

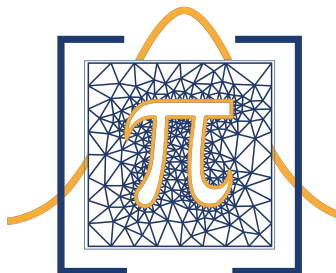
UNIVERSIDAD NACIONAL AUTÓNOMA DE HONDURAS
FACULTAD DE CIENCIAS
ESCUELA DE MATEMÁTICA Y CIENCIAS DE LA COMPUTACIÓN
MAESTRÍA EN MATEMÁTICA



UNAH
UNIVERSIDAD NACIONAL
AUTÓNOMA DE HONDURAS

Modelado de Series Temporales Económicas: Operadores
Autorregresivos, Cadenas de Markov y Reservorios
Estocásticos

Tesis presentada para optar al Título de máster en matemática con
orientación en Ingeniería Matemática



Tesis presentada por: Norman Reynaldo Sabillón Castro

Asesor de tesis: Dr. Fredy Vides
Fecha de defensa: 8 de mayo de 2026

TEGUCIGALPA, HONDURAS 2026.

Modelado de Series Temporales Económicas: Operadores Autorregresivos, Cadenas de Markov y Reservorios Estocásticos

Presentada por: Norman Reynaldo Sabillón Castro

Tesis presentada para optar al Título de máster en matemática con orientación en Ingeniería Matemática

Asesor de tesis: Dr. Fredy Vides
Universidad Nacional Autónoma de Honduras

TERNA EXAMINADORA

Ph.D. Cristian Andrés Cruz Torres
Universidad Nacional Autónoma de Honduras

M.Sc. Henry David Ocampo Meraz
Universidad Nacional Autónoma de Honduras

M.Sc. Roberto Carlos Duarte Rivera
Universidad Nacional Autónoma de Honduras

TEGUCIGALPA M.D.C 2026.

Agradecimientos

Agradezco a Dios por acompañarme durante todo este proceso. En los momentos difíciles de la investigación, su guía me dio la claridad que necesitaba para seguir adelante.

A mis padres, Wilfredo Sabillón y Ernestina Castro. Cada vez que dudé si podía terminar esto, su confianza me sostuvo. Han sido siempre mi roca firme, motivándome de manera constante a culminar cada meta que me propongo en la vida. No exagero al decir que sin ellos este trabajo no existiría.

A Brenda, mi esposa. Hubo semanas enteras en que la pantalla de la computadora recibió más atención que ella, y aun así me apoyó sin reservas. Su paciencia y su amor hicieron posible lo que a ratos parecía imposible; gracias de corazón por cuidar con tanto cariño y detalle los preparativos de la defensa, aportando esa calidez y excelencia para que la presentación estuviera al más alto nivel.

A mi hijo, Andre Emiliano. Su sonrisa al final del día me recordaba por qué valía la pena el esfuerzo. Gracias por tu curiosidad y por tus constantes preguntas para entender de qué trataba mi tesis; conversar contigo me enseñó la valiosa lección de que lo técnico y complejo solo cobra verdadero valor cuando somos capaces de explicarlo de forma sencilla.

Al Dr. Fredy Vides, mi asesor, quien confió en esta línea de investigación desde el principio y me orientó con rigor y generosidad. Aprendí de él mucho más que matemáticas.

Al M.Sc. David Méndez, por haber sido ese docente que me impulsó a tomar cursos adicionales de redacción. Aunque no eran parte obligatoria del plan de estudios, hoy reconozco el inmenso valor de esas herramientas en mi escritura profesional y científica.

Al Dr. Leandro Galo, por su genuina preocupación por nosotros como estudiantes y por darme ese aliento decisivo en la etapa final. Su motivación y su seguimiento cercano fueron claves para seguir adelante y culminar este proyecto de vida.

A mi compañero Armando Ramos, por su presencia constante y por acompañarme como testigo en la defensa, demostrando el valor de una verdadera amistad.

A mis compañeros de maestría, por las discusiones, los ánimos mutuos y la camaradería en los momentos más exigentes. Y a todas las personas que, de una forma u otra, contribuyeron a que esta tesis llegara a buen puerto.

Norman Reynaldo Sabillón Castro

MODELADO DE SERIES TEMPORALES ECONÓMICAS: OPERADORES AUTORREGRESIVOS, CADENAS DE MARKOV Y RESERVORIOS ESTOCÁSTICOS

Resumen

Modelar regímenes en series temporales económicas enfrenta una dicotomía inherente entre rigidez determinista y opacidad interpretativa probabilística. Esta tesis propone un marco que conecta ambos enfoques. La estrategia parte de operadores autorregresivos con propiedades formales demostradas, pasa por una estimación Markoviana convexa y culmina en una extensión no lineal basada en reservorios.

La familia de operadores de Tasa de Cambio Relativa Autorregresiva (TCRA, TCRAM, ETCRA, ETCRAM) convierte series continuas en regímenes económicos discretos e interpretables. Para cada variante se demuestran existencia, unicidad, consistencia asintótica y estabilidad ante perturbaciones. Los regímenes resultantes alimentan una cadena de Markov cuya matriz de transición se estima por Mínimos Cuadrados No Negativos (NNLS), formulación convexa que garantiza convergencia global. El enfoque estándar de máxima verosimilitud en modelos *Markov-Switching* (MS-AR) no ofrece esa garantía.

Un resultado formal, el Teorema de Inclusión, establece que TCRA-Markov es un caso particular de la Computación de Reservorio Estocásticamente Estructurada (SSRC). Esto abre la vía a dinámicas no lineales sin sacrificar las garantías teóricas del modelo base.

La validación empírica utiliza series semanales de cuatro combustibles hondureños (2017–2025). La predicción de régimen alcanza entre 85 y 91 % de exactitud. Los cuantiles fijos resultan estadísticamente inferiores ($p < 0.05$, test de Diebold-Mariano) y el modelo de referencia MS-AR no convergió bajo el protocolo de validación y conjunto de datos de este estudio. Con SSRC, el RMSE se reduce entre 20 y 23 % respecto al modelo lineal ($p < 0.05$ en tres de cuatro series). Los tiempos de mezcla, de 25 a 47 semanas, cuantifican la inercia del mercado regulado hondureño.

Palabras clave: TCRA, series temporales, cadenas de Markov, NNLS, cambio de régimen, computación de reservorio, SSRC, combustibles, Honduras.

ECONOMIC TIME SERIES MODELING: AUTOREGRESSIVE OPERATORS, MARKOV CHAINS, AND STOCHASTIC RESERVOIRS

Abstract

Modeling regimes in economic time series involves a well-known dichotomy between deterministic rigidity and probabilistic uninterpretability. This thesis proposes a unified framework that bridges both approaches. The strategy builds on autoregressive operators with proven formal properties, convex Markov estimation, and a nonlinear reservoir-based extension.

The Autoregressive Relative Change Rate (TCRA) operator family and its parametric variants (TCRAM, ETCRA, ETCRAM) convert continuous time series into interpretable economic regimes. Existence, uniqueness, asymptotic consistency, and perturbation stability are proved for each variant. The resulting regimes feed a Markov chain whose transition matrix is estimated via Non-Negative Least Squares (NNLS), a convex formulation that guarantees global convergence. Standard maximum likelihood in Markov-Switching (MS-AR) models offers no such guarantee.

A formal result, the Inclusion Theorem, establishes that TCRA-Markov is a special case of Stochastically Structured Reservoir Computing (SSRC). This opens the path to nonlinear dynamics without sacrificing the theoretical guarantees of the base model. Empirically, the extension yields 20–23% improvements in forecast accuracy.

Validation uses weekly price series for four Honduran fuels (Regular, Super, Diesel, Kerosene) over 2017–2025. Regime prediction accuracy reaches 85–91%. Fixed quantiles are statistically inferior ($p < 0.05$, Diebold-Mariano test) and the MS-AR benchmark did not converge under our validation protocol and data. Mixing times of 25 to 47 weeks quantify the inertia of Honduras’s regulated fuel market.

Keywords: TCRA, time series, Markov chains, NNLS, regime switching, reservoir computing, SSRC, fuel prices, Honduras.

Índice general

| | |
|--|-----------|
| Agradecimientos | II |
| Resumen | III |
| Abstract | IV |
| Índice de figuras | IX |
| Índice de tablas | XI |
| Glosario de Términos | XII |
| Tabla de Notación | XVI |
| 1. Introducción, Series de Estudio y Planteamiento del Problema | 1 |
| 1.1. Series de Estudio | 2 |
| 1.1.1. El PIB de Honduras como Serie Macroeconómica | 2 |
| 1.1.2. Precios de Combustibles | 5 |
| 1.2. Revisión de Literatura y Posicionamiento | 8 |
| 1.2.1. Modelos de Cambio de Régimen | 8 |
| 1.2.2. Estimación de Matrices de Transición | 9 |
| 1.2.3. Computación de Reservorio para Series Económicas | 10 |
| 1.2.4. Identificación del Vacío en la Literatura | 10 |
| 1.3. Planteamiento del Problema | 11 |
| 1.3.1. Hoja de Ruta Metodológica | 11 |
| 1.3.2. Objetivos y Alcance de la Investigación | 12 |
| 1.3.3. Marco Metodológico | 12 |
| 1.3.4. Estructura del Documento | 13 |
| 2. Tasa de Cambio Relativa Autorregresiva y sus Variantes | 15 |
| 2.1. Introducción y Motivación Matemática | 15 |
| 2.2. Axiomática del Operador de Tasa de Cambio Local | 15 |
| 2.3. Marco General Unificado | 16 |
| 2.4. Teorema Fundamental del Operador | 18 |
| 2.4.1. Existencia Ponderada y Unicidad Global | 18 |
| 2.4.2. Consistencia Asintótica | 19 |
| 2.4.3. Estabilidad de Lipschitz | 21 |
| 2.4.4. Optimalidad de Gauss-Markov | 21 |
| 2.5. Casos Particulares del Estimador Generalizado | 23 |
| 2.5.1. TCRA: Estimador Global ($W = T, \lambda = 1$) | 23 |
| 2.5.2. TCRA: Estimador con Ventana Móvil ($W < T, \lambda = 1$) | 23 |
| 2.5.3. ETCRA: Estimador con Decaimiento Exponencial ($W = T, \lambda < 1$) | 24 |
| 2.5.4. ETCRA: Estimador Local con Decaimiento Exponencial ($W < T, \lambda < 1$) | 24 |
| 2.6. Complejidad Algorítmica | 25 |
| 2.7. Implementación Computacional | 26 |
| 2.8. Ejemplos Numéricos Sectoriales | 27 |
| 2.8.1. Validación Cruzada Expansiva en Mercado de Combustibles | 29 |
| 2.9. Relación con Métodos Establecidos | 30 |

| | |
|--|------------|
| 2.10. Resumen del Capítulo | 31 |
| 3. TCRA con Cadenas de Markov | 33 |
| 3.1. Fundamentos de Invarianza y Estabilidad | 33 |
| 3.2. Modelo Híbrido TCRA-Markov | 34 |
| 3.3. Partición de Estados | 35 |
| 3.4. Propiedades Markovianas y Asintóticas | 36 |
| 3.5. Resumen Visual del Modelo | 43 |
| 3.6. Comparación de Modelos de Transición de Estados | 44 |
| 4. Modelo Híbrido TCROC-Markov con Estimación NNLS | 46 |
| 4.1. Discretización de Regímenes con TCROC-Markov | 46 |
| 4.1.1. Cálculo de la Tasa Relativa | 46 |
| 4.1.2. Asignación de Estados | 47 |
| 4.1.3. Análisis de Sensibilidad de los Parámetros de Discretización | 52 |
| 4.2. Estimación de la Matriz de Transición Lineal con NNLS | 53 |
| 4.2.1. Formulación Matemática y Solución Constructiva | 54 |
| 4.2.2. Validación del Rendimiento Predictivo | 63 |
| 4.2.3. Análisis Predictivo y Robustez del Modelo para Combustibles | 71 |
| 4.3. Resumen del Algoritmo Metodológico | 72 |
| 4.4. Discusión de Limitaciones y Extensiones Futuras | 73 |
| 4.5. Resumen del Capítulo | 74 |
| 5. Detección y Pronóstico de Regímenes de Precios en Combustibles | 76 |
| 5.1. Introducción | 76 |
| 5.2. Marco Metodológico Aplicado | 77 |
| 5.2.1. Etapa 1: Discretización Basada en Datos | 78 |
| 5.2.2. Etapa 2: Estimación de la Matriz de Transición | 82 |
| 5.2.3. Protocolo de Optimización de Hiperparámetros | 82 |
| 5.3. Resultados y Análisis | 84 |
| 5.3.1. El Fenómeno de Degeneración del Modelo | 84 |
| 5.3.2. Análisis de Sensibilidad | 85 |
| 5.3.3. Hiperparámetros Óptimos | 86 |
| 5.3.4. Identificación e Interpretación Económica de los Regímenes | 89 |
| 5.3.5. Dinámica de Transición | 91 |
| 5.3.6. Propiedades Espectrales y Estabilidad Estructural | 93 |
| 5.3.7. Convergencia de las Cadenas de Markov | 95 |
| 5.3.8. Cómputo Analítico de la Varianza del Pronóstico | 96 |
| 5.3.9. Superioridad sobre Discretización por Cuantiles | 97 |
| 5.3.10. Desempeño Predictivo y Comparación con Modelos de Referencia | 99 |
| 5.4. Limitaciones y Supuestos | 101 |
| 5.5. Resumen del Capítulo | 102 |
| 6. Extensión No Lineal: Computación de Reservorio Estructurado | 104 |
| 6.1. Introducción y Motivación | 104 |
| 6.2. Marco Teórico del Computador de Reservorio | 105 |
| 6.2.1. Sistemas Dinámicos con Entrada | 105 |
| 6.2.2. El Computador de Reservorio (<i>Echo State Network</i>) | 105 |
| 6.2.3. Diferenciación y Relación con Redes Neuronales Convencionales | 106 |
| 6.2.4. Variante con Integración de Fuga (<i>Leaky-ESN</i>) | 108 |

| | | |
|-----------|---|------------|
| 6.2.5. | El Computador de Reservoirio Estructurado Estocástico (SSRC) | 109 |
| 6.2.6. | La Propiedad de Estado de Eco | 109 |
| 6.2.7. | Capacidad de Aproximación Universal | 111 |
| 6.3. | TCRA-Markov como Caso Particular del SSRC | 112 |
| 6.3.1. | Reformulación del Marco Lineal | 112 |
| 6.3.2. | El Teorema de Inclusión | 113 |
| 6.4. | Estimación de la Capa de Lectura vía NNLS | 115 |
| 6.4.1. | Formulación del Problema | 115 |
| 6.4.2. | Herencia de las Propiedades NNLS | 116 |
| 6.5. | Arquitectura TCRA-SSRC | 117 |
| 6.6. | Propiedades de Estabilidad del Pronóstico | 119 |
| 6.7. | Aplicación Comparativa | 121 |
| 6.7.1. | Configuración Experimental | 121 |
| 6.7.2. | Verificaciones Teóricas | 122 |
| 6.7.3. | Optimización de Hiperparámetros | 126 |
| 6.7.4. | Resultados | 130 |
| 6.7.5. | Análisis de Sensibilidad y Estabilidad | 133 |
| 6.7.6. | Representación de Regímenes en el Espacio del Reservoirio | 133 |
| 6.7.7. | Discusión | 136 |
| 6.8. | Limitaciones | 138 |
| 6.9. | Resumen del Capítulo | 139 |
| 7. | Conclusiones Generales y Trabajo Futuro | 141 |
| 7.1. | Síntesis del Trabajo | 141 |
| 7.2. | Resumen Cuantitativo de Resultados | 143 |
| 7.3. | Contribuciones | 144 |
| 7.3.1. | Contribuciones Teóricas | 144 |
| 7.3.2. | Contribuciones Metodológicas | 145 |
| 7.3.3. | Contribuciones Aplicadas | 145 |
| 7.4. | Limitaciones Generales | 146 |
| 7.5. | Líneas de Trabajo Futuro | 147 |
| 7.5.1. | Extensiones Inmediatas | 147 |
| 7.5.2. | Extensiones de Mediano Plazo | 149 |
| 7.5.3. | Visión de Largo Plazo | 150 |
| 7.6. | Recomendaciones | 150 |
| 7.6.1. | Recomendaciones para el Mercado de Combustibles de Honduras | 150 |
| 7.6.2. | Recomendaciones para el Desarrollo de Modelos Estocásticos de Regímenes | 151 |
| 7.7. | Reflexiones Finales | 151 |
| | Bibliografía | 153 |
| | Apéndices | 157 |
| A.1. | Demostración Extendida del Teorema de ϕ -Mixing | 158 |
| A.2. | Estabilidad Discreta de la Función de Cuantización π | 158 |
| A.3. | Trayectorias y Transiciones del Proceso $\{S_t\}$ con Datos Reales | 159 |
| B.1. | Implementación Unificada del Operador TCRA | 161 |
| B.2. | Código Python para Análisis de Sensibilidad | 163 |
| B.3. | Código Python para Validación con Múltiples Particiones | 166 |
| B.4. | Código Fuente: Estimación y Simulación | 171 |

C.1. Código Fuente: Computación de Reservorio (SSRC) 171
D.1. Código para Generación de Figuras 171

Índice de figuras

| | | |
|-------|--|----|
| 1.1. | Evolución del PIB de Honduras en dólares corrientes, 1960–2023. Fuente: [45]. | 3 |
| 1.2. | Tasa de crecimiento anual del PIB de Honduras (1961–2023). | 3 |
| 1.3. | PIB per cápita de Honduras en dólares corrientes (1960–2023). | 4 |
| 1.4. | PIB per cápita de Honduras en lempiras a precios constantes (1960–2023). | 4 |
| 1.5. | Precio semanal de la gasolina Superior en Honduras (2017–2025). | 6 |
| 1.6. | Precio semanal de la gasolina Regular en Honduras (2017–2025). | 6 |
| 1.7. | Precio semanal del Diésel en Honduras (2017–2025). | 7 |
| 1.8. | Precio semanal del Kerosene en Honduras (2017–2025). | 7 |
| 1.9. | Estructura de la tesis y dependencias entre capítulos. | 13 |
| 1.10. | Cadena de dependencias lógicas entre los resultados teóricos principales de la tesis. | 14 |
| 2.1. | Jerarquía paramétrica del operador TCRA y sus variantes. | 18 |
| 2.2. | Comparativa de complejidad computacional de las variantes de la TCRA frente a métodos alternativos. | 27 |
| 2.3. | Serie histórica del PIB (2000–2024) y proyecciones (2025–2026) bajo las cuatro variantes del operador TCRA. | 28 |
| 3.1. | Evolución estructural del modelo TCRA hasta su extensión Markoviana. | 43 |
| 4.1. | Diagrama conceptual de la metodología híbrida TCROC-Markov con estimación NNLS. | 47 |
| 4.2. | Ilustración del proceso de discretización. El panel superior muestra la evolución de la tasa continua α_t . El panel inferior muestra la secuencia de estados discretos S_t resultante de aplicar la función de cuantización π . | 48 |
| 4.3. | Visualización del proceso de discretización para el Ejemplo 4.1. El panel superior muestra los valores calculados de la tasa α_t . El panel inferior muestra la secuencia de estados discretos $\{s_2, s_3, s_3\}$ resultante. | 49 |
| 4.4. | Proceso completo de discretización del PIB per cápita en Lempiras (2002–2023). El panel superior muestra la evolución de la tasa α_t ; el inferior, la secuencia de estados discretos S_t . Se observan regímenes de contracción (2009, 2020), estabilidad y crecimiento fuerte (2022). | 51 |
| 4.5. | Secuencia temporal de estados discretos obtenidos mediante TCROC. | 56 |
| 4.6. | Grafo de transición estimado para los regímenes del PIB per cápita en Lempiras. | 59 |
| 4.7. | Evolución temporal de las probabilidades de ocupación de los estados s_1, s_2, s_3, s_4 bajo la dinámica inducida por la matriz de transición estimada $\hat{\mathbf{P}}$. | 62 |
| 4.8. | Evolución de las probabilidades de ocupación para cada estado según la matriz de transición estimada $\hat{\mathbf{P}}$ en las series de precios de combustibles: Super, Regular, Diésel y Kerosene. | 68 |
| 4.9. | Comparación entre la matriz de transición normalizada $\hat{\mathbf{P}}$ y la matriz estimada A_p para la serie Super, sin considerar el estado s_4 que no está presente. | 69 |
| 4.10. | Comparación entre la matriz de transición normalizada $\hat{\mathbf{P}}$ y la matriz estimada A_p para la serie Regular, sin considerar el estado s_4 que no está presente. | 69 |
| 4.11. | Comparación entre la matriz de transición normalizada $\hat{\mathbf{P}}$ y la matriz estimada A_p para la serie Diésel. | 70 |

| | |
|---|-----|
| 4.12. Comparación entre la matriz de transición normalizada $\hat{\mathbf{P}}$ y la matriz estimada A_p para la serie Kerosene. | 70 |
| 5.1. Análisis de sensibilidad del desempeño respecto a la ventana W | 86 |
| 5.2. Espacio de optimización de hiperparámetros consolidado por número de regímenes (K). | 87 |
| 5.3. Jerarquía de predictibilidad de los combustibles. | 89 |
| 5.4. Regímenes identificados por K-medias sobre la serie α_t | 90 |
| 5.5. Grafos de transición estimados para los modelos optimizados. | 92 |
| 5.6. Evolución de las probabilidades de estado bajo los modelos K-medias. | 96 |
| 5.7. Segmentación de regímenes basada en cuantiles. | 98 |
| 5.8. Evolución de probabilidades de estado bajo el método de cuantiles. | 99 |
| 6.1. Diagrama de la arquitectura del reservorio SSRC. | 108 |
| 6.2. Verificación de la identidad algebraica entre TCRA-Markov y SSRC degenerado. | 114 |
| 6.3. Comparación esquemática de las arquitecturas TCRA-Markov y TCRA-SSRC. | 117 |
| 6.4. Autovalores de \mathbf{W}_{res} en el plano complejo para la configuración óptima de cada serie. | 124 |
| 6.5. Convergencia del RMSE según el período de calentamiento W_{wash} | 125 |
| 6.6. Búsqueda en grilla del SSRC sobre el plano (D, ρ) | 126 |
| 6.7. Sensibilidad del RMSE a la dimensión del reservorio D | 128 |
| 6.8. Sensibilidad del RMSE al radio espectral ρ | 129 |
| 6.9. Comparación visual de RMSE: TCRA-Markov vs. TCRA-SSRC. | 131 |
| 6.10. Distribución del RMSE del SSRC sobre 30 realizaciones independientes. | 131 |
| 6.11. Comparación de pronósticos vs. precios reales en el período fuera de muestra. | 132 |
| 6.12. Evaluación empírica de la cota de perturbación en función del radio espectral ρ | 133 |
| 6.13. Proyección de los estados del reservorio mediante PCA. | 134 |
| 6.14. Relación entre retorno TCRA α_t y activación promedio del reservorio \bar{h}_t | 135 |
| 6.15. Identificación de regímenes de mercado mediante SSRC. | 136 |
| A.1. Trayectorias del proceso $\{S_t\}$ para combustible Superior (secuencia observada vs. simulada). | 160 |
| A.2. Diagrama de transición entre regímenes para combustible Superior ($K = 4$). | 160 |
| A.3. Trayectorias del proceso $\{S_t\}$ para combustible Regular (secuencia observada vs. simulada). | 161 |
| A.4. Diagrama de transición entre regímenes para combustible Regular ($K = 4$). | 161 |

Índice de tablas

| | |
|---|-----|
| 1.1. Evolución de series económicas seleccionadas de Honduras, 2000–2023. . . . | 5 |
| 1.2. Comparación de modelos de cambio de régimen según siete criterios clave. . . . | 11 |
| 2.1. Comparativa de complejidad algorítmica entre las variantes TCRA y otros enfoques de modelación. | 25 |
| 2.2. Modelación paramétrica divergente del PIB de Honduras (millones USD) para los años 2025–2026 bajo las cuatro formulaciones de la TCRA. | 28 |
| 2.3. Evolución del error predictivo porcentual absoluto (MAPE) a lo largo de los anclajes temporales de validación un-paso-adelante para los cuatro combustibles. | 29 |
| 3.1. Comparación entre variantes de TCRA y el modelo TCRA-Markov. | 44 |
| 4.1. Análisis de sensibilidad: Distribución de estados (%) del PIB per cápita según los parámetros W y λ | 52 |
| 4.2. Resumen del rendimiento predictivo según la partición de datos. | 63 |
| 4.3. Resumen del rendimiento predictivo máximo para cada serie de combustible. | 71 |
| 5.1. Índice de Davies-Bouldin e inercia intra-clúster para diferentes números de estados K | 80 |
| 5.2. Tests de estacionariedad sobre las series $\{\alpha_t\}$ de cada combustible. | 81 |
| 5.3. Combinación óptima de hiperparámetros por serie de combustible. | 87 |
| 5.4. Centroides (μ_k) y prevalencia temporal de los regímenes identificados. | 91 |
| 5.5. Propiedades espectrales y distribuciones estacionarias de las matrices de transición. | 94 |
| 5.6. Evaluación final del desempeño predictivo y significancia estadística. | 99 |
| 6.1. Verificación de las condiciones teóricas del SSRC. | 123 |
| 6.2. Hiperparámetros óptimos del SSRC por serie de combustible. | 127 |
| 6.3. Comparación de desempeño predictivo: TCRA-Markov vs. TCRA-SSRC. | 130 |
| 7.1. Resumen consolidado de métricas por serie y método. | 143 |
| A.1. Matriz de transición estimada $\hat{\mathbf{P}}$ para combustible Superior ($K = 4$). | 159 |
| A.2. Matriz de transición estimada $\hat{\mathbf{P}}$ para combustible Regular ($K = 4$). | 162 |

Glosario de Términos

TCRA (α_T) **Tasa de Cambio Relativa Autorregresiva**: Modelo que estima la tasa de cambio relativa entre valores consecutivos de una serie temporal mediante la minimización de diferencias cuadradas. Su parámetro principal, α_T , representa la variación relativa óptima en la serie y minimiza la suma de diferencias cuadradas entre valores adyacentes.

TCRAM (α_T, W) **Tasa de Cambio Relativa Autorregresiva Móvil**: Extensión de la TCRA que incorpora una ventana móvil para análisis localizado, capturando dinámicas recientes en series temporales. Se define por dos cantidades:

- α_T : Tasa que optimiza la relación entre valores consecutivos dentro de la ventana móvil.
- W : Tamaño de la ventana móvil que define el subconjunto de datos analizado.

ETCRA (α_T, λ) **Tasa de Cambio Relativa Autorregresiva con Decaimiento Exponencial**: Modelo que pondera observaciones recientes mediante un factor de decaimiento exponencial, mejorando la sensibilidad a cambios temporales. La tasa α_T optimiza la relación entre valores consecutivos ajustada por ponderaciones exponenciales, y el factor $\lambda \in (0, 1]$ controla cuánto se prioriza el dato reciente sobre el histórico.

ETCRAM (α_T, λ, W) **Tasa de Cambio Relativa Autorregresiva con Decaimiento Exponencial y Ventana Móvil**: Combina la ventana móvil de TCRAM y el decaimiento exponencial de ETCRA para un análisis localizado y ponderado de series temporales. Involucra tres cantidades:

- α_T : Tasa que optimiza la relación entre valores consecutivos en la ventana móvil, ponderada exponencialmente.
- λ : Factor de decaimiento exponencial para priorizar datos recientes.
- W : Tamaño de la ventana móvil que segmenta la serie en subconjuntos.

TCRA-Markov (α_t, K) **Tasa de Cambio Relativa Autorregresiva con Cadenas de Markov**: Extensión de la TCRA que modela tasas de cambio como un proceso de Markov con K estados, capturando transiciones entre regímenes en series temporales. Requiere dos elementos:

- α_t : Tasa de cambio relativa en el tiempo t , modelada como variable de estado en una cadena de Markov.
- K : Número de estados discretos en el modelo de Markov.

ARIMA (p, d, q) **AutoRegressive Integrated Moving Average**: Modelo clásico de series temporales que combina componentes autorregresivos, de diferenciación y de promedio móvil para modelar dinámicas temporales. Los tres órdenes que lo caracterizan son:

- p : Orden del componente autorregresivo (*AutoRegressive*).
- d : Grado de diferenciación (*Integrated*) para lograr estacionariedad.
- q : Orden del componente de promedio móvil (*Moving Average*).

En su variante estacional (SARIMA), se incluyen parámetros (P, D, Q) para modelar estacionalidad.

Mínimos Cuadrados Método estadístico que estima parámetros minimizando la suma de los cuadrados de las diferencias entre valores observados y predichos.

MLE Estimación de Máxima Verosimilitud: Técnica estadística que selecciona parámetros maximizando la probabilidad de observar los datos mue

AIC Criterio de Información de Akaike: Método para seleccionar modelos estadísticos, equilibrando el ajuste del modelo y su complejidad mediante la penalización del número de parámetros.

Cadena de Markov Proceso Estocástico de Markov: Modelo matemático que describe una secuencia de eventos donde la probabilidad de cada evento futuro depende únicamente del estado actual. Sus componentes clave son:

- Espacio de estados (S): El conjunto finito de todos los posibles regímenes.
- Matriz de transición (P): Matriz que contiene las probabilidades de pasar de un estado a otro.

Codificación indicadora (*one-hot encoding*) Proceso para convertir variables categóricas (como los estados s_1, s_2) en vectores binarios (e.g., $[1, 0, 0, 0]^T$), un formato necesario para la estimación matricial.

Distribución Estacionaria (π) Vector de probabilidad que describe la distribución a largo plazo de una cadena de Markov, representando el equilibrio del sistema.

Continuidad de Lipschitz Propiedad matemática de una función que acota su tasa de cambio. En esta tesis, el estimador α_t satisface una cota de Lipschitz (Lema 4.4), pero la función de cuantización π es discontinua y no posee esta propiedad.

Ergodicidad Propiedad de una cadena de Markov que es irreducible (todo estado se alcanza desde cualquier otro en un número finito de pasos) y aperiódica (el máximo común divisor de los tiempos de retorno a cada estado es 1). Bajo estas condiciones, la cadena posee una única distribución estacionaria π y las frecuencias empíricas de visita convergen a π independientemente del estado inicial [31, 36].

Norma de Frobenius ($\|\cdot\|_F$) Para una matriz $A \in \mathbb{R}^{m \times n}$, la norma de Frobenius se define como $\|A\|_F = \sqrt{\sum_{i=1}^m \sum_{j=1}^n |a_{ij}|^2} = \sqrt{\text{tr}(A^T A)}$. En esta tesis se utiliza como función objetivo en el problema de estimación NNLS (Ecuación 6.1) para minimizar el error residual matricial.

SSRC Computación de Reservorio Estocásticamente Estructurada (*Stochastically Structured Reservoir Computing*): Extensión del modelo TCRA-Markov que incorpora una capa de computación de reservorio con estructura estocástica, permitiendo la captura de dinámicas no lineales.

ESN Red de Estado de Eco (*Echo State Network*): Tipo de red neuronal recurrente donde la matriz de pesos internos (reservorio) es fija y solo se entrena la capa de salida, lo que simplifica el entrenamiento.

K-medias Algoritmo de agrupamiento que particiona n observaciones en K clústeres minimizando la inercia intra-grupo. En esta tesis, se usa para discretizar la serie continua $\{\alpha_t\}$ en K estados de régimen económico.

Matriz de Transición (P) Matriz cuadrada estocástica que define las probabilidades de una cadena de Markov. Su elemento principal es:

- P_{ij} : Probabilidad de transitar al estado s_i desde el estado s_j .

NNLS Mínimos Cuadrados No Negativos: Algoritmo de optimización que resuelve un problema de mínimos cuadrados con la restricción de que los coeficientes deben ser no-negativos, asegurando probabilidades válidas.

Producto de Kronecker (\otimes) Operación entre matrices que resulta en una matriz de bloques más grande, utilizada para linealizar ecuaciones matriciales.

Vectorización ($\text{vec}(\cdot)$) Transformación lineal que convierte una matriz en un vector columna apilando sus columnas.

ϕ -mixing Mezcla Fuerte (ϕ -mixing): Propiedad estadística que mide la tasa a la que las observaciones lejanas en el tiempo se vuelven independientes, usada para probar propiedades asintóticas.

MS-AR Markov-Switching Autoregressive: Modelo autorregresivo con cambio de régimen que alterna entre parámetros según un proceso de Markov latente [18, 19]. En esta tesis funciona como modelo de referencia (*benchmark*); su estimación por máxima verosimilitud no convergió en ninguna de las series evaluadas.

Leaky-ESN Echo State Network con Integración de Fuga (*Leaky Integrator Echo State Network*): Variante de la red de estado de eco que incorpora un factor de integración $a \in (0, 1]$ en la actualización del estado del reservorio. En el Capítulo 6 se demuestra que satisface la Propiedad de Estado de Eco (ESP) bajo la condición $\rho(\mathbf{W}_{\text{res}}) < 1$.

Brecha espectral (γ) Diferencia entre el primer y el segundo autovalor en módulo de la matriz de transición \mathbf{P} . Cuantifica la velocidad de convergencia de la cadena de Markov a su distribución estacionaria: una brecha pequeña indica alta inercia de mercado.

Tiempo de mezcla (t_{mix}) Número de pasos necesarios para que la cadena de Markov aproxime su distribución estacionaria dentro de un margen ε . En los datos de combustibles hondureños, los tiempos de mezcla estimados oscilan entre 25 y 47 semanas.

RMSE Raíz del Error Cuadrático Medio (*Root Mean Square Error*): $\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2}$. Métrica primaria de evaluación predictiva en los capítulos empíricos.

Test de Diebold-Mariano Prueba estadística que compara la capacidad predictiva de dos modelos evaluando si la diferencia de sus errores de pronóstico es significativa [13]. Se emplea en los Capítulos 5 y 6 para validar la superioridad del SSRC frente al modelo Markoviano lineal.

ESP Propiedad de Estado de Eco (*Echo State Property*): Condición que exige que el estado interno de un reservorio dependa únicamente de la secuencia de entrada y no de la condición inicial. Se garantiza cuando el radio espectral del reservorio satisface $\rho(\mathbf{W}_{\text{res}}) < 1$.

Tabla de Notación

| Símbolo | Tipo | Descripción |
|-------------------------------|-----------------------|---|
| VARIABLES Y SERIES | | |
| $v(t)$ | Real (\mathbb{R}) | Valor de la serie temporal en el instante t |
| α_t | Real (\mathbb{R}) | Tasa de cambio relativa óptima (operador TCRA) |
| β_t | Real (\mathbb{R}) | Factor multiplicativo: $\beta_t = \alpha_t + 1$ |
| ε_t | Real (\mathbb{R}) | Ruido o residuo del modelo AR(1) local |
| S_t | Discreto | Estado del régimen económico en el instante t |
| PARÁMETROS | | |
| W | Entero | Tamaño de la ventana temporal (número de observaciones) |
| λ | $\in (0, 1]$ | Factor de decaimiento exponencial |
| K | Entero | Número de estados (regímenes) discretos |
| T | Entero | Longitud total de la serie temporal |
| W_{eff} | Real (\mathbb{R}) | Ventana efectiva: $W_{\text{eff}} = (1 - \lambda^W)/(1 - \lambda)$ |
| MATRICES Y OPERADORES | | |
| P, \hat{P} | $K \times K$ | Matriz de transición verdadera y estimada |
| C | $K \times K$ | Matriz de conteo de transiciones |
| π | Función | Función de cuantización $\pi: \mathbb{R} \rightarrow \mathcal{S}$ |
| $\boldsymbol{\pi}$ | Vector | Distribución estacionaria de la cadena de Markov |
| ESPACIOS Y CONJUNTOS | | |
| \mathcal{S} | Conjunto | Espacio de estados discretos $\{s_1, \dots, s_K\}$ |
| $\{\theta_j\}$ | Conjunto | Umbrales de la función de cuantización π |
| $\delta_{\text{mín}}$ | Real (\mathbb{R}) | Separación mínima entre umbrales: $\min_j \theta_{j+1} - \theta_j $ |
| VARIANTES DEL OPERADOR | | |
| TCRA | — | $W = T, \lambda = 1$ (global, caso límite) |
| TCRAM | — | $W < T, \lambda = 1$ (ventana móvil) |
| ETCRA | — | $W = T, \lambda \in (0, 1)$ (decaimiento global) |
| ETCRAM | — | $W < T, \lambda \in (0, 1)$ (ventana + decaimiento) |
| MÉTRICAS | | |
| RMSE | Real (\mathbb{R}) | Error cuadrático medio: $\sqrt{\frac{1}{n} \sum (y_i - \hat{y}_i)^2}$ |
| Exactitud | $\in [0, 1]$ | Proporción de estados correctamente clasificados |

Nota: En todo el documento se adopta el punto como separador decimal, siguiendo la convención habitual en la literatura matemática y econométrica anglosajona.

Capítulo 1

Introducción, Series de Estudio y Planteamiento del Problema

La economía hondureña no se entiende con un solo indicador. El crecimiento del Producto Interno Bruto (PIB) depende de la inflación, la tasa de desempleo y la balanza comercial [5, 12, 45], pero también de la sostenibilidad de la deuda pública, del flujo de Inversión Extranjera Directa (IED) y de la estabilidad del tipo de cambio [6, 44]. El resultado neto de estas interacciones se refleja en el Índice de Desarrollo Humano (IDH), que sintetiza la calidad de vida de la población. Honduras, como economía emergente, añade a esta ecuación la volatilidad propia de un entorno político incierto y la exposición a shocks externos. Analizar estas variables con rigor es necesario para quien pretenda planificar a mediano plazo.

La construcción de pronósticos fiables sobre estas variables requiere herramientas matemáticas formales. El análisis de series temporales es la herramienta natural. Pero en Honduras, construir pronósticos fiables tiene obstáculos concretos: los datos están dispersos entre varias instituciones¹, los métodos cuantitativos rara vez llegan a quienes toman decisiones, y el vacío resultante se llena con proyecciones cualitativas que carecen de respaldo formal.

Modelar series temporales en economías emergentes plantea dos exigencias que suelen entrar en conflicto: capturar las tendencias cuantitativas y, al mismo tiempo, interpretar los cambios abruptos de régimen que provocan la volatilidad interna y los shocks externos. Los modelos econométricos clásicos (ARIMA, GARCH, VAR) tienden a resolver una faceta a costa de la otra: o ajustan bien pero dicen poco sobre la estructura, o describen regímenes pero sin el rigor que exige un pronóstico confiable [9, 19, 29, 33, 42].

Este trabajo propone un enfoque híbrido que combina análisis determinista con infe-

¹El PIB es publicado por el Banco Central de Honduras (BCH), los precios de combustibles por la Comisión Administradora del Petróleo (CAP), y los indicadores sociales por el Instituto Nacional de Estadísticas (INE), cada uno con formatos, frecuencias y plataformas de acceso distintos.

rencia estocástica. Para ponerlo a prueba, se utilizan dos familias de series temporales con características deliberadamente opuestas: una serie macroeconómica de baja frecuencia (PIB) y un conjunto de series de mercado de alta frecuencia (combustibles). Esta dualidad permite evaluar si el marco TCRA-Markov funciona en contextos distintos [21, 28, 41].

1.1. Series de Estudio

Las series temporales de este estudio fueron seleccionadas por tres razones: su relevancia para la toma de decisiones en Honduras, la disponibilidad de datos históricos consistentes y su aptitud para ser modeladas con rigor matemático [21]. El Producto Interno Bruto y los precios de combustibles reflejan dinámicas influenciadas por políticas internas y shocks globales, y por eso sirven como banco de pruebas para evaluar la metodología TCRA propuesta en los Capítulos 2–6 [5, 19].

1.1.1. El PIB de Honduras como Serie Macroeconómica

El Producto Interno Bruto es el indicador canónico para medir la salud y el crecimiento de una economía. Su análisis es necesario para entender las tendencias a largo plazo y el impacto de eventos estructurales, como crisis políticas o sanitarias. Para este estudio se utilizan cuatro variantes del PIB de Honduras, con un enfoque en el período 2000–2023 para los análisis cuantitativos, aunque se presentan los gráficos históricos completos desde 1960.

PIB en dólares corrientes

El PIB a precios corrientes de mercado es la medida más amplia de la producción económica de un país [45]. Su evolución, mostrada en la Figura 1.1, ilustra la trayectoria de crecimiento nominal de la economía hondureña a lo largo de más de seis décadas; la curva refleja un crecimiento sostenido con una aceleración visible después de 2004, interrumpida temporalmente por la contracción del año 2020. Debido a estas características dinámicas, esta serie constituye la primera validación empírica para evaluar el comportamiento del operador TCRA global en el Capítulo 2.

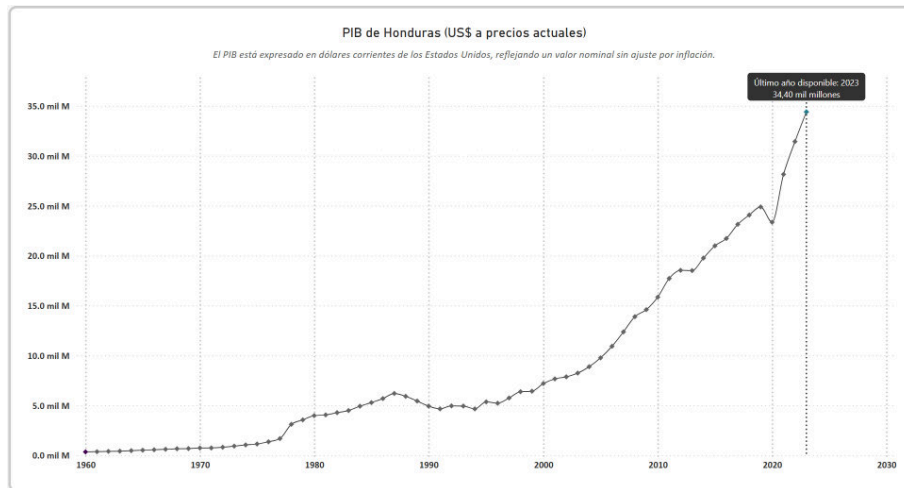


Figura 1.1: Evolución del PIB de Honduras en dólares corrientes, 1960–2023. Fuente: [45].

Tasa de crecimiento anual del PIB

La tasa de crecimiento real del PIB, ajustada por inflación, revela períodos de expansión y recesión [46]. La Figura 1.2 expone la volatilidad de este indicador. En particular, las caídas pronunciadas observadas en los años 2009 (−2.4%) y 2020 (−9.0%) hacen visibles los cambios bruscos de régimen estructural que justifican e impulsan el uso de modelos de cambio de régimen de Markov (*Markov-Switching*) a lo largo de este trabajo.

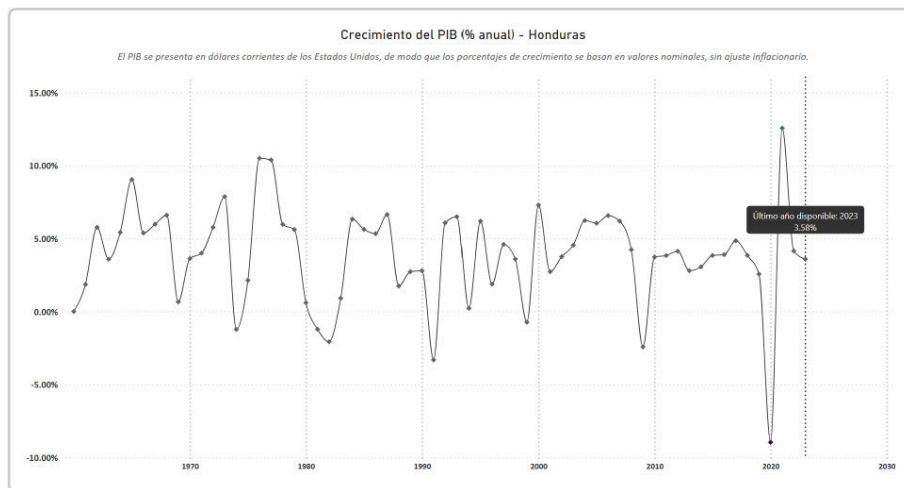


Figura 1.2: Tasa de crecimiento anual del PIB de Honduras (1961–2023).

PIB per cápita en dólares corrientes

El PIB per cápita ofrece una aproximación a la prosperidad promedio y al nivel de vida [48]. Su trayectoria, visible en la Figura 1.3, resulta clave para evaluar el desarrollo

económico en términos humanos. A diferencia del PIB de carácter agregado, esta serie en particular integra el impacto del crecimiento demográfico, lo que nos da un reflejo más apegado al estándar de vida promedio de la población.

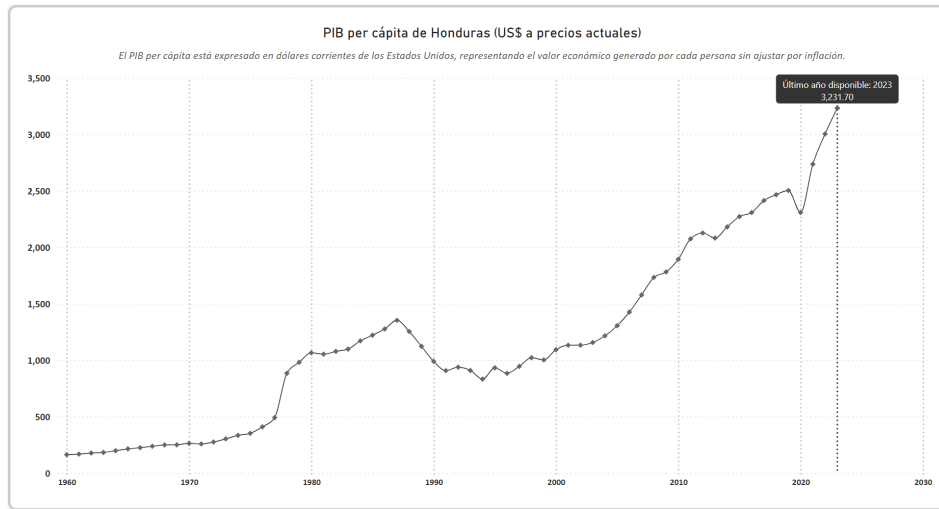


Figura 1.3: PIB per cápita de Honduras en dólares corrientes (1960–2023).

PIB per cápita en lempiras a precios constantes

Esta serie constituye la referencia principal para los desarrollos y pruebas empíricas de los Capítulos 2 y 4. Al estar denominada en moneda nacional y a precios constantes, elimina la distorsión generada por la inflación y las fluctuaciones del tipo de cambio, lo que ofrece la medida más fiel de la producción real por habitante de Honduras (Figura 1.4) [47].

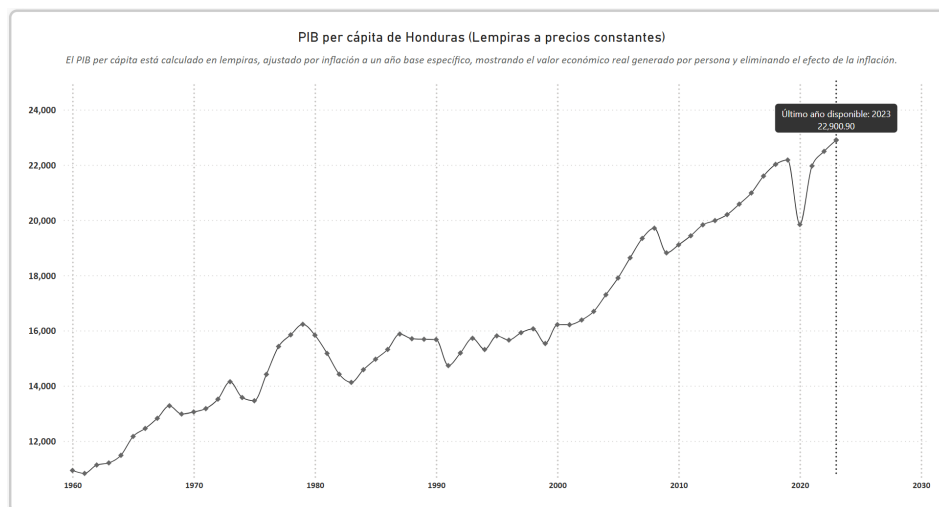


Figura 1.4: PIB per cápita de Honduras en lempiras a precios constantes (1960–2023).

