; Cui *et al.*, 2012), o bien el clutter se modela como un proceso aleatorio esféricamente invariante (Conte y Longo, 1987). Como se vió en el capítulo 2 algunas de estas distribuciones logran un buen ajuste a la distribución del clutter, sin embargo son independientes del tiempo y en muchos casos el entorno del radar puede cambiar abruptamente (Li *et al.*, 2009) llevando a una degradación en el desempeño del radar en un escenario real. Por este motivo se propone modelar el clutter como un proceso que presente heteroscedasticidad condicional autorregresiva generalizada (GARCH por sus siglas en inglés) (Engle, 1982; Bollerslev, 1986). Estos modelos utilizan la historia del proceso para mejorar su caracterización en el instante actual y las predicciones futuras. Han sido utilizados para modelar registros de índices financieros cuyas varianzas fluctúan temporalmente y recientemente también han sido usados en el modelado de ruido producido bajo el agua en aplicaciones de sonar (Amiri *et al.*, 2007). Dos de sus principales características son que sus funciones de densidad de pro-

babilidad presenta colas pesadas, propiedad buscada en las distribuciones mencionadas para el modelado de clutter (Anastassopoulos *et al.*, 1999; Shnidman, 1999; Sangston y Gerlach, 1994; Weinberg, 2011) y el agrupamiento de la volatilidad, es decir grandes cambios tienden a ser seguidos de grandes cambios y pequeños cambios tienden a ser seguidos de pequeños cambios, característica compatible con varios tipos de clutter en entornos naturales.

En la sección 3.2 se da una breve introducción a los procesos GARCH y se describe el método de cuasi-máxima verosimilitud para la estimación de sus parámetros. Este procedimiento de estimación ha sido descrito en detalle por Berkes *et al.* (2003); Straumann (2005) y ha sido utilizado en diferentes aplicaciones (Amiri *et al.*, 2007; Mousazadeh y Cohen, 2010).

En problemas de estimación paramétrica, bajo ciertas condiciones la cota de Cramér-Rao (CRB por sus siglas en inglés) es la cota mínima para la varianza del error de un estimador insesgado (Kay, 1993). De este modo, la CRB es una herramienta que permite evaluar la precisión con la que se puede estimar un parámetro desconocido, como así también realizar evaluaciones de desempeño de distintos estimadores. Para su cálculo es necesario conocer la distribución de las variables aleatorias involucradas. Una de las dificultades que presentan los procesos GARCH es justamente que no existe una expresión explícita para su función densidad de probabilidad. Sin embargo, bajo ciertas condiciones es posible determinar sus momentos. Por este motivo en la sección 3.3 se presenta una aproximación de la cota de Cramér-Rao para la estimación de parámetros de un proceso ARCH(1), que solo requiere conocer algunos momentos de la variable aleatoria considerada (Pascual *et al.*, 2013a). Esta aproximación resulta de considerar la distribución de máxima entropía dado un conjunto de momentos conocidos y de determinar la CRB a partir de ella. Es atractivo emplear la distribución de máxima entropía pues se trata de la distribución que realiza la mínima cantidad de suposiciones, sujeta a las restricciones de los momentos (Dacunha-Castelle y Gamboa, 1990). Se realiza una comparación de las dos variantes presentadas en (von Ellenrieder y Muravchik, 2008) para su cálculo, a partir de un ejemplo numérico en un caso particular. Para tal fin se deducen las expresiones necesarias de los momentos y sus derivadas respecto de los parámetros del proceso a estimar. Finalmente se verifica la aproximación realizada para la fdp en el cálculo de la CRB considerando las propiedades asintóticas del estimador de cuasi-máxima verosimilitud (QMLE por sus siglas en inglés).

En la sección 3.4 se realiza el ajuste del modelo GARCH a los datos de radar IPIX descritos en la sección 2.2.2. Con el objetivo de tener una idea del número de muestras necesarias en el proceso de ajuste se realiza un estudio del error de estimación del QMLE en función del número de muestras del vector de observaciones. Este análisis revela que para lograr un error de estimación razonable es necesario utilizar un número de muestras espaciales que puede resultar mayor al disponible en algunos escenarios con datos reales. Para sortear esta dificultad en la sección se presenta una solución alternativa utilizando distintas realizaciones temporales independientes (Pascual *et al.*, 2011). Análogamente se estudia el error de estimación en función del número de realizaciones así como también del número de muestras de cada una de ellas. Finalmente, en la sección 3.4.3 se realiza el ajuste a datos reales de clutter marítimo, utilizando

la propuesta alternativa mencionada, y se muestran los resultados obtenidos (Pascual *et al.*, 2011, 2013b).

Por último en la sección 3.5 se desarrolla un esquema de detección basado en un modelo GARCH para el clutter (Pascual *et al.*, 2013b). En este caso se modela el clutter considerando las realizaciones en rango (o tiempo rápido), es decir que se tiene una realización para cada pulso transmitido y la detección se lleva a cabo celda a celda para todas las celdas de rango en cada pulso. Este esquema posee la ventaja de que no requiere esperar el arribo de las reflexiones de múltiples pulsos para realizar la decisión. Se lleva a cabo un análisis teórico del desempeño del esquema de detección y por medio de simulaciones Monte Carlo se evalúan sus probabilidades de detección y falsa alarma. Finalmente se pone a prueba el detector en una situación real utilizando las mediciones marítimas reales. El desempeño del detector se compara con dos detectores que funcionan con la misma filosofía. Uno de ellos surge de considerar que las muestras de clutter poseen distribución Gaussiana o bien que su intensidad tiene distribución Rayleigh, mientras que para el tercer detector se supone que la intensidad de las muestras del clutter tiene distribución Weibull. Para valores bajos de probabilidad de falsa alarma el detector GARCH presenta una mayor probabilidad de detección independientemente del valor de relación señal a clutter que se considere. En condiciones de buena relación señal a clutter o para valores altos de probabilidad de falsa alarma los tres detectores presentan un desempeño semejante, es decir alcanzan aproximadamente la misma probabilidad de detección (Pascual *et al.*, 2013b).

3.2. Procesos GARCH

Los procesos aleatorios representados a través de series temporales que presentan heteroscedasticidad condicional autorregresiva (ARCH por sus siglas en inglés) fueron propuestos por Engle (1982) y han sido ampliamente utilizados en econometría para modelar diversos índices financieros. Se caracterizan por ser condicionalmente Gaussianos, con varianza condicional variante en el tiempo. Posteriormente Bollerslev (1986) introdujo una extensión de los procesos ARCH a los procesos GARCH, de manera similar a la extensión de un proceso AR, en series temporales, al proceso ARMA general. Estos permiten una estructura de retardos más flexible, en particular para estructuras de memorias largas. En general, se dice que un proceso aleatorio presenta heteroscedasticidad cuando su varianza cambia en el tiempo. En el caso de los procesos GARCH el hecho de que presenten heteroscedasticidad condicional implica una dependencia con observaciones del pasado inmediato y que sea autorregresiva describe un mecanismo de realimentación que incorpora información del pasado del proceso para describir su varianza presente.

3.2.1. Estructura

Un proceso GARCH(p, q) se define como (Bollerslev, 1986)

$$v_r = \sigma_r z_r \quad z_r \sim \mathcal{N}(0, 1) \text{ iid}, \quad (3.1)$$

$$\sigma_r^2 = k + \sum_{j=1}^p \alpha_j \sigma_{r-j}^2 + \sum_{j=1}^q \beta_j v_{r-j}^2, \quad (3.2)$$

donde k , α_j and β_j son los coeficientes del proceso. El proceso GARCH(p, q) modela las observaciones como un proceso de ruido blanco gaussiano con varianza condicional (o instantánea) variable (3.1), y el valor actual de la varianza condicional, σ_r^2 , es una función lineal de sus p valores pasados, $\sigma_{r-1}^2, \dots, \sigma_{r-p}^2$, más una combinación lineal de q valores pasados de la amplitud de las observaciones al cuadrado, $v_{r-1}^2, \dots, v_{r-q}^2$, (3.2).

En la figura 3.1 se muestran realizaciones de dos procesos GARCH particulares, junto a sus varianzas condicionales. En ella se aprecian las características impulsiva y de agrupamiento de volatilidad que presentan estos procesos, y el efecto que produce cada coeficiente sobre su comportamiento.

En general, es posible asegurar propiedades deseables para estos procesos imponiendo condiciones sobre sus coeficientes. Las que surgen de manera natural para asegurar que la varianza condicional sea siempre positiva son

$$k > 0 \quad \alpha_j \geq 0, \quad j = 1, \dots, p \quad \beta_j \geq 0, \quad j = 1, \dots, q, \quad (3.3)$$

Definiendo a ψ_r como toda la información disponible hasta el instante r , es decir, σ_ρ^2 y v_ρ para $\rho \leq r - 1$, entonces

$$v_r | \psi_r \sim \mathcal{N}(0, \sigma_r^2). \quad (3.4)$$

Esto justifica el nombre de varianza condicional utilizado para σ_r^2 . A partir de (3.4) se puede demostrar que $\mathbb{E}\{v_r\} = 0$ (Bollerslev, 1986). Además, si se cumple que

$$\sum_{j=1}^p \alpha_j + \sum_{j=1}^q \beta_j < 1, \quad (3.5)$$

su varianza no condicional es finita e igual a (Bollerslev, 1986)

$$\mathbb{V}\{v_r\} = \frac{k}{1 - \sum_{j=1}^p \alpha_j - \sum_{j=1}^q \beta_j}. \quad (3.6)$$

resultando v_r estacionario en sentido amplio.

Para este tipo de procesos no es posible obtener una expresión explícita de su densidad de probabilidad. A pesar de esto, existen condiciones similares a (3.5) que aseguran la existencia de los momentos de cualquier orden del proceso (Engle, 1982; Bollerslev, 1986).

3.2.2. Estimación de cuasi-máxima verosimilitud

Uno de los objetivos de este trabajo consiste en modelar el clutter como un proceso GARCH, es decir realizar un ajuste de un modelo GARCH a muestras de señales proveniente de clutter

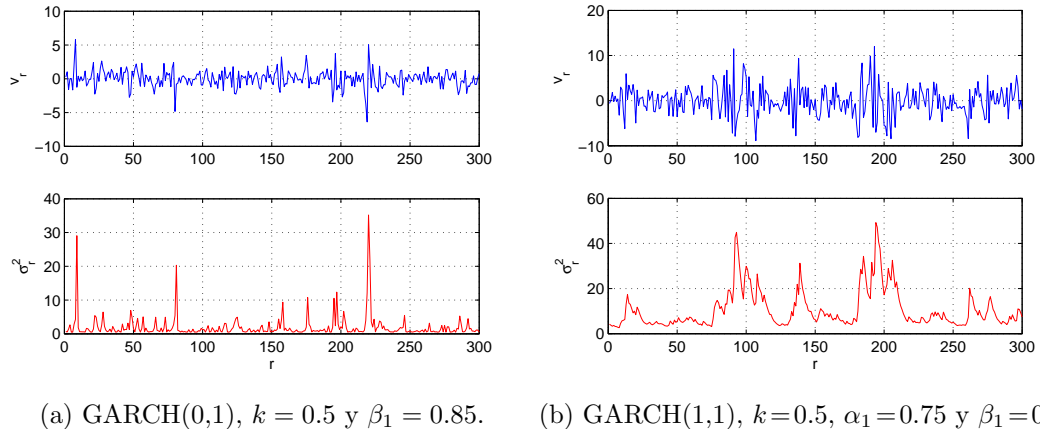


Figura 3.1: Realizaciones de dos casos particulares de proceso GARCH y sus varianzas condicionales.

recogidas por el radar. El ajuste consiste en estimar, a partir de los datos medidos, los valores de los coeficientes de la varianza condicional del proceso.

La manera natural de realizar dicha estimación sería a través del método de máxima verosimilitud. Sin embargo, en el caso de un vector $\mathbf{v} = [v_1 \cdots v_R]^T$, cuyos elementos corresponden a observaciones de un proceso GARCH(p, q), no existe una expresión explícita para su densidad de probabilidad, debido a que no se conoce la distribución de $[\sigma_1 \cdots \sigma_R]^T$. Para evitar esta dificultad en su lugar se utiliza la función de verosimilitud condicional de \mathbf{v} dadas las primeras $n = \max\{p, q\}$ observaciones, v_0, \dots, v_{1-n} (Shumway y Stoffer, 2006), la cual se puede escribir como

$$f(\mathbf{v}|v_0, \dots, v_{1-n}) = \prod_{r=1}^R f(v_r|v_{r-1}, \dots, v_{r-n}), \quad (3.7)$$

donde, a partir de (3.4), $f(v_r|v_{r-1}, \dots, v_{r-n})$ está dada por

$$f(v_r|v_{r-1}, \dots, v_{r-n}) = \frac{1}{\sqrt{2\pi\sigma_r^2}} e^{-\frac{v_r^2}{2\sigma_r^2}}. \quad (3.8)$$

Es importante destacar que para iniciar la recursión se requieren conocer los valores de $\sigma_0^2, \dots, \sigma_{1-n}^2$. En la bibliografía se sugieren diferentes alternativas para hacerlo; en el presente trabajo se optó por asignarles un valor igual a la varianza muestral de las realizaciones del proceso con las que se cuenta (Bollerslev, 1986). El problema es asimilable a conocer las muestras anteriores al inicio de la toma de observaciones en un modelo ARMA (Box *et al.*, 1994).

Por último, reemplazando (3.8) en (3.7) y tomando logaritmo natural se obtiene la expresión de la función de verosimilitud logarítmica

$$\ln(f(\mathbf{v}|v_0, \dots, v_{1-n})) = -\frac{R}{2} \ln(2\pi) - \frac{1}{2} \sum_{r=1}^R \left[\ln(\sigma_r^2) + \frac{v_r^2}{\sigma_r^2} \right], \quad (3.9)$$

donde $\boldsymbol{\theta} = [k \ \alpha_1 \ \cdots \ \alpha_p \ \beta_1 \ \cdots \ \beta_q]^T$ es el vector de parámetros que deseamos estimar.

Entonces, se define el estimador de cuasi-máxima verosimilitud, $\hat{\boldsymbol{\theta}}$, como el valor de $\boldsymbol{\theta}$ que maximiza la fdp condicional (3.7) ó, equivalentemente, que minimiza

$$\mathcal{L}_1(\boldsymbol{\theta}) = \frac{1}{2} \sum_{r=1}^R \left[\ln(\sigma_r^2) + \frac{v_r^2}{\sigma_r^2} \right] = \sum_{r=1}^R \ell_r(\boldsymbol{\theta}), \quad (3.10)$$

sujeto a las restricciones sobre los coeficientes (3.3) y (3.5). En síntesis, matemáticamente, se puede escribir de la siguiente manera

$$\begin{aligned} \hat{\boldsymbol{\theta}} &= \operatorname{argmin}_{\boldsymbol{\theta} \in \Theta} \mathcal{L}_1(\boldsymbol{\theta}), \end{aligned} \quad (3.11)$$

donde $\Theta \subset (0, \infty) \times \Omega$ con $\Omega = \{[\alpha_1 \cdots \alpha_p]^T \in [0, 1)^p \wedge [\beta_1 \cdots \beta_q]^T \in [0, 1)^q / \sum_{j=1}^p \alpha_j + \sum_{j=1}^q \beta_j < 1\}$.

3.2.3. Propiedades asintóticas del QMLE

Sea un proceso GARCH(p, q) estacionario, dado por (3.1) y (3.2), con $\boldsymbol{\theta}_v = [k^v \ \alpha_1^v \ \cdots \ \alpha_p^v \ \beta_1^v \ \cdots \ \beta_q^v]^T \in \Theta$ como el verdadero valor del vector de parámetros y considerando que los polinomios $\alpha^v(x) = \alpha_1^v x + \cdots + \alpha_p^v x^p$ y $\beta^v(x) = 1 - \beta_1^v x - \cdots - \beta_q^v x^q$ no poseen ceros comunes, entonces el QMLE $\hat{\boldsymbol{\theta}}$ obtenido de (3.11) es consistente (Berkes *et al.*, 2003), es decir

$$\hat{\boldsymbol{\theta}} \xrightarrow{cp1} \boldsymbol{\theta}_v \text{ cuando } R \rightarrow \infty. \quad (3.12)$$

Además, dado que en este caso las innovaciones z_r tienen distribución normal y suponiendo que $\boldsymbol{\theta}_v$ se encuentra en el interior de Θ , es posible demostrar que el estimador $\hat{\boldsymbol{\theta}}$ es, también, asintóticamente normal (Straumann, 2005)

$$\sqrt{R}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}) \xrightarrow{d} \mathcal{N}(0, \mathbf{F}^{-1}) \text{ cuando } R \rightarrow \infty, \quad (3.13)$$

con

$$\mathbf{F} = \mathbb{E} \left\{ \frac{\partial^2 \ell_0(\boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^T} \Big|_{\boldsymbol{\theta}=\boldsymbol{\theta}_v} \right\}, \quad (3.14)$$

donde $\ell_0(\boldsymbol{\theta})$ es $\ell_r(\boldsymbol{\theta})$ para r arbitrario.

Reemplazando la expresión de la derivada de $\ell_0(\boldsymbol{\theta})$, (3.14) se puede escribir como

$$\mathbf{F} = \frac{1}{2} \mathbb{E} \left\{ \frac{1}{\sigma_0^4(\boldsymbol{\theta}_v)} \left(\frac{\partial \sigma_0}{\partial \boldsymbol{\theta}} \Big|_{\boldsymbol{\theta}=\boldsymbol{\theta}_v} \right)^T \left(\frac{\partial \sigma_0}{\partial \boldsymbol{\theta}} \Big|_{\boldsymbol{\theta}=\boldsymbol{\theta}_v} \right) \right\}. \quad (3.15)$$

Existe la limitación de no poder determinar la matriz de covarianza \mathbf{F} de la distribución asintótica. Sin embargo, \mathbf{F} posee un estimador fuertemente consistente, dado por (Straumann, 2005)

$$\frac{1}{R} \frac{\partial^2 \mathcal{L}_1(\boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^T} \Big|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}} \xrightarrow{cp1} \mathbf{F} \text{ cuando } R \rightarrow \infty. \quad (3.16)$$

Lo que establece (3.16) es que cuando el número de muestras, R , es grande se puede realizar la aproximación

$$\mathbb{E} \left\{ \frac{\partial^2 \ell_0(\boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^T} \Big|_{\boldsymbol{\theta}=\boldsymbol{\theta}_v} \right\} \approx \frac{1}{R} \sum_{r=1}^R \frac{\partial^2 \ell_r(\boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^T} \Big|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}}. \quad (3.17)$$

3.3. Aproximación de la Cota de Crámer-Rao

En problemas de estimación paramétrica, bajo ciertas condiciones la cota de Cramér-Rao es una cota inferior para la matriz de covarianza del error de un estimador insesgado, $\hat{\boldsymbol{\theta}}$, de un parámetro, $\boldsymbol{\theta}$, la cual se define como (Kay, 1993)

$$\mathbb{E}\{(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})^T\} \geq \mathbf{CRB} = \mathbf{J}^{-1}, \quad (3.18)$$

donde $\mathbf{A} \geq \mathbf{B}$ significa que la matriz $\mathbf{A} - \mathbf{B}$ es positiva semi-definida. La matriz \mathbf{J} es la matriz de información de Fisher, cuyos elementos están dados por (Kay, 1993)

$$[\mathbf{J}]_{ij} \triangleq \mathbb{E} \left\{ \frac{\partial \ln(f(\mathbf{x}; \boldsymbol{\theta}))}{\partial \theta_i} \frac{\partial \ln(f(\mathbf{x}; \boldsymbol{\theta}))}{\partial \theta_j} \right\} \quad (3.19)$$

$$[\mathbf{J}]_{ij} = -\mathbb{E} \left\{ \frac{\partial^2 \ln(f(\mathbf{x}; \boldsymbol{\theta}))}{\partial \theta_i \partial \theta_j} \right\}, \quad (3.20)$$

donde $f(\mathbf{x}; \boldsymbol{\theta})$ es la función densidad de probabilidad conjunta de los datos $\mathbf{x} = [x_1 \ x_2 \ \cdots \ x_N]^T$.

La CRB es una herramienta que permite determinar la máxima precisión obtenible en la estimación de un parámetro desconocido cuando se usan estimadores insesgados, como así también realizar evaluaciones de desempeño de distintos estimadores. Como se observa en (3.19), o en (3.20), su cálculo depende directamente de la fdp de los datos. Sin embargo, existen casos donde la fdp no se conoce. Un ejemplo donde esto ocurre es el de los procesos GARCH. En particular, en esta sección se aborda el problema de aproximar la CRB en la estimación de los parámetros de un proceso ARCH (Pascual *et al.*, 2013a), cuya función densidad de probabilidad no tiene una expresión explícita. Para ello se seguirá la técnica esbozada en la Introducción: calcular la CRB a partir de la distribución de máxima entropía; y esta a partir de momentos.

3.3.1. Distribución de máxima entropía

Se conoce como problema de momentos al problema de hallar o aproximar una función de densidad de probabilidad dado un conjunto de momentos conocidos y existen diferentes enfoques para resolverlo (Blinnikov y Moessner, 1998; Talenti, 1987; Ormoneit y White, 1999). Una solución óptima desde el punto de vista de la teoría de la información es obtener la distribución de máxima entropía cuyos momentos sean iguales a los momentos conocidos. La entropía diferencial de una variable aleatoria X con fdp $f_X(x)$ está dada por

$$h(X) = \mathbb{E}\{-\ln(f_X(x))\}. \quad (3.21)$$

Se considera que los M primeros momentos no centrales existen y son conocidos, es decir $\eta_n = \mathbb{E}\{X^n\}$, $n = 1, \dots, M$, donde $\mathbb{E}\{\cdot\} = \int \cdot f_X(x) dx$ es el operador esperanza. Luego, la solución de máxima entropía al problema de momentos es un problema de optimización y su solución se obtiene minimizando el siguiente funcional

$$J(f, \boldsymbol{\lambda}) = \int_S \ln(f(x)) f(x) dx + \sum_{n=0}^M \lambda_n \left(\eta_n - \int_S x^n f(x) dx \right), \quad (3.22)$$

donde $f(x)$ es una fdp arbitraria, $S \in \mathbb{R}$ es el soporte de la fdp y $\boldsymbol{\lambda} = [\lambda_0 \cdots \lambda_M]^T$ son los multiplicadores de Lagrange para la restricción de los momentos η_n , $n = 0, \dots, M$, con $\eta_0 = 1$. La solución a este problema está dada por (Dacunha-Castelle y Gamboa, 1990)

$$f_{ME}(x) = e^{-\sum_{n=0}^M \lambda_n x^n}, \quad x \in S, \quad (3.23)$$

donde $f_{ME}(x)$ es la densidad de máxima entropía que aproxima a la fdp $f_X(x)$. Es importante destacar que no se conoce una expresión cerrada para los multiplicadores de Lagrange λ_n , deben calcularse numéricamente (Ormoneit y White, 1999).

3.3.2. Cota de Cramér-Rao

Dada una variable aleatoria X , con fdp $f_X(x; \boldsymbol{\theta})$, donde $\boldsymbol{\theta}$ es un vector de parámetros de la distribución, se desea calcular la matriz de información de Fisher de la distribución de máxima entropía $f_{ME}(x; \boldsymbol{\theta})$ que aproxima a $f_X(x; \boldsymbol{\theta})$. A partir de (3.19) y (3.23), los elementos de la matriz de información de Fisher para este problema se pueden escribir como (von Ellenrieder y Muravchik, 2008)

$$\begin{aligned} [\mathbf{J}]_{ij} &= \sum_{m=0}^M \sum_{k=0}^M \frac{\partial \lambda_m}{\partial \theta_i} \frac{\partial \lambda_k}{\partial \theta_j} \mathbb{E} \{ x^{m+k} \} \\ [\mathbf{J}]_{ij} &= \left(\frac{\partial \boldsymbol{\lambda}}{\partial \theta_i} \right)^T \mathbf{M} \left(\frac{\partial \boldsymbol{\lambda}}{\partial \theta_j} \right), \end{aligned} \quad (3.24)$$

donde \mathbf{M} es una matriz Hankel de $(M+1) \times (M+1)$ dada por

$$\mathbf{M} = \begin{bmatrix} 1 & m_1 & \cdots & m_M \\ m_1 & m_2 & \cdots & m_{M+1} \\ \vdots & \vdots & \ddots & \vdots \\ m_M & m_{M+1} & \cdots & m_{2M} \end{bmatrix}, \quad (3.25)$$

siendo m_n los momentos no centrales de la distribución de máxima entropía $f_{ME}(x; \boldsymbol{\theta})$, es decir $m_n = \int_S x^n f_{ME}(x; \boldsymbol{\theta}) dx$. La sensibilidad de los multiplicadores de Lagrange $\boldsymbol{\lambda}$ con el parámetro θ_i puede obtenerse como

$$\frac{\partial \boldsymbol{\lambda}}{\partial \theta_i} = \frac{\partial \boldsymbol{\lambda}}{\partial \mathbf{m}} \frac{\partial \mathbf{m}}{\partial \theta_i} = \left(\frac{\partial \mathbf{m}}{\partial \boldsymbol{\lambda}} \right)^{-1} \frac{\partial \mathbf{m}}{\partial \theta_i}, \quad (3.26)$$

donde $\mathbf{m} = [1 \cdots m_M]^T$ y se puede demostrar que $\frac{\partial \mathbf{m}}{\partial \boldsymbol{\lambda}} = -\mathbf{M}$ (von Ellenrieder y Muravchik, 2008). Entonces, de (3.26) se obtiene que $\frac{\partial \boldsymbol{\lambda}}{\partial \theta_i} = -\mathbf{M}^{-1} \frac{\partial \mathbf{m}}{\partial \theta_i}$. Reemplazando en (3.24) resulta

$$[\mathbf{J}]_{ij} = \left(\frac{\partial \mathbf{m}}{\partial \theta_i} \right)^T \mathbf{M}^{-1} \left(\frac{\partial \mathbf{m}}{\partial \theta_j} \right). \quad (3.27)$$

Los elementos de la matriz \mathbf{M} son los primeros $2M$ momentos de la distribución de máxima entropía $f_{ME}(x)$. Los primeros M son conocidos e iguales a los de la distribución verdadera $f_X(x)$, es decir $m_n = \eta_n$ para $n = 0, \dots, M$. Sin embargo, los momentos de orden mayor,

$n = M + 1, \dots, 2M$, deben calcularse para completar la matriz \mathbf{M} . Para llevar a cabo esto, es necesario obtener los valores de los multiplicadores de Lagrange $\boldsymbol{\lambda}$. Una vez obtenida la distribución de máxima entropía (3.23), los momentos de mayor orden pueden calcularse utilizando técnicas de integración numérica.

Alternativamente, en la sección 3.3.4 se describe una manera de evitar el cálculo de los multiplicadores de Lagrange.

3.3.3. Algoritmo numérico

El objetivo es describir un algoritmo que permita determinar los multiplicadores de Lagrange numéricamente y con estos calcular los elementos de la matriz \mathbf{M} dada por (3.25).

Reemplazando (3.23) en (3.22) se obtiene la siguiente expresión para el funcional

$$J_f(\boldsymbol{\lambda}) = 1 - \sum_{m=0}^M \lambda_m \eta_m - \int_S e^{-\sum_{m=0}^M \lambda_m x^m} dx, \quad (3.28)$$

el cual debe minimizarse respecto de los multiplicadores de Lagrange, $\boldsymbol{\lambda}$.

Debido a que la minimización del funcional (3.28) es un problema convexo (von Ellenrieder y Muravchik, 2008), para hallar su solución se utiliza el método de optimización de Newton (Sun y Yuan, 2006).

Aproximando $J_f(\boldsymbol{\lambda})$ por medio de una cuadrática $q(\boldsymbol{\lambda})$, a través del desarrollo en serie de Taylor de segundo orden alrededor de $\boldsymbol{\lambda}^{(k)}$

$$q(\boldsymbol{\lambda}) = J_f(\boldsymbol{\lambda}^{(k)}) + \nabla J_f(\boldsymbol{\lambda}^{(k)}) (\boldsymbol{\lambda} - \boldsymbol{\lambda}^{(k)}) + \frac{1}{2} (\boldsymbol{\lambda} - \boldsymbol{\lambda}^{(k)})^T \nabla^2 J_f(\boldsymbol{\lambda}^{(k)}) (\boldsymbol{\lambda} - \boldsymbol{\lambda}^{(k)}), \quad (3.29)$$

donde los elementos de $\nabla J_f(\boldsymbol{\lambda})$ y $\nabla^2 J_f(\boldsymbol{\lambda})$ resultan

$$[\nabla J_f(\boldsymbol{\lambda})]_i = \frac{\partial J_f(\boldsymbol{\lambda})}{\partial \lambda_i} = g_i(\boldsymbol{\lambda}) - \eta_i, \quad (3.30)$$

$$[\nabla^2 J_f(\boldsymbol{\lambda})]_{ij} = \frac{\partial^2 J_f(\boldsymbol{\lambda})}{\partial \lambda_j \partial \lambda_i} = \frac{\partial g_i(\boldsymbol{\lambda})}{\partial \lambda_j} = -g_{i+j}(\boldsymbol{\lambda}), \quad (3.31)$$

con

$$g_i(\boldsymbol{\lambda}) = \int_S x^i e^{-\sum_{n=0}^M \lambda_n x^n} dx. \quad (3.32)$$

Luego, el siguiente iterado $\boldsymbol{\lambda}^{(k+1)}$ se obtiene de resolver el sistema $\nabla q(\boldsymbol{\lambda}) = 0$, resultando la ecuación de actualización

$$\boldsymbol{\lambda}^{(k+1)} = \boldsymbol{\lambda}^{(k)} + [\nabla^2 J_f(\boldsymbol{\lambda}^{(k)})]^{-1} \nabla J_f(\boldsymbol{\lambda}^{(k)}). \quad (3.33)$$

El proceso se repite hasta que el paso de actualización sea despreciable. Finalmente el algoritmo calcula los momentos desconocidos, necesarios para completar la matriz \mathbf{M} , que es utilizada para evaluar la matriz de información de Fisher en (3.27).

Cabe destacar que este es un problema mal condicionado y se deben tener algunas precauciones desde el punto de vista numérico. En tal sentido para invertir el Hessiano de la aproximación cuadrática en (3.33) se utiliza preconditionamiento, de la misma manera que von Ellenrieder y Muravchik (2008). Es decir, el sistema $\mathbf{A}\mathbf{x} = \mathbf{b}$ se resuelve como $\mathbf{x} = \mathbf{P}(\mathbf{PAP})^{-1}(\mathbf{Pb})$ donde, en este caso, \mathbf{P} es una matriz diagonal cuyos elementos son $1/\eta_n$, $n = 0, \dots, M$.

3.3.4. Aproximación alternativa

Otra posibilidad es evitar el cálculo de los multiplicadores de Lagrange y aproximar la matriz de información de Fisher con la expresión obtenida para la distribución de máxima entropía, que escrita en forma matricial resulta

$$\mathbf{J} = \left(\frac{\partial \boldsymbol{\eta}}{\partial \boldsymbol{\theta}} \right)^T \mathbf{M}^{-1} \left(\frac{\partial \boldsymbol{\eta}}{\partial \boldsymbol{\theta}} \right) \quad (3.34)$$

pero ahora $[\mathbf{M}]_{ij} = \eta_{i+j-2}$ y dado que solo se conocen M momentos, el tamaño de la matriz \mathbf{M} es de $(M/2 + 1) \times (M/2 + 1)$, para M par (von Ellenrieder y Muravchik, 2008). Cabe destacar que esta matriz solo depende de los momentos conocidos y ninguno de ellos debe ser aproximado. Por otro lado solo se necesita la sensibilidad respecto del parámetro $\boldsymbol{\theta}$ de los primeros $M/2$ momentos. Lógicamente, al no utilizarse la información de los momentos de orden $1 + M/2$ a M , la calidad de la aproximación se resiente.

3.3.5. Resultados

En esta sección se presentan los resultados de evaluar la aproximación de la cota de Cramér-Rao para un proceso ARCH de primer orden, el cual coincide con un proceso GARCH(0,1) y se define como (Engle, 1982)

$$v_r = \sigma_r z_r \quad z_t \sim \mathcal{N}(0, 1) \text{ iid} \quad (3.35)$$

$$\sigma_r^2 = k + \beta_1 v_{r-1}^2, \quad (3.36)$$

A partir de (3.4), entonces $v_r | v_{r-1} \sim \mathcal{N}(0, \sigma_r^2)$, lo que permite demostrar que los momentos impares de v_r son nulos y que sus momentos pares condicionados a v_{r-1} están dados por (Papoulis, 1991)

$$\begin{aligned} \mathbb{E} \{ v_r^{2i} | v_{r-1} \} &= \sigma_r^{2i} (2i - 1)!! \\ \mathbb{E} \{ v_r^{2i} | v_{r-1} \} &= (k + \beta_1 v_{r-1}^2)^i (2i - 1)!!, \end{aligned} \quad (3.37)$$

donde $i = 1, 2, \dots$ y $(2i - 1)!! = \prod_{j=1}^i (2j - 1)$.

Por otro lado, el siguiente teorema asegura la existencia de los momentos pares del proceso ARCH de primer orden (Engle, 1982), que como se vera más adelante no varían con el tiempo.

Teorema 1 *Dado un entero n , entonces existen los momentos hasta el orden $2n$ del proceso ARCH de primer orden, con $k > 0$ y $\beta_1 \geq 0$, si y sólo si se satisface la condición*

$$\beta_1^n (2n - 1)!! < 1. \quad (3.38)$$

En las condiciones del teorema, si se define el vector $\mathbf{w}_r = [v_r^2 \ v_r^4 \ \dots \ v_r^{2n}]^T$, los momentos condicionados se pueden escribir como

$$\mathbb{E} \{ \mathbf{w}_r | v_{r-1} \} = \mathbf{b} + \mathbf{A} \mathbf{w}_{r-1}, \quad (3.39)$$

donde los elementos de la matriz \mathbf{A} y del vector \mathbf{b} son función solo de los coeficientes del proceso y se obtienen de (3.37). Teniendo en cuenta que $\mathbb{E}\{\mathbf{w}_r\} = \mathbb{E}\{\mathbb{E}\{\mathbf{w}_r|v_{r-1}\}\}$ y que los momentos no varían con el tiempo, es decir $\mathbb{E}\{\mathbf{w}_{r-1}\} = \mathbb{E}\{\mathbf{w}_r\}$, resulta la siguiente expresión para los momentos

$$\mathbb{E}\{\mathbf{w}_r\} = (\mathbf{I} - \mathbf{A})^{-1}\mathbf{b}, \quad (3.40)$$

donde \mathbf{I} es la matriz identidad.

Derivando (3.40) respecto de los coeficientes del proceso se obtiene la sensibilidad de los momentos respecto de los parámetros

$$\frac{\partial \mathbb{E}\{\mathbf{w}_r\}}{\partial \theta_i} = (\mathbf{I} - \mathbf{A})^{-1} \left[\frac{\partial \mathbf{A}}{\partial \theta_i} \mathbb{E}\{\mathbf{w}_r\} + \frac{\partial \mathbf{b}}{\partial \theta_i} \right], \quad (3.41)$$

donde $i = 1, 2$ con $\theta_1 = k$ y $\theta_2 = \beta_1$.

A partir de (3.40) y (3.41) se puede aproximar la CRB a través de (3.27) utilizando los algoritmos descritos (Pascual *et al.*, 2013a). A continuación se realiza una comparación de las dos variantes presentadas para su cálculo, a partir de un ejemplo numérico en un caso particular. Los valores de los coeficientes elegidos son $k = 1$ y $\beta_1 = 0.13$, lo que asegura la existencia de momentos del proceso hasta orden 20. Para resolver las integrales numéricamente se utilizó la fórmula de cuadratura de Gauss-Hermite (Abramowitz y Stegun, 1972) con 67 puntos, debido a que el soporte de la densidad del proceso es $(-\infty, \infty)$.

En la figura 3.2a se muestra el histograma del proceso ARCH junto con aproximaciones de máxima entropía de su densidad para diferente número de momentos conocidos y en la figura 3.2b se presenta la diferencia entre el histograma y las aproximaciones de la densidad. Notar que independientemente del número de momentos conocidos, para el número de realizaciones utilizadas para construir el histograma no se distinguen diferencias apreciables entre ambas aproximaciones. Además, teniendo en cuenta que la amplitud del error es pequeña en relación a la amplitud de las densidades, se puede afirmar que la aproximación de la densidad es aceptable.

La cota se evaluó considerando que ambos coeficientes son desconocidos. Se varió el número de momentos conocidos de 4 a 10, utilizando (3.40) y (3.41) se calcularon los momentos y sus derivadas respectivamente y se aproximó la matriz de información de Fisher utilizando los algoritmos descritos en la sección 3.3.3 y en la sección 3.3.4. Finalmente la CRB se obtuvo invirtiendo esta matriz (3.18). Las figuras 3.3a y 3.3b muestran la raíz cuadrada de la CRB de los coeficientes k y β_1 en función del número de momentos conocidos, respectivamente. Se puede observar que ambas aproximaciones tienden al mismo valor de la cota para cada coeficiente. El segundo enfoque converge más rápido que el primero, sin embargo se debe tener presente que este método utiliza el doble de momentos calculados de manera analítica. Por otro lado, la diferencia que se observa entre ambos métodos cuando el número de momentos conocidos es 10 se debe a que el algoritmo numérico calcula de forma precisa hasta el momento de orden 18 del proceso y el error que comete en el de orden 20 es mayor.

Finalmente, cabe destacar que el valor de la cota para los dos coeficientes es elevado, teniendo en cuenta que nos da el valor mínimo que puede tomar la varianza para estimadores

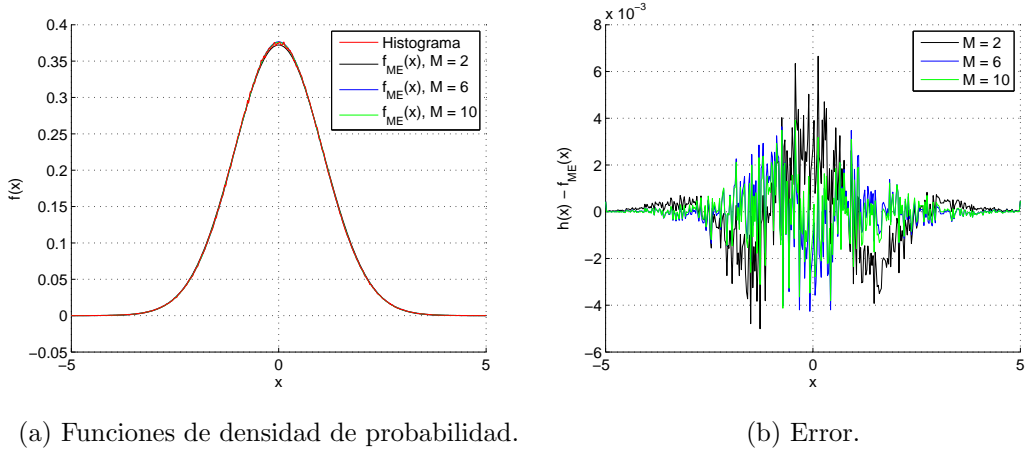


Figura 3.2: Histograma del proceso ARCH de primer orden y su aproximación de la fdp a través de la distribución de máxima entropía.

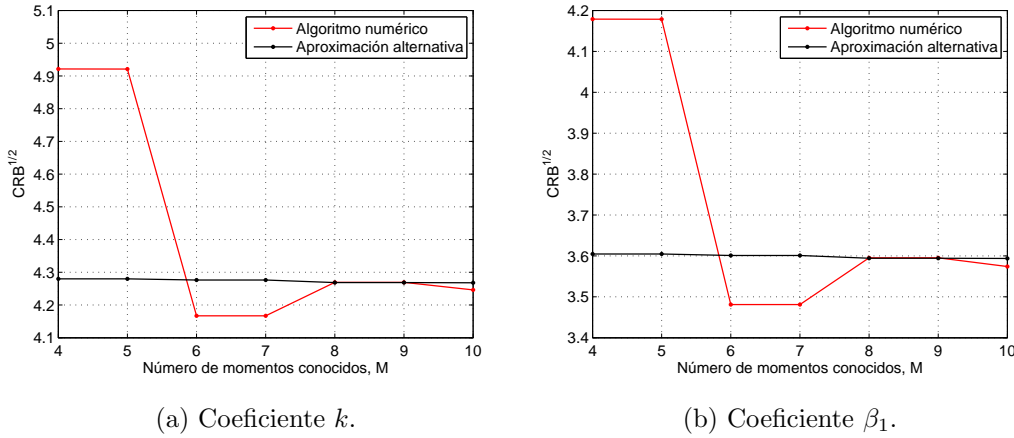


Figura 3.3: CRB para la estimación de los coeficientes del proceso ARCH de primer orden.

insegados. Esto se debe a que la aproximación de la matriz de información de Fisher presentada se desarrolló contemplando una sola muestra temporal. Hacerlo para un conjunto de muestras implica tener que resolver numéricamente integrales multidimensionales de la densidad de probabilidad conjunta de las muestras. Sin embargo, si se cuenta con R muestras independientes de un proceso con la misma distribución, la matriz de información de Fisher es simplemente R veces la matriz de información de Fisher de una de las muestras. Para obtener la cota en estas condiciones basta dividir por \sqrt{R} los resultados presentados.