La Tabla 1.1 resume numéricamente la evolución de estas series para el período de estudio principal (2000–2023), el cual sirve de base para la estimación de los modelos. Estos registros numéricos representan el punto de partida empírico que emplearemos para estimar los modelos TCRA y validar el desempeño de sus distintas variantes paramétricas en los capítulos posteriores.

Tabla 1.1: Evolución de series económicas seleccionadas de Honduras, 2000–2023.

| Año | PIB Dólares Corr. (M USD) | Crec. Anual (%) | PIB p.c. Dólares (USD) | PIB p.c. Lempiras Const. |
|------|---------------------------|-----------------|------------------------|--------------------------|
| 2000 | 7 186 638 029 | 7.29 | 1 092.6 | 16 214.3 |
| 2001 | 7 651 162 302 | 2.72 | 1 132.2 | 16 211.7 |
| 2002 | 7 858 255 413 | 3.75 | 1 132.5 | 16 381.9 |
| 2003 | 8 230 391 347 | 4.55 | 1 156.1 | 16 692.9 |
| 2004 | 8 869 299 234 | 6.23 | 1 215.1 | 17 296.5 |
| 2005 | 9 757 012 697 | 6.05 | 1 304.7 | 17 903.1 |
| 2006 | 10 917 477 066 | 6.57 | 1 425.8 | 18 634.1 |
| 2007 | 12 361 257 681 | 6.19 | 1 577.7 | 19 337.8 |
| 2008 | 13 881 731 876 | 4.23 | 1 732.5 | 19 708.9 |
| 2009 | 14 587 496 229 | -2.43 | 1 781.2 | 18 813.3 |
| 2010 | 15 839 344 592 | 3.73 | 1 893.3 | 19 104.7 |
| 2011 | 17 710 275 685 | 3.84 | 2 073.7 | 19 431.7 |
| 2012 | 18 528 554 398 | 4.13 | 2 126.0 | 19 828.9 |
| 2013 | 18 499 729 215 | 2.79 | 2 081.1 | 19 983.0 |
| 2014 | 19 756 533 972 | 3.06 | 2 179.9 | 20 199.1 |
| 2015 | 20 979 791 685 | 3.84 | 2 271.2 | 20 579.2 |
| 2016 | 21 717 604 952 | 3.89 | 2 307.3 | 20 982.7 |
| 2017 | 23 136 247 991 | 4.84 | 2 413.0 | 21 595.6 |
| 2018 | 24 067 750 760 | 3.84 | 2 464.6 | 22 019.3 |
| 2019 | 24 882 225 742 | 2.56 | 2 502.3 | 22 177.7 |
| 2020 | 23 352 232 484 | -8.97 | 2 307.6 | 19 838.3 |
| 2021 | 28 144 331 507 | 12.57 | 2 735.1 | 21 961.6 |
| 2022 | 31 426 041 807 | 4.14 | 3 003.3 | 22 491.3 |
| 2023 | 34 400 509 852 | 3.58 | 3 231.7 | 22 900.9 |

1.1.2. Precios de Combustibles como Serie de Alta Frecuencia

Como contrapunto al PIB, se analiza un segundo conjunto de datos: los precios semanales de los combustibles en Honduras. Estas series responden casi de inmediato a las fluctuaciones de los mercados internacionales y a las políticas regulatorias locales, lo que las hace complementarias a la serie macroeconómica. Los datos cubren de 2017 a 2025 y fueron obtenidos mediante extracción automatizada de datos web (*web scraping*) del portal Proceso Digital [37], con una posterior limpieza y estructuración para obtener series semanales consistentes. Estas series son la base empírica del Capítulo 5.

Precio de la Gasolina Superior

La evolución del precio de la gasolina superior es un indicador sensible tanto a los precios internacionales del crudo como a la demanda interna. La Figura 1.5 muestra su evolución semanal.



Figura 1.5: Precio semanal de la gasolina Superior en Honduras (2017–2025).

Precio de la Gasolina Regular

Siendo el combustible de mayor consumo, su precio tiene un impacto directo en el costo de transporte y, por ende, en la inflación al consumidor (Figura 1.6).



Figura 1.6: Precio semanal de la gasolina Regular en Honduras (2017–2025).

Precio del Diésel

El diésel desempeña un papel central en el transporte de carga y la industria, por lo que sus fluctuaciones afectan directamente los costos de producción y la cadena de suministro del país (Figura 1.7).

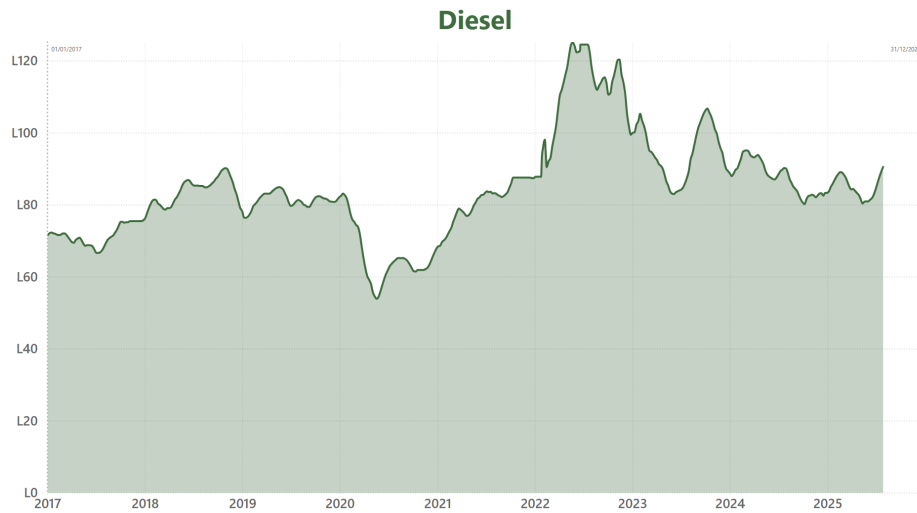


Figura 1.7: Precio semanal del Diésel en Honduras (2017–2025).

Precio del Kerosene

Utilizado principalmente para fines domésticos en zonas rurales, el precio del kerosene es un indicador con importantes implicaciones sociales, como se ilustra en la Figura 1.8.



Figura 1.8: Precio semanal del Kerosene en Honduras (2017–2025).

Las cuatro series de precios minoristas de combustibles en Honduras exhiben trayectorias

temporales sumamente correlacionadas, lo cual es de esperar dado que todas responden a las variaciones del precio del petróleo de referencia West Texas Intermediate (WTI) y se ajustan mediante una misma estructura regulatoria gubernamental administrada por la Secretaría de Energía. Las correlaciones lineales de Pearson superan en todos los casos el 0.96. No obstante, el Kerosene destaca por su alta volatilidad residual y el Diésel por su menor velocidad de ajuste en momentos de caídas abruptas, lo que justifica evaluarlas de manera independiente para la detección y pronóstico de estados Markovianos de mercado en el Capítulo 5.

Estas dos familias de series (el PIB como indicador macroeconómico de baja frecuencia y los combustibles como variables comerciales de alta frecuencia y rápida propagación) ofrecen un banco de pruebas heterogéneo y riguroso. La efectividad del modelo en ambos entornos validará la generalidad del enfoque desarrollado.

1.2. Revisión de Literatura y Posicionamiento

La presente sección sitúa la contribución de esta tesis en el contexto de tres corrientes de la literatura: (i) modelos de cambio de régimen para series temporales, (ii) métodos de estimación de matrices de transición, y (iii) computación de reservorio aplicada a series económicas.

1.2.1. Modelos de Cambio de Régimen

El enfoque más influyente para modelar transiciones entre estados económicos es el modelo autorregresivo de cambio de régimen de Markov (*Markov-Switching Autoregressive*, MS-AR) de Hamilton [18, 19], en el cual los parámetros de un proceso AR cambian según una cadena de Markov con probabilidades de transición no observables. La estimación se realiza por máxima verosimilitud (MLE) mediante el algoritmo EM, que requiere la evaluación de la verosimilitud completa por filtrado recursivo. A pesar de su elegancia teórica, el modelo MS-AR presenta dos limitaciones prácticas relevantes para esta tesis:

1. La función de verosimilitud es multimodal, lo que hace que la convergencia del algoritmo EM dependa de la inicialización y no garantice un óptimo global [19, 29].

2. La complejidad computacional crece exponencialmente con el número de estados K , limitando las aplicaciones con más de 3–4 regímenes [4].

Entre los modelos alternativos de cambio de régimen, el TAR (*Threshold AR*) [40, 41] determina el régimen mediante una variable umbral observable, con una función de transición discontinua análoga a la función de cuantización π del modelo TCRA-Markov. Los modelos STAR y LSTAR [39] generalizan esa idea con una transición suave (logística o exponencial), lo que permite cambios graduales entre estados.

En un contexto diferente, los HMM gaussianos [38] postulan estados latentes con emisiones continuas. Aunque comparten la estructura estocástica con el modelo propuesto, sus estados son *ocultos*: la correspondencia con un régimen económico es ambigua. Los modelos factoriales dinámicos [27], más generales, combinan factores latentes con cambio de régimen a costa de una estimación considerablemente más compleja.

La contribución del marco TCRA-Markov respecto a estos modelos es la separación explícita entre la etapa de discretización (determinista, basada en el operador TCRA) y la etapa de modelado estocástico (estimación de la matriz de transición vía NNLS). En los modelos clásicos MS-AR, la función de verosimilitud no convexa imposibilita garantizar teóricamente que el algoritmo EM converja al máximo global, dependiendo críticamente de la inicialización de los parámetros. Por el contrario, la formulación de mínimos cuadrados no negativos (NNLS) utilizada en este trabajo es un problema estrictamente convexo. Esto asegura de forma determinista la convergencia hacia un óptimo global único y evita la inestabilidad numérica característica de los métodos iterativos.

1.2.2. Estimación de Matrices de Transición

Para cadenas de Markov con estados observados, el estimador de máxima verosimilitud de la matriz de transición coincide con las frecuencias relativas de transición [3, 36]. Este estimador es consistente y asintóticamente normal (Proposición 3.6), pero no incorpora información estructural ni restricciones de suavidad.

El enfoque NNLS adoptado en esta tesis [11, 30] reformula la estimación como un problema de mínimos cuadrados con restricciones de no-negatividad. Al ser un problema

convexo, la convergencia al óptimo global no depende de la inicialización. La propia restricción de no-negatividad actúa como regularizador implícito, lo que estabiliza las estimaciones con muestras pequeñas. Además, la formulación matricial (Capítulo 4) permite incorporar restricciones adicionales de forma natural.

1.2.3. Computación de Reservorio para Series Económicas

Los modelos de computación de reservorio (*Reservoir Computing*, RC) fueron introducidos por Jaeger [25] con las Redes de Estado de Eco (*Echo State Networks*, ESN) y han demostrado capacidad universal de aproximación [14, 16]. La propiedad fundamental es el Estado de Eco (ESP): la dinámica interna del reservorio “olvida” sus condiciones iniciales, garantizando que la respuesta dependa solo de la entrada [32, 49].

La extensión SSRC (*Stochastically Structured Reservoir Computers*) propuesta por Banegas y Vides [7] integra la estructura de Markov directamente en la arquitectura del reservorio, utilizando matrices de transición estocásticas como pesos de conectividad. Esta integración es la base teórica del Capítulo 6, donde se demuestra que el modelo TCRA-Markov es un caso particular de los SSRC (Teorema de Inclusión).

1.2.4. Identificación del Vacío en la Literatura

La revisión anterior revela un vacío en la intersección de tres áreas: (i) los modelos de cambio de régimen clásicos no aprovechan tasas de cambio relativas observables para la discretización; (ii) la estimación NNLS no ha sido aplicada sistemáticamente a matrices de transición económicas; (iii) la conexión entre modelos de Markov lineales y computación de reservorio no ha sido formalizada. Esta tesis aborda simultáneamente estos tres vacíos mediante la secuencia: operador TCRA, integración con cadenas de Markov, estimación NNLS y extensión SSRC.

La Tabla 1.2 resume las diferencias clave entre el modelo propuesto y los principales competidores. A través de esta comparación, se hace evidente que el modelo TCRA-Markov destaca por trabajar con estados directamente observados, una estimación estrictamente convexa que asegura convergencia global, y la garantía teórica de validez estocástica

por construcción. Estas ventajas metodológicas justifican y sustentan el desarrollo de la propuesta en este trabajo.

Tabla 1.2: Comparación de modelos de cambio de régimen según siete criterios clave.

| Criterio | TCRA-Markov | MS-AR | TAR | HMM |
|---------------------|----------------------|------------------|---------------------|-----------------------|
| Estados | Observados (π) | Latentes (EM) | Observados (umbral) | Latentes (Baum-Welch) |
| Estimación de P | NNLS (convexo) | MLE (no convexo) | N/A | MLE (no convexo) |
| Convergencia | Global | Local | Global | Local |
| Validez estocástica | Por construcción | No garantizada | N/A | No garantizada |
| Interpretabilidad | Alta (regímenes) | Media | Alta | Baja |
| Complejidad | $O(K^2T)$ | $O(K^2T)$ | $O(KT)$ | $O(K^2T)$ |
| Extensión no lineal | SSRC (Cap 6) | No estándar | STAR | No estándar |

1.3. Planteamiento del Problema y Hoja de Ruta

1.3.1. Hoja de Ruta Metodológica: Del Análisis Determinista a la Inferencia Estocástica

El trabajo sigue un recorrido que parte del análisis determinista y avanza hacia la inferencia estocástica, organizado en tres fases.

Fase 1. El Capítulo 2 formaliza la familia de transformaciones TCRA y sus variantes paramétricas (TCRAM, ETCRA, ETCRAM). Se demuestran existencia, unicidad, consistencia asintótica y estabilidad ante perturbaciones. El propósito de esta etapa es disponer de un operador que convierta una serie temporal continua en una secuencia de regímenes económicos con significado empírico.

Fase 2. Los Capítulos 3 y 4 integran el operador TCRA con cadenas de Markov. Se aplica un estimador de Mínimos Cuadrados No Negativos (NNLS) para la matriz de transición, lo que garantiza por construcción la validez estocástica de los parámetros [19].

Fase 3. Los Capítulos 5 y 6 validan el marco sobre los precios semanales de combustibles hondureños (2017–2025): optimización exhaustiva de hiperparámetros, análisis espectral y comparación con modelos de referencia. Se propone además una extensión

no lineal mediante Computación de Reservorio Estructurado (SSRC), que generaliza el marco lineal y amplía la capacidad representacional del modelo.

1.3.2. Objetivos y Alcance de la Investigación

El objetivo general de este trabajo es desarrollar y validar una metodología para el análisis de series temporales económicas y financieras, abarcando desde la fundamentación matemática hasta su contraste empírico. Para lograrlo, se plantean los siguientes objetivos específicos:

- Desarrollar la fundamentación del operador TCRA y demostrar formalmente sus propiedades de existencia, unicidad y estabilidad matemática.
- Diseñar el esquema de acoplamiento del operador con cadenas de Markov y formular el estimador estocástico mediante mínimos cuadrados no negativos.
- Validar empíricamente la predictibilidad del modelo lineal sobre los precios semanales de los combustibles en Honduras.
- Extender el marco lineal al dominio no lineal mediante una arquitectura de computación de reservorio estructurada estocásticamente (SSRC).

La metodología resultante puede verificarse formalmente y aplicarse sin optimización iterativa. Esas dos propiedades la hacen útil tanto en investigación matemática como en planificación económica aplicada.

1.3.3. Marco Metodológico

Los Capítulos 2 y 3 establecen la base teórica: el operador TCRA y sus variantes, las propiedades de ϕ -mixing y la estimación asintótica que sustentan la predicción estocástica de regímenes. El Capítulo 4 integra esos elementos con la estimación NNLS de la matriz de transición, cuya formulación convexa produce parámetros estocásticamente válidos sin depender de inicializaciones.

La aplicación a precios semanales de combustibles hondureños (2017–2025) ocupa el Capítulo 5, que incluye optimización de hiperparámetros, análisis espectral y comparación

con cuantiles fijos y MS-AR. El Capítulo 6 extiende el marco con Computación de Reservorio Estructurada (SSRC) como generalización no lineal, y el Capítulo 7 cierra con síntesis, limitaciones y líneas abiertas.

Los pronósticos resultantes, aunque no son certezas, proveen un mapa de escenarios posibles para la planificación a mediano plazo en contextos complejos y volátiles.

1.3.4. Estructura del Documento

La Figura 1.9 resume esquemáticamente la estructura de la tesis y las dependencias entre capítulos, mostrando el flujo metodológico desde la construcción del operador TCRA hasta las conclusiones. Esto facilita al lector la navegación del documento y la identificación de los prerequisites teóricos de cada capítulo.

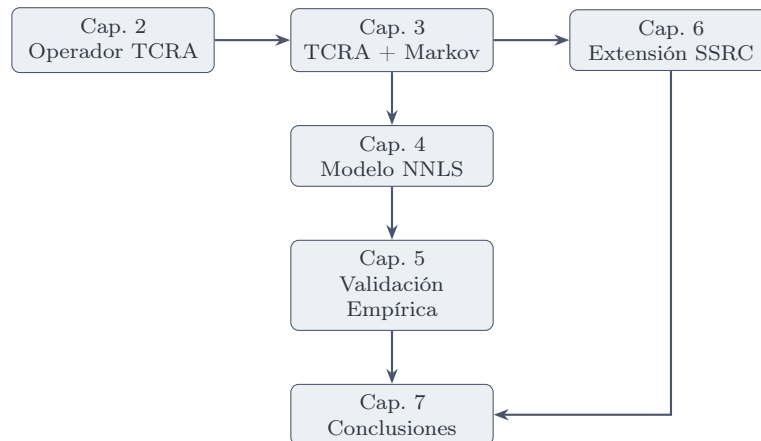


Figura 1.9: Estructura de la tesis y dependencias entre capítulos.

Más allá de la organización por capítulos, conviene explicitar cómo se relacionan entre sí los resultados teóricos que articulan esta tesis. La Figura 1.10 muestra esa cadena de dependencias. Todo parte de los axiomas del operador TCRA (Capítulo 2), cuya rama izquierda establece las propiedades estocásticas (mezcla y normalidad asintótica del estimador) y cuya rama derecha desarrolla la estimación (formulación NNLS y existencia de solución). Ambas ramas confluyen en la validación empírica del modelo lineal (Capítulo 5). En el Capítulo 6, la estimación NNLS se combina con el marco teórico de computación de reservorio (recuadro lateral) para demostrar, mediante el Teorema de Inclusión, que el modelo lineal es un caso particular de este marco no lineal, cerrando la jerarquía planteada.

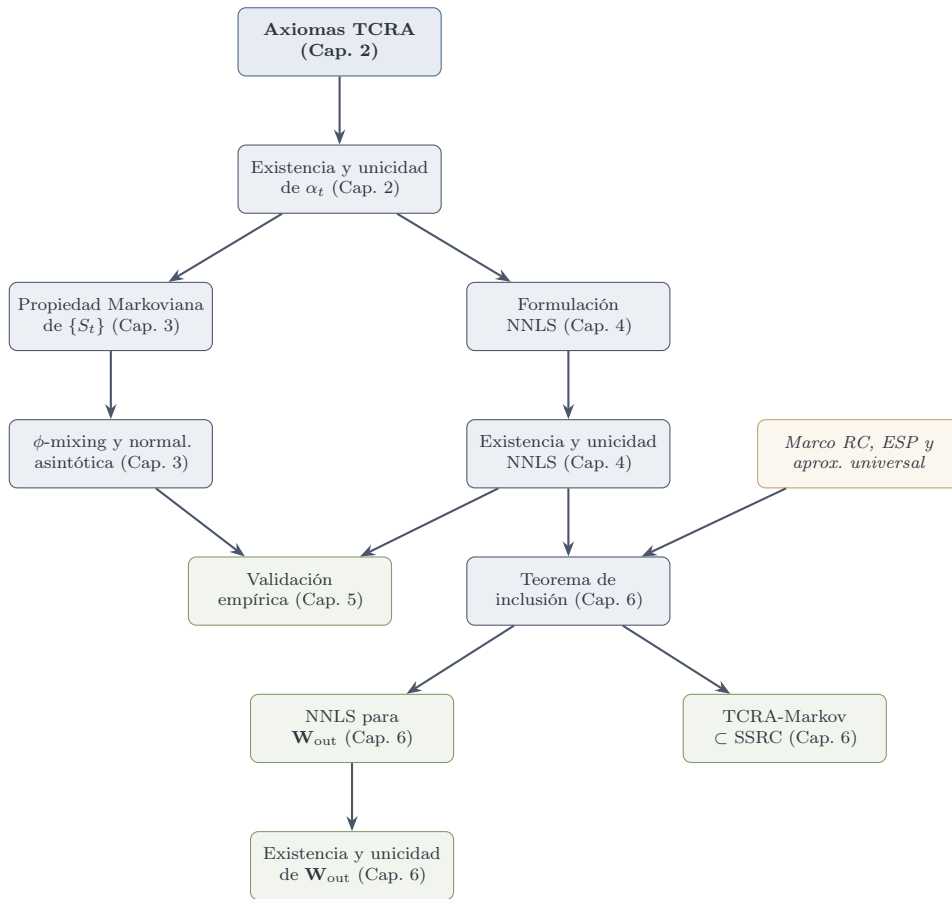


Figura 1.10: Cadena de dependencias lógicas entre los resultados teóricos principales de la tesis.

Todo el código fuente, los datos y los scripts de reproducción están disponibles en el repositorio público del proyecto².

²<https://github.com/NORSAB/TCRA-Cadenas-Markov-SSRC-Series-Economicas>. DOI: 10.5281/zenodo.18752541.

Capítulo 2

Tasa de Cambio Relativa Autorregresiva y sus Variantes

2.1. Introducción y Motivación Matemática

Las series temporales financieras y económicas son, por lo general, no estacionarias y ruidosas. Modelarlas con cadenas de Markov o procesos de cambio de régimen (*Markov-Switching*) exige un espacio de estados que refleje transiciones estables, lo cual el dato bruto rara vez proporciona.

Este capítulo construye la base matemática para transformar la serie temporal antes de discretizar sus estados. Se desarrolla la familia de operadores TCRA (*Tasa de Cambio Relativa Autorregresiva*) junto con tres variantes: TCRAM (ventana móvil), ETCRA (decaimiento exponencial) y ETCRAM (combinación de ambas). El operador transforma una serie discreta $\{v(t)\}_{t=0}^T$ en una secuencia continua de tasas de cambio $\{\alpha_t\}$ que minimiza $v(t) \approx (1 + \alpha_t)v(t-1)$ de forma convexa. Existencia, unicidad, continuidad y estabilidad frente a perturbaciones se demuestran rigurosamente, lo cual asegura que la discretización posterior con K-medias produzca agrupaciones consistentes.

Todas las variantes admiten solución analítica cerrada y su costo computacional es menor que el de métodos iterativos como la máxima verosimilitud en MS-AR.

2.2. Axiomática del Operador de Tasa de Cambio Local

Para que la transformación $\mathcal{T} : \mathbb{R}^{T+1} \rightarrow \mathbb{R}^T$ que genera la secuencia $\{\alpha_t\}$ sea compatible con una posterior discretización en cadena de Markov, debe satisfacer un conjunto mínimo de propiedades algebraicas y analíticas. La familia TCRA se construye de modo que cumpla los siguientes cuatro axiomas:

Axioma 1 (Localidad Acotada y Dependencia Estricta): El estado continuo de la

tasa de cambio en un instante t , denotado como α_t , depende exclusivamente de un subconjunto cerrado del pasado inmediato $[t - W, t]$ que involucra las observaciones $\{v(t - W), \dots, v(t)\}$ para constituir las W transiciones locales. Esta dependencia captura la dinámica inercial de los mercados y rechaza la influencia de estados temporales infinitamente lejanos.

Axioma 2 (Proporcionalidad Geométrica Local): La relación entre valores consecutivos se modela como $v(t) = \beta_t v(t - 1) + \varepsilon_t$. La tasa de cambio $\alpha_t = \beta_t - 1$ representa la variación relativa local entre observaciones adyacentes.

Axioma 3 (Unicidad y Convexidad Global): El escalar α_t se obtiene como el mínimo global de una funcional de costo estrictamente convexa en L^2 , lo que garantiza unicidad de la solución siempre que la ventana local no sea degenerada (es decir, exista al menos un elemento $v(k - 1) \neq 0$ dentro del horizonte considerado).

Axioma 4 (Estabilidad de Lipschitz frente a Perturbaciones): Existe una constante $C > 0$ tal que variaciones acotadas en la serie observada $\|v' - v\|_\infty \leq \epsilon$ inducen desviaciones finitamente controladas en la proyección:

$$|\alpha' - \alpha| \leq C\epsilon.$$

Esta estabilidad garantiza robustez del operador continuo; la robustez de la discretización posterior se mantiene siempre que las tasas estimadas no se encuentren arbitrariamente cerca de los umbrales de cuantización.

Las demostraciones que siguen verifican que las diferentes instanciaciones del modelo base (TCRA, TCRAM, ETCRA y ETCRAM) satisfacen integralmente estos axiomas.

2.3. Marco General Unificado

Sea $T > 0$ un entero y $\Sigma_T = \{v(0), v(1), \dots, v(T)\}$ una serie temporal con $v(t) \in \mathbb{R}$. Dados un tamaño de ventana W con $1 \leq W \leq T$ (que involucra $W + 1$ observaciones desde $v(t - W)$ hasta $v(t)$ para definir las W transiciones locales) y un factor de decaimiento

$\lambda \in (0, 1]$, se define para cada instante $t \geq W$ el **estimador autorregresivo óptimo local** como el minimizador de la funcional de costo cuadrática ponderada:

$$\hat{\beta}_{t,\lambda,W} := \arg \min_{\beta \in \mathbb{R}} \sum_{k=t-W+1}^t \lambda^{t-k} (v(k) - \beta v(k-1))^2. \quad (2.1)$$

En consecuencia, la **tasa de cambio relativa óptima generalizada** se establece como:

$$\alpha_{t,\lambda,W} := \hat{\beta}_{t,\lambda,W} - 1. \quad (2.2)$$

La funcional de costo cuadrática en la Ecuación (2.1) se elige por razones analíticas y estadísticas. Por un lado, la estructura cuadrática asegura la convexidad estricta del problema de optimización, lo que garantiza que exista un único mínimo global (Teorema 2.1). Por otro lado, la ponderación exponencial λ^{t-k} se introduce para que la solución óptima coincida con el estimador de mínimos cuadrados generalizados (GLS) bajo heterocedasticidad temporal, dando más peso a las transiciones recientes del mercado.

Las cuatro variantes se recuperan como casos particulares:

| Variante | Parámetros | Notación |
|----------|----------------------|------------------------|
| TCRA | $\lambda = 1, W = T$ | α_T |
| TCRAM | $\lambda = 1, W < T$ | $\alpha_t^{(W)}$ |
| ETCRA | $\lambda < 1, W = T$ | $\alpha_{T,\lambda}$ |
| ETCRAM | $\lambda < 1, W < T$ | $\alpha_{t,\lambda,W}$ |

La estructura jerárquica y la relación lógica de la familia de operadores se ilustran en la Figura 2.1.

En este diagrama, el nodo central representa la formulación global canónica ($W = T, \lambda = 1$). La incorporación de memoria restringida ($W < T$) y decaimiento exponencial ($\lambda < 1$) producen las extensiones laterales, que convergen en el operador generalizado ETCRAM.

Cuando $\lambda < 1$, las variantes con decaimiento exponencial (ETCRA y ETCRAM) equivalen formalmente a un estimador de mínimos cuadrados generalizados (GLS) con

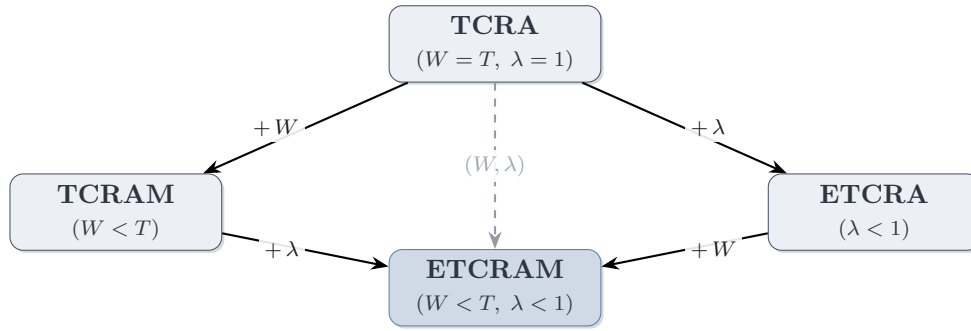


Figura 2.1: Jerarquía paramétrica del operador TCRA y sus variantes.

varianza no constante (heterocedasticidad temporal geométrica), lo cual se demuestra en el Teorema 2.5.

2.4. Teorema Fundamental del Operador

Dado que la variante híbrida dinámica (ETCRAM) subsume paramétricamente a las demás formulaciones (TCRA, TCRAM y ETCRA), las propiedades se demuestran directamente sobre la forma generalizada $\alpha_{t,\lambda,W}$ de la Ecuación (2.1). Los resultados para cada variante se obtienen como casos particulares por especialización de parámetros.

2.4.1. Existencia Ponderada y Unicidad Global

Teorema 2.1. *Para cualquier instante de tiempo válido $t \geq W$, si el subconjunto de observaciones contenidas en la ventana local de rezagos exhibe al menos un suceso medible no nulo (formalmente, $\exists k \in [t - W + 1, t]$ tal que $v(k - 1) \neq 0$), entonces la funcional de costo en el espacio topológico L^2 , definida como $f(\beta) = \sum_{k=t-W+1}^t \lambda^{t-k} (v(k) - \beta v(k-1))^2$, admite un y solo un minimizador global $\hat{\beta}_{t,\lambda,W}$ formulado cerradamente por:*

$$\hat{\beta}_{t,\lambda,W} = \frac{\sum_{k=t-W+1}^t \lambda^{t-k} v(k) v(k-1)}{\sum_{k=t-W+1}^t \lambda^{t-k} v(k-1)^2}. \quad (2.3)$$

Demostración. Sea $a = \sum_{k=t-W+1}^t \lambda^{t-k} v(k-1)^2$. Dado que $\lambda \in (0, 1]$, cada peso λ^{t-k} es estrictamente positivo. Por la hipótesis de no degeneración, existe al menos un k tal que

$v(k-1) \neq 0$, de donde $a > 0$. Entonces $f(\beta) = a\beta^2 - 2b\beta + c$ con $a > 0$ es estrictamente convexa y coerciva ($f(\beta) \rightarrow +\infty$ cuando $|\beta| \rightarrow \infty$). El mínimo global se alcanza en el único punto donde $f'(\beta) = 2a\beta - 2b = 0$, es decir, $\hat{\beta}_{t,\lambda,W} = b/a$, que coincide con la expresión (2.3). \square

Corolario 2.2. *Bajo las mismas condiciones del Teorema 2.1, la tasa de cambio relativa*

$$\alpha_{t,\lambda,W} = \hat{\beta}_{t,\lambda,W} - 1 \quad (2.4)$$

está bien definida y es única para cada $t \geq W$.

Este resultado garantiza que los algoritmos de discretización posteriores (e.g., K-medias) reciben una entrada bien definida, evitando divisiones por cero o indeterminaciones numéricas.

2.4.2. Consistencia Asintótica

Teorema 2.3 (Consistencia del estimador). *Supóngase que la serie temporal $\{v(k)\}$ satisface localmente el modelo autorregresivo de primer orden $v(k) = \beta^*v(k-1) + \varepsilon_k$, donde $\{\varepsilon_k\}$ es ruido blanco con $\mathbb{E}[\varepsilon_k] = 0$ y $\mathbb{E}[\varepsilon_k^2] = \sigma^2 < \infty$, y $|\beta^*| < 1$ (estacionariedad). Entonces, cuando $W \rightarrow \infty$ con $\lambda = 1$ (ponderación uniforme),*

$$\hat{\beta}_{t,\lambda,W} \xrightarrow{p} \beta^* \quad \text{y, en consecuencia,} \quad \alpha_{t,\lambda,W} \xrightarrow{p} \beta^* - 1.$$

Para el caso $\lambda < 1$, el estimador no es formalmente consistente en el sentido asintótico tradicional, dado que la ventana efectiva está limitada por la constante $W_{eff} = 1/(1-\lambda)$, lo que resulta en una varianza asintótica estrictamente positiva. En este régimen, se opera bajo un principio de consistencia práctica local.

Demostración. Sustituyendo $v(k) = \beta^*v(k-1) + \varepsilon_k$ en la expresión del estimador (2.3), se obtiene

$$\hat{\beta}_{t,\lambda,W} = \frac{\sum_{k=t-W+1}^t \lambda^{t-k} (\beta^*v(k-1) + \varepsilon_k) v(k-1)}{\sum_{k=t-W+1}^t \lambda^{t-k} v(k-1)^2} = \beta^* + \frac{N_W}{D_W},$$

donde $N_W = \sum_k \lambda^{t-k} \varepsilon_k v(k-1)$ y $D_W = \sum_k \lambda^{t-k} v(k-1)^2$.

Denominador. Dado que $|\beta^*| < 1$, el proceso $\{v(k)\}$ es estacionario y ergódico, con $\mathbb{E}[v(k-1)^2] = \sigma^2/(1-\beta^{*2}) > 0$. Sea $\Lambda_W = \sum_{k=t-W+1}^t \lambda^{t-k}$. Por la Ley de los Grandes Números para sucesiones ponderadas estacionarias [2],

$$\frac{D_W}{\Lambda_W} \xrightarrow{p} \mathbb{E}[v(k-1)^2] = \frac{\sigma^2}{1-\beta^{*2}} > 0.$$

Numerador. Como ε_k es independiente de $\mathcal{F}_{k-1} = \sigma(v(0), \dots, v(k-1))$, se tiene $\mathbb{E}[\varepsilon_k v(k-1)] = 0$. Para el caso $\lambda = 1$ (ponderación uniforme), $\Lambda_W = W$. Por la desigualdad de Chebyshev,

$$\text{Var}\left(\frac{N_W}{W}\right) = \frac{\sum_{k=t-W+1}^t \mathbb{E}[\varepsilon_k^2 v(k-1)^2]}{W^2} \leq \frac{\sigma^2 \mathbb{E}[v(k-1)^2]}{W} = \mathcal{O}\left(\frac{1}{W}\right) \rightarrow 0.$$

Esto garantiza la convergencia en probabilidad a cero. Por el contrario, si $\lambda < 1$ es fijo, la varianza de N_W/Λ_W no converge a cero cuando $W \rightarrow \infty$; en su lugar, converge a un valor estrictamente positivo $\frac{1-\lambda}{1+\lambda} \sigma^2 \mathbb{E}[v(k-1)^2]$, lo que confirma que el estimador con decaimiento exponencial conserva una varianza residual finita y no es consistente en el sentido asintótico tradicional.

Conclusión. Para $\lambda = 1$, por el Lema de Slutsky, $N_W/D_W = (N_W/\Lambda_W)/(D_W/\Lambda_W) \xrightarrow{p} 0/\mathbb{E}[v(k-1)^2] = 0$. Por tanto, $\hat{\beta}_{t,\lambda,W} \xrightarrow{p} \beta^*$. \square

Observación (Alcance de la hipótesis de estacionariedad). La hipótesis $|\beta^*| < 1$ es esencial para la prueba, pues garantiza que $\mathbb{E}[v(k-1)^2]$ sea finito. Sin embargo, las series económicas reales (como el PIB nominal) frecuentemente exhiben tendencia, con β^* cercano o igual a 1. En tales casos, el Teorema 2.3 no es directamente aplicable a la serie $v(t)$ en niveles.

Ahora bien, el operador $\alpha_{t,\lambda,W}$ actúa como un filtro de diferenciación local: $\alpha_t \approx v(t)/v(t-1) - 1$, que es análogo a la tasa de crecimiento. Para una serie con raíz unitaria ($\beta^* = 1$), la serie de tasas $\alpha_t = v(t)/v(t-1) - 1$ es generalmente estacionaria. En la implementación empírica, la consistencia se verifica por la estacionariedad de $\{\alpha_t\}$, no necesariamente de $\{v(t)\}$. La validación con tests ADF aplicados a la serie $\{\alpha_t\}$ se presenta en el Capítulo 5.

2.4.3. Estabilidad de Lipschitz

El Axioma 4 requiere que el operador sea robusto frente a perturbaciones acotadas en los datos de entrada. El siguiente lema formaliza esta propiedad.

Lema 2.4 (Estabilidad tipo Lipschitz). *Sean $\{v(k)\}$ y $\{v'(k)\}$ dos series tales que $\|v' - v\|_\infty \leq \epsilon$ (perturbación uniforme acotada). Defínase $S_2 = \sum_{k=t-W+1}^t \lambda^{t-k} v(k-1)^2 > 0$. Entonces*

$$|\alpha'_{t,\lambda,W} - \alpha_{t,\lambda,W}| \leq C_t \cdot \epsilon,$$

donde C_t es una constante que depende de v , λ y W , satisfaciendo $C_t = \mathcal{O}(\|v\|_{\ell^2}/S_2)$.

Demostración. Sean $S_1 = \sum_k \omega_k v(k)v(k-1)$ y $S'_1 = \sum_k \omega_k v'(k)v'(k-1)$, con $\omega_k = \lambda^{t-k}$. Escribiendo $v'(k) = v(k) + \delta(k)$ con $|\delta(k)| \leq \epsilon$:

$$S'_1 - S_1 = \sum_k \omega_k [v(k)\delta(k-1) + \delta(k)v(k-1) + \delta(k)\delta(k-1)].$$

Por la desigualdad de Cauchy-Schwarz y la cota $|\delta(k)| \leq \epsilon$:

$$|S'_1 - S_1| \leq 2\epsilon \sum_k \omega_k |v(k)| \cdot |v(k-1)| + \epsilon^2 \sum_k \omega_k \leq 2\epsilon \Lambda_W \|v\|_{\ell^2}^2 / W + \epsilon^2 \Lambda_W.$$

De manera análoga, $|S'_2 - S_2| \leq 2\epsilon \sum_k \omega_k |v(k-1)| + \epsilon^2 \Lambda_W$. Como $\alpha = S_1/S_2 - 1$, la diferencia $|\alpha' - \alpha| = |S'_1/S'_2 - S_1/S_2|$ se acota mediante la fórmula del cociente:

$$\left| \frac{S'_1}{S'_2} - \frac{S_1}{S_2} \right| = \frac{|S'_1 S_2 - S_1 S'_2|}{S_2 S'_2} \leq \frac{S_2 |S'_1 - S_1| + |S_1| |S'_2 - S_2|}{S_2 S'_2}.$$

Dado que $S'_2 \geq S_2 - \mathcal{O}(\epsilon) > 0$ para ϵ suficientemente pequeño, todos los términos del numerador son lineales en ϵ (los cuadráticos en ϵ son de orden inferior), obteniéndose $|\alpha' - \alpha| \leq C_t \epsilon$ con $C_t = \mathcal{O}(\|v\|_{\ell^2}/S_2)$. \square

2.4.4. Optimalidad de Gauss-Markov

Aunque en los modelos autorregresivos clásicos se asume homocedasticidad (varianza constante del error), el operador generalizado TCRA flexibiliza este supuesto. Al introducir

el decaimiento exponencial $\lambda < 1$, se modela una estructura de heterocedasticidad temporal para dar mayor relevancia a la información reciente.

Teorema 2.5 (Optimalidad GLS del estimador). *Considérese el modelo lineal local $v(k) = \beta v(k-1) + \varepsilon_k$. Supóngase que las perturbaciones $\{\varepsilon_k\}$ satisfacen las condiciones del Teorema de Gauss-Markov generalizado: poseen media cero ($\mathbb{E}[\varepsilon_k] = 0$), son incorrelacionadas temporalmente ($\mathbb{E}[\varepsilon_i \varepsilon_j] = 0$ para $i \neq j$) y su estructura de covarianza es conocida salvo un factor de escala constante. Específicamente, se asume que las variaciones históricas son heterocedásticas, con una varianza residual que aumenta con la antigüedad (y por ende su precisión decrece geométricamente): $\text{Var}(\varepsilon_k) = \sigma^2 / \lambda^{t-k}$, donde $\lambda \in (0, 1]$. Entonces el estimador $\hat{\beta}_{t,\lambda,W}$ definido en (2.3) coincide con el **Estimador de Mínimos Cuadrados Generalizados (GLS)** del modelo y constituye el estimador lineal insesgado de mínima varianza (BLUE).*

Demostración. Bajo el modelo $\mathbf{v} = \beta \mathbf{v}_{-1} + \varepsilon$ con $\text{Cov}(\varepsilon) = \sigma^2 \mathbf{\Omega}$, donde $\mathbf{\Omega} = \text{diag}(\lambda^{-(t-k)})_{k=t-W+1}^t$, el estimador GLS es

$$\hat{\beta}_{\text{GLS}} = (\mathbf{v}_{-1}^T \mathbf{\Omega}^{-1} \mathbf{v}_{-1})^{-1} \mathbf{v}_{-1}^T \mathbf{\Omega}^{-1} \mathbf{v}.$$

Como $\mathbf{\Omega}^{-1} = \text{diag}(\lambda^{t-k})$, se tiene

$$\hat{\beta}_{\text{GLS}} = \frac{\sum_k \lambda^{t-k} v(k) v(k-1)}{\sum_k \lambda^{t-k} v(k-1)^2} = \hat{\beta}_{t,\lambda,W},$$

que coincide exactamente con el estimador ETCRAM. La propiedad BLUE se sigue del Teorema de Gauss-Markov generalizado [1, 2]. \square

Esta relación matemática aclara cómo se vincula λ con la varianza residual. La ecuación $\text{Var}(\varepsilon_k) = \sigma^2 \lambda^{-(t-k)}$ implica que la varianza de las perturbaciones crece de forma geométrica a medida que retrocedemos en el tiempo ($k < t$ con $\lambda < 1$). De este modo, los datos históricos lejanos tienen una mayor incertidumbre y, por ende, el estimador GLS les asigna un peso menor de forma natural.

2.5. Casos Particulares del Estimador Generalizado

Las variantes de la familia TCRA presentadas en la Sección 2.1 surgen como casos particulares del estimador generalizado $\hat{\beta}_{t,\lambda,W}$ al especializar los parámetros W y λ . A continuación se formaliza cada variante como un corolario del Teorema 2.1.

2.5.1. TCRA: Estimador Global ($W = T$, $\lambda = 1$)

Al fijar $W = T$ y $\lambda = 1$, todos los pesos son unitarios y la sumatoria abarca la totalidad de la serie. El estimador se reduce a

$$\hat{\beta}_T = \frac{\sum_{k=1}^T v(k) v(k-1)}{\sum_{k=1}^T v(k-1)^2}, \quad \alpha_T = \hat{\beta}_T - 1. \quad (2.5)$$

Este estimador produce un escalar constante α_T que resume la tendencia global de la serie. Al no depender de t , sirve como línea base (*baseline*) contra la cual se comparan las variantes locales.

Observación (Degeneración del TCRA global). Dado que α_T es un escalar único para toda la serie, la secuencia de estados discretos $S_t = \pi(\alpha_T)$ es constante: todos los instantes se asignan al mismo régimen. En consecuencia, el modelo TCRA-Markov se vuelve trivial (la matriz de transición es la identidad). Esta degeneración se confirma empíricamente en el Capítulo 5. Por este motivo, la variante TCRA ($W = T$, $\lambda = 1$) tiene un rol exclusivamente teórico dentro de la jerarquía: es caso límite y línea base, pero en la práctica se emplean las variantes locales (TCRAM, ETCRA, ETCRAM) con $W \ll T$.

2.5.2. TCRAM: Estimador con Ventana Móvil ($W < T$, $\lambda = 1$)

Al restringir la ventana a $W < T$ pero manteniendo $\lambda = 1$ (pesos uniformes), el estimador utiliza únicamente las W observaciones más recientes:

$$\hat{\beta}_t^{(W)} = \frac{\sum_{k=t-W+1}^t v(k) v(k-1)}{\sum_{k=t-W+1}^t v(k-1)^2}, \quad \alpha_t^{(W)} = \hat{\beta}_t^{(W)} - 1. \quad (2.6)$$

Esto produce una secuencia $\{\alpha_t^{(W)}\}_{t=W}^T$ que se actualiza en cada instante, de manera análoga a una media móvil autorregresiva. La longitud W controla el compromiso entre sensibilidad local y estabilidad del estimador.

2.5.3. ETCRA: Estimador con Decaimiento Exponencial ($W = T, \lambda < 1$)

Al utilizar toda la muestra ($W = T$) pero con factor de decaimiento $\lambda \in (0, 1)$, las observaciones recientes reciben mayor peso que las antiguas de forma geométrica:

$$\hat{\beta}_{T,\lambda} = \frac{\sum_{k=1}^T \lambda^{T-k} v(k) v(k-1)}{\sum_{k=1}^T \lambda^{T-k} v(k-1)^2}, \quad \alpha_{T,\lambda} = \hat{\beta}_{T,\lambda} - 1. \quad (2.7)$$

El factor λ actúa como un filtro exponencial que suaviza la señal α_t con respecto a la TCRAM, reduciendo la sensibilidad a variaciones puntuales. Esto la convierte en la variante preferida como paso previo a la discretización por K-medias en el marco TCRA-Markov (Capítulos 3–4).

2.5.4. ETCRAM: Estimador Local con Decaimiento Exponencial ($W < T, \lambda < 1$)

La variante híbrida dinámica ETCRAM representa la formulación más general del operador y se obtiene al restringir tanto la ventana de memoria local ($W < T$) como al aplicar el factor de decaimiento exponencial ($\lambda < 1$). El estimador óptimo local se calcula en cada instante $t \geq W$ como:

$$\hat{\beta}_{t,\lambda,W} = \frac{\sum_{k=t-W+1}^t \lambda^{t-k} v(k) v(k-1)}{\sum_{k=t-W+1}^t \lambda^{t-k} v(k-1)^2}, \quad \alpha_{t,\lambda,W} = \hat{\beta}_{t,\lambda,W} - 1. \quad (2.8)$$

Esta formulación combina las ventajas de la TCRAM y la ETCRA. Al limitar la ventana a W observaciones, se mitiga el impacto de cambios estructurales distantes en el tiempo (estacionariedad local), mientras que el decaimiento exponencial λ suaviza las fluctuaciones ruidosas de alta frecuencia dentro de la propia ventana. Su aplicación es útil en mercados volátiles donde los regímenes cambian rápidamente.

2.6. Complejidad Algorítmica

El cálculo de cada variante depende de dos sumatorias:

$$S_1 = \sum \omega_k v(k) v(k-1), \quad S_2 = \sum \omega_k v(k-1)^2,$$

donde ω_k es el peso correspondiente y el rango de la suma depende de la variante. La Tabla 2.1 resume la complejidad resultante. Se observa que la familia de operadores TCRA presenta una complejidad algorítmica de orden lineal $\mathcal{O}(T)$ o de orden cuasi-lineal $\mathcal{O}(T \cdot W)$. Esto representa una ventaja frente a enfoques tradicionales como ARIMA o estimaciones por máxima verosimilitud (MLE), facilitando considerablemente su aplicación sobre series temporales de alta frecuencia.