Para verificar los resultados se puede pensar en un método del tipo de Monte Carlo, evaluando muchas realizaciones del estimador de máxima verosimilitud de los coeficientes utilizando un número grande de muestras y calcular su error cuadrático medio. Como este estimador es asintóticamente eficiente, el error tiende al valor de la CRB. Sin embargo, en la sección 3.2.2 se vio que en realidad no se conoce el MLE y que se utiliza una densidad condicional para estimar los parámetros. Más aún, asintóticamente la matriz de covarianza del estimador no tiende a la inversa de la matriz de información de Fisher buscada. Concretamente, la aproximación (3.17)

| Elemento | De (3.17) | De (3.45) | RMSE |
|--------------------------|-----------|-----------|--------|
| $[\mathbf{F}^{-1}]_{11}$ | 1,8956 | 1,8959 | 1,8968 |
| $[\mathbf{F}^{-1}]_{22}$ | 1,3676 | 1,3675 | 1,4066 |

 Tabla 3.1: Diagonal de las aproximaciones de \mathbf{F}^{-1} .

corresponde a

$$-\mathbb{E} \left\{ \frac{\partial^2 \ln(f(v_r|v_{r-1}))}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^T} \right\} = \frac{1}{2} \mathbb{E} \left\{ \frac{1}{\sigma_r^4} \frac{\partial \sigma_r^2}{\partial \boldsymbol{\theta}} \frac{\partial \sigma_r^2}{\partial \boldsymbol{\theta}^T} \right\}, \quad (3.42)$$

y no a la matriz de información de Fisher

$$\mathbf{J}(\boldsymbol{\theta}) = -\mathbb{E} \left\{ \frac{\partial^2 \ln(f(v_r; \boldsymbol{\theta}))}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^T} \right\}, \quad (3.43)$$

que es la matriz que se aproxima con la solución al problema de momentos. Dichas matrices se vinculan de la siguiente manera

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{\partial^2 \ln(f(v_r, v_{r-1}))}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^T} f(v_r, v_{r-1}) dv_r dv_{r-1} = \mathbb{E} \left\{ \frac{\partial^2 \ln(f(v_r|v_{r-1}))}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^T} \right\} - \mathbf{J}(\boldsymbol{\theta}). \quad (3.44)$$

Por este motivo, para validar los resultados se evaluó (3.42) por un lado, mediante (3.17) luego de realizar la estimación de los parámetros con un conjunto de 200000 muestras del proceso, generadas con los valores de los coeficientes utilizados en esta sección para calcular la cota, y posteriormente se calculó su inversa. Por otro lado, se aproximó (3.42) utilizando la fdp de máxima entropía obtenida cuando el número de momentos conocidos es $M = 10$ e integración numérica, es decir

$$-\mathbb{E} \left\{ \frac{\partial^2 \ln(f(v_r|v_{r-1}))}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^T} \right\} \approx \frac{1}{2} \int_{-\infty}^{\infty} \left[\frac{1}{\sigma_r^4} \frac{\partial \sigma_r^2}{\partial \boldsymbol{\theta}} \frac{\partial \sigma_r^2}{\partial \boldsymbol{\theta}^T} \right] f_{ME}(v_{r-1}) dv_{r-1}, \quad (3.45)$$

y también se calculó su inversa. Por último, se evaluaron 1000 realizaciones del QMLE de los coeficientes y se calculó la raíz cuadrada del error cuadrático medio (RMSE por sus siglas en inglés). En la tabla 3.1 se presentan los valores de los elementos de la diagonal de las matrices aproximadas y de los valores de la raíz cuadrada de los errores cuadráticos medios. Se puede observar que los resultados obtenidos por los tres caminos son aproximadamente iguales.

Estrictamente este último análisis permite concluir que la fdp de máxima entropía es una buena aproximación de la fdp del proceso y a partir de ello se puede inferir que también lo es la aproximación de la CRB. Por otro lado, se observa que el error de estimación obtenido es menor que lo que predice la CRB (aunque son del mismo orden). Esto se debe a que se está utilizando la densidad condicional para evaluar el estimador, que cuenta con la información de muestras previas y por lo tanto menor incertidumbre. Nótese que \mathbf{F} es equivalente a la matriz de información de Fisher, pero con la densidad de probabilidad condicional en lugar de la fdp del proceso.

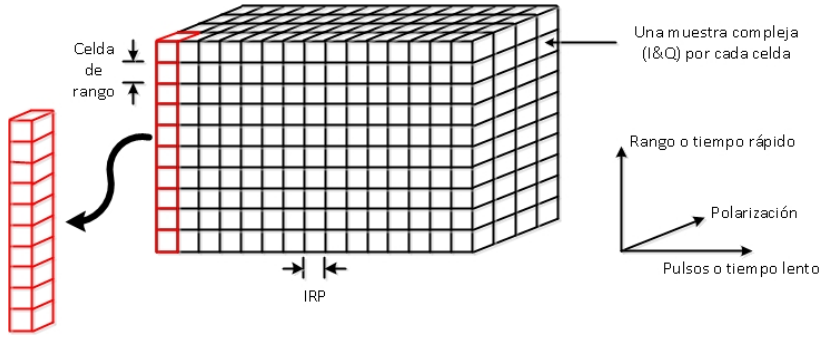


Figura 3.4: Modelado del clutter a través de procesos GARCH.

3.4. Modelado del clutter

El objetivo de esta sección es ajustar un proceso GARCH a datos reales de radar. Para esto se definen las muestras del proceso como las respuestas correspondientes a distintas celdas de rango para cada pulso transmitido, siendo la respuesta a cada uno de estos pulsos una realización diferente del proceso. De acuerdo con la convención adoptada para el almacenamiento de los datos que se describió en la sección 2.2.1, para una matriz correspondiente a una dada polarización, cada columna corresponde a una realización del proceso GARCH que modela los datos y las muestras de cada columna corresponden a muestras del proceso (figura 3.4).

Antes de continuar cabe destacar que los datos con los que se cuenta se componen de valores complejos, correspondientes a las componentes en fase y cuadratura de la señal recibida por el radar. Esto implica que el modelo que deseamos ajustar a los datos también debe ser complejo. La extensión compleja del modelo GARCH se define como (Mousazadeh y Cohen, 2010)

$$v_r = \sigma_r z_r \quad z_r \sim \mathcal{CN}(0, 1) \text{ iid}, \quad (3.46)$$

$$\sigma_r^2 = k + \sum_{j=1}^p \alpha_j \sigma_{r-j}^2 + \sum_{j=1}^q \beta_j |v_{r-j}|^2, \quad (3.47)$$

donde \mathcal{CN} denota distribución normal circularmente compleja. Esto quiere decir que ahora $v_r | \psi_r \sim \mathcal{CN}(0, \sigma_r^2)$ y por ende las expresiones de las densidades varían ligeramente respecto de las presentadas en la sección 3.2. En particular, el funcional que debe minimizarse para obtener el QMLE queda dado por

$$\mathcal{L}_2(\boldsymbol{\theta}) = \sum_{r=1}^R \left[\ln(\sigma_r^2) + \frac{|v_r|^2}{\sigma_r^2} \right]. \quad (3.48)$$

En el apéndice A se deduce la expresión del gradiente de $\mathcal{L}_2(\boldsymbol{\theta})$, que será necesaria para resolver el problema de optimización para hallar los estimados del vector de parámetros $\boldsymbol{\theta}$.

3.4.1. Método alternativo para la estimación de los coeficientes

El número de muestras requerido por el QMLE descrito en la sección 3.2.2 para lograr buenos estimados de los coeficientes en ocasiones suele ser elevado en relación con el que se

dispone en una situación práctica. Por este motivo en esta sección se propone una modificación al método de estimación de los coeficientes, la cual consiste en el uso de varias realizaciones del proceso de corta duración en lugar de una larga (Pascual *et al.*, 2011).

Denotando como $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_N$ a N realizaciones independientes de un proceso GARCH, entonces la función de verosimilitud condicional será el producto de las N densidades marginales condicionales de las diferentes realizaciones, es decir

$$f(\mathbf{v}_1, \dots, \mathbf{v}_N | v_{01}, \dots, v_{0N}, \dots, v_{1-n,1}, \dots, v_{1-n,N}) = \prod_{t=1}^N f(\mathbf{v}_t | v_{0t}, \dots, v_{1-n,t}), \quad (3.49)$$

donde v_{it} es el elemento i -ésimo de \mathbf{v}_t y cada fdp $f(\mathbf{v}_t | v_{0t}, \dots, v_{1-n,t})$ es de la forma de (3.7)

$$f(\mathbf{v}_t | v_{0t}, \dots, v_{1-n,t}) = \prod_{r=1}^R f(v_{rt} | v_{r-1,t}, \dots, v_{r-n,t}). \quad (3.50)$$

Para un proceso GARCH complejo $f(v_{rt} | v_{r-1,t}, \dots, v_{r-n,t})$ está dada por

$$f(v_{rt} | v_{r-1,t}, \dots, v_{r-n,t}) = \frac{1}{\pi \sigma_{rt}^2} e^{-\frac{|v_{rt}|^2}{\sigma_{rt}^2}}. \quad (3.51)$$

Reemplazando (3.50) y (3.51) en (3.49) y tomando logaritmo natural, la nueva función de verosimilitud logarítmica resulta

$$\ln(f(\mathbf{v}_1, \dots, \mathbf{v}_N | v_{01}, \dots, v_{0N}, \dots, v_{1-n,1}, \dots, v_{1-n,N})) = -R \ln(\pi) - \sum_{t=1}^N \sum_{r=1}^R \left[\ln(\sigma_{rt}^2) + \frac{|v_{rt}|^2}{\sigma_{rt}^2} \right]. \quad (3.52)$$

Finalmente, el estimador alternativo se obtiene como el valor de $\boldsymbol{\theta}$ que maximiza la fdp condicional (3.50) ó, equivalentemente, que minimiza

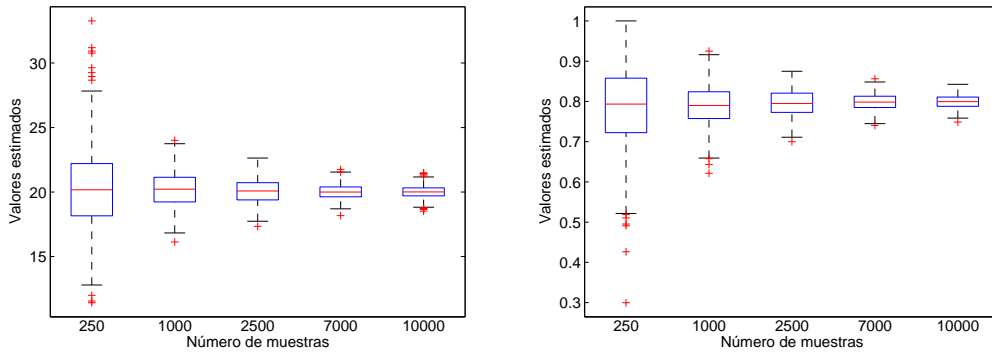
$$\mathcal{L}_3(\boldsymbol{\theta}) = \sum_{t=1}^N \sum_{r=1}^R \left[\ln(\sigma_{rt}^2) + \frac{|v_{rt}|^2}{\sigma_{rt}^2} \right], \quad (3.53)$$

sujeto a las restricciones mencionadas.

En el apéndice A también se da una expresión para el gradiente de $\mathcal{L}_3(\boldsymbol{\theta})$, cuya deducción es análoga a la de $\mathcal{L}_2(\boldsymbol{\theta})$.

3.4.2. Error de estimación

A fin de cuantificar el número de muestras necesarias para llevar a cabo la estimación al momento de utilizar las mediciones, en esta sección se realiza un análisis mediante simulaciones numéricas del error de estimación en función del número de muestras del vector de observaciones. Las simulaciones consisten en la generación de realizaciones independientes de un proceso GARCH complejo de distinto largo, R , y la estimación de los coeficientes, resolviendo el problema de optimización (3.11) con el funcional $\mathcal{L}_2(\boldsymbol{\theta})$ dado por (3.48) en lugar de $\mathcal{L}_1(\boldsymbol{\theta})$, es decir hallando el QMLE de los coeficientes por medio del método original. La minimización del funcional se resolvió utilizando el algoritmo de conjuntos activos (*Active Set Algorithm*)



(a) Coeficiente k .

(b) Coeficiente β_1 .

Figura 3.5: Estimados de los coeficientes de un proceso GARCH(0,1) en función del número de muestras.

de MATLAB®, implementado a partir del método SQP (*Sequential Quadratic Programming*), (Fletcher, 1987). El procedimiento se repitió 500 veces en cada caso, es decir se obtuvieron 500 valores del estimador para cada largo del vector de observaciones. Solo se incluyen los resultados correspondientes al proceso GARCH(0,1), sin embargo se obtienen resultados similares cuando el análisis se repite para diferentes órdenes del proceso (p, q) , en Pascual *et al.* (2011) se puede encontrar un análisis análogo para un proceso GARCH(1,1). Los valores de los coeficientes con los que se generaron los datos son $k^v = 20$ y $\beta_1^v = 0.8$. Cabe destacar que los resultados no fueron contrastados con la aproximación de la CRB presentada, debido a que el valor de β_1 elegido solo permite calcular el momento de segundo orden del proceso a causa de la restricción de existencia que pesa sobre los momentos.

En la figura 3.5 se presentan los resultados obtenidos. En ella se indica, para cada coeficiente y para cada valor de R , el rango de valores que tomaron los estimados y su valor medio. A su vez en la tabla 3.2 se muestran los valores de la media, de la desviación estándar y de la raíz cuadrada del error cuadrático medio de los estimados de los coeficientes k y β_1 . Los resultados obtenidos verifican el comportamiento asintótico mencionado en la sección 3.2.3. No obstante es importante destacar que revelan la necesidad de utilizar, para lograr un error de estimación razonable, un número de muestras espaciales que puede resultar mayor al disponible en situaciones realistas. En particular los conjuntos con datos de los que se disponen constan de realizaciones con 61 observaciones en promedio.

Con motivo de establecer una comparación entre el método original de estimación y el método alternativo propuesto en la sección 3.4.1 se llevaron a cabo simulaciones numéricas para calcular media, desviación estándar y error cuadrático medio de los estimados. Análogamente al análisis anterior, las simulaciones consisten en la generación de $N = 40$ realizaciones independientes de $R = 250$ muestras de un proceso GARCH(0, 1) complejo y la estimación de sus coeficientes resolviendo el problema de optimización (3.11) con $\mathcal{L}_3(\theta)$ dada por (3.53). Los valores de los coeficientes son nuevamente $k^v = 20$ y $\beta_1^v = 0.8$ y también el experimento se repitió 500 veces.

| R | Coficiente | Media | Desv. estándar | RMSE |
|-------|------------|---------|----------------|--------|
| 250 | k | 20.2676 | 3.2321 | 3.2400 |
| | β_1 | 0.7874 | 0.1045 | 0.1051 |
| 1000 | k | 20.2198 | 1.3927 | 1.4086 |
| | β_1 | 0.7904 | 0.0485 | 0.0494 |
| 2500 | k | 20.0780 | 0.9642 | 0.9664 |
| | β_1 | 0.7959 | 0.0339 | 0.0341 |
| 7000 | k | 20.0364 | 0.5613 | 0.5619 |
| | β_1 | 0.7986 | 0.0198 | 0.0198 |
| 10000 | k | 20.0121 | 0.4891 | 0.4887 |
| | β_1 | 0.7995 | 0.0159 | 0.0159 |

Tabla 3.2: Estimados de los coeficientes de un proceso GARCH(0,1).

| N | R | Coficiente | Media | Desv. estándar | RMSE |
|-----|-------|------------|---------|----------------|--------|
| 40 | 250 | k | 20.0211 | 0.4738 | 0.4738 |
| | | β_1 | 0.7990 | 0.0153 | 0.0153 |
| 1 | 10000 | k | 20.0121 | 0.4891 | 0.4887 |
| | | β_1 | 0.7995 | 0.0159 | 0.0159 |

Tabla 3.3: Comparación de los métodos de estimación.

Es importante destacar que para otros órdenes del proceso se obtienen resultados similares, en Pascual *et al.* (2011) también se estudia este método para un proceso GARCH(1,1). En la tabla 3.3 se resumen los resultados. En esta tabla también se incluyen los mismos parámetros calculados por el método de estimación original con una realización de 10000 muestras. Notar que los estimados obtenidos por este método presentan un comportamiento similar al que se obtiene utilizando el método original cuando el número de muestras total es el mismo en ambos casos. Además de mostrar que son insesgados, los órdenes de magnitud del error son iguales. Esto permite determinar el número de realizaciones independientes, de largo fijo, necesarias para un dado error de estimación. Por lo tanto, nos basaremos en estos resultados para utilizar esta alternativa para estimar los coeficientes en el ajuste del proceso GARCH a los datos reales de clutter. Sin embargo no se puede hablar de consistencia para estos estimadores debido a que por el momento no se cuenta con ningún resultado teórico que la fundamente.

3.4.3. Ajuste a los datos reales

Se presentan a continuación los resultados del ajuste de un modelo del tipo GARCH a datos reales. Por ajuste se hace referencia al proceso de elección de los órdenes del proceso y la estimación de sus coeficientes, utilizando como observaciones mediciones reales de clutter marítimo. Los datos utilizados corresponden a la polarización VV del conjunto de datos *stare0*

descrito en la sección 2.2.2. Las realizaciones del proceso son tomadas en rango, es decir que se cuenta con una realización por cada pulso transmitido. Sin embargo, debido a que se busca ajustar el modelo GARCH solo a muestras correspondientes a clutter, se considerará que las realizaciones se extienden hasta la celda de rango 45, de modo de excluir las celdas que contienen al objetivo. Por otra parte, en el proceso de estimación solo se utilizarán las primeras 1024 realizaciones (o pulsos). Los pulsos restantes del conjunto de datos serán utilizados para evaluar el desempeño del esquema de detección.

En principio, no hay motivos para considerar que las realizaciones son independientes, hipótesis necesaria para realizar la estimación a partir del método alternativo propuesto en la sección 3.4.1. Sin embargo, se puede pensar que transcurrido un cierto tiempo el oleaje va cambiando haciendo válida esta consideración. Con el objetivo de tener una idea cuantitativa, un parámetro a considerar es la estimación de la correlación estadística de los datos presentada en la sección 2.2.2. Los resultados muestran que luego de 30 ms, o equivalentemente luego de 6 pulsos, la función de correlación es aproximadamente cero, es decir que las muestras de dos pulsos separados por un tiempo mayor al mencionado no están correlacionadas. Si bien esto no necesariamente implica que exista independencia, si fija un límite inferior al tiempo entre pulsos para que la misma pueda existir. Se supondrá entonces que las realizaciones separadas por 6 pulsos o más serán independientes.

En estas condiciones la estimación de los coeficientes se realizó utilizando el método alternativo. Se tomaron $N = 156$ realizaciones separadas cada 6 pulsos, las cuales como se mencionó constaron de $R = 45$ muestras. A priori no se conocen los órdenes del proceso que se busca ajustar, por este motivo se realizó el ajuste para diferentes órdenes comenzando por $p = 0$ y $q = 1$. Inicialmente se contemplo la utilización de algún criterio de selección de modelos como el criterio de información Bayesiano (BIC por sus siglas en inglés) para determinar los órdenes. Sin embargo, esta idea se desestimó por dos motivos: en primer lugar, porque la deducción de la fórmula del BIC hace uso del hecho que el gradiente de la función de verosimilitud es cero para $\theta = \hat{\theta}$ (Konishi y Kitagawa, 2008), lo cual, en este caso, no es necesariamente cierto dado que $\hat{\theta}$ es un máximo restringido de esta función; el segundo motivo es que al realizar el ajuste para los órdenes superiores los estimadores de los coeficientes correspondientes a k y β_1 toman los mismos valores que para el proceso GARCH(0,1) y los estimados de los nuevos coeficientes, en general son cero o toman valores estadísticamente despreciables. Este comportamiento fue común no solo para los estimadores calculados con este conjunto de datos, sino también en ajustes realizados con otras mediciones disponibles (Pascual *et al.*, 2011). Los valores de los estimados \hat{k} y $\hat{\beta}_1$ resultantes del ajuste del proceso GARCH(0,1) son 5.3675×10^{-3} y 0.7463, respectivamente (Pascual *et al.*, 2013b).

Estos procesos son de tipo impulsivo. Un valor de \hat{k} pequeño indica que la varianza condicional del piso de ruido será pequeña y el valor de $\hat{\beta}_1$ es el responsable de la tasa y amplitud con la que se producen los impulsos, cuanto mayor sea β_1 mayor será la probabilidad de que ocurran. Al ser un proceso con órdenes (0,1) la varianza condicional solo depende del último valor de las observaciones del proceso, y los impulsos serán de corta duración a menos que

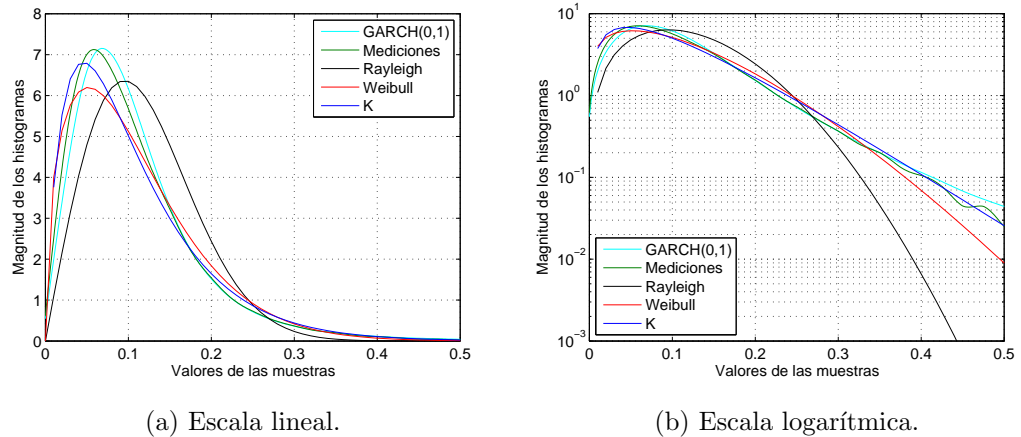


Figura 3.6: Histogramas de la magnitud del clutter del conjunto de datos *stare0* para la polarización VV, de la magnitud del proceso GARCH(0,1) y funciones densidad de probabilidad resultantes del ajuste.

ocurra una ráfaga de innovaciones con valores considerables.

En la figura 3.6 se muestran el histograma de la magnitud del proceso GARCH(0,1) con los valores de los coeficientes resultantes del ajuste, junto con el histograma de la magnitud del conjunto de datos *stare0* para la polarización VV y las fdp de las distribuciones Rayleigh, Weibull y K con los parámetros obtenidos sección 2.3.4 para el mismo conjunto de datos. Por un lado se puede destacar el parecido entre los histogramas de los datos reales y el del proceso GARCH. Sin embargo, la distribución K sigue siendo la mejor aproxima la cola del histograma de los datos. Además, es importante mencionar que los datos generados con la densidad normal (o magnitud Rayleigh) tienen una probabilidad muy baja de tomar valores por encima de tres veces el valor de su desviación estándar, cuando en realidad las mediciones de clutter tienen mayor probabilidad de tomar estos valores, tal como se discutió en el párrafo anterior. Pensando en el diseño de un sistema de detección es de esperar que el desempeño del mismo mejore con la representación de los datos utilizando un proceso GARCH o K, teniendo en cuenta que si se utiliza una distribución normal la probabilidad de tener una falsa detección sería mayor dada la impulsividad del clutter no modelada por los procesos Gaussianos. En el capítulo 4 se verá que la deducción de un esquema de detección a partir del modelado de los datos con distribución K presenta dificultades en el tratamiento matemático llevando a soluciones subóptimas, que poseen un peor desempeño aún frente a un detector obtenido modelando los datos con una distribución Gaussiana.

3.5. Esquema de detección

El problema de detección se formula como un test de hipótesis en la forma descrita en la sección 2.4. En este caso particular debido a que el modelo de los datos es en rango, la decisión se lleva a cabo muestra a muestra. Si se denota como y_r a la muestra compleja a la salida del

demodulador en cuadratura, el procedimiento de detección está dado por la decisión entre dos hipótesis H_0 y H_1 una vez que y_r fue recibida para la celda de rango bajo prueba

$$\begin{aligned} H_0 : y_r &= v_r \\ H_1 : y_r &= x_r + v_r. \end{aligned} \tag{3.54}$$

Bajo la hipótesis nula H_0 se supone que los datos consisten solo de clutter v_r el cual es modelado como un proceso GARCH(p, q). Se considera que el ruido electrónico es despreciable o es parte del modelo de clutter. Bajo la hipótesis H_1 se supone que la medición es el resultado de la combinación de clutter v_r y reflexiones en un objetivo x_r . Debido a que y_r es una muestra tomada a la salida del filtro adaptado se puede pensar que x_r es una constante proporcional a la energía de la señal. Sin embargo, un conocimiento perfecto de la señal en el receptor implica conocer el rango del objetivo de manera muy precisa. Debido a que en general es difícil de alcanzar esta precisión para frecuencias en el orden de las microondas, un modelo más realista es considerar que es conocido salvo por un factor que da cuenta del error de sincronismo $e^{j\theta}$, donde θ es una variable aleatoria uniformemente distribuida en $(0, 2\pi]$ (Richards, 2005). En resumen, el modelo de señal resulta $x_t = \mathcal{E}e^{j\theta}$, donde \mathcal{E} es una constante determinística. Se supone que la distribución del clutter es la misma para ambas hipótesis.

Existen diferentes cuestiones que llevan a que el problema se aparte del test de hipótesis clásico descrito en 2.4.1. La primera es el hecho de que no se cuenta con una expresión explícita para la fdp de un proceso GARCH. Para salvar esta dificultad, se puede considerar en su lugar la fdp condicional dada las observaciones del proceso hasta la celda de rango bajo prueba y_{r-1}, \dots, y_1 de manera análoga a lo que se hace en el proceso de estimación de sus coeficientes. Esta es una consideración razonable debido a que, excepto para la primera celda con el modelo GARCH(0,1), las muestras previas a la celda de rango bajo prueba son conocidas. Si se define $\boldsymbol{\psi}_r = [y_{r-1} \cdots y_1]^T$, entonces se reemplaza (2.16) por la *relación de verosimilitud condicional* $\Lambda(y_r|\boldsymbol{\psi}_r)$, dada por

$$\Lambda(y_r|\boldsymbol{\psi}_r) = \frac{f_y(y_r|\boldsymbol{\psi}_r; H_1)}{f_y(y_r|\boldsymbol{\psi}_r; H_0)}. \tag{3.55}$$

Por otro lado, no se tiene un conocimiento perfecto de las densidades de probabilidad condicionales. Respecto del modelo de clutter, no se conocen los valores de los coeficientes del proceso GARCH y tampoco se conoce la fase del modelo de señal. Esto lleva a un test de hipótesis compuestas.

Siguiendo el enfoque del test de relación de verosimilitud generalizado (GLRT) se deben maximizar las fdp's para cada hipótesis respecto de los parámetros desconocidos. Sin embargo, la aplicación del GLRT para los coeficientes del proceso GARCH lleva a un problema de maximización no-lineal multi-dimensional para el cual no existe una expresión cerrada de su solución. Por este motivo se propone un planteo alternativo que se basa en considerar que los valores de los coeficientes son conocidos. Luego, en la implementación del algoritmo estos serán estimados como se describió en la sección 3.4.1 utilizando un conjunto de datos secundario que solo contengan muestras correspondientes a clutter. Posteriormente se analizará por medio de

simulaciones numéricas los alcances de esta consideración. Por otro lado, la fase de la señal se supuso aleatoria, por lo que se adopta el enfoque Bayesiano del test de hipótesis compuestas descrito en la sección 2.4.3.

Si $f_\theta(\theta)$ representa la fdp a priori de θ , entonces las fdp's condicionales para cada hipótesis están dadas por

$$f_y(y_r|\boldsymbol{\psi}_r; H_i) = \int f_{y\theta}(y_r, \theta|\boldsymbol{\psi}_r; H_i)d\theta = \int f_{y|\theta}(y_r|\boldsymbol{\psi}_r, \theta; H_i)f_\theta(\theta)d\theta \quad i = 0, 1. \quad (3.56)$$

A partir de (3.4) y (3.54) resulta

$$f_{y|\theta}(y_r|\boldsymbol{\psi}_r, \theta; H_0) = \frac{1}{\pi\sigma_r^2}e^{-|y_r|^2/\sigma_r^2} \quad (3.57)$$

$$f_{y|\theta}(y_r|\boldsymbol{\psi}_r, \theta; H_1) = \frac{1}{\pi\sigma_r^2}e^{-|y_r - \mathcal{E}e^{j\theta}|^2/\sigma_r^2}, \quad (3.58)$$

De (3.57) se observa que $f_y(y_r|\boldsymbol{\psi}_r; H_0) = f_{y|\theta}(y_r|\boldsymbol{\psi}_r, \theta; H_0)$ debido a que no depende de θ . Además, análogamente al caso Gaussiano (Richards, 2005), luego de integrar (3.56) respecto de θ para la hipótesis H_1 se obtiene

$$f_y(y_r|\boldsymbol{\psi}_r; H_1) = \frac{1}{\pi\sigma_r^2}e^{-(|y_r|^2 - \mathcal{E}^2)/\sigma_r^2}I_0\left(\frac{2\mathcal{E}|y_r|}{\sigma_r^2}\right), \quad (3.59)$$

donde $I_0(\cdot)$ es la función de Bessel modificada de primer tipo y de orden cero (Abramowitz y Stegun, 1972). Entonces, reemplazando $f_y(y_r|\boldsymbol{\psi}_r; H_1)$ y $f_y(y_r|\boldsymbol{\psi}_r; H_0)$ en (3.55) se obtiene la siguiente regla de decisión

$$\ln\left[I_0\left(\frac{2\mathcal{E}|y_r|}{\sigma_r^2}\right)\right] \underset{H_0}{\overset{H_1}{\geq}} \ln(\eta) + \frac{\mathcal{E}^2}{\sigma_r^2} = \lambda. \quad (3.60)$$

donde η es el umbral con el que se compara la relación de verosimilitud para tomar la decisión y λ es el umbral resultante de operar sobre la expresión de $\Lambda(y_r|\boldsymbol{\psi}_r)$.

Debido a que $\ln[I_0(\cdot)]$ es monótonamente creciente, se obtienen los mismos resultados en la detección comparando simplemente su argumento con un umbral modificado. Luego, la regla de decisión se convierte a

$$|y_r| \underset{H_0}{\overset{H_1}{\geq}} \lambda'. \quad (3.61)$$

Con este procedimiento se obtuvo un criterio de decisión, pero resta hallar el valor del umbral, η o λ' , que asegure que la probabilidad de falsa alarma no exceda un dado valor o nivel de significación del test.

La manera clásica de determinar el valor del umbral, consiste en obtener una expresión para la probabilidad de falsa alarma, P_{FA} , a partir de (2.17), que luego se invierte para hallar el umbral en términos de P_{FA} . Esto implica que debe conocerse la distribución de $u_r = |y_r|$ bajo la hipótesis H_0 . Nuevamente, debido a que no se cuenta con una expresión explícita de la fdp de y_r o de v_r , se recurre a la probabilidad de falsa alarma condicional dadas y_{r-1}, \dots, y_1 , $P_{FA}|\boldsymbol{\psi}_r$.

Bajo la hipótesis H_0 la fdp condicional de u_r dado $\boldsymbol{\psi}_r$, $f_u(u_r|\boldsymbol{\psi}_r; H_0)$, es Rayleigh con parámetro $\sigma_r^2/2$, entonces $P_{FA}|\boldsymbol{\psi}_r$ resulta

$$P_{FA}|\boldsymbol{\psi}_r = \int_{\lambda'}^{\infty} f_u(u_r|\boldsymbol{\psi}_r; H_0)du_r = e^{-\lambda'^2/\sigma_r^2}. \quad (3.62)$$

Finalmente, de (3.62) se obtiene la expresión para el umbral λ'

$$\lambda' = \sqrt{-\sigma_r^2 \ln(P_{FA}|\psi_r)}. \quad (3.63)$$

Una vez que se fija $P_{FA}|\psi_r$ a partir de (3.63) se obtiene el valor para el umbral para cada celda de rango. Con este procedimiento se llegó a un umbral adaptivo λ' que depende de las muestras de clutter previas para cada celda de rango bajo prueba a través de σ_r . Esto puede ser un problema cuando se debe decidir en una celda de rango luego de una detección, debido a que la señal recibida por el radar corresponde a señal más clutter y no es posible separar la contribución del clutter. En este trabajo esto no significa un problema, debido a que se consideran objetivos aislados. Sin embargo, el algoritmo de decisión puede modificarse para objetivos extendidos cambiando la manera en que se fija el umbral luego de una presunta detección. Una alternativa sería reiniciar el algoritmo de detección cada vez que el estadístico supera el umbral utilizando la varianza no condicional o bien fijando el umbral luego de una presunta detección en base a las últimas muestras para las cuales no se produjo detección.

La probabilidad de detección, P_D no puede evaluarse analíticamente por los motivos expuestos para la P_{FA} y se calculará por medio de simulaciones numéricas en la sección 3.5.1.

En la deducción del algoritmo de decisión se supuso que los coeficientes del proceso son conocidos, sin embargo esto no es cierto en una situación real. Por este motivo, en la práctica el *detector GARCH* está dado por

$$|y_r| \underset{H_0}{\overset{H_1}{\gtrless}} \sqrt{-\hat{\sigma}_r^2 \ln(P_{FA}|\psi_r)}. \quad (3.64)$$

donde

$$\hat{\sigma}_r^2 = \hat{k} + \sum_{j=1}^p \hat{\alpha}_j \hat{\sigma}_{r-j}^2 + \sum_{j=1}^q \hat{\beta}_j |v_{r-j}|^2 \quad (3.65)$$

siendo \hat{k} , $\hat{\alpha}_j$ y $\hat{\beta}_j$ los estimados de los coeficientes del proceso que se obtienen por medio del método descrito en la sección 3.4.1. En la sección 3.5.2 se realiza un análisis de sensibilidad, donde se estudia como afectan los errores en los coeficientes en el desempeño del detector.

3.5.1. Probabilidad de detección

Debido a que no es posible hallar una expresión analítica para la probabilidad de detección del detector GARCH, se realizaron simulaciones Monte Carlo para estimar P_D para diferentes valores de relación señal a clutter (SCR por sus siglas en inglés) y un dado valor de P_{FA} . Al mismo tiempo se evaluó la P_{FA} de modo de verificar la expresión (3.62). Estas simulaciones consistieron en generar realizaciones de un proceso GARCH(0,1) de R muestras, donde cada muestra corresponde a una celda de rango. Para los coeficientes se fijaron los mismos valores que los utilizados en la sección 3.4.2 lo que da una varianza no condicional igual a 100. En una celda de rango, elegida al azar, se sumó una muestra de señal modelada como se describió en la sección 3.5. El umbral dado por (3.63) fue evaluado para el valor de $P_{FA}|\psi_t = 10^{-3}$. Notar que para estos órdenes del proceso el umbral solo depende de la muestra previa a la celda

de rango bajo prueba, es decir $\psi_r = v_{r-1}$. Por este motivo se utiliza y_{r-1} en lugar de v_{r-1} , independientemente de si el estadístico excede o no el umbral de detección en la celda de rango previa. Para calcular el primer valor del umbral se utiliza la desviación estándar del proceso obtenida a partir de (3.6). Luego, las muestras del estadístico fueron comparadas con su umbral correspondiente y la P_D y la P_{FA} fueron estimadas por medio de sus frecuencias relativas.

El número de realizaciones independientes del proceso se eligió lo suficientemente grande como para asegurar que las probabilidades de detección y de falsa alarma sean estimadas de manera precisa siguiendo la regla deducida por Kay (1998) y que se describe a continuación. Sea P la probabilidad que se desea estimar y \hat{P} su estimado, luego para asegurar que su error relativo

$$e = \frac{|\hat{P} - P|}{P} \quad (3.66)$$

no sea mayor que ϵ el $100(1 - \delta)\%$ del tiempo, se debe elegir el número de ensayos independientes, K , de modo de satisfacer

$$K \geq \frac{[Q^{-1}(\delta/2)]^2(1 - P)}{\epsilon^2 P} \quad (3.67)$$

donde $Q^{-1}(\cdot)$ es la función Q inversa. En estas condiciones, para estimar $P_{FA} = 10^{-3}$ con un error relativo menor que 0.01 (1%) el 95% del tiempo, a partir de (3.67), K debe ser mayor que 38×10^6 .

Para estimar los valores de la probabilidad de detección se utilizaron 1×10^6 realizaciones, cada una de largo $R = 200$ muestras. En consecuencia el número de ensayos para estimar P_D es 1×10^6 , sin embargo esto no es un problema y esta dentro de lo predicho por (3.67) debido a que los valores de P_D son al menos dos órdenes de magnitud mayores que el de P_{FA} . Las simulaciones se repitieron con el propósito de estimar la P_D para diferentes valores de SCR, manteniendo la potencia de clutter constante y variando el valor de la energía de la señal, \mathcal{E} .

Por su parte, el número de realizaciones consideradas para estimar P_{FA} fue de 76×10^6 lo que da como resultado un número de ensayos igual a 152×10^8 , es decir tres órdenes de magnitud mayor que lo predicho por (3.67), lo que puede considerarse aceptable aún cuando en cada realización las muestras no satisfagan la hipótesis de independencia.

A modo de comparación, también se consideró el detector resultante de suponer que los datos poseen distribución normal (o equivalentemente su intensidad posee distribución Rayleigh, ver sección 2.3.1), al que se denominará *detector Gaussiano*. Considerando que el clutter posee distribución normal con media nula y varianza σ^2 y el modelo de señal descrito en la sección 3.5, la regla de decisión del detector Gaussiano resulta igual a (3.61), pero con el umbral dado por

$$\lambda'_g = \sqrt{-\sigma^2 \ln(P_{FA})}. \quad (3.68)$$

Notar que en este caso el estadístico coincide con el del detector GARCH, pero el umbral es constante. Su deducción es análoga a la del detector GARCH, excepto que en este caso se utilizan las fdp's del modelo de los datos en lugar de las fdp's condicionales (Richards, 2005).

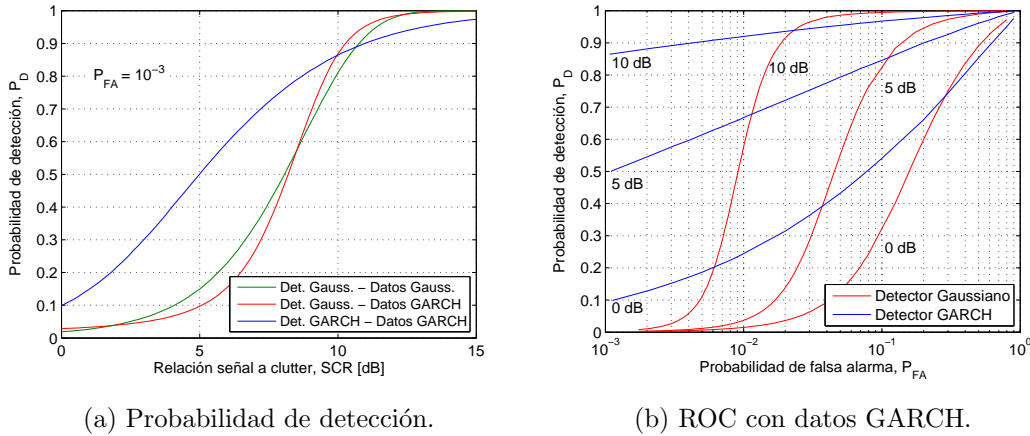


Figura 3.7: Desempeño del detector GARCH y comparación con el detector Gaussiano, utilizando datos sintéticos.

| Detector - Tipo de dato | P_{FA} |
|---------------------------------------|------------------------|
| Detector GARCH - Datos GARCH | 1.10×10^{-3} |
| Detector Gaussiano - Datos GARCH | 16.70×10^{-3} |
| Detector Gaussiano - Datos Gaussianos | 0.99×10^{-3} |

Tabla 3.4: Probabilidad de falsa alarma de los detectores GARCH y Gaussiano.

Se repitieron las simulaciones para determinar la P_D y la P_{FA} con el detector Gaussiano. Sin embargo, para este detector las simulaciones se realizaron de dos maneras diferentes. En primer lugar se utilizó el detector Gaussiano con clutter generado a partir de una distribución Gaussiana, y en segundo término, se utilizó con clutter generado con un proceso GARCH(0,1). En ambos casos la varianza se fijó igual a 100.

En la figura 3.7a se muestran los valores estimados de la P_D por medio de las simulaciones numéricas para ambos detectores. Se observa que la P_D del detector GARCH es mayor que la del detector Gaussiano para valores bajos y medios de SCR (menores a 10 dB). En la tabla 3.4 se muestran los valores estimados de la P_{FA} , donde se puede ver que para clutter GARCH el detector Gaussiano presenta una P_{FA} mayor a la especificada. Es importante destacar que debido a que los procesos Gaussianos son un caso particular de los procesos GARCH (GARCH(0,0)), entonces el detector GARCH con datos Gaussianos presentará el mismo desempeño que el detector Gaussiano. En efecto, se puede observar que si la varianza es constante en el tiempo, el umbral del detector GARCH (3.63) es igual al umbral del detector Gaussiano (3.68).

De manera similar, es posible construir la característica operativa del receptor (curva ROC, por sus siglas en inglés) utilizando datos GARCH para ambos detectores. Cada par (P_D, P_{FA}) corresponde a un punto de la curva y variando el umbral se obtienen los diferentes puntos de la curva. En la figura 3.7b se muestran estas curvas para los dos detectores y para tres valores de SCR. Se puede observar la degradación en el desempeño que sufre el detector Gaussiano

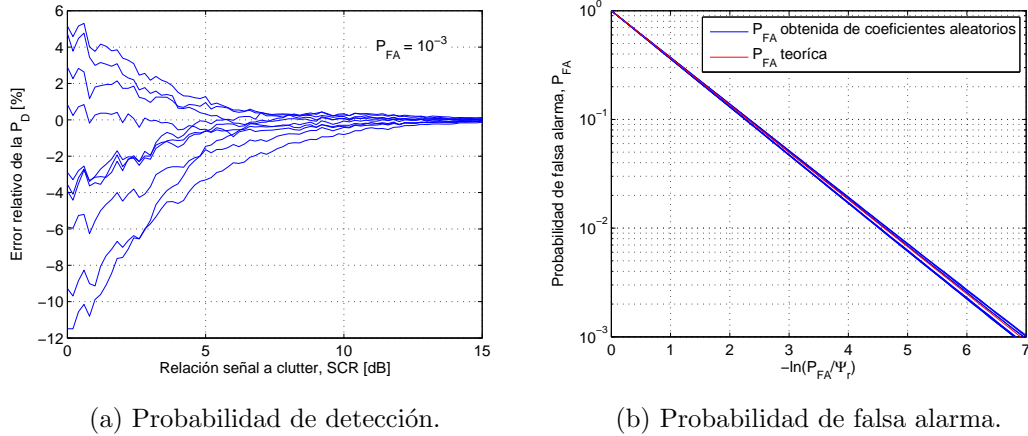


Figura 3.8: Efecto del error en la estimación de los coeficientes sobre el desempeño del detector GARCH.

comparado con el detector GARCH para valores bajos de probabilidad de falsa alarma.

3.5.2. Análisis de sensibilidad

Una de las características deseables para un detector en aplicaciones de radar es que las variaciones de cualquiera de los parámetros desconocidos no afecte significativamente su desempeño, es decir sus probabilidades de detección y falsa alarma. En particular en la deducción del detector GARCH se supuso que los coeficientes del proceso eran conocidos y en la práctica sus valores son estimados. Por este motivo, en presencia de errores de estimación o bien ante pequeños cambios en los parámetros debido a la dinámica del fenómeno que modelan, se espera que el desempeño del detector no varíe notablemente. En el caso del detector GARCH, una sensibilidad baja a las variaciones de los parámetros permitiría estimar los coeficientes con un número reducido de muestras o bien mantener el valor de los coeficientes por períodos de tiempo prolongados antes de actualizarlos.

Para cuantificar la sensibilidad del detector a la variación de sus parámetros se realizaron simulaciones numéricas que consistieron en calcular la P_D y la P_{FA} , considerando que los coeficientes presentan errores de estimación. Dado que no es posible hallar la distribución verdadera de los estimadores de los coeficientes, se supuso que poseían distribución normal. Se consideró nuevamente un proceso GARCH(0,1) y para la media y la varianza de la distribución de cada coeficiente se tomaron los valores obtenidos en el análisis asintótico realizado en la sección 3.4.2, para $R = 7000$ muestras, es decir $k \sim \mathcal{CN}(20, (0.5613)^2)$ y $\beta_1 \sim \mathcal{CN}(0.8, (0.0198)^2)$. En el caso de la probabilidad de detección la simulación se repitió 10 veces, es decir para 10 valores diferentes de los coeficientes. En la figura 3.8a se muestra el error relativo de P_D , $(P_D - \hat{P}_D)/P_D$, obtenido para diferentes valores de SCR, donde se fijó $P_{FA}|\psi_r = 10^{-3}$ para la determinación del umbral. En el caso de la probabilidad de falsa alarma la simulación se repitió 100 veces. La media y la desviación estándar de los estimados de P_{FA} fue 1.10×10^{-3} y 1.34×10^{-4} respectivamente.

Por otro lado, el cálculo de la probabilidad de falsa alarma se repitió variando el valor de $P_{FA}|\psi_r$ y estimando P_{FA} a través de su frecuencia relativa para distintas realizaciones de coeficientes. En la figura 3.8b se presentan las curvas obtenidas junto con la curva teórica dada por (3.62).

Los resultados obtenidos muestran que el detector GARCH es robusto a pequeños cambios en los valores de sus parámetros, al menos desde un punto de vista práctico. Al mismo tiempo la figura 3.8b permite concluir que no posee tasa de falsa alarma constante (CFAR, por sus siglas en inglés). Sin embargo, las perturbaciones de los coeficientes producen pequeñas variaciones en P_{FA} y nuevamente dichas variaciones no son significativas desde un punto de vista práctico.

3.5.3. Resultados

Con el fin de evaluar el desempeño de los detectores en una situación real se utilizó nuevamente el conjunto de datos *stare0* descrito en la sección 2.2.2. En este caso se consideraron los últimos 1024 pulsos o realizaciones (los primeros 1024 fueron utilizados como datos secundarios para estimar los coeficientes del proceso). Para tener control sobre la SCR, en primer lugar se consideró que las realizaciones se extienden hasta la celda de rango 45 de modo de excluir las celdas que contienen el objetivo real, y se insertó un objetivo sintético con el modelo propuesto en la sección 3.5.

En base a los resultados obtenidos en el ajuste a los datos reales presentado en la sección 3.4.3, se consideró un detector GARCH de órdenes (0,1). Es decir, a partir de (3.64) y de (3.65), el umbral está dado por

$$\lambda' = \sqrt{-(\hat{k} + \hat{\beta}_1|v_{r-1}|^2)\ln(P_{FA}|\psi_r)}. \quad (3.69)$$

donde \hat{k} y $\hat{\beta}_1$ son los estimados de los coeficientes obtenidos en la sección 3.4.3.

En estas pruebas nuevamente se consideró el detector Gaussiano. Para este detector el umbral está dado por (3.68) salvo que en este caso debe reemplazarse el valor de la varianza, σ^2 , por su estimado $\hat{\sigma}^2$ análogamente a lo que se hace con los coeficientes para el detector GARCH. En particular para obtener $\hat{\sigma}^2$ el método elegido es el máxima verosimilitud y su expresión viene dada por (2.10).

Por otro lado también se incluyó un tercer detector basado en un algoritmo CFAR para un modelo Weibull para la intensidad del clutter (Ravid y Levanon, 1992), al que se denominará *detector Weibull*. Para este detector se supone que el estadístico es la intensidad de las muestras y que la distribución de la intensidad de clutter es Weibull con parámetros conocidos. Luego, el umbral se obtiene a partir del cálculo de la probabilidad de falsa alarma. Debido a que en la práctica los valores de los parámetros no se conocen son reemplazados por valores estimados. Finalmente, el umbral para el detector Weibull está dado por (Ravid y Levanon, 1992)

$$\lambda'_w = \hat{b}\rho^{1/\hat{c}}, \quad (3.70)$$

donde \hat{b} y \hat{c} son los estimados de máxima verosimilitud de los parámetros de la distribución Weibull que se obtienen a partir de las ecuaciones (2.11) y (2.12), y el coeficiente ρ es función de

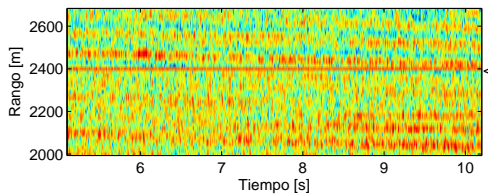
| Objetivo | Detector | P_D | P_{FA} |
|-----------|-----------|--------|----------|
| Sintético | GARCH | 0.4619 | 0.0056 |
| | Gaussiano | 0.0439 | 0.0113 |
| | Weibull | 0.0029 | 0.0027 |
| Real | GARCH | 0.2725 | 0.0048 |
| | Gaussiano | 0.2578 | 0.0093 |
| | Weibull | 0.1592 | 0.0022 |

Tabla 3.5: Tasas empíricas de detección y falsa alarma.

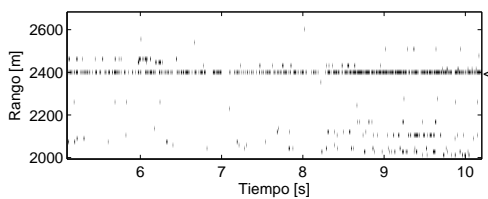
la probabilidad de falsa alarma deseada y del número de muestras utilizadas en la estimación de los coeficientes (Ravid y Levanon, 1992). En el caso en que el número de muestras es suficientemente grande $\rho = -\ln(P_{FA})$.

Los parámetros de los detectores Gaussiano y Weibull fueron estimados utilizando el mismo conjunto de datos que se empleó en la estimación de los coeficientes del detector GARCH. Los umbrales de los tres detectores se fijaron para la $P_{FA} = 10^{-3}$ y se fijó la SCR = 5 dB. En la figura 3.9 se muestra el mapa de detección para cada uno de los detectores. Los píxeles negros corresponden a las celdas de rango donde el estadístico es mayor que el umbral. El objetivo se encuentra ubicado en la celda de rango 27 correspondiente a 2391 m, la misma se indica sobre el lado derecho de las figuras por medio del marcador “<”.

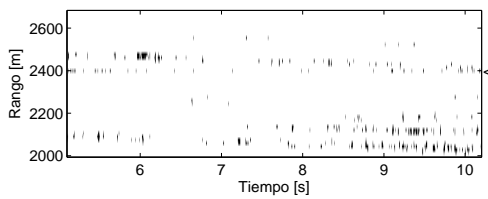
Se observa que el detector GARCH encuentra la verdadera posición del objetivo con solo unas pocas detecciones falsas a pesar de las fuertes reflexiones que presenta el clutter, que se pueden apreciar en la figura 3.9a. También se aprecia que el detector Gaussiano posee la tasa de detección más baja y la tasa de falsa alarma más alta. Como se mencionó, la alta tasa de falsa alarma se debe a que la impulsividad del clutter no es contemplada por el modelo Gaussiano. Estos resultados nuevamente permiten concluir que el modelo GARCH para el clutter es un mejor modelo que el Gaussiano. La distribución Weibull modela mejor que la Gaussiana las colas pesadas que presenta la distribución del clutter. Sin embargo, para tener en cuenta la impulsividad asociada a las colas pesadas, el umbral de detección toma un valor mayor que el umbral del detector Gaussiano. En estas condiciones el detector Weibull reduce la probabilidad de falsa alarma a expensas de la reducción de la probabilidad de detección. En el caso del detector GARCH esto no ocurre debido a que, como ya se mencionó, adapta su umbral en base a la celda de rango previa. En la tabla 3.5 se presentan las tasas de falsa alarma y de detección para cada uno de los algoritmos. La tasa de falsa alarma empírica se define como el número de píxeles detectados, excluyendo los píxeles correspondientes al objetivo, sobre el total de píxeles del mapa de detección, nuevamente excluyendo los píxeles correspondientes al objetivo. Análogamente, la tasa de detección empírica se define como el número de píxeles detectados en la celda de rango donde se ubica el objetivo sobre el total de píxeles para esta celda de rango.



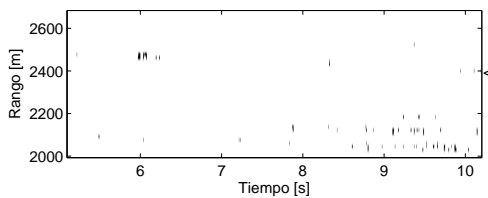
(a) Magnitud de las mediciones.



(b) Detector GARCH.

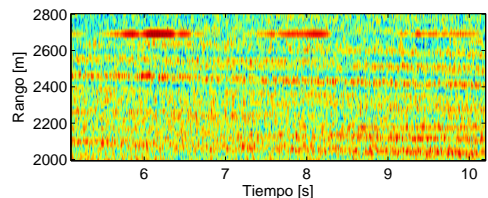


(c) Detector Gaussiano.

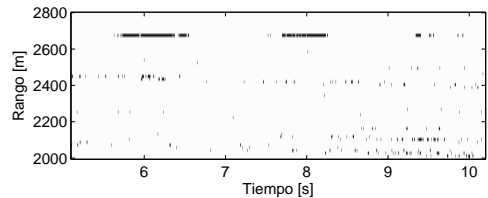


(d) Detector Weibull.

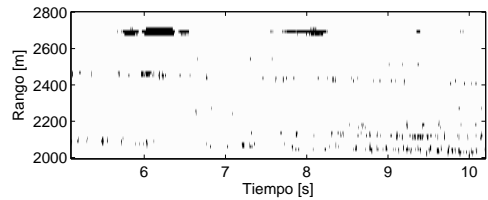
Figura 3.9: Mapa de detección en rango y tiempo con objetivo sintético para el conjunto de datos *stare0*, polarización VV. La ubicación del objetivo es la celda 27, correspondiente a 2391 m.



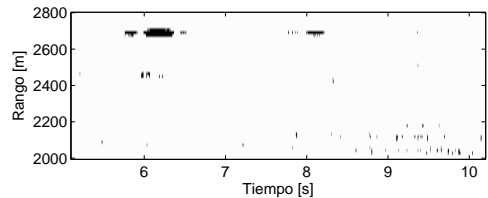
(a) Magnitud de las mediciones.



(b) Detector GARCH.



(c) Detector Gaussiano.



(d) Detector Weibull.

Figura 3.10: Mapa de detección en rango y tiempo con objetivo real para el conjunto de datos *stare0*, polarización VV. La ubicación del objetivo es 2691 m.

Se realizó una segunda prueba en la misma línea de la anterior, utilizando nuevamente las últimos 1024 pulsos del conjunto de datos, pero en esta oportunidad se incluyeron las celdas de rango que contienen el objetivo real. Se corrieron el detector GARCH, el detector Gaussiano y el detector Weibull con los mismos valores de los parámetros que la prueba anterior. En la figura 3.10 se muestra el mapa de detección y en la tabla 3.5 se muestran las tasas de falsa alarma y detección empíricas para cada uno de los detectores con el objetivo real. En este caso se observa que los detectores GARCH y Gaussiano poseen una buena tasa de detección del objetivo, pero el detector GARCH presenta una tasa de falsa alarma menor respecto del detector Gaussiano. Estos resultados son consecuencia de que con el objetivo real la SCR es mayor que la SCR fijada para la prueba anterior. En el caso del detector Weibull nuevamente se observa que presenta una tasa de falsa alarma menor que los otros detectores a expensas de una reducción en la tasa de detección.

Por otra parte, se puede observar que el objetivo real afecta a más de una celda de rango y en el mapa de detección del detector GARCH resultante luce como si existiera enmascaramiento del objetivo. Esto se debe al modo en que se evalúa el umbral luego de que el estadístico lo supera. Sin embargo, esto no es un problema debido a que el objetivo afecta a dos o tres celdas de rango porque la resolución del radar es de 30 m y el intervalo de muestreo en rango es de 15 m. Este fenómeno puede ser un inconveniente cuando dos objetivos se encuentran muy próximos entre sí. En este caso, sería conveniente utilizar una regla diferente para evaluar el umbral en la celda de rango posterior a una presunta detección.

Notar que las tasas de falsa alarma se apartan de los valores teóricos especificados. Esto es de esperarse debido a que en realidad el clutter real no es generado por ninguno de los modelos considerados. Desde un punto de vista práctico, el interés radica en no superar la P_{FA} prefijada, no en alcanzarla exactamente. Como es lógico, esto es así en la medida que ello no signifique disminuir la P_D . La distribución Weibull presenta un buen ajuste de las colas de la distribución y por este motivo es el modelo cuya tasa empírica mejor aproxima la probabilidad de falsa alarma. Sin embargo, como es el detector que presenta la tasa de detección más baja esto no necesariamente implica que sea la mejor distribución para modelar los datos en el problema de detección. Una comparación justa entre los modelos propuestos no debería hacerse solo basada en la calidad del ajuste de las funciones de densidad de probabilidad, sino que debe compararse el desempeño general de la solución al problema de detección. Con el objetivo de llevar a cabo esto, se calculó la curva ROC empírica, es decir el gráfico de P_D en función de P_{FA} calculada a partir del conjunto de datos (Hurtado y Nehorai, 2008).

El modo en el que opera el detector GARCH presenta dos escalas de tiempo: una adaptación lenta dada por la estimación de los parámetros del modelo, realizada como un proceso por lotes; y una adaptación rápida inherente al modelo, en la cual el umbral en cada detección es afectado por la muestra correspondiente a la celda de rango previa a la celda bajo prueba. Entonces, para que la comparación sea justa con los detectores clásicos se incluyen dos versiones de los detectores Gaussiano y Weibull: una versión de adaptación rápida y una versión de adaptación lenta. Los detectores con adaptación lenta corresponden a la versión que se utilizó hasta el

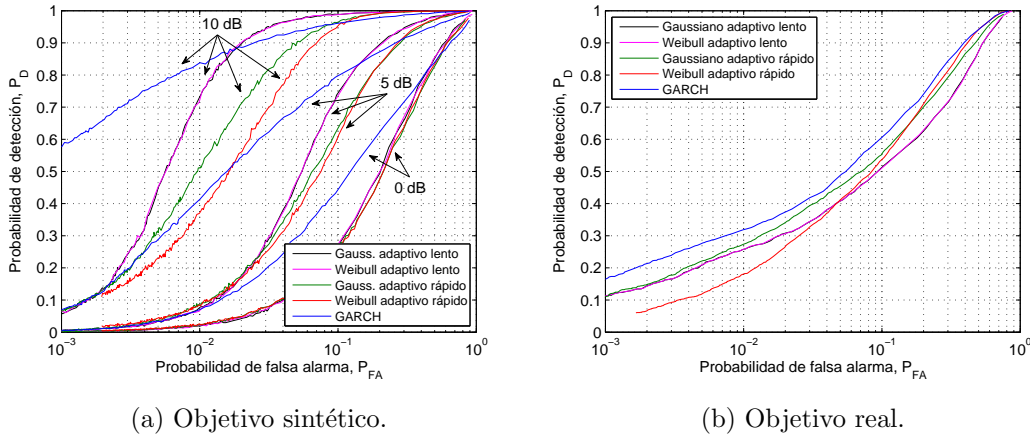


Figura 3.11: Curvas ROC empíricas para el conjunto de datos *stare0*, polarización VV.

momento. Estiman los parámetros de las distribuciones Gaussiana o Weibull de la misma manera que el detector GARCH, es decir utilizando un gran bloque de datos con reflexiones de múltiples pulsos de radar, y determinando los umbrales resultantes para la detección a partir de esa estimación, mientras que se almacenan datos para la próxima actualización del modelo. Por contrapartida, los detectores con adaptación rápida estiman los parámetros de las distribuciones Gaussiana o Weibull a partir de las reflexiones correspondientes a un único pulso de radar con un conjunto de celdas de rango secundarias alrededor de la celda de rango bajo prueba, con una guarda alrededor de esta. Estas versiones de detectores con adaptación rápida buscan imitar la variación temporal del umbral del detector GARCH.

En la figura 3.11a se muestra la curva ROC empírica con datos reales de clutter y objetivo sintético para diferentes valores de SCR. En la figura 3.11b se presenta la curva ROC empírica obtenida con datos de clutter y objetivo reales. Se puede observar que las curvas ROC de los detectores Gaussiano y Weibull con adaptación lenta se superponen. De hecho la misma curva ROC empírica se obtendría para un detector relacionado a cualquier distribución invariante en el tiempo que implique la comparación con un umbral fijo o constante. Las curvas ROC teóricas pueden diferir, pero las curvas empíricas se obtienen simplemente variando el umbral independientemente de la probabilidad de falsa alarma teórica asociada a cada valor del umbral. En cambio, en los detectores Gaussiano y Weibull con adaptación rápida los umbrales cambian en cada decisión, sin embargo su desempeño se degrada debido a que sus parámetros son estimados con un error considerable a causa del pequeño número de muestras utilizado. En el caso del detector Weibull deben estimarse dos parámetros con el mismo conjunto de datos pequeño, con lo cual este efecto se agrava aún más. Si se considera el desempeño en la detección como parámetro de comparación, a partir de las figuras 3.11a y 3.11b, nuevamente podemos concluir que el modelo GARCH es un mejor modelo para los datos que los otros dos presentados en este trabajo.

3.6. Conclusiones

Las olas de cresta espumosa que se forman en el océano a menudo producen picos de amplitud en las señales recibidas por los radares. Este tipo de respuesta impulsiva no es correctamente modelada por procesos Gaussianos. En este capítulo se propuso utilizar procesos GARCH para una mejor representación de las características del clutter marítimo. Al mismo tiempo se desarrolló un esquema de detección basado en este tipo de modelo.

En primer lugar se dió una breve descripción de la estructura de los procesos GARCH y del método de cuasi-máxima verosimilitud para la estimación de sus parámetros. Debido a que no existe una expresión explícita para su fdp se calculó la cota de Carmér-Rao para un proceso ARCH, a partir de la aproximación de su densidad de probabilidad por la de máxima entropía dado un conjunto de momentos conocidos. Se dedujeron las expresiones de los momentos y sus derivadas, necesarias para aproximar la CRB. A partir de un ejemplo numérico se observó que aún con un número reducido de momentos el algoritmo converge al valor que aproxima la cota.

Para validar los resultados se hizo uso de las propiedades asintóticas del estimador QMLE, resultante de maximizar la densidad de probabilidad condicional del proceso ARCH. Esta validación mostró que la aproximación de la densidad del proceso por la de máxima entropía es una buena aproximación y además quedó de manifiesto que el error que se obtiene estimando los parámetros con el QMLE es menor que lo predicho por la CRB. Esto se debe a que con el QMLE se está utilizando la densidad condicional en lugar de la densidad del proceso, y la información extra aportada por conocer la fdp condicional aumenta la información de Fisher y por consiguiente disminuye la CRB.

Debido a que el modelo propuesto considera las realizaciones en rango, el número de muestras necesarias para obtener un error de estimación razonable no es suficiente en una situación práctica. Para resolver este problema se presentó un método de estimación alternativo que utiliza secuencias de datos cortas, sin pérdida en el desempeño. Este procedimiento emplea varias realizaciones independientes del proceso y se mostró que presenta un comportamiento similar al método de estimación original. Utilizando este algoritmo se realizó el ajuste de un proceso GARCH a las mediciones reales del radar IPIX y se compararon los resultados con los de los ajustes realizados en el capítulo 2 a las distribuciones comúnmente utilizadas en la literatura.

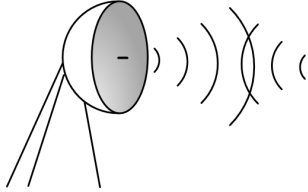
Se desarrolló un algoritmo de detección basado en un modelo GARCH para el clutter. Una de sus ventajas es que es un esquema de detección adaptivo cuyo umbral depende de las muestras de clutter previas para cada celda de rango bajo prueba. De este modo es capaz de capturar las variaciones rápidas del ruido en el tiempo. Por otro lado, es importante destacar que si el clutter es Gaussiano el detector GARCH se reduce a un detector Gaussiano, es decir el procedimiento de estimación dará un valor no nulo de \hat{k} , igual a la varianza del clutter y valores próximos a cero para los estimados de los demás coeficientes.

Asimismo, se evaluaron las probabilidades de detección y falsa alarma por medio de simulaciones Monte Carlo. Se mostró que para valores bajos de SCR el detector GARCH presenta un

mejor desempeño que un detector Gaussiano, es decir presenta una probabilidad de detección mayor con una tasa de falsa alarma menor. Además, este detector presenta la ventaja de decidir pulso a pulso, esto significa que no necesita esperar el arribo de las reflexiones de múltiples pulsos para tomar la decisión. Sin embargo, como se verá en el capítulo 4 extender estas ideas incorporando información de varios pulsos mejorará el desempeño, sin necesidad de hipótesis sobre estacionariedad.

Con respecto a los resultados obtenidos utilizando los datos de clutter real, se mostró que el detector GARCH presenta un mejor desempeño que los detectores Gaussiano y Weibull, excepto para valores altos de SCR y alta probabilidad de falsa alarma. En efecto, bajo estas últimas condiciones favorables los tres detectores alcanzan aproximadamente la misma probabilidad de detección. El éxito del detector GARCH se puede explicar a través de su variación temporal. Tomando la intensidad de la muestra como el estadístico para los tres detectores, en el caso del detector GARCH el umbral en cada comparación es función de la intensidad de la muestra previa, pero los parámetros del modelo no cambian rápidamente. Este comportamiento no puede obtenerse con los detectores clásicos en los cuales existe una relación de compromiso entre la calidad en la estimación de los parámetros de la distribución y la velocidad de adaptación.

Finalmente, pensando en una aplicación en tiempo real el algoritmo de detección posee una carga computacional práctica. Si bien requiere las reflexiones de múltiples pulsos para realizar la actualización de los estimados de los coeficientes, se mostró que el detector es robusto respecto a pequeños errores en los valores de sus parámetros. De este modo, no hay necesidad de estimar los parámetros frecuentemente. En realidad, los coeficientes cambian cuando la condiciones ambientales cambian drásticamente. Esta flexibilidad permite esperar las muestras de señal necesarias para realizar la estimación mientras continua la ejecución del detector. Por lo tanto, comparando solo los algoritmos de detección, el detector GARCH posee una carga computacional ligeramente mayor que los detectores Gaussiano y Weibull con adaptación lenta debido a que sus umbrales son constantes. Sin embargo, posee una carga computacional menor que los detectores Gaussiano y Weibull con adaptación rápida debido a que el cálculo de sus umbrales requiere la estimación de los parámetros de la distribución para cada celda de rango.



Capítulo 4

Detección en rango y tiempo

4.1. Introducción

En las aplicaciones de radar, para lograr el desempeño deseado, es importante modelar apropiadamente las señales resultantes de la contribución de las reflexiones en el entorno del objetivo de interés. Estas reflexiones no deseadas, o clutter, son en muchos casos de naturaleza no homogénea y su estadística es desconocida. A causa de su naturaleza cambiante, es frecuente considerar el empleo de algoritmos adaptivos. Debido a su tratabilidad matemática, los esquemas de detección adaptiva en radar basados en modelos de clutter Gaussiano han sido extensamente investigados (Reed *et al.*, 1974; Kelly, 1986; Bose y Steinhardt, 1995; Sheikhi *et al.*, 1998; Kraut y Scharf, 1999). Una de las soluciones más conocidas es el test de relación de verosimilitud generalizado (GLRT por sus siglas en inglés) (Kelly, 1986), diseñado considerando un modelo de clutter Gaussiano con media nula y matriz de covarianza desconocida. En este esquema de detección se utilizan dos conjuntos de datos de entrada los cuales se denominan primario y secundario. La posibilidad de que exista señal presente solo se contempla para los datos primarios, mientras que se supone que los datos secundarios solo se componen de muestras de clutter, independientes e idénticamente distribuidas de las componentes de clutter de los datos primarios. En el caso del detector de relación de verosimilitud autorregresiva generalizada (ARGLR por sus siglas en inglés) (Sheikhi *et al.*, 1998) el clutter es modelado a través de un proceso autorregresivo (AR) de modo de imponerle una estructura a su matriz de covarianza. En este caso el detector se ajusta a si mismo utilizando únicamente el conjunto de datos primario.

Sin embargo, existe evidencia experimental que demuestra que el modelo Gaussiano no representa adecuadamente al clutter en muchas situaciones prácticas de interés (Haykin *et al.*, 1991; Farina *et al.*, 1997). De este modo dependiendo de la aplicación se han propuesto modelos de clutter basados en distribuciones tales como la Log-normal, la Weibull, la K (Sangston y Gerlach, 1994; Shnidman, 1999), en densidades de probabilidad compuestas generalizadas (Anastassopoulos *et al.*, 1999; Shang y Song, 2011; Sangston *et al.*, 2012) o en procesos aleatorios esféricamente invariantes (Conte y Longo, 1987). El problema de detección para algunos de estos modelos de clutter no Gaussianos también ha sido estudiado (Conte *et al.*, 1996; Gini *et al.*, 1995; Gini y Greco, 1999). El detector adaptivo lineal-cuadrático (ALQ por sus siglas en inglés) (Gini y Greco, 1999) es una solución sub-óptima basada en un modelo de clutter Gaussiano compuesto. La formulación del problema para la deducción de este detector es análoga a la del detector GLRT, y también requiere los conjuntos de datos primario y secundario.

En el capítulo 3 se propuso un esquema de detección modelando el clutter a través de un proceso GARCH. El detector resultante mostró un desempeño superior a detectores basados en modelos Gaussiano y Weibull para el clutter en diferentes situaciones prácticas. Sin embargo, el detector GARCH tiene la desventaja de que no permite incorporar información de varios pulsos de radar en la decisión debido a que el proceso GARCH fue utilizado para modelar el clutter solo en rango (o tiempo rápido). Los procesos GARCH multidimensionales han sido utilizados en el modelado de coeficientes wavelet en problemas de detección de anomalías en aplicaciones de sonar (Noiboar y Cohen, 2007) y en la supresión de speckle en imágenes de radar de apertura sintética (Amirmazlaghani *et al.*, 2009).

En este capítulo se extienden las ideas previas proponiendo un modelo de clutter que combina un proceso GARCH en dos dimensiones (GARCH-2D) con un proceso AR con el objetivo de incorporar la información de múltiples pulsos en la decisión y se deriva el detector adecuado para este modelo de clutter (Pascual *et al.*, 2014c). La parte GARCH-2D del modelo preserva la propiedad impulsiva de los procesos GARCH, es decir este modelo preserva las colas pesadas de la fdp, y el modelo AR en las innovaciones permite modelar la correlación pulso a pulso existente en los datos, como se vió en la sección 2.2.2.

En la sección 4.2 se describe el modelo AR-GARCH-2D y se presenta un método de estimación para sus parámetros. En la sección 4.3 se desarrolla un algoritmo de detección adaptivo en rango y tiempo basado en el modelo AR-GARCH-2D. Se lleva a cabo un análisis teórico para determinar su probabilidad de falsa alarma. Por medio de simulaciones numéricas en la sección 4.4 se calcula su probabilidad de detección para evaluar el desempeño del detector AR-GARCH-2D. Análogamente a lo realizado para el detector GARCH, se analiza como afecta un error de estimación de sus parámetros en su desempeño. A pesar de que el detector resultante no es de tasa de falsa alarma constante (CFAR) con respecto a los coeficientes del proceso, se muestra que en situaciones prácticas presenta un comportamiento robusto, es decir la probabilidad de falsa alarma no cambia significativamente cuando varían los valores de los coeficientes. En la sección 4.5 se examina el desempeño de detectores AR-GARCH-2D con diferentes órdenes del modelo en una situación real utilizando los datos medidos y se aplica un

criterio de selección de modelos para elegir el modelo AR-GARCH-2D que mejor ajusta a los datos. Finalmente, se compara el detector resultante con los detectores GLRT (Kelly, 1986), ALQ (Gini y Greco, 1999) y ARGLR (Sheikhi *et al.*, 1998) utilizando los datos medidos de clutter marítimo.

4.2. Modelado del clutter

4.2.1. Procesos AR-GARCH-2D

Sea el proceso doblemente indexado v_{rt} un proceso aleatorio de dos dimensiones que representa el clutter. El primer índice, r , corresponde a la dimensión tiempo rápido (o rango) y el segundo índice, t , corresponde a la dimensión tiempo lento (o pulsos). Luego, basado en un proceso GARCH-2D (Amirmazlaghani *et al.*, 2009) se define un proceso AR-GARCH-2D complejo, v_{rt} , como (Pascual *et al.*, 2014c)

$$v_{rt} = \sigma_{rt} z_{rt} + \sum_{j=1}^M a_{0j} v_{r,t-j}, \quad z_{rt} \sim \mathcal{CN}(0, 1) \text{ i.i.d.}, \quad (4.1)$$

$$\sigma_{rt}^2 = k + \sum_{i,j \in \Lambda_1} \alpha_{ij} \sigma_{r-i,t-j}^2 + \sum_{i,j \in \Lambda_2} \beta_{ij} |v_{r-i,t-j}|^2, \quad (4.2)$$

donde a_{0j} son los coeficientes de la parte autorregresiva del proceso, σ_{rt}^2 es la varianza condicional, dada por (4.2), siendo k el término independiente que fija el piso de la varianza, α_{ij} los coeficientes de la parte autorregresiva de la varianza condicional y β_{ij} los coeficientes de las intensidades al cuadrado de los valores pasados del proceso. Para asegurar que la varianza condicional, σ_{rt}^2 , sea siempre positiva se deben imponer las siguientes restricciones

$$k > 0 \quad \alpha_{ij} \geq 0, (i, j) \in \Lambda_1 \quad \beta_{ij} \geq 0, (i, j) \in \Lambda_2. \quad (4.3)$$

Sea también

$$\Lambda_1 = \{(i, j) / 0 \leq i \leq p_1, 0 \leq j \leq p_2, (i, j) \neq (0, 0)\}, \quad (4.4)$$

$$\Lambda_2 = \{(i, j) / 0 \leq i \leq q_1, 0 \leq j \leq q_2, (i, j) \neq (0, 0)\}, \quad (4.5)$$

donde p_1 , p_2 , q_1 y q_2 son los órdenes del proceso. El proceso AR-GARCH-2D definido arriba difiere de los procesos GARCH-2D propuestos en (Amirmazlaghani *et al.*, 2009) en la inclusión del segundo término en el lado derecho de la igualdad de (4.1), es decir el término de la parte autorregresiva en el segundo índice de tiempo. De este modo, mientras que el proceso GARCH-2D es no correlacionado, el proceso AR-GARCH-2D permite modelar la correlación pulso a pulso, preservando la propiedad impulsiva de los procesos GARCH. Los procesos AR-GARCH-2D modelan las observaciones como procesos Gaussianos correlacionados con varianza condicional variante en el tiempo (4.1). Para cada ubicación (r, t) , las vecindades de la observación y de la varianza condicional juegan un rol en el valor actual de la varianza condicional (4.2), lo que lleva al agrupamiento de las variaciones. En este modelo el número de

coeficientes en la varianza condicional aumenta drásticamente con los órdenes p_1 , p_2 , q_1 y q_2 lo cual obstaculiza la estimación precisa de los coeficientes de proceso. Para hacer que el modelo sea tratable a los fines prácticos es necesario imponerle estructura adicional. Como se vió en la sección 3.4.3 para un modelo GARCH-1D, un proceso GARCH con solo dos coeficientes, k y el primer β , es suficiente para capturar la mayor parte del comportamiento estadístico del clutter. Por este motivo de aquí en adelante el análisis se restringirá al modelo

$$v_{rt} = \sigma_{rt} z_{rt} + \sum_{j=1}^M a_{0j} v_{r,t-j}, \quad z_{rt} \sim \mathcal{CN}(0, 1) \text{ i.i.d.}, \quad (4.6)$$

$$\sigma_{rt}^2 = k + \sum_{j=0}^1 \alpha_{1j} \sigma_{r-1,t-j}^2 + \sum_{j=0}^1 \beta_{1j} |v_{r-1,t-j}|^2, \quad (4.7)$$

donde los coeficientes α_{01} y β_{01} fueron excluidos a fin de obtener una expresión cerrada para el detector que se deducirá en la sección 4.3. Los coeficientes α 's imponen cierta suavidad sobre σ_{rt}^2 mientras que los coeficientes β 's determinan su grado de dependencia con las observaciones pasadas. Restringiendo esta dependencia a la celda de rango y pulso previos solo se capturará la dinámica dominante de la varianza condicional. Probablemente el clutter producido por la mayor parte de los entornos sea adecuadamente modelado con esta dependencia de vecinos más cercanos. Por otro lado si algún entorno exhibe clutter con una dinámica más compleja, un modelo con mayor número de coeficientes no necesariamente implicará un mejor ajuste en la práctica, ya que un mayor número de coeficientes llevará a mayores errores de estimación para un dado conjunto de datos. Además, se restringe M a que varíe solo de 0 a 3, debido a que un modelo AR de tercer orden para el clutter a mostrado ser un buen compromiso entre fidelidad del modelo y tratabilidad matemática (Greco *et al.*, 2004).

Sea $\psi_{r-1,t}$ el conjunto de toda la información hasta la celda de rango $r-1$ y el instante t , es decir $\sigma_{\rho\tau}^2$ y $v_{\rho\tau}$ para $\rho \leq r-1$ y $\tau \leq t$. Luego, si se condiciona v_{rt} a $(\psi_{r-1,t}; \{v_{r,t-j}\}_{j=1}^M)$ de (4.6) se observa que σ_{rt} y $\sum_{j=1}^M a_{0j} v_{r,t-j}$ están dados y solo $z_{rt} \sim \mathcal{CN}(0, 1)$ es aleatorio, entonces

$$v_{rt} | (\psi_{r-1,t}; \{v_{r,t-j}\}_{j=1}^M) \sim \mathcal{CN} \left(\sum_{j=1}^M a_{0j} v_{r,t-j}; \sigma_{rt}^2 \right). \quad (4.8)$$

Notar que el conjunto de muestras del proceso $\{v_{r,t-j}\}_{j=1}^M$ no se encuentra incluido en $\psi_{r-1,t}$.

Este modelo puede diverger debido a que involucra ecuaciones autorregresivas. Para evitar esta situación es posible imponer condiciones sobre los coeficientes del proceso con el fin de asegurar que resulte estacionario en sentido amplio y con varianza finita, análogamente a los procesos GARCH-1D (Bollerslev, 1986) o a los procesos GARCH-2D (Noiboar y Cohen, 2007). Se puede demostrar que $\mathbb{E}\{v_{rt}\} = 0$ y que la varianza no condicional es finita e igual a

$$\mathbb{V}\{v_{rt}\} = \frac{k}{1 - \sum_{j=0}^1 \alpha_{1j} - \sum_{j=0}^1 \beta_{1j} - \left(1 - \sum_{j=0}^1 \alpha_{1j}\right) \phi_M}, \quad (4.9)$$

siempre y cuando se satisfaga la condición

$$\sum_{j=0}^1 \alpha_{1j} + \sum_{j=0}^1 \beta_{1j} + \left(1 - \sum_{j=0}^1 \alpha_{1j}\right) \phi_M < 1, \quad (4.10)$$

donde ϕ_M es una constante que depende del orden M y de los coeficientes a_{0j} . Para $M = 0, 1, 2, 3$ las expresiones de ϕ_M son respectivamente

$$\phi_0 = 0, \quad (4.11)$$

$$\phi_1 = |a_{01}|^2, \quad (4.12)$$

$$\phi_2 = \sum_{j=1}^2 |a_{0j}|^2 + 2\text{Re} \left\{ \frac{a_{01}a_{02}^*(a_{01} + a_{01}^*a_{02})}{1 - |a_{02}|^2} \right\}, \quad (4.13)$$

$$\begin{aligned} \phi_3 = \sum_{j=1}^3 |a_{0j}|^2 + 2\text{Re} \{ & a_{01}a_{02}a_{03}^* + \\ & + [a_{01}(a_{02}^* + a_{01}a_{03}^*) + a_{03}^*(a_{02} + a_{01}^*a_{03})] \frac{a_{01} + a_{02}^*a_{03} + \frac{(a_{01}^* + a_{02}a_{03}^*)(a_{02} + a_{01}^*a_{03})}{1 - |a_{03}|^2}}{1 - |a_{03}|^2 - \frac{|a_{02} + a_{01}^*a_{03}|^2}{1 - |a_{03}|^2}} \}, \end{aligned} \quad (4.14)$$

donde $*$ denota complejo conjugado y $\text{Re}\{\cdot\}$ la parte real de un número complejo. En el apéndice B se presenta la deducción ϕ_M para $M = 2$, con un procedimiento similar es posible obtener las expresiones correspondientes a los demás valores de M .

Como en el caso de los procesos GARCH-1D y GARCH-2D, no existe una expresión explícita para la función densidad de probabilidad de los procesos AR-GARCH-2D.

4.2.2. Estimación de parámetros

Para ajustar el modelo AR-GARCH-2D al clutter se debe utilizar un conjunto de datos que no incluyan reflexiones en un objetivo. Suponiendo que un radar pulsado monoestático transmite un tren de T pulsos coherente y que el receptor muestrea la señal reflejada de cada pulso a la salida del filtro adaptado formando R celdas de rango, siguiendo el esquema de la figura 2.1. Las muestras de la envolvente compleja a la salida del demodulador en cuadratura pueden agruparse en una matriz $\mathbf{V} \in \mathbb{C}^{R \times T}$ con elementos complejos $[\mathbf{V}]_{rt} = v_{rt}$, donde las filas corresponden a la dimensión rango y las columnas corresponden a la dimensión tiempo lento, como se describe en la figura 4.1.