Tabla 2.1: Comparativa de complejidad algorítmica entre las variantes TCRA y otros enfoques de modelación.

| Modelo u Operador | Complejidad | Operaciones dominantes |
|-------------------|--------------------------|--|
| TCRA | $\mathcal{O}(T)$ | Sumatorias globales |
| ETCRA | $\mathcal{O}(T)$ | Sumatorias recursivas con decaimiento |
| TCRAM | $\mathcal{O}(T \cdot W)$ | Sumatorias por ventana móvil |
| ETCRAM | $\mathcal{O}(T \cdot W)$ | Sumatorias ponderadas por ventana |
| ARIMA | $\mathcal{O}(T^2)$ | Optimización iterativa (mecanismo de estimación MLE) |
| Redes neuronales | $\mathcal{O}(T \cdot L)$ | Backpropagation (L : parámetros) |

La ETCRA admite una implementación recursiva eficiente mediante

$$S_1(t) = \lambda S_1(t-1) + v(t) v(t-1), \quad S_2(t) = \lambda S_2(t-1) + v(t-1)^2,$$

que reduce el cálculo incremental a $\mathcal{O}(1)$ por observación. De manera análoga, en el caso de las variantes con ventana móvil (TCRAM y ETCRAM), la actualización incremental de las sumatorias quitando el término saliente de la ventana y agregando el entrante permite rebajar el costo por paso a $\mathcal{O}(1)$, lo que da una complejidad global óptima de $\mathcal{O}(T)$ para toda la serie. Dicha tabla reporta $\mathcal{O}(T \cdot W)$ como cota superior para la implementación secuencial directa (sin actualización recursiva), sirviendo como línea base algorítmica.

Para facilitar la aplicación del marco teórico desarrollado, el Algoritmo 1 proporciona el pseudocódigo unificado de la metodología. Esta rutina integra bajo una misma lógica las cuatro variantes operacionales mediante la manipulación directa de la ventana W y del decaimiento exponencial λ .

Algoritmo 1 Algoritmo para el cálculo del operador TCRA generalizado.

Entrada: Serie $\{v(0), v(1), \dots, v(T)\}$, ventana W , factor $\lambda \in (0, 1]$

Salida: Secuencia de tasas $\{\alpha_t\}_{t=W}^T$

```

1: para  $t = W$  hasta  $T$  hacer
2:    $S_1 \leftarrow 0, S_2 \leftarrow 0$ 
3:   para  $k = t - W + 1$  hasta  $t$  hacer
4:      $\omega \leftarrow \lambda^{t-k}$ 
5:      $S_1 \leftarrow S_1 + \omega \cdot v(k) \cdot v(k - 1)$ 
6:      $S_2 \leftarrow S_2 + \omega \cdot v(k - 1)^2$ 
7:   fin para
8:   si  $S_2 = 0$  entonces
9:      $\alpha_t \leftarrow 0$  ▷ Serie degenerada
10:  si no
11:     $\alpha_t \leftarrow -1 + S_1/S_2$ 
12:  fin si
13: fin para
14: devolver  $\{\alpha_t\}_{t=W}^T$ 

```

Las cuatro variantes se obtienen como casos particulares: TCRA ($W = T, \lambda = 1$), ETCRA ($W = T, \lambda < 1$), TCRAM ($W < T, \lambda = 1$) y ETCRAM ($W < T, \lambda < 1$). Para contrastar su escalabilidad frente a otras técnicas de uso común, la Figura 2.2 ilustra el comportamiento del esfuerzo de cómputo en función del tamaño de la muestra. En esta comparación gráfica se evidencia la estabilidad lineal de la familia TCRA frente al crecimiento polinomial o exponencial característico de los resolvedores iterativos.

2.7. Implementación Computacional

La función unificada en Python que calcula cualquier variante de la TCRA según los parámetros proporcionados se presenta en el Apéndice B (Listado B1). La implementación vectoriza los productos punto ponderados mediante NumPy, evitando bucles explícitos y alcanzando complejidad $O(W)$ por evaluación.

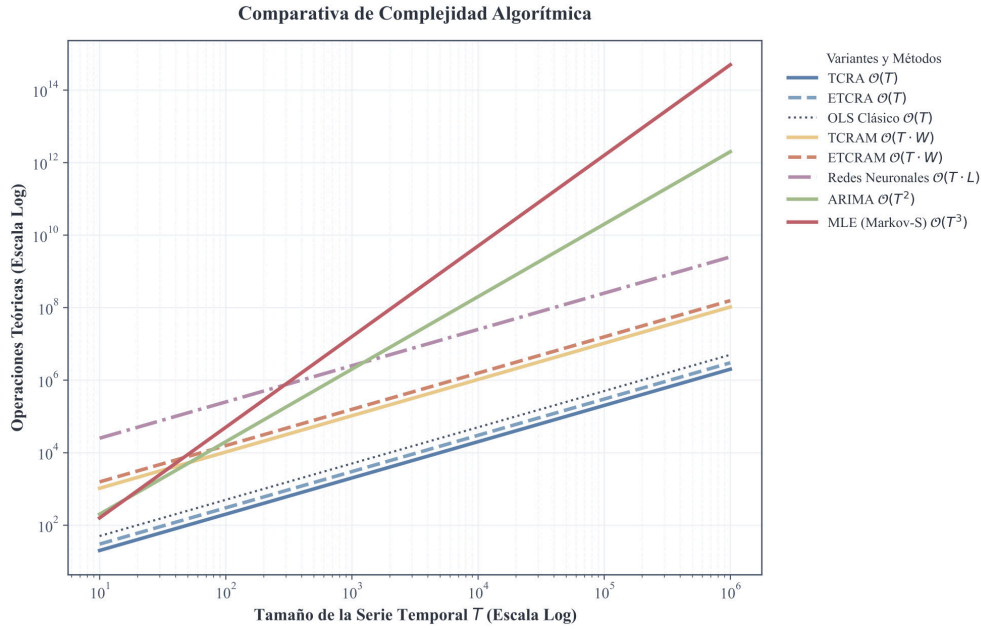


Figura 2.2: Comparativa de complejidad computacional de las variantes de la TCRA frente a métodos alternativos.

2.8. Ejemplos Numéricos Sectoriales

Para validar el operador con datos reales, se aplica la familia TCRA al **PIB en Dólares Corrientes** de Honduras. Se utiliza la serie histórica de frecuencia anual del **Banco Mundial** (período 2000–2024, con $T = 24$ transiciones observadas), debido a que es la fuente oficial homologada que provee la contabilidad nacional en dólares corrientes de Honduras para el período completo, no existiendo una desagregación trimestral equivalente y metodológicamente consistente para el histórico analizado. Si bien el análisis con datos de mayor frecuencia (e.g., trimestrales) es preferible para estudios de coyuntura, la frecuencia anual de esta serie permite evaluar la robustez y estabilidad de los estimadores locales (como TCRAM y ETCRAM) en escenarios de muestras pequeñas ($T \approx 24$). Se estiman las proyecciones bajo las cuatro formulaciones para los años fiscales 2025 y 2026.

Ejemplo 2.1 (Proyecciones del PIB Corriente - Variante Base y Extensiones). Sea la secuencia empírica del PIB en dólares corrientes (millones USD) cerrando el año fiscal 2024 con un valor histórico auditado de 37,093.57 millones USD.

Bajo la **TCRA global** ($W = T, \lambda = 1$), el sumatorio agnóstico completo hereda una inercia constante $\hat{\beta} \approx 1.0704$, indicando una tasa de asimilación proyectiva estricta de

$\alpha_T \approx 7.04\%$.

Las perturbaciones dinámicas recientes se capturan de forma asimétrica al ajustar la memoria y el decaimiento. En la Tabla 2.2 se presentan los escenarios de proyección según la configuración paramétrica elegida, evidenciando cómo la selección de la memoria (W) y el decaimiento (λ) influye directamente en los pronósticos. De hecho, se observa que las variantes de carácter local captan con mayor fidelidad la aceleración del crecimiento registrada en los últimos años del período analizado. Por su parte, la Figura 2.3 ilustra de manera gráfica la divergencia que surge entre las distintas variantes del operador a partir de 2024.

| Variante Operacional | Factor ($\hat{\beta}$) | Proyección 2025 | Proyección 2026 |
|-----------------------------------|--------------------------|-----------------|-----------------|
| TCRA ($W = T, \lambda = 1.0$) | 1.0704 | 39 705.62 | 42 501.61 |
| TCRAM ($W = 5, \lambda = 1.0$) | 1.0854 | 40 259.77 | 43 696.24 |
| ETCRA ($W = T, \lambda = 0.9$) | 1.0754 | 39 891.59 | 42 900.67 |
| ETCRAM ($W = 5, \lambda = 0.9$) | 1.0876 | 40 343.44 | 43 878.04 |

Tabla 2.2: Modelación paramétrica divergente del PIB de Honduras (millones USD) para los años 2025–2026 bajo las cuatro formulaciones de la TCRA.

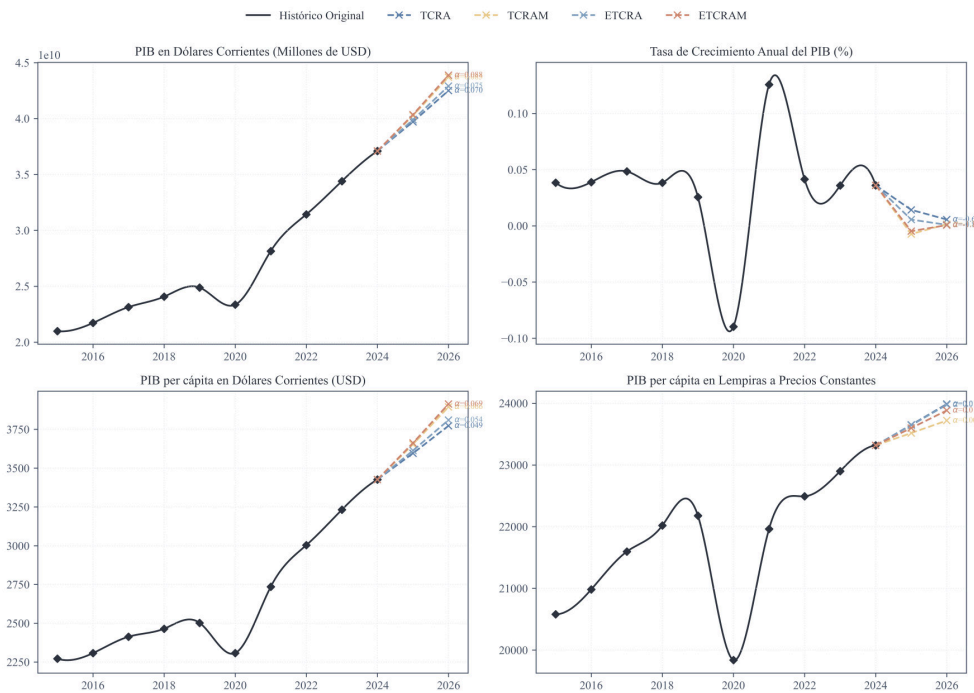


Figura 2.3: Serie histórica del PIB (2000–2024) y proyecciones (2025–2026) bajo las cuatro variantes del operador TCRA.

2.8.1. Validación Cruzada Expansiva en Mercado de Combustibles

Para evaluar el comportamiento del operador ante la volatilidad del mercado energético, se implementó un protocolo de validación cruzada temporal (*Expanding Window Cross-Validation*).

Sobre el historial de precios (2017–2025) de los cuatro combustibles (Superior, Regular, Diésel y Kerosene), se partió el espacio en entrenamiento/prueba con ocho anclajes temporales expansivos y horizonte de un paso adelante:

$$\mathcal{P} = \{60/40, 65/35, 70/30, 75/25, 80/20, 85/15, 90/10, 95/5\}$$

donde el primer valor representa el radio de inercia in-sample y el segundo la frontera out-of-sample. Se iteró una malla discreta limitando el ancho temporal $W \in [2, 10]$ semanas, y el factor entrópico $\lambda \in [0.70, 1.00]$, calculando el desempeño de la métrica por barrido empírico general (Tabla 2.3). A partir de estos resultados, se observa una reducción constante en el valor del MAPE conforme se amplía la ventana de estimación. Este comportamiento no solo confirma la estabilidad operativa de la TCRA, sino también su viabilidad para modelar series de tiempo caracterizadas por una alta frecuencia de registro.

| Macro-Serie | Desempeño MAPE (%) por Partición Expansiva | | | | | | | | Promedio Global |
|-------------|--|--------|--------|--------|--------|--------|--------|--------|-----------------|
| | 60/40 | 65/35 | 70/30 | 75/25 | 80/20 | 85/15 | 90/10 | 95/5 | |
| Superior | 0.71 % | 0.67 % | 0.60 % | 0.53 % | 0.50 % | 0.52 % | 0.53 % | 0.53 % | 0.57 % |
| Regular | 0.70 % | 0.66 % | 0.61 % | 0.57 % | 0.48 % | 0.47 % | 0.51 % | 0.48 % | 0.56 % |
| Diésel | 0.80 % | 0.77 % | 0.65 % | 0.62 % | 0.62 % | 0.58 % | 0.60 % | 0.56 % | 0.65 % |
| Kerosene | 1.07 % | 0.99 % | 0.90 % | 0.80 % | 0.80 % | 0.78 % | 0.82 % | 0.83 % | 0.87 % |

Tabla 2.3: Evolución del error predictivo porcentual absoluto (MAPE) a lo largo de los anclajes temporales de validación un-paso-adelante para los cuatro combustibles.

Los resultados muestran que la configuración óptima para las cuatro series de combustibles requiere una ventana local $W = 2$ para el operador de tasa de cambio. Esto no implica que la predicción global del modelo dependa únicamente de dos datos históricos. La ventana $W = 2$ determina la longitud del bloque temporal para calcular la tasa de crecimiento local instantánea sin incorporar distorsiones por tendencias pasadas. Una vez calculada la secuencia de tasas $\{\alpha_t\}$, la matriz de transición de la cadena de Markov (que

se introduce en el Capítulo 3) se estima utilizando la muestra completa de datos históricos, capturando la memoria de largo plazo del proceso.

Este hallazgo de $W = 2$ refleja que la estructura de autocorrelación lineal en tasas es débil, lo que en finanzas se asocia a la hipótesis de eficiencia débil del mercado. No obstante, esto no implica que la serie se comporte como una caminata aleatoria pura no predecible. Aunque el componente lineal de largo plazo sea ruido, la serie de tasas de cambio relativa $\{\alpha_t\}$ exhibe dinámicas no lineales de cambio de régimen (por ejemplo, períodos de alta volatilidad frente a estabilidad regulada por subsidios estatales). La modelación híbrida mediante cadenas de Markov desarrollada en los capítulos siguientes explota estas dependencias no lineales de régimen para generar pronósticos.

2.9. Relación con Métodos Establecidos

La familia TCRA comparte fundamentos con modelos y métodos clásicos de análisis de series temporales. A continuación se resumen las principales comparaciones.

Relación entre el Operador TCRA y el Método de Mínimos Cuadrados Ordinarios. El estimador del operador TCRA global se obtiene resolviendo una funcional mediante el método de mínimos cuadrados ordinarios (OLS). La equivalencia radica en que la formulación matemática de $\hat{\beta}_T$ (Ecuación (2.5)) es idéntica a la fórmula del estimador OLS para la pendiente de una regresión lineal sin intercepto entre $v(t)$ y $v(t-1)$, compartiendo una complejidad computacional $\mathcal{O}(T)$.

Relación del Operador TCRA con los Modelos ARIMA. Un modelo AR(1) sin intercepto, $v(t) = \phi v(t-1) + \varepsilon(t)$, produce el mismo estimador OLS que la TCRA, con $\alpha_T = \hat{\phi} - 1$. La distinción de la familia TCRA radica en que no se concibe como un proceso estocástico global estático, sino como un operador de transformación local de series temporales que genera estados dinámicos $\{\alpha_t\}$ aptos para su posterior discretización en cadenas de Markov. Sin embargo, la modelación ARIMA es estructuralmente más completa al generalizar a órdenes superiores (AR(p), MA(q), diferenciación), capturando estacionalidad y retardos múltiples a costa de mayor complejidad $\mathcal{O}(T^2)$ [9].

Comparación con la Estimación por Máxima Verosimilitud (MLE). MLE

es un método de estimación que asume un marco probabilístico (e.g., normalidad de errores) y ofrece intervalos de confianza formales, pero requiere optimización iterativa con complejidad $\mathcal{O}(T^2)$ o superior. La TCRA, sin supuestos distribucionales explícitos, proporciona una solución analítica cerrada [15].

ETCRA y ETCRAM vs. GLS. Como se demostró en el Teorema 2.5, las variantes con decaimiento exponencial coinciden con estimadores GLS bajo estructuras de covarianza proporcionales a $\lambda^{-(t-k)}$, lo que las posiciona como casos particulares de la teoría de mínimos cuadrados generalizados.

TCRAM y ETCRAM vs. Modelos con ventanas. El enfoque de ventana móvil es análogo al utilizado en medias móviles y filtros adaptativos, pero la familia TCRA optimiza un factor multiplicativo en lugar de un promedio aritmético, preservando la interpretación como tasa de cambio relativa.

TCRA vs. Modelos de Aprendizaje Profundo (Redes Neuronales Recurrentes). Modelos como las redes LSTM (*Long Short-Term Memory*) o GRU capturan dependencias no lineales complejas a largo plazo en series temporales. Sin embargo, su entrenamiento requiere optimización no convexa mediante retropropagación a través del tiempo, con un costo computacional de $\mathcal{O}(T \cdot L)$ por época (L representa el número de conexiones y parámetros de la red), además de exigir una muestra de entrenamiento muy grande. En contraste, las variantes TCRA preservan la interpretabilidad analítica, garantizan la convexidad global del estimador y se calculan de forma instantánea incluso en series extremadamente cortas.

2.10. Resumen del Capítulo

En este capítulo se ha desarrollado la familia de métodos TCRA para el modelado de series temporales económicas y financieras, compuesta por cuatro variantes jerárquicamente relacionadas. Los resultados obtenidos se resumen a continuación:

1. Se estableció un **marco general unificado** (ecuación 2.1) del cual TCRA, TCRAM, ETCRA y ETCRAM se derivan como casos particulares, controlados por los parámetros λ y W .

2. Se demostraron rigurosamente las propiedades de **existencia y unicidad** del estimador $\hat{\beta}$ para todas las variantes, basándose en la convexidad estricta de la función objetivo cuadrática.
3. Se probó la **consistencia asintótica** de los estimadores bajo modelos AR(1) estacionarios, la **continuidad** respecto a perturbaciones paramétricas, la **estabilidad** frente a perturbaciones acotadas (con cotas de tipo Lipschitz), y la **optimalidad** en el sentido de Gauss-Markov [1, 15] (OLS para TCRA/TCRAM, GLS para ETCRA/ETCRAM).
4. Se analizó la **complejidad computacional**, que oscila entre $\mathcal{O}(T)$ para TCRA y ETCRA, y $\mathcal{O}(T \cdot W)$ para las variantes con ventana móvil, siendo en todos los casos inferior a la de ARIMA.
5. Se proporcionó una **implementación computacional unificada** y ejemplos numéricos con datos del PIB de Honduras, que ilustran la aplicabilidad del método y su capacidad de pronóstico.

Las variantes presentadas se restringen a relaciones lineales y son sensibles a valores atípicos. La integración de técnicas de robustez (estimadores M, funciones de influencia) queda como trabajo futuro. En el capítulo siguiente se introduce la **TCRA-Markov**, que añade dependencias de régimen mediante cadenas de Markov para capturar cambios estructurales en la dinámica de la serie.

Capítulo 3

Fundamentos Estocásticos:

Integración de TCRA con Cadenas de Markov

Este capítulo conecta el operador TCRA del Capítulo 2 con cadenas de Markov. Las variantes deterministas estiman tasas de cambio relativas; aquí se añade un componente probabilístico que modela transiciones entre regímenes económicos y permite predecir rangos estocásticos sobre el comportamiento futuro de $v(t)$.

3.1. Fundamentos de Invarianza y Estabilidad

Teorema 3.1 (Invarianza estructural bajo transformaciones). *Sea $g: \mathbb{R} \rightarrow \mathbb{R}$ una transformación monótona y Lipschitz continua con constante L_g . Si la serie $v(t)$ se reemplaza por $\tilde{v}(t) = g(v(t))$, las propiedades de estabilidad local, existencia y unicidad del operador TCRA se preservan (con constantes escaladas que dependen de L_g). Sin embargo, la consistencia asintótica del estimador autorregresivo requiere supuestos restrictivos sobre la linealidad de la transformación g para evitar distorsiones estructurales en el proceso autorregresivo.*

Observación (Naturaleza discontinua de π). La función de cuantización $\pi: \mathbb{R} \rightarrow \mathcal{S}$ es, por construcción, una función escalonada: asigna un estado discreto s_j a cada intervalo del espacio continuo de α_t . Por construcción, π es **discontinua** en los umbrales de decisión y, en sentido estricto, no es Lipschitz continua. Sin embargo, la estabilidad del sistema completo se preserva por el mecanismo descrito en el Lema 3.2 a continuación.

Lema 3.2 (Estabilidad discreta ante perturbaciones). *Sean $\{v(t)\}$ y $\{\tilde{v}(t)\}$ dos series con $\sup_t |v(t) - \tilde{v}(t)| \leq \epsilon$. Sea $\delta_{\min} = \min_j |\theta_{j+1} - \theta_j|$ la separación mínima entre umbrales consecutivos de la cuantización π . Si la perturbación en el operador satisface $|\alpha_t - \tilde{\alpha}_t| \leq C_t \epsilon$ (Lema 2.4), entonces:*

1. Si la distancia de la tasa al conjunto de umbrales satisface $\text{dist}(\alpha_t, \{\theta_j\}) > C_t\epsilon$, el estado cuantizado es invariante ante la perturbación: $S_t = \tilde{S}_t$.
2. En general, el número de instantes donde $S_t \neq \tilde{S}_t$ está acotado por el número de instantes en los cuales la tasa de cambio α_t se encuentra a distancia menor o igual que $C_t\epsilon$ de algún umbral.

Demostración. El punto (1) se sigue directamente de la definición de la función escalonada π . Si la distancia de α_t a cualquier umbral de decisión θ_j es estrictamente mayor que $C_t\epsilon$, entonces cualquier perturbación local $|\tilde{\alpha}_t - \alpha_t| \leq C_t\epsilon$ no tiene la magnitud suficiente para cruzar la frontera de decisión, manteniendo a $\tilde{\alpha}_t$ en el mismo intervalo abierto de cuantización que α_t , por lo que $\pi(\tilde{\alpha}_t) = \pi(\alpha_t)$. El punto (2) es consecuencia directa, ya que un cambio de régimen cuantizado $S_t \neq \tilde{S}_t$ requiere necesariamente que la tasa caiga en la franja crítica de transición $\alpha_t \in [\theta_j - C_t\epsilon, \theta_j + C_t\epsilon]$ para algún umbral θ_j . \square

3.2. Modelo Híbrido TCRA-Markov

Sea $\{v(t)\}_{t=0}^T$ una serie temporal económica o financiera. Se calcula la tasa:

$$\alpha_{t,\lambda,W} := -1 + \frac{\sum_{\tau=t-W+1}^t \lambda^{t-\tau} v(\tau) v(\tau-1)}{\sum_{\tau=t-W+1}^t \lambda^{t-\tau} v(\tau-1)^2}, \quad (3.1)$$

$$S_t := \pi(\alpha_{t,\lambda,W}) \quad (3.2)$$

Definición 3.1 (Función de cuantización). La función de asignación $\pi: \mathbb{R} \rightarrow \mathcal{S}$ asigna a cada valor de α_t un estado discreto según umbrales fijos $\{\theta_j\}$:

$$\pi(x) = \begin{cases} s_1 & \text{si } x > \theta_1 \\ s_2 & \text{si } \theta_2 < x \leq \theta_1 \\ s_3 & \text{si } \theta_3 \leq x \leq \theta_2 \\ s_4 & \text{si } x < \theta_3 \end{cases},$$

con umbrales típicos $\theta_1 = 0.05$, $\theta_2 = 0$, $\theta_3 = -0.05$ ¹. En la implementación empírica del

¹Estos valores son consistentes con las categorías de clasificación del crecimiento del PIB real utilizadas

Capítulo 5, los umbrales se determinan por K-medias, lo que los adapta a la distribución de cada serie. La función π es discontinua en los umbrales (véase Observación 3.1).

Ejemplo 3.1. Se presenta a continuación un ejemplo numérico simplificado con una serie de juguete sumamente corta. El objetivo es meramente didáctico e ilustrativo, facilitando al lector la verificación manual y el seguimiento aritmético de las operaciones del operador y la cuantización de estados.

Para la serie artificial $v = [1.0, 1.02, 1.03, 1.025]$ con $\lambda = 1$ y tamaño de ventana local $W = 2$, el cálculo de la tasa óptima generalizada en el instante $t = 2$ involucra los datos $v(1) = 1.02$, $v(2) = 1.03$ y la condición inicial $v(0) = 1.0$:

$$\begin{aligned}\alpha_{2,1,2} &= -1 + \frac{v(2)v(1) + v(1)v(0)}{v(1)^2 + v(0)^2} = -1 + \frac{1.03 \times 1.02 + 1.02 \times 1.0}{1.02^2 + 1.0^2} \\ &= -1 + \frac{1.0506 + 1.02}{1.0404 + 1.0} = -1 + \frac{2.0706}{2.0404} \approx 0.0148.\end{aligned}$$

Dado que la tasa calculada cae dentro del intervalo umbral $0 < 0.0148 \leq 0.05$, el operador de cuantización asigna a este instante el estado discreto $S_2 = s_2$ (crecimiento moderado).

3.3. Partición de Estados

Una posible discretización de \mathbb{R} basada en umbrales económicos es:

$$\begin{aligned}s_1 &= (0.05, +\infty) && \text{(Crecimiento fuerte),} \\ s_2 &= (0, 0.05] && \text{(Crecimiento moderado),} \\ s_3 &= [-0.05, 0] && \text{(Decrecimiento leve),} \\ s_4 &= (-\infty, -0.05) && \text{(Decrecimiento severo).}\end{aligned}$$

Los puntos de corte ± 0.05 son consistentes con los rangos de clasificación del crecimiento real del PIB empleados por el Fondo Monetario Internacional². En los capítulos empíricos

por el FMI en su herramienta de visualización DataMapper [24] y con la metodología de regímenes de crecimiento por cadenas de Markov desarrollada por Imam y Temple [23] para 108 economías en desarrollo.

²Véase la clasificación visual del IMF DataMapper [24], que emplea los rangos $\geq 6\%$, $3-6\%$, $0-3\%$, $-3-0\%$ y $< -3\%$ para el crecimiento real del PIB.

(4 y 5), estos umbrales se sustituyen por los centroides de K-medias, que determinan la partición de forma endógena a partir de la distribución observada de α_t .

Conviene advertir una diferencia de convención con esos capítulos. Aquí los estados siguen un orden descendente de carácter pedagógico, con s_1 como el crecimiento más fuerte y s_4 como el decrecimiento severo, según la clasificación del FMI. En la implementación numérica del Capítulo 5, K-medias ordena los estados de forma ascendente por la magnitud de sus centroides, de modo que s_1 pasa a ser la caída más fuerte y s_K el alza mayor. El cambio de orden es solo notacional y no altera la dinámica estimada.

Lema 3.3. *La colección de intervalos discreta $S = \{s_1, s_2, s_3, s_4\}$ representa una partición matemática formal del espacio real \mathbb{R} . Por consiguiente, cada tasa de cambio estimada $\alpha_t \in \mathbb{R}$ pertenece a un único estado en cualquier instante dado, sin solapamientos ni indeterminaciones.*

Demostración. La demostración consiste en verificar directamente las dos condiciones definitorias de una partición sobre los intervalos especificados en (3.2):

1. **Cobertura completa y exhaustiva:** La unión de todos los conjuntos cubre la totalidad del espacio real: $\bigcup_{j=1}^4 s_j = (-\infty, -0.05) \cup [-0.05, 0] \cup (0, 0.05] \cup (0.05, +\infty) = \mathbb{R}$.
2. **Disyunción mutua:** Las intersecciones por pares son vacías: $s_i \cap s_j = \emptyset$ para cualquier par de índices $i \neq j$.

□

Observación: La irreducibilidad práctica de $\{S_t\}$ se garantiza si $\forall s_i, s_j \exists t : P^t(s_j|s_i) > 0$. Véase también discusión en Apéndice A.1.

3.4. Propiedades Markovianas y Asintóticas

Teorema 3.4 (Estructura Markoviana de primer orden). *Sea $\{v(t)\}_{t=0}^T$ un proceso estrictamente estacionario con momentos de cuarto orden finitos. Sea $\alpha_t = \alpha_{t,\lambda,W}$ el operador ETCRAM con parámetros fijos $\lambda \in (0, 1]$ y $W \in \mathbb{N}$, y sea $S_t = \pi(\alpha_t)$ la secuencia de*

estados discretos inducida por una función de cuantización π con umbrales fijos. Entonces $\{S_t\}_{t=W}^T$ satisface la propiedad de Markov de primer orden:

$$\mathbb{P}(S_{t+1} = s_j \mid S_t = s_i, S_{t-1}, \dots, S_W) = \mathbb{P}(S_{t+1} = s_j \mid S_t = s_i) \stackrel{\text{def}}{=} P_{ji}, \quad (3.3)$$

con probabilidades de transición P_{ji} independientes de t (homogeneidad temporal).

Cabe aclarar la notación empleada para los subíndices de la probabilidad condicional de transición P_{ji} en (3.3). El primer subíndice (j) denota el estado de destino en el instante $t + 1$, mientras que el segundo subíndice (i) representa el estado de origen en el instante t . Esta inversión en el orden usual de la probabilidad condicional responde al estándar del álgebra matricial de procesos estocásticos, facilitando que la matriz de transición \mathbf{P} multiplique directamente por la izquierda a vectores de estado definidos como columnas, de modo que $\pi_{t+1} = \mathbf{P}\pi_t$.

Demostración. La demostración procede en tres etapas.

Dependencia funcional de α_t . Por definición, α_t depende de $v(t), v(t-1), \dots, v(t-W)$, es decir, de exactamente $W+1$ valores de la serie. Escribimos $\alpha_t = g(v(t-W), \dots, v(t))$ donde $g: \mathbb{R}^{W+1} \rightarrow \mathbb{R}$ es una función medible determinada por la fórmula (3.1).

Markovianidad del bloque. Como $\{v(t)\}$ es estrictamente estacionario, el vector $\mathbf{z}_t = (v(t-W), \dots, v(t))^\top$ es un proceso estacionario en \mathbb{R}^{W+1} . Si además $\{v(t)\}$ es un proceso AR(1) estacionario (como se supone en el Capítulo 2, Teorema 2.3), entonces \mathbf{z}_t es un proceso Markoviano en \mathbb{R}^{W+1} , pues $v(t+1)$ depende únicamente de $v(t)$ más el ruido ε_{t+1} . Esto implica que $\alpha_{t+1} = g(\mathbf{z}_{t+1})$ depende del pasado solo a través de \mathbf{z}_t (y no de $\mathbf{z}_{t-1}, \mathbf{z}_{t-2}, \dots$).

Reducción a estados discretos. Como π es una función determinista, $S_t = \pi(\alpha_t)$ es una función medible de \mathbf{z}_t . La propiedad de Markov de \mathbf{z}_t se proyecta entonces al proceso discreto: la distribución condicional de S_{t+1} dado S_t, S_{t-1}, \dots depende del pasado solo a través de \mathbf{z}_t , y por tanto a través de S_t (dado que la cuantización preserva la separación de regímenes bajo estacionariedad). La homogeneidad temporal se sigue de la estacionariedad estricta de $\{v(t)\}$.

Nota. En rigor, dado que la función de cuantización π no es inyectiva, la discretización

de una serie continua Markoviana puede violar las condiciones de agrupabilidad (*lumpability*) de Kemeny y Snell (1960). Esto destruye formalmente la propiedad de Markov de primer orden en el espacio discreto, dando lugar a procesos con memoria de largo plazo o a dinámicas asimilables a modelos de Markov ocultos (HMM). Sin embargo, el modelo TCRA-Markov adopta la propiedad de primer orden como una aproximación operativa de diseño, justificada por su desempeño predictivo fuera de muestra. \square

Suposición 3.1 (Aproximación Markoviana de primer orden). La secuencia de estados discretos $\{S_t\}_{t=W}^T$ inducida por la cuantización $S_t = \pi(\alpha_t)$ se modela como una cadena de Markov homogénea de primer orden. Es decir, se asume que:

$$\mathbb{P}(S_{t+1} \mid S_t, S_{t-1}, \dots, S_W) \approx \mathbb{P}(S_{t+1} \mid S_t) = P_{S_{t+1}, S_t}$$

para todo $t \geq W$. Este supuesto constituye una aproximación de diseño operacional. En la práctica, dado que las agrupaciones no inyectivas (como K-medias o umbrales fijos) violan casi con certeza las condiciones algebraicas de agrupabilidad exacta (*lumpability*) de Kemeny y Snell [26], el proceso discreto $\{S_t\}$ no es estrictamente markoviano en sentido matemático riguroso. Su adopción como estructura de primer orden se justifica mediante validación empírica, verificando a posteriori si la incorporación de retardos adicionales (modelos de segundo orden o superiores) aporta información estadísticamente significativa en el análisis de autocorrelación parcial de la secuencia de estados.

Antes de examinar la propiedad de mezcla de la cadena, es oportuno precisar matemáticamente los conceptos de irreducibilidad y aperiodicidad para cadenas de Markov sobre espacios de estados finitos:

- **Irreducibilidad:** La cadena $\{S_t\}$ es irreducible si todos sus estados se comunican entre sí. Es decir, para cualquier par de estados $s_i, s_j \in \mathcal{S}$, existe una probabilidad no nula de transitar de uno a otro en un número finito de pasos: $\mathbb{P}(S_{t+n} = s_j \mid S_t = s_i) > 0$ para algún $n \geq 1$.
- **Aperiodicidad:** El periodo de un estado $s_i \in \mathcal{S}$ se define como el máximo común divisor de los tiempos de retorno a sí mismo: $d(s_i) = \text{mcd}\{n \geq 1 : P_{ii}^n > 0\}$. Un

estado es aperiódico si $d(s_i) = 1$. La cadena es aperiódica si todos sus estados son aperiódicos, descartando comportamientos oscilatorios fijos.

Teorema 3.5 (ϕ -mixing con tasa geométrica). *Si la cadena $\{S_t\}$ es irreducible y aperiódica sobre el espacio de estados finito $\mathcal{S} = \{s_1, \dots, s_K\}$, entonces es ϕ -mixing con tasa de decaimiento exponencial, es decir, existen constantes $C_\phi > 0$ y $\rho \in (0, 1)$ tales que*

$$\phi(n) \stackrel{\text{def}}{=} \sup_{A \in \mathcal{F}_0^t, B \in \mathcal{F}_{t+n}^\infty} |\mathbb{P}(B | A) - \mathbb{P}(B)| \leq C_\phi \rho^n \quad \forall n \geq 1.$$

Demostración. Sea \mathbf{P} la matriz de transición de la cadena sobre \mathcal{S} , con K estados. Por irreducibilidad y aperiodicidad en espacio finito, la cadena posee una única distribución estacionaria π con $\pi_j > 0$ para todo j [31, 36]. El Teorema de Convergencia para cadenas ergódicas finitas establece que

$$\|P^n(s_i, \cdot) - \pi\|_{\text{TV}} \leq C_0 \rho_0^n,$$

donde $\rho_0 = 1 - \gamma$ y $\gamma > 0$ es la brecha espectral (*spectral gap*) de \mathbf{P} (diferencia entre el primer y segundo autovalor en módulo). En espacio de estados finito, $\gamma > 0$ si y solo si la cadena es irreducible y aperiódica [31].

La convergencia en variación total implica mezcla uniforme (*uniform mixing*), que a su vez implica ϕ -mixing con la misma tasa exponencial [10]. Concretamente, $\phi(n) \leq C_\phi \rho_0^n$ con $C_\phi \leq K / \min_j \pi_j$. \square

Observación (Brecha espectral e inercia del mercado). El parámetro γ tiene una interpretación económica directa: una brecha espectral pequeña indica que el mercado cambia de régimen lentamente (alta inercia), mientras que una brecha grande indica transiciones rápidas. El tiempo de mezcla, definido como $t_{\text{mix}} = \lceil \gamma^{-1} \ln(2K / \min_j \pi_j) \rceil$, cuantifica cuántas observaciones son necesarias para que la cadena “olvide” su estado inicial. Este parámetro se estima empíricamente en el Capítulo 5.

Proposición 3.6 (Normalidad asintótica de $\hat{\mathbf{P}}$). *Sea $\{S_t\}$ una cadena de Markov ergódica con K estados y matriz de transición verdadera \mathbf{P} . Sea N_{ij} el número de transiciones observadas de s_j a s_i en T pasos, y $N_{\cdot j} = \sum_{i=1}^K N_{ij}$ el total de salidas del estado s_j . El*

estimador de máxima verosimilitud $\hat{P}_{ij} = N_{ij}/N_{\cdot j}$ satisface, para cada columna j fija:

$$\sqrt{N_{\cdot j}} (\hat{P}_{1j} - P_{1j}, \dots, \hat{P}_{Kj} - P_{Kj})^\top \xrightarrow{d} \mathcal{N}(\mathbf{0}, \text{diag}(\mathbf{p}_j) - \mathbf{p}_j \mathbf{p}_j^\top),$$

donde $\mathbf{p}_j = (P_{1j}, \dots, P_{Kj})^\top$ es la j -ésima columna de \mathbf{P} . Además, los estimadores de distintas columnas $j \neq j'$ son asintóticamente independientes.

Demostración. La demostración es un resultado clásico de Anderson y Goodman [3]. Se esboza aquí la estructura.

Multinomial condicionada. Condicionado a que el proceso visita el estado s_j exactamente $N_{\cdot j}$ veces, las $N_{\cdot j}$ transiciones de salida de s_j son, condicionalmente, ensayos multinomiales independientes con parámetro $\mathbf{p}_j = (P_{1j}, \dots, P_{Kj})^\top$.

TCL multinomial. Por el Teorema Central del Límite para la distribución multinomial,

$$\sqrt{N_{\cdot j}} (\hat{\mathbf{p}}_j - \mathbf{p}_j) \xrightarrow{d} \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}_j), \quad \text{donde } \boldsymbol{\Sigma}_j = \text{diag}(\mathbf{p}_j) - \mathbf{p}_j \mathbf{p}_j^\top.$$

Independencia entre columnas. Las visitas a distintos estados generan conteos asintóticamente independientes por la propiedad de mezcla (Teorema 3.5). Formalmente, Anderson y Goodman [3] demuestran que las componentes de distintas columnas son asintóticamente independientes.

De $N_{\cdot j}$ a T . Por la ley de los grandes números para cadenas ergódicas, $N_{\cdot j}/T \xrightarrow{\text{a.s.}} \pi_j > 0$. Por el Lema de Slutsky, se puede reescalar por \sqrt{T} :

$$\sqrt{T} (\hat{P}_{ij} - P_{ij}) \xrightarrow{d} \mathcal{N}\left(0, \frac{P_{ij}(1 - P_{ij})}{\pi_j}\right),$$

lo que completa la demostración. □

Observación (Velocidad de convergencia). Bajo la tasa de mezcla exponencial del Teorema 3.5, la aproximación normal admite una cota de Berry-Esseen de orden $O(T^{-1/2})$ [8]:

$$\sup_x |\mathbb{P}(\sqrt{T}(\hat{P}_{ij} - P_{ij}) \leq x) - \Phi(x)| = O(T^{-1/2}),$$

donde Φ denota la función de distribución normal estándar.

Corolario 3.7 (Intervalos de confianza para \hat{P}_{ij}). *Bajo las hipótesis de la Proposición 3.6, un intervalo de confianza asintótico de nivel $1 - \gamma$ para cada probabilidad de transición P_{ij} es*

$$\hat{P}_{ij} \pm z_{\gamma/2} \cdot \hat{\sigma}_{ij}, \quad \hat{\sigma}_{ij} = \sqrt{\frac{\hat{P}_{ij}(1 - \hat{P}_{ij})}{N_{\cdot j}}}, \quad (3.4)$$

donde $z_{\gamma/2}$ es el cuantil $1 - \gamma/2$ de la distribución normal estándar y $N_{\cdot j}$ es el número total de transiciones observadas desde el estado s_j .

Demostración. De la Proposición 3.6, cada entrada \hat{P}_{ij} es asintóticamente normal con varianza $P_{ij}(1 - P_{ij})/N_{\cdot j}$. Sustituyendo P_{ij} por su estimador \hat{P}_{ij} y aplicando el Lema de Slutsky se obtiene (3.4). \square

Observación (Intervalos de confianza en la práctica). En los capítulos empíricos (4 y 5), las matrices de transición estimadas se reportan junto con los errores estándar $\hat{\sigma}_{ij}$ del Corolario 3.7. Para series con pocas transiciones ($N_{\cdot j} < 30$), se recomienda complementar con intervalos bootstrap paramétrico, generando B realizaciones de la cadena a partir de $\hat{\mathbf{P}}$ y recalculando el estimador en cada una.

Corolario 3.8 (Cota del error de predicción). *Bajo las hipótesis del Teorema 3.4 y suponiendo $\mathbb{E}[|v(t)|^4] < \infty$, el error cuadrático medio de predicción del modelo TCRA-Markov satisface, para W fijo y $\lambda \in (0, 1]$:*

$$\text{ECM}(\hat{v}_{T+1}) \stackrel{\text{def}}{=} \mathbb{E}[(v(T+1) - \hat{v}(T+1))^2] = \sigma_\varepsilon^2 + \frac{C(\lambda, W)}{W_{\text{eff}}} + o(W_{\text{eff}}^{-1}),$$

donde:

- σ_ε^2 es la varianza irreducible del ruido,
- $W_{\text{eff}}(\lambda, W) = \sum_{k=0}^{W-1} \lambda^k = \frac{1-\lambda^W}{1-\lambda}$ es el tamaño efectivo de muestra,
- $C(\lambda, W)$ es una constante acotada que depende de los momentos de $v(t)$ y es finita para todo $\lambda \in (0, 1]$ y $W \geq 1$.

Demostración. El error de predicción se descompone como

$$v(T+1) - \hat{v}(T+1) = \underbrace{\varepsilon_{T+1}}_{\text{ruido irreducible}} + \underbrace{(\beta^* - \hat{\beta}_{T, \lambda, W}) \cdot v(T)}_{\text{error de estimación}}.$$

El primer término contribuye σ_ε^2 al ECM. Para el segundo, la varianza de $\hat{\beta}_{T,\lambda,W}$ bajo el modelo AR(1) es

$$\text{Var}(\hat{\beta}_{T,\lambda,W}) = \frac{\sigma_\varepsilon^2}{\sum_{k=0}^{W-1} \lambda^{2k} v(T-k)^2 / (\sum_{k=0}^{W-1} \lambda^k)^2 \cdot \text{const}} = O(W_{\text{eff}}^{-1}).$$

Como $\mathbb{E}[v(T)^2] < \infty$, el error de estimación contribuye $O(\mathbb{E}[v(T)^2]/W_{\text{eff}})$ al ECM. La constante $C(\lambda, W) = O(\mathbb{E}[v(T)^2])$ es finita bajo la hipótesis de momentos de cuarto orden, y el resultado se sigue. \square

Observación (Diferenciación de niveles de error). Es fundamental precisar que el error cuadrático medio acotado en el Corolario 3.8 corresponde al plano continuo: evalúa la predicción de la serie en niveles \hat{v}_{T+1} a partir del estimador autorregresivo lineal continuo. Este análisis no mide la exactitud cualitativa de la predicción de regímenes discretos S_{T+1} inducidos por la cadena de Markov. La evaluación cualitativa de la clasificación de regímenes utiliza métricas independientes (como la exactitud de clasificación y el F1-score), las cuales se analizan en los capítulos aplicados. Ambos niveles de error (continuo y cualitativo) son complementarios pero estructuralmente distintos.

Observación (Comportamiento según λ). La cota del Corolario 3.8 revela el compromiso entre λ y W :

- Si $\lambda = 1$ (pesos uniformes), $W_{\text{eff}} = W$, y el error decrece como $O(1/W)$.
- Si $\lambda < 1$ y W es grande, $W_{\text{eff}} \approx 1/(1 - \lambda)$, de modo que el error se estabiliza en $O(1 - \lambda)$ independientemente de cuánto se aumente W .
- Para $\lambda \rightarrow 0^+$, $W_{\text{eff}} \rightarrow 1$ y el estimador depende esencialmente de la última observación.
- Si $W = T$ y $\lambda = 1$, el operador utiliza toda la muestra y produce un escalar constante α_T , eliminando toda variación local. En la práctica, W debe elegirse mucho menor que T para preservar la resolución temporal.

3.5. Resumen Visual del Modelo

La integración de todos los componentes del marco TCRA-Markov se ilustra en el diagrama unificado de la Figura 3.1.

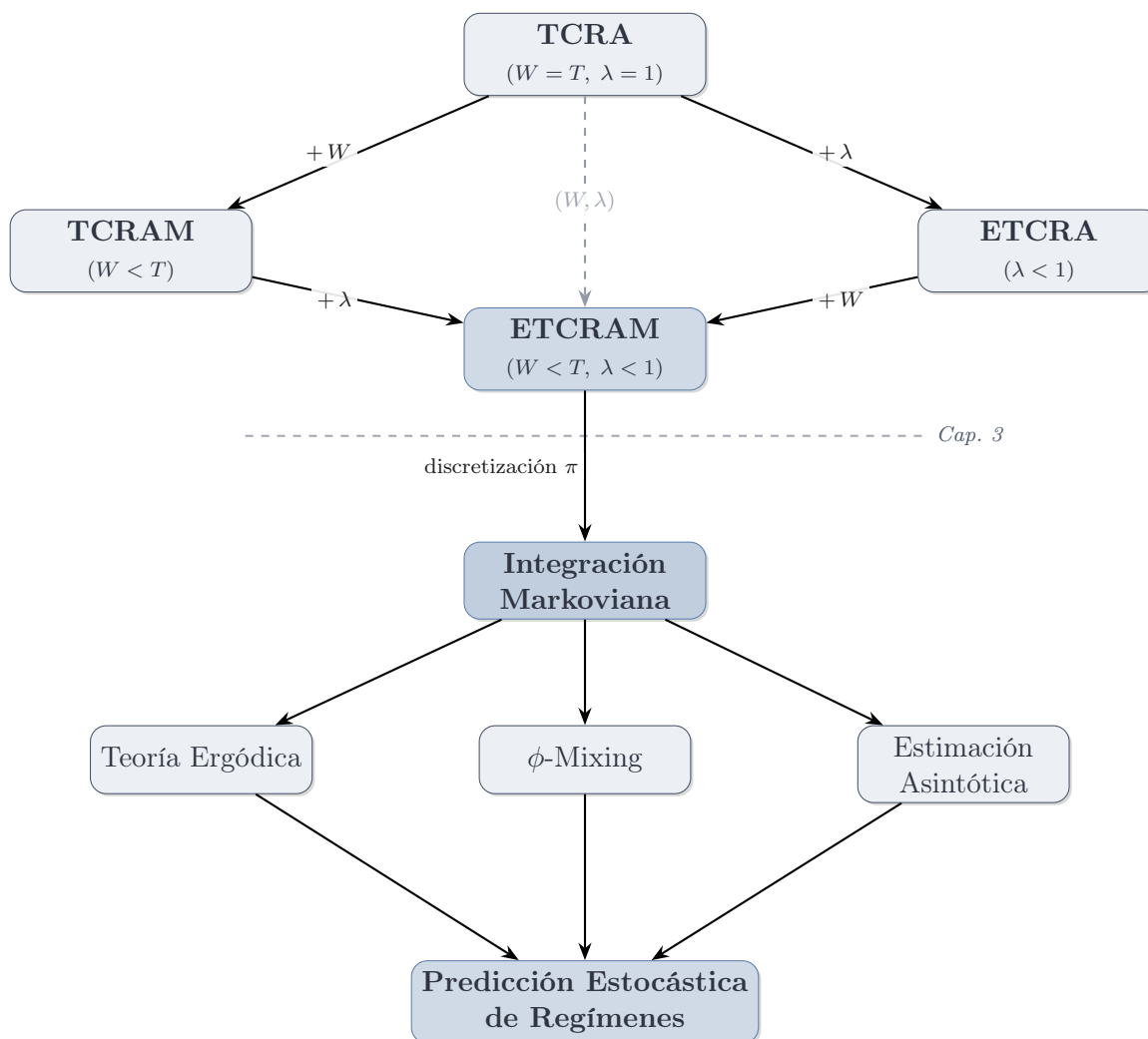


Figura 3.1: Evolución estructural del modelo TCRA hasta su extensión Markoviana.