Una manera de realizar la estimación sería utilizando el método de máxima verosimilitud, pero en este caso tampoco se cuenta con una expresión explícita para la fdp de un proceso AR-GARCH-2D. Para superar esta dificultad, análogamente a lo realizado para los procesos GARCH, se considera en su lugar la función de verosimilitud condicional $f(\mathbf{V}|\boldsymbol{\psi}_{R-1,-M}; \boldsymbol{\psi}_{0T}; \mathbf{v}_{1-M}; \cdots; \mathbf{v}_0)$ que puede escribirse como

$$f(\mathbf{V}|\boldsymbol{\psi}_{R-1,-M}; \boldsymbol{\psi}_{0T}; \mathbf{v}_{1-M}; \cdots; \mathbf{v}_0) = \prod_{r=1}^R \prod_{t=1}^T f(v_{rt}|\boldsymbol{\psi}_{r-1,t}; \{v_{t-j}\}_{j=1}^M), \quad (4.15)$$

donde $\mathbf{v}_t = [v_{1t} \ v_{2t} \ \cdots \ v_{Rt}]^T$ y las fdp's condicionales $f(v_{rt}|\boldsymbol{\psi}_{r-1,t}; \{v_{t-j}\}_{j=1}^M)$ están dadas por (4.8). Luego, se define la función $\mathcal{L}_4(\boldsymbol{\theta})$ como

$$\mathcal{L}_4(\boldsymbol{\theta}) = \sum_{r=1}^R \sum_{t=1}^T \left[\ln(\sigma_{rt}^2) + \frac{1}{\sigma_{rt}^2} \left| v_{rt} - \sum_{j=1}^M a_{0j}v_{r,t-j} \right|^2 \right], \quad (4.16)$$

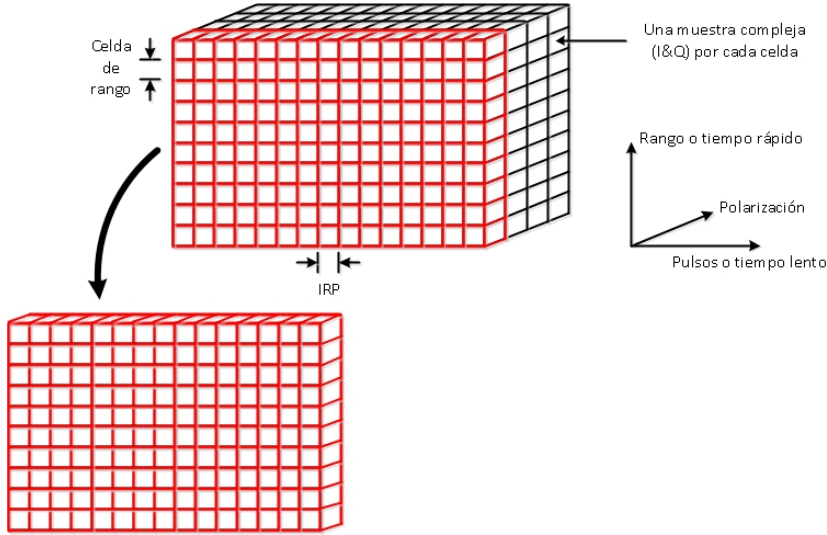


Figura 4.1: Modelado del clutter a través de procesos AR-GARCH-2D.

donde $\boldsymbol{\theta} = [k \ \alpha_{10} \ \alpha_{11} \ \beta_{10} \ \beta_{11} \ \text{Re}\{a_{01}\} \ \text{Im}\{a_{01}\} \ \cdots \ \text{Re}\{a_{0M}\} \ \text{Im}\{a_{0M}\}]^T$ es el vector de parámetros que se desea estimar. La función de verosimilitud logarítmica condicional $-\ln(f(\mathbf{V}|\boldsymbol{\psi}_{R-1,-M}; \boldsymbol{\psi}_{0T}; \mathbf{v}_{1-M}; \cdots; \mathbf{v}_0))$ difiere de $\mathcal{L}_4(\boldsymbol{\theta})$ solo en una constante.

Finalmente, el QMLE, $\hat{\boldsymbol{\theta}}$, es el valor de $\boldsymbol{\theta}$ que maximiza (4.15) o, equivalentemente, que minimiza $\mathcal{L}_4(\boldsymbol{\theta})$ sujeta a las restricciones dadas por (4.3) y (4.10). En resumen, se puede escribir

$$\hat{\boldsymbol{\theta}} = \underset{\boldsymbol{\theta} \in \Theta}{\text{argmin}} \mathcal{L}_4(\boldsymbol{\theta}), \quad (4.17)$$

donde Θ es el conjunto de valores de $\boldsymbol{\theta}$ que satisfacen las restricciones (4.3) y (4.10).

Como en el caso de los funcionales presentados en el capítulo 3, en el apéndice A se presenta una expresión para el gradiente de $\mathcal{L}_4(\boldsymbol{\theta})$, que será necesaria para resolver el problema de optimización.

4.3. Esquema de detección

Para una dada celda de rango r , pueden agruparse N muestras complejas correspondientes a pulsos consecutivos $\{y_{rt}\}_{t=1}^N$ en un vector $\mathbf{y}^r = [y_{r1} \ y_{r2} \ \cdots \ y_{rN}]^T$ de dimensión $N \times 1$. Luego, el procedimiento de detección está dado por un test de hipótesis binario entre las hipótesis H_0 y H_1 una vez que \mathbf{y}^r ha sido medido para la celda de rango bajo prueba

$$\begin{aligned} H_0 : \mathbf{y}^r &= \mathbf{v}^r \\ H_1 : \mathbf{y}^r &= \mathbf{s}^r + \mathbf{v}^r. \end{aligned} \quad (4.18)$$

Bajo la hipótesis nula H_0 , se supone que las mediciones consisten solo de clutter \mathbf{v}^r . Es importante destacar que se considera que el ruido electrónico es despreciable o bien que se encuentra absorbido por el modelo de clutter. Bajo la hipótesis H_1 , se supone que las mediciones son una combinación resultante de clutter y los ecos producidos en un objetivo, \mathbf{s}^r . El vector de señal

se modela como $\mathbf{s}^r = \gamma \mathbf{p}^r$, donde γ es un escalar complejo determinístico pero desconocido, mientras que \mathbf{p}^r es un vector complejo perfectamente conocido cuyos elementos están dados por $[\mathbf{p}^r]_t = p_{rt} = e^{j\Omega t}$, con $\Omega = 2\pi f_d T_s$ donde f_d es la frecuencia Doppler y T_s es el intervalo de repetición de pulso del radar (Gini y Greco, 1999). Se considera que la estadística del clutter es la misma para ambas hipótesis y se modela como un proceso AR-GARCH-2D.

El GLRT descrito en la sección 2.4.2 involucra el cociente de las funciones de verosimilitud bajo cada una de las hipótesis, evaluadas en sus respectivos máximos -dados por los MLEs de todos los parámetros desconocidos correspondientes para cada hipótesis (Kay, 1998). Debido a que no existe una expresión explícita para la distribución de un proceso AR-GARCH-2D, entonces tampoco existe una expresión para la relación de verosimilitud requerida para obtener la regla de decisión. Por lo tanto, en este caso no se puede utilizar el enfoque clásico del GLRT. En su lugar, nuevamente se desarrolla el test en base a las funciones de verosimilitud condicionales dadas las observaciones del proceso hasta la celda de rango bajo prueba (Pascual *et al.*, 2013b, 2014c).

Las fdp's condicionales están dadas por

$$f_i(\mathbf{y}^r | \boldsymbol{\psi}_{r-1,N}; \{y_{r,t-M}\}_{t=1}^M) = \prod_{t=1}^N f_i(y_{rt} | \boldsymbol{\psi}_{r-1,t}; \{y_{r,t-j}\}_{j=1}^M), \quad (4.19)$$

con $i = 0, 1$. Bajo la hipótesis nula, a partir de (4.8) se obtiene

$$f_0(y_{rt} | \boldsymbol{\psi}_{r-1,t}; \{y_{r,t-j}\}_{j=1}^M) = \frac{1}{\pi \sigma_{rt}^2} e^{-\frac{1}{\sigma_{rt}^2} |y_{rt} - \sum_{j=1}^M a_{0j} y_{r,t-j}|^2}. \quad (4.20)$$

Bajo la hipótesis H_1 , los elementos de \mathbf{y}^r pueden escribirse como $y_{rt} = \gamma p_{rt} + v_{rt}$ y de (4.6) resulta

$$y_{rt} = \gamma p_{rt} + \sum_{j=1}^M a_{0j} v_{r,t-j} + \sigma_{rt} z_{rt}. \quad (4.21)$$

Expresando las muestras de clutter como $v_{r,t-j} = y_{r,t-j} - \gamma p_{r,t-j}$ para $j = 1, \dots, M$, a partir de (4.21), se obtiene

$$y_{rt} = \sum_{j=1}^M a_{0j} y_{r,t-j} + \gamma \left(p_{rt} - \sum_{j=1}^M a_{0j} p_{r,t-j} \right) + \sigma_{rt} z_{rt}. \quad (4.22)$$

De este modo, $f_1(y_{rt} | \boldsymbol{\psi}_{r-1,t}; \{y_{r,t-j}\}_{j=1}^M)$ es también una fdp normal pero con valor diferente para su media

$$f_1(y_{rt} | \boldsymbol{\psi}_{r-1,t}; \{y_{r,t-j}\}_{j=1}^M) = \frac{1}{\pi \sigma_{rt}^2} e^{-\frac{1}{\sigma_{rt}^2} |y_{rt} - \sum_{j=1}^M a_{0j} y_{r,t-j} - \gamma (p_{rt} - \sum_{j=1}^M a_{0j} p_{r,t-j})|^2}. \quad (4.23)$$

Siguiendo el enfoque del GLRT debería maximizarse $f_i(\mathbf{y}^r | \boldsymbol{\psi}_{r-1,N}; \{y_{r,t-M}\}_{t=1}^M)$ con respecto a los parámetros desconocidos $\boldsymbol{\theta}$ y γ , de manera separada para cada $i = 0, 1$. Sin embargo, la aplicación de esta metodología para los coeficientes del proceso AR-GARCH-2D lleva a un problema de maximización no-lineal multi-dimensional para el cual no existe una expresión

cerrada de su solución. Por este motivo, se propone la solución alternativa utilizada en la sección 3.5 para la detección en rango, la cual consiste en considerar que los valores de los coeficientes del proceso AR-GARCH-2D son conocidos. Luego, en la implementación del algoritmo serán estimados como se explicó en la sección 4.2.2, utilizando un conjunto de datos secundario que solo contenga muestras de clutter.

Bajo esta consideración el único parámetro desconocido es γ , y solo aparece en la fdp de la hipótesis H_1 .

A continuación se deducirá el valor de γ que maximiza $f_1(\mathbf{y}^r | \boldsymbol{\psi}_{r-1,N}; \{y_{r,t-M}\}_{t=1}^M)$ a partir de la condición necesaria de primer orden que debe satisfacer para ser un máximo local (Fletcher, 1987). $f_1(\mathbf{y}^r | \boldsymbol{\psi}_{r-1,N}; \{y_{r,t-M}\}_{t=1}^M)$ es una función real del parámetro complejo γ y es analítica respecto a γ y a γ^* , donde γ^* denota el conjugado de γ . Entonces, una condición necesaria y suficiente para que tenga un punto estacionario es que

$$\frac{\partial f_1(\mathbf{y}^r | \boldsymbol{\psi}_{r-1,N}; \{y_{r,t-M}\}_{t=1}^M)}{\partial \gamma} = 0, \quad (4.24)$$

donde γ^* es tratado como una constante en la derivada parcial (Brandwood, 1983).

De manera equivalente, se pueden obtener los puntos estacionarios de $\ln(f_1(\mathbf{y}^r | \boldsymbol{\psi}_{r-1,N}; \{y_{r,t-M}\}_{t=1}^M))$, ya que $\ln\{\cdot\}$ es una función monótonamente creciente. $f_1(\mathbf{y}^r | \boldsymbol{\psi}_{r-1,N}; \{y_{r,t-M}\}_{t=1}^M)$ se obtiene reemplazando (4.23) en (4.19) y se toma $\ln\{\cdot\}$ para hallar la expresión de $\ln(f_1(\mathbf{y}^r | \boldsymbol{\psi}_{r-1,N}; \{y_{r,t-M}\}_{t=1}^M))$

$$\ln(f_1) = \sum_{t=1}^N \left[-\ln(\pi \sigma_{rt}^2) - \frac{1}{\sigma_{rt}^2} \left| y_{rt} - \sum_{j=1}^M a_{0j} y_{r,t-j} - \gamma \left(p_{rt} - \sum_{j=1}^M a_{0j} p_{r,t-j} \right) \right|^2 \right], \quad (4.25)$$

donde $\ln(f_1)$ denota $\ln(f_1(\mathbf{y}^r | \boldsymbol{\psi}_{r-1,N}; \{y_{r,t-M}\}_{t=1}^M))$ por cuestiones de espacio. Luego, a partir de (4.25) su derivada con respecto a γ está dada por

$$\frac{\partial \ln(f_1)}{\partial \gamma} = \sum_{t=1}^N \frac{1}{\sigma_{rt}^2} \left(p_{rt} - \sum_{j=1}^M a_{0j} p_{r,t-j} \right) \left[\left(y_{rt} - \sum_{j=1}^M a_{0j} y_{r,t-j} \right) - \gamma \left(p_{rt} - \sum_{j=1}^M a_{0j} p_{r,t-j} \right) \right]^*. \quad (4.26)$$

Entonces, (4.26) es cero cuando $\gamma = \hat{\gamma}$ resultando

$$\sum_{t=1}^N \frac{1}{\sigma_{rt}^2} \left(p_{rt} - \sum_{j=1}^M a_{0j} p_{r,t-j} \right) \left(y_{rt} - \sum_{j=1}^M a_{0j} y_{r,t-j} \right)^* = \hat{\gamma}^* \sum_{t=1}^N \frac{1}{\sigma_{rt}^2} \left| p_{rt} - \sum_{j=1}^M a_{0j} p_{r,t-j} \right|^2. \quad (4.27)$$

A partir de (4.27) se obtiene la expresión para $\hat{\gamma}$

$$\hat{\gamma} = \frac{\sum_{t=1}^N \frac{1}{\sigma_{rt}^2} \left(p_{rt} - \sum_{j=1}^M a_{0j} p_{r,t-j} \right)^* \left(y_{rt} - \sum_{j=1}^M a_{0j} y_{r,t-j} \right)}{\sum_{t=1}^N \frac{1}{\sigma_{rt}^2} \left| p_{rt} - \sum_{j=1}^M a_{0j} p_{r,t-j} \right|^2}, \quad (4.28)$$

que es valor de γ que maximiza $f_1(\mathbf{y}^r | \boldsymbol{\psi}_{r-1,N}; \{y_{r,t-M}\}_{t=1}^M)$.

Por lo tanto, se utiliza la relación de verosimilitud condicional, $\Lambda(\mathbf{y}^r)$, y se establece la siguiente regla de decisión

$$\Lambda(\mathbf{y}^r) = \frac{f_1(\mathbf{y}^r, \hat{\gamma} | \boldsymbol{\psi}_{r-1,N}; \{y_{r,t-M}\}_{t=1}^M)}{f_0(\mathbf{y}^r | \boldsymbol{\psi}_{r-1,N}; \{y_{r,t-M}\}_{t=1}^M)} \underset{H_0}{\overset{H_1}{\gtrless}} \lambda, \quad (4.29)$$

donde λ es el umbral de decisión que debe ser determinado. Reemplazando las expresiones de las fdp's y de $\hat{\gamma}$ en (4.29), es posible reescribir la regla de decisión como

$$\left| \sum_{t=1}^N \frac{1}{\sigma_{rt}^2} \left(p_{rt} - \sum_{j=1}^M a_{0j} p_{r,t-j} \right) \left(y_{rt} - \sum_{j=1}^M a_{0j} y_{r,t-j} \right) \right| \underset{H_0}{\overset{H_1}{\gtrless}} \eta, \quad (4.30)$$

donde el umbral η está dado por

$$\eta = \sqrt{\lambda \sum_{t=1}^N \frac{1}{\sigma_{rt}^2} \left| p_{rt} - \sum_{j=1}^M a_{0j} p_{r,t-j} \right|^2}. \quad (4.31)$$

En el enfoque clásico, el siguiente paso sería obtener una expresión de la probabilidad de falsa alarma, P_{FA} , en términos del umbral, η or λ , y luego invertirla para fijar el umbral en función de P_{FA} . Este procedimiento requiere conocer la distribución del estadístico en (4.30) bajo la hipótesis H_0 . Nuevamente, se presenta el problema de que no se cuenta con una expresión explícita de la fdp de v_{rt} o y_{rt} . Por este motivo, análogamente a lo que hizo en la sección 3.5, para determinar el umbral se recurre a la probabilidad de falsa alarma condicional dado $\boldsymbol{\psi}_{r-1,N}$, es decir $P_{FA} | \boldsymbol{\psi}_{r-1,N}$. Teniendo en cuenta que bajo la hipótesis H_0 $y_{rt} = v_{rt}$, entonces

$$y_{rt} - \sum_{j=1}^M a_{0j} y_{r,t-j} = v_{rt} - \sum_{j=1}^M a_{0j} v_{r,t-j} = \sigma_{rt} z_{rt}, \quad (4.32)$$

y en estas condiciones, el estadístico se reduce a

$$\left| \sum_{t=1}^N \frac{1}{\sigma_{rt}} \left(p_{rt} - \sum_{j=1}^M a_{0j} p_{r,t-j} \right) z_{rt} \right| = \left| \sum_{t=1}^N w_{rt} \right|, \quad (4.33)$$

donde

$$w_{rt} | \boldsymbol{\psi}_{r-1,N} \sim \mathcal{CN} \left(0; \frac{1}{\sigma_{rt}} \left| p_{rt} - \sum_{j=1}^M a_{0j} p_{r,t-j} \right|^2 \right). \quad (4.34)$$

Si se define $S_r = \sum_{t=1}^N w_{rt}$ y $\Gamma = |S_r|$, luego $S_r | \boldsymbol{\psi}_{r-1,N}$ es la suma de N variables aleatorias complejas con distribución normal e independientes. Entonces, de (4.34) $S_r | \boldsymbol{\psi}_{r-1,N} \sim \mathcal{CN}(0, \sigma_s^2)$, con

$$\sigma_s^2 = \sum_{t=1}^N \frac{1}{\sigma_{rt}^2} \left| p_{rt} - \sum_{j=1}^M a_{0j} p_{r,t-j} \right|^2. \quad (4.35)$$

Por lo tanto, $\Gamma | \boldsymbol{\psi}_{r-1,N}$ tiene distribución Rayleigh y $P_{FA} | \boldsymbol{\psi}_{r-1,N}$ está dada por

$$P_{FA} | \boldsymbol{\psi}_{r-1,N} = \int_{\eta}^{\infty} f_{\Gamma}(\Gamma | \boldsymbol{\psi}_{r-1,N}; H_0) d\Gamma = e^{-\eta^2 / \sigma_s^2}. \quad (4.36)$$

Finalmente, a partir de (4.31) y (4.36) se obtiene el umbral λ

$$\lambda = -\ln(P_{FA}|\psi_{r-1,N}). \quad (4.37)$$

La probabilidad de detección, P_D , no puede evaluarse analíticamente y será calculada en las secciones 4.4.2 y 4.5 por medio de simulaciones Monte Carlo, utilizando datos de clutter sintéticos y reales, respectivamente.

En la deducción del algoritmo de decisión se consideró que los coeficientes del proceso eran conocidos. Sin embargo, esto no es cierto en una situación real. De este modo, en la práctica el detector AR-GARCH-2D está dado por

$$\frac{\left| \sum_{t=1}^N \frac{1}{\hat{\sigma}_{rt}^2} \left(p_{rt} - \sum_{j=1}^M \hat{a}_{0j} p_{r,t-j} \right) \left(y_{rt} - \sum_{j=1}^M \hat{a}_{0j} y_{r,t-j} \right) \right|^2}{\sum_{t=1}^N \frac{1}{\hat{\sigma}_{rt}^2} \left| p_{rt} - \sum_{j=1}^M \hat{a}_{0j} p_{r,t-j} \right|^2} \underset{H_0}{\overset{H_1}{\gtrless}} \lambda, \quad (4.38)$$

donde

$$\hat{\sigma}_{rt}^2 = \hat{k} + \sum_{j=0}^1 \hat{\alpha}_{1j} \hat{\sigma}_{r-1,t-j}^2 + \sum_{j=0}^1 \hat{\beta}_{1j} |y_{r-1,t-j}|^2, \quad (4.39)$$

y \hat{k} , $\hat{\alpha}_{1j}$, $\hat{\beta}_{1j}$, para $j = 0, 1$ y \hat{a}_{0j} , para $j = 1, \dots, M$ son los estimados de los coeficientes del proceso. Para tener en cuenta las limitaciones que pueda generar esta consideración, en la sección 4.4.3 se realizará un análisis de sensibilidad en el cual se estudia como afectan errores en los valores de los coeficientes sobre el desempeño del detector.

4.4. Simulaciones numéricas

En esta sección se lleva a cabo un análisis numérico del método de estimación de los coeficientes de un proceso AR-GARCH-2D y del desempeño de detección de su detector asociado. Para las simulaciones numéricas, el análisis se restringe a uno de los procesos de la familia representada por el modelo AR-GARCH-2D. Los datos sintéticos utilizados en esta sección son generados por medio del siguiente modelo

$$v_{rt} = (0.45 - j0.2)v_{r,t-1} + (0.1 + j0.3)v_{r,t-2} + \sigma_{rt}z_{rt}, \quad (4.40)$$

$$\sigma_{rt}^2 = 1 + 0.1\sigma_{r-1,t-1}^2 + 0.5|v_{r-1,t}|^2, \quad (4.41)$$

con $z_{rt} \sim \mathcal{CN}(0,1)$ e i.i.d. Los valores de los coeficientes se eligieron, dentro del conjunto de valores que satisfacen la restricción (4.10), de modo de que tanto la correlación como la heterocedasticidad presenten efectos significativos sobre el proceso.

4.4.1. Estimación

A fin de verificar que los estimados de los coeficientes del proceso satisfacen ciertas propiedades deseadas y también para contar con una idea del error de la estimación al momento de utilizar los datos reales, se realizó un análisis de la calidad de la estimación en función del

| Coef. | $R \times T$ | | | | |
|---------------|--------------------|--------------------|--------------------|--------------------|--------------------|
| | 68×2048 | 68×4096 | 136×2048 | 68×8192 | 136×4096 |
| k | 1.0043 | 1.0045 | 1.0009 | 1.0033 | 1.0008 |
| α_{11} | 0.0986 | 0.0985 | 0.1001 | 0.0984 | 0.0995 |
| β_{10} | 0.4997 | 0.4996 | 0.4994 | 0.4999 | 0.5000 |
| a_{01} | $0.4497 - j0.1994$ | $0.4497 - j0.1998$ | $0.4496 - j0.1997$ | $0.4499 - j0.1999$ | $0.4499 - j0.1998$ |
| a_{02} | $0.1000 + j0.2996$ | $0.1001 + j0.2998$ | $0.1001 + j0.2996$ | $0.1000 + j0.2999$ | $0.1000 + j0.2999$ |

Tabla 4.1: Media de los estimados de los coeficientes del proceso AR-GARCH-2D.

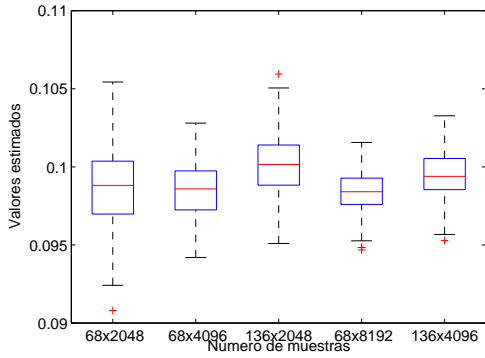
| Coef. | $R \times T$ | | | | |
|---------------|-----------------------|----------------------|----------------------|----------------------|----------------------|
| | 68×2048 | 68×4096 | 136×2048 | 68×8192 | 136×4096 |
| k | 10.7×10^{-3} | 8.7×10^{-3} | 8.2×10^{-3} | 5.7×10^{-3} | 5.9×10^{-3} |
| α_{11} | 2.5×10^{-3} | 1.7×10^{-3} | 2.0×10^{-3} | 1.4×10^{-3} | 1.4×10^{-3} |
| β_{10} | 2.7×10^{-3} | 2.1×10^{-3} | 1.8×10^{-3} | 1.2×10^{-3} | 1.3×10^{-3} |
| a_{01} | 2.0×10^{-3} | 1.5×10^{-3} | 1.4×10^{-3} | 1.1×10^{-3} | 1.1×10^{-3} |
| a_{02} | 2.1×10^{-3} | 1.5×10^{-3} | 1.4×10^{-3} | 1.0×10^{-3} | 1.0×10^{-3} |

Tabla 4.2: Desviación estándar de los estimados de los coeficientes del proceso AR-GARCH-2D.

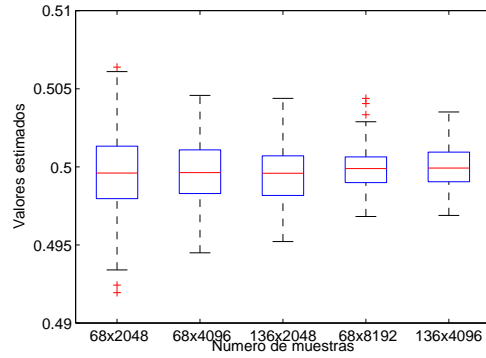
número de muestras de la matriz de observación \mathbf{V} . Este estudio se llevó a cabo por medio de simulaciones numéricas que consistieron en la generación de realizaciones independientes del proceso de diferente tamaño, $R \times T$, y la estimación de los coeficientes del modelo para cada estimación. El problema de estimación, descrito por (4.17), se resolvió utilizando el algoritmo de conjuntos activos de MATLAB®. Este procedimiento se repitió 200 veces para cada caso, es decir, se obtuvieron 200 valores de los estimados de los coeficientes para cada tamaño del conjunto de datos sintético. Las condiciones iniciales fueron elegidas aleatoriamente dentro del conjunto de valores de los coeficientes que satisfacen las restricciones (4.3) y (4.10). No se hicieron pruebas de concavidad del problema, pero el algoritmo convergió para todos los valores iniciales utilizados.

En las tablas 4.1 y 4.2 y en las figuras 4.3 y 4.2 se muestran los valores de la media y la desviación estándar de los estimados en función del tamaño del conjunto de datos, $R \times T$. La media presenta un pequeño sesgo que decrece a medida que se incrementa el número de muestras utilizadas en la estimación. Se puede observar que algunos coeficientes son más sensibles a una dimensión que a otras; por ejemplo el sesgo del estimador de α_{11} decrece rápidamente cuando se incrementa el número de muestras en rango, pero no presenta el mismo comportamiento con las muestras en tiempo lento. La desviación estándar de los estimados disminuye con el número de muestras, como se ve en la tabla 4.2. Notar que los resultados verifican el comportamiento asintótico mencionado con una tasa de convergencia cercana a la raíz cuadrada del número de muestras. En estos ejemplos, se observa que el comportamiento

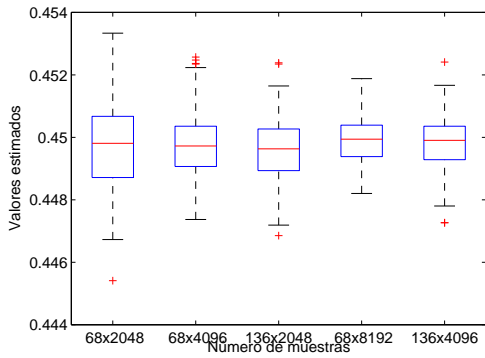
del QMLE se asemeja al de un estimador asintóticamente insesgado y consistente.



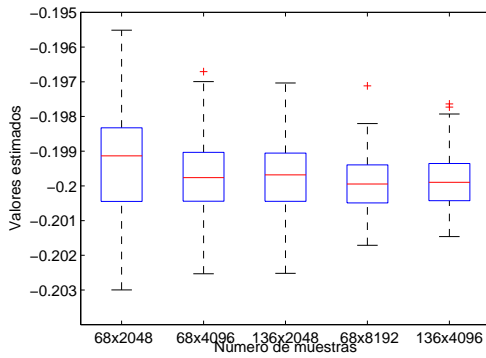
(a) Coeficiente α_{11} .



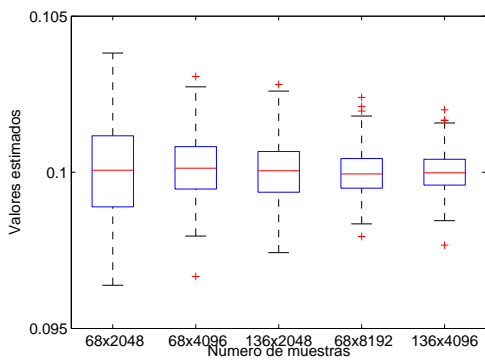
(b) Coeficiente β_{10} .



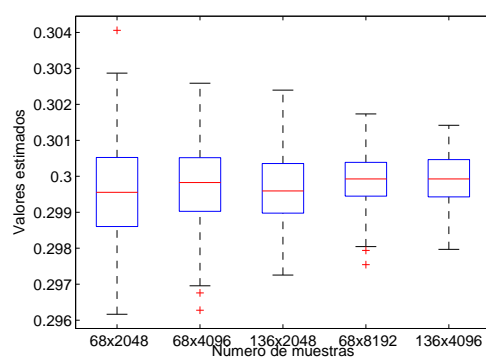
(c) Parte real del coeficiente a_{01} .



(d) Parte imaginaria del coeficiente a_{01} .



(e) Parte real del coeficiente a_{02} .



(f) Parte imaginaria del coeficiente a_{02} .

Figura 4.2: Estimados de los coeficientes del proceso AR-GARCH-2D simulado en función del número de muestras.

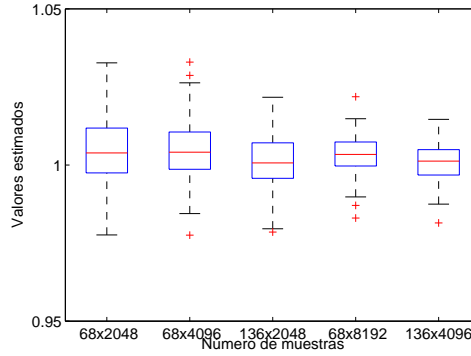


Figura 4.3: Estimados del coeficiente k , del proceso AR-GARCH-2D simulado, en función del número de muestras.

4.4.2. Probabilidad de detección

Con el fin de evaluar el desempeño del detector AR-GARCH-2D se realizaron simulaciones de Monte Carlo donde se estimó la probabilidad de detección, P_D , para diferentes valores de relación señal a clutter (SCR). Se generaron realizaciones del proceso AR-GARCH-2D dado por (4.40) y (4.41) y en una celda elegida aleatoriamente se agregó un objetivo sintético con el modelo propuesto en la sección 4.3. La señal sigue el modelo de objetivo Swerling I (Richards, 2005), es decir γ es una variable aleatoria compleja normal circularmente simétrica, con media nula y varianza $\mathbb{E}\{|\gamma|^2\} = \sigma_\gamma^2$, tomando un valor diferente para cada instante de decisión. Se definió $\text{SCR} = \sigma_\gamma^2 / \hat{\sigma}^2$, donde $\hat{\sigma}^2$ es la varianza muestral de los datos de clutter. La frecuencia Doppler del objetivo f_d se fijó de modo que $f_d T_s = 0.5$ y el umbral λ se eligió a partir de (4.37) para $P_{FA} = 10^{-2}$ y $P_{FA} = 10^{-3}$. Por cada valor de P_{FA} se repitió el procedimiento para $N = 4, 8$ y 16 pulsos. La P_D se estimó por medio de la frecuencia relativa del estadístico superando el umbral.

Se generaron datos sintéticos de tamaño 68×8192 (igual al tamaño de los conjuntos de datos reales que se utilizarán en la sección 4.5) y se utilizó un número de conjuntos de datos independientes lo suficientemente grande para asegurar que la probabilidad de detección sea estimada de manera precisa, con la regla descrita en la sección 3.5.1. Las simulaciones se repitieron para diferentes valores de SCR, manteniendo la potencia de clutter constante y variando el valor de σ_γ^2 . En la figura 4.4a se muestran los valores de P_D estimados que se obtuvieron como resultado de las simulaciones. Como era de esperarse para un dado valor de SCR la probabilidad de detección aumenta con P_{FA} y con N .

Finalmente, se evaluaron curvas ROC para el detector AR-GARCH-2D. Para un dado valor de SCR la ROC se obtuvo variando el umbral, calculando la probabilidad de falsa alarma utilizando (4.37) y la probabilidad de detección por medio de su frecuencia relativa. La figura 4.4b muestra las curvas ROC para $N = 4, 8$ y 16 pulsos con $\text{SCR} = 0$ dB.

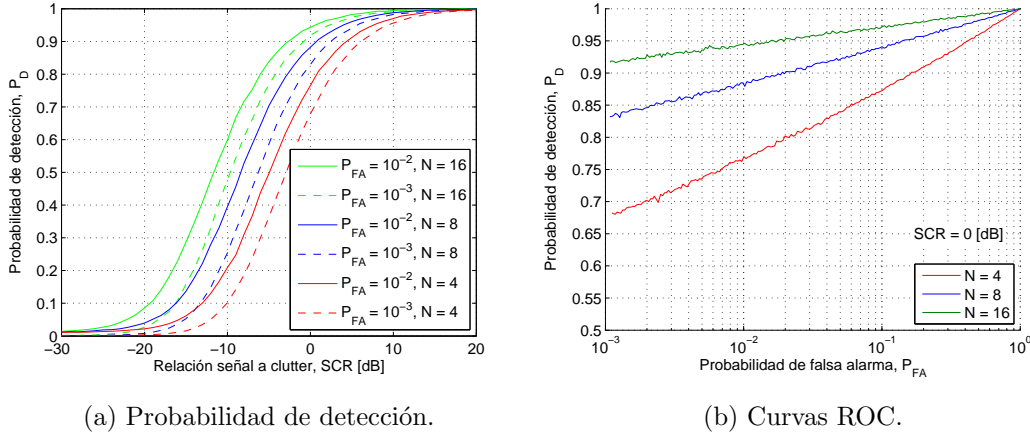


Figura 4.4: Desempeño del detector AR-GARCH-2D utilizando datos sintéticos.

4.4.3. Análisis de sensibilidad

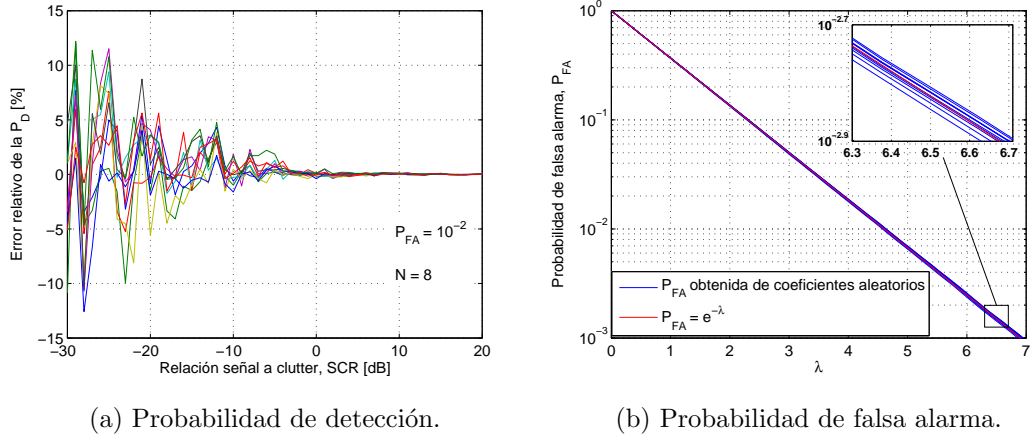
En la deducción del detector AR-GARCH-2D se supuso que los coeficientes del proceso eran conocidos. Sin embargo, en una situación real esto rara vez es cierto y deben ser estimados. Por este motivo, en esta sección se evalúa como afecta la varianza de los estimados de los coeficientes sobre el desempeño del detector. En presencia de un error de estimación o de pequeñas variaciones en los parámetros se espera que el desempeño del detector no varíe significativamente.

Se realizaron simulaciones generando los datos de clutter con los valores reales de los coeficientes y se evaluaron la P_D y la P_{FA} , considerando que en el detector los coeficientes fueron perturbados debido a errores en la estimación. Se generaron coeficientes aleatorios a partir de una distribución normal, cuya media y varianza corresponden a las de los estimados obtenidos en la sección 4.4.1 (tablas 4.1 y 4.2) para los conjuntos de datos sintéticos de tamaño 68×8192 .

El procedimiento para estimar la probabilidad de detección fue el mismo que se describió en la sección anterior, con $N = 8$ y λ se eligió para que $P_{FA} = 10^{-2}$. Las simulaciones se repitieron 10 veces, es decir, para diez conjuntos diferentes de coeficientes corruptos. En la figura 4.5a se muestra el error relativo de la P_D resultante para diferentes valores de SCR, donde la P_D obtenida en la sección 4.4.2 se tomó como el valor real. Se observa que el detector presenta un error bajo en P_D para un amplio rango de valores de SCR, y se vuelve despreciable para valores altos de SCR.

El enfoque para estimar la probabilidad de falsa alarma fue similar. En este caso el objetivo no se agregó a los datos de clutter. El estadístico se calculó con los coeficientes corruptos y se comparó con el umbral. La P_{FA} fue estimada por medio de su frecuencia relativa. Este procedimiento se repitió variando el umbral, λ , para obtener una curva de P_{FA} en función de λ .

Una característica deseable para un detector en aplicaciones de radar es que posea tasa de falsa alarma constante (CFAR), es decir que la dispersión de los parámetros desconocidos no afecte su probabilidad de falsa alarma. Entonces, en presencia de errores de estimación o



(a) Probabilidad de detección.

(b) Probabilidad de falsa alarma.

Figura 4.5: Efecto del error en la estimación de los coeficientes sobre el desempeño del detector AR-GARCH-2D.

de pequeñas variaciones en sus parámetros, el desempeño del detector no debe variar. En la figura 4.5b se presentan 10 curvas de P_{FA} obtenidas a partir de coeficientes aleatorios y la P_{FA} condicional teórica, dada por (4.37), en función de λ . Se puede observar que a pesar de que el detector AR-GARCH-2D no es CFAR con respecto a los coeficientes del proceso, es muy robusto en una situación práctica, es decir P_{FA} no cambia significativamente cuando los valores de los coeficientes varían.

4.4.4. Comparación de desempeño

En esta sección se consideran tres detectores adaptivos y se comparan sus desempeños con el del detector AR-GARCH-2D dado por (4.40) y (4.41), utilizando los datos sintéticos de clutter. Estos son el test de relación de verosimilitud generalizada (GLRT) (Kelly, 1986), el detector adaptivo lineal-cuadrático (ALQ) (Gini y Greco, 1999) y el detector de relación de verosimilitud autorregresiva generalizada (ARGLR) (Sheikhi *et al.*, 1998). Estos detectores surgen del test de hipótesis (4.18), con el mismo modelo de señal \mathbf{s}^r presentado en la sección 4.3 pero con diferentes modelos de clutter \mathbf{v}^r .

Para estimar la matriz de covarianza del clutter los detectores GLRT y ALQ utilizan K vectores $\{\mathbf{y}_n^r\}_{n=1}^K$ de datos secundarios correspondientes a K celdas de rango adyacentes a la celda de rango bajo prueba. Se supone que los datos secundarios comparten las mismas propiedades estadísticas con los datos \mathbf{y}^r de la celda de rango bajo prueba. El GLRT puede escribirse como (Kelly, 1986)

$$\frac{|(\mathbf{p}^r)^H \hat{\mathbf{M}}_g^{-1} \mathbf{y}^r|^2}{\left[1 + \frac{1}{K} (\mathbf{y}^r)^H \hat{\mathbf{M}}_g^{-1} \mathbf{y}^r\right] \left[(\mathbf{p}^r)^H \hat{\mathbf{M}}_g^{-1} \mathbf{p}^r\right]} \underset{H_0}{\overset{H_1}{\geq}} \lambda, \quad (4.42)$$

donde $\hat{\mathbf{M}}_g$ es el estimado muestral de la matriz de covarianza del clutter

$$\hat{\mathbf{M}}_g = \frac{1}{K} \sum_{n=1}^K \mathbf{y}_n^r (\mathbf{y}_n^r)^H. \quad (4.43)$$

Por su parte, el detector ALQ está dado por (Gini y Greco, 1999)

$$\frac{|(\mathbf{p}^r)^H \hat{\mathbf{M}}_a^{-1} \mathbf{y}^r|^2}{\left[(\mathbf{y}^r)^H \hat{\mathbf{M}}_a^{-1} \mathbf{y}^r \right] \left[(\mathbf{p}^r)^H \hat{\mathbf{M}}_a^{-1} \mathbf{p}^r \right]} \underset{H_0}{\overset{H_1}{\geq}} 1 - e^{\lambda/N}, \quad (4.44)$$

donde $\hat{\mathbf{M}}_a$ es el estimado muestral normalizado de la matriz de covarianza del clutter (NSCME por sus siglas en inglés) (Gini *et al.*, 1995)

$$\hat{\mathbf{M}}_a = \frac{N}{K} \sum_{n=1}^K \frac{\mathbf{y}_n^r (\mathbf{y}_n^r)^H}{(\mathbf{y}_n^r)^H \mathbf{y}_n^r}. \quad (4.45)$$

El ARGLR es un detector que resulta de modelar el clutter como un proceso AR. Este detector se ajusta por si mismo al entorno utilizando únicamente un conjunto de datos primarios y está dado por (Sheikhi *et al.*, 1998)

$$\frac{(\mathbf{u} - \mathbf{Y} \hat{\mathbf{a}}_0)^H (\mathbf{u} - \mathbf{Y} \hat{\mathbf{a}}_0)}{(\mathbf{u} - \mathbf{Y} \hat{\mathbf{a}}_1)^H \mathbf{H}^H \mathbf{H} (\mathbf{u} - \mathbf{Y} \hat{\mathbf{a}}_1)} \underset{H_0}{\overset{H_1}{\geq}} e^{\lambda/N}, \quad (4.46)$$

donde los vectores $\hat{\mathbf{a}}_0$ y $\hat{\mathbf{a}}_1$ son $(\mathbf{Y}^H \mathbf{Y})^{-1} (\mathbf{Y}^H \mathbf{u})$ y $(\mathbf{Y}^H \mathbf{H}^H \mathbf{H} \mathbf{Y})^{-1} (\mathbf{Y}^H \mathbf{H}^H \mathbf{H} \mathbf{u})$ respectivamente, con

$$\mathbf{Y} = \begin{bmatrix} y_{rm} & \cdots & y_{r1} \\ \vdots & \ddots & \vdots \\ y_{r,N-1} & \cdots & y_{r,N-m} \end{bmatrix}, \quad (4.47)$$

$\mathbf{u} = [y_{r,m+1} \cdots y_{rN}]^T$, \mathbf{H} es la matriz de proyección sobre el espacio nulo de $\Upsilon = [1 \ e^{j\Omega} \ \dots \ e^{j(N-m-1)\Omega}]^T$, y m es el orden del proceso AR que modela el clutter.