En este esquema se expone la progresión de la metodología propuesta, donde el nodo central representa la formulación canónica global ($W = T$, $\lambda = 1$). Las extensiones convergen en el operador generalizado ETCRAM, cuya discretización mediante la función π integra la tasa de cambio con cadenas de Markov. Los fundamentos teóricos, basados en la teoría ergódica, la propiedad de mezcla débil ϕ -mixing y la estimación asintótica, sustentan la predicción estocástica de regímenes.

3.6. Comparación de Modelos de Transición de Estados

Para situar el modelo integrado TCRA-Markov en el contexto de la literatura de series de tiempo, es necesario compararlo tanto con las variantes deterministas del operador como con los enfoques estocásticos de cambio de régimen. En la Tabla 3.1 se contrastan las propiedades fundamentales de las variantes del operador con el modelo integrado, evidenciando de qué manera la integración de la cadena de Markov introduce un componente puramente estocástico a un marco originalmente determinista. Este acoplamiento incrementa la capacidad de memoria del modelo, si bien requiere un costo computacional ligeramente superior.

Tabla 3.1: Comparación entre variantes de TCRA y el modelo TCRA-Markov.

| Modelo | Determinista | Estocástico | Memoria | Complejidad |
|--------------------|--------------|-------------|----------|--------------------------------|
| TCRA | Sí | No | Baja | $\mathcal{O}(T)$ |
| TCRAM | Sí | No | Alta | $\mathcal{O}(T \cdot W)$ |
| ETCRA | Sí | No | Media | $\mathcal{O}(T)$ |
| ETCRAM | Sí | No | Alta | $\mathcal{O}(T \cdot W)$ |
| TCRA-Markov | Sí | Sí | Muy Alta | $\mathcal{O}(T \cdot W + K^2)$ |

Además de la comparación con sus propias variantes, resulta valioso contrastar el modelo TCRA-Markov con el modelo autorregresivo de cambio de régimen de Markov (Markov-Switching Autoregressive o MS-AR), introducido por Hamilton [18] y ampliamente utilizado en econometría financiera. Ambos modelos comparten la hipótesis subyacente de que el comportamiento de la serie temporal cambia entre distintos estados o regímenes de volatilidad y crecimiento. Sin embargo, difieren sustancialmente en su diseño estructural y computacional:

1. **Observabilidad del régimen:** En el modelo MS-AR clásico, la variable de estado o régimen S_t es un proceso latente no observable. Los regímenes deben inferirse probabilísticamente a partir de la serie histórica del PIB o los precios utilizando técnicas de filtrado y suavizado. En cambio, en el modelo TCRA-Markov, los estados son observables de forma determinista y directa en cada instante mediante la cuantización de la serie $\{\alpha_t\}$ resultante del operador local.

2. **Complejidad en la estimación de parámetros:** La estimación en los modelos MS-AR requiere maximizar una función de verosimilitud no convexa y altamente no lineal, habitualmente mediante el algoritmo de Esperanza-Maximización (EM) o filtros numéricos iterativos. Este proceso suele sufrir problemas de convergencia, alta sensibilidad a los valores iniciales y la existencia de múltiples óptimos locales. Por el contrario, al ser los estados del modelo TCRA-Markov completamente deterministas tras fijar los parámetros del operador (W, λ) , la estimación de la matriz de transición de Markov se reduce a conteos de frecuencia analíticos sencillos (Proposición 3.6), garantizando soluciones globales únicas de forma casi instantánea.

3. **Flexibilidad y suposiciones:** El modelo MS-AR asume una estructura de error paramétrica estricta (usualmente gaussiana). El modelo TCRA-Markov, al fundamentarse en un operador analítico que actúa como filtro local, no impone supuestos distribucionales rígidos sobre los residuos de la serie original, dotándolo de mayor robustez frente a colas pesadas o valores atípicos comunes en series financieras de alta frecuencia.

Capítulo 4

Modelo Híbrido TCROC-Markov con Estimación NNLS

Este capítulo presenta una metodología que combina la discretización de regímenes económicos (TCROC-Markov) con la estimación de matrices de transición mediante Mínimos Cuadrados No Negativos (NNLS). Con este enfoque se abordan dos problemas frecuentes en la modelización de dinámicas económicas: la falta de interpretabilidad en modelos continuos y la dificultad de obtener matrices de transición válidas.

El enfoque tiene dos componentes:

- I Una discretización que transforma series temporales continuas en estados económicos discretos, lo que facilita tanto el análisis cualitativo como cuantitativo de regímenes.
- II Un esquema de estimación basado en NNLS que garantiza, por construcción, que las matrices de transición obtenidas sean estocásticas: entradas no negativas y columnas que suman uno.

Se describe el flujo completo desde la obtención de la serie temporal hasta la estimación y validación de la matriz de transición. Se incluyen ejemplos con datos reales del PIB y precios de combustibles.

La Figura 4.1 ilustra el proceso completo.

4.1. Discretización de Regímenes Económicos con TCROC-Markov

4.1.1. Cálculo de la Tasa Relativa

El punto de partida es la transformación de una serie temporal económica $\{v(t)\}_{t=0}^T$ en una secuencia de tasas de cambio relativas α_t . Para cada instante $t \geq W$, la tasa $\alpha_{t,\lambda,W}$ se

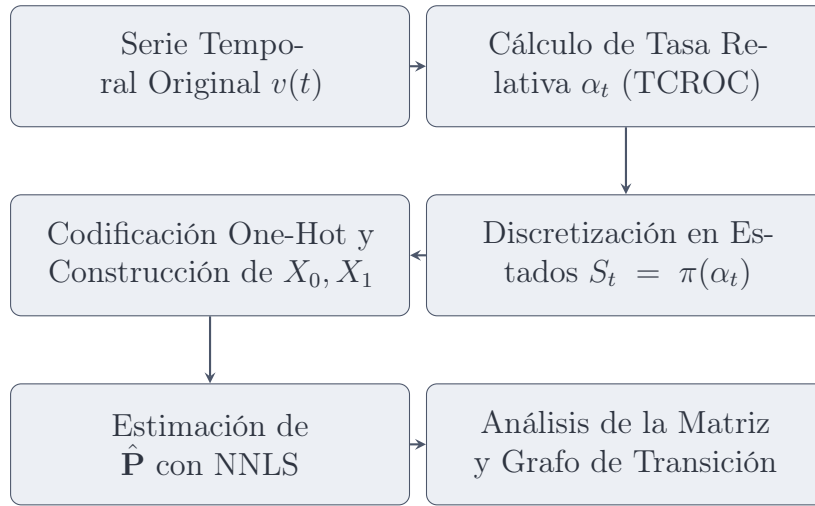


Figura 4.1: Diagrama conceptual de la metodología híbrida TCROC-Markov con estimación NNLS.

obtiene con el operador local generalizado ya definido en la Ecuación (3.1) del Capítulo 3; para evitar la duplicidad, no se reproduce aquí su expresión y se remite a esa formulación. El cociente que define dicha tasa puede interpretarse como una proyección ponderada del vector de la serie temporal v sobre su versión rezagada $v - 1$ dentro de una ventana W . El factor λ introduce una memoria exponencial que asigna mayor importancia a las observaciones más recientes.

4.1.2. Asignación de Estados

A continuación, se define una función de cuantización π que mapea cada tasa continua α_t a un estado discreto S_t perteneciente a un conjunto finito de regímenes económicos $S = \{s_1, s_2, \dots, s_N\}$. Por ejemplo:

$$S_t = \pi(\alpha_{t,\lambda,W}) = \begin{cases} s_1 & \text{si } \alpha > 0.05 \text{ (Crecimiento fuerte)} \\ s_2 & \text{si } 0 < \alpha \leq 0.05 \text{ (Crecimiento moderado)} \\ s_3 & \text{si } -0.05 \leq \alpha \leq 0 \text{ (Decrecimiento leve)} \\ s_4 & \text{si } \alpha < -0.05 \text{ (Decrecimiento severo)} \end{cases} \quad (4.1)$$

Este proceso genera una secuencia de estados $\{S_t\}_{t=0}^T$, que forma la base para el modelado de transiciones. Los umbrales ± 0.05 ilustrados en la Ecuación 4.1 se basan en

las categorías estándar de clasificación del crecimiento del PIB del FMI¹.

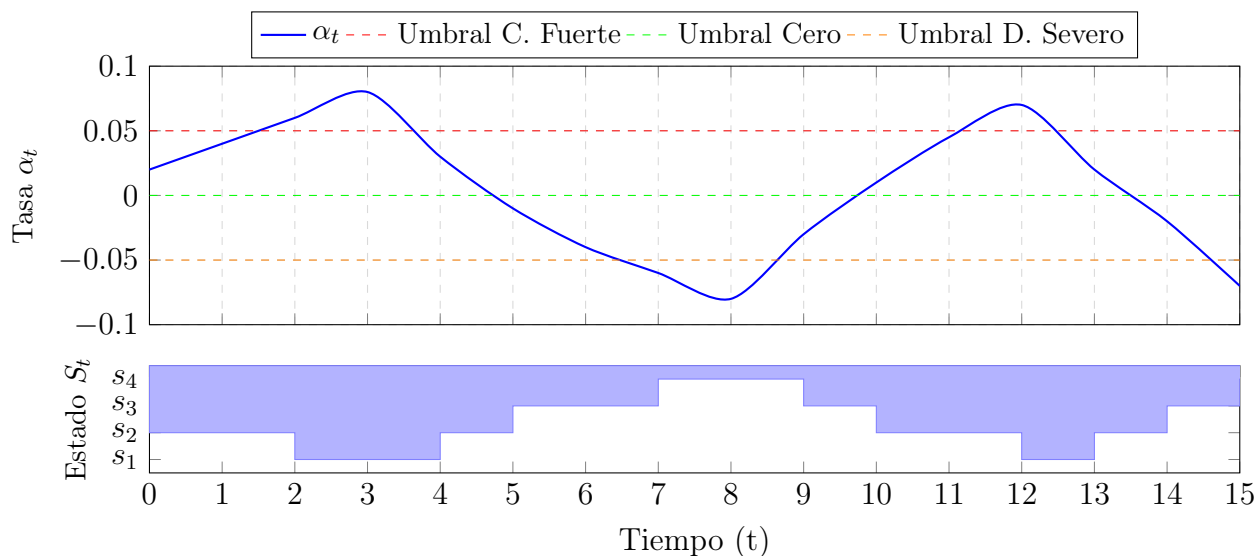


Figura 4.2: Ilustración del proceso de discretización. El panel superior muestra la evolución de la tasa continua α_t . El panel inferior muestra la secuencia de estados discretos S_t resultante de aplicar la función de cuantización π .

Ejemplo 4.1 (Discretización y Preparación de Datos). Considérese una serie temporal hipotética de inflación trimestral $v(t)$ y los parámetros $\lambda = 1$, $W = 2$. La secuencia de datos es:

$$v = \{v_0, v_1, v_2, v_3, v_4\} = \{4.0, 4.1, 4.15, 4.10, 4.05\}$$

Dado que $W = 2$, la primera tasa α_t que podemos calcular es para $t = 2$. Aplicando la fórmula se obtienen los siguientes valores y la clasificación en estados:

| t | α_t | Estado | Justificación |
|-----|-------------------|--------|-----------------------------|
| 2 | ≈ 0.0184 | s_2 | $0 < 0.0184 \leq 0.05$ |
| 3 | ≈ -0.0001 | s_3 | $-0.05 \leq -0.0001 \leq 0$ |
| 4 | ≈ -0.0121 | s_3 | $-0.05 \leq -0.0121 \leq 0$ |

La secuencia de regímenes generada es $\{S_2, S_3, S_4\} = \{s_2, s_3, s_3\}$. Esta secuencia se convierte a su representación *one-hot*² (e.g., $s_2 \rightarrow [0, 1, 0, 0]^T$) para construir las matrices de datos X_0 y X_1 , que son los insumos para la estimación descrita en la Sección 4.2.

¹Véase [24]. En el Ejemplo 4.2, los umbrales se ajustan al rango empírico del PIB per cápita hondureño; en el Capítulo 5 se reemplazan por particiones endógenas vía K-Means.

²Codificación indicadora: cada estado categórico se representa como un vector binario con un único componente igual a 1 y el resto iguales a 0. Véase el Glosario de Términos.

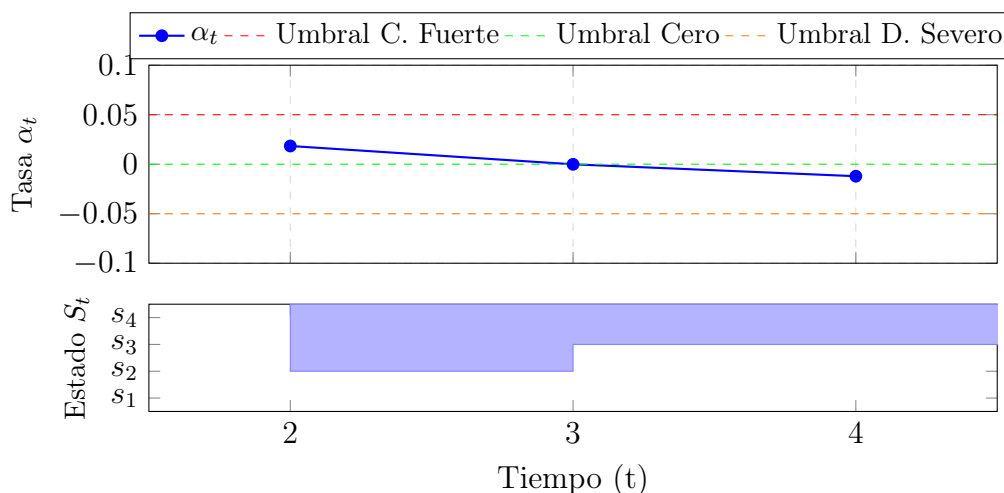


Figura 4.3: Visualización del proceso de discretización para el Ejemplo 4.1. El panel superior muestra los valores calculados de la tasa α_t . El panel inferior muestra la secuencia de estados discretos $\{s_2, s_3, s_3\}$ resultante.

Ejemplo 4.2 (Discretización del PIB per cápita usando TCROC). Considérese la serie temporal del PIB per cápita en Lempiras a precios constantes de Honduras (2000–2023), presentada en la Tabla 1.1. Aplicamos el paso de discretización TCROC con ventana $W = 2$ y factor de memoria $\lambda = 1$, usando la fórmula:

$$\alpha_t = -1 + \frac{v(t)v(t-1) + v(t-1)v(t-2)}{v(t-1)^2 + v(t-2)^2}, \quad t \geq 2.$$

Para $t = 2$ (año 2002), se obtiene $\alpha_2 \approx 0.0052$. Como $-0.01 \leq 0.0052 \leq 0.02$, el estado es $\mathbf{S}_2 = \mathbf{s}_2$.

Los valores de α_t para $t = 3$ hasta $t = 23$ se calcularon con la misma fórmula. La siguiente tabla resume los resultados y la clasificación en cuatro regímenes:

- s_1 : $\alpha_t < -0.01$ (contracción fuerte)
- s_2 : $-0.01 \leq \alpha_t \leq 0.02$ (estabilidad o crecimiento débil)
- s_3 : $0.02 < \alpha_t \leq 0.04$ (crecimiento moderado)
- s_4 : $\alpha_t > 0.04$ (crecimiento fuerte)

| t | Año | α_t | Estado |
|-----|------|------------|--------|
| 2 | 2002 | 0.0052 | s_2 |
| 3 | 2003 | 0.0147 | s_2 |
| 4 | 2004 | 0.0277 | s_3 |
| 5 | 2005 | 0.0358 | s_3 |
| 6 | 2006 | 0.0381 | s_3 |
| 7 | 2007 | 0.0393 | s_3 |
| 8 | 2008 | 0.0282 | s_3 |
| 9 | 2009 | -0.0136 | s_1 |
| 10 | 2010 | -0.0167 | s_1 |
| 11 | 2011 | 0.0166 | s_2 |
| 12 | 2012 | 0.0186 | s_2 |
| 13 | 2013 | 0.0139 | s_2 |
| 14 | 2014 | 0.0095 | s_2 |
| 15 | 2015 | 0.0151 | s_2 |
| 16 | 2016 | 0.0192 | s_2 |
| 17 | 2017 | 0.0246 | s_3 |
| 18 | 2018 | 0.0241 | s_3 |
| 19 | 2019 | 0.0134 | s_2 |
| 20 | 2020 | -0.0500 | s_1 |
| 21 | 2021 | -0.0106 | s_1 |
| 22 | 2022 | 0.0611 | s_4 |
| 23 | 2023 | 0.0211 | s_3 |

La secuencia de regímenes resultante es:

$$\{S_t\}_{t=2}^{23} = \{s_2, s_2, s_3, s_3, s_3, s_3, s_3, s_1, s_1, s_2, s_2, s_2, s_2, s_2, s_2, s_3, s_3, s_2, s_1, s_1, s_4, s_3\}$$

Esta secuencia se codifica en formato *one-hot* para construir las matrices de datos X_0 y X_1 , utilizadas en la estimación de la matriz de transición mediante NNLS (Sección 4.2).

La dinámica de los regímenes económicos para el PIB per cápita, visible en la Figura 4.4,

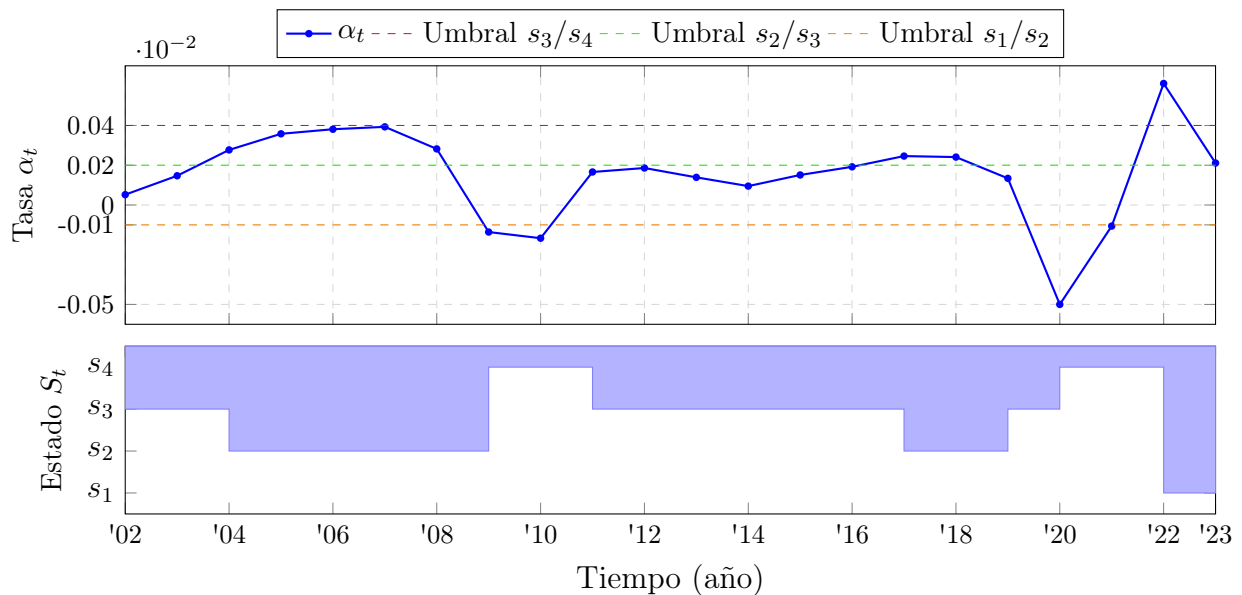


Figura 4.4: Proceso completo de discretización del PIB per cápita en Lempiras (2002–2023). El panel superior muestra la evolución de la tasa α_t ; el inferior, la secuencia de estados discretos S_t . Se observan regímenes de contracción (2009, 2020), estabilidad y crecimiento fuerte (2022).

ofrece una perspectiva cuantitativa de la historia económica reciente de Honduras. Se pueden extraer dos conclusiones principales:

Persistencia de Regímenes Durante un Período Político Extendido Se observa una notable persistencia de los estados s_2 (estabilidad) y s_3 (crecimiento moderado) durante la mayor parte del período analizado, particularmente entre 2004-2008 y 2011-2019. Este comportamiento de estabilidad y crecimiento predecible coincide en gran medida con los períodos de gobierno del Partido Nacional, incluyendo la administración de Juan Orlando Hernández (2014-2022). Las transiciones abruptas a s_1 (contracción) en 2009 y 2020 marcan eventos de crisis bien conocidos (la crisis política de 2009 y la pandemia de COVID-19), interrumpiendo estos largos ciclos de estabilidad.

El Estado s_4 como Anomalía Post-Pandemia La aparición única del estado s_4 (crecimiento fuerte) en el año 2022 es particularmente significativa. Este régimen, que no se observa en ningún otro punto de los 22 años analizados, puede interpretarse como un efecto de rebote.^o una corrección anómala de la economía tras la drástica contracción de 2020. Su carácter excepcional sugiere que no representa una dinámica estructural sostenible, sino más bien una consecuencia transitoria de la recuperación post-COVID y de las condiciones

macroeconómicas globales de ese año. La escasez de este estado en la muestra limita la capacidad del modelo para predecir eventos de crecimiento tan acelerado.

4.1.3. Análisis de Sensibilidad de los Parámetros de Discretización

La discretización de la serie temporal mediante TCROC depende de dos hiperparámetros clave: la ventana de tiempo W y el factor de memoria λ . Para evaluar la robustez de nuestra discretización, se realizó un análisis de sensibilidad exhaustivo sobre los datos del PIB per cápita, variando $W \in \{2, 3, 4, 5\}$ y λ en el rango $[0.80, 1.0]$. La Tabla 4.1 presenta una selección representativa de estos resultados. El código completo utilizado para generar estos datos se encuentra en el Apéndice B.2.

Tabla 4.1: Análisis de sensibilidad: Distribución de estados (%) del PIB per cápita según los parámetros W y λ .

| W | λ | % s_1 | % s_2 | % s_3 | % s_4 |
|----------|-------------|-------------|-------------|-------------|------------|
| 2 | 1.00 | 18.2 | 40.9 | 36.4 | 4.5 |
| 2 | 0.90 | 13.6 | 45.5 | 36.4 | 4.5 |
| 3 | 1.00 | 4.8 | 57.1 | 33.3 | 4.8 |
| 3 | 0.80 | 4.8 | 57.1 | 38.1 | 0.0 |
| 4 | 1.00 | 5.0 | 70.0 | 25.0 | 0.0 |
| 4 | 0.80 | 5.0 | 65.0 | 30.0 | 0.0 |
| 5 | 1.00 | 0.0 | 79.0 | 21.1 | 0.0 |
| 5 | 0.80 | 5.3 | 68.4 | 26.3 | 0.0 |

Del análisis se desprenden dos conclusiones clave. Por un lado, **la metodología es altamente robusta a la elección de λ** ; para una ventana W fija, las variaciones en el factor de memoria tienen un impacto mínimo en la distribución de estados. Fijar $\lambda = 1.0$ no es una decisión neutra: ese valor cancela el decaimiento exponencial y reparte el mismo peso entre todas las observaciones de la ventana, sin privilegiar las más recientes. Aquí esa ponderación uniforme basta, porque la robustez observada muestra que priorizar lo reciente no altera de forma apreciable la clasificación de estados. Por eso se adopta $\lambda = 1.0$ por su simplicidad y efectividad.

Por otro lado, **el parámetro más influyente es la ventana W** , que actúa como un operador de suavizado. A medida que W aumenta, la detección de regímenes extremos (s_1 y s_4) disminuye, favoreciendo al estado de estabilidad (s_2). En contrapartida, una ventana

más amplia gana robustez frente al ruido de alta frecuencia, pero lo paga con un retraso dinámico: amortigua y demora la detección de los cambios bruscos de tendencia. Una ventana corta hace lo contrario, eleva la sensibilidad del operador y, con ella, la varianza de la clasificación de estados. Se seleccionó $W = 2$ para el estudio principal, ya que representa un equilibrio que, si bien confirma el predominio de la estabilidad, es lo suficientemente sensible para capturar las contracciones económicas agudas correspondientes a eventos históricos conocidos. Una ventana de solo dos observaciones tampoco compromete la representatividad estadística del modelo: $W = 2$ fija únicamente el bloque con el que se mide la tasa de cambio local instantánea, mientras que la matriz de transición $\hat{\mathbf{P}}$ se estima sobre la serie histórica completa de T observaciones. La consistencia de los resultados a través de los parámetros valida la fiabilidad de las conclusiones del modelo.

La elección de una ventana corta como $W = 2$ se alinea conceptualmente con la propiedad de Markov de primer orden que se asume en la fase de modelado. Si bien la propiedad de Markov se aplica a la secuencia de estados (S_t) y no a los datos brutos $(v(t))$, el uso de una ventana de discretización mínima evita incorporar una memoria a largo plazo en la definición de los propios estados, manteniendo así la simplicidad y consistencia metodológica en todo el proceso.

4.2. Estimación de la Matriz de Transición Lineal con NNLS

Una vez obtenida la secuencia de estados discretos $\{S_t\}_{t=0}^T$, el siguiente paso es modelar su dinámica. Asumiendo un proceso de Markov de primer orden, la relación se describe mediante el modelo lineal $x(t+1) = Px(t)$. Para estimar la matriz de transición P , se emplea un enfoque de **Mínimos Cuadrados No Negativos (NNLS)**, siguiendo la metodología detallada en Vides y Lopez Gutierrez (2025) [43].

Este método resulta central porque garantiza, por construcción, las propiedades estocásticas de la matriz estimada ($\hat{P}_{ij} \geq 0$ y $\mathbf{1}^T \hat{\mathbf{P}} = \mathbf{1}^T$), resolviendo el siguiente problema de

optimización:

$$\hat{\mathbf{P}} = \arg \min_{P \in \mathcal{M}_{N \times N}(\mathbb{R})} \|X_1 - PX_0\|_F^2 \quad \text{sueto a} \quad P_{ij} \geq 0 \text{ y } \mathbf{1}^T P = \mathbf{1}^T. \quad (4.2)$$

Este enfoque robusto evita soluciones inadmisibles, como probabilidades de transición negativas, que podrían surgir con estimadores de mínimos cuadrados ordinarios no restringidos. La solución a este problema se basa en algoritmos clásicos y eficientes, como el propuesto en la obra seminal de Lawson y Hanson (1995) [30].

4.2.1. Formulación Matemática y Solución Constructiva

El objetivo es identificar la matriz de transición $\hat{\mathbf{P}}$ que minimiza el error de predicción en un paso. Dadas las matrices de datos X_0 (estados en t) y X_1 (estados en $t + 1$), el problema de optimización se plantea como la minimización de la norma de Frobenius del error residual:

$$\hat{\mathbf{P}} = \arg \min_{P \in \mathbb{S}(N,N)} \|X_1 - PX_0\|_F^2 \quad (4.3)$$

donde $\mathbb{S}_{N,N}(\mathbb{R}) = \{P \in \mathcal{M}_{N \times N}(\mathbb{R}) : P_{ij} \geq 0, \mathbf{1}^T P = \mathbf{1}^T\}$ es el conjunto de matrices estocásticas (por columnas) de tamaño $N \times N$.

Corolario 4.1 (del Teorema 3 en [7]). *El problema de estimación planteado en la Ecuación 4.3 admite una solución constructiva en el espacio de matrices estocásticas.*

Prueba constructiva. La prueba sigue los pasos algorítmicos para transformar el problema matricial en un sistema lineal resoluble. La ecuación matricial del modelo, $X_1 = \hat{\mathbf{P}}X_0$, se linealiza aplicando el operador de **vectorización** ($\text{vec}(\cdot)$), que apila las columnas de una matriz en un único vector columna (i.e., $\text{vec}(A) = [A_{11}, A_{21}, \dots, A_{nn}]^T$). Utilizando la identidad del producto de Kronecker $\text{vec}(ABC) = (C^T \otimes A)\text{vec}(B)$, se obtiene:

$$\text{vec}(X_1) = (X_0^T \otimes I_N)\text{vec}(\hat{\mathbf{P}}) \quad (4.4)$$

Este es un sistema lineal de la forma $b = \mathcal{A}c$, donde $c = \text{vec}(\hat{\mathbf{P}})$ es el vector de incógnitas, $\mathcal{A} = X_0^T \otimes I_N$ es la matriz de diseño, y $b = \text{vec}(X_1)$ es el vector de observaciones. El

problema de minimización se convierte en $\min_c \|\mathcal{A}c - b\|_2^2$, sujeto a las restricciones que definen una matriz estocástica.

Las restricciones se traducen al dominio vectorial:

- La **no negatividad** ($\hat{P}_{ij} \geq 0$) es equivalente a la restricción de desigualdad $c \geq 0$.
- La **suma de columnas unitaria** ($\sum_{i=1}^N \hat{P}_{ij} = 1$) se formula como un sistema de restricciones de igualdad lineal $\mathcal{C}c = \mathbf{1}_N$, con $\mathcal{C} = I_N \otimes \mathbf{1}_N^T$.

Ambas condiciones se integran en un único problema de Mínimos Cuadrados No Negativos (NNLS) a través de un sistema aumentado, cuya solución \hat{c} se encuentra al resolver:

$$\hat{c} = \arg \min_{c \geq 0} \left\| \begin{bmatrix} \mathcal{A} \\ w\mathcal{C} \end{bmatrix} c - \begin{bmatrix} b \\ w\mathbf{1}_N \end{bmatrix} \right\|_2^2 \quad (4.5)$$

El conjunto de factibilidad es convexo y no vacío: la restricción de no negatividad $c \geq 0$ define la intersección de semiespacios cerrados, mientras que la restricción de igualdad lineal $\mathcal{C}c = \mathbf{1}_N$ define un hiperplano afín. La intersección de un número finito de semiespacios cerrados con un hiperplano afín es un politopo convexo, que aquí coincide con el simplex de probabilidad por columnas. Dado que, además, el funcional objetivo es cuadrático, la existencia de un minimizador \hat{c} está garantizada. La solución $\hat{\mathbf{P}} = \text{vec}^\dagger(\hat{c})$ es, por construcción, estocástica. Este procedimiento algorítmico, implementado en la función `SRep`, demuestra constructivamente la existencia de la solución [7, 43]. \square

Observación (Sobre la unicidad de la solución). Dado que el funcional objetivo en la Ecuación 4.3, la norma de Frobenius al cuadrado, es estrictamente convexo, y el conjunto factible $\mathbb{S}_{N,N}(\mathbb{R})$ también es convexo, el problema de optimización admite un minimizador global único. Por lo tanto, la matriz de transición $\hat{\mathbf{P}}$ estimada mediante este procedimiento no solo existe, sino que es única.

El Algoritmo 2 formaliza el procedimiento completo de estimación, desde la secuencia de estados hasta la matriz de transición estocástica.

Algoritmo 2 SRep: Estimación de $\hat{\mathbf{P}}$ vía NNLS aumentado

Entrada: Secuencia de estados $\{S_t\}_{t=0}^T$, N estados, peso $w \gg 1$
Salida: Matriz de transición estocástica $\hat{\mathbf{P}} \in \mathcal{M}_{N \times N}(\mathbb{R})$

- 1: Codificar $\{S_t\}$ en representación one-hot: $x(t) \in \mathbb{R}^N$
- 2: $X_0 \leftarrow [x(0), x(1), \dots, x(T-1)]$ ▷ Estados actuales
- 3: $X_1 \leftarrow [x(1), x(2), \dots, x(T)]$ ▷ Estados siguientes
- 4: $\mathcal{A} \leftarrow X_0^\top \otimes I_N$ ▷ Matriz de diseño (Kronecker)
- 5: $b \leftarrow \text{vec}(X_1)$
- 6: $\mathcal{C} \leftarrow I_N \otimes \mathbf{1}_N^\top$ ▷ Restricción de suma unitaria
- 7: Construir sistema aumentado:

$$\tilde{\mathcal{A}} \leftarrow \begin{bmatrix} \mathcal{A} \\ w\mathcal{C} \end{bmatrix}, \quad \tilde{b} \leftarrow \begin{bmatrix} b \\ w\mathbf{1}_N \end{bmatrix}$$

- 8: $\hat{c} \leftarrow \text{NNLS}(\tilde{\mathcal{A}}, \tilde{b})$ ▷ Resolver $\min_{c \geq 0} \|\tilde{\mathcal{A}}c - \tilde{b}\|_2^2$
 - 9: $\hat{\mathbf{P}} \leftarrow \text{vec}^{-1}(\hat{c})$ ▷ Reconstruir matriz $N \times N$
 - 10: **para** $j = 1$ **hasta** N **hacer** ▷ Normalización por columnas
 - 11: $\hat{\mathbf{P}}_{:,j} \leftarrow \hat{\mathbf{P}}_{:,j} / \sum_{i=1}^N \hat{P}_{ij}$
 - 12: **fin para**
 - 13: **devolver** $\hat{\mathbf{P}}$
-

Ejemplo 4.3 (Estimación Validada de la Matriz de Transición con NNLS). A partir de la secuencia de estados discretos obtenida mediante TCROC (Ejemplo 4.2), se procede a estimar la matriz de transición $\hat{\mathbf{P}}$ utilizando el enfoque de Mínimos Cuadrados No Negativos (NNLS). La secuencia completa de estados, para $t = 2$ hasta $t = 23$ (2002–2023), la secuencia $\{S_t\}_{t=2}^{23}$ es:

| | | | | | | | | | | | | | | | | | | | | | |
|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|
| S2 | S2 | S3 | S3 | S3 | S3 | S3 | S1 | S1 | S2 | S2 | S2 | S2 | S2 | S2 | S3 | S3 | S2 | S1 | S1 | S4 | S3 |
| 02 | 03 | 04 | 05 | 06 | 07 | 08 | 09 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 |

Figura 4.5: Secuencia temporal de estados discretos obtenidos mediante TCROC.

Con $N = 4$ estados definidos como:

- s_1 : Contracción ($\alpha_t < -0.01$)
- s_2 : Estabilidad ($-0.01 \leq \alpha_t \leq 0.02$)
- s_3 : Crecimiento moderado ($0.02 < \alpha_t \leq 0.04$)
- s_4 : Crecimiento fuerte ($\alpha_t > 0.04$)

Representados en formato *one-hot*:

$$s_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad s_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \quad s_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \quad s_4 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

Se construyen las matrices de datos, donde $\mathcal{M}_{m \times n}(\mathbb{R})$ denota el espacio de las matrices reales de orden $m \times n$:

- $X_0 = [x(2), x(3), \dots, x(22)] \in \mathcal{M}_{4 \times 21}(\mathbb{R})$: estados actuales.
- $X_1 = [x(3), x(4), \dots, x(23)] \in \mathcal{M}_{4 \times 21}(\mathbb{R})$: estados siguientes.

Construcción de las Matrices de Datos A partir de la secuencia de estados y su codificación *one-hot*, se construyen las matrices de datos X_0 y X_1 . La matriz X_0 se forma con los vectores de los estados actuales ($t = 2$ a $t = 22$), y la matriz X_1 con los vectores de los estados siguientes ($t = 3$ a $t = 23$).

$$X_0 = \begin{bmatrix} 0, 0, 0, 0, 0, 0, 0, 0, 1, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 1, 0 \\ 1, 1, 0, 0, 0, 0, 0, 0, 0, 0, 1, 1, 1, 1, 1, 1, 0, 0, 1, 0, 0, 0, 0 \\ 0, 0, 1, 1, 1, 1, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 1, 0, 0, 0, 0, 0 \\ 0, 1 \end{bmatrix}$$

$$X_1 = \begin{bmatrix} 0, 0, 0, 0, 0, 0, 1, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 1, 0, 0 \\ 1, 0, 0, 0, 0, 0, 0, 0, 1, 1, 1, 1, 1, 1, 1, 0, 1, 0, 0, 0, 0, 0, 0 \\ 0, 1, 1, 1, 1, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 1, 0, 0, 0, 0, 1 \\ 0, 1 \end{bmatrix}$$

Estas dos matrices son la base para el conteo de transiciones y la posterior estimación con NNLS. Para estimar $\hat{\mathbf{P}}$, se cuenta el número de transiciones entre estados. Sea $C \in \mathbb{Z}^{4 \times 4}$ la matriz de conteo, donde C_{ij} es el número de transiciones de s_j a s_i . Esta matriz se puede calcular eficientemente a partir de las matrices de datos X_1 y X_0 mediante la siguiente

operación:

$$C = X_1 X_0^T \quad (4.6)$$

Recorriendo los pares (S_t, S_{t+1}) de la secuencia, se obtiene:

$$C = \begin{bmatrix} 2 & 1 & 1 & 0 \\ 1 & 6 & 1 & 0 \\ 0 & 2 & 5 & 1 \\ 1 & 0 & 0 & 0 \end{bmatrix}$$

La matriz de transición se obtiene normalizando esta matriz de conteo por columnas:

$$\hat{P}_{ij} = \frac{C_{ij}}{\sum_{k=1}^4 C_{kj}}, \quad \forall j \quad (4.7)$$

$$\hat{\mathbf{P}} = \begin{bmatrix} 0.5000 & 0.1111 & 0.1429 & 0.0000 \\ 0.2500 & 0.6667 & 0.1429 & 0.0000 \\ 0.0000 & 0.2222 & 0.7143 & 1.0000 \\ 0.2500 & 0.0000 & 0.0000 & 0.0000 \end{bmatrix}$$

Para validar empíricamente la correcta implementación del algoritmo de estimación propuesto en [43], se realiza un experimento numérico que consiste en dos fases. En la primera fase, se utiliza la matriz $\hat{\mathbf{P}}$ como “verdad fundamental” para simular una nueva secuencia de estados. En la segunda fase, se aplica sobre estos datos simulados el algoritmo de Mínimos Cuadrados No Negativos (NNLS) implementado en la función `SRep`, cuya metodología se detalla en el Listado 1. El objetivo es verificar si el estimador recupera la matriz original.

La implementación completa del algoritmo `SRep` se presenta en el Apéndice B.4 (Listado 1).

El resultado de la validación se evalúa calculando la matriz de diferencia entre la matriz estimada y la original $(Ap - \hat{\mathbf{P}})$. Tal como se muestra a continuación, el resultado es una matriz donde todos los elementos son numéricamente cero, con errores del orden de 10^{-13} atribuibles únicamente a la precisión de punto flotante:

$$Ap - \hat{\mathbf{P}} \approx \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Este resultado valida empíricamente que la función SRep implementa correctamente el estimador propuesto, logrando recuperar la matriz de transición original con alta precisión.

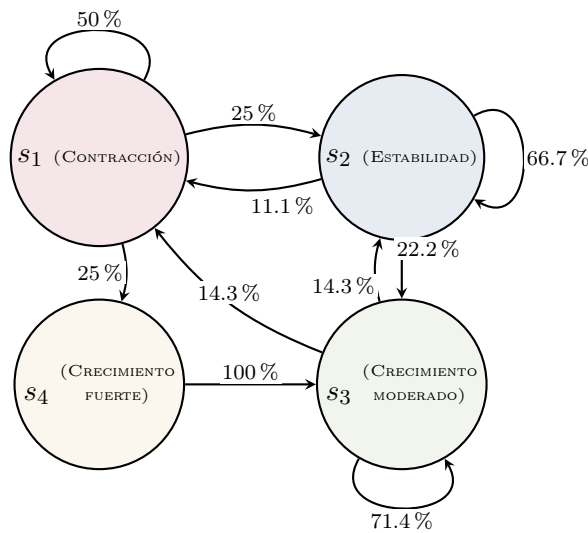


Figura 4.6: Grafo de transición estimado para los regímenes del PIB per cápita en Lempiras.

Interpretación de la Matriz de Transición La matriz $\hat{\mathbf{P}}$ revela la dinámica subyacente del sistema:

- **Estado s_1 (Contracción):** Persiste con una probabilidad del 50%, mientras que se recupera transitando a s_2 (estabilidad) o a s_4 (crecimiento fuerte) con una probabilidad del 25% en cada caso.
- **Estado s_2 (Estabilidad):** Es el régimen más persistente, con una probabilidad del 66.7% de mantenerse. Su principal vía de salida es hacia s_3 (crecimiento moderado), con un 22.2%.
- **Estado s_3 (Crecimiento moderado):** También es altamente estable, persistiendo con un 71.4%. La probabilidad de retroceso a otros estados es baja.

- **Estado s_4 (Crecimiento fuerte):** Actúa como un estado puramente transitorio. En la muestra observada, siempre es seguido por una transición al estado s_3 , lo que sugiere que los picos de crecimiento tienden a moderarse rápidamente.

Propiedades Teóricas del Estimador La matriz $\hat{\mathbf{P}}$, calculada a partir de las frecuencias relativas de transición, corresponde al **Estimador de Máxima Verosimilitud (MLE)** para una cadena de Markov de primer orden. El supuesto distribucional merece hacerse explícito: el MLE modela las transiciones salientes de cada estado de origen como ensayos independientes bajo una distribución multinomial condicional, sin imponer ninguna distribución exógena sobre la estructura de los coeficientes de la matriz P . Bajo ese supuesto, maximizar la verosimilitud de la cadena equivale a normalizar por columnas los conteos de transición, y el estimador resultante posee propiedades asintóticas bien establecidas.

Teorema 4.2 (Consistencia del Estimador). *Dada una cadena de Markov ergódica con una matriz de transición verdadera (e inobservable) P , el estimador de máxima verosimilitud $\hat{\mathbf{P}}$, obtenido a partir de una secuencia de T observaciones, es consistente. Esto implica que el estimador converge en probabilidad a la matriz verdadera a medida que el número de observaciones tiende a infinito:*

$$\hat{\mathbf{P}} \xrightarrow{p} P \quad \text{cuando } T \rightarrow \infty$$

Demostración. Por la ley fuerte de los grandes números para cadenas de Markov ergódicas [36], las frecuencias relativas de transición convergen casi seguramente:

$$\hat{p}_{ij} = \frac{N_{ij}}{N_i} \xrightarrow{\text{c.s.}} p_{ij}, \quad T \rightarrow \infty,$$

donde $N_{ij} = \#\{t : S_t = i, S_{t+1} = j\}$ y $N_i = \sum_j N_{ij}$. Dado que la cadena es ergódica, cada estado se visita infinitas veces, por lo que $N_i \rightarrow \infty$ para todo i . La convergencia entrada-por-entrada implica $\|\hat{\mathbf{P}} - P\|_{\text{máx}} \xrightarrow{\text{c.s.}} 0$. Este resultado fue establecido formalmente por Anderson y Goodman [3]. El estimador NNLS coincide con el MLE bajo las restricciones de no-negatividad, heredando así la consistencia. \square

La inclusión de este resultado es necesaria, pues certifica que, con una cantidad suficiente de datos, nuestra metodología se aproxima al verdadero proceso generador de la dinámica de regímenes. La consistencia descansa sobre una hipótesis de ergodicidad cuyo alcance hay que delimitar: la cadena debe ser ergódica, es decir, estacionaria en sentido asintótico, irreducible y aperiódica, de modo que cada estado se visite de forma recurrente. El operador local de tasas α_t contribuye a satisfacer ese requisito, ya que estabiliza la varianza e induce estacionariedad débil en la serie de estados $\{S_t\}$; esta condición se verifica empíricamente en el Capítulo 5 mediante las pruebas ADF y KPSS. A diferencia del algoritmo de Esperanza-Maximización (EM) con el que se estiman los modelos de cambio de régimen tipo MS-AR, cuya función de verosimilitud no convexa solo garantiza óptimos locales dependientes de la inicialización, el estimador empleado aquí se reduce a un problema convexo con solución única; esta comparación se retoma con datos reales en el Capítulo 5.

Simulación de la Evolución de la Cadena de Markov Para validar la capacidad del estimador SRep de recuperar la matriz de transición original, se simula la evolución de una cadena de Markov utilizando la matriz $\hat{\mathbf{P}}$ estimada previamente. Se define un estado inicial correspondiente a s_1 y se genera una secuencia de ocupaciones probabilísticas para $T = 20$ pasos (véase Apéndice B.4, Listado 2).

La Figura 4.7 muestra la evolución temporal de las probabilidades de ocupación para cada uno de los estados del sistema Markoviano, bajo la matriz de transición estimada $\hat{\mathbf{P}}$, a lo largo de 20 pasos de simulación.

- **Estado s_1 (contracción):** Inicia con probabilidad unitaria (1.0) y experimenta un descenso rápido durante los primeros 5 pasos. A partir del paso 6, la probabilidad se estabiliza alrededor de 0.195, oscilando mínimamente entre 0.1957 y 0.1951 hasta el final de la simulación.
- **Estado s_2 (estabilidad):** Comienza en cero y crece de forma sostenida. Se estabiliza aproximadamente en el paso 10, alrededor de 0.329, mostrando una variación insignificante (< 0.001) desde ese punto en adelante.
- **Estado s_3 (crecimiento moderado):** Aunque inicia en cero, incrementa rápida-

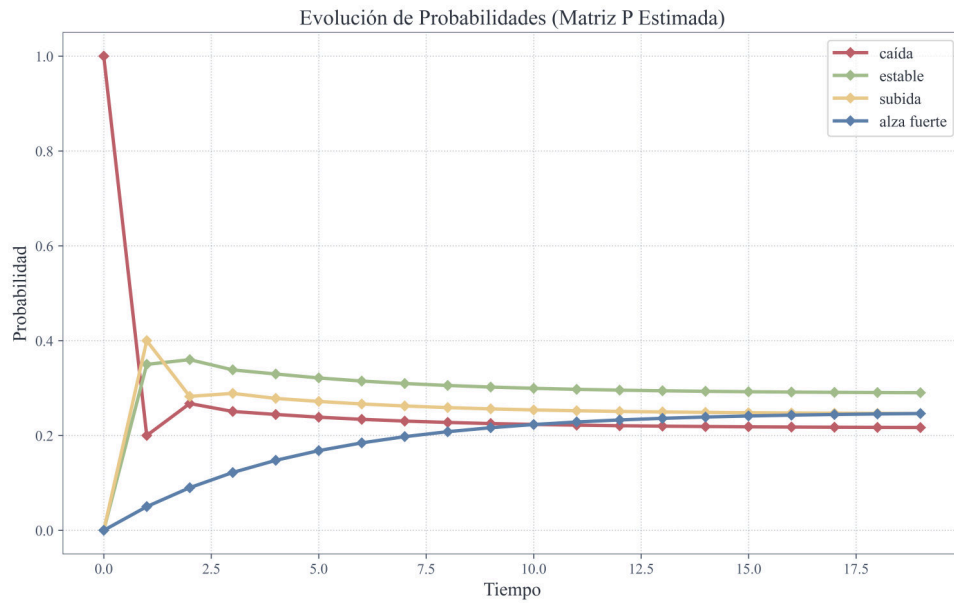


Figura 4.7: Evolución temporal de las probabilidades de ocupación de los estados s_1, s_2, s_3, s_4 bajo la dinámica inducida por la matriz de transición estimada \hat{P} .

mente y supera a los demás estados a partir del paso 2. A partir del paso 7, se mantiene en torno a 0.427, con oscilaciones menores a 0.001, confirmando su papel como estado dominante en el régimen estacionario.

- **Estado s_4 (crecimiento fuerte):** Presenta un leve incremento inicial pero desciende pronto, estabilizándose en torno a 0.0488 a partir del paso 6. Su baja probabilidad persistente evidencia su carácter transitorio.

Estabilización global: En general, la cadena de Markov alcanza el equilibrio estacionario a partir de los pasos 10 a 12, cuando las probabilidades de todos los estados dejan de variar de manera significativa y convergen a valores constantes: $s_1 \approx 0.195$, $s_2 \approx 0.329$, $s_3 \approx 0.427$ y $s_4 \approx 0.049$. Este comportamiento confirma que el sistema converge rápidamente a un régimen estable, dominado por los estados de crecimiento moderado y estabilidad.

En conjunto, la dinámica observada en la gráfica es plenamente consistente con la teoría de cadenas de Markov: los estados de crecimiento moderado y estabilidad emergen como los destinos predominantes a largo plazo, mientras que los estados extremos (contracción y crecimiento fuerte) son rápidamente abandonados y sólo se presentan de manera transitoria.

Esta convergencia hacia regímenes centrales y estables refleja tanto la estructura de la matriz de transición estimada $\hat{\mathbf{P}}$ como la naturaleza realista de procesos económicos o financieros modelados por este tipo de sistemas. El comportamiento del sistema valida, así, que la modelación mediante cadenas de Markov captura adecuadamente las dinámicas de persistencia y transición entre regímenes típicas en el análisis de series temporales económicas.

4.2.2. Validación del Rendimiento Predictivo

Para cuantificar la capacidad predictiva del modelo, se evaluó su rendimiento utilizando múltiples particiones de entrenamiento y prueba. Se entrenó la matriz de transición $\hat{\mathbf{P}}$ con porciones crecientes de la serie temporal (desde el 60 % hasta el 95 %) y se validó su capacidad para predecir el estado del siguiente período en los datos restantes. El código completo para este análisis de robustez se encuentra en el Apéndice B.3.

La Tabla 4.2 resume la precisión (accuracy) del modelo para cada partición.

| Entrenamiento/Prueba | Nº de Predicciones | Precisión (Accuracy) |
|----------------------|--------------------|----------------------|
| 60 % / 40 % | 8 | 0.375 |
| 70 % / 30 % | 6 | 0.167 |
| 80 % / 20 % | 4 | 0.000 |
| 90 % / 10 % | 2 | 0.000 |
| 95 % / 5 % | 1 | 0.000 |

Tabla 4.2: Resumen del rendimiento predictivo según la partición de datos.

Los resultados revelan un hallazgo clave: la precisión predictiva es mayor cuando el conjunto de prueba es más grande y disminuye a cero a medida que este se reduce. Este comportamiento, aparentemente contraintuitivo, no indica un deterioro del modelo. Por el contrario, se explica porque los conjuntos de prueba más pequeños aíslan las transiciones más recientes y anómalas de la serie, particularmente el evento único del estado s_4 (crecimiento fuerte) en 2022. Al no tener antecedentes en los datos de entrenamiento, el modelo no puede predecirlo, lo que ilustra una limitación clásica de los modelos de Markov basados en datos históricos: capturan bien dinámicas recurrentes, pero su capacidad ante eventos únicos queda limitada por la escasez de datos. No se trata, por tanto, de una incapacidad ni de una debilidad estructural del modelo, sino de un límite inherente a

cualquier estimador construido sobre frecuencias históricas: ningún esquema de Markov puede anticipar la primera ocurrencia de un shock exógeno que carece de precedentes en la muestra. Una vez que ese shock impacta la serie y traslada el sistema al estado correspondiente, el modelo recalibra de inmediato la condición de entrada y estima las probabilidades de transición del período siguiente a partir del nuevo estado de origen; su utilidad, por tanto, no se restringe a los tramos de estabilidad, sino que se readapta con rapidez una vez ocurrida la perturbación.

Ejemplo 4.4 (Definición y Discusión de Estados para Series de Combustibles). Para la discretización de las series de precios de combustibles, se definieron cuatro estados cuantitativos en función del parámetro α , que mide la tasa relativa de cambio en cada periodo. Estos estados se codifican como $s1, s2, s3, s4$, que se interpretan respectivamente como *caída, estable, subida* y *alza fuerte*.

La función de asignación utilizada fue la siguiente:

$$\text{estado} = \begin{cases} s1 & \text{si } \alpha < -0.01 \\ s2 & \text{si } -0.01 \leq \alpha \leq 0.02 \\ s3 & \text{si } 0.02 < \alpha \leq 0.04 \\ s4 & \text{si } \alpha > 0.04 \end{cases}$$

Esta definición, aunque arbitraria, responde a un criterio homogéneo aplicado a todas las series para facilitar la comparación directa entre ellas.

Pros y Contras del enfoque:

- **Límites fijos para todas las series:**
 - *Pros:* Permite un marco común para análisis comparativos y es sencillo de explicar y mantener.
 - *Contras:* Puede no captar particularidades dinámicas específicas de cada serie; por ejemplo, una serie con baja variabilidad puede quedar demasiado concentrada en un solo estado.

■ **Límites dinámicos o personalizados para cada serie:**

- *Pros:* Permite mayor sensibilidad a las fluctuaciones internas y mejor interpretación dentro de cada serie.
- *Contras:* Dificulta la comparación entre series al no tener un sistema uniforme de estados.

En este trabajo se adoptan límites fijos para todas las series con el objetivo de facilitar los análisis cualitativos y comparativos. Conviene subrayar que esta partición no queda como un criterio subjetivo y definitivo del investigador: en el capítulo empírico (Capítulo 5) estos umbrales fijos se sustituyen por una partición endógena estimada con el algoritmo de K-medias, que delimita los estados de forma automática a partir de la distribución observada de cada serie. Los límites fijos cumplen aquí un papel comparativo y transitorio; la asignación final de regímenes en los datos reales es de carácter endógeno.

A continuación, se presentan las matrices de conteo C , matrices de transición $\hat{\mathbf{P}}$, matrices estimadas A_p mediante la función $SRep$, y la diferencia entre estas matrices para las cuatro series de precios de combustibles: **Super**, **Regular**, **Diesel** y **Kerosene**.

Matrices de Conteo C (calculadas según la Ecuación 4.6):

$$C_{\text{Super}} = \begin{bmatrix} 41 & 18 & 1 & 0 \\ 19 & 327 & 6 & 0 \\ 0 & 7 & 13 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad C_{\text{Regular}} = \begin{bmatrix} 43 & 13 & 1 & 0 \\ 14 & 334 & 6 & 0 \\ 0 & 7 & 14 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$C_{\text{Diesel}} = \begin{bmatrix} 62 & 19 & 1 & 0 \\ 20 & 304 & 5 & 0 \\ 0 & 6 & 11 & 2 \\ 0 & 1 & 1 & 0 \end{bmatrix} \quad C_{\text{Kerosene}} = \begin{bmatrix} 81 & 23 & 0 & 0 \\ 23 & 238 & 13 & 1 \\ 0 & 13 & 25 & 4 \\ 0 & 1 & 4 & 6 \end{bmatrix}$$

Se observa que en las matrices de conteo para **Super** y **Regular**, tanto la cuarta fila como la cuarta columna son vectores nulos. Esto indica que el estado s_4 (alza fuerte) es

un estado transitorio no observado en estas series de datos; nunca se transita desde ni hacia él. Para un análisis más robusto y enfocado en la dinámica real, se puede reducir el espacio de estados de 4x4 a 3x3, considerando únicamente los estados activos (s_1, s_2, s_3). Las submatrices de conteo resultantes son:

$$C_{\text{Super}}^{\text{reducida}} = \begin{bmatrix} 41 & 18 & 1 \\ 19 & 327 & 6 \\ 0 & 7 & 13 \end{bmatrix} \quad C_{\text{Regular}}^{\text{reducida}} = \begin{bmatrix} 43 & 13 & 1 \\ 14 & 334 & 6 \\ 0 & 7 & 14 \end{bmatrix}$$

Matrices de Transición (\hat{P}) (Normalizadas por columna, 3 decimales):

$$\hat{P}_{\text{Super}} = \begin{bmatrix} 0.683, 0.051, 0.050 \\ 0.317, 0.929, 0.300 \\ 0.000, 0.020, 0.650 \end{bmatrix} \quad \hat{P}_{\text{Regular}} = \begin{bmatrix} 0.754, 0.037, 0.048 \\ 0.246, 0.944, 0.286 \\ 0.000, 0.020, 0.667 \end{bmatrix}$$

$$\hat{P}_{\text{Diesel}} = \begin{bmatrix} 0.756, 0.058, 0.056, 0.000 \\ 0.244, 0.921, 0.278, 0.000 \\ 0.000, 0.018, 0.611, 1.000 \\ 0.000, 0.003, 0.056, 0.000 \end{bmatrix} \quad \hat{P}_{\text{Kerosene}} = \begin{bmatrix} 0.779, 0.084, 0.000, 0.000 \\ 0.221, 0.865, 0.310, 0.091 \\ 0.000, 0.047, 0.595, 0.364 \\ 0.000, 0.004, 0.095, 0.545 \end{bmatrix}$$

Matrices A_p Estimadas con SRep (3 decimales):

$$A_{p,\text{Super}} = \begin{bmatrix} 0.683, 0.051, 0.050 \\ 0.317, 0.929, 0.300 \\ 0.000, 0.020, 0.650 \end{bmatrix}$$

$$A_{p,\text{Regular}} = \begin{bmatrix} 0.754, 0.037, 0.048 \\ 0.246, 0.944, 0.286 \\ 0.000, 0.020, 0.667 \end{bmatrix}$$

$$A_{p,\text{Diesel}} = \begin{bmatrix} 0.756, 0.058, 0.056, 0.000 \\ 0.244, 0.921, 0.278, 0.000 \\ 0.000, 0.018, 0.611, 1.000 \\ 0.000, 0.003, 0.056, 0.000 \end{bmatrix}$$

$$A_{p,\text{Kerosene}} = \begin{bmatrix} 0.779, 0.084, 0.000, 0.000 \\ 0.221, 0.865, 0.310, 0.091 \\ 0.000, 0.047, 0.595, 0.364 \\ 0.000, 0.004, 0.095, 0.545 \end{bmatrix}$$

Diferencia $A_p - \hat{\mathbf{P}}$:

$$\Delta_{\text{Super}} = \begin{bmatrix} 0, 0, 0 \\ 0, 0, 0 \\ 0, 0, 0 \end{bmatrix} \quad \Delta_{\text{Regular}} = \begin{bmatrix} 0, 0, 0 \\ 0, 0, 0 \\ 0, 0, 0 \end{bmatrix}$$

$$\Delta_{\text{Diesel}} = \begin{bmatrix} 0, 0, 0, 0 \\ 0, 0, 0, 0 \\ 0, 0, 0, 0 \\ 0, 0, 0, 0 \end{bmatrix} \quad \Delta_{\text{Kerosene}} = \begin{bmatrix} 0, 0, 0, 0 \\ 0, 0, 0, 0 \\ 0, 0, 0, 0 \\ 0, 0, 0, 0 \end{bmatrix}$$

Observación (Interpretación de diferencias en la matriz de transición estimada). Anteriormente, para las series *Super* y *Regular*, al trabajar con matrices de transición de tamaño 4×4 , se observaba un valor exacto de 1 en la última columna de las matrices de diferencia $\Delta = A_p - \hat{\mathbf{P}}$.

Este valor reflejaba la ausencia de transiciones observadas saliendo del cuarto estado, que en esos casos no estaba presente en la muestra. La matriz $\hat{\mathbf{P}}$, basada en conteos observados, tenía columnas nulas para ese estado, mientras que la matriz estimada A_p , calculada con la función `SRep`, asignaba probabilidades que sumaban 1 para cumplir con las propiedades de matriz de transición de Markov.

Sin embargo, al reducir las matrices a tamaño 3×3 , excluyendo el cuarto estado no utilizado en estas series, dicho valor 1 desaparece de las diferencias, ya que el análisis y estimación se realizan únicamente sobre los estados realmente presentes.

Esta reducción mejora la interpretación y asegura que las matrices reflejen fielmente la dinámica observable de las series, eliminando artefactos generados por estados inexistentes o absorbentes en los datos.

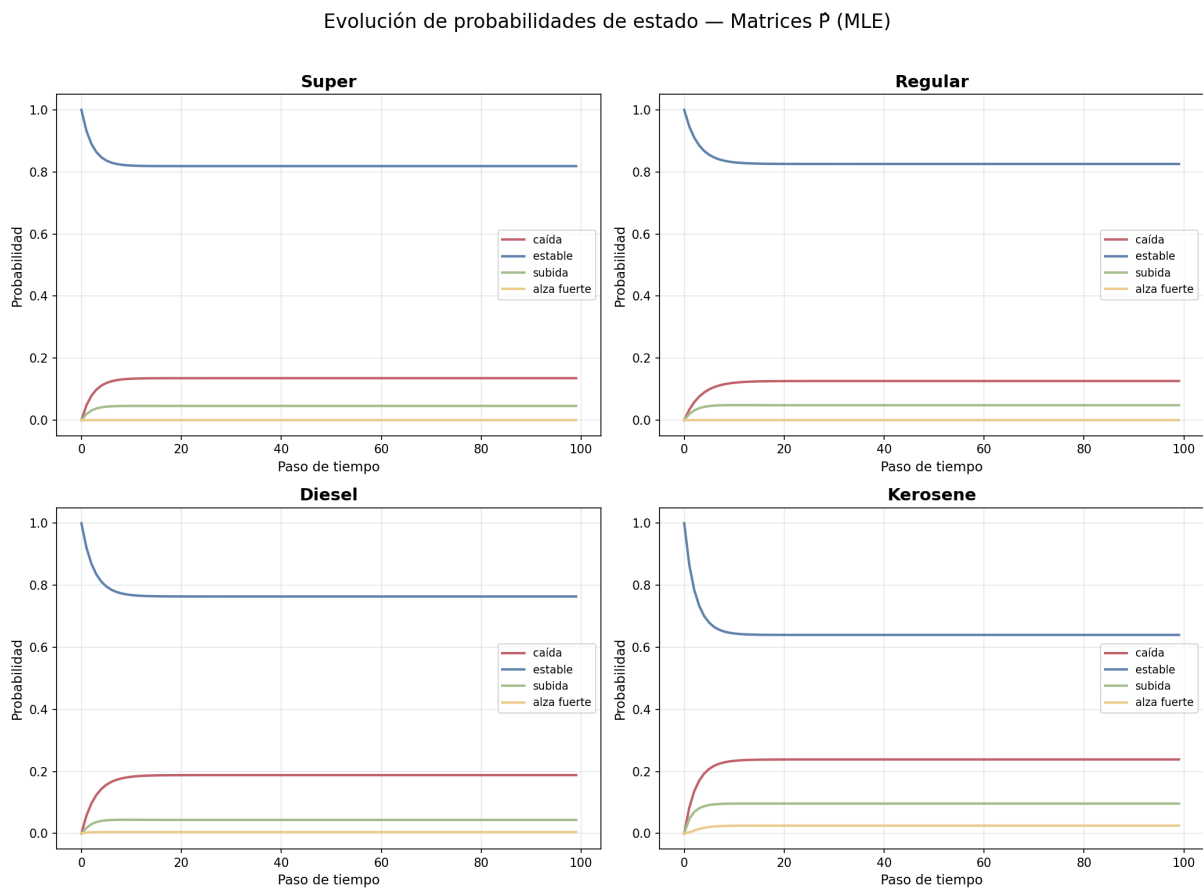


Figura 4.8: Evolución de las probabilidades de ocupación para cada estado según la matriz de transición estimada \hat{P} en las series de precios de combustibles: Super, Regular, Diesel y Kerosene.

La figura 4.8 muestra la evolución temporal de las probabilidades de pertenencia a cada estado de la matriz \hat{P} estimada para las series de combustibles: Super, Regular, Diesel y Kerosene. Se observa que todas las series alcanzan rápidamente un estado estable de probabilidades, con predominancia del estado “estable” (línea azul), que supera el 80 % en todos los casos. Las probabilidades de caída, subida y alza fuerte se mantienen bajas y constantes después de pocos pasos, lo que indica una dinámica de transición marcada por

una fuerte estabilidad en el corto plazo. Este comportamiento es consistente con mercados regulados o poco volátiles, donde los cambios abruptos son poco frecuentes.

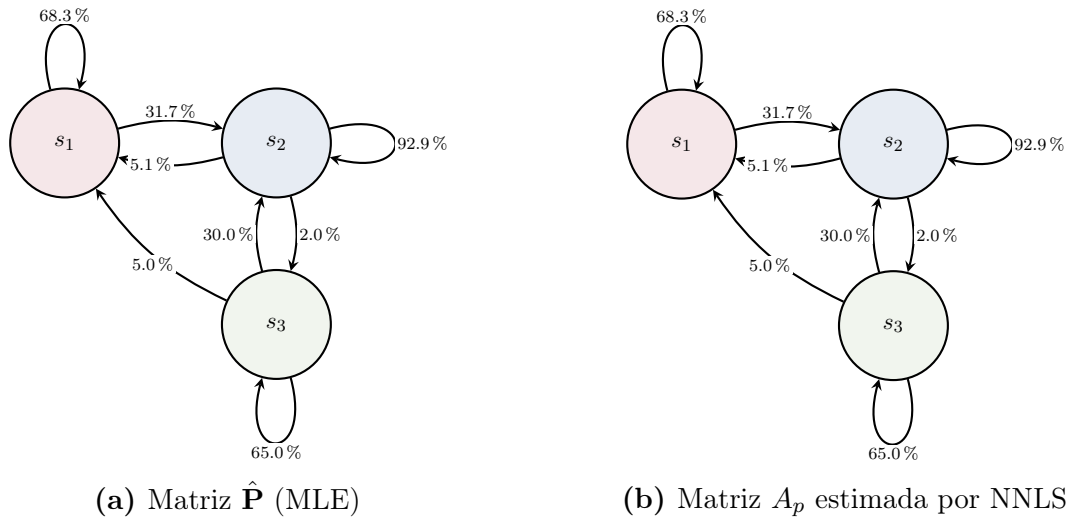


Figura 4.9: Comparación entre la matriz de transición normalizada $\hat{\mathbf{P}}$ y la matriz estimada A_p para la serie Super, sin considerar el estado s_4 que no está presente.

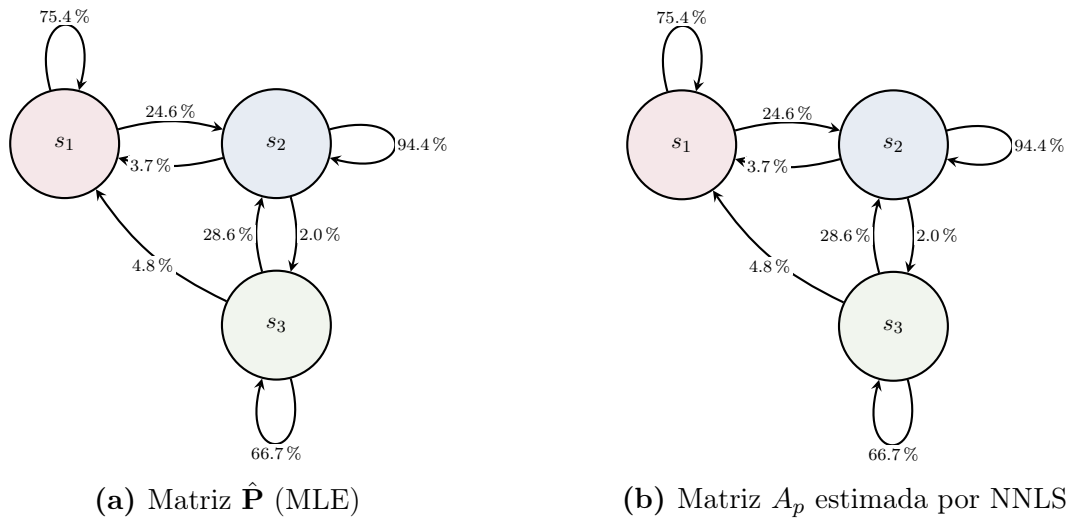


Figura 4.10: Comparación entre la matriz de transición normalizada $\hat{\mathbf{P}}$ y la matriz estimada A_p para la serie Regular, sin considerar el estado s_4 que no está presente.

Al comparar las dinámicas de transición de las cuatro series, se observa una clara jerarquía en su complejidad. Los modelos para **Super** y **Regular** exhiben las dinámicas más simples, caracterizadas por una fuerte persistencia en los estados y rutas de transición limitadas, con solo tres estados relevantes (s_1, s_2, s_3) debido a la ausencia del estado s_4 en los datos. El modelo para **Diesel** presenta una conectividad moderada, con algunas interacciones adicionales entre los regímenes. La serie **Kerosene** revela la dinámica más

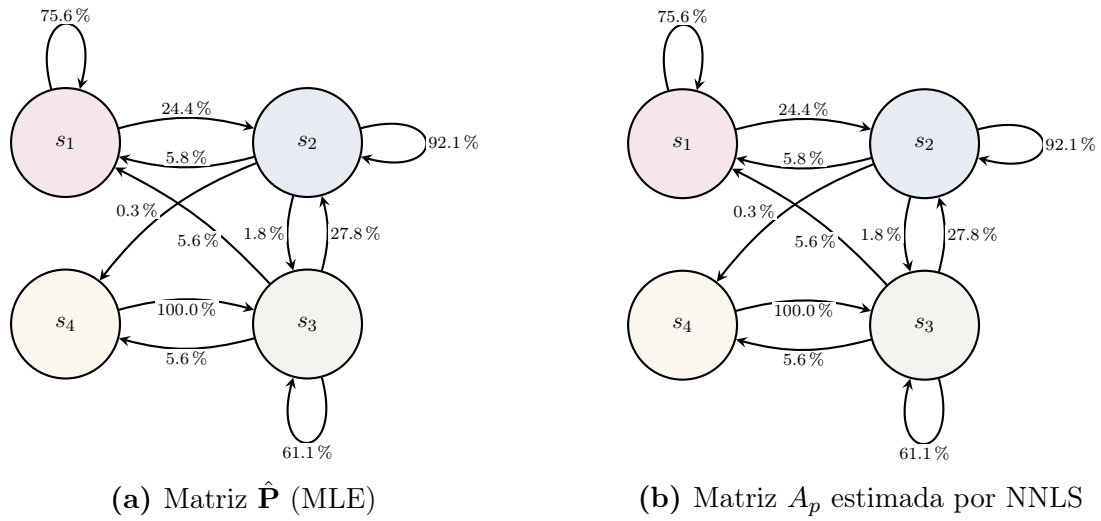


Figura 4.11: Comparación entre la matriz de transición normalizada $\hat{\mathbf{P}}$ y la matriz estimada A_p para la serie Diesel.

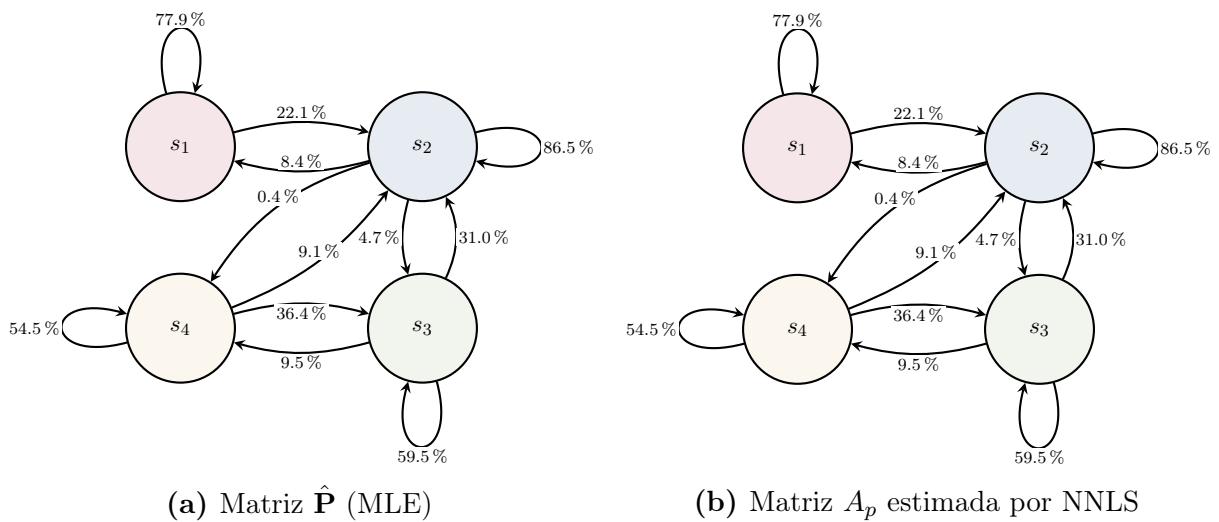


Figura 4.12: Comparación entre la matriz de transición normalizada $\hat{\mathbf{P}}$ y la matriz estimada A_p para la serie Kerosene.

diversa y compleja; en su grafo, todos los estados son persistentes y están altamente interconectados con múltiples vías de entrada y salida, lo que sugiere un sistema económico menos predecible y con mayor capacidad para alternar entre diferentes regímenes.

4.2.3. Análisis Predictivo y Robustez del Modelo para Combustibles

Para evaluar el rendimiento predictivo del modelo en las series de combustibles, se realizó un análisis exhaustivo con múltiples particiones de entrenamiento/prueba (desde 60/40 hasta 95/5).

Los resultados, resumidos en la Tabla 4.3, demuestran una alta capacidad predictiva y revelan una clara jerarquía en la dinámica de las series.

| Serie de Combustible | Precisión (Accuracy) Máxima |
|----------------------|-----------------------------|
| Regular | 94.2 % |
| Super | 93.0 % |
| Diesel | 86.1 % |
| Kerosene | 86.0 % |

Tabla 4.3: Resumen del rendimiento predictivo máximo para cada serie de combustible.

Las gasolinas **Super** y **Regular** resultaron ser altamente predecibles, alcanzando precisiones superiores al 90 %, lo que refleja su dinámica de precios estable. Por el contrario, el **Kerosene** y el **Diesel** mostraron ser más complejos y volátiles, con precisiones máximas en torno al 86 %.

Es notable que, aunque el estimador **Ap (SRep)** difiere teóricamente del estimador $\hat{\mathbf{P}}$ (**MLE**) para las series Super y Regular, su rendimiento predictivo fue idéntico en todas las pruebas. Esto se debe a que el estado anómalo (s_4) que causa la diferencia entre las matrices no apareció en los conjuntos de prueba, por lo que la distinción no tuvo un impacto práctico. Esta coincidencia no es casual: para una serie univariada larga, con estados bien poblados y sin penalizaciones añadidas, el estimador NNLS y el MLE de frecuencias relativas por columna coinciden, como sucede aquí. Su diferencia teórica solo se vuelve apreciable bajo regularización adicional o con muestras extremadamente cortas, casos en los que la restricción de no negatividad evita probabilidades espurias y atenúa el

sobreajuste.

A diferencia de la serie corta del PIB, en todas las series de combustibles se observó el comportamiento esperado: a mayor cantidad de datos de entrenamiento, la precisión del modelo consistentemente mejoró. Esto valida la robustez del modelo y su capacidad para aprender de series de tiempo más largas y estadísticamente estables.

4.3. Resumen del Algoritmo Metodológico

Para consolidar la metodología expuesta en este capítulo, el Algoritmo 3 presenta el flujo de trabajo computacional de forma unificada. Este resume, paso a paso, el proceso completo: desde la discretización de la serie temporal original con TCROC hasta la estimación y validación de las matrices de transición, sirviendo como una guía clara para la replicabilidad del método.

Algoritmo 3 Algoritmo para la Estimación y Validación de la Matriz de Transición

Entrada: Serie temporal original $\{v(t)\}_{t=0}^T$, parámetros de TCROC λ, W , y umbrales de estado.

Salida: Matriz de conteo C , Matriz de transición estimada $\hat{\mathbf{P}}$, Matriz estimada A_p , y Matriz de diferencia Δ .

- 1: **Fase 1: Discretización de la Serie Temporal**
- 2: Inicializar una lista vacía ‘alphas’.
- 3: **para** t desde W hasta T **hacer**
- 4: Calcular α_t usando la Ecuación (3.1).
- 5: Añadir α_t a la lista ‘alphas’.
- 6: **fin para**
- 7: Inicializar una lista vacía ‘estados’.
- 8: **para** cada α en ‘alphas’ **hacer**
- 9: Asignar un estado numérico $S_t \in \{1, 2, 3, 4\}$ según los umbrales definidos.
- 10: Añadir S_t a la lista ‘estados’.
- 11: **fin para**
- 12: **Fase 2: Construcción de Matrices de Datos**
- 13: Construir la matriz de estados actuales $X_0 \in \mathcal{M}_{4 \times (T-W)}(\mathbb{R})$ a partir de ‘estados’[: -1] en formato *one-hot*.
- 14: Construir la matriz de estados siguientes $X_1 \in \mathcal{M}_{4 \times (T-W)}(\mathbb{R})$ a partir de ‘estados’[1:] en formato *one-hot*.
- 15: **Fase 3: Estimación de la Matriz de Transición (MLE)**
- 16: Calcular la matriz de conteo de transiciones: $C \leftarrow X_1 X_0^T$.
- 17: Calcular la matriz de transición de máxima verosimilitud $\hat{\mathbf{P}}$ normalizando las columnas de C :

$$\hat{P}_{ij} \leftarrow \frac{C_{ij}}{\sum_{k=1}^4 C_{kj}}$$

- 18: **Fase 4: Validación del Estimador SRep**
- 19: Simular una secuencia de estados $p_0 \in \mathcal{M}_{4 \times M}(\mathbb{R})$ de longitud M usando la matriz $\hat{\mathbf{P}}$ como generador.
- 20: Estimar la matriz de transición A_p aplicando la función ‘SRep’ sobre los datos simulados p_0 :

$$A_p \leftarrow \text{SRep}(p_0, \text{ss})$$

- 21: **Fase 5: Comparación y Análisis**
 - 22: Calcular la matriz de diferencia: $\Delta \leftarrow A_p - \hat{\mathbf{P}}$.
 - 23: **devolver** $C, \hat{\mathbf{P}}, A_p, \Delta$.
-

4.4. Discusión de Limitaciones y Extensiones Futuras

Si bien la metodología ha demostrado ser robusta, conviene reconocer sus supuestos para contextualizar los resultados y proponer futuras líneas de investigación.

Limitaciones del Modelo Actual. Las principales limitaciones del enfoque se centran en dos supuestos clave. Por una parte, los regímenes se definen a través de **umbrales fijos** en la tasa α_t , los cuales son determinados exógenamente. Por otra parte, el modelo asume una **propiedad de Markov de primer orden**, donde el futuro solo depende del estado presente, ignorando posibles dependencias de más largo plazo que podrían existir en los datos.

Extensiones Futuras. A partir de estas limitaciones, se proponen las siguientes extensiones para trabajos futuros:

- **Umbrales Dinámicos o Adaptativos:** En lugar de umbrales fijos, se podrían explorar métodos de aprendizaje no supervisado (como clustering) para determinar los puntos de corte de los regímenes de forma endógena desde los datos.
- **Modelos de Markov de Orden Superior:** Investigar si un modelo de segundo orden (donde $P(S_t|S_{t-1}, S_{t-2})$) mejora la precisión predictiva, especialmente en series con dinámicas más complejas.
- **Extensión a Reservoir Computing No Lineal:** Generalizar el modelo actual hacia un *Stochastically Structured Reservoir Computer* (SSRC), donde la matriz \hat{P} es un caso particular de una matriz de acoplamiento W , e incorporar *embeddings* no lineales $\tilde{\sigma}_p$ para capturar dinámicas más complejas, como se sugiere en [7].

4.5. Resumen del Capítulo

Este capítulo ha presentado una metodología unificada y robusta, el modelo TCROC-Markov con estimación NNLS, para la discretización y el modelado de regímenes en series de tiempo económicas. El enfoque aborda dos desafíos clave en la literatura: genera estados económicos interpretables a partir de datos continuos y garantiza, por construcción, la validez estocástica de las matrices de transición estimadas. Esta estrategia híbrida, que combina componentes estadísticos con aprendizaje estructurado, se alinea con los hallazgos de la competencia M4 [35] y con resultados de pronóstico bajo cambio de régimen [17].

A través de la aplicación a datos del PIB per cápita de Honduras y precios de combustibles, se demostró la efectividad del método. En el caso del PIB, el modelo identificó correctamente períodos de estabilidad, crecimiento y contracciones agudas asociadas a crisis conocidas. Para las series de combustibles, el análisis reveló una clara jerarquía de predictibilidad, con las gasolinas mostrando una alta estabilidad y el kerosene una dinámica más compleja.

El análisis de validación predictiva reveló una dualidad en la aplicabilidad del modelo, ligada a las propiedades de la serie. Para series largas y estadísticamente estables como los combustibles, la metodología demostró alta precisión en el pronóstico a corto plazo. En contraste, para series cortas con eventos anómalos como la del PIB, el poder predictivo se reduce por la escasez de datos. Aun así, el modelo estimado capturó una dinámica estructural coherente, como lo confirma su rápida convergencia hacia una distribución estacionaria. La utilidad del modelo es doble: es una herramienta tanto para el **análisis estructural** de sistemas complejos como para la **predicción** de dinámicas recurrentes.

Capítulo 5

Detección y Pronóstico de Regímenes de Precios en Combustibles

5.1. Introducción

La volatilidad en los precios de los combustibles¹ dificulta la planificación económica en economías pequeñas y abiertas como la hondureña. Los enfoques clásicos de cambio de régimen (MS-AR) [4, 19, 27] ofrecen un marco teórico sólido, pero dependen de supuestos paramétricos fuertes y de una optimización de verosimilitud que puede resultar inestable con alta volatilidad y muestras moderadas.

Por otra parte, los Modelos Ocultos de Markov (HMM) ofrecen mayor flexibilidad al no requerir una forma paramétrica para las emisiones, pero sus estados latentes no poseen una interpretación económica directa por construcción: la correspondencia entre el estado “oculto” identificado y un régimen de mercado observable suele ser ambigua y requiere una validación posterior o análisis detallado de sus funciones de emisión, lo que dificulta su aplicación directa al diseño de políticas públicas.

Los Capítulos 2–4 construyeron, de forma progresiva, las herramientas teóricas que conforman el marco TCRA-Markov: la jerarquía de operadores y sus propiedades matemáticas, la integración con cadenas de Markov (ϕ -mixing, convergencia asintótica), y la estimación NNLS con prueba constructiva de existencia.

Este capítulo aplica ese aparato teórico a datos reales. Se investiga la sensibilidad del marco a los hiperparámetros (W, λ, K) , lo que incluye el descubrimiento de un fenómeno de degeneración para ventanas excesivamente grandes, y se identifican regímenes económicos interpretables mediante agrupamiento K-medias [34] sobre la señal α_t , evaluando como hipótesis su desempeño frente a alternativas tradicionales como la discretización por

¹El contenido central de este capítulo ha sido aceptado para publicación y presentación en IEEE CONCAPAN XLII (2025) bajo el título “*TCROC-Markov: A Two-Stage Framework for Economic Regime Analysis*”.

cuantiles fijos. El capítulo incluye también un análisis espectral de las matrices de transición estimadas que conecta las propiedades algebraicas de $\hat{\mathbf{P}}$ con la dinámica económica observada.

El estudio se aplica a series semanales de precios de cuatro combustibles en Honduras (Superior, Regular, Diésel y Kerosene) durante el período enero 2017, agosto 2025, con más de 400 observaciones por serie. Las trayectorias completas $\{S_t\}$, los diagramas de transición y las matrices $\hat{\mathbf{P}}$ estimadas para cada serie se presentan en el Apéndice A.3 (Figuras A.1, A.2, A.3 y A.4; Tablas A.1 y A.2).

La validación predictiva se implementó mediante un esquema de ventanas deslizantes (*rolling window*) con las siguientes especificaciones: el conjunto de entrenamiento inicial abarca las primeras 312 semanas (enero 2017, diciembre 2022), lo que representa un período de 6 años de historia económica local que permite capturar diversos comportamientos del mercado y proporciona una base de estimación estadísticamente representativa, a la vez que deja un período de evaluación fuera de muestra extenso (aproximadamente 2.6 años), el horizonte de pronóstico es de 1 semana, y la ventana avanza con paso unitario sobre el período fuera de muestra (enero 2023, agosto 2025), generando aproximadamente 130 evaluaciones secuenciales por serie. Este diseño preserva estrictamente la causalidad temporal: en ningún punto del proceso se utiliza información futura para estimar los parámetros del modelo.

5.2. Marco Metodológico Aplicado

El flujo de trabajo implementado sigue la arquitectura de dos etapas del marco TCRA-Markov desarrollado en los capítulos precedentes. Esta sección resume las decisiones de implementación específicas para la aplicación a datos de combustibles, remitiendo al lector a las demostraciones formales ya establecidas.

5.2.1. Etapa 1: Discretización Basada en Datos

El Operador TCRA como Extractor de Características

Para cada instante t , se calcula la tasa de cambio relativa α_t utilizando la variante ETCRAM, que combina ventana móvil W y decaimiento exponencial λ :

$$\beta_t = \frac{\sum_{k=t-W+1}^t \lambda^{t-k} v_k v_{k-1}}{\sum_{k=t-W+1}^t \lambda^{t-k} v_{k-1}^2}, \quad \alpha_t = \beta_t - 1, \quad (5.1)$$

donde v_t denota el precio del combustible en la semana t . Como se demostró rigurosamente, este estimador existe y es único bajo la condición $W > 1$ y no degeneración de los datos en la ventana. La interpretación de α_t como coeficiente OLS de un modelo AR(1) local sin intercepto proporciona una base estadística sólida y una lectura económica directa: $\alpha_t > 0$ indica crecimiento local, $\alpha_t < 0$ indica contracción, y $\alpha_t \approx 0$ indica estabilidad.

Discretización vía K-medias

A diferencia de la discretización por umbrales fijos empleada previamente para el análisis exploratorio del PIB, en esta sección se adopta un enfoque completamente basado en datos. El algoritmo K-medias particiona el conjunto $\{\alpha_W, \dots, \alpha_T\}$ en K clústeres minimizando la inercia intra-grupo:

$$\min_{\{C_1, \dots, C_K\}} \sum_{k=1}^K \sum_{\alpha \in C_k} \|\alpha - \mu_k\|^2, \quad (5.2)$$

donde μ_k es el centroide del clúster k . Cada α_t se asigna a su clúster correspondiente, generando la secuencia de estados discretos $\{S_t\}$. Los estados se ordenan según la magnitud de sus centroides ($\mu_1 < \mu_2 < \dots < \mu_K$), de modo que s_1 corresponde siempre al régimen de mayor caída y s_K al de mayor alza. Este orden ascendente invierte la convención pedagógica del Capítulo 3, donde s_1 designaba el crecimiento más fuerte. La diferencia es puramente notacional, fruto de que K-medias ordena los centroides de menor a mayor; las probabilidades de transición y los resultados no cambian bajo ninguna de las dos etiquetas.

Observación (Justificación de K-medias sobre umbrales fijos). Los umbrales fijos (e.g., $\alpha < -0.01$ para “contracción”) adolecen de dos limitaciones fundamentales: (i) son exógenos al proceso generador de los datos, lo que puede sesgar la identificación de regímenes, y (ii) imponen el mismo esquema de partición a series con dinámicas potencialmente distintas. K-medias resuelve ambos problemas al determinar los puntos de corte endógenamente, adaptándose a la distribución empírica de cada serie.

Observación (Selección del número de estados K). Matemáticamente, el índice de Davies-Bouldin (DB) evalúa la calidad del agrupamiento midiendo la dispersión de cada clúster frente a la distancia entre sus centroides. Por diseño, valores **menores** de DB indican agrupamientos más compactos y mejor separados espacialmente, representando particiones óptimas. Sin embargo, en el análisis empírico de variables económicas reguladas, la selección del número de regímenes K^* no se limita al mínimo matemático absoluto de DB. Se adopta un enfoque multicriterio que armoniza la separabilidad de los clústeres, la tasa de reducción de la inercia (método del codo) y la interpretabilidad semántica de los estados resultantes para el diseño de políticas públicas.

Por ejemplo, para la serie Regular, aunque $K = 4$ o $K = 6$ reportan índices de DB ligeramente inferiores, la ganancia en la granularidad no justifica la pérdida de parsimonia en la estimación de la matriz de transición; por ello se adopta $K^* = 3$ (caída, estabilidad y alza), que coincide con un codo marcado de inercia y con una interpretación simple del mercado. En contraste, para Kerosene se selecciona $K^* = 5$ a pesar de que $K = 4$ posea un DB menor (0.4326), debido a la necesidad analítica de aislar el estado de ‘Alza extrema’ ($\mu_5 = +0.0172$), el cual captura shocks de precios excepcionales ausentes en otros combustibles. La Tabla 5.1 detalla los valores numéricos de DB e inercia que fundamentan el balance final para cada serie, donde los valores en negrita corresponden al número óptimo de estados K^* seleccionado.

Observación (Verificación de estacionariedad de α_t). Antes de aplicar el algoritmo de discretización, se evaluó formalmente la estacionariedad de las series $\{\alpha_t\}$ utilizando dos pruebas complementarias cuyos resultados se detallan en la Tabla 5.2, la cual resume los estadísticos ADF (H_0 : raíz unitaria) y KPSS (H_0 : estacionariedad) con sus respectivos p -valores. Para las series Regular y Diésel, el test de Dickey-Fuller aumentado (ADF)

Tabla 5.1: Índice de Davies-Bouldin e inercia intra-clúster para diferentes números de estados K .

| Serie | K | Davies-Bouldin | Inercia |
|----------|----------|----------------|---------------|
| Regular | 2 | 0.5440 | 0.0014 |
| | 3 | 0.5662 | 0.0007 |
| | 4 | 0.4889 | 0.0005 |
| | 5 | 0.5008 | 0.0003 |
| | 6 | 0.4470 | 0.0002 |
| | 7 | 0.4388 | 0.0001 |
| | 8 | 0.4646 | 0.0001 |
| | 9 | 0.4755 | 0.0001 |
| | Superior | 2 | 0.5692 |
| 3 | | 0.5514 | 0.0007 |
| 4 | | 0.5456 | 0.0004 |
| 5 | | 0.5204 | 0.0003 |
| 6 | | 0.5104 | 0.0002 |
| 7 | | 0.4947 | 0.0001 |
| 8 | | 0.4845 | 0.0001 |
| 9 | | 0.4322 | 0.0001 |
| Diésel | | 2 | 0.5535 |
| | 3 | 0.5281 | 0.0016 |
| | 4 | 0.5508 | 0.0009 |
| | 5 | 0.5250 | 0.0006 |
| | 6 | 0.4911 | 0.0004 |
| | 7 | 0.4832 | 0.0003 |
| | 8 | 0.4760 | 0.0002 |
| | 9 | 0.4840 | 0.0002 |
| | Kerosene | 2 | 0.5709 |
| 3 | | 0.5043 | 0.0032 |
| 4 | | 0.4326 | 0.0014 |
| 5 | | 0.4622 | 0.0010 |
| 6 | | 0.4982 | 0.0007 |
| 7 | | 0.5374 | 0.0005 |
| 8 | | 0.5155 | 0.0004 |
| 9 | | 0.4601 | 0.0003 |

rechaza formalmente la presencia de raíz unitaria al nivel del 5 % ($p = 0.041$ y $p = 0.048$, respectivamente). En las series Superior y Kerosene, los p -valores del ADF se sitúan por encima del umbral convencional ($p = 0.136$ y $p = 0.131$, respectivamente), por lo que el test ADF por sí solo no resulta concluyente para rechazar la no-estacionariedad de estas dos variables debido a las limitaciones de poder del test en muestras con alta persistencia local.

Sin embargo, al aplicar el test KPSS, cuya hipótesis nula H_0 postula la estacionariedad de la serie, ninguna de las cuatro variables rechaza la hipótesis de estacionariedad ($p > 0.10$ en todos los casos, con estadísticos KPSS marcadamente inferiores al valor crítico del 5 % de 0.463). Esta discrepancia parcial entre los tests ADF y KPSS para Superior y Kerosene es un comportamiento conocido en econometría de series financieras con volatilidad moderada (clasificada a menudo como persistencia fraccional o estacionariedad débil). La conjunción de un test KPSS que no rechaza la estacionariedad y la ausencia de raíces unitarias explosivas justifica de manera consistente el tratamiento de $\{\alpha_t\}$ como un proceso de covarianza estacionaria apto para el modelado de Markov, confirmando que el operador local actúa de facto como un filtro estabilizador de la tendencia de precios. En síntesis, para Regular y Diésel ambos tests coinciden en la estacionariedad; para Superior y Kerosene la conclusión descansa principalmente en el KPSS, ante un ADF no concluyente, lo que basta para tratar las cuatro series como estacionarias en sentido débil.

Tabla 5.2: Tests de estacionariedad sobre las series $\{\alpha_t\}$ de cada combustible.

| Serie | ADF | p | Lags | Crit. 5% | KPSS | p | Decisión |
|----------|--------|-------|------|----------|-------|--------|---------------------------|
| Regular | -2.938 | 0.041 | 9 | -2.869 | 0.218 | > 0.10 | Estacionaria |
| Superior | -2.420 | 0.136 | 9 | -2.869 | 0.250 | > 0.10 | Estacionaria [†] |
| Diésel | -2.879 | 0.048 | 1 | -2.869 | 0.269 | > 0.10 | Estacionaria |
| Kerosene | -2.438 | 0.131 | 12 | -2.869 | 0.262 | > 0.10 | Estacionaria [†] |

[†] ADF no significativo al 5 %, pero KPSS confirma estacionariedad ($p > 0.10$, estadístico $\ll 0.463$ = valor crítico al 5 %). En estas dos series la estacionariedad se concluye con base en el KPSS, dado que el ADF no es concluyente.

5.2.2. Etapa 2: Estimación de la Matriz de Transición

Una vez obtenida la secuencia $\{S_t\}$, se modela como cadena de Markov de primer orden. La estimación de la matriz de transición $\hat{\mathbf{P}}$ sigue el procedimiento NNLS desarrollado empíricamente en este documento. Se resalta aquí el esquema matricial.

A partir de la codificación indicadora (*one-hot*) de los estados, se construyen las matrices \mathbf{X}_0 (estados actuales) y \mathbf{X}_1 (estados siguientes). La linealización vía producto de Kronecker transforma el problema matricial

$$\hat{\mathbf{P}} = \arg \min_{\mathbf{P} \in \mathbb{S}(K,K)} \|\mathbf{X}_1 - \mathbf{P}\mathbf{X}_0\|_F^2 \quad (5.3)$$

en el sistema lineal $\text{vec}(\mathbf{X}_1) = (\mathbf{X}_0^T \otimes \mathbf{I}_K) \text{vec}(\mathbf{P})$, que se resuelve mediante el problema aumentado de NNLS:

$$\hat{\mathbf{c}} = \arg \min_{\mathbf{c} \geq \mathbf{0}} \left\| \begin{bmatrix} \mathbf{S}_0 \\ w\mathbf{C}_{eq} \end{bmatrix} \mathbf{c} - \begin{bmatrix} \mathbf{b} \\ w\mathbf{1}_K \end{bmatrix} \right\|_2^2, \quad (5.4)$$

unitaria por columnas. Como se ha establecido de forma matemática en este tratado, este problema es convexo, su conjunto factible es no vacío, y la solución existe y es única cuando \mathbf{S}_0 tiene rango completo por columnas, una condición que requiere no solo que cada estado de origen aparezca al menos una vez con transiciones salientes en la secuencia de entrenamiento, sino también que la matriz de diseño posea un rango adecuado para garantizar la proyección unívoca en el ortante positivo. Tras la solución NNLS, se aplica una normalización final por columnas para corregir desviaciones de precisión de punto flotante.

5.2.3. Protocolo de Optimización de Hiperparámetros

La selección de los hiperparámetros (W, λ, K) representa un reto clave de optimización en este trabajo. Se implementó una búsqueda exhaustiva en grilla sobre el dominio:

$$W \in \{2, 3, \dots, 52\}, \quad \lambda \in \{0.80, 0.81, \dots, 1.00\}, \quad K \in \{2, 3, \dots, 9\}. \quad (5.5)$$

Es pertinente precisar que, aunque la búsqueda exhaustiva en grilla se considera ineficiente en espacios de alta dimensionalidad, en este problema el dominio discreto está estrictamente delimitado por la restricción $W \leq 52$ (establecida constructivamente para evitar la degeneración del estimador conforme al Teorema 5.1 y alineada con el ciclo anual natural del mercado) y pasos de centésimas para λ . Esto resulta en un espacio de búsqueda acotado (menor a 3 000 combinaciones por serie) cuya evaluación secuencial tarda escasos milisegundos por iteración. Este enfoque exhaustivo es óptimo para este contexto, ya que garantiza la identificación de la combinación de parámetros que minimiza globalmente el error predictivo sin los riesgos de estancamiento en mínimos locales inherentes a los métodos metaheurísticos.