En la figura 4.6 se muestran las curvas de la probabilidad de falsa alarma, P_{FA} , en función de λ para los cuatro detectores con $N = 8$ pulsos. Estas curvas fueron obtenidas usando 10 conjuntos de datos sintéticos de 100×8000 , generados a través de (4.40) y (4.41). A partir de estos resultados se fijó λ con $P_{FA} = 10^{-2}$ y $P_{FA} = 10^{-3}$ para cada detector y se calcularon las curvas de P_D en función de la SCR, con el modelo de señal descrito en la sección 4.4.2.

En la figura 4.7 se presentan las curvas de P_D en función de SCR con el umbral determinado con $P_{FA} = 10^{-2}$ y $P_{FA} = 10^{-3}$, respectivamente. Para los detectores GLRT y ALQ se utilizaron $K = 16$ vectores de datos secundarios y para el detector ARGLR se consideró un modelo AR de primer orden para el clutter, es decir $m = 1$, debido a que su desempeño se degrada cuando el orden aumenta, como consecuencia de que sus coeficientes son pobremente estimados dado el pequeño número de muestras de los datos primarios.

El desempeño del detector basado en el modelo AR-GARCH-2D es el mejor entre todas las alternativas analizadas. Este comportamiento era el esperado, debido a que los datos sintéticos utilizados fueron generados empleando justamente el modelo AR-GARCH-2D. Los demás detectores presentan un desempeño similar entre ellos para $P_{FA} = 10^{-2}$ e inferior para el ARGLR respecto del GLRT y el ALQ cuando $P_{FA} = 10^{-3}$. Esto se debe a que el detector ARGLR estima los parámetros empleando solo el conjunto de datos primario, cuyo tamaño resulta insuficiente cuando el umbral se fija con un valor de probabilidad de falsa alarma pequeño produciendo el deterioro en su desempeño. Este tipo de experimento se repitió para un

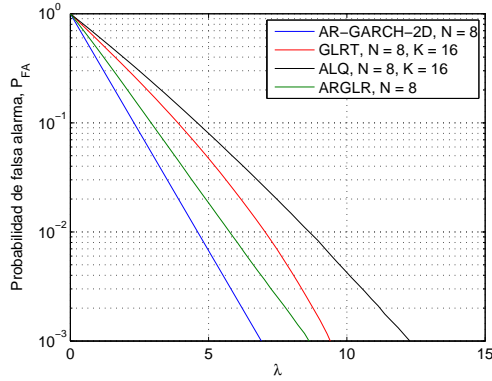


Figura 4.6: Probabilidad de falsa alarma en función del umbral de los detectores, utilizando datos de clutter sintéticos.

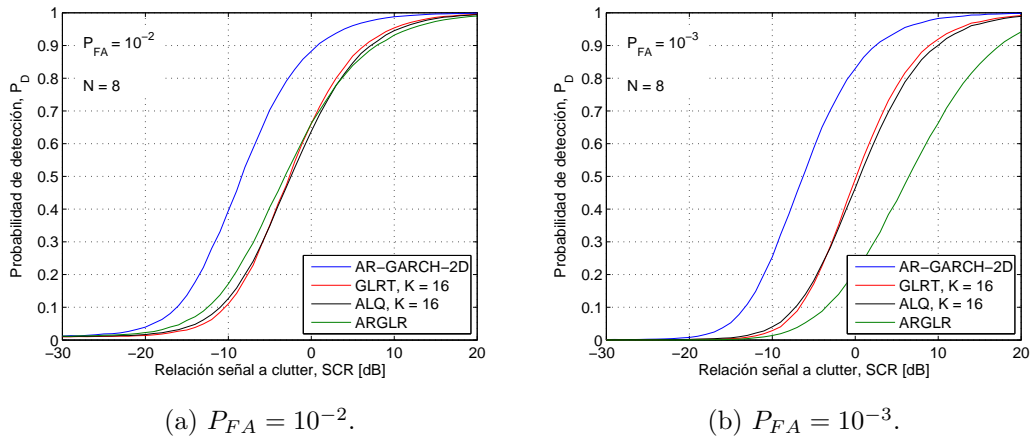


Figura 4.7: Probabilidad de detección de los detectores en función de la SCR.

número de pulsos, N , mayor y se obtuvieron comportamientos similares, con una mejora en el detector ARGLR, inclusive en algunos casos superando al GLRT y al ALQ, a causa de que el conjunto de datos primarios era de mayor tamaño.

4.5. Análisis de desempeño con mediciones reales

En esta sección se evalúa el desempeño del detector AR-GARCH-2D en una situación práctica utilizando datos reales de clutter. En esta oportunidad se utilizaron los conjuntos de datos *stare6*, *stare7* y *stare8* del radar IPIX, cuyas características se describen en la sección 2.2.2. Los tres conjuntos de datos se componen de $R = 68$ celdas de rango, para $T = 8192$ pulsos transmitidos y no existe objetivo presente. Los resultados que se presentan corresponden a la polarización vertical (VV).

4.5.1. Ajuste del modelo a los datos de clutter

En el modelado del clutter como un proceso AR-GARCH-2D, es de interés ajustar el proceso a las muestras de las mediciones de clutter. Como en el caso de los modelos GARCH estudiados en el capítulo 3, el ajuste incluye la estimación de los coeficientes del proceso como la elección de los órdenes del proceso, es decir el número óptimo de coeficientes. Para llevar a cabo el ajuste se utiliza el conjunto de datos *stare6*; los conjuntos de datos restantes serán utilizados para evaluar el desempeño del algoritmo de detección. La estimación de los coeficientes se realizó utilizando la metodología descrita en la sección 4.4.1 en las simulaciones numéricas. Debido a que los órdenes del proceso, es decir el número de coeficientes, no son conocidos de antemano, la estimación de los coeficientes se repitió para diferentes órdenes. Para cada valor de M entre 0 y 3 se consideraron todas las posibles combinaciones de los coeficientes de la varianza condicional, suponiendo que estos pueden ser o no iguales a cero. Teniendo en cuenta que para un modelo GARCH los coeficientes β 's no pueden ser simultáneamente cero, entonces los modelos posibles son 48. Sin embargo, cuando se realizó el ajuste, para algunos órdenes del modelo los estimados correspondientes a ciertos coeficientes tomaron los mismos valores que en el caso de otros ajustes al proceso AR-GARCH-2D, y los estimados de los nuevos coeficientes no resultaron estadísticamente significativos. Eliminando estos casos, el número de modelos sobrevivientes se reduce a 27; 8 corresponden a modelos para los cuales $M = 1, 2, 3$ y los 3 restantes corresponden a $M = 0$.

Una comparación justa entre los modelos propuestos no debería hacerse solo en base a la calidad del ajuste, sino que también debe hacerse teniendo en cuenta el desempeño en el problema de detección. Por este motivo, antes de adoptar un criterio de selección de modelos para determinar el proceso que presenta el compromiso óptimo entre complejidad y ajuste a las mediciones de clutter, se llevó a cabo un análisis de desempeño del problema de detección para los 27 modelos AR-GARCH-2D resultantes del ajuste.

Para evaluar el desempeño de los detectores dados por (4.38) se utilizaron los conjuntos de datos *stare7* y *stare8* que no fueron utilizados en el procedimiento de estimación, con $N = 8$ pulsos. Se calculó la P_{FA} y la P_D de la misma manera que se describió en la sección 4.4 pero usando los datos de clutter reales en lugar del clutter sintético.

En la figura 4.8 se muestran las curvas de la probabilidad de falsa alarma empírica, P_{FA} , para los diferentes detectores AR-GARCH-2D, correspondientes a los 27 modelos obtenidos del ajuste, y la P_{FA} condicional teórica, dada por (4.37), en función del umbral λ . La probabilidad de falsa alarma empírica se define como el número de decisiones que exceden el umbral dividido por el número total de decisiones $2RT/N$. Se utiliza el valor de M para agrupar los modelos debido a que, como se observa en la figura, los procesos con igual orden en la parte autorregresiva presentan un comportamiento similar. Notar que los procesos para los cuales $M = 3$ aproximan mejor la probabilidad de falsa alarma que los otros procesos.

Con el objetivo de evaluar la probabilidad de detección P_D se sumó un objetivo sintético a los datos reales de clutter marítimo, con el modelo descrito en la sección 4.4.2. Nuevamente,

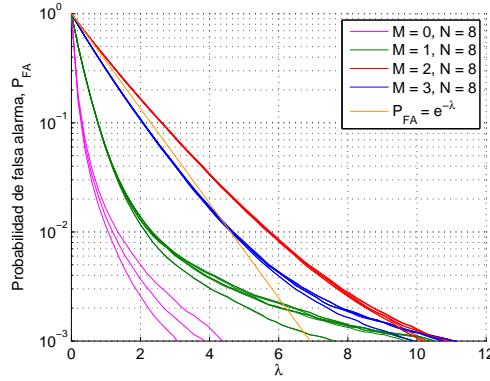


Figura 4.8: Probabilidad de falsa alarma en función del umbral de los detectores AR-GARCH-2D, utilizando datos de clutter del radar IPIX.

se define $SCR = \sigma_\gamma^2 / \hat{\sigma}^2$, donde ahora $\hat{\sigma}^2$ es la varianza muestral de los datos reales de clutter. Para cada detector el umbral λ se eligió para $P_{FA} = 10^{-2}$, en base a los resultados de la figura 4.8. En la figura 4.9a se muestran las curvas de las P_D empíricas para los diferentes detectores AR-GARCH-2D en función de la SCR. La tasa de detección empírica se definió como el número de detecciones en la celda de rango donde se ubica el objetivo dividido por el número total de decisiones en esta celda de rango. Esto se repitió para 16 celdas de rango diferentes tomadas aleatoriamente de cualquiera de los dos conjuntos de datos utilizados, dando un total de $16T/N$ decisiones. Los resultados muestran que el desempeño de los procesos con órdenes $M = 2$ y $M = 3$ es similar y es mejor que el desempeño de los modelos restantes.

Finalmente, se calcularon las curvas ROC para los detectores AR-GARCH-2D. En este caso las ROC se obtuvieron utilizando los datos reales de clutter, variando el umbral y evaluando las tasas de detección y falsa alarma empíricas. En la figura 4.9b se presentan las curvas ROC resultantes con los datos de clutter reales y un objetivo sintético con $SCR = -15$ dB. Nuevamente se observan mejores resultados para los modelos con la parte autorregresiva de segundo y tercer orden.

Con el propósito de evaluar cual de los modelos AR-GARCH-2D presenta un mejor compromiso entre complejidad y calidad del ajuste se utilizó el criterio de información Bayesiano (BIC) (Konishi y Kitagawa, 2008). Ignorando un término constante que comparten todos los modelos, a partir (4.16) el estadístico BIC puede escribirse como

$$BIC = 2\mathcal{L}_{4i}(\hat{\theta}_i) + \nu_i \ln(2RT), \quad (4.48)$$

donde $\hat{\theta}_i$ es el estimador de máxima verosimilitud del vector de parámetros ν_i -dimensional θ_i del i -ésimo modelo, para $i = 1, \dots, 27$, y RT es el número de observaciones complejas. El criterio elige el modelo que posee el mínimo BIC. A diferencia de los modelos del capítulo 3 aquí cobra sentido utilizar el BIC debido que se busca seleccionar uno entre 27 procesos posibles, y los estimados de los coeficientes en este caso no son un máximo restringido de la función de verosimilitud.

Para los datos de clutter del radar IPIX, el modelo AR-GARCH-2D con BIC mínimo es uno

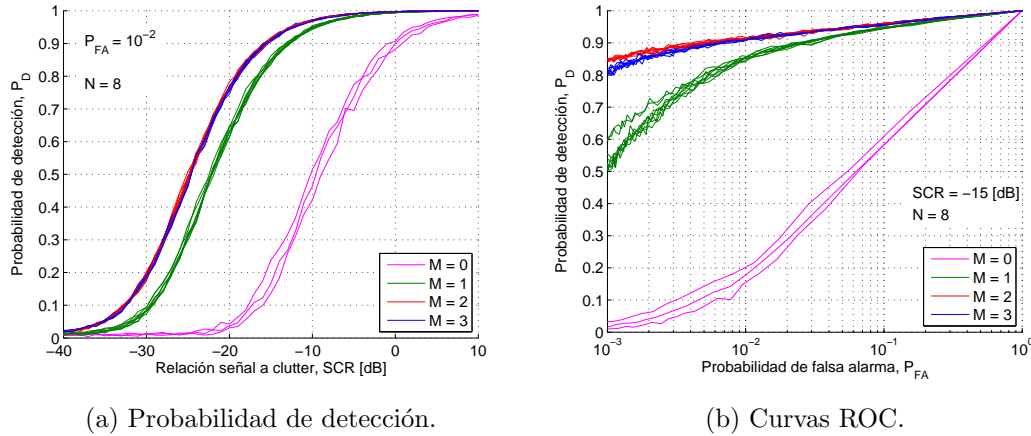


Figura 4.9: Desempeño de los detectores AR-GARCH-2D, utilizando datos de clutter del radar IPIX y objetivo sintético.

de los modelos que mejor desempeño mostró, es decir, con $M = 3$, cuyos coeficientes están dados por

$$v_{rt} = (1.049 - j1.079)v_{r,t-1} + (0.097 + j0.441)v_{r,t-2} + (0.063 + j0.151)v_{r,t-3} + \hat{\sigma}_{rt}z_{rt}, \quad (4.49)$$

$$\hat{\sigma}_{rt}^2 = 1.244 \times 10^{-4} + 0.022\hat{\sigma}_{r-1,t-1}^2 + 0.012|v_{r-1,t}|^2, \quad (4.50)$$

con $z_{rt} \sim \mathcal{CN}(0, 1)$ e i.i.d.

4.5.2. Comparación de desempeño

En esta sección se consideran los detectores adaptivos GLRT, ALQ y ARGLR descritos en la sección 4.4.4, y se comparan sus desempeños con el del detector AR-GARCH-2D seleccionado, en esta oportunidad utilizando las mediciones reales de clutter.

En la figura 4.10 se muestran las curvas de la probabilidad de falsa alarma, P_{FA} , en función de λ para los cuatro detectores con $N = 8$ pulsos. Estas curvas fueron obtenidas usando los conjuntos de datos *stare7* y *stare8* del radar IPIX. A partir de estos resultados se fijó λ con $P_{FA} = 10^{-2}$ para cada detector y se calcularon las curvas de P_D en función de la SCR, con el modelo de señal descrito en la sección 4.4.2. También se evaluó la ROC empírica para una $SCR = -15$ dB.

En las figuras 4.11a y 4.11b se presentan las curvas de P_D en función de SCR y las curvas ROC, respectivamente. Nuevamente, para los detectores GLRT y ALQ se utilizaron $K = 16$ vectores de datos secundarios y para el detector ARGLR se consideró un modelo AR de primer orden para el clutter, es decir $m = 1$. El desempeño del detector basado en el modelo AR-GARCH-2D es el mejor entre todas las alternativas analizadas, sugiriendo que el modelo AR-GARCH-2D es el que mejor representa los datos de clutter marítimo frente a los otros modelos examinados. Notar que los detectores ARGLR y ALQ muestran aproximadamente el mismo desempeño y el GLRT supera al detector ALQ, como había sido mostrado por Gini *et al.* (2000).

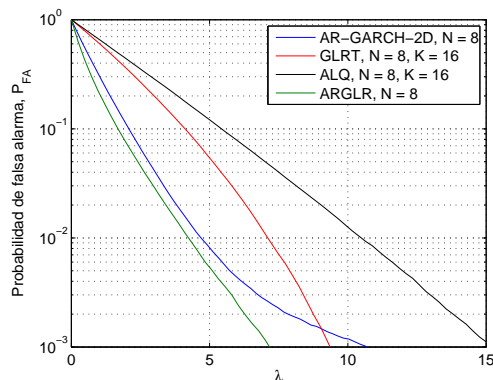
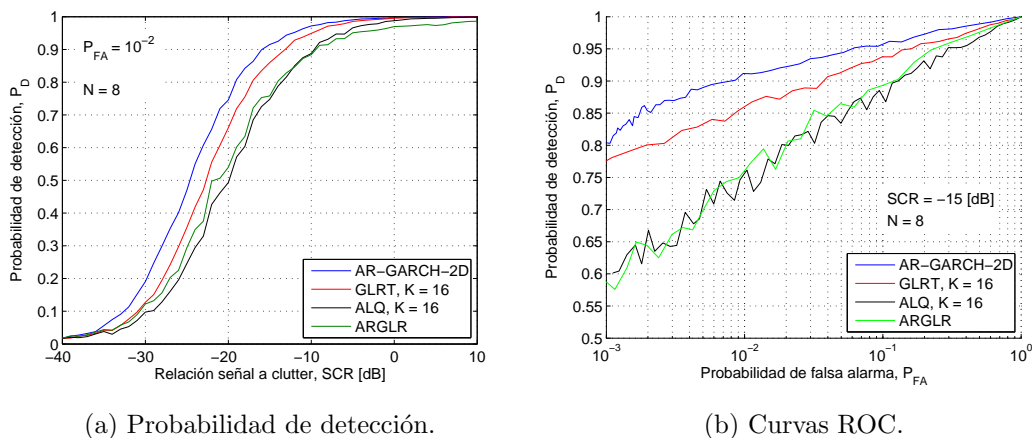


Figura 4.10: Probabilidad de falsa alarma en función del umbral, utilizando datos de clutter del radar IPIX.



(a) Probabilidad de detección.

(b) Curvas ROC.

Figura 4.11: Comparación de desempeño entre los detectores, utilizando datos de clutter del radar IPIX y objetivo sintético.

4.6. Discusión

Hasta aquí se analizó la calidad del ajuste de los datos al proceso AR-GARCH-2D solo a través del desempeño del detector resultante para este modelo de clutter. En esta sección se estudia cuán representativo es el proceso AR-GARCH-2D obtenido de la parámetros estadísticos de los datos.

En la figura 4.12 se muestran los histogramas de la magnitud de los datos correspondientes al conjunto de datos *stare6* y de la magnitud de muestras generadas a partir del proceso AR-GARCH-2D seleccionado por medio del BIC en la sección 4.5.1. En estos gráficos se aprecia claramente que la representación del histograma del modelo se aparta considerablemente del de las mediciones. Este comportamiento se debe a que los estimados de los coeficientes α_{01} y β_{10} de la varianza condicional del proceso (4.50) toman valores pequeños. Por este motivo, para los valores de las mediciones el valor instantáneo de la varianza se aparta poco del valor de k , dando como resultado un proceso poco impulsivo cuyo histograma no modela correctamente las colas pesadas de la densidad de los datos.

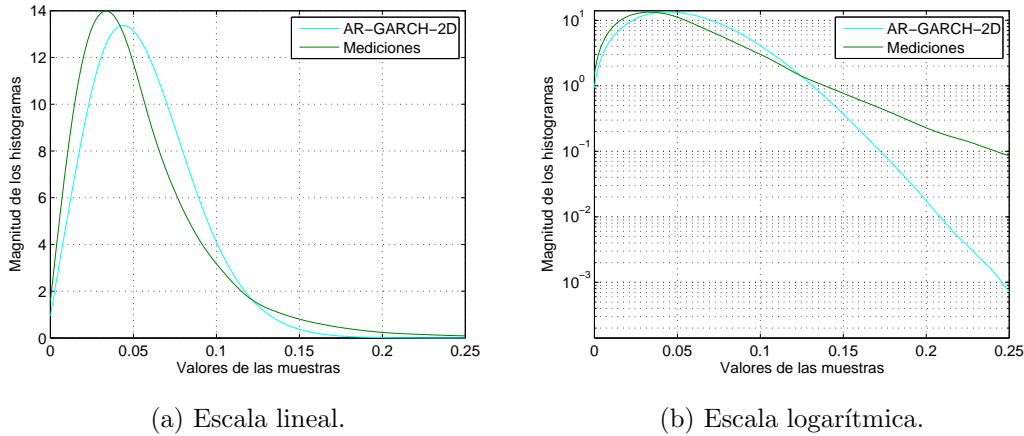


Figura 4.12: Histogramas de la magnitud del clutter del conjunto de datos *stare6* para la polarización VV y de la magnitud del proceso AR-GARCH-2D.

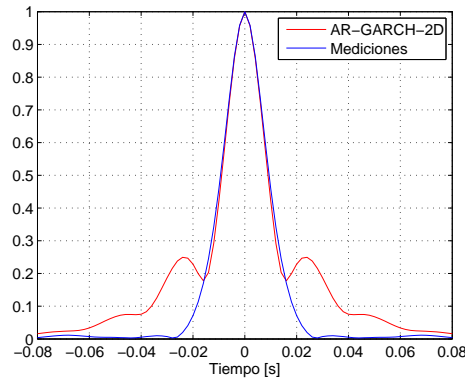


Figura 4.13: Correlaciones estimadas correspondientes al conjunto de datos *stare6* para la polarización VV y al proceso AR-GARCH-2D.

Sin embargo, los detectores AR-GARCH-2D y GLRT mostraron un desempeño superior respecto del ALQ. La causa de este comportamiento se encuentra en la modelización de la correlación que presentan los datos. En la figura 4.13 se muestra la estimación de la correlación del conjunto de datos *stare6* y la correlación que presenta el modelo AR-GARCH-2D. Se puede observar que para pulsos cercanos la correlación del modelo aproxima bien a la de los datos.

En la figura 4.13 también se observa que los datos presentan una fuerte correlación entre pulsos próximos para una dada celda de rango. Por este motivo el procedimiento de ajuste al proceso AR-GARCH-2D arroja como resultado coeficientes de la parte autorregresiva con valores importantes, limitando el valor que pueden tomar los estimados de los coeficientes de la varianza condicional a causa de la restricción impuesta para asegurar la estabilidad del proceso (4.10).

Finalmente, debido a que los valores de los coeficientes α_{01} y β_{10} son pequeños, se realizó un análisis comparativo entre el modelo AR-GARCH-2D y un modelo Gaussiano, autorregresivo en la dimensión tiempo lento de orden tres, es decir con tres coeficientes. El modelo se describe

en el apéndice C, donde también se deducen las expresiones para estimar sus parámetros. En primer lugar se realizó el ajuste del modelo al conjunto de datos *stare6* dando como resultado valores para los coeficientes de la parte autorregresiva del orden de los obtenidos para el proceso AR-GARCH-2D y un valor para la varianza próximo al valor de k . En segundo lugar se estudió el desempeño de su detector asociado, que es de la misma forma de (4.38) pero con el valor de la varianza constante e igual al valor estimado. Como era de esperar en base a lo mencionado a lo largo de esta sección, el desempeño de este detector resultó similar al del AR-GARCH-2D. Sin embargo, la evaluación del BIC para este modelo da un valor mayor no solo que el del modelo AR-GARCH-2D seleccionado, sino también que el de aquellos modelos AR-GARCH-2D con orden $M = 2$. En este punto también se presenta una contradicción entre calidad de ajuste y desempeño en la detección. Se puede pensar que ante desempeños comparables lo conveniente para estos conjuntos de datos sería emplear el modelo autorregresivo, debido a la clara ventaja que presenta frente al modelo AR-GARCH-2D en el número de operaciones, tanto en el algoritmo de estimación y como en el de detección.

Es importante destacar que los resultados obtenidos con los datos reales no van en detrimento del modelo propuesto. Por un lado, porque la evaluación del modelo no debe hacerse solo a partir de la calidad de un ajuste a mediciones, que como en esta oportunidad esto puede resultar engañoso. El mismo debe hacerse considerando además el desempeño del detector asociado al modelo. Por otra parte, los resultados reafirman las ventajas mencionadas cuando se presentaron los procesos AR-GARCH-2D. Este tipo de modelo poseen la versatilidad de modelar tanto la correlación de los datos como su impulsividad. Para el conjunto de mediciones del que se dispone, la correlación pulso a pulso es el fenómeno que domina su representación en lo que hace al desempeño de los detectores y el proceso no se ve limitado para modelarlo. Sin embargo, lo que hace atractivo a estos procesos es que en aquellas circunstancias en las que la impulsividad sea tal que controle el comportamiento de los datos, el mismo posee la estructura necesaria para modelarla, cosa que no ocurriría si por ejemplo se adoptara el uso del detector asociado al modelo autorregresivo.

4.7. Conclusiones

En este capítulo se propuso el uso de un proceso GARCH-2D autorregresivo para una mejor caracterización de datos de clutter marítimo. Este modelo preserva la fdp de colas pesadas de los procesos GARCH y permite considerar la correlación pulso a pulso.

En base al método de estimación comúnmente utilizado para estimar los parámetros de los procesos GARCH, se derivó un algoritmo de estimación para los coeficientes de los procesos AR-GARCH-2D y se verificaron las propiedades asintóticas por medio de simulaciones numéricas.

También se presentó un algoritmo de detección basado en este modelo de clutter. Una de las ventajas del detector AR-GARCH-2D es que es un algoritmo rápidamente adaptivo dado que la detección depende no solo de la vecindad de la dimensión tiempo lento de la celda de rango

bajo prueba, sino también en la vecindad de la dimensión tiempo rápido. Por este motivo, es capaz de captar las variaciones rápidas y el comportamiento heterogéneo del clutter.

Se derivó una expresión para la probabilidad de falsa alarma del detector AR-GARCH-2D. Dada su complejidad matemática, la probabilidad de detección se evaluó recurriendo a simulaciones numéricas. Además, se analizó cómo los errores en la estimación de los coeficientes repercuten en el desempeño del detector. En este aspecto el detector mostró un comportamiento robusto en situaciones prácticas.

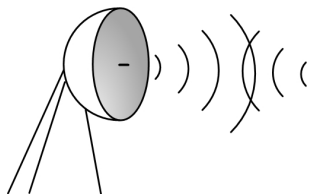
Se evaluaron las probabilidades de detección y falsa alarma por medio de simulaciones de Monte Carlo utilizando datos marítimos reales e inyectando un objetivo sintético, para el detector AR-GARCH-2D con diferentes órdenes de proceso, cuyos coeficientes resultaron del ajuste a dichos datos. Se mostró que los detectores para los cuales el orden de la parte autorregresiva, M , es 2 o 3 presentan un mejor desempeño que los detectores para los cuales M es 0 o 1, es decir una mayor probabilidad de detección con una tasa de falsa alarma menor. Luego, se utilizó un criterio de selección de modelos para elegir aquel que presente el compromiso óptimo entre complejidad y calidad del ajuste.

Además, se comparó el desempeño del detector AR-GARCH-2D elegido con el de los detectores GLRT, ALQ y ARGLR utilizando datos reales de clutter. Se mostró que el detector propuesto supera a los otros tres detectores. El mismo comportamiento se observó para otros conjuntos de datos y para diferentes números de pulsos N involucrados en la decisión. El GLRT es el segundo mejor detector, superando al detector ALQ diseñado suponiendo un modelo Gaussiano-compuesto para el clutter. En el caso del detector ARGLR, su desempeño se degrada drásticamente cuando el número de pulsos integrados decrece o el orden del modelo aumenta a causa de que este detector se ajusta al entorno utilizando solo un conjunto de datos primario.

El modelo AR-GARCH-2D que se propuso para el clutter y se estudió con datos de radar marítimo, debería desempeñarse mejor en aquellos entornos donde el fenómeno que domina la representación estadística de los datos es la impulsividad del clutter, es decir que la función de densidad de probabilidad del clutter posea colas pesadas. Cuando esto no ocurre, los coeficientes α 's y β 's de la parte GARCH del modelo no serán estadísticamente significativos, y el detector presentará un comportamiento similar al de un detector basado en un modelo de clutter Gaussiano-autorregresivo con datos secundarios. Por otro lado, si el entorno es tal que no existe correlación entre pulsos consecutivos, la parte AR del modelo será despreciable y el detector funcionará como un detector GARCH-2D. En el caso de un entorno de clutter no-impulsivo y no-correlacionado el detector se reduce al caso de un detector basado en un modelo de clutter blanco y Gaussiano, aunque con un conjunto de datos secundario de dimensiones excesivas para estimar su varianza.

Para una aplicación en tiempo real, el algoritmo de detección posee una carga computacional práctica. Existen dos tipos diferentes de cálculos que deben realizarse: la estimación de los coeficientes y la evaluación del estadístico en la detección. El último no involucra un costo computacional alto y a diferencia de los demás detectores analizados no requiere resolver

ningún sistema de ecuaciones lineales. En cambio, la estimación de los coeficientes requiere una mayor carga computacional. Sin embargo, esta puede realizarse como un proceso secundario (*batch process*) y como en el caso de detector GARCH-1D del capítulo anterior, los valores de los coeficientes no requieren ser actualizados muy frecuentemente. De hecho, en los resultados presentados, la estimación y la detección se realizaron con diferentes conjuntos de datos, con 2 y 5 minutos de retraso. La estimación de los coeficientes con los conjuntos de datos *stare7* o *stare8* arroja valores similares a los mostrados a lo largo del capítulo, sugiriendo que solo frente a un cambio importante en las condiciones del mar es necesaria una actualización de los mismos.



Estimación secuencial de parámetros de procesos GARCH

5.1. Introducción

En los capítulos 3 y 4 se presentaron modelos para el clutter en aplicaciones de radar, basados en procesos de tipo GARCH. La característica de estos procesos es que la varianza en el instante actual depende de su historia (Bollerslev, 1986). A partir de estos modelos para el clutter se obtuvieron esquemas de detección cuyos umbrales dependen de la varianza condicional de la celda de rango bajo prueba (Pascual *et al.*, 2013b, 2014c). En los enfoques propuestos en una primera etapa se estiman los coeficientes del proceso y luego, en la etapa de detección, los estimados resultantes son utilizados para la actualización del valor de la varianza condicional en la fórmula del umbral.

Siguiendo el enfoque de estimación Bayesiano, en este capítulo se presenta un estimador lineal de mínimo error cuadrático medio (LMMSE por sus siglas en inglés) para la varianza condicional. La idea que se persigue es hacerse de un mecanismo que permita actualizar su valor de manera secuencial, en cada instante de decisión. Como se vió, la estimación de los coeficientes empleando el QMLE es un problema de optimización no lineal computacionalmente intenso. Con el estimador LMMSE se pretende reemplazar este proceso, o bien evitar que se requiera su utilización a menudo.

La deducción del algoritmo es análoga a la del filtro de Kalman (Simon, 2006), pero en este problema las matrices del sistema son aleatorias. En la sección 5.2 se deriva el estimador

LMMSE para la varianza condicional considerando que los valores de los coeficientes son conocidos. En este planteo el modelo de estados resultante es un caso particular de un sistema lineal de tiempo discreto con parámetros aleatorios. El problema de estimación para este tipo de modelos de estados ha sido tratado en la literatura (Koning, 1984; Luo *et al.*, 2012). Sin embargo, en este trabajo la ecuación de estados cuenta con un término que no considera Koning (1984), el cual se puede pensar como una acción de control haciendo un paralelismo con un sistema de control representado en variables de estado. Por otra parte, este término no es independiente de la matriz de estados aleatoria.

Considerar que los coeficientes son conocidos en la deducción del estimador LMMSE no es una hipótesis cierta desde un punto de vista práctico. Sin embargo, la utilidad de este estimador radica, por un lado en la posibilidad de utilizarlo para actualizar la varianza condicional en cada instante de decisión una vez que los coeficientes fueron estimados con el QMLE, permitiendo de este modo corregir errores debido a inexactitudes en sus valores. Por otra parte, en la sección 5.3 se considera el caso general en el que se supone que los coeficientes son desconocidos. Este contexto es el de la estimación conjunta de estados y parámetros (Ljung, 1999). Como en el filtro de Kalman, los parámetros son agregados al vector de estados, lo que convierte el sistema en no lineal. Se deducen entonces las expresiones de lo que se denomina el estimador LMMSE extendido que permite la estimación conjunta de la varianza condicional y sus coeficientes. El procedimiento usado en este punto es similar al del filtro de Kalman extendido (Simon, 2006) y es aquí donde también son de utilidad las fórmulas que conforman el estimador LMMSE.

Finalmente, ambas versiones de los estimadores LMMSE se implementan en MATLAB® y se verifica su desempeño por medio de simulaciones numéricas empleando datos sintéticos de los procesos. Es importante destacar que si bien dichos estimadores muestran el comportamiento deseado en ningún caso se realizaron análisis de convergencia.

5.2. Estimador LMMSE

5.2.1. Formulación del problema

Se considera un proceso GARCH(p, q) como se definió en el capítulo 3 y que se transcribe a continuación

$$v_r = \sigma_r z_r, \tag{5.1}$$

$$\sigma_r^2 = k + \sum_{j=1}^p \alpha_j \sigma_{r-j}^2 + \sum_{j=1}^q \beta_j v_{r-j}^2, \tag{5.2}$$

donde z_r es un proceso aleatorio con momentos no centrales $m_n = \mathbb{E}\{|z_r|^n\}$ y k , α_j y β_j son los coeficientes de la varianza condicional σ_r^2 . A lo largo de la tesis el interés se centró en las distribuciones normal, $\mathcal{N}(0, 1)$, o normal circularmente compleja, $\mathcal{CN}(0, 1)$, para z_r . Sin embargo, la deducción que se presenta a continuación es válida para cualquier distribución de z_r con media nula y momentos finitos hasta cuarto orden.

A partir de (5.1) resulta $|v_r|^2 = \sigma_r^2 |z_r|^2$ y reemplazando (5.2) en esta última expresión se obtiene

$$|v_r|^2 = k|z_r|^2 + \sum_{j=1}^p \alpha_j |z_r|^2 \sigma_{r-j}^2 + \sum_{j=1}^q \beta_j |z_r|^2 |v_{r-j}|^2. \quad (5.3)$$

En base a (5.2) y (5.3) se puede escribir la siguiente ecuación matricial

$$\begin{bmatrix} \sigma_r^2 \\ \sigma_{r-1}^2 \\ \vdots \\ \sigma_{r-p+1}^2 \\ |v_r|^2 \\ |v_{r-1}|^2 \\ \vdots \\ |v_{r-q+1}|^2 \end{bmatrix} = \begin{bmatrix} \alpha_1 & \cdots & \alpha_{p-1} & \alpha_p & \beta_1 & \cdots & \beta_{q-1} & \beta_q \\ 1 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & 1 & 0 & 0 & \cdots & 0 & 0 \\ \alpha_1 |z_r|^2 & \cdots & \alpha_{p-1} |z_r|^2 & \alpha_p |z_r|^2 & \beta_1 |z_r|^2 & \cdots & \beta_{q-1} |z_r|^2 & \beta_q |z_r|^2 \\ 0 & \cdots & 0 & 0 & 1 & \cdots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & 0 & 0 & 0 & \cdots & 1 & 0 \end{bmatrix} \begin{bmatrix} \sigma_{r-1}^2 \\ \sigma_{r-2}^2 \\ \vdots \\ \sigma_{r-p}^2 \\ |v_{r-1}|^2 \\ |v_{r-2}|^2 \\ \vdots \\ |v_{r-q}|^2 \end{bmatrix} + \begin{bmatrix} k \\ 0 \\ \vdots \\ 0 \\ k|z_r|^2 \\ 0 \\ \vdots \\ 0 \end{bmatrix}. \quad (5.4)$$

Notar que (5.4) tiene la forma de una ecuación de estados, que permite escribir el siguiente sistema de tiempo discreto

$$\mathbf{x}_r = \mathbf{F}_{r-1} \mathbf{x}_{r-1} + \mathbf{u}_{r-1} + \mathbf{w}_{r-1} \quad (5.5)$$

$$y_r = \mathbf{h}^T \mathbf{x}_r + n_r, \quad (5.6)$$

donde $\mathbf{x}_r = [\sigma_r^2 \ \sigma_{r-1}^2 \ \cdots \ \sigma_{r-p+1}^2 \ |v_r|^2 \ |v_{r-1}|^2 \ \cdots \ |v_{r-q+1}|^2]^T$ de dimensión $(p+q) \times 1$, es el vector de estados, $\mathbf{u}_{r-1} = [1 \ 0 \ \cdots \ 0 \ k|z_r|^2 \ 0 \ \cdots \ 0]^T$ de dimensión $(p+q) \times 1$, $\mathbf{h}^T = [0 \ 0 \ \cdots \ 0 \ 1 \ 0 \ \cdots \ 0]$ de dimensión $1 \times (p+q)$, \mathbf{w}_r y n_r son procesos de ruido blanco de media nula, no correlacionados entre ellos, con $\mathbb{E}\{\mathbf{w}_r \mathbf{w}_r^H\} = \mathbf{Q}_r$ y $\mathbb{E}\{n_r^2\} = \rho_r$ y la matriz de estados \mathbf{F}_{r-1} de dimensión $(p+q) \times (p+q)$, está dada por

$$\mathbf{F}_{r-1} = \begin{bmatrix} \alpha_1 & \cdots & \alpha_{p-1} & \alpha_p & \beta_1 & \cdots & \beta_{q-1} & \beta_q \\ 1 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & 1 & 0 & 0 & \cdots & 0 & 0 \\ \alpha_1 |z_r|^2 & \cdots & \alpha_{p-1} |z_r|^2 & \alpha_p |z_r|^2 & \beta_1 |z_r|^2 & \cdots & \beta_{q-1} |z_r|^2 & \beta_q |z_r|^2 \\ 0 & \cdots & 0 & 0 & 1 & \cdots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & 0 & 0 & 0 & \cdots & 1 & 0 \end{bmatrix}. \quad (5.7)$$

El sistema resultante posee la forma de un modelo lineal de estados, sin embargo en este caso la matriz de estados \mathbf{F}_{r-1} y el vector \mathbf{u}_{r-1} son aleatorios. El objetivo es deducir un estimador LMMSE para el vector de estado \mathbf{x}_r basado en el conocimiento de los coeficientes del proceso y de las mediciones y_r . El primer elemento de \mathbf{x}_r es la varianza condicional del proceso σ_r^2 , que es el parámetro de especial interés.

5.2.2. Deducción del estimador

En esta sección se deriva el estimador óptimo lineal para el vector de estados del sistema lineal de tiempo discreto con parámetros estocásticos. Para ello se define la siguiente factorización de \mathbf{F}_{r-1} y \mathbf{u}_{r-1}

$$\mathbf{F}_{r-1} = \bar{\mathbf{F}}_{r-1} + \mathbf{F}'_{r-1} \quad (5.8)$$

$$\mathbf{u}_{r-1} = \bar{\mathbf{u}}_{r-1} + \mathbf{u}'_{r-1}, \quad (5.9)$$

donde $\bar{\mathbf{F}}_{r-1} = \mathbb{E}\{\mathbf{F}_{r-1}\}$ y $\bar{\mathbf{u}}_{r-1} = \mathbb{E}\{\mathbf{u}_{r-1}\}$ de modo que la matriz \mathbf{F}'_{r-1} y el vector \mathbf{u}'_{r-1} son aleatorios con $\mathbb{E}\{\mathbf{F}'_{r-1}\} = \mathbf{0}$ y $\mathbb{E}\{\mathbf{u}'_{r-1}\} = \mathbf{0}$, respetivamente.

En la deducción del estimador LMMSE para el vector de estados \mathbf{x}_r se utiliza su interpretación geométrica, basada en el concepto de ortogonalidad en el espacio de Hilbert (Kay, 1993). Esta formulación supone que \mathbf{x}_r e y_r poseen media nula. Por este motivo, se deben definir las variables aleatorias $\mathbf{x}'_r = \mathbf{x}_r - \mathbb{E}\{\mathbf{x}_r\}$ e $y'_r = y_r - \mathbb{E}\{y_r\}$ y considerar la estimación de \mathbf{x}'_r por medio de una combinación lineal de las mediciones $\{y'_i\}_{i=1}^r$. De (5.5) y (5.6), $\mathbb{E}\{\mathbf{x}_r\}$ y $\mathbb{E}\{y_r\}$ son respectivamente

$$\mathbb{E}\{\mathbf{x}_r\} = \bar{\mathbf{x}}_r = \bar{\mathbf{F}}_{r-1}\bar{\mathbf{x}}_{r-1} + \bar{\mathbf{u}}_{r-1} \quad (5.10)$$

$$\mathbb{E}\{y_r\} = \bar{y}_r = \mathbf{h}^T \bar{\mathbf{x}}_r. \quad (5.11)$$

De la diferencia entre (5.5) y (5.10) y entre (5.6) y (5.11) se obtienen expresiones para \mathbf{x}'_r e y'_r respectivamente

$$\mathbf{x}'_r = \mathbf{F}_{r-1}\mathbf{x}'_{r-1} + \mathbf{F}'_{r-1}\bar{\mathbf{x}}_{r-1} + \mathbf{u}'_{r-1} + \mathbf{w}_{r-1} \quad (5.12)$$

$$y'_r = \mathbf{h}^T \mathbf{x}'_r. \quad (5.13)$$

El objetivo es obtener los coeficientes de la combinación lineal que minimice el error cuadrático medio (MSE por sus siglas en inglés)

$$\text{mse} = \mathbb{E}\{(x'_{jr} - \hat{x}'_{jr|r})^2\}, \quad j = 1, \dots, p + q \quad (5.14)$$

donde $x'_{jr} = [\mathbf{x}'_r]_j$ es el j -ésimo elemento del vector \mathbf{x}'_r y $\hat{x}'_{jr|r} = [\hat{\mathbf{x}}'_{r|r}]_j$, siendo $\hat{\mathbf{x}}'_{r|r}$ el estimador LMMSE de \mathbf{x}'_r dadas las mediciones hasta el instante r , es decir $\{y'_i\}_{i=1}^r$.

Con este enfoque el estimador LMMSE de x'_{jr} es la proyección ortogonal de x'_{jr} en el subespacio generado por $\{y'_i\}_{i=1}^r$, con el producto interno entre x e y definido como $\langle x, y \rangle = \mathbb{E}\{xy\}$ (Kay, 1993). Por otro lado, las innovaciones dadas por

$$\Delta y'_i = y'_i - \hat{y}'_{i|i-1}, \quad i = 1, \dots, r \quad (5.15)$$

conforman una base ortogonal de este subespacio, donde $\hat{y}'_{i|i-1}$ es el estimador LMMSE de y'_i dadas las mediciones hasta el instante $i - 1$. Luego, $\hat{\mathbf{x}}'_{r|r}$ se puede escribir como

$$\hat{\mathbf{x}}'_{r|r} = \sum_{i=1}^r \frac{\mathbb{E}\{\mathbf{x}'_r \Delta y'_i\}}{\mathbb{E}\{(\Delta y'_i)^2\}} \Delta y'_i, \quad (5.16)$$

que a su vez se puede factorizar de la siguiente manera

$$\hat{\mathbf{x}}'_{r|r} = \sum_{i=1}^{r-1} \frac{\mathbb{E}\{\mathbf{x}'_r \Delta y'_i\}}{\mathbb{E}\{(\Delta y'_i)^2\}} \Delta y'_i + \frac{\mathbb{E}\{\mathbf{x}'_r \Delta y'_r\}}{\mathbb{E}\{(\Delta y'_r)^2\}} \Delta y'_r. \quad (5.17)$$

Teniendo en cuenta que

$$\hat{\mathbf{x}}'_{r|r-1} = \sum_{i=1}^{r-1} \frac{\mathbb{E}\{\mathbf{x}'_r \Delta y'_i\}}{\mathbb{E}\{(\Delta y'_i)^2\}} \Delta y'_i, \quad (5.18)$$

resulta

$$\hat{\mathbf{x}}'_{r|r} = \hat{\mathbf{x}}'_{r|r-1} + \mathbf{k}_r (y'_r - \hat{y}'_{r|r-1}), \quad (5.19)$$

donde el vector de ganancias \mathbf{k}_r está dado por

$$\mathbf{k}_r = \frac{\mathbb{E}\{\mathbf{x}'_r \Delta y'_r\}}{\mathbb{E}\{(\Delta y'_r)^2\}}. \quad (5.20)$$

Para hallar una expresión de $\hat{\mathbf{x}}'_{r|r-1}$ se debe evaluar $\mathbb{E}\{\mathbf{x}'_r \Delta y'_i\}$. Reemplazando (5.12) dentro de la esperanza y distribuyendo los términos se obtiene

$$\mathbb{E}\{\mathbf{x}'_r \Delta y'_i\} = \mathbb{E}\{\mathbf{F}_{r-1} \mathbf{x}'_{r-1} \Delta y'_i\} + \mathbb{E}\{\mathbf{F}'_{r-1} \bar{\mathbf{x}}_{r-1} \Delta y'_i\} + \mathbb{E}\{\mathbf{u}'_{r-1} \Delta y'_i\} + \mathbb{E}\{\mathbf{w}_{r-1} \Delta y'_i\}. \quad (5.21)$$

Debido a que $\Delta y'_i$ es independiente de \mathbf{F}'_{r-1} , de \mathbf{u}'_{r-1} y de \mathbf{w}_{r-1} y teniendo en cuenta que estos tres últimos poseen media cero, el único término del lado derecho de la igualdad de (5.21) que no se anula es el primero. Asimismo, \mathbf{F}_{r-1} es independiente de \mathbf{x}'_{r-1} y de $\Delta y'_i$, por lo tanto

$$\mathbb{E}\{\mathbf{x}'_r \Delta y'_i\} = \bar{\mathbf{F}}_{r-1} \mathbb{E}\{\mathbf{x}'_{r-1} \Delta y'_i\}. \quad (5.22)$$

Finalmente, reemplazando (5.22) en (5.18) se obtiene

$$\hat{\mathbf{x}}'_{r|r-1} = \sum_{i=1}^{r-1} \bar{\mathbf{F}}_{r-1} \frac{\mathbb{E}\{\mathbf{x}'_{r-1} \Delta y'_i\}}{\mathbb{E}\{(\Delta y'_i)^2\}} \Delta y'_i = \bar{\mathbf{F}}_{r-1} \hat{\mathbf{x}}'_{r-1|r-1} \quad (5.23)$$

Notar que $\hat{y}'_{r|r-1}$ es de la forma de $\hat{\mathbf{x}}'_{r|r}$ en (5.16), es decir

$$\hat{y}'_{r|r-1} = \sum_{i=1}^{r-1} \frac{\mathbb{E}\{y'_r \Delta y'_i\}}{\mathbb{E}\{(\Delta y'_i)^2\}} \Delta y'_i. \quad (5.24)$$

Luego, reemplazando (5.13) en (5.24) se llega a una expresión que relaciona $\hat{y}'_{r|r-1}$ con $\hat{\mathbf{x}}'_{r|r-1}$

$$\hat{y}'_{r|r-1} = \sum_{i=1}^{r-1} \mathbf{h}^T \frac{\mathbb{E}\{\hat{\mathbf{x}}'_r \Delta y'_i\}}{\mathbb{E}\{(\Delta y'_i)^2\}} \Delta y'_i = \mathbf{h}^T \hat{\mathbf{x}}'_{r|r-1}. \quad (5.25)$$

Sumando $\bar{\mathbf{x}}_r$ a $\hat{\mathbf{x}}'_{r|r-1}$ e \bar{y}_r a $\hat{y}'_{r|r-1}$ se obtienen expresiones del tipo de (5.23) y (5.25) pero en términos de $\hat{\mathbf{x}}'_{r|r-1}$ y de $\hat{y}'_{r|r-1}$

$$\hat{\mathbf{x}}'_{r|r-1} = \bar{\mathbf{F}}_{r-1} \hat{\mathbf{x}}'_{r-1|r-1} + \bar{\mathbf{u}}_{r-1} \quad (5.26)$$

$$\hat{y}'_{r|r-1} = \mathbf{h}^T \hat{\mathbf{x}}'_{r|r-1}. \quad (5.27)$$

Solo resta determinar el vector de ganancias \mathbf{k}_r . Teniendo en cuenta que $\hat{\mathbf{x}}'_{r|r-1}$ pertenece al subespacio generado por las observaciones hasta el instante $r-1$ y que $\Delta y'_r$ es ortogonal a

dicho subespacio, entonces $\mathbb{E}\{\hat{\mathbf{x}}'_{r|r-1}\Delta y'_r\} = \mathbf{0}$. Esto permite reescribir el numerador de (5.20) en la forma

$$\mathbb{E}\{\hat{\mathbf{x}}'_{r|r-1}\Delta y'_r\} = \mathbb{E}\{(\mathbf{x}'_r - \hat{\mathbf{x}}'_{r|r-1})\Delta y'_r\}. \quad (5.28)$$

Reemplazando la expresión de $\Delta y'_r$ y distribuyendo se obtiene

$$\mathbb{E}\{\hat{\mathbf{x}}'_{r|r-1}\Delta y'_r\} = \mathbb{E}\{(\mathbf{x}'_r - \hat{\mathbf{x}}'_{r|r-1})(\mathbf{x}'_r - \hat{\mathbf{x}}'_{r|r-1})^T\}\mathbf{h} + \mathbb{E}\{(\mathbf{x}'_r - \hat{\mathbf{x}}'_{r|r-1})n_r\}. \quad (5.29)$$

Debido a que $\mathbf{x}'_r - \hat{\mathbf{x}}'_{r|r-1}$ y n_r son independientes y a que $\mathbb{E}\{n_r\} = 0$, entonces el término que contiene al ruido es cero, resultando

$$\mathbb{E}\{\hat{\mathbf{x}}'_{r|r-1}\Delta y'_r\} = \mathbf{M}_{r|r-1}\mathbf{h}, \quad (5.30)$$

donde

$$\mathbf{M}_{r|r-1} = \mathbb{E}\{(\mathbf{x}'_r - \hat{\mathbf{x}}'_{r|r-1})(\mathbf{x}'_r - \hat{\mathbf{x}}'_{r|r-1})^T\} = \mathbb{E}\{(\mathbf{x}_r - \hat{\mathbf{x}}_{r|r-1})(\mathbf{x}_r - \hat{\mathbf{x}}_{r|r-1})^T\}, \quad (5.31)$$

es la matriz de covarianza del error a priori.

Para el cálculo del denominador de (5.20) basta reemplazar $\Delta y'_r$ y operar

$$\mathbb{E}\{(\Delta y'_r)^2\} = \mathbf{h}^T \mathbb{E}\{(\mathbf{x}'_r - \hat{\mathbf{x}}'_{r|r-1})(\mathbf{x}'_r - \hat{\mathbf{x}}'_{r|r-1})^T\}\mathbf{h} + 2\mathbf{h}^T \mathbb{E}\{(\mathbf{x}'_r - \hat{\mathbf{x}}'_{r|r-1})n_r\} + \mathbb{E}\{n_r^2\}. \quad (5.32)$$

Con el argumento utilizado en el cálculo del numerador se puede concluir que es cero el segundo término del lado derecho de la igualdad en (5.32), resultando

$$\mathbb{E}\{(\Delta y'_r)^2\} = \mathbf{h}^T \mathbf{M}_{r|r-1} \mathbf{h} + \rho_r. \quad (5.33)$$

Del cociente entre (5.30) y (5.33) se obtiene una expresión para el vector de ganancias \mathbf{k}_r , análoga a la filtro de Kalman (Kay, 1993)

$$\mathbf{k}_r = \frac{\mathbf{M}_{r|r-1}\mathbf{h}}{\mathbf{h}^T \mathbf{M}_{r|r-1} \mathbf{h} + \rho_r}, \quad (5.34)$$

De (5.34) se puede apreciar que para evaluar el vector de ganancias es necesario contar con una fórmula que permita calcular la matriz de covarianza del error a priori. A partir de (5.5) y (5.26) se puede escribir

$$\mathbf{x}_r - \hat{\mathbf{x}}_{r|r-1} = \mathbf{F}'_{r-1}\mathbf{x}_{r-1} + \bar{\mathbf{F}}_{r-1}(\mathbf{x}_{r-1} - \hat{\mathbf{x}}_{r-1|r-1}) + \mathbf{u}'_{r-1} + \mathbf{w}_{r-1}. \quad (5.35)$$

Reemplazando (5.35) en (5.31), y luego de algunas operaciones algebraicas se obtiene una expresión para el cálculo de $\mathbf{M}_{r|r-1}$, la cual está dada por

$$\begin{aligned} \mathbf{M}_{r|r-1} = & \bar{\mathbf{F}}_{r-1}\mathbf{M}_{r-1|r-1}\bar{\mathbf{F}}_{r-1}^T + \mathbb{E}\{\mathbf{F}'_{r-1}\mathbf{x}_{r-1}\mathbf{x}_{r-1}^T\mathbf{F}'_{r-1}\} + \mathbb{E}\{\mathbf{F}'_{r-1}\mathbf{x}_{r-1}\mathbf{u}'_{r-1}\} + \\ & \mathbb{E}\{\mathbf{u}'_{r-1}\mathbf{x}_{r-1}^T\mathbf{F}'_{r-1}\} + \mathbb{E}\{\mathbf{u}'_{r-1}\mathbf{u}'_{r-1}\} + \mathbf{Q}_{r-1} \end{aligned} \quad (5.36)$$

donde $\mathbf{M}_{r|r}$ es la matriz de covarianza del error a posteriori

$$\mathbf{M}_{r|r} = \mathbb{E}\{(\mathbf{x}'_r - \hat{\mathbf{x}}'_{r|r})(\mathbf{x}'_r - \hat{\mathbf{x}}'_{r|r})^T\} = \mathbb{E}\{(\mathbf{x}_r - \hat{\mathbf{x}}_{r|r})(\mathbf{x}_r - \hat{\mathbf{x}}_{r|r})^T\}. \quad (5.37)$$

Es importante destacar que en la deducción de (5.36) muchos de los términos se anulan debido a que los factores involucrados en los productos dentro de las esperanzas son independientes entre sí, y a que algunos de ellos a su vez poseen media nula. Por otro lado, se observa que (5.36) es de la forma de $\mathbf{M}_{r|r-1}$ en Koning (1984), si se iguala \mathbf{u}'_{r-1} a cero como es la hipótesis en ese trabajo las expresiones coinciden.

La ecuación (5.36) es una fórmula general que permite evaluar $\mathbf{M}_{r|r-1}$ para cualquier matriz \mathbf{F}'_{r-1} y vector \mathbf{u}'_{r-1} . Sin embargo, para el problema particular planteado en este trabajo \mathbf{F}'_{r-1} y \mathbf{u}'_{r-1} presentan una estructura particular. A continuación se desarrolla una alternativa para evaluar los términos de (5.36) que involucran estas matrices con dicha estructura.

A partir de (5.8) y (5.9) se pueden escribir \mathbf{F}'_{r-1} y \mathbf{u}'_{r-1} , respectivamente, como

$$\mathbf{F}'_{r-1} = \mathbf{F}_{r-1} - \bar{\mathbf{F}}_{r-1} = (|z_r|^2 - \mathbb{E}\{|z_r|^2\})\Delta\bar{\mathbf{F}}_{r-1} \quad (5.38)$$

$$\mathbf{u}'_{r-1} = \mathbf{u}_{r-1} - \bar{\mathbf{u}}_{r-1} = (|z_r|^2 - \mathbb{E}\{|z_r|^2\})\Delta\bar{\mathbf{u}}_{r-1}, \quad (5.39)$$

donde $\Delta\bar{\mathbf{u}}_{r-1}$ y $\Delta\bar{\mathbf{F}}_{r-1}$ son determinísticos, con $\Delta\bar{\mathbf{u}}_{r-1} = [0 \ 0 \ \dots \ 0 \ k \ 0 \ \dots \ 0]^T$ y

$$\Delta\bar{\mathbf{F}}_{r-1} = \begin{bmatrix} 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 \\ \alpha_1 & \dots & \alpha_{p-1} & \alpha_p & \beta_1 & \dots & \beta_{q-1} & \beta_q \\ 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 \end{bmatrix}. \quad (5.40)$$

Con el objetivo de hallar una expresión para el término $\mathbb{E}\{\mathbf{F}'_{r-1}\mathbf{x}_{r-1}\mathbf{x}_{r-1}^T\mathbf{F}'_{r-1}{}^T\}$ una posibilidad es vectorizar la matriz $\mathbf{F}'_{r-1}\mathbf{x}_{r-1}\mathbf{x}_{r-1}^T\mathbf{F}'_{r-1}{}^T$ utilizando el operador $\text{vec}(\cdot)$. Dada una matriz \mathbf{A} de $R_A \times T_A$, $\text{vec}(\mathbf{A})$ es un vector de $R_A T_A \times 1$ compuesto por las columnas de \mathbf{A} apiladas (Harville, 1997). Debido a que se conocen las dimensiones de las matrices involucradas no existirán problemas para realizar el proceso inverso a la vectorización.

Para cualquier conjunto de matrices \mathbf{A} de $R_A \times T_A$, \mathbf{B} de $T_A \times T_B$ y \mathbf{C} de $T_B \times T_C$, el vector $\text{vec}(\mathbf{ABC})$ se puede expresar como

$$\text{vec}(\mathbf{ABC}) = (\mathbf{C}^T \otimes \mathbf{A})\text{vec}(\mathbf{B}), \quad (5.41)$$

donde \otimes denota el producto de Kronecker entre las matrices \mathbf{C}^T y \mathbf{A} , el que se define como

$$\mathbf{A} \otimes \mathbf{B} = \begin{bmatrix} a_{11}\mathbf{B} & a_{12}\mathbf{B} & \dots & a_{1T_a}\mathbf{B} \\ a_{21}\mathbf{B} & a_{22}\mathbf{B} & \dots & a_{2T_a}\mathbf{B} \\ \vdots & \vdots & \ddots & \vdots \\ a_{R_a1}\mathbf{B} & a_{R_a2}\mathbf{B} & \dots & a_{R_aT_a}\mathbf{B} \end{bmatrix}, \quad (5.42)$$

siendo $a_{ij} = [\mathbf{A}]_{ij}$ el elemento ij de la matriz \mathbf{A} .

A partir de la propiedad (5.41) se puede expresar

$$\mathbb{E}\{\text{vec}(\mathbf{F}'_{r-1}\mathbf{x}_{r-1}\mathbf{x}_{r-1}^T\mathbf{F}'_{r-1})\} = \mathbb{E}\{(\mathbf{F}'_{r-1} \otimes \mathbf{F}'_{r-1})\text{vec}(\mathbf{x}_{r-1}\mathbf{x}_{r-1}^T)\}. \quad (5.43)$$

La ventaja de esta representación es que permite factorizar la esperanza como producto de esperanzas debido a que \mathbf{F}'_{r-1} y \mathbf{x}_{r-1} son independientes, es decir

$$\mathbb{E}\{\text{vec}(\mathbf{F}'_{r-1}\mathbf{x}_{r-1}\mathbf{x}_{r-1}^T\mathbf{F}'_{r-1})\} = \mathbb{E}\{\mathbf{F}'_{r-1} \otimes \mathbf{F}'_{r-1}\}\text{vec}(\mathbf{P}_{r-1}). \quad (5.44)$$

donde $\mathbf{P}_r = \mathbb{E}\{\mathbf{x}_r\mathbf{x}_r^T\}$. Reemplazando (5.38) en $\mathbb{E}\{(\mathbf{F}'_{r-1} \otimes \mathbf{F}'_{r-1})\}$ resulta

$$\mathbb{E}\{(\mathbf{F}'_{r-1} \otimes \mathbf{F}'_{r-1})\} = \mathbb{E}\{(|z_r|^2 - \mathbb{E}\{|z_r|^2\})^2\}\Delta\bar{\mathbf{F}}_{r-1} \otimes \Delta\bar{\mathbf{F}}_{r-1}. \quad (5.45)$$

Definiendo

$$\gamma = \mathbb{E}\{(|z_r|^2 - \mathbb{E}\{|z_r|^2\})^2\} = m_4 - m_2^2, \quad (5.46)$$

finalmente resulta

$$\mathbb{E}\{\text{vec}(\mathbf{F}'_{r-1}\mathbf{x}_{r-1}\mathbf{x}_{r-1}^T\mathbf{F}'_{r-1})\} = \gamma(\Delta\mathbf{F}_{r-1} \otimes \Delta\mathbf{F}_{r-1})\text{vec}(\mathbf{P}_{r-1}). \quad (5.47)$$

Invirtiendo el proceso de vectorización se obtiene la expresión que permite calcular $\mathbb{E}\{\mathbf{F}'_{r-1}\mathbf{x}_{r-1}\mathbf{x}_{r-1}^T\mathbf{F}'_{r-1}\}$, la cual está dada por

$$\mathbb{E}\{\mathbf{F}'_{r-1}\mathbf{x}_{r-1}\mathbf{x}_{r-1}^T\mathbf{F}'_{r-1}\} = \gamma\Delta\mathbf{F}_{r-1}\mathbf{P}_{r-1}\Delta\mathbf{F}_{r-1}^T. \quad (5.48)$$

El procedimiento para calcular $\mathbb{E}\{\mathbf{F}'_{r-1}\mathbf{x}_{r-1}\mathbf{u}'_{r-1}\}$ es el mismo. A partir de la propiedad (5.41) se puede expresar

$$\mathbb{E}\{\text{vec}(\mathbf{F}'_{r-1}\mathbf{x}_{r-1}\mathbf{u}'_{r-1})\} = \mathbb{E}\{(\mathbf{u}'_{r-1} \otimes \mathbf{F}'_{r-1})\mathbf{x}_{r-1}\}. \quad (5.49)$$

Debido a que tanto \mathbf{u}'_{r-1} como \mathbf{F}'_{r-1} son independientes de \mathbf{x}_{r-1} entonces (5.49) puede escribirse como

$$\mathbb{E}\{\text{vec}(\mathbf{F}'_{r-1}\mathbf{x}_{r-1}\mathbf{u}'_{r-1})\} = \mathbb{E}\{\mathbf{u}'_{r-1} \otimes \mathbf{F}'_{r-1}\}\bar{\mathbf{x}}_{r-1}. \quad (5.50)$$

Reemplazando (5.38) y (5.39) en el término que involucra el producto de Kronecker en (5.50) se obtiene

$$\mathbb{E}\{(\mathbf{u}'_{r-1} \otimes \mathbf{F}'_{r-1})\} = \gamma\Delta\bar{\mathbf{u}}_{r-1} \otimes \Delta\bar{\mathbf{F}}_{r-1}, \quad (5.51)$$

resultando

$$\mathbb{E}\{\text{vec}(\mathbf{F}'_{r-1}\mathbf{x}_{r-1}\mathbf{u}'_{r-1})\} = \gamma(\Delta\bar{\mathbf{u}}_{r-1} \otimes \Delta\bar{\mathbf{F}}_{r-1})\bar{\mathbf{x}}_{r-1}. \quad (5.52)$$

Invitiendo el proceso de vectorización lleva a la siguiente expresión para evaluar $\mathbb{E}\{\mathbf{F}'_{r-1}\mathbf{x}_{r-1}\mathbf{u}'_{r-1}\}$

$$\mathbb{E}\{\mathbf{F}'_{r-1}\mathbf{x}_{r-1}\mathbf{u}'_{r-1}\} = \gamma\Delta\mathbf{F}_{r-1}\bar{\mathbf{x}}_{r-1}\Delta\mathbf{u}'_{r-1}^T. \quad (5.53)$$

Por último, el término $\mathbb{E}\{\mathbf{u}'_{r-1}\mathbf{u}'_{r-1}\}$ se evalúa simplemente reemplazando \mathbf{u}'_{r-1} por (5.39), dando como resultado

$$\mathbb{E}\{\mathbf{u}'_{r-1}\mathbf{u}'_{r-1}\} = \gamma\Delta\mathbf{u}_{r-1}\Delta\mathbf{u}_{r-1}^T. \quad (5.54)$$

Por lo tanto, reemplazando (5.48), (5.53) y (5.54) en (5.36) se obtiene la fórmula que permite calcular $\mathbf{M}_{r|r-1}$ para problemas donde \mathbf{F}'_{r-1} y \mathbf{u}'_{r-1} son de la forma de (5.38) y (5.39), respectivamente

$$\begin{aligned} \mathbf{M}_{r|r-1} = & \bar{\mathbf{F}}_{r-1} \mathbf{M}_{r-1|r-1} \bar{\mathbf{F}}_{r-1}^T + \gamma (\Delta \mathbf{F}_{r-1} \mathbf{P}_{r-1} \Delta \mathbf{F}_{r-1}^T + \Delta \mathbf{F}_{r-1} \bar{\mathbf{x}}_{r-1} \Delta \mathbf{u}_{r-1}^T + \\ & \Delta \mathbf{u}_{r-1} \bar{\mathbf{x}}_{r-1}^T \Delta \mathbf{F}_{r-1}^T + \Delta \mathbf{u}_{r-1} \Delta \mathbf{u}_{r-1}^T) + \mathbf{Q}_{r-1} \end{aligned} \quad (5.55)$$

Se puede observar como difiere (5.36), o (5.55), respecto de las ecuaciones del filtro del Kalman original. Esto se debe a la naturaleza aleatoria de la matrices del sistema a través de z_r .

Por último, se requieren fórmulas recursivas para evaluar $\mathbf{M}_{r|r}$ y \mathbf{P}_r . Es posible demostrar que $\mathbf{M}_{r|r}$ puede computarse como en el caso del filtro de Kalman (Simon, 2006). Para ello es necesario reemplazar (5.19) en (5.37), a partir de lo que resulta

$$\begin{aligned} \mathbf{M}_{r|r} = & \mathbb{E}\{(\mathbf{x}'_r - \hat{\mathbf{x}}'_{r|r-1})(\mathbf{x}'_r - \hat{\mathbf{x}}'_{r|r-1})^T\} - \mathbb{E}\{(\mathbf{x}'_r - \hat{\mathbf{x}}'_{r|r-1})\Delta y'_r\} \mathbf{k}_r^T - \\ & \mathbf{k}_r \mathbb{E}\{\Delta y'_r (\mathbf{x}'_r - \hat{\mathbf{x}}'_{r|r-1})^T\} + \mathbf{k}_r \mathbb{E}\{(\Delta y'_r)^2\} \mathbf{k}_r^T, \end{aligned} \quad (5.56)$$

donde el primer término es $\mathbf{M}_{r|r-1}$, las esperanzas del segundo y del tercero corresponden al numerador de \mathbf{k}_r y su traspuesto, respectivamente, y la esperanza del último es el denominador de \mathbf{k}_r . Haciendo uso de (5.30) $\mathbf{M}_{r|r}$ se puede reescribir como

$$\mathbf{M}_{r|r} = \mathbf{M}_{r|r-1} - \mathbf{M}_{r|r-1} \mathbf{h} \mathbf{k}_r^T - \mathbf{k}_r \mathbf{h}^T \mathbf{M}_{r|r-1} + \mathbf{M}_{r|r-1} \mathbf{h} \mathbf{k}_r^T. \quad (5.57)$$

Finalmente, la fórmula para $\mathbf{M}_{r|r}$ es

$$\mathbf{M}_{r|r} = (\mathbf{I} - \mathbf{k}_r \mathbf{h}^T) \mathbf{M}_{r|r-1}, \quad (5.58)$$

donde \mathbf{I} es la matriz de identidad de dimensión $(p+q) \times (p+q)$.

Por su parte, la expresión para \mathbf{P}_r surge de manera similar a la deducción de $\mathbf{M}_{r|r-1}$. Sustituyendo la expresión de \mathbf{x}_r en $\mathbb{E}\{\mathbf{x}_r \mathbf{x}_r^T\}$, muchos de los términos resultantes una vez distribuidos se anulan utilizando las mismas consideraciones empleadas en el caso de $\mathbf{M}_{r|r-1}$. En conclusión, la fórmula para \mathbf{P}_r está dada por

$$\begin{aligned} \mathbf{P}_r = & \bar{\mathbf{F}}_{r-1} \mathbf{P}_{r-1} \bar{\mathbf{F}}_{r-1}^T + \bar{\mathbf{F}}_{r-1} \bar{\mathbf{x}}_{r-1} \bar{\mathbf{u}}_{r-1}^T + \bar{\mathbf{u}}_{r-1} \bar{\mathbf{x}}_{r-1}^T \bar{\mathbf{F}}_{r-1}^T + \bar{\mathbf{u}}_{r-1} \bar{\mathbf{u}}_{r-1}^T + \mathbb{E}\{\mathbf{F}'_{r-1} \mathbf{x}_{r-1} \mathbf{x}_{r-1}^T \mathbf{F}'_{r-1}^T\} + \\ & \mathbb{E}\{\mathbf{F}'_{r-1} \mathbf{x}_{r-1} \mathbf{u}'_{r-1}^T\} + \mathbb{E}\{\mathbf{u}'_{r-1} \mathbf{x}_{r-1}^T \mathbf{F}'_{r-1}^T\} + \mathbb{E}\{\mathbf{u}'_{r-1} \mathbf{u}'_{r-1}^T\} + \mathbf{Q}_{r-1}, \end{aligned} \quad (5.59)$$

para el caso general y

$$\begin{aligned} \mathbf{P}_r = & \bar{\mathbf{F}}_{r-1} \mathbf{P}_{r-1} \bar{\mathbf{F}}_{r-1}^T + \bar{\mathbf{F}}_{r-1} \bar{\mathbf{x}}_{r-1} \bar{\mathbf{u}}_{r-1}^T + \bar{\mathbf{u}}_{r-1} \bar{\mathbf{x}}_{r-1}^T \bar{\mathbf{F}}_{r-1}^T + \bar{\mathbf{u}}_{r-1} \bar{\mathbf{u}}_{r-1}^T + \gamma (\Delta \mathbf{F}_{r-1} \mathbf{P}_{r-1} \Delta \mathbf{F}_{r-1}^T + \\ & \Delta \mathbf{F}_{r-1} \bar{\mathbf{x}}_{r-1} \Delta \mathbf{u}_{r-1}^T + \Delta \mathbf{u}_{r-1} \bar{\mathbf{x}}_{r-1}^T \Delta \mathbf{F}_{r-1}^T + \Delta \mathbf{u}_{r-1} \Delta \mathbf{u}_{r-1}^T) + \mathbf{Q}_{r-1}, \end{aligned} \quad (5.60)$$

en las situación en que \mathbf{F}'_{r-1} y \mathbf{u}'_{r-1} son de la forma (5.38) y (5.39), respectivamente. Nuevamente, estas ecuaciones se asemejan a las del filtro del Kalman, con un término adicional.

Esto completa la deducción del estimador LMMSE para la varianza condicional de un proceso GARCH. A continuación se resumen la ecuaciones obtenidas en el orden en que deben implementarse. Una vez que se inicializan $\bar{\mathbf{x}}_0$, $\hat{\mathbf{x}}_{0|0}$, \mathbf{P}_0 y $\mathbf{M}_{0|0}$ para $r \geq 1$ los pasos son los siguientes

- 1) **Estimado a priori:** actualizar $\hat{\mathbf{x}}_{r|r-1}$ y $\hat{y}_{r|r-1}$ a partir de (5.26) y (5.27) respectivamente.
- 2) **Matriz de covarianza del error a priori:** calcular la matriz $\mathbf{M}_{r|r-1}$ por medio de (5.55).
- 3) **Ganancia:** evaluar el vector de ganancias \mathbf{k}_r utilizando la ecuación (5.34).
- 4) **Estimado a posteriori:** actualizar $\hat{\mathbf{x}}_{r|r}$. Reemplazando (5.10) y (5.11) en (5.19) se llega a la expresión que permite llevar a cabo este paso

$$\hat{\mathbf{x}}_{r|r} = \hat{\mathbf{x}}_{r|r-1} + \mathbf{k}_r(y_r - \hat{y}_{r|r-1}), \quad (5.61)$$

- 5) **Matriz de covarianza del error a posteriori:** actualizar $\mathbf{M}_{r|r}$ por medio de (5.58).
- 6) **Media y valor cuadrático medio:** evaluar $\bar{\mathbf{x}}_r$ y \mathbf{P}_r utilizando las ecuaciones (5.10) y (5.60), respectivamente.

5.2.3. Simulaciones numéricas

En esta sección se analiza una aplicación del estimador LMMSE de la varianza condicional para un proceso GARCH(1,1). De este modo, el modelo de estados se reduce a

$$\begin{bmatrix} x_{1t} \\ x_{2t} \end{bmatrix} = \begin{bmatrix} \alpha_1 & \beta_1 \\ \alpha_1 |z_t|^2 & \beta_1 |z_t|^2 \end{bmatrix} \begin{bmatrix} x_{1,t-1} \\ x_{2,t-1} \end{bmatrix} + \begin{bmatrix} 1 \\ |z_t|^2 \end{bmatrix} k + \begin{bmatrix} w_{1t} \\ w_{2t} \end{bmatrix} \quad (5.62)$$

$$y_t = [0 \ 1] \begin{bmatrix} x_{1,t} \\ x_{2,t} \end{bmatrix} + n_t. \quad (5.63)$$

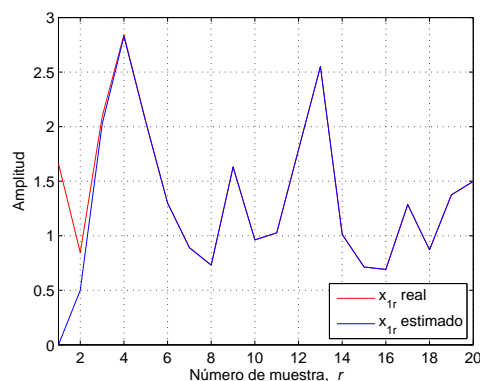


Figura 5.1: Estimado de la varianza condicional a partir de datos sin ruido y estado inicial desconocido.

Se considera que $z_r \sim \mathcal{N}(0, 1)$, por lo tanto $m_2 = 1$, $m_4 = 3$ y $\gamma = 2$. Con el objetivo de evaluar el desempeño del estimador LMMSE propuesto, el mismo se pone a prueba utilizando datos sintéticos generados a partir del proceso GARCH(1,1). Los coeficientes del proceso son $k = 0.5$, $\alpha_1 = 0.2$ y $\beta_1 = 0.3$. En primer lugar se corre el estimador con datos sin ruido, pero

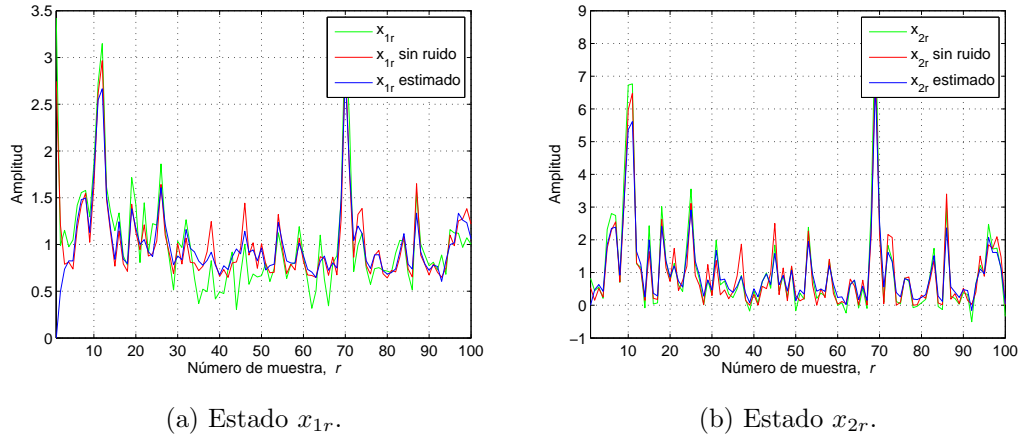


Figura 5.2: Estimado LMMSE a partir de datos con ruido y estado inicial desconocido.

con el estado inicial desconocido. En la figura 5.1 se muestra que, como es de esperar, el estimado converge rápidamente al valor verdadero de la varianza condicional. Luego, se corre el algoritmo con ruido de estado, \mathbf{w}_r , y de observación, n_r , donde la varianza de ruido de observación fue $\rho_r = k^2/9$ y la matriz de covarianza del ruido de estado fue $\mathbf{Q}_r = \frac{k^2}{9}\mathbf{I}$. Las figuras 5.2 y 5.3 muestran los estimados de los estados y el error de estimación, respectivamente para una realización de los datos. Para la varianza condicional se observa que los estimados aproximan su verdadero valor. Además, se calculó el estimador LMMSE para un conjunto de datos de 1000 muestras de largo y se evaluó la varianza de la varianza condicional (x_{1t} , figura 5.2a) y la varianza del error de estimación (figura 5.3a). Este experimento se repitió para 1000 realizaciones independientes y obtuvo la media de estas varianzas. El valor medio de la varianza de x_{1t} y el valor medio de la varianza del error de estimación resultó 0.304 y 0.029 respectivamente. Finalmente, se ejecutó el estimador LMMSE para numerosas realizaciones de 1000 muestras de largo y se evaluó la raíz cuadrada del error cuadrático medio para cada instante. La figura 5.4 muestra la RMSE de la varianza condicional. Sus valores se encuentran alrededor de 0.17, sin embargo para algunos instantes se aparta de este valor debido al comportamiento impulsivo del proceso y de su varianza condicional.

5.3. Estimador LMMSE extendido

5.3.1. Formulación del problema

El estimador LMMSE para la varianza condicional desarrollado en la sección anterior no tiene sentido desde un punto de vista práctico, debido a que no es una situación realista que los valores de sus coeficientes se conozcan. Sin embargo, en primera instancia se formuló esa hipótesis para confirmar la forma del estimador en ese caso, debido a que difiere del filtro de Kalman tradicional y como paso previo a complejizar el algoritmo incorporando la estimación de los parámetros del modelo. Por este motivo, a continuación se considera que los coeficientes son desconocidos y que se desean estimar simultáneamente con la varianza condicional del

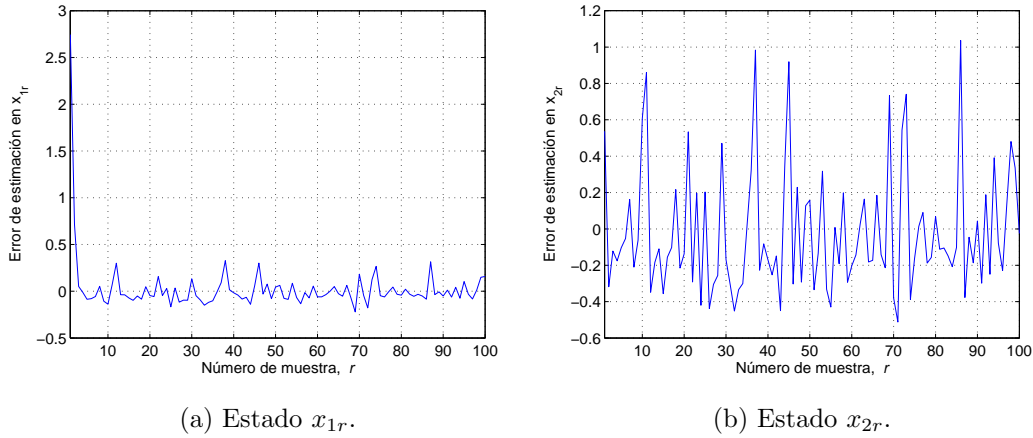


Figura 5.3: Error en la estimación LMMSE a partir de datos con ruido y estado inicial desconocido.

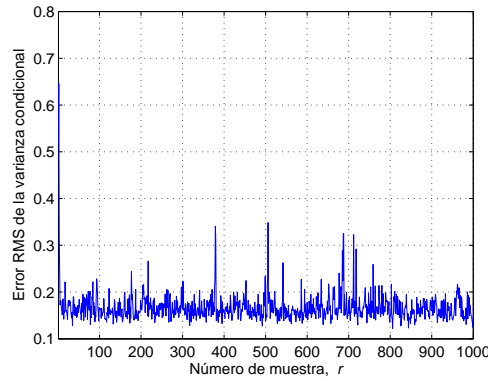


Figura 5.4: RMSE en la estimación LMMSE de la varianza condicional.

proceso.

Con el objetivo de simplificar las expresiones, la deducción se restringirá a un proceso GARCH(1,1) de la forma

$$v_r = \sigma_r z_r, \tag{5.64}$$

$$\sigma_r^2 = k + \alpha_1 \sigma_{r-1}^2 + \beta_1 v_{r-1}^2, \tag{5.65}$$

donde nuevamente z_r es un proceso aleatorio con momentos no centrales $m_n = \mathbb{E}\{|z_r|^n\}$. Sin embargo, los resultados a los que se arriban pueden ser fácilmente extendidos a un proceso GARCH con órdenes p y q arbitrarios.