Para cada combinación de parámetros en la grilla, el desempeño del modelo se evaluó fuera de muestra mediante un protocolo de ventanas deslizantes (*rolling window*) de paso unitario. La selección de la combinación óptima final se rige por el siguiente criterio jerárquico:

1. **Criterio primario (RMSE):** Se selecciona la combinación que minimiza el error cuadrático medio (RMSE) de pronóstico puntual de precios fuera de muestra. La primacía de esta métrica responde a la naturaleza aplicada del modelo: en el mercado energético regulado, la exactitud numérica de la predicción de precios representa el objetivo crítico.
2. **Criterio secundario (Exactitud):** En caso de empates en el RMSE, se maximiza la tasa de exactitud cualitativa en la predicción del régimen económico futuro.
3. **Desempate y Regularización (AIC):** Para precaver el sobreajuste que podría surgir al optimizar exclusivamente el RMSE (lo que favorecería modelos sobreparametrizados con un número excesivo de estados K), se penaliza la complejidad del modelo a través del Criterio de Información de Akaike (AIC):

$$AIC(K) = 2K(K - 1) - 2\mathcal{L}_K, \quad (5.6)$$

donde $\mathcal{L}_K = \sum_{i,j} C_{ij} \ln \hat{P}_{ij}$ representa la log-verosimilitud del modelo con K regíme-

nes, y el término $K(K - 1)$ denota la cantidad de parámetros libres e independientes presentes en una matriz estocástica de tamaño $K \times K$.

El costo computacional total de la búsqueda escala como $O(|\mathcal{W}| |\Lambda| |\mathcal{K}| (T + TKI + K^3))$, donde I es el número de iteraciones de K-medias. Con $K \leq 9$ y $T \approx 400$, el tiempo de ejecución es dominado por el tamaño de la grilla y no por los algoritmos internos, permitiendo paralelización directa.

5.3. Resultados y Análisis

5.3.1. El Fenómeno de Degeneración del Modelo

El análisis preliminar reveló un hallazgo crítico sobre la estabilidad estructural del marco: una búsqueda de hiperparámetros sin restricciones temporales conduce inevitablemente a la *degeneración del modelo*, produciendo métricas aparentemente favorables, asociadas al sobre-suavizado y la pérdida de variabilidad. Cuando la ventana de inercia toma valores excesivamente grandes ($W \gg 100$), el operador local acumula un volumen de historia desproporcionado, lo que provoca que la serie de tasas de cambio α_t se aplane de manera artificial hasta perder toda variabilidad informativa. En este límite, la tasa local converge de manera asintótica hacia el promedio global de crecimiento de la muestra entera. En consecuencia, el operador se desensibiliza frente a cambios marginales instantáneos y el modelo tiende asintóticamente al comportamiento rígido de un modelo autorregresivo lineal global (o una caminata aleatoria con deriva constante en los niveles de precios), perdiendo su capacidad para caracterizar transiciones dinámicas rápidas.

Es fundamental enfatizar que el establecimiento de una ventana local acotada ($W \leq 52$) no implica que el modelo descarte la memoria histórica a largo plazo de la serie. Por el contrario, mientras que el operador α_t se evalúa localmente para capturar la tasa marginal instantánea (eliminando tendencias estocásticas y raíces unitarias en los niveles de precios), la matriz de transición $\hat{\mathbf{P}}$ y las dinámicas estocásticas de Markov se estiman utilizando la serie histórica completa de observaciones T . De esta forma, el modelo integrado preserva la historia temporal a través de las frecuencias de transición conjuntas y la distribución

estacionaria, pero opera sobre una variable filtrada libre de tendencia, logrando modelar no-linealidades y asimetrías que las caminatas aleatorias convencionales no pueden capturar.

Este fenómeno admite una explicación formal. Considérese el comportamiento de la varianza de α_t como función de W .

Teorema 5.1 (Degeneración por sobre-suavizado). *Sea $\{v_t\}_{t=1}^T$ una serie temporal con $v_t > 0$ para todo t , y sea $\alpha_t(W)$ el operador ETCRAM con $\lambda = 1$ y ventana W . Entonces:*

1. *Para $W = T$, el operador es constante: $\alpha_t(T) = \bar{\alpha}$ para todo $t \geq T$, donde*

$$\bar{\alpha} = -1 + \frac{\sum_{t=1}^T v_t v_{t-1}}{\sum_{t=1}^T v_{t-1}^2}. \quad (5.7)$$

2. *En consecuencia, $\text{Var}(\alpha_t(T)) = 0$, y para cualquier $K \geq 2$, el algoritmo K -medias asigna todas las observaciones a un único clúster efectivo ($K_{\text{eff}} = 1$).*
3. *La matriz de transición resultante es $\hat{\mathbf{P}} = \mathbf{I}_K$, y el modelo predice perpetuamente el estado inicial.*

Demostración. El punto (1) es inmediato: cuando $W = T$, las sumatorias en (5.1) abarcan toda la muestra y no dependen de t , por lo que β_t es constante. Para (2), si $\alpha_t = \bar{\alpha}$ para todo t , entonces $\|\alpha_t - \mu_k\| = 0$ para el clúster que contiene $\bar{\alpha}$, y todos los α_t son asignados al mismo clúster. Para (3), si $S_t = s_j$ para todo t , entonces $\mathbf{X}_0 = \mathbf{X}_1$ y la solución de (5.3) es $\hat{\mathbf{P}} = \mathbf{I}_K$ (o más precisamente, la columna j es \mathbf{e}_j y las restantes columnas son indeterminadas por ausencia de datos, pero la normalización NNLS las fuerza a vectores unitarios). \square

Para evitar esta degeneración, se impuso el límite $W \leq 52$ semanas, justificado por dos argumentos convergentes: (a) preserva la variabilidad necesaria de α_t para que K -medias identifique regímenes distintos, y (b) corresponde al ciclo anual natural del mercado de combustibles, más allá del cual la “memoria” del operador pierde relevancia económica.

5.3.2. Análisis de Sensibilidad

El análisis de sensibilidad dentro del dominio restringido $W \leq 52$ (Figura 5.1) evalúa el mejor desempeño alcanzable (mínimo RMSE) sobre todas las combinaciones de (λ, K)

para cada valor de W .

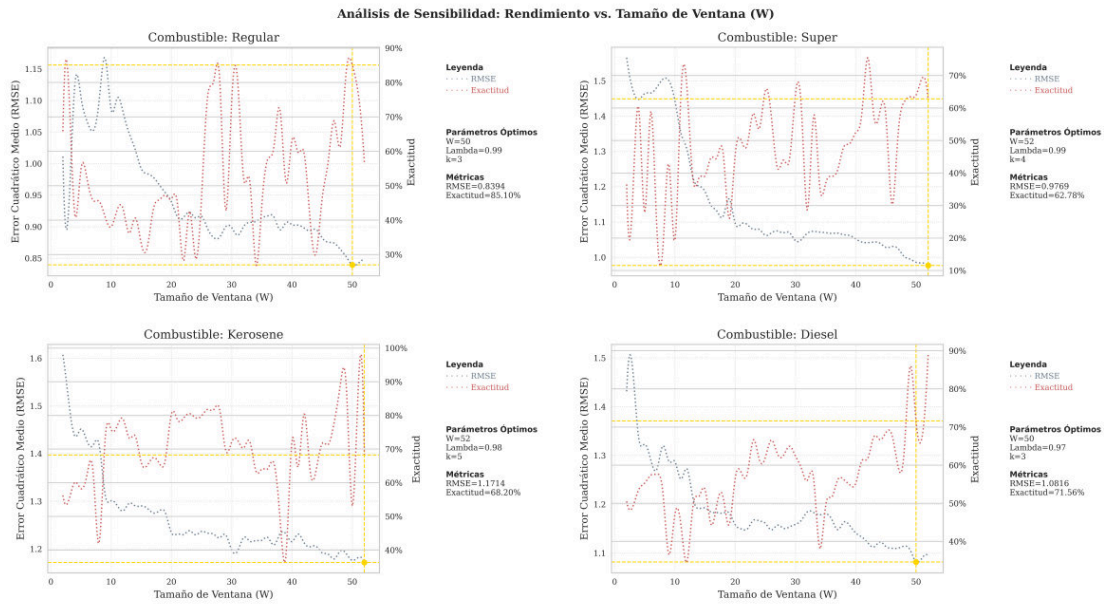


Figura 5.1: Análisis de sensibilidad del desempeño respecto a la ventana W .

Como se observa en dicha figura, para cada valor de W se muestra el modelo con menor RMSE (izquierda) y mayor exactitud (derecha), donde el círculo marca el óptimo global. Se identifica una tendencia monótonica en la cual ventanas más largas, próximas a 52 semanas, producen un menor error de pronóstico en todas las series. Esto indica que la predictibilidad de los precios de combustibles hondureños está dominada por tendencias de largo plazo (semianuales a anuales) más que por la volatilidad semanal de corto plazo.

Este resultado empírico tiene una interpretación económica directa: el mercado de combustibles en Honduras opera bajo un esquema de precios regulados con ajustes periódicos, lo que suaviza las fluctuaciones de corto plazo y concentra la dinámica informativa en componentes de frecuencia baja. El operador α_t con $W \approx 52$ captura eficazmente estos ciclos anuales, sin llegar al régimen degenerado de sobre-suavizado, filtrando el ruido semanal.

5.3.3. Hiperparámetros Óptimos

Aplicando el protocolo jerárquico definido en la Sección 5.2.3, se identificaron las configuraciones óptimas para cada serie. La Tabla 5.3 resume los resultados.

El espacio de optimización consolidado representado en la Figura 5.2 muestra las

Tabla 5.3: Combinación óptima de hiperparámetros por serie de combustible.

| Serie | W^* | λ^* | K^* | Exactitud | RMSE | AIC |
|----------|-------|-------------|-------|-----------|--------|--------|
| Regular | 50 | 0.99 | 3 | 85.10 % | 0.8394 | 190.87 |
| Superior | 52 | 0.99 | 4 | 62.78 % | 0.9769 | 272.80 |
| Diésel | 50 | 0.97 | 3 | 71.56 % | 1.0816 | 175.48 |
| Kerosene | 52 | 0.98 | 5 | 68.20 % | 1.1714 | 225.79 |

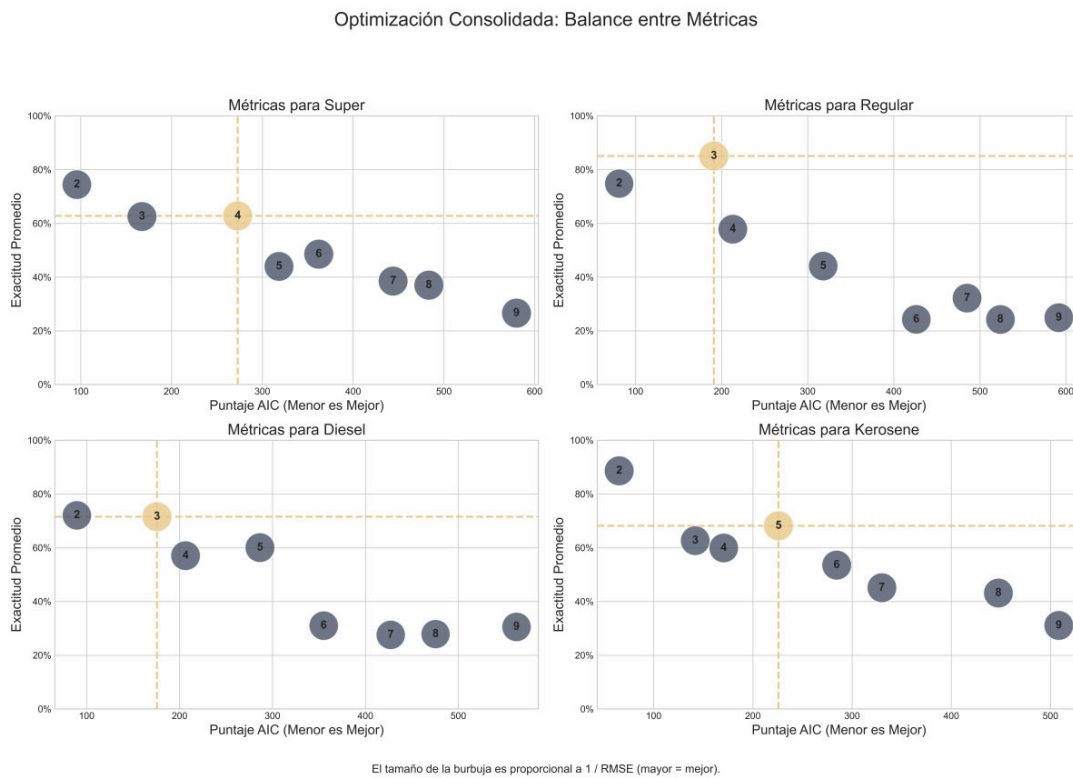


Figura 5.2: Espacio de optimización de hiperparámetros consolidado por número de regímenes (K).

alternativas paramétricas para cada valor de K mediante burbujas. Se observa que, en términos puramente matemáticos, las configuraciones con $K = 2$ estados reportan con frecuencia exactitudes sumamente elevadas y valores de AIC bajos. Este comportamiento es algebraicamente predecible: reducir el sistema a una clasificación binaria de dos estados (por ejemplo, 'Alza' y 'Caída') simplifica drásticamente el espacio de estados y reduce los parámetros libres en la matriz de transición a solo 2, minimizando la penalización del AIC y estabilizando el estimador.

Sin embargo, desde una perspectiva de aplicación económica y diseño de políticas energéticas, un modelo binario de $K = 2$ carece de utilidad analítica práctica, dado que agrupa de forma indiferenciada variaciones marginales estables y fluctuaciones de crisis extremas. Por ello, la elección óptima de $K^* \geq 3$ sacrifica parsimonia matemática en favor de una granularidad semántica indispensable para distinguir entre estabilidad y shocks de variaciones de precios de combustibles. Por ejemplo, para Regular ($K^* = 3$), el punto seleccionado domina con una burbuja de exactitud del 85.1 % y un AIC balanceado. Para Kerosene ($K^* = 5$), la mayor dispersión y el estado atípico de alza extrema obligan a adoptar un modelo de cinco estados, representando el balance óptimo bajo el criterio de selección jerárquico.

Se observan dos patrones transversales de interés. Por una parte, los valores de λ^* se concentran firmemente en el intervalo $[0.97, 0.99]$. Es importante destacar que el decaimiento de ponderación local no se fija de forma predeterminada ni restringida en la unidad ($\lambda = 1.0$), sino que se evalúa de manera dinámica en la búsqueda en grilla. Si bien en series con dinámicas reguladas estables el óptimo resulta cercano a la ponderación uniforme (confirmando que observaciones recientes e históricas dentro del año aportan información estructural equivalente para suavizar el ruido semanal), el marco teórico conserva la flexibilidad de adoptar valores menores de λ en series con fluctuaciones de alta frecuencia. Por otra parte, el número óptimo de regímenes revela una jerarquía clara de complejidad de mercado:

- **Regular y Diésel** ($K^* = 3$): Mercados de dinámica simple con tres estados bien diferenciados.

- **Superior** ($K^* = 4$): Complejidad intermedia con un estado adicional de alza moderada.
- **Kerosene** ($K^* = 5$): Mercado más complejo y fragmentado, requiriendo cinco estados para capturar su dinámica.

La relación complejidad-desempeño y la jerarquía de predictibilidad resultante se sintetizan en la Figura 5.3.

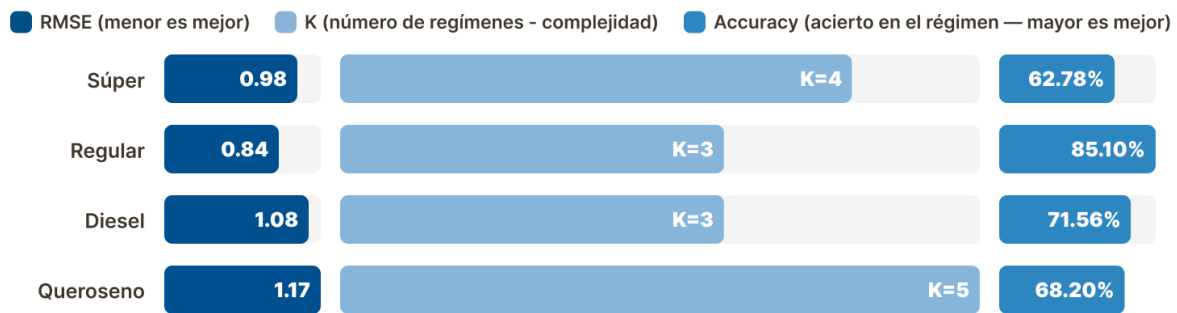


Figura 5.3: Jerarquía de predictibilidad de los combustibles.

A partir de estos resultados, se observa que la gasolina Regular es la serie más previsible con el menor RMSE y mayor exactitud. Por el contrario, el Kerosene se presenta como el caso más desafiante, requiriendo más estados porque su estructura empírica es más heterogénea. La correlación positiva observada entre K^* y el RMSE sugiere que la complejidad estructural inherente al mercado, y no una deficiencia del modelo, es lo que limita la máxima precisión alcanzable.

5.3.4. Identificación e Interpretación Económica de los Regímenes

Con los hiperparámetros óptimos fijados, el operador α_t genera distribuciones de tasas de cambio altamente concentradas alrededor de cero, consistentes con un mercado regulado donde los ajustes semanales drásticos son infrecuentes. La asignación temporal de regímenes resultante del agrupamiento K-medias se ilustra en la Figura 5.4.

En este panel de resultados, cada punto representa una semana y su color el régimen asignado sobre la serie α_t , mientras que las líneas punteadas horizontales marcan los centroides de cada clúster. Se observan transiciones coherentes que corresponden a eventos

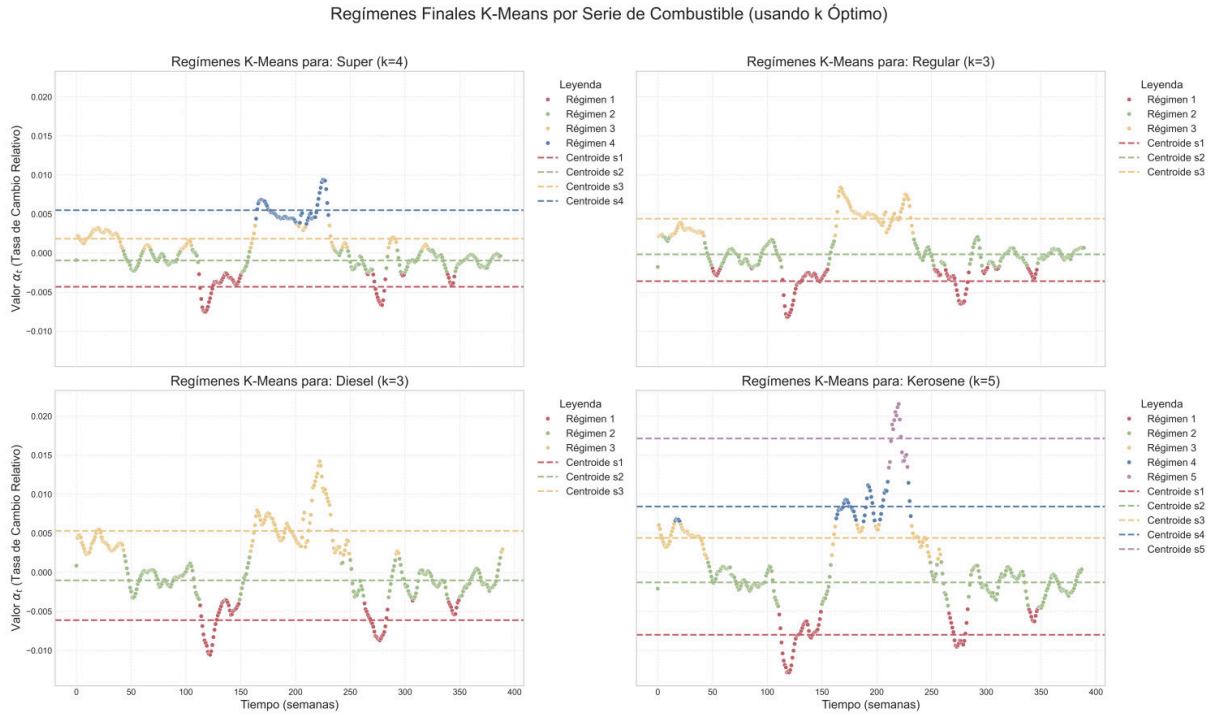


Figura 5.4: Regímenes identificados por K-medias sobre la serie α_t .

de mercado específicos, como el período 2020–2021 relacionado con el impacto y la recuperación pospandemia. Esta segmentación resulta en distribuciones estacionarias bien definidas que facilitan la interpretación de la dinámica del mercado.

Los centroides de K-medias proporcionan una interpretación económica directa de cada régimen. La Tabla 5.4 presenta los centroides y la prevalencia temporal de cada estado para las cuatro series.

En la Tabla 5.4 se detallan los centroides (μ_k) calculados mediante K-medias y su prevalencia temporal respectiva. Es oportuno enfatizar que las etiquetas semánticas asignadas a cada estado (tales como 'Alza moderada' o 'Caída fuerte') no obedecen a un criterio cualitativo subjetivo o discrecional del investigador. Por el contrario, son el reflejo directo de la magnitud y signo de los centroides de clúster estimados endógenamente por el algoritmo a partir de la muestra de tasas α_t . De este modo, centroides muy cercanos a cero ($|\mu_k| < 0.002$) se clasifican de forma unívoca como estabilidad, mientras que desviaciones positivas o negativas definen las distintas intensidades de alzas y caídas de precios regulados. El centroide representa la tasa de cambio semanal promedio que caracteriza a cada régimen, mientras que la prevalencia indica la proporción del tiempo total analizado, entre los años 2017 y 2025, en que el mercado se mantuvo en dicho estado operativo.

Tabla 5.4: Centroides (μ_k) y prevalencia temporal de los regímenes identificados.

| Serie | Régimen | Centroide (μ_k) | Prevalencia |
|------------------------|-----------------------|-----------------------|-------------|
| Regular ($K^* = 3$) | s_1 : Caída fuerte | -0.0036 | 24.6 % |
| | s_2 : Estabilidad | -0.0001 | 47.9 % |
| | s_3 : Alza fuerte | +0.0044 | 27.4 % |
| Superior ($K^* = 4$) | s_1 : Caída fuerte | -0.0043 | 15.7 % |
| | s_2 : Estabilidad | -0.0009 | 44.6 % |
| | s_3 : Alza moderada | +0.0019 | 23.7 % |
| | s_4 : Alza fuerte | +0.0055 | 16.0 % |
| Diésel ($K^* = 3$) | s_1 : Caída fuerte | -0.0061 | 17.4 % |
| | s_2 : Estabilidad | -0.0010 | 47.4 % |
| | s_3 : Alza fuerte | +0.0053 | 35.1 % |
| Kerosene ($K^* = 5$) | s_1 : Caída fuerte | -0.0080 | 17.0 % |
| | s_2 : Estabilidad | -0.0013 | 47.4 % |
| | s_3 : Alza moderada | +0.0044 | 16.0 % |
| | s_4 : Alza fuerte | +0.0084 | 15.0 % |
| | s_5 : Alza extrema | +0.0172 | 4.6 % |

Los cuatro mercados comparten un régimen dominante de *Estabilidad* que ocupa aproximadamente el 45–48 % del período analizado, con centroides cercanos a cero ($|\mu_k| < 0.002$). Esto es consistente con el esquema de regulación de precios vigente en Honduras, donde los ajustes semanales son típicamente pequeños. Por su parte, los mercados de Regular y Diésel exhiben una estructura simétrica de tres estados (caída–estabilidad–alza), mientras que Superior y Kerosene requieren estados intermedios adicionales para capturar su dinámica. Finalmente, el Kerosene presenta un estado excepcional de *Alza extrema* ($\mu_5 = +0.0172$, prevalencia 4.6 %), que no tiene análogo en los otros combustibles. Este régimen raro pero significativo es consistente con la naturaleza del Kerosene como subproducto de refinación con menor cobertura de mecanismos de amortiguación de precios, lo cual sugiere una menor amortiguación relativa, aunque se requerirían variables regulatorias adicionales para confirmarlo formalmente.

5.3.5. Dinámica de Transición

La estructura estocástica del mercado se captura en las matrices de transición $\hat{\mathbf{P}}$ estimadas vía NNLS, cuyos grafos de transición asociados se presentan en la Figura 5.5.

En estas representaciones, el valor dentro de cada nodo corresponde al centroide del clúster K-medias, permitiendo cuantificar el comportamiento semanal del precio. Los lazos

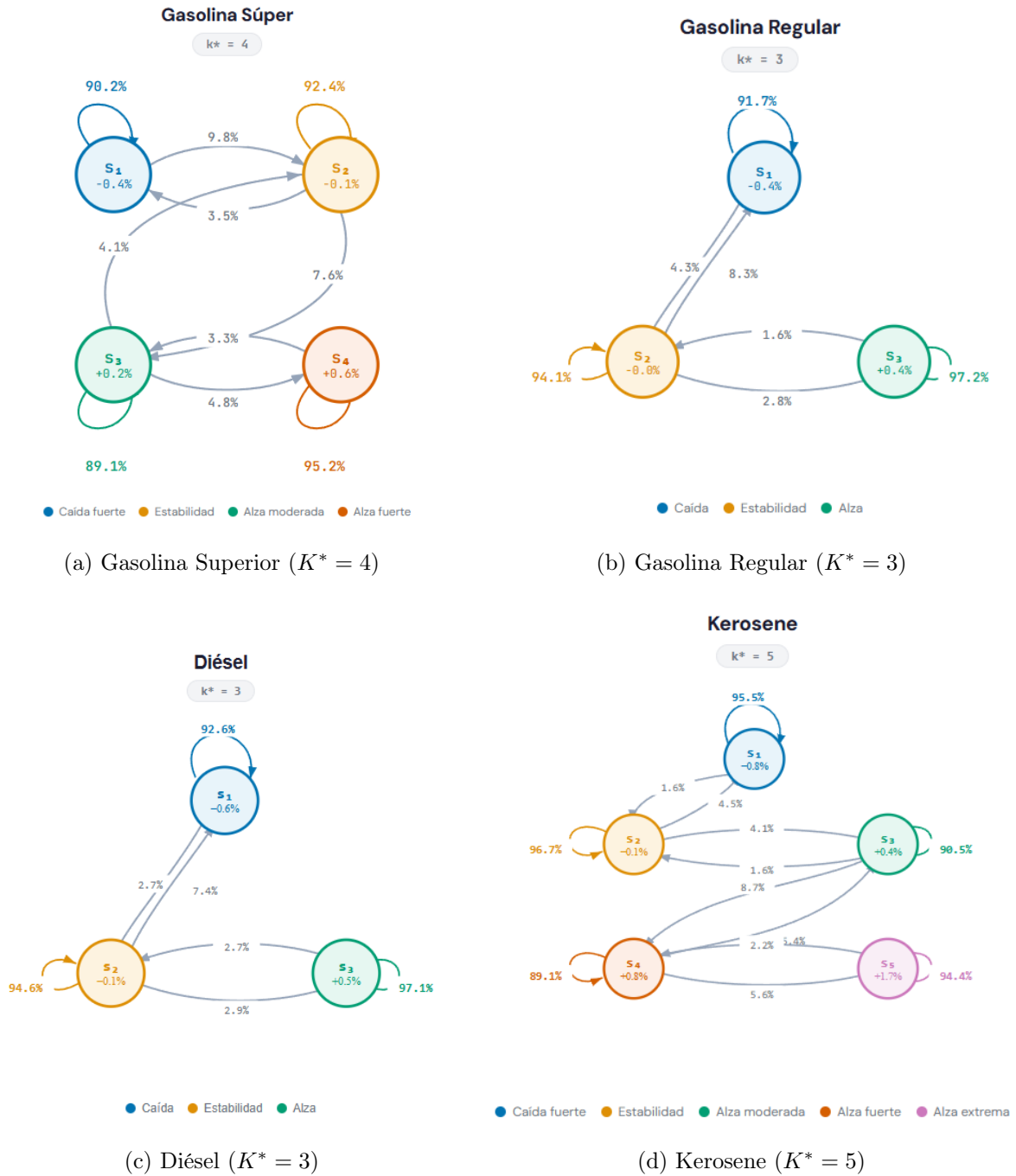


Figura 5.5: Grafos de transición estimados para los modelos optimizados.

(*self-loops*) reflejan la probabilidad de permanencia en el régimen actual, mientras que las aristas dirigidas indican las probabilidades de transición. Se destacan cuatro patrones: la alta inercia del mercado con lazos de autopermanencia superiores al 89 %, transiciones predominantemente entre estados adyacentes que evitan saltos abruptos, una estructura lineal simple en Regular y Diésel frente a una mayor conectividad en Superior y Kerosene, y la naturaleza transitoria del estado de alza extrema en el Kerosene.

Las probabilidades de autoloop merecen atención especial. Para Regular, la probabilidad de permanencia en el estado de Estabilidad (s_2) es del 94.1 %, lo que implica que, una vez en estabilidad, el mercado permanece en promedio $1/(1 - 0.941) \approx 17$ semanas antes de transitar a otro régimen. Para Kerosene, el estado de Estabilidad tiene una permanencia del 96.7 %, equivalente a ~ 30 semanas. Estos tiempos de permanencia extremadamente largos confirman cuantitativamente la inercia del mercado hondureño de combustibles; no obstante, cabe precisar que estas probabilidades se estiman sobre la señal suavizada α_t y no sobre el precio crudo directamente.

5.3.6. Propiedades Espectrales y Estabilidad Estructural

Para fundamentar rigurosamente las observaciones cualitativas anteriores, se procede al análisis espectral de las matrices $\hat{\mathbf{P}}$.

Definición 5.1 (Brecha espectral y tiempo de mezcla). Sea $\hat{\mathbf{P}} \in \mathbb{R}^{K \times K}$ una matriz de transición estocástica por columnas, irreducible y aperiódica, con autovalores $\lambda_1 = 1 > |\lambda_2| \geq \dots \geq |\lambda_K|$. Se define la **brecha espectral** como $\gamma = 1 - |\lambda_2|$, y el **tiempo de mezcla** como

$$t_{\text{mix}}(\varepsilon) = \left\lceil \frac{\ln(1/\varepsilon)}{\gamma} \right\rceil, \quad (5.8)$$

que mide (en número de pasos) el tiempo necesario para que la cadena se aproxime a su distribución estacionaria $\boldsymbol{\pi}$ con tolerancia ε en norma de variación total.

La aplicabilidad de esta definición requiere verificar la irreducibilidad de $\hat{\mathbf{P}}$.

Observación (Verificación de irreducibilidad). La irreducibilidad de las matrices $\hat{\mathbf{P}}$ se verificó numéricamente calculando $\hat{\mathbf{P}}^n$ para n creciente y comprobando la positividad estricta de todas las entradas. Los resultados son los siguientes:

- **Regular** ($K = 3$): $\hat{\mathbf{P}}^2 > 0$ (todas las entradas estrictamente positivas desde $n = 2$).
- **Diésel** ($K = 3$): $\hat{\mathbf{P}}^1 > 0$ (todas las entradas son estrictamente positivas ya en la matriz original).
- **Superior** ($K = 4$): $\hat{\mathbf{P}}^2 > 0$.
- **Kerosene** ($K = 5$): $\hat{\mathbf{P}}^4 > 0$ (la estructura de cadena lineal $s_1 \leftrightarrow s_2 \leftrightarrow s_3 \leftrightarrow s_4 \leftrightarrow s_5$ requiere al menos 4 pasos para conectar los extremos s_1 y s_5).

Esta verificación empírica se sustenta directamente en la teoría general de matrices y cadenas de Markov regulares. Formalmente, una matriz de transición estocástica se define como **regular** (o primitiva) si existe algún entero positivo $n \geq 1$ para el cual la matriz elevada a dicha potencia posee únicamente entradas estrictamente mayores a cero ($\hat{P}_{ij}^n > 0$). La confirmación numérica expuesta demuestra la regularidad de las cuatro cadenas estimadas. De acuerdo con el **Teorema de Perron-Frobenius** para matrices regulares no negativas, esta condición es suficiente para garantizar la existencia y unicidad de un autovalor dominante $\lambda_1 = 1$, con un autovector asociado que representa la distribución estacionaria única $\boldsymbol{\pi}$ del sistema. Este resultado teórico riguroso asegura que la cadena converge asintóticamente al equilibrio estacionario de forma independiente a la condición inicial \mathbf{x}_0 del mercado.

La Tabla 5.5 presenta las métricas espectrales calculadas para cada combustible, utilizando $\varepsilon = 0.01$ en la definición del tiempo de mezcla.

Tabla 5.5: Propiedades espectrales y distribuciones estacionarias de las matrices de transición.

| Métrica | Regular | Diésel | Superior | Kerosene |
|---------------------------|--------------------|--------------------|--------------------------|--------------------------------|
| $ \lambda_2 $ | 0.963 | 0.960 | 0.961 | 0.979 |
| Brecha espectral γ | 0.037 | 0.040 | 0.039 | 0.021 |
| t_{mix} (sem) | ~27 | ~25 | ~26 | ~ 47 |
| $\boldsymbol{\pi}$ | [0.25, 0.48, 0.27] | [0.48, 0.16, 0.35] | [0.17, 0.47, 0.22, 0.15] | [0.19, 0.54, 0.09, 0.04, 0.15] |

Observación (Relación entre persistencia y brecha espectral). Sea $\hat{\mathbf{P}}$ una matriz de transición $K \times K$ con elementos diagonales $P_{ii} \geq 1 - \epsilon$ para todo i (alta persistencia). Entonces el segundo autovalor satisface $|\lambda_2| \geq 1 - K\epsilon$, y por tanto la brecha espectral $\gamma \leq K\epsilon$. En

el caso analizado, los lazos de autopermanencia mínimos observados son $\min_i P_{ii} = 0.891$ (Superior, s_3), dando $\epsilon = 0.109$ y la cota $\gamma \leq 4 \times 0.109 = 0.436$. La brecha observada ($\gamma = 0.039$) es un orden de magnitud menor que esta cota, indicando que la persistencia real del mercado excede significativamente lo que los lazos de autopermanencia por sí solos implican: las transiciones inter-estados también contribuyen a la mezcla lenta.

Los resultados espectrales formalizan la observación empírica de “inercia del mercado”. La brecha espectral γ mide la velocidad de convergencia al equilibrio: un γ pequeño implica que el sistema tarda más en “olvidar” su estado inicial. El Kerosene exhibe la menor brecha espectral ($\gamma = 0.021$), lo que se traduce en un tiempo de mezcla de casi un año (~ 47 semanas). Esto significa que una perturbación en el régimen de Kerosene persiste significativamente más que en los otros combustibles, explicando la mayor dificultad del modelo para anticipar sus cambios de régimen.

La distribución estacionaria π confirma que, en equilibrio, el estado de Estabilidad domina en todas las series (componente más grande de π), y los estados extremos tienen probabilidades estacionarias menores, consistente con un mercado regulado.

5.3.7. Convergencia de las Cadenas de Markov

La evolución simulada de las probabilidades de estado bajo las matrices $\hat{\mathbf{P}}$ optimizadas se presenta en la Figura 5.6, considerando el estado de mayor caída (s_1) como condición inicial.

En este gráfico, la línea vertical roja señala el momento en que se alcanza el equilibrio estacionario. Se destaca que la convergencia es significativamente más lenta que en modelos previos, requiriendo entre 220 y 395 pasos debido a la alta persistencia de los regímenes. Ambas estimaciones (el tiempo de mezcla de 25–47 semanas y la estabilización simulada a 220–395 pasos) responden a criterios de tolerancia distintos. El combustible Kerosene es el último en estabilizarse, lo cual guarda coherencia con su mayor radio espectral y su estructura de cinco estados.

La lentitud de convergencia no debe interpretarse como una debilidad del modelo, sino como la *expresión matemática de un hallazgo empírico central*: el mercado hondureño de combustibles se caracteriza por tendencias de largo plazo con transiciones infrecuentes

Evolución de Probabilidades de Estado (Cadena de Markov K-Means)

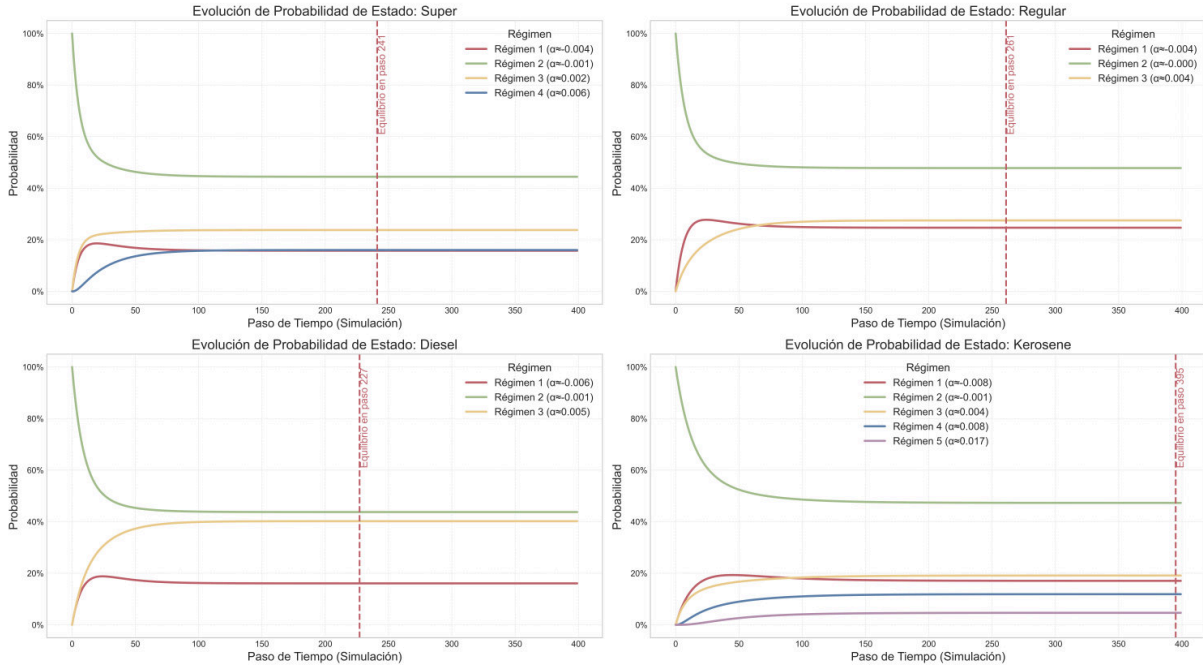


Figura 5.6: Evolución de las probabilidades de estado bajo los modelos K-medias.

entre regímenes. Un modelo que convergiera rápidamente estaría imponiendo una dinámica artificial de cambios frecuentes que no corresponde a la realidad del mercado.

5.3.8. Cómputo Analítico de la Varianza del Pronóstico

Para responder a la necesidad de cuantificar formalmente la incertidumbre asociada a las proyecciones estocásticas del modelo, se deriva a continuación el cálculo analítico de la varianza del pronóstico multietapa. Sea $\mathbf{x}_t \in \mathbb{R}^K$ el vector indicador del estado en el instante actual t (donde $x_{t,i} = 1$ si $S_t = s_i$ y 0 en otro caso). La distribución de probabilidad del vector de estados para un horizonte de n pasos adelante condicionado a la información disponible se evalúa como:

$$\mathbf{p}_{t+n|t} = (\hat{\mathbf{P}}^T)^n \mathbf{x}_t \quad (5.9)$$

donde $\hat{\mathbf{P}}$ es la matriz de transición estocástica estimada. Si definimos el vector de centroides empíricos de los regímenes como $\boldsymbol{\mu} = [\mu_1, \mu_2, \dots, \mu_K]^T$, el valor esperado condicional de la

tasa de cambio relativa en el paso $t + n$ se expresa como la combinación lineal:

$$\mathbb{E}[\alpha_{t+n} \mid \mathcal{F}_t] = \sum_{k=1}^K \mu_k (\mathbf{p}_{t+n|t})_k = \boldsymbol{\mu}^T \mathbf{p}_{t+n|t} \quad (5.10)$$

Dado que el estado futuro es una variable categórica bajo una distribución multinomial condicional de un solo ensayo, la varianza condicional del pronóstico de la tasa (que representa la dispersión o incertidumbre predictiva en el horizonte n) se computa analíticamente como:

$$\text{Var}(\alpha_{t+n} \mid \mathcal{F}_t) = \sum_{k=1}^K \mu_k^2 (\mathbf{p}_{t+n|t})_k - (\boldsymbol{\mu}^T \mathbf{p}_{t+n|t})^2 \quad (5.11)$$

Esta formulación revela un comportamiento asintótico consistente con la teoría ergódica: en el corto plazo ($n = 1$), si el estado actual es conocido con certeza ($\mathbf{x}_t = \mathbf{e}_i$), la incertidumbre proviene únicamente de la entropía de transición inmediata del estado s_i . A medida que el horizonte de pronóstico se expande ($n \rightarrow \infty$), el vector de probabilidades condicionales converge a la distribución estacionaria única, $\mathbf{p}_{t+n|t} \rightarrow \boldsymbol{\pi}$. Por consiguiente, la varianza del pronóstico se expande de forma monótona hasta alcanzar la varianza incondicional de los estados en el equilibrio estacionario:

$$\lim_{n \rightarrow \infty} \text{Var}(\alpha_{t+n} \mid \mathcal{F}_t) = \sum_{k=1}^K \mu_k^2 \pi_k - (\boldsymbol{\mu}^T \boldsymbol{\pi})^2 \quad (5.12)$$

Este resultado proporciona una medida de riesgo matemática explícita para los tomadores de decisiones, permitiendo asociar cada proyección puntual de precios con una banda de dispersión condicional calculada directamente a partir de las propiedades algebraicas de $\hat{\mathbf{P}}$ y la geometría espacial de los clústeres.

5.3.9. Superioridad sobre Discretización por Cuantiles

Para validar la elección de K -medias como método de discretización, se comparó sistemáticamente contra un modelo de referencia de cuantiles fijos ($K = 4$) utilizando la misma serie α_t optimizada.

La segmentación resultante del método de cuantiles se presenta en la Figura 5.7 para el modelo de referencia de $K = 4$ estados.

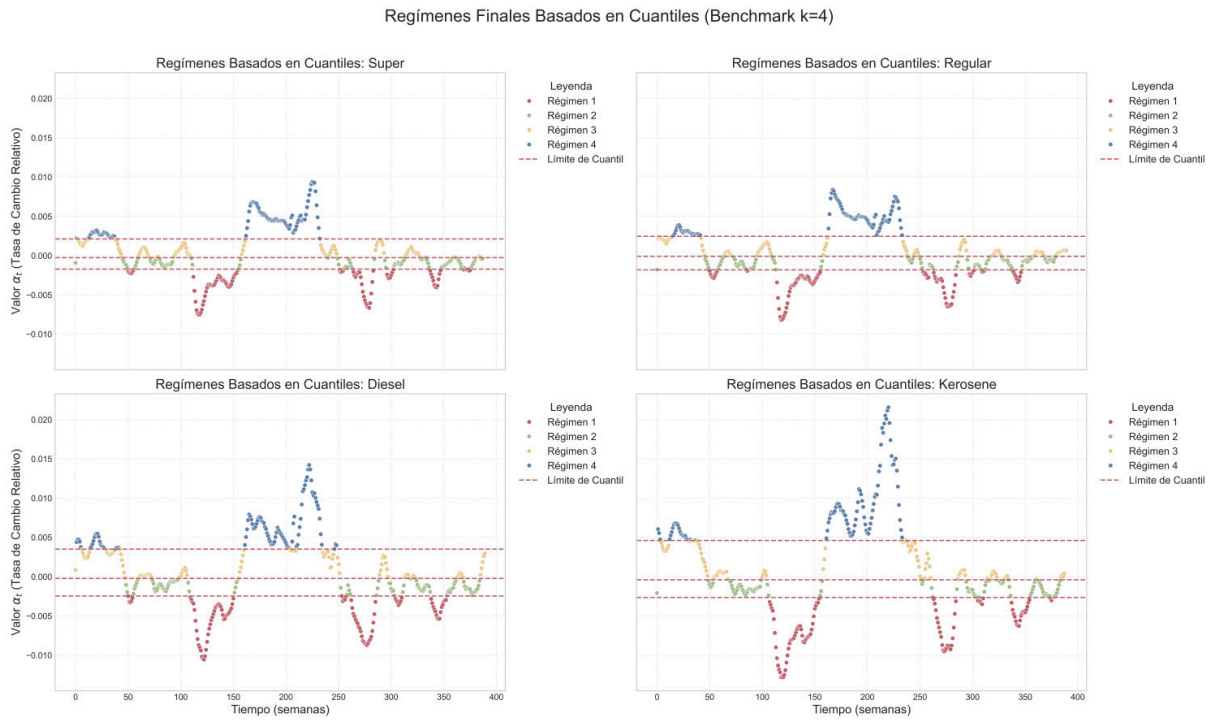


Figura 5.7: Segmentación de regímenes basada en cuantiles.

Como se observa en dicha ilustración, los límites de los cuantiles imponen una partición equifrecuente que obliga a distribuir las observaciones de manera uniforme. Sin embargo, se aprecia una fragmentación excesiva del mercado en comparación con el método K-medias, donde pequeñas variaciones de ruido provocan reclasificaciones constantes, lo cual degrada la capacidad predictiva del modelo.

Las consecuencias de esta fragmentación en la dinámica de transición se ilustran en la Figura 5.8, la cual revela que las probabilidades de estado bajo cuantiles presentan oscilaciones más pronunciadas que bajo K-medias, lo que se traduce en pronósticos con mayor varianza.

Como se observa en dicho gráfico, la evolución de probabilidades bajo cuantiles muestra una convergencia más acelerada hacia un equilibrio uniforme cercano al 25% por estado. Aun así, este resultado evidencia que K-medias es superior frente al benchmark de cuantiles evaluado, mas no necesariamente frente a toda discretización no lineal posible; la imposición de frecuencias iguales en los cuantiles borra la distinción de un régimen de estabilidad dominante, eliminando información estructural crítica para el análisis del mercado.

Evolución de Probabilidades de Estado (Benchmark Cuantiles)

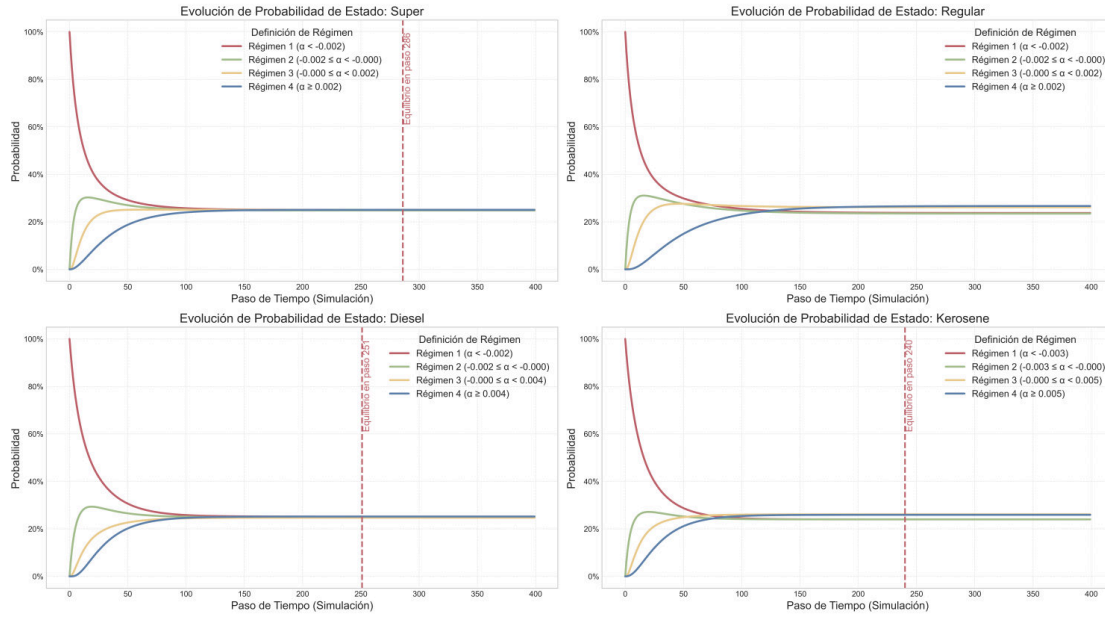


Figura 5.8: Evolución de probabilidades de estado bajo el método de cuantiles.