Como no se conocen los coeficientes, con el objetivo de estimarlos, se aumenta el vector de

estados con estos parámetros, es decir

$$\mathbf{x}_r = \begin{bmatrix} x_{1r} \\ x_{2r} \\ x_{3r} \\ x_{4r} \\ x_{5r} \end{bmatrix} = \begin{bmatrix} \sigma_r^2 \\ |v_r|^2 \\ k \\ \alpha_1 \\ \beta_1 \end{bmatrix}. \quad (5.66)$$

Debido a que los coeficientes son constantes, se modelan de la siguiente manera

$$k_r = k_{r-1} + w_{k,r-1} \quad (5.67)$$

$$\alpha_{1r} = \alpha_{1,r-1} + w_{\alpha,r-1} \quad (5.68)$$

$$\beta_{1r} = \beta_{1,r-1} + w_{\beta,r-1}, \quad (5.69)$$

donde $w_{k,r}$, $w_{\alpha,r}$ y $w_{\beta,r-1}$ son términos de ruido artificiales que permiten que el filtro modifique el estado de k , de α_1 y de β_1 respectivamente. Como en cualquier esquema de estimación conjunta estados-parámetros, la varianza del ruido regulará la velocidad de adaptación del valor de estos parámetros.

En estas condiciones y haciendo uso de (5.3) para un proceso GARCH(1,1), el sistema de estados aumentado puede escribirse como

$$\mathbf{x}_r = \mathbf{f}(\mathbf{x}_{r-1}, \mathbf{w}_{r-1}) = \begin{bmatrix} x_{3,r-1} + x_{4,r-1}x_{1,r-1} + x_{5,r-1}x_{2,r-1} \\ |z_r|^2(x_{3,r-1} + x_{4,r-1}x_{1,r-1} + x_{5,r-1}x_{2,r-1}) \\ x_{3,r-1} \\ x_{4,r-1} \\ x_{5,r-1} \end{bmatrix} + \begin{bmatrix} w_{1r} \\ w_{2r} \\ w_{3r} \\ w_{4r} \\ w_{5r} \end{bmatrix} \quad (5.70)$$

$$y_r = \mathbf{h}^T \mathbf{x}_r + n_r = [0 \ 1 \ 0 \ 0 \ 0] \mathbf{x}_r + n_r, \quad (5.71)$$

donde \mathbf{w}_r y n_r son procesos de ruido blanco de media nula, no correlacionados entre ellos, con $\mathbb{E}\{\mathbf{w}_r \mathbf{w}_r^H\} = \mathbf{Q}_r$ y $\mathbb{E}\{|n_r|^2\} = \rho_r$.

Notar que $\mathbf{f}(\mathbf{x}_{r-1}, \mathbf{w}_{r-1})$ es una función no lineal del estado aumentado \mathbf{x}_{r-1} . Por lo tanto, para estimar \mathbf{x}_r debe utilizarse una versión extendida del estimador LMMSE presentado, análoga al filtro de Kalman extendido (Simon, 2006). El procedimiento consiste en linealizar el problema empleando el desarrollo en serie de Taylor alrededor de un valor nominal del estado. Estos valores nominales consisten en estimaciones anteriores de la evolución de los estados del sistema.

El desarrollo en serie de Taylor de la función $\mathbf{f}(\mathbf{x}, \mathbf{w})$ alrededor de $\mathbf{x} = \hat{\mathbf{x}}_{r-1|r-1}$ y $\mathbf{w} = \mathbf{0}$, truncada al término de primer orden está dado por

$$\mathbf{f}(\mathbf{x}, \mathbf{w}) \approx \mathbf{f}(\hat{\mathbf{x}}_{r-1|r-1}, \mathbf{0}) + \left. \frac{\partial \mathbf{f}}{\partial \mathbf{x}} \right|_{\substack{\mathbf{x}=\hat{\mathbf{x}}_{r-1|r-1} \\ \mathbf{w}=\mathbf{0}}} (\mathbf{x} - \hat{\mathbf{x}}_{r-1|r-1}) + \left. \frac{\partial \mathbf{f}}{\partial \mathbf{w}} \right|_{\substack{\mathbf{x}=\hat{\mathbf{x}}_{r-1|r-1} \\ \mathbf{w}=\mathbf{0}}} \mathbf{w}, \quad (5.72)$$

donde $\frac{\partial \mathbf{f}}{\partial \mathbf{x}}$ y $\frac{\partial \mathbf{f}}{\partial \mathbf{w}}$ son matrices de 5×5 cuyos elementos son $[\frac{\partial \mathbf{f}}{\partial \mathbf{x}}]_{ij} = \frac{\partial f_i}{\partial x_j}$ y $[\frac{\partial \mathbf{f}}{\partial \mathbf{w}}]_{ij} = \frac{\partial f_i}{\partial w_j}$ respectivamente.

Definiendo

$$\mathbf{F}_{r-1} = \left. \frac{\partial \mathbf{f}}{\partial \mathbf{x}} \right|_{\substack{\mathbf{x}=\hat{\mathbf{x}}_{r-1|r-1} \\ \mathbf{w}=\mathbf{0}}} = \begin{bmatrix} \hat{x}_{4,r-1|r-1} & \hat{x}_{5,r-1|r-1} & 1 & \hat{x}_{1,r-1|r-1} & \hat{x}_{2,r-1|r-1} \\ |z_r|^2 \hat{x}_{4,r-1|r-1} & |z_r|^2 \hat{x}_{5,r-1|r-1} & |z_r|^2 & |z_r|^2 \hat{x}_{1,r-1|r-1} & |z_r|^2 \hat{x}_{2,r-1|r-1} \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}, \quad (5.73)$$

y teniendo en cuenta que

$$\left. \frac{\partial \mathbf{f}}{\partial \mathbf{w}} \right|_{\substack{\mathbf{x}=\hat{\mathbf{x}}_{r-1|r-1} \\ \mathbf{w}=\mathbf{0}}} = \mathbf{I}, \quad (5.74)$$

entonces (5.72) puede expresarse como

$$\mathbf{f}(\mathbf{x}, \mathbf{w}) \approx \mathbf{f}(\hat{\mathbf{x}}_{r-1|r-1}, \mathbf{0}) + \mathbf{F}_{r-1}(\mathbf{x} - \hat{\mathbf{x}}_{r-1|r-1}) + \mathbf{w}. \quad (5.75)$$

Esta aproximación de $\mathbf{f}(\mathbf{x}, \mathbf{w})$ permite plantear el siguiente sistema lineal de tiempo discreto

$$\mathbf{x}_r = \mathbf{F}_{r-1} \mathbf{x}_{r-1} + \mathbf{u}_{r-1} + \mathbf{w}_{r-1} \quad (5.76)$$

$$y_r = \mathbf{h}^T \mathbf{x}_r + n_r, \quad (5.77)$$

donde se definió

$$\mathbf{u}_{r-1} = \mathbf{f}(\hat{\mathbf{x}}_{r-1|r-1}, \mathbf{0}) - \mathbf{F}_{r-1} \hat{\mathbf{x}}_{r-1|r-1}. \quad (5.78)$$

5.3.2. Dedución del estimador

El sistema representado por (5.76) y (5.77) es de la misma forma del que se utilizó en la deducción del estimador LMMSE bajo la consideración de coeficientes conocidos.

Continuando con el paralelismo con el filtro de Kalman extendido, una vez que el sistema de estados es linealizado, se sustituyen las matrices obtenidas en las ecuaciones del filtro de Kalman y se deriva el algoritmo que permite estimar el vector de estados (Simon, 2006). En este procedimiento se supone que todas las esperanzas involucradas en la deducción son condicionadas a $\hat{\mathbf{x}}_{r-1|r-1}$.

En este caso se procederá de la misma manera, y es importante destacar que bajo esta consideración las esperanzas se tomarán respecto de los términos de las matrices aleatorias en los que aparece z_r .

Antes de utilizar los resultados obtenidos en la sección anterior se deben factorizar la matriz \mathbf{F}_{r-1} y el vector \mathbf{u}_{r-1} en la forma dada por (5.8) y (5.9). Para ello, la media de \mathbf{F}_{r-1} es

$$\bar{\mathbf{F}}_{r-1} = \begin{bmatrix} \hat{x}_{4,r-1|r-1} & \hat{x}_{5,r-1|r-1} & 1 & \hat{x}_{1,r-1|r-1} & \hat{x}_{2,r-1|r-1} \\ \mathbb{E}\{|z_r|^2\} \hat{x}_{4,r-1|r-1} & \mathbb{E}\{|z_r|^2\} \hat{x}_{5,r-1|r-1} & \mathbb{E}\{|z_r|^2\} & \mathbb{E}\{|z_r|^2\} \hat{x}_{1,r-1|r-1} & \mathbb{E}\{|z_r|^2\} \hat{x}_{2,r-1|r-1} \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}, \quad (5.79)$$

lo que permite expresar \mathbf{F}'_{r-1} en la forma de (5.38) con

$$\Delta\bar{\mathbf{F}}_{r-1} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ \hat{x}_{4,r-1|r-1} & \hat{x}_{5,r-1|r-1} & 1 & \hat{x}_{1,r-1|r-1} & \hat{x}_{2,r-1|r-1} \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}. \quad (5.80)$$

Por otro lado, la media de \mathbf{u}_{r-1} es

$$\bar{\mathbf{u}}_{r-1} = \bar{\mathbf{f}}(\hat{\mathbf{x}}_{r-1|r-1}, \mathbf{0}) - \bar{\mathbf{F}}_{r-1}\hat{\mathbf{x}}_{r-1|r-1}, \quad (5.81)$$

donde

$$\bar{\mathbf{f}}(\hat{\mathbf{x}}_{r-1|r-1}, \mathbf{0}) = \mathbb{E}\{\mathbf{f}(\hat{\mathbf{x}}_{r-1|r-1}, \mathbf{0})\} = \begin{bmatrix} \hat{x}_{3,r-1} + \hat{x}_{4,r-1}\hat{x}_{1,r-1} + \hat{x}_{5,r-1}\hat{x}_{2,r-1} \\ \mathbb{E}\{|z_r|^2\}(\hat{x}_{3,r-1} + \hat{x}_{4,r-1}\hat{x}_{1,r-1} + \hat{x}_{5,r-1}\hat{x}_{2,r-1}) \\ \hat{x}_{3,r-1} \\ \hat{x}_{4,r-1} \\ \hat{x}_{5,r-1} \end{bmatrix}. \quad (5.82)$$

A partir de (5.81) \mathbf{u}'_{r-1} se puede escribir como

$$\mathbf{u}'_{r-1} = \mathbf{f}'(\hat{\mathbf{x}}_{r-1|r-1}, \mathbf{0}) - \mathbf{F}'_{r-1}\hat{\mathbf{x}}_{r-1|r-1}, \quad (5.83)$$

donde $\mathbf{f}'(\hat{\mathbf{x}}_{r-1|r-1}, \mathbf{0}) = \mathbf{f}(\hat{\mathbf{x}}_{r-1|r-1}, \mathbf{0}) - \bar{\mathbf{f}}(\hat{\mathbf{x}}_{r-1|r-1}, \mathbf{0})$, que también puede factorizarse en la forma

$$\mathbf{f}'(\hat{\mathbf{x}}_{r-1|r-1}, \mathbf{0}) = (|z_r|^2 - \mathbb{E}\{|z_r|^2\})\Delta\bar{\mathbf{f}}_{r-1}, \quad (5.84)$$

con

$$\Delta\bar{\mathbf{f}}_{r-1} = \begin{bmatrix} 0 \\ \hat{x}_{3,r-1} + \hat{x}_{4,r-1}\hat{x}_{1,r-1} + \hat{x}_{5,r-1}\hat{x}_{2,r-1} \\ 0 \\ 0 \\ 0 \end{bmatrix}. \quad (5.85)$$

Finalmente (5.83) puede reescribirse en la forma de (5.39), siendo

$$\Delta\bar{\mathbf{u}}_{r-1} = \Delta\bar{\mathbf{f}}_{r-1} - \Delta\bar{\mathbf{F}}_{r-1}\hat{\mathbf{x}}_{r-1|r-1} = \begin{bmatrix} 0 \\ -(\hat{x}_{4,r-1}\hat{x}_{1,r-1} + \hat{x}_{5,r-1}\hat{x}_{2,r-1}) \\ 0 \\ 0 \\ 0 \end{bmatrix}. \quad (5.86)$$

Utilizando estas expresiones junto con las ecuaciones del estimador LMMSE estándar de la sección 5.2, se obtiene el denominado LMMSE extendido por su analogía con el filtro de Kalman. Una vez que se inicializan $\bar{\mathbf{x}}_0$, $\hat{\mathbf{x}}_{0|0}$, \mathbf{P}_0 y $\mathbf{M}_{0|0}$ para $r \geq 1$, los pasos del algoritmo se pueden resumir de la siguiente manera.

- 1) **Matrices del sistema:** evaluar las matrices $\bar{\mathbf{F}}_{r-1}$ y $\Delta\bar{\mathbf{F}}_{r-1}$ y los vectores $\bar{\mathbf{u}}_{r-1}$ y $\Delta\bar{\mathbf{u}}_{r-1}$ por medio de (5.79), (5.80), (5.81) y (5.86) respectivamente.
- 2) **Estimado a priori:** actualizar $\hat{y}_{r|r-1}$ a partir de (5.27) y $\hat{\mathbf{x}}_{r|r-1}$, que para este caso su expresión se reduce a

$$\hat{\mathbf{x}}_{r|r-1} = \mathbf{f}(\hat{\mathbf{x}}_{r-1|r-1}, \mathbf{0}), \quad (5.87)$$
 de reemplazar (5.81) en (5.26).
- 3) **Matriz de covarianza del error a priori:** calcular la matriz $\mathbf{M}_{r|r-1}$ por medio de (5.55).
- 4) **Ganancia:** evaluar el vector de ganancias \mathbf{k}_r utilizando la ecuación (5.34).
- 5) **Estimado a posteriori:** actualizar $\hat{\mathbf{x}}_{r|r}$ haciendo uso de (5.61).
- 6) **Matriz de covarianza del error a posteriori:** actualizar $\mathbf{M}_{r|r}$ por medio de (5.58).
- 7) **Media y valor cuadrático medio:** evaluar $\bar{\mathbf{x}}_r$ y \mathbf{P}_r utilizando las ecuaciones (5.10) y (5.60), respectivamente.

5.3.3. Simulaciones numéricas

Con el objetivo de estudiar el desempeño del estimador LMMSE extendido, se implementó el algoritmo y se analizó utilizando datos sintéticos. Al igual que en la sección 5.2.3 $z_r \sim \mathcal{N}(0, 1)$, lo que implica que $\gamma = 2$. Los valores de los coeficientes elegidos son $k = 0.5$, $\alpha_1 = 0.2$ y $\beta_1 = 0.3$. En este caso no se sumó ruido a las observaciones, debido a que el mismo es absorbido por el modelo GARCH. Sin embargo, se consideró un valor pequeño para la varianza del ruido de observación y se tomó el valor del error cuadrático medio de cada estado para los elementos de la diagonal de la matriz de covarianza \mathbf{Q}_r .

La varianza condicional se inicializó con el valor resultante de promediar 50 muestras del proceso elevados al cuadrado, tomadas previamente al registro de datos utilizado en la implementación del algoritmo. De este conjunto de datos previo al registro utilizado, se tomó la última muestra para inicializar el segundo estado. Por otra parte, el valor del estado correspondiente a k se inicializó con el valor utilizado con el mismo fin para la varianza condicional. Por último, en el caso de los estados correspondientes a α_1 y β_1 sus valores iniciales fueron elegidos al azar en el intervalo entre 10^{-3} y 10^{-1} . En la figura 5.5 se presentan los resultados obtenidos por medio de esta simulación para un conjunto de datos de 5000 muestras.

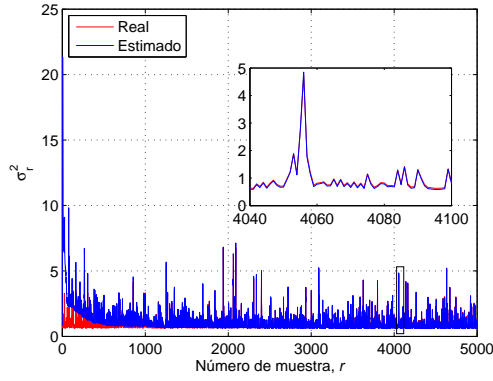
Las figuras 5.5a y 5.5b muestran los valores de la varianza condicional y el error cuadrático medio de la estimación. Se puede observar que rápidamente el valor de la varianza condicional estimado tiende a su verdadero valor reduciendo el error. Es importante destacar que si bien en este trabajo no se realizó un análisis de convergencia del método, este fue el comportamiento obtenido también en numerosos ensayos cuyos resultados no fueron incluidos. En la figura 5.5b se puede apreciar que el error se reduce, pero en determinados instantes presenta picos de

mayor valor, por ejemplo alrededor de la muestra $r = 2000$. Este fenómeno se produce por la naturaleza impulsiva del proceso y de la varianza condicional. Dependiendo de la realización, en este aspecto sí existen diferencias con otros ensayos, debido a que en algunos casos los picos ocurren con mayor frecuencia o bien poseen valores mayores, dependiendo de los valores de los coeficientes α_1 y β_1 .

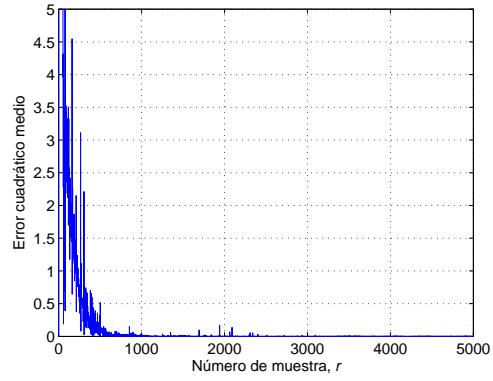
En las figuras 5.5c, 5.5d, 5.5e y 5.5f se presentan los estimados de los estados restantes junto con sus verdaderos valores. Notar que aún cuando la estimación de los coeficientes se aparta del verdadero valor, en este caso dicho efecto se aprecia particularmente en la estimación de α_1 , la aproximación de la varianza condicional es buena.

La idea que se persigue con este filtro no es lograr una estimación exacta en el valor de los coeficientes, sino obtener un estimador que permita actualizar sus valores ante variaciones de los mismos, manteniendo una estimación aceptable de la varianza condicional, que es el parámetro utilizado en la detección. Por este motivo se realizaron simulaciones con realizaciones del proceso generadas con variaciones en los parámetros. En particular, en la figura 5.6 se presentan los resultados de una simulación realizada empleando un registro de datos sintéticos de 30000 muestras. En la generación de los datos se varió el coeficiente k a partir de la muestra 5000, desde el valor inicial 0.5 a 2 por medio de una ley exponencial.

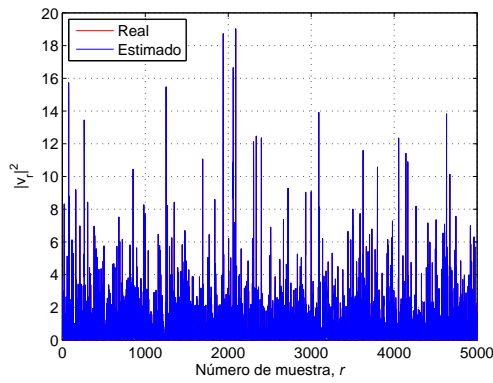
Notar que si bien el valor estimado del coeficiente α_1 no toma valores cercanos al verdadero antes de la muestra 5000, el error en la estimación de la varianza condicional se reduce. A partir del instante en que comienza a variar el valor de k los estimados de los coeficientes de α_1 y β_1 se apartan de sus verdaderos valores para compensar su variación. Una vez que el estimado de k aproxima mejor su valor entonces también lo hacen los estimados de α_1 y β_1 . Durante el transitorio el error se incrementa y comienza a decrecer lentamente pasado el mismo.



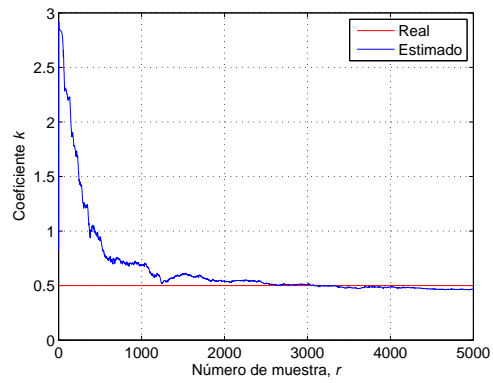
(a) Estado x_{1r} .



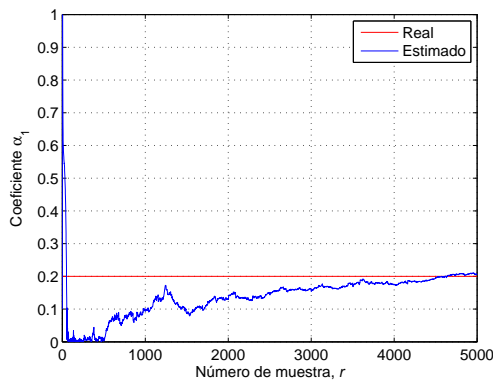
(b) MSE estado x_{1r} .



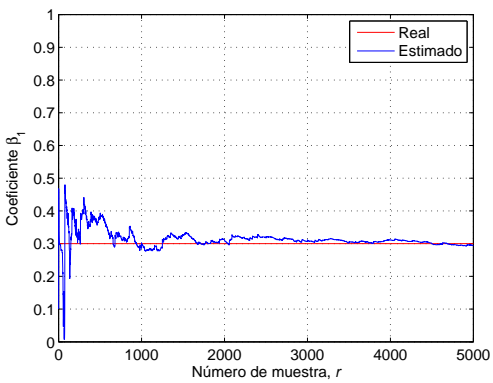
(c) Estado x_{2r} .



(d) Estado x_{3r} .



(e) Estado x_{4r} .



(f) Estado x_{5r} .

Figura 5.5: Resultados de la implementación del estimador LMMSE extendido empleando datos sintéticos.

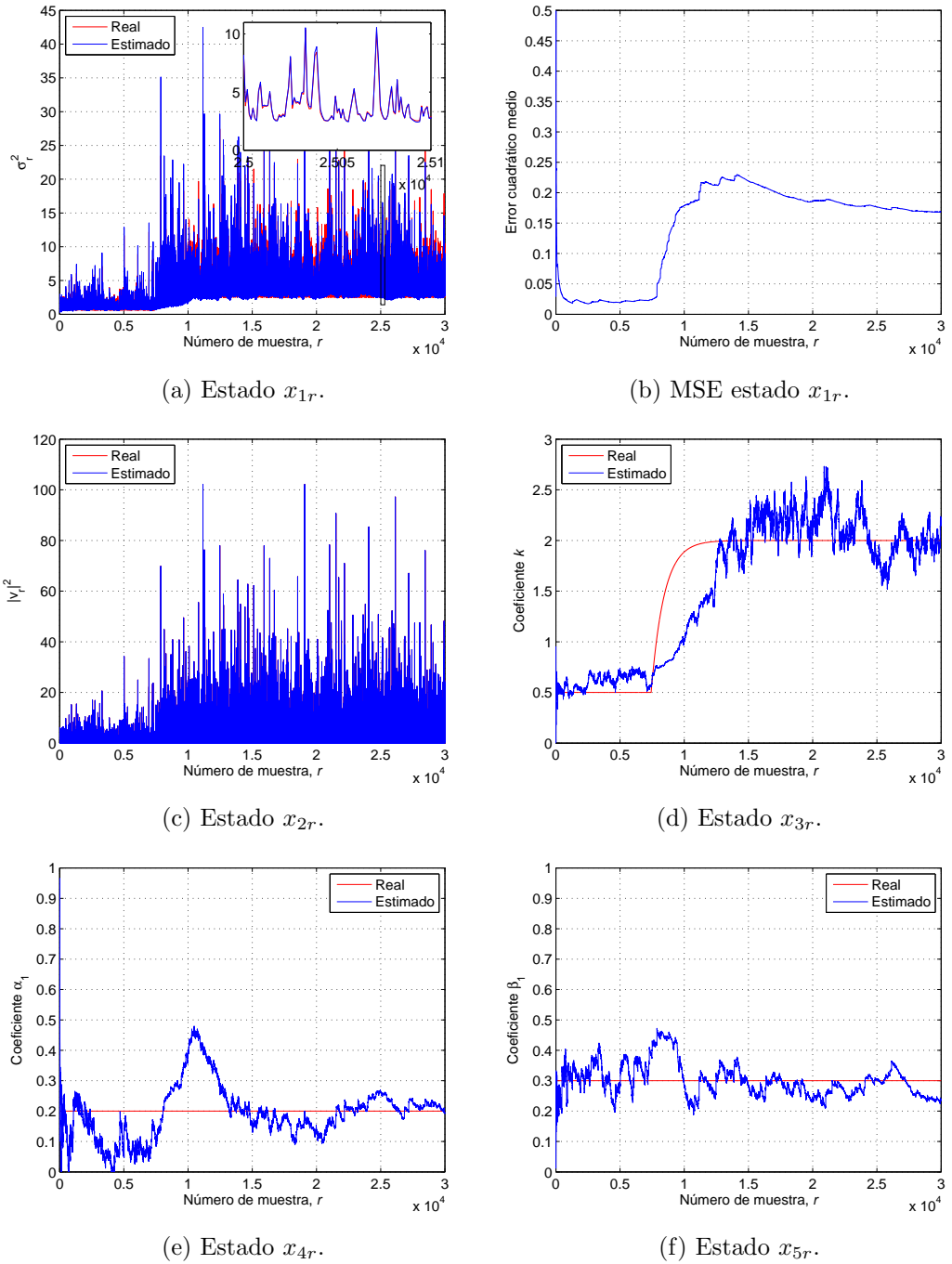


Figura 5.6: Resultados de la implementación del estimador LMMSE extendido con variaciones en el parámetro k .

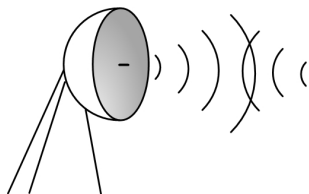
5.4. Conclusiones

En este capítulo se propuso un algoritmo de estimación secuencial para los parámetros de un proceso GARCH. El mismo es un estimador lineal de mínimo error cuadrático medio que surge a partir del enfoque de estimación Bayesiano.

En primer lugar se consideró el caso en que los coeficientes del proceso son conocidos. Se desarrollaron las ecuaciones del modelo de estados, las cuales presentan la particularidad de poseer parámetros aleatorios. A través de la interpretación geométrica del estimador LMMSE se dedujeron las ecuaciones que describen el algoritmo. Las matrices del sistema que resultan del modelo GARCH presentan una estructura especial. Por este motivo se dan tanto las expresiones generales para un estimador de este tipo como las particulares para el caso de interés.

Mediante simulaciones numéricas se verificó el desempeño del estimador utilizando datos sintéticos para un proceso GARCH(1,1). Si bien la hipótesis de coeficientes conocidos no es realista, este estimador sienta las bases para la deducción del estimador cuando se considera que no los son. Alternativamente, puede ser posible estudiar la estadística de las innovaciones. En el caso del filtro de Kalman un valor incorrecto de las matrices del sistema producen una desviación en la estadística de las innovaciones respecto de su comportamiento teórico (Simon, 2006). Un análisis similar para el estimador LMMSE de la varianza condicional de un proceso GARCH podría dar un mecanismo para decidir el momento en que es necesario actualizar los estimados de los coeficientes.

Asimismo, se presentó la versión extendida del estimador LMMSE. Análogamente al filtro de Kalman, éste aparece como solución a la estimación de estados en un sistema no lineal, a partir de su linealización empleando el desarrollo en serie de Taylor truncada al término de primer orden. En el caso de los procesos GARCH el modelo de estados no lineal surge al considerar que los coeficientes son desconocidos y los mismos se agregan al vector de estados para estimarlos conjuntamente con la varianza condicional. También se deducen las fórmulas que conforman el algoritmo del estimador LMMSE extendido y se verificó su comportamiento mediante simulaciones numéricas. Para este último caso se hicieron dos experiencias. En la primera se generó la realización de datos utilizando un conjunto de coeficientes constantes, mientras que en la segunda se varió uno de los coeficientes de modo de comprobar que el algoritmo efectivamente es capaz de seguir dichas variaciones. Es importante destacar que si bien se obtuvieron resultados de este tipo con otros conjuntos de datos, no se realizaron análisis de convergencia de estos estimadores.



Capítulo 6

Conclusiones

En esta tesis se presentaron diferentes aportes relacionados con el modelado del clutter y con el desarrollo de algoritmos de detección en aplicaciones de radar.

En primer lugar se propuso la utilización de los procesos GARCH para el modelado del clutter. Estos procesos poseen la desventaja de no contar con una expresión explícita para su función densidad de probabilidad. Por este motivo, se comenzó estudiando el método de estimación de cuasi-máxima verosimilitud, habitualmente utilizado en la estimación de sus coeficientes. Con el objetivo de evaluar el desempeño del estimador QMLE se dedujo una aproximación de la cota de Cramér-Rao para su matriz de covarianza. La misma se basa en la utilización de la función densidad de probabilidad de máxima entropía obtenida a partir de un conjunto de momentos conocidos. Los resultados obtenidos fueron validados mediante un ejemplo numérico. Sin embargo, existe una restricción sobre los valores que pueden tomar los coeficientes de un proceso GARCH para que sus momentos sean finitos, lo que limita la utilización de esta herramienta.

Debido a la estructura que poseen los procesos GARCH, se los utilizó considerando que sus realizaciones modelan los datos tomados en rango. Esto presenta el inconveniente que las muestras son insuficientes para alcanzar un error de estimación aceptable. Por este motivo, se propuso un método alternativo para la estimación de sus coeficientes. Se estudiaron las propiedades estadísticas de este estimador mediante simulaciones numéricas y se compararon con las del procedimiento de estimación QMLE, mostrando que los métodos presentan un desempeño similar, ambos muestran un comportamiento asintóticamente consistente. Posteriormente, empleando el método de estimación alternativo, se realizó el ajuste de un proceso

GARCH a mediciones reales de clutter marítimo. De la comparación entre los histogramas del proceso y de las mediciones se observó que el modelo representa adecuadamente las colas pesadas que presenta la densidad de probabilidad de los datos.

Basado en este modelo para el clutter se desarrolló un esquema de detección. Para ello se utilizó una versión adaptada del test de la relación de verosimilitud, la cual consistió simplemente en reemplazar las funciones de densidad para cada hipótesis por densidades de probabilidad condicionales, dadas las muestras anteriores del proceso. El detector resultante presenta la característica de poseer el umbral de detección adaptivo, debido a que depende de la varianza instantánea del proceso. Se dedujo una expresión para la probabilidad de falsa alarma y se evaluó la probabilidad de detección a través de simulaciones numéricas debido a que no es posible hallarla analíticamente.

El detector GARCH no es de falsa alarma constante respecto de los coeficientes del proceso que modela el clutter. Por ello, se realizó un estudio del efecto de variaciones en los valores de los coeficientes sobre las probabilidades de falsa alarma y de detección. El mismo consistió en perturbar los valores de los primeros y evaluar dichas probabilidades. El orden de las variaciones de los coeficientes se fijó en base a los errores de estimación obtenidos para un conjunto de datos del mismo tamaño que el de las mediciones disponibles. Los resultados de este análisis arrojaron que el detector presenta un comportamiento robusto en una situación práctica.

Se estudió también el desempeño del detector GARCH utilizando datos reales y se los comparó con el desempeño de dos detectores del mismo tipo, uno basado en una distribución Gaussiana para la amplitud de los datos y el otro diseñado empleando una distribución Weibull para su intensidad. Para ello, se evaluaron las probabilidades de detección en función de la relación señal a clutter y las curvas ROC. Se mostró que el detector GARCH presenta un mejor desempeño que los otros detectores, excepto para valores altos de SCR y alta probabilidad de falsa alarma. Bajo estas condiciones favorables los tres detectores alcanzan aproximadamente la misma probabilidad de detección. El éxito del detector GARCH se puede explicar a través de su variación temporal. Tomando la intensidad de la muestra como el estadístico para los tres detectores, en el caso del detector GARCH el umbral en cada comparación es función de la intensidad de la muestra previa, pero los parámetros del modelo no cambian rápidamente. Este comportamiento no puede obtenerse con los detectores clásicos en los que existe una relación de compromiso entre la calidad en la estimación de los parámetros de la distribución y la velocidad de adaptación.

En segundo lugar, se propuso un modelo para el clutter al que se denominó AR-GARCH-2D. Este modelo combina un proceso GARCH en dos dimensiones con un proceso autorregresivo, con lo que conserva las características impulsivas de los procesos GARCH y permite modelar la correlación pulso a pulso que presentan los datos. Se dedujo un algoritmo de estimación para los coeficientes de este tipo de procesos, con un procedimiento análogo al utilizado para los procesos GARCH. Por medio de simulaciones numéricas se analizaron las propiedades de dichos estimadores. Finalmente, se realizó el ajuste de dicho modelo a un conjunto de datos disponible.

También se derivó un algoritmo de detección utilizando el modelo AR-GARCH-2D para el clutter. La ventaja de este detector frente al anterior es la posibilidad de incorporar múltiples pulsos en el instante de decisión. Se obtuvo una expresión para su probabilidad de falsa alarma y debido a su complejidad matemática la probabilidad de detección se evaluó mediante simulaciones numéricas, empleando un proceso representativo de la familia. De la misma forma que para el caso GARCH, se estudió la sensibilidad en el desempeño de este detector ante variaciones en los valores de los coeficientes. Este esquema de detección tampoco es CFAR respecto a sus parámetros, pero también mostró un comportamiento robusto en una situación práctica. Además, se comparó el desempeño del detector AR-GARCH-2D con el de los detectores GLRT, ALQ y ARGLR utilizando datos sintéticos. Como era de esperar, con este tipo de datos el detector AR-GARCH-2D presentó un desempeño superior que los demás detectores.

El ajuste de los procesos AR-GARCH-2D a los datos reales arrojó como resultado una familia de modelos para el clutter, con distintos órdenes. Previamente a la utilización de un criterio de selección de modelos para elegir aquel que presente el compromiso óptimo entre complejidad y calidad del ajuste, se analizaron comparativamente sus desempeños. Se mostró que los detectores para los cuales el orden de la parte autorregresiva, M , es 2 o 3 presentan una mayor probabilidad de detección con una tasa de falsa alarma menor que los detectores para los cuales M es 0 o 1, además de un mejor ajuste de la probabilidad de falsa alarma en relación a la predicción teórica. Finalmente se eligió uno de los modelos empleando el criterio de información Bayesiano (BIC).

Además, se comparó el desempeño del detector AR-GARCH-2D elegido con el de los detectores GLRT, ALQ y ARGLR utilizando datos reales de clutter. Se mostró que el detector propuesto supera a los otros tres detectores. Sin embargo, el modelo AR-GARCH-2D resultante para el conjunto de datos del que se dispone, presenta la característica de no poseer colas pesadas representando de manera inexacta la densidad de los datos. Por este motivo, se hizo un análisis de sus propiedades y se comparó con un detector asociado a un modelo autorregresivo. Ambos detectores presentaron desempeños similares. Esto muestra la flexibilidad de los procesos AR-GARCH-2D, que se adapta a este modelo mas sencillo sin mermar desempeño.

Los resultados obtenidos permiten concluir que cuando se estudian esquemas de detección en radar la evaluación de los modelos para los datos debe llevarse a cabo no solo en base a la calidad del ajuste, sino también teniendo en cuenta el desempeño de los detectores asociados a dichos modelos.

Por último, se desarrolló un método para estimar secuencialmente la varianza condicional de los procesos GARCH, con el objetivo de contar con una herramienta que permita actualizar la varianza en cada instante de decisión de los detectores. En primer lugar se supuso que los coeficientes son conocidos y se describieron las ecuaciones que describen el proceso como un modelo de estados, con la particularidad que posee parámetros aleatorios. Se dedujo un estimador lineal de mínimo error cuadrático medio para la varianza condicional de manera análoga al filtro de Kalman. La hipótesis de coeficientes conocidos en la deducción del estimador LMMSE no es realista desde un punto de vista práctico, lo que implica que el interés del

planteo presentado en este punto es de carácter teórico. Sin embargo, en primera instancia se formuló esa hipótesis para dilucidar aspectos desconocidos aún en el caso de coeficientes conocidos, debido a que el problema se aparta del filtro de Kalman tradicional y como paso previo a complejizar el algoritmo incorporando la estimación de los parámetros del modelo.

En una segunda etapa, se planteó el caso general en el que se consideró que los coeficientes son desconocidos. En estas condiciones el sistema resultante, además de poseer parámetros aleatorios, es no lineal. Se dedujeron las expresiones para el estimador LMMSE que permite la estimación conjunta de la varianza condicional y los coeficientes del proceso. Por medio de simulaciones numéricas se evaluó su funcionamiento para dos situaciones particulares. En el primer caso los valores de los coeficientes se mantuvieron constantes y en el segundo se perturbó uno de ellos. Los resultados obtenidos muestran que siguen las variaciones de los coeficientes y aún cuando sus valores no son estimados de manera exacta, los mismos se adaptan de manera de estimar correctamente la varianza condicional, reduciendo su error de estimación. Es importante destacar que no se hicieron análisis de convergencia para ninguno de los estimadores LMMSE propuestos y que los estudios se redujeron a ensayos numéricos.

Existen líneas de investigación que pueden continuarse más allá de los resultados mostrados en esta tesis, así como resultados parciales que requieren mayor confirmación teórica y experimental.

Una posible extensión del trabajo presentado sería el desarrollo de un esquema de detección que permita incorporar la información polarimétrica en las decisiones. Esto puede llevarse a cabo combinando modelos AR-GARCH-2D en cada polarización a través de una estructura del tipo de la de los procesos GARCH multivariados (Bauwens *et al.*, 2006).

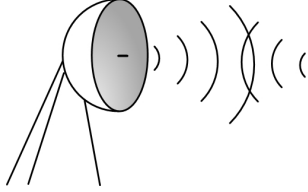
El método propuesto para la estimación secuencial de la varianza condicional se desarrolló para un proceso GARCH convencional. Sin embargo de acuerdo con la estructura de los datos reales no puede ser utilizado directamente con los esquemas de detección desarrollados en la tesis. En este aspecto dicho estimador debería adaptarse, en particular para su uso con un modelo para el clutter del tipo AR-GARCH-2D, donde además de los coeficientes de la varianza condicional aparecen los parámetros de la parte autorregresiva.

Por otro lado, sería deseable desarrollar una demostración teórica de la convergencia de los estimadores lineales de mínimo error cuadrático medio para la varianza condicional, que hasta el momento solo se verificó numéricamente. Una alternativa sería la de adaptar la propuesta de Ljung (1999) para modelos lineales con parámetros constantes al caso de matrices de sistema aleatorias.

De los resultados presentados en los análisis con el conjunto particular de datos reales utilizados, se desprende que el parámetro que domina el desempeño de los detectores es la correlación de los datos en tiempo lento, por sobre la caracterización de la impulsividad que presenta el clutter. Sin embargo la escasa disponibilidad de conjuntos de datos reales útiles para los experimentos que deberían realizarse, no permite establecer afirmaciones concluyentes. En este sentido es deseable ampliar dicho estudio a diferentes conjuntos de mediciones de modo de realizar un análisis más abarcativo del tema.

Los procesos GARCH también podrían utilizarse en radares de apertura sintética (SAR). Existen diferentes modelos para las imágenes SAR, algunos basados en principios físicos y otros basados en enfoques matemáticos o que surgen a partir del análisis de mediciones. Un modelo que combina ambos enfoques es el modelo de clutter multiplicativo (Oliver y Quegan, 2004). En este caso los datos son modelados como el producto de dos variables aleatorias. La primera se denomina *speckle* y surge de la contribución de numerosos reflectores aleatorios dentro de una celda de resolución. La otra se denomina *textura* y se puede interpretar como las fluctuaciones de la sección transversal de radar (*radar cross section*). En general, el *speckle* se modela mediante una distribución Gaussiana, mientras que la *textura* es modelada utilizando diversas distribuciones (Anastassopoulos *et al.*, 1999) o de manera más general como la contribución de una mezcla de distribuciones (Oliver y Quegan, 2004).

El modelo GARCH-2D puede ser interpretado como un caso particular de modelo multiplicativo dinámico dado que se define como el producto entre un proceso Gaussiano y un proceso aleatorio independiente, cuyas muestras al cuadrado se modelan como una serie temporal autorregresiva. En este aspecto sería de interés estudiar la representación de datos SAR a través de procesos de tipo GARCH-2D y diseñar algoritmos de clasificación a partir de ellos. Para llevar a cabo esto último una alternativa es adecuar el método de clasificación para imágenes a través del criterio de máximo a posteriori (MAP) por medio del algoritmo EM (Expectation-Maximization) basado en un modelo de mezcla de procesos GARCH-2D para los datos (Pascual *et al.*, 2014b).



Apéndice A

Gradientes de los funcionales utilizados

En este apéndice se deducen las expresiones de los funcionales utilizados en los problemas de optimización para obtener los estimados de cuasi-máxima verosimilitud.