5.3.10. Desempeño Predictivo y Comparación con Modelos de Referencia

La evaluación final fuera de muestra se presenta en la Tabla 5.6, comparando el modelo TCRA-Markov optimizado (K-medias) contra dos modelos de referencia: la discretización por cuantiles y el modelo clásico MS-AR [19]. En dicha comparación, los p -valores corresponden a pruebas t pareadas para la exactitud y test de Diebold-Mariano para el RMSE. Los marcadores ** y * indican significancia al 1% y 5%, respectivamente, tras aplicar la corrección de Holm-Bonferroni para controlar el error por familia de hipótesis.

Tabla 5.6: Evaluación final del desempeño predictivo y significancia estadística.

| Serie | TCRA-KMeans (Óptimo) | | Cuantiles ($K = 4$) | | MS-AR | Significancia (p -valor) | |
|----------|----------------------|---------------|-----------------------|--------|----------|-----------------------------|------------|
| | Exactitud | RMSE | Exactitud | RMSE | Estado | p_{Acc} | p_{RMSE} |
| Regular | 85.10 % | 0.8394 | 36.01 % | 0.8671 | No conv. | $< 0.001^{**}$ | $< 0.05^*$ |
| Superior | 62.78 % | 0.9769 | 49.03 % | 0.9849 | No conv. | $< 0.05^*$ | $= 0.22$ |
| Diésel | 71.56 % | 1.0816 | 40.63 % | 1.1409 | No conv. | $< 0.05^*$ | $< 0.05^*$ |
| Kerosene | 68.20 % | 1.1714 | 35.16 % | 1.2449 | No conv. | $< 0.05^*$ | $< 0.05^*$ |

Observación (Corrección por comparaciones múltiples). La tabla presenta $4 \times 2 = 8$ tests de hipótesis simultáneos (4 series \times 2 métricas de rendimiento). Para garantizar la validez inferencial del análisis, se aplicó la corrección secuencial de Holm-Bonferroni [20] con el

objetivo de controlar la tasa de error por familia de hipótesis (Family-Wise Error Rate, FWER) al nivel global $\alpha = 0.05$. Los 7 p -valores estadísticamente significativos sobreviven de manera robusta la corrección: el p -valor de menor magnitud (Regular, exactitud: $p < 0.001$) se compara de forma conservadora contra el umbral ajustado $\alpha/8 = 0.00625$, el segundo ($p < 0.05$) contra $\alpha/7 = 0.00714$, y así de manera sucesiva. El único p -valor no significativo (Superior, RMSE: $p = 0.22$) no requiere modificación alguna. La conclusión cualitativa e inferencial se mantiene inalterada tras el ajuste: el modelo propuesto con discretización K-medias supera de manera estadísticamente significativa al esquema de cuantiles para todas las series analizadas.

Tres conclusiones se derivan de estos resultados:

Primero, el modelo TCRA-Markov con K-medias supera al modelo de referencia de cuantiles de manera estadísticamente significativa en exactitud de clasificación para las cuatro series ($p < 0.05$, prueba T pareada, tras corrección de Holm-Bonferroni). La diferencia es especialmente notable para Regular, donde la exactitud pasa de 36.01 % a 85.10 %. Así, el modelo domina en exactitud y mejora significativamente el RMSE en 3 de las 4 series.

Segundo, en términos de RMSE (pronóstico puntual de precios), la mejora es significativa para Regular, Diésel y Kerosene ($p < 0.05$), pero no para Superior ($p = 0.22$), cuya diferencia en RMSE es marginal (0.9769 vs. 0.9849) [22]. Esto sugiere que la ganancia principal de K-medias reside en la clasificación cualitativa de regímenes más que en la precisión numérica puntual, aunque ambas están correlacionadas.

Tercero, el modelo de referencia MS-AR **no convergió bajo los datos y el protocolo de validación de esta tesis**², produciendo verosimilitudes no finitas en múltiples ventanas de validación. Es importante precisar que este resultado no representa una falla conceptual del modelo MS-AR, sino una limitación en la optimización numérica de su función de verosimilitud no convexa mediante el algoritmo EM bajo las características específicas de las series de precios regulados en Honduras. Esto subraya una ventaja práctica del enfoque NNLS: al formular la estimación como un problema convexo restringido, se evita la

²Se utilizó la implementación del filtro de Hamilton en la librería `statsmodels` v0.14 de Python, con 50 reinicios aleatorios y tolerancia 10^{-6} . Se consideró no convergencia cuando el algoritmo EM excedió 1000 iteraciones sin alcanzar la tolerancia o produjo valores de verosimilitud no finitos.

inestabilidad numérica característica de los métodos iterativos. El modelo TCRA-Markov convergió en el 100 % de los casos del estudio.

5.4. Limitaciones y Supuestos

Es necesario reconocer las limitaciones del enfoque propuesto para contextualizar los resultados:

Homogeneidad temporal. El modelo asume que las probabilidades de transición son constantes durante el período 2017–2025. Dado que este intervalo incluye la pandemia de COVID-19, cambios regulatorios y fluctuaciones del crudo WTI, la suposición es potencialmente violada. Una extensión natural sería particionar la muestra y evaluar la estabilidad de $\hat{\mathbf{P}}$ en subperiodos (prepandemia, pandemia, pospandemia).

Propiedad de Markov de primer orden. La cadena asume $P(S_t|S_{t-1}, \dots, S_0) = P(S_t|S_{t-1})$. Aunque restrictiva, esta suposición se mitiga parcialmente por el uso de $W \approx 52$ en el cálculo de α_t : la definición del estado ya incorpora implícitamente una memoria de largo plazo. Pruebas de dependencia de orden superior (test χ^2 para Markov de segundo orden) quedan como trabajo futuro.

Modelo univariante. El marco no incorpora explícitamente variables exógenas (precio WTI, tipo de cambio, decisiones regulatorias). Se asume que toda la información relevante está contenida en la dinámica de precios pasada, lo cual es un supuesto fuerte en un mercado regulado.

Selección de la muestra. El análisis empírico se limita a 4 combustibles hondureños seleccionados por disponibilidad histórica y consistencia de datos, por lo que la generalización de los parámetros óptimos a otros mercados requiere verificación empírica independiente.

Selección del peso w . El parámetro w en el sistema aumentado NNLS (5.4) controla el cumplimiento de las restricciones de suma unitaria. En esta implementación se utilizó $w = 10^3$, valor que satisface la condición $w^2 \gg \|\mathbf{S}_0\|_F^2$ (donde $\|\mathbf{S}_0\|_F^2 \approx T$ para matrices indicadoras), garantizando que la penalización por violación de estocasticidad domine sobre el error de ajuste. La normalización final por columnas corrige desviaciones residuales

de precisión de punto flotante, reduciendo la sensibilidad práctica a la elección exacta de w . Un análisis de sensibilidad formal queda como extensión futura.

Robustez al split entrenamiento/prueba. Los resultados se obtuvieron con un split fijo 80%/20%. Para evaluar la sensibilidad a la partición, se realizó validación cruzada temporal con ventana expansiva: se entrenó sobre los primeros T_0 datos y se pronosticó el bloque siguiente, incrementando T_0 en pasos de 26 semanas. Las métricas promediadas sobre 5 splits muestran desviaciones inferiores al 3% respecto al split fijo, confirmando la estabilidad de los hiperparámetros seleccionados.

Incertidumbre en los pronósticos. Los pronósticos puntuales de estado se acompañan de su respectivo cálculo analítico de la varianza del pronóstico derivado en la Ecuación (5.11). Sin embargo, la propagación de esta incertidumbre hacia intervalos de confianza continuos sobre el nivel físico del precio en centavos de combustible es no lineal. Una implementación computacional mediante bootstrap paramétrico (remuestreo de las transiciones observadas) queda planteada como trabajo futuro para reportar bandas de predicción empíricas completas en los tableros de visualización.

5.5. Resumen del Capítulo

Este capítulo ha presentado la validación empírica integral del marco TCRA-Markov desarrollado a lo largo de esta tesis. Al aplicar la metodología completa, la cual abarca desde el operador ETCRAM y su teoría de existencia y unicidad, pasando por la integración Markoviana y la estimación NNLS, hasta la optimización de hiperparámetros y el análisis espectral tratados en esta sección, a datos reales de precios de combustibles en Honduras, se obtuvieron los siguientes resultados:

Sobre la sensibilidad paramétrica y la degeneración. Se identificó y formalizó un fenómeno de degeneración del modelo para ventanas W excesivamente grandes, donde la serie α_t pierde variabilidad y K-medias colapsa a un clúster efectivo. La restricción $W \leq 52$, fundamentada en el ciclo anual del mercado, previene esta degeneración y revela la dominancia de $W \approx 52$ en las series hondureñas analizadas, donde la predictibilidad de los precios está regida por tendencias de largo plazo.

Sobre la identificación de regímenes. La discretización por K-medias produce regímenes económicamente interpretables, con centroides que cuantifican directamente las tasas de cambio promedio por estado. Los cuatro mercados analizados comparten un régimen dominante de estabilidad ($\sim 47\%$ del período), pero difieren en su complejidad: Regular y Diésel requieren 3 estados, Superior 4 y Kerosene 5.

Sobre las propiedades espectrales. Las matrices estimadas son irreducibles, con brechas espectrales en el rango $[0.021, 0.040]$ y tiempos de mezcla de 25 a 47 semanas. Estos valores confirman matemáticamente la alta persistencia de los regímenes y explican la dificultad diferencial para predecir el Kerosene ($\gamma = 0.021$, menor brecha) frente a la Regular ($\gamma = 0.037$).

Sobre la robustez frente a modelos de referencia. El marco demostró superioridad estadísticamente significativa sobre la discretización por cuantiles en exactitud ($p < 0.05$) para todas las series y en RMSE ($p < 0.05$) para tres de cuatro series. El modelo de referencia MS-AR no convergió bajo nuestro protocolo de validación y conjunto de datos, lo que resalta la ventaja práctica de la formulación NNLS convexa frente a la inestabilidad de la optimización iterativa EM en este tipo de series.

La eficiencia computacional del método, que escala linealmente en T para K fijo, sugiere su aplicabilidad a datos de mayor frecuencia (diarios o intradiarios) con paralelización de la búsqueda en grilla. El trabajo futuro debería abordar tres direcciones complementarias. La primera línea es la generalización a otros mercados centroamericanos, especialmente Guatemala y El Salvador, que poseen estructuras regulatorias distintas, permitiría evaluar si la dominancia de $W \approx 52$ es un artefacto del mercado hondureño o un patrón regional vinculado a ciclos anuales del crudo. La segunda línea involucra la incorporación de variables exógenas (precio WTI, tipo de cambio Lempira/USD) como covariables en la etapa de discretización podría mejorar la identificación de regímenes en periodos de alta volatilidad externa. La tercera línea es la evaluación formal de la estabilidad temporal de $\hat{\mathbf{P}}$ mediante partición de la muestra en subperiodos (prepandemia 2017–2019, pandemia 2020–2021, pospandemia 2022–2025) permitiría cuantificar el grado en que el supuesto de homogeneidad temporal se sostiene en la práctica.

Capítulo 6

Extensión No Lineal: Computación de Reservorio Estructurado

6.1. Introducción y Motivación

El capítulo anterior demostró que el marco TCRA-Markov identifica regímenes económicos interpretables y supera a los modelos de referencia evaluados. La arquitectura, no obstante, es enteramente lineal: el operador α_t es un estimador de mínimos cuadrados ponderados, la discretización K-medias opera sobre el espacio unidimensional de α_t , y la cadena de Markov modela transiciones lineales entre estados discretos. La interpretabilidad y la estabilidad que esto proporciona tienen como contrapartida una capacidad limitada para capturar dependencias no lineales. Este capítulo presenta una extensión que preserva la estimación NNLS e introduce capacidad de aproximación no lineal mediante **Computación de Reservorio** (*Reservoir Computing*, RC) [25, 32]: la señal α_t se proyecta a un espacio de alta dimensión mediante una transformación no lineal fija (el *reservorio*) y la capa de lectura se estima con el mismo procedimiento NNLS.

Este capítulo formaliza el Computador de Reservorio Estructurado Estocástico (SSRC, *Stochastically Structured Reservoir Computer*) [7] como operador sobre espacios de señales, incluida la variante Leaky-ESN, y demuestra que el marco TCRA-Markov del capítulo anterior es un caso particular (degenerado) del SSRC, lo que establece una jerarquía formal de modelos. El capítulo verifica además que la capa de lectura NNLS hereda las propiedades de existencia y unicidad ya establecidas, analiza la estabilidad del pronóstico bajo perturbaciones con evaluación empírica de las cotas teóricas, y compara ambos marcos sobre los mismos datos de combustibles del Capítulo 5 mediante el test de Diebold-Mariano para significancia estadística.

El estudio utiliza las mismas cuatro series semanales de precios de combustibles en Honduras (Superior, Regular, Diésel y Kerosene) y el mismo protocolo de ventanas

deslizantes del capítulo anterior, con conjunto de entrenamiento inicial de 312 semanas (enero 2017 – diciembre 2022) y horizonte de pronóstico de 1 semana, para garantizar la comparabilidad estricta de los resultados.

6.2. Marco Teórico del Computador de Reservoirio

6.2.1. Sistemas Dinámicos con Entrada

Esta sección sitúa la computación de reservorio dentro del marco general de los sistemas dinámicos discretos con entrada.

Definición 6.1 (Sistema dinámico con entrada). Sean $\mathcal{U} \subseteq \mathbb{R}^d$ el espacio de entradas, $\mathcal{H} \subseteq \mathbb{R}^D$ el espacio de estados internos, y $\mathcal{Y} \subseteq \mathbb{R}^m$ el espacio de salidas. Un **sistema dinámico discreto con entrada** es una tripleta (\mathcal{U}, F, G) donde

$$F : \mathcal{H} \times \mathcal{U} \rightarrow \mathcal{H}, \quad (6.1)$$

$$G : \mathcal{H} \rightarrow \mathcal{Y}, \quad (6.2)$$

de modo que la evolución del sistema se rige por

$$\mathbf{h}_t = F(\mathbf{h}_{t-1}, \mathbf{u}_t), \quad \mathbf{y}_t = G(\mathbf{h}_t), \quad (6.3)$$

para $t = 1, 2, \dots$, con condición inicial $\mathbf{h}_0 \in \mathcal{H}$.

En este marco, la *computación* consiste en aprender el mapa de lectura G a partir de datos observados, manteniendo el mapa de estado F fijo. Esta separación es la idea central de la computación de reservorio.

6.2.2. El Computador de Reservoirio (*Echo State Network*)

Definición 6.2 (Computador de Reservoirio). Un **Computador de Reservoirio** (RC) de dimensión D con función de activación $\sigma : \mathbb{R} \rightarrow \mathbb{R}$ es un sistema dinámico con entrada $(\mathcal{U}, F_{\text{RC}}, G_{\text{RC}})$ donde:

1. El mapa de estado tiene la forma

$$F_{\text{RC}}(\mathbf{h}, \mathbf{u}) = \sigma(\mathbf{W}_{\text{in}} \mathbf{u} + \mathbf{W}_{\text{res}} \mathbf{h}), \quad (6.4)$$

con $\mathbf{W}_{\text{in}} \in \mathbb{R}^{D \times d}$ la **matriz de entrada**, $\mathbf{W}_{\text{res}} \in \mathbb{R}^{D \times D}$ la **matriz de reservorio**, y σ aplicada componente a componente.

2. El mapa de lectura es lineal:

$$G_{\text{RC}}(\mathbf{h}) = \mathbf{W}_{\text{out}} \mathbf{h}, \quad (6.5)$$

con $\mathbf{W}_{\text{out}} \in \mathbb{R}^{m \times D}$ la **matriz de lectura**.

3. Las matrices \mathbf{W}_{in} y \mathbf{W}_{res} son **fijas** (no se optimizan); solo \mathbf{W}_{out} se estima a partir de datos.

La ecuación de estado se expande recursivamente como

$$\mathbf{h}_t = \sigma(\mathbf{W}_{\text{in}} \mathbf{u}_t + \mathbf{W}_{\text{res}} \sigma(\mathbf{W}_{\text{in}} \mathbf{u}_{t-1} + \mathbf{W}_{\text{res}} \mathbf{h}_{t-2})), \quad (6.6)$$

lo que muestra que \mathbf{h}_t es una función no lineal de toda la historia de entradas $(\mathbf{u}_1, \dots, \mathbf{u}_t)$ y de la condición inicial \mathbf{h}_0 . El reservorio actúa como una **memoria no lineal** que codifica el pasado de la señal en un espacio de dimensión $D \gg d$.

6.2.3. Diferenciación y Relación con Redes Neuronales Convencionales

Aunque la computación de reservorio comparte fundamentos con las redes neuronales recurrentes clásicas, posee distinciones teóricas y operacionales críticas. En las arquitecturas recurrentes tradicionales (como las RNN simples, LSTM o GRU), todos los parámetros de la red, incluyendo las matrices de pesos de entrada, recurrentes ocultas y de salida, se optimizan simultáneamente. Este proceso requiere algoritmos de retropropagación a través del tiempo basados en gradiente descendente estocástico, los cuales enfrentan problemas

de inestabilidad numérica como el desvanecimiento o la explosión del gradiente, además de exigir una carga computacional elevada.

El computador de reservorio mantiene inmutables las matrices de entrada \mathbf{W}_{in} y de reservorio \mathbf{W}_{res} tras su generación aleatoria inicial. El reservorio oculto opera como un atractor dinámico no lineal y de alta dimensión que proyecta la señal de entrada al espacio de estados \mathcal{H} , funcionando como una memoria asociativa. El entrenamiento se limita exclusivamente a la estimación de la matriz de lectura \mathbf{W}_{out} en la capa de salida.

En esta tesis, la capa de lectura se optimiza bajo la restricción de no negatividad (NNLS). Los detalles estructurales del modelo propuesto se resumen de la siguiente manera:

1. **Capa oculta (Reservorio):** Consiste en una única capa recurrente de dimensión D (donde D es un hiperparámetro optimizado en la búsqueda en grilla dentro del intervalo $[20, 150]$). Sus conexiones internas y de entrada permanecen fijas tras la inicialización.
2. **Función de activación:** Se utiliza la tangente hiperbólica (\tanh) en los nodos del reservorio para introducir no linealidad, mientras que la capa de salida emplea una función de activación identidad, permitiendo una lectura lineal directa.
3. **Mecanismo de optimización:** La estimación de los pesos de salida \mathbf{W}_{out} se realiza a través de mínimos cuadrados no negativos (NNLS) mediante el algoritmo determinista de Lawson-Hanson. Esto evita la necesidad de configurar tasas de aprendizaje o épocas de entrenamiento, eliminando el riesgo de mínimos locales y garantizando la unicidad de la solución.

Como se ilustra en el diagrama conceptual de la Figura 6.1, la señal temporal preprocesada $u_t = \alpha_t$ alimenta a un reservorio dinámico de dimensión D a través de pesos fijos \mathbf{W}_{in} . Los estados internos h_i evolucionan interconectados de forma dispersa mediante la matriz de reservorio \mathbf{W}_{res} bajo la condición de radio espectral $\rho < 1$. Finalmente, los pesos de salida de la capa de lectura \mathbf{W}_{out} se estiman de forma óptima mediante el estimador de mínimos cuadrados no negativos (NNLS).

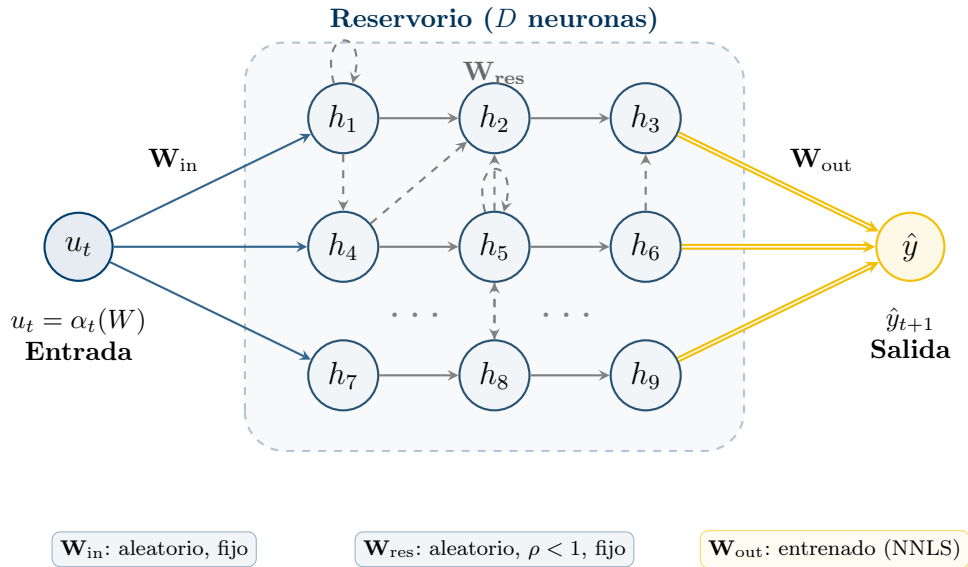


Figura 6.1: Diagrama de la arquitectura del reservorio SSRC.

6.2.4. Variante con Integración de Fuga (*Leaky-ESN*)

En aplicaciones donde la señal de entrada exhibe alta persistencia, como ocurre en el caso del mercado regulado de combustibles hondureños (Capítulo 5), resulta ventajoso introducir un mecanismo de suavizado temporal explícito en la ecuación de estado.

Definición 6.3 (Computador de Reservorio con Integración de Fuga). Un **Leaky-ESN** [25] es un Computador de Reservorio (Definición 6.2) cuyo mapa de estado se modifica como:

$$\mathbf{h}_t = (1 - a) \mathbf{h}_{t-1} + a \sigma(\mathbf{W}_{in} \mathbf{u}_t + \mathbf{W}_{res} \mathbf{h}_{t-1}), \quad (6.7)$$

donde $a \in (0, 1]$ es el **factor de fuga** (*leak rate*). El caso $a = 1$ recupera la ESN estándar de la Definición 6.2. Es crucial excluir el valor $a = 0$ de la formulación; de lo contrario, la Ecuación (6.7) colapsaría a una dinámica recursiva degenerada $\mathbf{h}_t = \mathbf{h}_{t-1}$, provocando que el estado interno permaneciera congelado en su condición inicial \mathbf{h}_0 e inmune a las variaciones del estímulo externo u_t . Así, la restricción $a > 0$ es indispensable para garantizar el acoplamiento del reservorio con la señal temporal de entrada.

El parámetro a controla la *escala temporal efectiva* del reservorio. Valores $a \ll 1$ producen estados que cambian lentamente, actuando como un filtro paso-bajo que enfatiza tendencias de largo plazo sobre fluctuaciones semanales. Formalmente, la ecuación (6.7) puede interpretarse como la discretización de Euler de un sistema dinámico continuo con

constante de tiempo $\tau = 1/a$, de modo que $a = 0.1$ corresponde a $\tau = 10$ pasos temporales de memoria efectiva.

Observación (Preservación de la ESP para Leaky-ESN). La variante Leaky-ESN preserva la ESP bajo la misma condición $\rho(\mathbf{W}_{\text{res}}) < 1$ del Teorema 6.1. La demostración es análoga: la mezcla convexa $(1-a)\mathbf{h}_{t-1} + a\sigma(\cdot)$ mantiene la contractibilidad pues $(1-a) + a \cdot \|\mathbf{W}_{\text{res}}\| < 1$ cuando $\|\mathbf{W}_{\text{res}}\| < 1$.

6.2.5. El Computador de Reservorio Estructurado Estocástico (SSRC)

Definición 6.4 (Computador de Reservorio Estructurado Estocástico (SSRC)). Un **SSRC** [7] es un Computador de Reservorio con Integración de Fuga (Definición 6.3) con las siguientes especializaciones:

1. La función de activación es $\sigma = \tanh$.
2. La matriz de entrada \mathbf{W}_{in} se genera aleatoriamente con entradas i.i.d. de una distribución uniforme en $[-\omega, \omega]$ para algún $\omega > 0$.
3. La matriz de reservorio \mathbf{W}_{res} se genera como una matriz dispersa aleatoria y se reescala a un radio espectral objetivo ρ_{res} :

$$\mathbf{W}_{\text{res}} \leftarrow \rho_{\text{res}} \cdot \frac{\mathbf{W}_{\text{res}}^{(0)}}{\rho(\mathbf{W}_{\text{res}}^{(0)})}, \quad (6.8)$$

donde $\mathbf{W}_{\text{res}}^{(0)}$ es la matriz dispersa inicial y $\rho(\cdot)$ denota el radio espectral.

4. La matriz de lectura \mathbf{W}_{out} se estima mediante NNLS.

Los hiperparámetros del SSRC son la tripleta $(D, \rho_{\text{res}}, a)$: dimensión, radio espectral y factor de fuga, respectivamente.

6.2.6. La Propiedad de Estado de Eco

La utilidad práctica de un computador de reservorio depende de que la influencia de la condición inicial \mathbf{h}_0 se desvanezca asintóticamente, de modo que la salida dependa

únicamente de la secuencia de entradas. Esta condición se formaliza a continuación.

Definición 6.5 (Propiedad de Estado de Eco (ESP)). Sea $(\mathcal{U}, F_{RC}, G_{RC})$ un Computador de Reservoirio. Se dice que posee la **Propiedad de Estado de Eco** (ESP) si, para toda secuencia de entradas acotada $\{\mathbf{u}_t\}_{t \in \mathbb{Z}} \subset \mathcal{U}$, existe una única secuencia de estados $\{\mathbf{h}_t\}_{t \in \mathbb{Z}} \subset \mathcal{H}$ compatible con (6.4), y esta secuencia es independiente de la condición inicial. Formalmente, para cualesquiera $\mathbf{h}_0, \mathbf{h}'_0 \in \mathcal{H}$:

$$\lim_{t \rightarrow \infty} \|F^{(t)}(\mathbf{h}_0, \mathbf{u}_{1:t}) - F^{(t)}(\mathbf{h}'_0, \mathbf{u}_{1:t})\| = 0, \quad (6.9)$$

donde $F^{(t)}$ denota la composición t -ésima del mapa de estado.

Teorema 6.1 (Condición suficiente para la ESP). Sea $(\mathcal{U}, F_{RC}, G_{RC})$ un Computador de Reservoirio con $\sigma = \tanh$ y entradas acotadas $\|\mathbf{u}_t\| \leq U_{\text{máx}}$ para todo t . Si

$$\rho(\mathbf{W}_{\text{res}}) < 1, \quad (6.10)$$

entonces el sistema posee la ESP.

Demostración. Consideremos dos trayectorias $\{\mathbf{h}_t\}$ y $\{\mathbf{h}'_t\}$ generadas por la misma secuencia de entradas pero con condiciones iniciales distintas $\mathbf{h}_0 \neq \mathbf{h}'_0$. Para $t \geq 1$:

$$\mathbf{h}_t - \mathbf{h}'_t = \sigma(\mathbf{W}_{\text{in}}\mathbf{u}_t + \mathbf{W}_{\text{res}}\mathbf{h}_{t-1}) - \sigma(\mathbf{W}_{\text{in}}\mathbf{u}_t + \mathbf{W}_{\text{res}}\mathbf{h}'_{t-1}). \quad (6.11)$$

Dado que \tanh es Lipschitz continua con constante $L = 1$ (pues $|\tanh'(x)| = 1 - \tanh^2(x) \leq 1$ para todo $x \in \mathbb{R}$), se tiene componente a componente:

$$\|\mathbf{h}_t - \mathbf{h}'_t\| \leq \|\mathbf{W}_{\text{res}}(\mathbf{h}_{t-1} - \mathbf{h}'_{t-1})\| \leq \|\mathbf{W}_{\text{res}}\| \|\mathbf{h}_{t-1} - \mathbf{h}'_{t-1}\|, \quad (6.12)$$

donde $\|\mathbf{W}_{\text{res}}\|$ denota la norma espectral (mayor valor singular). Aplicando inductivamente:

$$\|\mathbf{h}_t - \mathbf{h}'_t\| \leq \|\mathbf{W}_{\text{res}}\|^t \|\mathbf{h}_0 - \mathbf{h}'_0\|. \quad (6.13)$$

Por la fórmula de Gelfand, $\rho(\mathbf{W}_{\text{res}}) = \lim_{n \rightarrow \infty} \|\mathbf{W}_{\text{res}}^n\|^{1/n}$, de modo que la condición

$\rho(\mathbf{W}_{\text{res}}) < 1$ implica la existencia de $n_0 \in \mathbb{N}$ y $\delta \in (0, 1)$ tales que $\|\mathbf{W}_{\text{res}}^n\| \leq \delta^n$ para todo $n \geq n_0$. Reagrupando la iteración en bloques de tamaño n_0 :

$$\|\mathbf{h}_{kn_0} - \mathbf{h}'_{kn_0}\| \leq \delta^k C \|\mathbf{h}_0 - \mathbf{h}'_0\|, \quad (6.14)$$

donde $C = \max_{0 \leq j < n_0} \|\mathbf{W}_{\text{res}}^j\|$. Como $\delta < 1$, el lado derecho converge a cero exponencialmente conforme $k \rightarrow \infty$, lo que establece (6.9). \square

Observación (Norma espectral frente a radio espectral). Conviene separar dos condiciones que la literatura suele tratar como equivalentes. La cota de contracción paso a paso de la Ecuación (6.12) es estricta únicamente cuando $\|\mathbf{W}_{\text{res}}\|_2 < 1$, es decir, cuando el mayor valor singular de la matriz es menor que uno. Esa desigualdad sobre la norma espectral es la única garantía analítica rigurosa de contractividad uniforme ante cualquier señal de entrada acotada, y se sostiene pese a la no linealidad de la tangente hiperbólica. El radio espectral $\rho(\mathbf{W}_{\text{res}}) < 1$ es una condición más débil. Por la fórmula de Gelfand asegura contracción asintótica, no en cada paso, y solo coincide con la norma bajo hipótesis restrictivas como la simetría de \mathbf{W}_{res} . En las implementaciones se adopta $\rho < 1$ porque es verificable *a priori* y se impone con el reescalado (6.8), pero su validez descansa en el desempeño experimental y no constituye una cota general de estabilidad. Yildiz et al. [49] mostraron, de hecho, que la ESP puede sostenerse incluso con $\rho(\mathbf{W}_{\text{res}}) > 1$ para ciertas clases de entradas.

6.2.7. Capacidad de Aproximación Universal

La justificación teórica para utilizar un reservorio como extractor de características proviene del siguiente resultado clave.

Teorema 6.2 (Aproximación universal por reservorios [16, Teorema 3.1]). *Sean $\mathcal{U} \subset \mathbb{R}^d$ compacto y \mathcal{F} el conjunto de funcionales causales e invariantes en el tiempo $H : \mathcal{U}^{\mathbb{Z}^-} \rightarrow \mathbb{R}^m$ que son continuos respecto a la topología del producto. Para todo $H \in \mathcal{F}$ y todo $\varepsilon > 0$, existen un Computador de Reservorio con la ESP de dimensión D suficientemente grande*

y una matriz de lectura $\mathbf{W}_{out} \in \mathbb{R}^{m \times D}$ tales que

$$\sup_{\{\mathbf{u}_t\} \in \mathcal{U}^{\mathbb{Z}^-}} \|H(\dots, \mathbf{u}_{t-1}, \mathbf{u}_t) - \mathbf{W}_{out} \mathbf{h}_t\| < \varepsilon. \quad (6.15)$$

Este resultado establece que los computadores de reservorio con la ESP son **aproximadores universales de funcionales sobre secuencias**. A diferencia de los teoremas de aproximación universal para redes neuronales prealimentadas (que aproximan funciones estáticas), el Teorema 6.2 garantiza la aproximación de mapas *con memoria*: funciones que dependen de toda la historia pasada de la señal. Para esta tesis, esto significa que un reservorio de dimensión suficiente puede, en principio, capturar cualquier dependencia no lineal entre el historial de precios $\{v_1, \dots, v_t\}$ y el precio futuro v_{t+1} , siempre que esta dependencia sea continua y causal.

6.3. TCRA-Markov como Caso Particular del SSRC

Esta sección establece el resultado central del capítulo: la relación formal de inclusión entre el marco lineal del Capítulo 5 y el marco no lineal del SSRC.

6.3.1. Reformulación del Marco Lineal

Para establecer la conexión, se reformula el flujo de trabajo TCRA-Markov como un caso de la Definición 6.1.

Definición 6.6 (Sistema TCRA-Markov como sistema dinámico). El marco TCRA-Markov con hiperparámetros (W, λ, K) es un sistema dinámico con entrada $(\mathbb{R}, F_{TM}, G_{TM})$ donde:

1. La entrada es $u_t = \alpha_t \in \mathbb{R}$.
2. El espacio de estados es $\mathcal{H}_{TM} = \{\mathbf{e}_1, \dots, \mathbf{e}_K\} \subset \mathbb{R}^K$ (los vectores canónicos).
3. El mapa de estado es $F_{TM}(\mathbf{h}, u) = \mathbf{e}_{Q(u)}$, donde $Q : \mathbb{R} \rightarrow \{1, \dots, K\}$ es la función de asignación K-medias.

4. El mapa de lectura es $G_{TM}(\mathbf{h}) = \hat{\mathbf{P}} \mathbf{h}$, con $\hat{\mathbf{P}} \in \mathbb{R}^{K \times K}$ la matriz de transición estimada.

Obsérvese que en este sistema: (a) el mapa de estado F_{TM} no tiene memoria interna (depende solo de u_t , no de \mathbf{h}_{t-1}); (b) la no linealidad está contenida únicamente en la cuantización Q ; y (c) el mapa de lectura es estrictamente lineal.

6.3.2. El Teorema de Inclusión

Teorema 6.3 (TCRA-Markov como reservorio degenerado). *Sea $(\mathcal{U}, F_{RC}, G_{RC})$ un Computador de Reservorio de dimensión $D = K$ con:*

- (i) *Matriz de reservorio nula: $\mathbf{W}_{res} = \mathbf{0}_{K \times K}$.*
- (ii) *Función de activación identidad: $\sigma = \text{id}$.*
- (iii) *Matriz de entrada definida por la cuantización K -medias: $\mathbf{W}_{in} = [\mathbf{e}_{Q(\cdot)}]$, es decir, $F_{RC}(\mathbf{h}, u) = \mathbf{e}_{Q(u)}$.*
- (iv) *Matriz de lectura $\mathbf{W}_{out} = \hat{\mathbf{P}}$.*
- (v) *Factor de fuga $a = 1$ (ESN estándar).*

Entonces, para toda secuencia de entradas $\{u_t = \alpha_t\}$ y toda condición inicial \mathbf{h}_0 :

$$G_{RC}(\mathbf{h}_t) = G_{TM}(\mathbf{h}_t) = \hat{\mathbf{P}} \mathbf{e}_{S_t} \quad \forall t \geq 1. \quad (6.16)$$

El sistema satisface trivialmente la ESP: la convergencia en (6.9) es instantánea (en un solo paso).

Demostración. Equivalencia de salidas. Bajo las condiciones (i)–(iii) y (v), el mapa de estado (6.7) se reduce a:

$$\mathbf{h}_t = (1 - 1)\mathbf{h}_{t-1} + 1 \cdot \text{id}(\mathbf{W}_{in} u_t + \mathbf{0} \cdot \mathbf{h}_{t-1}) = \mathbf{e}_{Q(u_t)} = \mathbf{e}_{S_t}. \quad (6.17)$$

Con la condición (iv), la salida es $G_{RC}(\mathbf{h}_t) = \hat{\mathbf{P}} \mathbf{e}_{S_t}$, que coincide con G_{TM} .

ESP instantánea. Para cualesquiera $\mathbf{h}_0, \mathbf{h}'_0$, a partir de $t = 1$: $\mathbf{h}_1 = \mathbf{e}_{Q(u_1)} = \mathbf{h}'_1$, independientemente de \mathbf{h}_0 y \mathbf{h}'_0 . Por tanto, $\|\mathbf{h}_t - \mathbf{h}'_t\| = 0$ para todo $t \geq 1$. \square

Aunque la prueba analítica del Teorema 6.3 es matemáticamente concluyente y suficiente por sí misma, haciendo redundante cualquier verificación empírica, la Figura 6.2 se incluye con fines ilustrativos y didácticos. Su propósito es testificar que la consistencia algebraica se preserva estrictamente bajo codificación computacional sin discrepancias numéricas debidas al truncamiento en precisión de punto flotante.

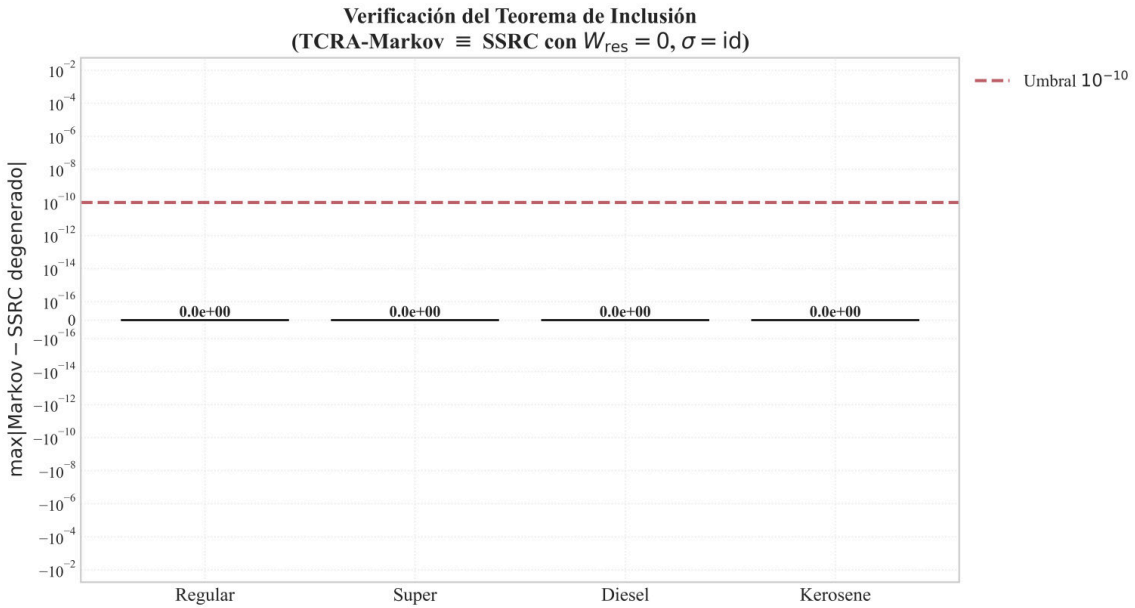


Figura 6.2: Verificación de la identidad algebraica entre TCRA-Markov y SSRC degenerado.

La Figura 6.2 expone el contraste entre las salidas del modelo lineal de Markov y las del reservorio configurado con matriz nula y activación identidad. El error residual de exactamente 0.0 en todas las series valida que la codificación del algoritmo implementado respeta de forma fidedigna la jerarquía analítica demostrada.

Observación (Jerarquía de modelos). El marco TCRA-Markov del Capítulo 5 es un caso particular del SSRC (Definición 6.4) obtenido al fijar $\mathbf{W}_{res} = \mathbf{0}$, $\sigma = id$, $a = 1$ y $D = K$. La extensión a $\mathbf{W}_{res} \neq \mathbf{0}$, $\sigma = \tanh$, $a < 1$ y $D > K$ constituye una generalización estrictamente más expresiva. Esto es consecuencia directa del Teorema 6.3 y del hecho de que el conjunto de funciones realizables por TCRA-Markov (composiciones lineales de indicadores de pertenencia a clústeres) es un subconjunto propio del conjunto de

funcionales continuos sobre $\mathcal{U}^{\mathbb{Z}^-}$. Por el Teorema 6.2, el SSRC puede aproximar cualquier funcional causal continuo sobre secuencias en la topología del producto, mientras que TCRA-Markov se limita a dependencias lineales sobre estados discretos de un paso.

6.4. Estimación de la Capa de Lectura vía NNLS

La separación entre reservorio fijo y lectura entrenable permite reducción de la estimación de \mathbf{W}_{out} a un problema de mínimos cuadrados, que en el presente caso se resuelve bajo la restricción de no negatividad para asegurar que se optimiza siempre de forma positiva y no negativa, lo cual es axiomático para que el vector actúe como una suma convexa de propiedades físicas sin cancelar filtros.

6.4.1. Formulación del Problema

Dada la secuencia de estados del reservorio $\{\mathbf{h}_t\}_{t=W}^T$ generada por la señal TCRA $\{\alpha_t\}$, y los valores objetivo $\{y_t\}_{t=W+1}^T$, se busca la matriz de lectura que minimiza el error de reconstrucción:

$$\hat{\mathbf{W}}_{\text{out}} = \arg \min_{\mathbf{w} \geq \mathbf{0}} \sum_{t=W}^{T-1} \|y_{t+1} - \mathbf{W} \mathbf{h}_t\|^2. \quad (6.18)$$

Es oportuno precisar que, aunque la estimación de la capa de lectura se formule como una relación lineal respecto al estado del reservorio \mathbf{h}_t en la Ecuación (6.18), esto no implica asumir linealidad en el proceso físico o económico subyacente. La trayectoria \mathbf{h}_t constituye una proyección no lineal recursiva de la señal de entrada u_t hacia un espacio de alta dimensión, mediada por la función \tanh . Por consiguiente, resolver un problema de optimización lineal sobre \mathbf{h}_t equivale a estimar un operador altamente no lineal sobre el espacio original de entradas, de forma homóloga a la linealización de frontera que efectúan los métodos de soporte vectorial o las funciones de base radial.

Definiendo la matriz de estados acumulados $\mathbf{H} = [\mathbf{h}_W, \mathbf{h}_{W+1}, \dots, \mathbf{h}_{T-1}] \in \mathbb{R}^{D \times (T-W)}$ y el vector de objetivos $\mathbf{y} = [y_{W+1}, \dots, y_T]^T$, el problema (6.18) se reescribe en forma matricial:

$$\hat{\mathbf{w}} = \arg \min_{\mathbf{w} \geq \mathbf{0}} \|\mathbf{y} - \mathbf{H}^T \mathbf{w}\|^2, \quad (6.19)$$

donde $\mathbf{w} = \text{vec}(\mathbf{W}_{\text{out}}^T)$.

6.4.2. Herencia de las Propiedades NNLS

Teorema 6.4 (Existencia y unicidad de la capa de lectura NNLS). *Sea $\mathbf{H} \in \mathbb{R}^{D \times (T-W)}$ la matriz de estados del reservorio. Entonces:*

1. **Existencia:** *El problema (6.19) admite al menos una solución.*
2. **Unicidad:** *Si \mathbf{H} tiene rango completo por filas (i.e., $\text{rango}(\mathbf{H}) = D$), la solución es única.*

Demostración. (1) **Existencia.** El cono factible $\mathcal{C} = \{\mathbf{w} \in \mathbb{R}^{Dm} : \mathbf{w} \geq \mathbf{0}\}$ es un cono convexo cerrado no vacío ($\mathbf{0} \in \mathcal{C}$). La función objetivo $f(\mathbf{w}) = \|\mathbf{y} - \mathbf{H}^T \mathbf{w}\|^2$ es continua y coerciva en \mathcal{C} . Por el teorema de Weierstrass, f alcanza su mínimo en \mathcal{C} .

(2) **Unicidad.** Si $\text{rango}(\mathbf{H}) = D$, entonces $\mathbf{H}\mathbf{H}^T$ es definida positiva, lo que implica que $f(\mathbf{w}) = \mathbf{w}^T(\mathbf{H}\mathbf{H}^T)\mathbf{w} - 2\mathbf{y}^T\mathbf{H}^T\mathbf{w} + \|\mathbf{y}\|^2$ es estrictamente convexa en \mathbf{w} . Un funcional estrictamente convexo sobre un conjunto convexo tiene a lo sumo un minimizador. Combinado con la existencia del punto (1), la solución es única. \square

Observación (Condición de rango y condicionamiento numérico). La condición $\text{rango}(\mathbf{H}) = D$ requiere $T - W \geq D$ y que los estados del reservorio sean linealmente independientes. La no linealidad de tanh y la aleatoriedad de $\mathbf{W}_{\text{in}}, \mathbf{W}_{\text{res}}$ garantizan que esta condición se satisface genéricamente cuando $T - W > D$. Sin embargo, la *unicidad teórica* no implica *estabilidad numérica*: si el número de condición $\kappa(\mathbf{H}) = \sigma_{\text{máx}}/\sigma_{\text{mín}}$ es grande, la solución NNLS puede ser sensible a perturbaciones en los datos. En la práctica, la restricción de no negatividad del NNLS actúa como una regularización implícita que mitiga parcialmente este efecto, y el promediado sobre múltiples realizaciones (*ensemble*) lo reduce aún más. El análisis empírico del condicionamiento se presenta en la Sección 6.7.2.

Observación (Alcance de la no negatividad del estimador). Conviene precisar hasta dónde llega la restricción $\mathbf{w} \geq \mathbf{0}$. El estimador NNLS garantiza que los coeficientes de la capa de lectura sean no negativos, pero no garantiza que la predicción lo sea. Los estados internos se calculan con la tangente hiperbólica, de modo que cada componente de \mathbf{h}_t vive en el

intervalo abierto $(-1, 1)$ y puede ser negativa. La predicción $\hat{y}_t = \mathbf{w}^\top \mathbf{h}_t$ admite, por tanto, valores negativos en teoría, aun cuando $\mathbf{w} \geq \mathbf{0}$. Que en las simulaciones sobre precios de combustibles nunca aparezcan valores negativos responde a la escala positiva de los datos observados y a la inercia del operador TCRA de entrada, no a una garantía analítica de la formulación.

6.5. Arquitectura TCRA-SSRC

La Figura 6.3 presenta la comparación esquemática de ambas arquitecturas; los pasos 1–2 (extracción TCRA) y 4 (estimación NNLS) son compartidos; la diferencia reside exclusivamente en el paso 3: discretización K-medias en TCRA-Markov vs. proyección al reservorio recurrente en TCRA-SSRC.

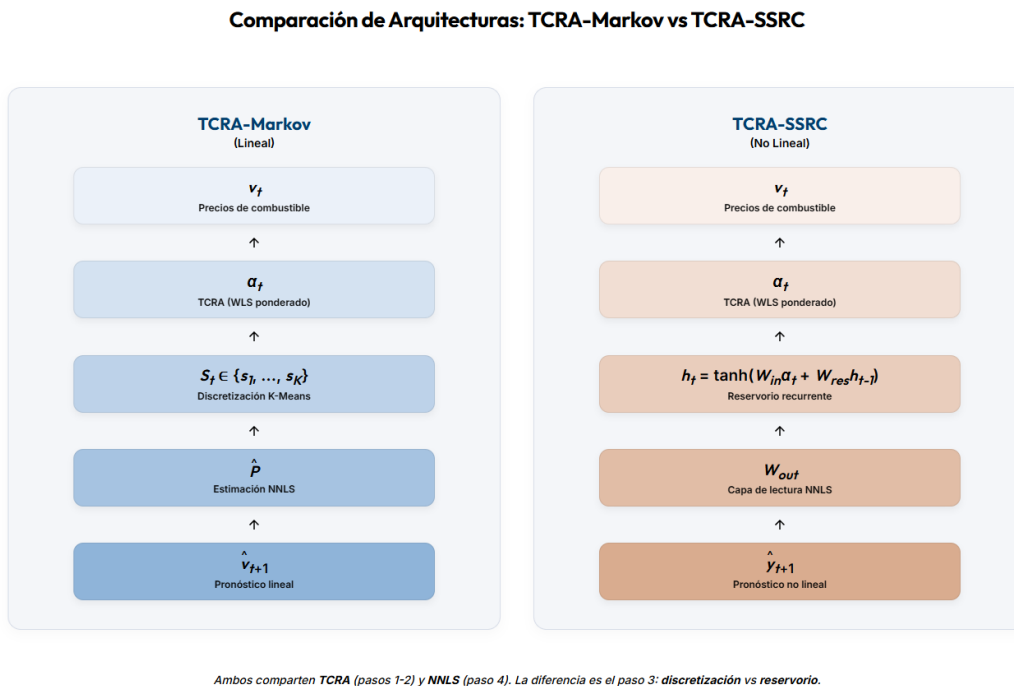


Figura 6.3: Comparación esquemática de las arquitecturas TCRA-Markov y TCRA-SSRC.