A.1. Gradiente de $\mathcal{L}_2(\boldsymbol{\theta})$

El funcional $\mathcal{L}_2(\boldsymbol{\theta})$ es utilizado en el capítulo 3 para estimar los coeficientes de un proceso GARCH. El mismo está dado por

$$\mathcal{L}_2(\boldsymbol{\theta}) = \sum_{r=1}^R \left[\ln(\sigma_r^2) + \frac{|v_r|^2}{\sigma_r^2} \right], \quad (\text{A.1})$$

con $\boldsymbol{\theta} = [k \ \alpha_1 \ \cdots \ \alpha_p \ \beta_1 \ \cdots \ \beta_q]^T$. Entonces, su gradiente es un vector que puede escribirse como

$$\nabla \mathcal{L}_2(\boldsymbol{\theta}) = \frac{\partial \mathcal{L}_2(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} = \sum_{r=1}^R \frac{1}{\sigma_r^2} \frac{\partial \sigma_r^2}{\partial \boldsymbol{\theta}} \left[1 + \frac{|v_r|^2}{\sigma_r^2} \right], \quad (\text{A.2})$$

de dimensión $(p + q + 1) \times 1$, donde $\frac{\partial \sigma_r^2}{\partial \boldsymbol{\theta}} = \left[\frac{\partial \sigma_r^2}{\partial k} \frac{\partial \sigma_r^2}{\partial \alpha_1} \cdots \frac{\partial \sigma_r^2}{\partial \alpha_p} \frac{\partial \sigma_r^2}{\partial \beta_1} \cdots \frac{\partial \sigma_r^2}{\partial \beta_q} \right]^T$, con

$$\frac{\partial \sigma_r^2}{\partial k} = 1 + \sum_{j=1}^p \alpha_j \frac{\partial \sigma_{r-j}^2}{\partial k} \quad (\text{A.3})$$

$$\frac{\partial \sigma_r^2}{\partial \alpha_i} = \sigma_{r-i}^2 + \sum_{j=1}^p \alpha_j \frac{\partial \sigma_{r-j}^2}{\partial \alpha_i}, \quad i = 1, \dots, p \quad (\text{A.4})$$

$$\frac{\partial \sigma_r^2}{\partial \beta_i} = |v_{r-i}|^2 + \sum_{j=1}^p \alpha_j \frac{\partial \sigma_{r-j}^2}{\partial \beta_i}, \quad i = 1, \dots, q. \quad (\text{A.5})$$

A.2. Gradiente de $\mathcal{L}_3(\boldsymbol{\theta})$

$\mathcal{L}_3(\boldsymbol{\theta})$ es el funcional que se minimiza para hallar los estimados de los coeficientes de un proceso GARCH empleando varias realizaciones. Su expresión es

$$\mathcal{L}_3(\boldsymbol{\theta}) = \sum_{t=1}^N \sum_{r=1}^R \left[\ln(\sigma_r^2) + \frac{|v_r|^2}{\sigma_r^2} \right], \quad (\text{A.6})$$

donde $\boldsymbol{\theta} = [k \ \alpha_1 \ \cdots \ \alpha_p \ \beta_1 \ \cdots \ \beta_q]^T$. El gradiente de $\mathcal{L}_3(\boldsymbol{\theta})$ también es de dimensión $(p+q+1) \times 1$ y su expresión es

$$\nabla \mathcal{L}_3(\boldsymbol{\theta}) = \frac{\partial \mathcal{L}_3(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} = \sum_{t=1}^N \sum_{r=1}^R \frac{1}{\sigma_r^2} \frac{\partial \sigma_r^2}{\partial \boldsymbol{\theta}} \left[1 + \frac{|v_r|^2}{\sigma_r^2} \right]. \quad (\text{A.7})$$

En este caso $\frac{\partial \sigma_r^2}{\partial \boldsymbol{\theta}}$ es de la misma forma que la del gradiente de la sección anterior y sus elementos se obtienen a partir de ecuaciones análogas a (A.3), (A.4) y (A.5).

A.3. Gradiente de $\mathcal{L}_4(\boldsymbol{\theta})$

El funcional $\mathcal{L}_4(\boldsymbol{\theta})$ es utilizado en el capítulo 4 para estimar los coeficientes de un proceso AR-GARCH-2D. Su expresión está dada por

$$\mathcal{L}_4(\boldsymbol{\theta}) = \sum_{r=1}^R \sum_{t=1}^T \left[\ln(\sigma_{rt}^2) + \frac{1}{\sigma_{rt}^2} \left| v_{rt} - \sum_{j=1}^M a_{0j} v_{r,t-j} \right|^2 \right], \quad (\text{A.8})$$

donde $\boldsymbol{\theta} = [k \ \alpha_{10} \ \alpha_{11} \ \beta_{10} \ \beta_{11} \ \text{Re}\{a_{01}\} \ \text{Im}\{a_{01}\} \ \cdots \ \text{Re}\{a_{0M}\} \ \text{Im}\{a_{0M}\}]^T$ es el vector de parámetros que se desea estimar, de dimensión $(5+2M) \times 1$. Luego, el gradiente de $\mathcal{L}_4(\boldsymbol{\theta})$ está dado por $\nabla \mathcal{L}_4(\boldsymbol{\theta}) = \frac{\partial \mathcal{L}_4(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}} = \left[\frac{\partial \mathcal{L}_4}{\partial k} \ \frac{\partial \mathcal{L}_4}{\partial \alpha_{10}} \ \frac{\partial \mathcal{L}_4}{\partial \alpha_{11}} \ \frac{\partial \mathcal{L}_4}{\partial \beta_{10}} \ \frac{\partial \mathcal{L}_4}{\partial \beta_{11}} \ \cdots \ \frac{\partial \mathcal{L}_4}{\partial \text{Re}\{a_{01}\}} \ \frac{\partial \mathcal{L}_4}{\partial \text{Im}\{a_{01}\}} \ \cdots \ \frac{\partial \mathcal{L}_4}{\partial \text{Re}\{a_{0M}\}} \ \frac{\partial \mathcal{L}_4}{\partial \text{Im}\{a_{0M}\}} \right]^T$.

Llamando θ_i al elemento i -ésimo de $\boldsymbol{\theta}$, es decir $[\boldsymbol{\theta}]_i = \theta_i$. Entonces para $i = 1, 2, 3, 4, 5$ los elementos de $\nabla \mathcal{L}_4(\boldsymbol{\theta})$ se pueden escribir de la siguiente manera

$$\frac{\partial \mathcal{L}_4(\boldsymbol{\theta})}{\partial \theta_i} = \sum_{r=1}^R \sum_{t=1}^T \frac{1}{\sigma_{rt}^2} \frac{\partial \sigma_{rt}^2}{\partial \theta_i} \left[1 + \frac{1}{\sigma_{rt}^2} \left| v_{rt} - \sum_{j=1}^M a_{0j} v_{r,t-j} \right|^2 \right], \quad (\text{A.9})$$

donde

$$\frac{\partial \sigma_{rt}^2}{\partial \theta_1} = \frac{\partial \sigma_{rt}^2}{\partial k} = 1 + \alpha_{10} \frac{\partial \sigma_{r-1,t}^2}{\partial k} + \alpha_{11} \frac{\partial \sigma_{r-1,t-1}^2}{\partial k}, \quad (\text{A.10})$$

$$\frac{\partial \sigma_{rt}^2}{\partial \theta_2} = \frac{\partial \sigma_{rt}^2}{\partial \alpha_{10}} = \sigma_{r-1,t}^2 + \alpha_{10} \frac{\partial \sigma_{r-1,t}^2}{\partial \alpha_{10}} + \alpha_{11} \frac{\partial \sigma_{r-1,t-1}^2}{\partial \alpha_{10}}, \quad (\text{A.11})$$

$$\frac{\partial \sigma_{rt}^2}{\partial \theta_3} = \frac{\partial \sigma_{rt}^2}{\partial \alpha_{11}} = \sigma_{r-1,t-1}^2 + \alpha_{10} \frac{\partial \sigma_{r-1,t}^2}{\partial \alpha_{11}} + \alpha_{11} \frac{\partial \sigma_{r-1,t-1}^2}{\partial \alpha_{11}}, \quad (\text{A.12})$$

$$\frac{\partial \sigma_{rt}^2}{\partial \theta_4} = \frac{\partial \sigma_{rt}^2}{\partial \beta_{10}} = |v_{r-1,t}|^2 + \alpha_{10} \frac{\partial \sigma_{r-1,t}^2}{\partial \beta_{10}} + \alpha_{11} \frac{\partial \sigma_{r-1,t-1}^2}{\partial \beta_{10}}, \quad (\text{A.13})$$

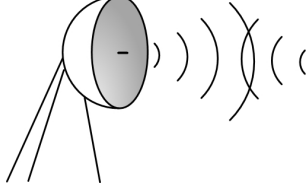
$$\frac{\partial \sigma_{rt}^2}{\partial \theta_5} = \frac{\partial \sigma_{rt}^2}{\partial \beta_{11}} = |v_{r-1,t-1}|^2 + \alpha_{10} \frac{\partial \sigma_{r-1,t}^2}{\partial \beta_{11}} + \alpha_{11} \frac{\partial \sigma_{r-1,t-1}^2}{\partial \beta_{11}}. \quad (\text{A.14})$$

Mientras que cuando $i = 5 + 2n - 1$ con $n = 1, \dots, M$, es decir par y mayor que 5, los elementos de $\nabla \mathcal{L}_4(\boldsymbol{\theta})$ son de la forma

$$\frac{\partial \mathcal{L}_4(\boldsymbol{\theta})}{\partial \theta_i} = \frac{\partial \mathcal{L}_4(\boldsymbol{\theta})}{\partial \text{Re}\{a_{0n}\}} = - \sum_{r=1}^R \sum_{t=1}^T \frac{2}{\sigma_{rt}^2} \text{Re} \left\{ v_{r,t-n} \left(v_{rt} - \sum_{j=1}^M a_{0j} v_{r,t-j} \right)^* \right\}, \quad (\text{A.15})$$

y, finalmente, si $i = 5 + 2n$ con $n = 1, \dots, M$, es decir impar y mayor que 5, resultan

$$\frac{\partial \mathcal{L}_4(\boldsymbol{\theta})}{\partial \theta_i} = \frac{\partial \mathcal{L}_4(\boldsymbol{\theta})}{\partial \text{Im}\{a_{0n}\}} = \sum_{r=1}^R \sum_{t=1}^T \frac{2}{\sigma_{rt}^2} \text{Im} \left\{ v_{r,t-n} \left(v_{rt} - \sum_{j=1}^M a_{0j} v_{r,t-j} \right)^* \right\}. \quad (\text{A.16})$$



Apéndice B

Condición de varianza finita para proceso AR-GARCH-2D

En este apéndice se presenta la deducción para la varianza no condicional, dada por (4.9), de un proceso AR-GARCH-2D de la forma de (4.6) y (4.7) con $M = 2$.

El proceso considerado puede escribirse como

$$v_{rt} = \sigma_{rt}z_{rt} + a_{01}v_{r,t-1} + a_{02}v_{r,t-2}, \quad z_{rt} \sim \mathcal{CN}(0, 1) \text{ i.i.d.}, \quad (\text{B.1})$$

$$\sigma_{rt}^2 = k + \alpha_{10}\sigma_{r-1,t}^2 + \alpha_{11}\sigma_{r-1,t-1}^2 + \beta_{10}|v_{r-1,t}|^2 + \beta_{11}|v_{r-1,t-1}|^2. \quad (\text{B.2})$$

Luego, por la propiedad de linealidad de la esperanza, la media de v_{rt} está dada por

$$\mu_v = \mathbb{E}\{v_{rt}\} = \mathbb{E}\{\sigma_{rt}z_{rt}\} + a_{01}\mathbb{E}\{v_{r,t-1}\} + a_{02}\mathbb{E}\{v_{r,t-2}\}. \quad (\text{B.3})$$

Teniendo en cuenta que σ_{rt} y z_{rt} son independientes y bajo la hipótesis de estacionaridad en sentido amplio resulta

$$\mu_v = \mathbb{E}\{\sigma_{rt}\}\mathbb{E}\{z_{rt}\} + a_{01}\mu_v + a_{02}\mu_v. \quad (\text{B.4})$$

Sabiendo que $\mathbb{E}\{z_{rt}\} = 0$, a partir de (B.4) se obtiene que $\mu_v = \mathbb{E}\{v_{rt}\} = 0$.

En estas condiciones la varianza de v_{rt} es $\sigma_v^2 = \mathbb{E}\{|v_{rt}|^2\} = \mathbb{E}\{v_{rt}v_{rt}^*\}$, es decir

$$\sigma_v^2 = \mathbb{V}\{v_{rt}\} = \mathbb{E}\{[(a_{01}v_{r,t-1} + a_{02}v_{r,t-2}) + \sigma_{rt}z_{rt}][(a_{01}v_{r,t-1} + a_{02}v_{r,t-2})^* + \sigma_{rt}z_{rt}^*]\}. \quad (\text{B.5})$$

Distribuyendo el producto dentro de la esperanza y haciendo uso su propiedad de linealidad,

se puede escribir

$$\begin{aligned} \sigma_v^2 = & \mathbb{E}\{|a_{01}v_{r,t-1} + a_{02}v_{r,t-2}|^2\} + \mathbb{E}\{\sigma_{rt}z_{rt}(a_{01}v_{r,t-1} + a_{02}v_{r,t-2})^*\} + \\ & + \mathbb{E}\{(a_{01}v_{r,t-1} + a_{02}v_{r,t-2})\sigma_{rt}z_{rt}^*\} + \mathbb{E}\{\sigma_{rt}^2|z_{rt}|^2\}. \end{aligned} \quad (\text{B.6})$$

Teniendo en cuenta que z_{rt} es independiente de σ_{rt} , de $v_{r,t-1}$ y de $v_{r,t-2}$ y además como su media es nula, es decir $\mathbb{E}\{z_{rt}\} = 0$, entonces los términos cruzados en (B.6) son cero. Asimismo, en virtud de que $\mathbb{E}\{|z_{rt}|^2\} = 1$, (B.6) se reduce a

$$\sigma_v^2 = \mathbb{E}\{|a_{01}v_{r,t-1} + a_{02}v_{r,t-2}|^2\} + \mathbb{E}\{\sigma_{rt}^2\}. \quad (\text{B.7})$$

Tomando esperanza de la ecuación (B.2) se obtiene $\mathbb{E}\{\sigma_{rt}^2\}$. Haciendo uso de la propiedad de linealidad, esta se puede escribir de la siguiente manera

$$\mathbb{E}\{\sigma_{rt}^2\} = k + \alpha_{10}\mathbb{E}\{\sigma_{r-1,t}^2\} + \alpha_{11}\mathbb{E}\{\sigma_{r-1,t-1}^2\} + \beta_{10}\mathbb{E}\{|v_{r-1,t}|^2\} + \beta_{11}\mathbb{E}\{|v_{r-1,t-1}|^2\}. \quad (\text{B.8})$$

Considerando nuevamente la hipótesis de estacionareidad, de (B.8) se obtiene la expresión para $\mathbb{E}\{\sigma_{rt}^2\}$, la cual está dada por

$$\mathbb{E}\{\sigma_{rt}^2\} = \frac{k + (\beta_{10} + \beta_{11})\sigma_v^2}{1 - \alpha_{10} - \alpha_{11}}. \quad (\text{B.9})$$

Para obtener σ_v^2 resta calcular $\mathbb{E}\{|a_{01}v_{r,t-1} + a_{02}v_{r,t-2}|^2\}$. Este término puede expresarse como

$$\mathbb{E}\{|a_{01}v_{r,t-1} + a_{02}v_{r,t-2}|^2\} = \mathbb{E}\{(a_{01}v_{r,t-1} + a_{02}v_{r,t-2})(a_{01}v_{r,t-1} + a_{02}v_{r,t-2})^*\}. \quad (\text{B.10})$$

Distribuyendo el producto dentro de la esperanza

$$\begin{aligned} \mathbb{E}\{|a_{01}v_{r,t-1} + a_{02}v_{r,t-2}|^2\} = & |a_{01}|^2\mathbb{E}\{|v_{r,t-1}|^2\} + a_{01}a_{02}^*\mathbb{E}\{v_{r,t-1}v_{r,t-2}^*\} + \\ & + a_{01}^*a_{02}\mathbb{E}\{v_{r,t-2}v_{r,t-1}^*\} + |a_{02}|^2\mathbb{E}\{|v_{r,t-2}|^2\}. \end{aligned} \quad (\text{B.11})$$

Teniendo en cuenta que la función de autocorrelación de v_{rt} en tiempo lento se define como

$$R_{vv}[\tau] = \mathbb{E}\{v_{r,t+\tau}v_{rt}^*\}, \quad (\text{B.12})$$

y que $\sigma_v^2 = \mathbb{E}\{|v_{r,t-1}|^2\} = \mathbb{E}\{|v_{r,t-2}|^2\}$, entonces

$$\mathbb{E}\{|a_{01}v_{r,t-1} + a_{02}v_{r,t-2}|^2\} = (|a_{01}|^2 + |a_{02}|^2)\sigma_v^2 + 2\text{Re}\{a_{01}a_{02}^*R_{vv}[1]\}. \quad (\text{B.13})$$

$R_{vv}[1]$ se obtiene reemplazando $v_{r,t-1}$ a partir de la relación que define el proceso (B.1), es decir

$$R_{vv}[1] = \mathbb{E}\{v_{r,t-1}v_{r,t-2}^*\} = \mathbb{E}\{(\sigma_{r,t-1}z_{r,t-1} + a_{01}v_{r,t-2} + a_{02}v_{r,t-3})v_{r,t-2}^*\}, \quad (\text{B.14})$$

y se puede reescribir como

$$R_{vv}[1] = \mathbb{E}\{\sigma_{r,t-1}z_{r,t-1}v_{r,t-2}^*\} + a_{01}\mathbb{E}\{|v_{r,t-2}|^2\} + a_{02}\mathbb{E}\{v_{r,t-3}v_{r,t-2}^*\}, \quad (\text{B.15})$$

donde el primer término se anula por los motivos ya enunciados a lo largo de la deducción, permitiendo expresar $R_{vv}[1]$ en la forma

$$R_{vv}[1] = a_{01}\sigma_v^2 + a_{02}R_{vv}[-1]. \quad (\text{B.16})$$

Debido a que la función de autocorrelación posee simetría Hermítica, luego

$$R_{vv}[-1] = R_{vv}^*[1] = a_{01}^*\sigma_v^2 + a_{02}^*R_{vv}[1]. \quad (\text{B.17})$$

Sustituyendo (B.17) en (B.16)

$$R_{vv}[1] = \frac{a_{01} + a_{02}a_{01}^*}{1 - |a_{02}|^2}\sigma_v^2. \quad (\text{B.18})$$

Reemplazando (B.18) en (B.13) resulta

$$\mathbb{E}\{|a_{01}v_{r,t-1} + a_{02}v_{r,t-2}|^2\} = \left[(|a_{01}|^2 + |a_{02}|^2) + 2\text{Re} \left\{ \frac{a_{01}a_{02}^*(a_{01} + a_{02}a_{01}^*)}{1 - |a_{02}|^2} \right\} \right] \sigma_v^2 = \phi_2\sigma_v^2, \quad (\text{B.19})$$

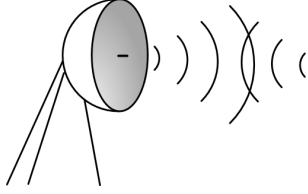
donde

$$\phi_2 = |a_{01}|^2 + |a_{02}|^2 + 2\text{Re} \left\{ \frac{a_{01}a_{02}^*(a_{01} + a_{01}^*a_{02})}{1 - |a_{02}|^2} \right\}. \quad (\text{B.20})$$

Finalmente, reemplazando (B.9) y (B.19) en (B.7) y operando algebraicamente se obtiene la expresión para la varianza no condicional de v_{rt}

$$\sigma_v^2 = \mathbb{V}\{v_{rt}\} = \frac{k}{1 - \alpha_{10} - \alpha_{11} - \beta_{10} - \beta_{11} - (1 - \alpha_{10} - \alpha_{11})\phi_2}. \quad (\text{B.21})$$

Empleando un procedimiento análogo al descrito en este apéndice es posible obtener expresiones de ϕ_M para otros valores de M , así como también la varianza no condicional del proceso v_{rt} .



Apéndice C

Modelo Autorregresivo

En este apéndice se deducen las expresiones para la estimación de los parámetros de un proceso autorregresivo, definido como

$$v_{rt} = \sigma z_{rt} + \sum_{j=1}^M a_{0j} v_{r,t-j}, \quad z_{rt} \sim \mathcal{CN}(0,1) \text{ i.i.d.}, \quad (\text{C.1})$$

donde a_{0j} son los coeficientes de la parte autorregresiva del proceso y la varianza σ^2 en este caso es constante.

C.1. Estimación de parámetros

Partiendo de una matriz de datos \mathbf{V} de dimensiones $R \times T$, que contiene R realizaciones independientes del proceso, cada una de T muestras de largo, los estimadores de los parámetros surgen de la maximización de la densidad condicional $f(\mathbf{V}|\mathbf{v}_{1-M}; \dots; \mathbf{v}_0)$, siguiendo un procedimiento análogo al utilizado para los procesos AR-GARCH-2D. Esta densidad puede expresarse como

$$f(\mathbf{V}|\mathbf{v}_{1-M}; \dots; \mathbf{v}_0) = \prod_{r=1}^R \prod_{t=1}^T f(v_{rt}|\{v_{t-j}\}_{j=1}^M), \quad (\text{C.2})$$

donde $\mathbf{v}_t = [v_{1t} \ v_{2t} \ \dots \ v_{Rt}]^T$ y las fdp's condicionales $f(v_{rt}|\{v_{t-j}\}_{j=1}^M)$ están dadas por

$$f(v_{rt}|\{v_{t-j}\}_{j=1}^M) = \frac{1}{\pi\sigma^2} e^{-\frac{1}{\sigma^2} \left| v_{rt} - \sum_{j=1}^M a_{0j} v_{r,t-j} \right|^2} \quad (\text{C.3})$$

A partir de (C.2) se puede definir la función $\mathcal{L}_5(\boldsymbol{\theta})$

$$\mathcal{L}_5(\boldsymbol{\theta}) = RT \ln(\sigma^2) + \frac{1}{\sigma^2} \sum_{r=1}^R \sum_{t=1}^T \left| v_{rt} - \sum_{j=1}^M a_{0j} v_{r,t-j} \right|^2, \quad (\text{C.4})$$

donde el vector de parámetros a estimar es $\boldsymbol{\theta} = [\sigma^2 \operatorname{Re}\{a_{01}\} \operatorname{Im}\{a_{01}\} \cdots \operatorname{Re}\{a_{0M}\} \operatorname{Im}\{a_{0M}\}]^T$. La función de verosimilitud logarítmica condicional $-\ln(f(\mathbf{V}|\mathbf{v}_{1-M}; \cdots; \mathbf{v}_0))$ difiere de $\mathcal{L}_5(\boldsymbol{\theta})$ solo en una constante. Por lo tanto, los estimados de los parámetros pueden obtenerse como el valor de $\boldsymbol{\theta}$ que maximiza $\mathcal{L}_5(\boldsymbol{\theta})$.

A continuación se deducirán las expresiones de los estimadores de los coeficientes a_{0j} y de σ^2 a partir de la condición necesaria que deben satisfacer para ser un máximo local de $\mathcal{L}_5(\boldsymbol{\theta})$. Por tratarse de un proceso Gaussiano en este caso es posible obtener expresiones analíticas cerradas. Debido a que los coeficientes a_{0j} son complejos, al momento de derivar respecto de ellos debe seguirse el mismo procedimiento que para el parámetro γ en (4.24). Entonces, a partir de las condiciones

$$\left. \frac{\partial \mathcal{L}_5(\boldsymbol{\theta})}{\partial a_{0i}} \right|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}} = 0, \quad i = 1, \dots, M \quad (\text{C.5})$$

se obtiene el siguiente conjunto de ecuaciones lineales

$$\sum_{r=1}^R \sum_{t=1}^T v_{rt} v_{r,t-i}^* = \sum_{j=1}^M \hat{a}_{0j} \sum_{r=1}^R \sum_{t=1}^T v_{r,t-i}^* v_{r,t-j}, \quad i = 1, \dots, M \quad (\text{C.6})$$

que permite hallar los valores de los estimadores, \hat{a}_{0j} , de los coeficientes a_{0j} .

Por otra parte, a partir de la condición

$$\left. \frac{\partial \mathcal{L}_5(\boldsymbol{\theta})}{\partial \sigma^2} \right|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}} = 0, \quad (\text{C.7})$$

se obtiene la expresión del estimador de σ^2 , que está dado por

$$\hat{\sigma}^2 = \frac{1}{RT} \sum_{r=1}^R \sum_{t=1}^T \left| v_{rt} - \sum_{j=1}^M \hat{a}_{0j} v_{r,t-j} \right|^2. \quad (\text{C.8})$$

Bibliografía

- Abramowitz, M. y I. A. Stegun. *Handbook of Mathematical Functions*. Dover Publications, New York, 9 edición (1972).
- Amiri, H., H. Amindavar, y M. Kamarei. Underwater noise modeling and direction-finding based on heteroscedastic time series. *EURASIP Journal on Advances in Signal Processing*, (71528) (2007).
- Amirmazlaghani, M., H. Amindavar, y A. Moghaddamjoo. Speckle suppression in SAR images using the 2-D GARCH model. *IEEE Transactions on Image Processing*, 18(2):250–259 (2009).
- Anastassopoulos, V., G. A. Lampropoulos, A. Drosopoulos, y M. Rey. High resolution radar clutter statistics. *IEEE Transactions on Aerospace and Electronic Systems*, 35(1):43–60 (1999).
- Bakker, R. y B. Currie. *The McMaster IPIX radar sea clutter database*. Página web, disponible en: <http://soma.ece.mcmaster.ca/ipix/> (2001).
- Bauwens, L., S. Laurent, y J. Rombouts. Multivariate GARCH models: a survey. *Journal of Applied Econometrics*, 21(1):79–109 (2006).
- Berkes, I. y L. Horváth. The rate of consistency of the quasi-maximum likelihood estimator. *Statistics & Probability Letters*, 61(2):133–143 (2003).
- Berkes, I. y L. Horváth. The efficiency of the estimators of the parameters in GARCH processes. *The Annals of Statistics*, 32(2):633–655 (2004).
- Berkes, I., L. Horváth, y P. Kokoszka. GARCH processes: Structure and estimation. *Bernoulli*, 9(2):201–227 (2003).
- Blinnikov, S. y R. Moessner. Expansions for nearly Gaussian distributions. *Astronomy and Astrophysics*, 130:193–205 (1998).
- Bollerslev, T. Generalized autoregressive conditional heteroscedasticity. *Journal of Econometrics*, 31(3):307–327 (1986).

- Bose, S. y A. Steinhardt. A maximal invariant framework for adaptive detection with structured and unstructured covariance matrices. *IEEE Transactions on Signal Processing*, 43(9):2164–2175 (1995).
- Box, G., G. Jenkins, y G. Reinsel. *Time Series Analysis, Forecasting and Control*. Prentice Hall, Upper Saddle River, NJ, 3 edición (1994).
- Brandwood, D. H. A complex gradient operator and its application in adaptive array theory. *IEE Proceedings F in Communications, Radar and Signal Processing*, 130(1):11–16 (1983).
- Burden, R. L. y J. D. Faires. *Numerical Analysis*. Brooks/Cole, 20 Channel Center Street, Boston, 9 edición (2011).
- Conte, E. y M. Longo. Characterisation of radar clutter as a spherically invariant random process. *IEE Proceedings in Communications, Radar and Signal Processing*, 134(2):191–197 (1987).
- Conte, E., M. Lops, y G. Ricci. Adaptive matched filter detection in spherically invariant noise. *IEEE Signal Processing Letters*, 3(8):248–250 (1996).
- Cui, G., L. Kong, X. Yang, y J. Yang. Distributed target detection with polarimetric MIMO radar in compound-Gaussian clutter. *Digital Signal Processing*, 22(3):430–438 (2012).
- Dacunha-Castelle, D. y F. Gamboa. Maximum d’entropie et problème des moments. *Annales de l’institut Henri Poincaré (B) Probabilités et Statistiques*, 26(4):567–596 (1990).
- Ding, Z., C. Granger, y R. Engle. A long memory property of stock market returns and a new model. *Journal of Empirical Finance*, 1(1):83–106 (1993).
- Drosopoulos, A. *Description of the OHGR database*. Technical note 94-14, Defence Research Establishment Ottawa (1994).
- Engle, R. F. Autoregressive conditional heteroscedasticity with estimates of the variance of UK inflation. *Econometrica*, 50(4):987–1008 (1982).
- Farina, A., F. Gini, M. Greco, y L. Verrazzani. High resolution sea clutter data: statistical analysis of recorded live data. *IEE Proceedings in Radar, Sonar and Navigation*, 144(3):121–130 (1997).
- Fletcher, R. *Practical Methods of Optimization*. John Wiley & Sons, New York (1987).
- Francq, C. y J. Zakoian. Maximum likelihood estimation of pure GARCH and ARMA-GARCH processes. *Bernoulli*, 10(4):605–637 (2004).
- Francq, C. y J. Zakoian. Quasi-maximum likelihood estimation in GARCH processes when some coefficients are equal to zero. *Stochastic Processes and their Applications*, 117(9):1265–1284 (2007).

- Gini, F. y M. Greco. Texture modeling and validation using recorded high resolution sea clutter data. En *Proceedings of the IEEE Radar Conference, 2001.*, págs. 387–392. Atlanta, GA (2001).
- Gini, F., M. Greco, M. Diani, y L. Verrazzani. Performance analysis of two adaptive radar detectors against non-Gaussian real sea clutter data. *IEEE Transactions on Aerospace and Electronic Systems*, 36(4):1429–1439 (2000).
- Gini, F., M. Greco, A. Farina, y P. Lombardo. Optimum and mismatched detection against K-distributed plus Gaussian clutter. *IEEE Transactions on Aerospace and Electronic Systems*, 34(3):860–876 (1998).
- Gini, F., M. Greco, y L. Verrazzani. Detection problem in mixed clutter environment as a Gaussian problem by adaptive preprocessing. *Electronics Letters*, 31(14):1189–1190 (1995).
- Gini, F. y M. V. Greco. Suboptimum approach to adaptive coherent radar detection in compound-Gaussian clutter. *IEEE Transactions on Aerospace and Electronic Systems*, 35(3):1095–1104 (1999).
- Goldstein, G. B. False alarm regulation in log-normal and weibull clutter. *IEEE Transactions on Aerospace and Electronic Systems*, 9:84–92 (1973).
- Gouriéroux, C. *ARCH models and financial applications*. Springer-Verlag, New York (1997).
- Greco, M., F. Bordoni, y F. Gini. X-band sea-clutter nonstationarity: influence of long waves. *IEEE Journal of Oceanic Engineering*, 29(2):269–283 (2004).
- Harville, D. A. *Matrix Algebra From a Statistician's Perspective*. Springer-Verlag, New York (1997).
- Haykin, S. Radar signal processing. *IEEE ASSP Magazine*, 2:2–18 (1985).
- Haykin, S., C. Krasnor, T. J. Nohara, B. W. Currie, y D. Hamburger. A coherent dual-polarized radar for studying the ocean environment. *IEEE Transactions on Geoscience and Remote Sensing*, 29(1):189–191 (1991).
- Hurtado, M. y A. Nehorai. Polarimetric detection of targets in heavy inhomogeneous clutter. *IEEE Transactions on Signal Processing*, 56(4):1349–1361 (2008).
- Jakeman, E. y P. Pusey. A model for Non-Rayleigh sea echo. *IEEE Transactions on Antennas and Propagation*, 24:806–814 (1976).
- Kay, S. M. *Fundamentals of Statistical Signal Processing, Estimation Theory*. Prentice Hall, Upper Saddle River, NJ (1993).
- Kay, S. M. *Fundamentals of Statistical Signal Processing, Detection Theory*. Prentice Hall, Upper Saddle River, NJ (1998).

- Kelly, E. J. An adaptive detection algorithm. *IEEE Transactions on Aerospace and Electronic Systems*, AES-22(1):115–127 (1986).
- Koning, W. D. Optimal estimation of linear discrete-time systems with stochastic parameters. *Automatica*, 20(1):113–115 (1984).
- Konishi, S. y G. Kitagawa. *Information Criteria and Statistical Modeling*. Springer, New York (2008).
- Kraut, S. y L. Scharf. The CFAR adaptive subspace detector is a scale-invariant GLRT. *IEEE Transactions on Signal Processing*, 47(9):2538–2541 (1999).
- Lebedev, N. *Special functions and their applications*. Prentice Hall, Englewood Cliffs, NJ (1965).
- Lee, S. y B. Hansen. Asymptotic theory for the GARCH(1,1) quasi-maximum likelihood estimator. *Econometric Theory*, 10(1):29–52 (1994).
- Li, Y., W. Moran, S. Sira, A. Papandreou-Suppappola, y D. Morrell. Monte-Carlo based estimation methods for rapidly-varying sea clutter. En *Proceedings of the 13th IEEE Digital Signal Processing Workshop and 5th IEEE Signal Processing Education Workshop, 2009. DSP/SPE 2009.*, págs. 256–261. Marco Island, FL (2009).
- Ljung, L. *System Identification, Theory for the User*. Prentice Hall, Upper Saddle River, NJ (1999).
- Luo, Y., Y. Zhu, D. Luo, J. Zhou, E. Song, y D. Wang. Globally optimal multisensor distributed random parameter matrices kalman filtering fusion with applications. *Sensors*, 8(12):8086–8103 (2008).
- Luo, Y., Y. Zhu, X. Shen, y E. Song. Novel data association algorithm based on integrated random coefficient matrices kalman filtering. *IEEE Transactions on Aerospace and Electronic Systems*, 48(1):144–158 (2012).
- Ma, J. y S. Sun. Optimal linear estimators for systems with random sensor delays, multiple packet dropouts and uncertain observations. *IEEE Transactions on Signal Processing*, 59(11):5181–5192 (2011).
- Mahafza, B. R. y A. Z. Elsherbeni. *MATLAB Simulations for Radar Systems Design*. Chapman & Hall/CRC, Boca Raton (2004).
- Maio, A. D., G. Foglia, E. Conte, y A. Farina. CFAR behavior of adaptive detectors: an experimental analysis. *IEEE Transactions on Aerospace and Electronic Systems*, 41(1):233–251 (2005).
- Marcum, J. A statistical theory of target detection by pulsed radar. *IRE Transactions on Information Theory*, 6(2):59–267 (1960).

- Moayed, M., Y. Foo, y Y. Soh. Adaptive kalman filtering in networked systems with random sensor delays, multiple packet dropouts and missing measurements. *IEEE Transactions on Signal Processing*, 58(3):1577–1588 (2010).
- Mousazadeh, S. y I. Cohen. Simultaneous parameter estimation and state smoothing of complex GARCH process in the presence of additive noise. *Signal Processing*, 90(11):2947–2953 (2010).
- Neuweiler, G. *The biology of bats*. Oxford University Press, 198 Madison Avenue, New York (2000).
- Noiboar, A. y I. Cohen. Anomaly detection based on wavelet domain GARCH random field modeling. *IEEE Transactions on Geoscience and Remote Sensing*, 45(5):1361–1373 (2007).
- Oliver, C. y S. Quegan. *Understanding Synthetic Aperture Radar Images*. SciTech Publishing, Inc., Raleigh, NC 27613 (2004).
- Ormoneit, D. y H. White. An efficient algorithm to compute maximum entropy densities. *Econometrics Reviews*, 18:127–140 (1999).
- Papoulis, A. *Probability, Random Variables, and Stochastic Processes*. McGraw-Hill, New York, 3 edición (1991).
- Pascual, J., N. von Ellenrieder, M. Hurtado, y C. H. Muravchik. Modelado de clutter de radar utilizando procesos GARCH. En *Actas de la XIV Reunión de Trabajo Procesamiento de la Información y Control (RPIC 2011)*, págs. 537–547. Oro Verde, Santa Fe, Argentina (2011).
- Pascual, J., N. von Ellenrieder, y C. Muravchik. Conditional variance LMMSE estimator for a GARCH process clutter model. En *Proceedings of the 8th IEEE Sensor Array and Multichannel Signal Processing Workshop (SAM 2014)*, págs. 309–312. A Coruña, Spain (2014a).
- Pascual, J., N. von Ellenrieder, y C. H. Muravchik. Aproximación de la cota de Cramér–Rao para la estimación de parámetros de un proceso ARCH. En *Actas de la XV Reunión de Trabajo Procesamiento de la Información y Control (RPIC 2013)*, págs. 13–18. San Carlos de Bariloche, Río Negro, Argentina (2013a).
- Pascual, J. P., J. I. F. Michelli, N. von Ellenrieder, M. Hurtado, J. Areta, y C. H. Muravchik. Image classification by means of CEM algorithm based on a GARCH-2D data model. *IEEE Latin America Transactions*, 12(5):877–882 (2014b).
- Pascual, J. P., N. von Ellenrieder, M. Hurtado, y C. H. Muravchik. Radar detection algorithm for GARCH clutter model. *Digital Signal Processing*, 23(4):1255–1264 (2013b).
- Pascual, J. P., N. von Ellenrieder, M. Hurtado, y C. H. Muravchik. Adaptive radar detection algorithm based on an autoregressive GARCH-2D clutter model. *IEEE Transactions on Signal Processing*, 62(15):3822–3832 (2014c).

- Peebles, P. Z. *Radar Principles*. John Wiley & Sons, New York (1998).
- Ravid, R. y N. Levanon. Maximum-likelihood CFAR for Weibull background. *IEE Proceedings-F, Radar and Signal Processing*, 139(3):256–264 (1992).
- Reed, I., J. Mallett, y L. Brennan. Rapid convergence rate in adaptive arrays. *IEEE Transactions on Aerospace and Electronic Systems*, AES-10(6):853–863 (1974).
- Richards, M. A. *Fundamentals of Radar Signal Processing*. McGraw-Hill, New York (2005).
- Richmond, C. A note on non-Gaussian adaptive array detection and signal parameter estimation. *IEEE Signal Processing Letters*, 2(8):251–252 (1996).
- Sangston, K. y K. Gerlach. Coherent detection of radar targets in a non-Gaussian background. *IEEE Transactions on Aerospace and Electronic Systems*, 30(2):330–340 (1994).
- Sangston, K., F. Gini, y M. Greco. Coherent radar target detection in heavy-tailed compound-Gaussian clutter. *IEEE Transactions on Aerospace and Electronic Systems*, 48(1):64–77 (2012).
- Sangston, K., F. Gini, M. Greco, y A. Farina. Structures for radar detection in compound Gaussian clutter. *IEEE Transactions on Aerospace and Electronic Systems*, 35(2):445–458 (1999).
- Sekine, M., T. Musha, Y. Tomita, T. Hagiwara, T. Irabu, y E. Kiuchi. On weibull-distributed weather clutter. *IEEE Transactions on Aerospace and Electronic Systems*, AES-15(6):824–830 (1979).
- Shang, X. y H. Song. Radar detection based on compound-Gaussian model with inverse gamma texture. *IET Radar, Sonar and Navigation*, 5(3):315–321 (2011).
- Sheikhi, A., M. Nayebi, y M. Aref. Adaptive detection algorithm for radar signals in autoregressive interference. *IEE Proceedings in Radar, Sonar and Navigation*, 145(5):309–314 (1998).
- Shnidman, D. A. Generalized radar clutter model. *IEEE Transactions on Aerospace and Electronic Systems*, 35(3):857–865 (1999).
- Shumway, R. H. y D. S. Stoffer. *Time Series Analysis and Its Applications With R Examples*. Springer, New York, 2 edición (2006).
- Simon, D. *Optimal State Estimation - Kalman, H_∞ , and Nonlinear Approaches*. John Wiley & Sons, New York (2006).
- Skolnik, M. *Introduction to Radar Systems*. McGraw-Hill, New York, 3 edición (2001).
- Straumann, D. *Estimation in Conditionally Heteroscedastic Time Series Models*, tomo 181 de *Lecture Notes in Statistics*. Springer-Verlag, Berlin (2005).

- Sun, S. y J. Ma. Linear estimation for networked control systems with random transmission delays and packet dropouts. *Information Sciences*, 269(0):349–365 (2014).
- Sun, S., L. Xie, W. Xiao, y Y. C. Soh. Optimal linear estimation for systems with multiple packet dropouts. *Automatica*, 44(5):1333–1342 (2008).
- Sun, W. y Y. X. Yuan. *Optimization Theory and Methods, Nonlinear Programming*. Springer, New York (2006).
- Swerling, P. Probability of detection for fluctuating targets. *IRE Transactions on Information Theory*, 6(2):269–308 (1960).
- Talenti, G. Recovering a function from a finite number of moments. *Inverse Problems*, 3:501–517 (1987).
- Van Trees, H. L. *Detection, Estimation and Modulation Theory. Part I: Detection, Estimation and Linear Modulation Theory*. John Wiley & Sons, New York (2001).
- von Ellenrieder, N. y C. H. Muravchik. Maximum-entropy moments-based approximations of the Cramér–Rao bound. En *Proceedings of the 28th International Workshop on Bayesian Inference and Maximum Entropy Methods in Science and Engineering*, págs. 165–172 (2008).
- Weinberg, G. V. Assessing Pareto fit to high-resolution high-grazing-angle sea clutter. *Electronics Letters*, 47(8):516–517 (2011).
- Weiss, A. Asymptotic theory for ARCH models: Estimation and testing. *Econometric Theory*, 2(1):107–131 (1986).
- Yaz, E. y R. E. Skelton. Parametrization of all linear compensators for discrete-time stochastic parameter systems. *Automatica*, 30(6):945–955 (1994).
- Zha, D. y T. Qiu. Underwater sources location in non-gaussian impulsive noise environments. *Digital Signal Processing*, 16(2):149–163 (2006).