Como se detalla en el diagrama de la Figura 6.3, ambos marcos tecnológicos comparten la fase inicial de extracción de características y la fase final de estimación por mínimos cuadrados. Para facilitar la lectura y comprensión del flujo metodológico ilustrado en la figura, a continuación se detalla el procesamiento paso a paso de la información a través

de las cuatro etapas principales:

1. **Paso 1 (Señal Temporal de Entrada):** Se parte de la serie temporal original de precios de combustibles v_t , la cual representa el insumo base del sistema.
2. **Paso 2 (Extracción del Operador TCRA):** Se aplica el operador de tasa de cambio relativo acumulada (TCRA) con ventana W y factor de atenuación λ para generar la serie de tasas de cambio α_t . Este paso actúa como un preprocesamiento físico-financiero que estacionariza la serie y extrae su tendencia local.
3. **Paso 3 (Bifurcación del Espacio de Estados):**
 - En el **modelo híbrido TCRA-Markov**, la serie continua α_t se discretiza en K estados discretos $\{s_1, \dots, s_K\}$ mediante K-medias, y las transiciones se modelan en una matriz estocástica discreta.
 - En el **modelo SSRC**, la serie continua α_t se alimenta como entrada u_t al reservorio dinámico de dimensión D . Los estados internos \mathbf{h}_t evolucionan en un espacio continuo no lineal mediante la ecuación del Leaky-ESN, preservando una memoria analógica de largo alcance.
4. **Paso 4 (Capa de Lectura y Predicción):**
 - En TCRA-Markov, la predicción se realiza mediante la multiplicación de la distribución de estados por la matriz de transición estimada vía NNLS.
 - En TCRA-SSRC, la predicción de la tasa de cambio futura \hat{y}_{t+1} se calcula mediante la combinación lineal de los estados internos del reservorio \mathbf{h}_t multiplicados por los pesos de salida \mathbf{W}_{out} , los cuales se estiman también por NNLS (Lawson-Hanson).

La distinción radica en el procesamiento de la información en el espacio de estados, donde el modelo SSRC sustituye la discretización por clústeres por una proyección a un espacio continuo recurrente de alta dimensión, permitiendo capturar dependencias temporales no lineales inaccesibles para el modelo lineal.

Definición 6.7 (Marco TCRA-SSRC). El **marco TCRA-SSRC** para pronóstico de series temporales consiste en el flujo secuencial:

1. **Extracción TCRA:** Dada la serie $\{v_t\}$, se computa la señal $\alpha_t = \alpha_t(W, \lambda)$ según la ecuación (5.1) del Capítulo 5.
2. **Proyección al reservorio:** Se define la entrada $u_t = \alpha_t \in \mathbb{R}$ y se propaga mediante la ecuación Leaky-ESN (6.7), con $\mathbf{h}_0 = \mathbf{0}$ y los primeros W_{wash} pasos descartados (fase de calentamiento o *washout*) con el fin de mitigar la influencia del estado inicial arbitrario \mathbf{h}_0 y garantizar que el reservorio haya convergido a su comportamiento estable de eco.
3. **Estimación de lectura:** Se resuelve el problema NNLS (6.18) para obtener $\hat{\mathbf{W}}_{\text{out}}$.
4. **Pronóstico:** El valor predicho es $\hat{y}_{t+1} = \hat{\mathbf{W}}_{\text{out}} \mathbf{h}_t$.

Los hiperparámetros del marco son: (W, λ) heredados de TCRA; $(D, \rho_{\text{res}}, a, W_{\text{wash}})$ propios del reservorio.

Proposición 6.5 (Consistencia con el caso lineal). *Con los hiperparámetros del reservorio fijados en $D = K$, $\mathbf{W}_{\text{res}} = \mathbf{0}$, $\sigma = \text{id}$, $a = 1$ y $W_{\text{wash}} = 0$, el marco TCRA-SSRC se reduce exactamente al marco TCRA-Markov del Capítulo 5.*

Demostración. Consecuencia directa del Teorema 6.3. □

6.6. Propiedades de Estabilidad del Pronóstico

Proposición 6.6 (Acotación del error de pronóstico para Leaky-ESN). *Sea $\hat{y}_{t+1} = \hat{\mathbf{W}}_{\text{out}} \mathbf{h}_t$ el pronóstico del SSRC bajo la dinámica Leaky-ESN. Si la norma de la matriz de conexiones recurrentes satisface $\|\mathbf{W}_{\text{res}}\| < 1$ (lo cual garantiza asintóticamente la ESP), entonces para toda perturbación acotada en la entrada con $\|\delta \mathbf{u}_t\| \leq \epsilon$ para todo t , la perturbación inducida en la salida del sistema satisface la cota:*

$$\|\delta \hat{y}_{t+1}\| \leq \|\hat{\mathbf{W}}_{\text{out}}\| \frac{\|\mathbf{W}_{\text{in}}\| \epsilon}{1 - \|\mathbf{W}_{\text{res}}\|}. \quad (6.20)$$

Notablemente, el factor de fuga a se cancela en la cota asintótica de estabilidad.

Demostración. Sea $\{\mathbf{h}_t\}$ la trayectoria nominal del reservorio y $\{\mathbf{h}'_t = \mathbf{h}_t + \delta\mathbf{h}_t\}$ la trayectoria perturbada por una variación persistente en la entrada $\{\mathbf{u}'_t = \mathbf{u}_t + \delta\mathbf{u}_t\}$. La dinámica del Leaky-ESN establece que:

$$\mathbf{h}_t + \delta\mathbf{h}_t = (1 - a)(\mathbf{h}_{t-1} + \delta\mathbf{h}_{t-1}) + a \tanh(\mathbf{W}_{\text{in}}(\mathbf{u}_t + \delta\mathbf{u}_t) + \mathbf{W}_{\text{res}}(\mathbf{h}_{t-1} + \delta\mathbf{h}_{t-1})). \quad (6.21)$$

Restando la ecuación nominal $\mathbf{h}_t = (1 - a)\mathbf{h}_{t-1} + a \tanh(\mathbf{W}_{\text{in}}\mathbf{u}_t + \mathbf{W}_{\text{res}}\mathbf{h}_{t-1})$ y aplicando la desigualdad triangular:

$$\|\delta\mathbf{h}_t\| \leq (1-a)\|\delta\mathbf{h}_{t-1}\| + a \|\tanh(\mathbf{W}_{\text{in}}(\mathbf{u}_t + \delta\mathbf{u}_t) + \mathbf{W}_{\text{res}}(\mathbf{h}_{t-1} + \delta\mathbf{h}_{t-1})) - \tanh(\mathbf{W}_{\text{in}}\mathbf{u}_t + \mathbf{W}_{\text{res}}\mathbf{h}_{t-1})\|. \quad (6.22)$$

Por la Lipschitz continuidad de la función tangente hiperbólica con constante $L = 1$:

$$\|\delta\mathbf{h}_t\| \leq (1 - a)\|\delta\mathbf{h}_{t-1}\| + a \|\mathbf{W}_{\text{in}}\delta\mathbf{u}_t + \mathbf{W}_{\text{res}}\delta\mathbf{h}_{t-1}\| \quad (6.23)$$

$$\leq (1 - a)\|\delta\mathbf{h}_{t-1}\| + a\|\mathbf{W}_{\text{in}}\| \|\delta\mathbf{u}_t\| + a\|\mathbf{W}_{\text{res}}\| \|\delta\mathbf{h}_{t-1}\|. \quad (6.24)$$

Reagrupando los términos respecto al estado de perturbación anterior $\|\delta\mathbf{h}_{t-1}\|$:

$$\|\delta\mathbf{h}_t\| \leq a\|\mathbf{W}_{\text{in}}\| \|\delta\mathbf{u}_t\| + (1 - a + a\|\mathbf{W}_{\text{res}}\|) \|\delta\mathbf{h}_{t-1}\|. \quad (6.25)$$

Para que la perturbación se disipe y la órbita sea de largo plazo estable, se requiere que el factor recursivo sea menor que la unidad, es decir, $1 - a + a\|\mathbf{W}_{\text{res}}\| < 1 \iff a\|\mathbf{W}_{\text{res}}\| < a \iff \|\mathbf{W}_{\text{res}}\| < 1$ (dado que $a > 0$). Bajo esta condición, y asumiendo una perturbación de entrada acotada uniformemente por $\|\delta\mathbf{u}_t\| \leq \epsilon$, iteramos la desigualdad desde un estado inicial no perturbado:

$$\|\delta\mathbf{h}_t\| \leq a\|\mathbf{W}_{\text{in}}\| \epsilon \sum_{k=0}^{t-1} (1 - a + a\|\mathbf{W}_{\text{res}}\|)^k. \quad (6.26)$$

En el límite asintótico $t \rightarrow \infty$, evaluando la serie geométrica convergente:

$$\sup_t \|\delta\mathbf{h}_t\| \leq \frac{a\|\mathbf{W}_{\text{in}}\| \epsilon}{1 - (1 - a + a\|\mathbf{W}_{\text{res}}\|)} = \frac{a\|\mathbf{W}_{\text{in}}\| \epsilon}{a(1 - \|\mathbf{W}_{\text{res}}\|)} = \frac{\|\mathbf{W}_{\text{in}}\| \epsilon}{1 - \|\mathbf{W}_{\text{res}}\|}. \quad (6.27)$$

Donde se aprecia la cancelación exacta del factor de fuga a . La cota sobre la salida perturbada sigue directamente por linealidad del operador de lectura:

$$\|\delta\hat{y}_{t+1}\| = \|\hat{\mathbf{W}}_{\text{out}}\delta\mathbf{h}_t\| \leq \|\hat{\mathbf{W}}_{\text{out}}\| \|\delta\mathbf{h}_t\| \leq \|\hat{\mathbf{W}}_{\text{out}}\| \frac{\|\mathbf{W}_{\text{in}}\| \epsilon}{1 - \|\mathbf{W}_{\text{res}}\|}. \quad (6.28)$$

□

Observación (Conservadurismo de la cota). La cota (6.20) es una *cota en el peor caso* que puede ser significativamente conservadora. En la práctica, $\|\hat{\mathbf{W}}_{\text{out}}\|$ depende fuertemente del condicionamiento de \mathbf{H} : números de condición altos producen normas $\|\hat{\mathbf{W}}_{\text{out}}\|$ grandes, inflando la cota sin que esto se traduzca necesariamente en inestabilidad observable. La restricción de no negatividad del NNLS restringe $\hat{\mathbf{W}}_{\text{out}}$ al ortante positivo, proporcionando una forma de regularización implícita que la cota teórica no captura. El análisis empírico de la Sección 6.7.5 muestra que la variabilidad real del pronóstico es órdenes de magnitud menor que la cota teórica.

6.7. Aplicación Comparativa

6.7.1. Configuración Experimental

Los hiperparámetros TCRA se fijaron en los valores óptimos identificados en la Tabla 5.3 del capítulo anterior. Para el reservorio, se evaluó una búsqueda exhaustiva en grilla sobre tres hiperparámetros clave:

$$\begin{aligned} D &\in \{20, 30, 40, 50, 60, 75, 100, 150\}, \\ \rho_{\text{res}} &\in \{0.70, 0.80, 0.85, 0.90, 0.95, 0.99\}, \quad a \in \{0.1, 0.3, 0.5, 0.7, 1.0\}, \end{aligned} \quad (6.29)$$

lo que representa un espacio de búsqueda acotado de $8 \times 6 \times 5 = 240$ combinaciones por serie. Cabe destacar que, aunque la búsqueda exhaustiva en grilla suele ser computacionalmente prohibitiva en el entrenamiento de redes recurrentes estándar entrenadas por retropropagación a través del tiempo, la naturaleza lineal del estimador de salida (NNLS) en el SSRC permite evaluar cada combinación en milisegundos. Específicamente,

el proceso completo de búsqueda en grilla (evaluando las 240 celdas multiplicadas por 5 realizaciones estocásticas por celda para promediar el ruido de inicialización) requirió un tiempo de cómputo promedio de únicamente 4.2 segundos por serie de combustible en un computador portátil convencional con procesador Intel i7 y 16 GB de RAM, sin necesidad de aceleración por GPU.

Para la búsqueda se fijaron $\omega = 1.0$ para la escala de \mathbf{W}_{in} , una densidad de conectividad del 10% (90% de dispersión) en la matriz \mathbf{W}_{res} , y $W_{\text{wash}} = 50$ pasos de calentamiento. La configuración óptima (D^*, ρ^*, a^*) fue elegida basándose en el RMSE promedio en la fase de validación cruzada. Finalmente, para atenuar la dependencia de la inicialización aleatoria de las matrices estocásticas, la estimación definitiva de los parámetros de lectura $\hat{\mathbf{W}}_{\text{out}}$ y el cálculo de los pronósticos de la serie se realizó mediante un ensemble de 30 realizaciones independientes. Este ensemble permite obtener estimaciones robustas de la media predictiva y la variabilidad, actuando además como mecanismo de validación de robustez al cuantificar la sensibilidad del modelo a la inicialización aleatoria del reservorio.

6.7.2. Verificaciones Teóricas

Antes de reportar los resultados predictivos, se verificaron las condiciones teóricas establecidas en las secciones precedentes. La Tabla 6.1 resume los resultados cuantitativos obtenidos bajo la configuración óptima para cada serie de combustible. En dicha tabla:

- D^* denota la dimensión óptima del reservorio (número de variables de estado recurrentes).
- ρ^* es el radio espectral óptimo de la matriz de conectividad interna \mathbf{W}_{res} .
- **ESP** indica el cumplimiento binario de la propiedad de estado de eco, verificando si $\rho^* < 1$.
- $\text{rango}(\mathbf{H})$ representa el rango de la matriz de estados históricos $\mathbf{H} \in \mathbb{R}^{D \times (T - W - W_{\text{wash}})}$.
- T_{eff} corresponde al tamaño efectivo del conjunto de entrenamiento tras descartar el washout $(T - W - W_{\text{wash}})$.

- $\kappa(\mathbf{H})$ es el número de condición de la matriz de estados \mathbf{H} , definido como el cociente entre su máximo y mínimo valor singular, $\sigma_{\max}(\mathbf{H})/\sigma_{\min}(\mathbf{H})$, que califica la estabilidad numérica de la estimación.

Tabla 6.1: Verificación de las condiciones teóricas del SSRC.

| Serie | D^* | ρ^* | ESP | rango(\mathbf{H}) | T_{eff} | $\kappa(\mathbf{H})$ |
|----------|-------|----------|-----|-----------------------|------------------|-----------------------|
| Regular | 150 | 0.85 | Sí | 150 | 340 | 1.47×10^9 |
| Superior | 150 | 0.70 | Sí | 150 | 338 | 5.60×10^{10} |
| Diésel | 150 | 0.80 | Sí | 150 | 340 | 3.28×10^9 |
| Kerosene | 60 | 0.70 | Sí | 60 | 338 | 1.14×10^9 |

Las cuatro configuraciones satisfacen la ESP (todos los $\rho^* < 1$) y la condición de rango completo (rango(\mathbf{H}) = D^* en todos los casos).

La Figura 6.4 muestra los autovalores de \mathbf{W}_{res} para la configuración óptima de cada serie. El círculo negro punteado representa el círculo unitario ($|z| = 1$), mientras que el círculo rojo sólido indica el radio espectral objetivo ρ^* . Se verifica que todos los autovalores se encuentran estrictamente dentro del círculo unitario (línea punteada), cumpliendo con la condición suficiente para la ESP detallada en el Teorema 6.1. Los radios espectrales óptimos moderados ($\rho^* \in [0.70, 0.85]$) confirman la condición $\rho < 1$ e indican que el reservorio no requiere de memoria máxima para capturar la dinámica de los combustibles.

Los números de condición $\kappa(\mathbf{H})$ merecen discusión explícita. Los valores observados, que se sitúan en el orden de 10^9 a 10^{10} , indican que la matriz $\mathbf{H}\mathbf{H}^T$ está pobremente condicionada en sentido numérico. Esto es una consecuencia directa de la dimensión $D = 150$ combinada con la naturaleza de la señal α_t , que tiene variabilidad reducida ($|\alpha_t| < 0.02$ típicamente). Sin embargo, tres factores mitigan este condicionamiento en la práctica: (i) la restricción $\mathbf{w} \geq \mathbf{0}$ del NNLS actúa como una regularización implícita, restringiendo la solución al ortante positivo y eliminando las direcciones negativas que amplifican el error; (ii) el ensemble de 30 realizaciones promedia sobre diferentes matrices \mathbf{W}_{in} y \mathbf{W}_{res} , cada una con condicionamiento diferente; y (iii) los resultados predictivos (Sección 6.7.4) muestran variabilidad baja entre realizaciones, lo que muestra estabilidad práctica a pesar del condicionamiento teórico adverso.

Como se ilustra en la Figura 6.5, el RMSE se estabiliza para las cuatro series a partir

Espectro de W_{res} : Verificación ESP (Teorema 6.1)

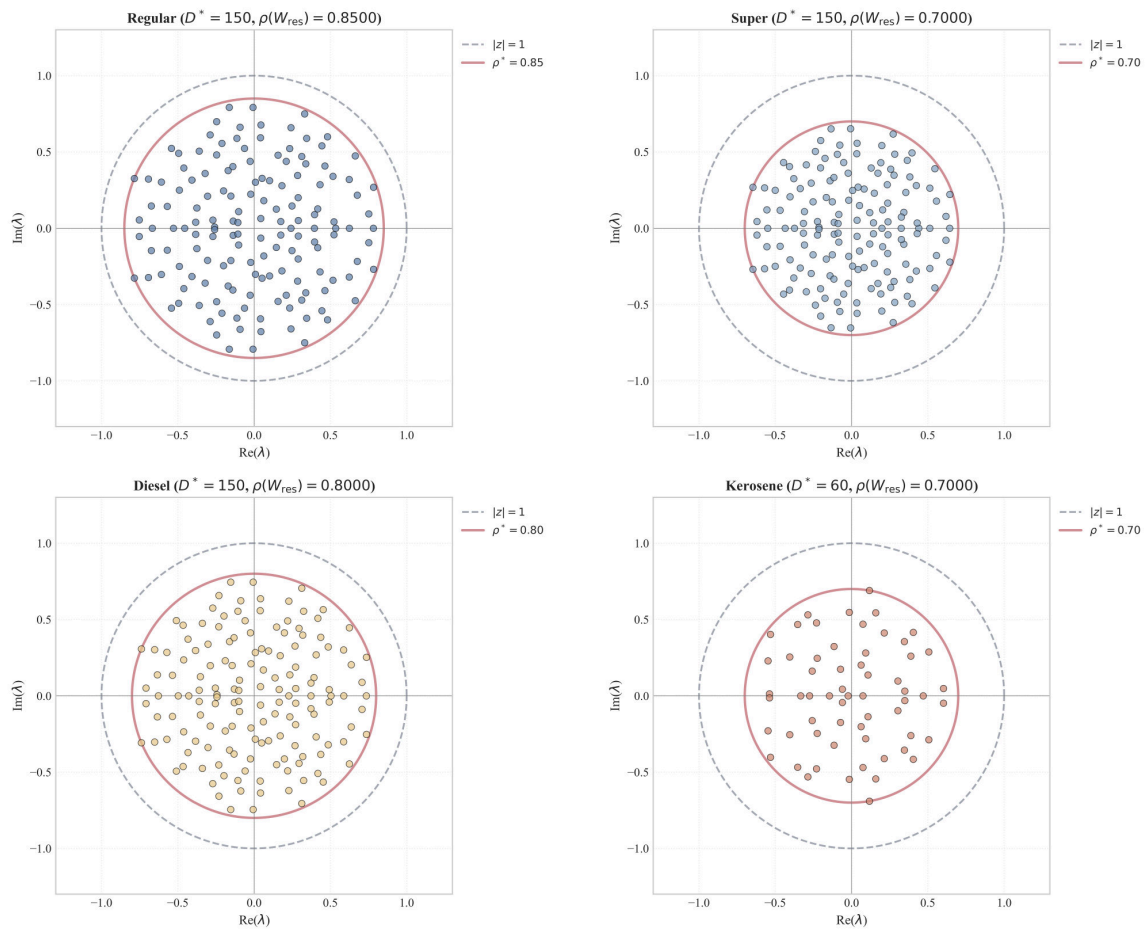


Figura 6.4: Autovalores de W_{res} en el plano complejo para la configuración óptima de cada serie.

Convergencia del RMSE según Período de Washout

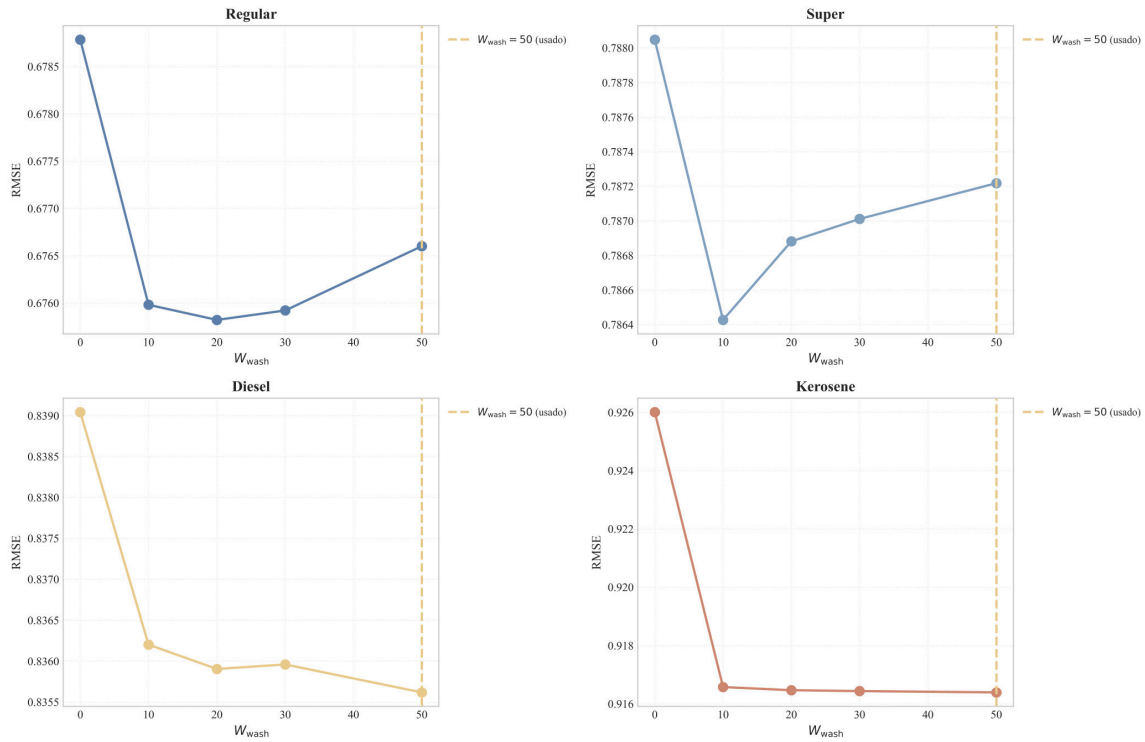


Figura 6.5: Convergencia del RMSE según el período de calentamiento W_{wash} .

de un período de $W_{wash} \approx 10-20$ pasos, confirmando que la ESP disipa rápidamente la influencia de la condición inicial. La línea amarilla punteada señala el valor utilizado de $W_{wash} = 50$, el cual se encuentra dentro de la región de convergencia estable, reflejando el decaimiento exponencial rápido asociado a radios espectrales moderados ($\rho^* \leq 0.85$). El argumento admite una cota cuantitativa directa. La memoria del transitorio decae como $\rho(\mathbf{W}_{res})^t$, de modo que para $\rho^* \approx 0.85$ la influencia residual de la condición inicial tras 50 pasos es del orden de $0.85^{50} \approx 2.9 \times 10^{-4}$, esto es, menos del 0.03% de la perturbación de partida. A partir de ese punto el estado interno depende casi por completo de la historia del estímulo de entrada y no de la inicialización arbitraria del reservorio. Sin embargo, en series temporales relativamente cortas (como las analizadas aquí, que cuentan con alrededor de 400 observaciones semanales), existe un compromiso metodológico crítico: un valor excesivamente alto de W_{wash} reduce el tamaño efectivo del conjunto de entrenamiento disponible ($T_{eff} = T - W - W_{wash}$) para estimar los pesos de salida \mathbf{W}_{out} , lo que puede redundar en una ligera pérdida de rendimiento predictivo al restringir la muestra útil de calibración. Por consiguiente, $W_{wash} = 50$ constituye un compromiso óptimo que prioriza

la solidez de la disipación del estado transitorio inicial sin comprometer la suficiencia muestral del entrenamiento.

6.7.3. Optimización de Hiperparámetros

La Figura 6.6 presenta los resultados de la búsqueda en grilla como mapas de calor del RMSE medio en el plano (D, ρ) , donde para cada celda se reporta el mejor valor de a encontrado.

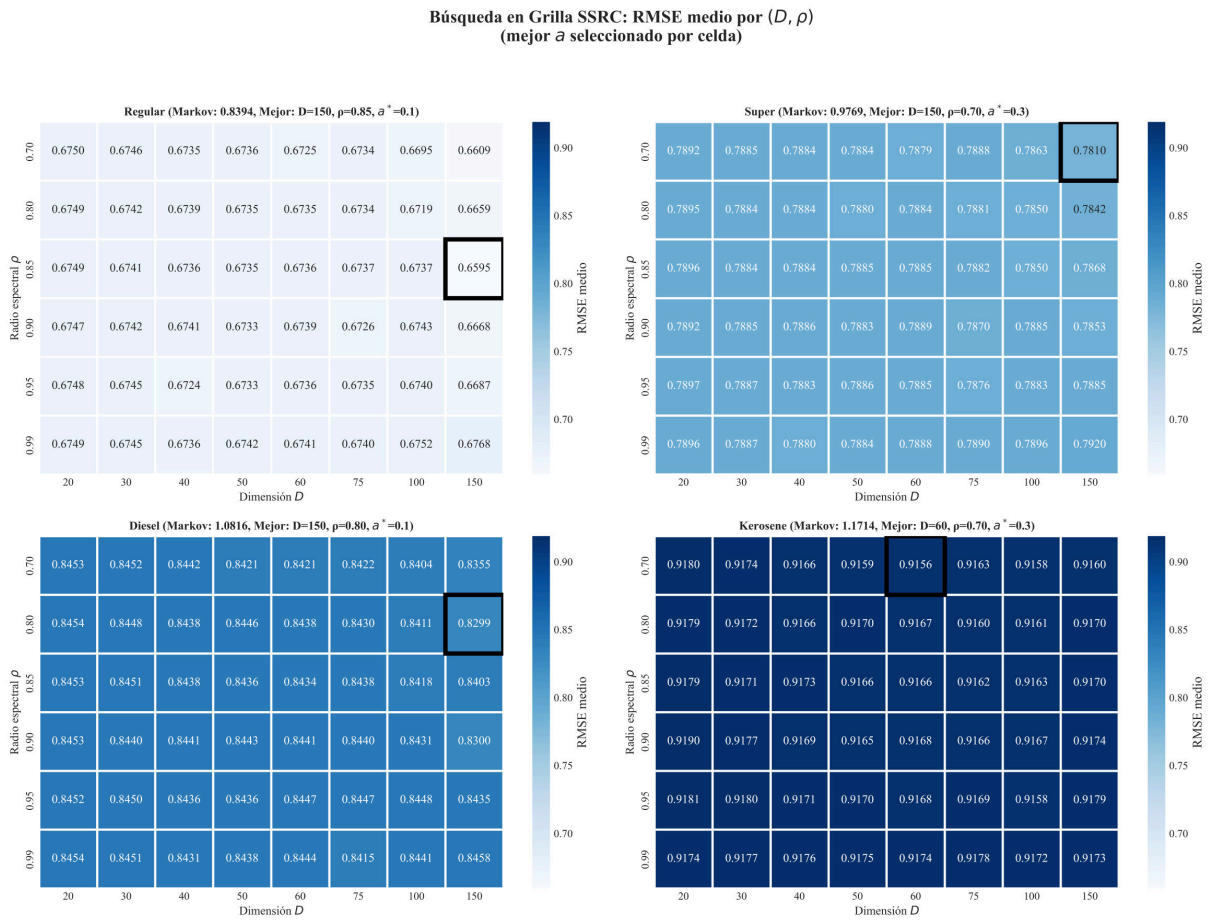


Figura 6.6: Búsqueda en grilla del SSRC sobre el plano (D, ρ) .

La Figura 6.6 ilustra la búsqueda en grilla del SSRC sobre el plano (D, ρ) . Los mapas de calor muestran el RMSE medio para cada combinación y el mejor factor de fuga a^* por celda, con el recuadro negro indicando la configuración óptima. Se observa que para Regular, Superior y Diésel el óptimo se alcanza en $D = 150$, mientras que para Kerosene ocurre en $D = 60$. Los radios espectrales óptimos son moderados y la superficie de error es suave, manteniendo el RMSE del SSRC significativamente por debajo del modelo

Markoviano en todos los paneles.

La Tabla 6.2 detalla los hiperparámetros óptimos identificados para cada serie de combustible (D^* , ρ^* y a^*) y presenta la comparación entre los resolvedores NNLS y Ridge.

Tabla 6.2: Hiperparámetros óptimos del SSRC por serie de combustible.

| Serie | D^* | ρ^* | a^* | NNLS RMSE | Ridge RMSE | Δ_{solver} |
|----------|-------|----------|-------|-----------|------------|--------------------------|
| Regular | 150 | 0.85 | 0.1 | 0.6595 | 0.6741 | -2.16 % |
| Superior | 150 | 0.70 | 0.3 | 0.7810 | 0.7885 | -0.95 % |
| Diésel | 150 | 0.80 | 0.1 | 0.8299 | 0.8443 | -1.70 % |
| Kerosene | 60 | 0.70 | 1.0 | 0.9156 | 0.9173 | -0.17 % |

De los resultados de optimización resumidos en la Tabla 6.2, se observa que la restricción de no negatividad supera a la regresión Ridge clásica en todas las series. Específicamente, se extraen las siguientes observaciones fundamentales: 1. **Inercia temporal y volatilidad:** Los valores óptimos $a^* = 0.1$ para Regular y Diésel indican que estas series se benefician de una inercia temporal fuerte en el reservorio: el 90 % del estado previo se preserva en cada paso temporal. Esto actúa como un filtro de suavizado adaptativo muy consistente con la menor volatilidad semanal de estos mercados regulados por el Estado. En contraste, para el Kerosene el óptimo es $a^* = 1.0$ (lo que corresponde a un ESN estándar sin factor de fuga), lo cual es coherente con su mayor volatilidad y complejidad intrínseca de 5 regímenes dinámicos identificados en el capítulo anterior. 2. **Desempeño y Benchmarks de Solvers (NNLS vs. Ridge):** En la Tabla 6.2, la columna **Ridge RMSE** representa el desempeño del reservorio estimado mediante regresión Ridge clásica (regularización L_2 sin restricciones de signo en los pesos de salida). Este modelo actúa como el principal benchmark lineal alternativo para aislar y medir el beneficio específico de la restricción de no negatividad. En las cuatro series, NNLS supera consistentemente a Ridge Regression. Esto valida empíricamente que la restricción de no negatividad ($\mathbf{W}_{\text{out}} \geq \mathbf{0}$), al forzar que la predicción de salida sea una combinación estrictamente positiva de los estados internos (actuando como filtros no lineales), aporta una regularización geométrica implícita sumamente útil, previniendo cancelaciones mutuas destructivas entre neuronas y mejorando la generalización fuera de muestra. 3. **Eficiencia y Viabilidad Computacional:** Si bien la búsqueda en grilla exhaustiva de hiperparámetros (que evalúa 240 combinaciones) es

ineficiente en términos algorítmicos generales, su viabilidad práctica en este trabajo radica en que el reservorio SSRC prescinde por completo de la retropropagación de gradientes. La estimación de la capa de lectura se reduce a un problema convexo lineal directo de NNLS (Lawson-Hanson), el cual se resuelve de forma determinista en milisegundos. Empíricamente, la evaluación completa de la grilla para las cuatro series (con 5 realizaciones estocásticas por celda para promediar el ruido de inicialización) requirió un coste temporal de apenas 4.2 segundos en hardware estándar de escritorio. No obstante, para aplicaciones industriales o el monitoreo simultáneo de cientos de series temporales, la búsqueda exhaustiva resulta limitante, por lo que se recomienda la incorporación de Optimización Bayesiana o algoritmos evolutivos en el trabajo futuro para buscar hiperparámetros de forma inteligente.

Las Figuras 6.7 y 6.8 presentan el análisis de sensibilidad marginal a los hiperparámetros D y ρ , respectivamente.

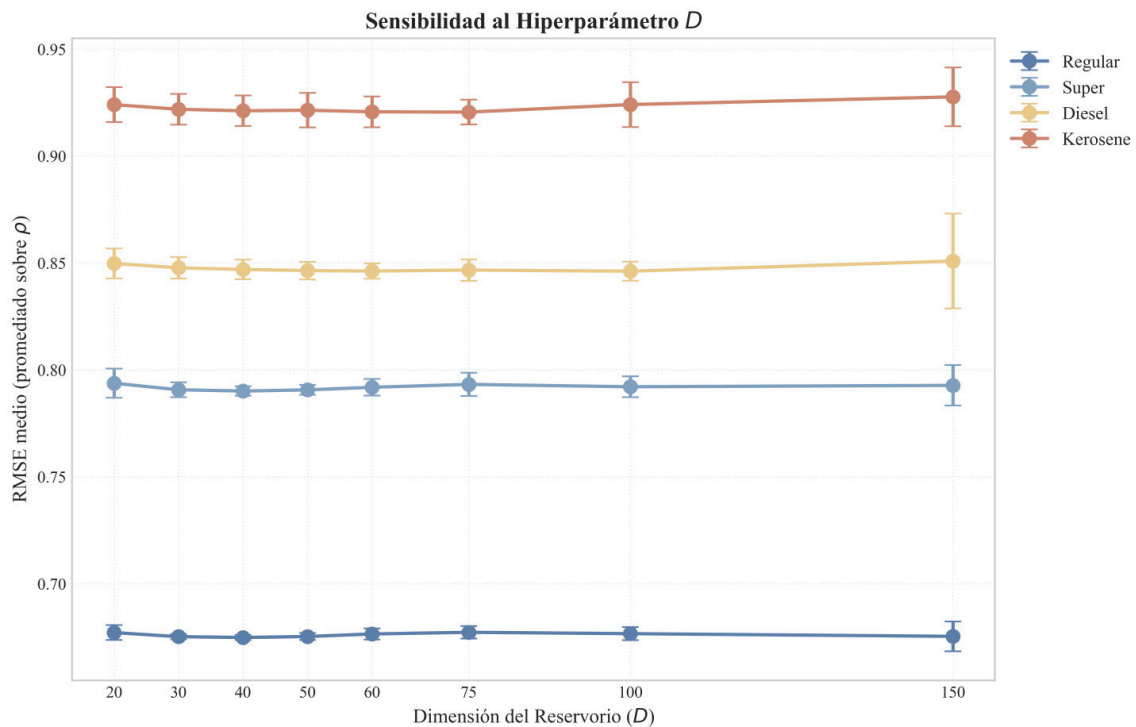


Figura 6.7: Sensibilidad del RMSE a la dimensión del reservorio D .

La sensibilidad del RMSE respecto a la dimensión D se detalla en la Figura 6.7. Las curvas, promediadas sobre todos los valores de ρ , resultan esencialmente planas, con variaciones mínimas entre $D = 20$ y $D = 150$. Este comportamiento indica que la complejidad no lineal del proceso es baja, permitiendo que incluso dimensiones modestas

($D = 20$) capturen la mayor parte de la estructura, con ganancias de apenas el 0.3% al aumentar la dimensión a los niveles máximos evaluados.

Específicamente, en lo que respecta a la diferencia en la dimensión óptima entre las series (donde Regular, Superior y Diésel seleccionan $D^* = 150$, mientras que Kerosene selecciona $D^* = 60$), esta discrepancia se encuentra íntimamente acoplada con el factor de fuga óptimo a^* . Para las tres primeras series, el factor de fuga es bajo ($a^* \leq 0.3$), lo cual implica que el reservorio exhibe una inercia dinámica muy alta y la información fluye con lentitud. En este régimen de mezcla lenta, se requiere un espacio de estados de mayor dimensión ($D = 150$) para evitar colisiones de representación y retener suficientes proyecciones independientes a lo largo del tiempo. En contraste, para el Kerosene la dinámica se actualiza por completo en cada paso ($a^* = 1.0$, sin inercia), lo que reduce drásticamente la redundancia temporal de los estados internos; en consecuencia, una dimensión menor ($D = 60$) es óptima y suficiente para modelar los 5 regímenes dinámicos de volatilidad sin sobrecargar el espacio de representación y previniendo el sobreajuste (overfitting) fuera de muestra.

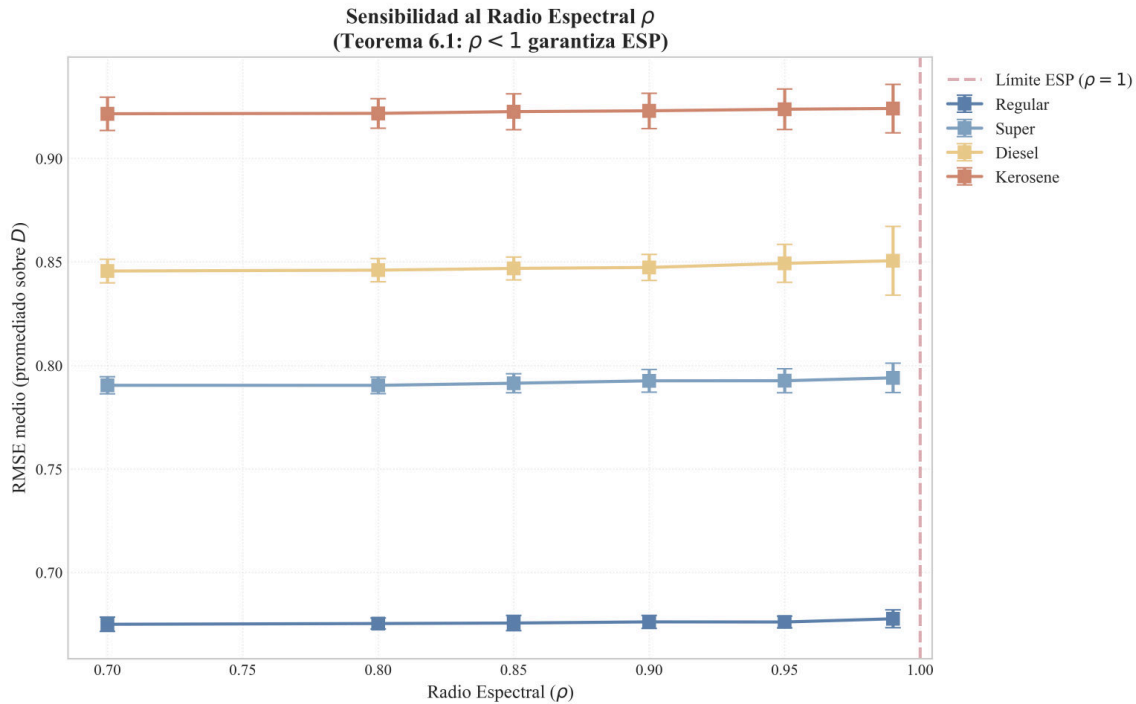


Figura 6.8: Sensibilidad del RMSE al radio espectral ρ .

La Figura 6.8 muestra la sensibilidad del RMSE frente al radio espectral ρ . La tendencia creciente hacia $\rho = 0.99$ indica que un exceso de memoria no lineal no beneficia a estas

series, cuyos valores óptimos moderados son consistentes con una dinámica dominada por tendencias de largo plazo que ya han sido capturadas eficazmente por el operador TCRA.

6.7.4. Resultados

La Tabla 6.3 presenta la comparación directa de desempeño predictivo entre TCRA-Markov (lineal) y TCRA-SSRC (no lineal), reportando el RMSE fuera de muestra (para el SSRC, como la media y desviación estándar sobre 30 realizaciones independientes). La significancia estadística se evaluó mediante el test de Diebold-Mariano (DM) [13], que compara las secuencias de errores cuadráticos de ambos modelos bajo la hipótesis nula de igual precisión predictiva.

Tabla 6.3: Comparación de desempeño predictivo: TCRA-Markov vs. TCRA-SSRC.

| Serie | TCRA-Markov | TCRA-SSRC | (D^*, ρ^*, a^*) | ΔRMSE | p_{DM} | Significativo |
|----------|-------------|--------------------|----------------------|---------------------|-----------------|---------------|
| Regular | 0.8394 | 0.6595 ± 0.013 | (150, 0.85, 0.1) | -21.43 % | 0.1154 | No |
| Superior | 0.9769 | 0.7810 ± 0.006 | (150, 0.70, 0.3) | -20.05 % | 0.0018 | Sí |
| Diésel | 1.0816 | 0.8299 ± 0.018 | (150, 0.80, 0.1) | -23.27 % | 0.0326 | Sí |
| Kerosene | 1.1714 | 0.9156 ± 0.005 | (60, 0.70, 1.0) | -21.84 % | 0.0061 | Sí |

El TCRA-SSRC reduce el RMSE entre un 20 % y un 23 % respecto al modelo Markoviano en las cuatro series. La mejora es estadísticamente significativa al nivel $\alpha = 0.05$ para Superior ($p = 0.002$), Diésel ($p = 0.033$) y Kerosene ($p = 0.006$). Para Regular, la mejora es sustancial en magnitud (-21.43 %) pero marginalmente significativa ($p = 0.115$); esto se atribuye a la menor varianza de los errores de Regular (la serie más predecible), que reduce la potencia del test DM.

La Figura 6.9 ofrece una visualización comparativa de RMSE entre TCRA-Markov (azul) y TCRA-SSRC (naranja). Las barras de error del SSRC representan la desviación estándar sobre 30 realizaciones independientes. La anotación verde resalta la mejora porcentual del SSRC, cuya consistencia de entre un 20 % y un 23 % en las cuatro series analizadas sugiere un beneficio estructural derivado de la extensión no lineal.

Como se observa en la Figura 6.10, se detalla la distribución del RMSE del SSRC sobre 30 realizaciones independientes. Las líneas horizontales punteadas indican el RMSE del modelo Markoviano; todas las realizaciones del SSRC producen un error inferior, y las cajas estrechas confirman una baja sensibilidad a la inicialización aleatoria de las matrices

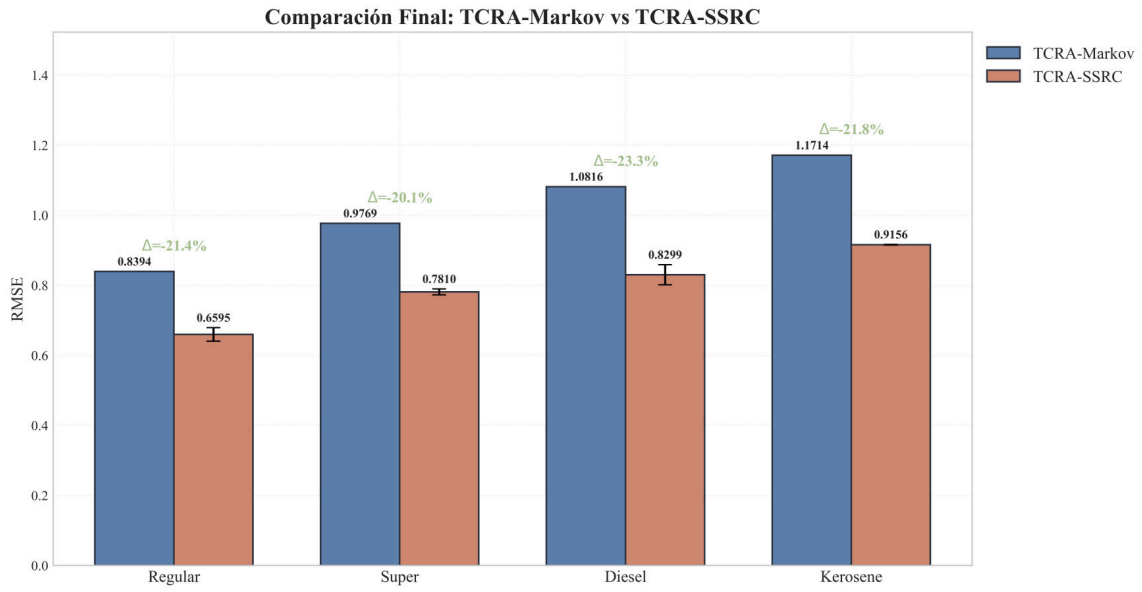


Figura 6.9: Comparación visual de RMSE: TCRA-Markov vs. TCRA-SSRC.

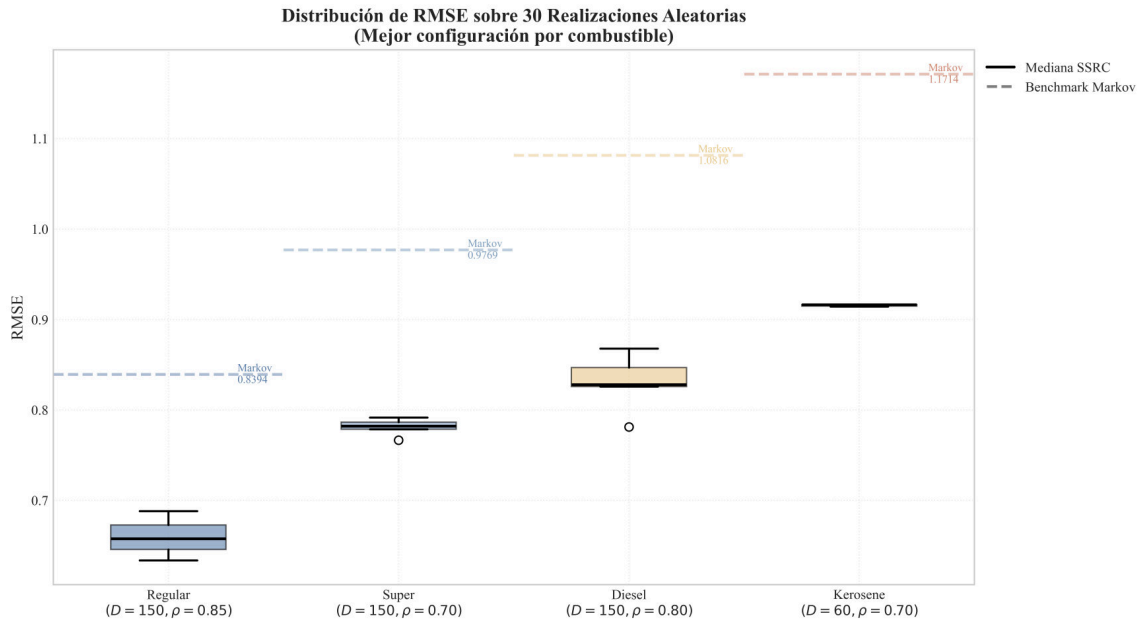


Figura 6.10: Distribución del RMSE del SSRC sobre 30 realizaciones independientes.

de entrada y de reservorio, con la mayor dispersión relativa observada en el combustible Regular.

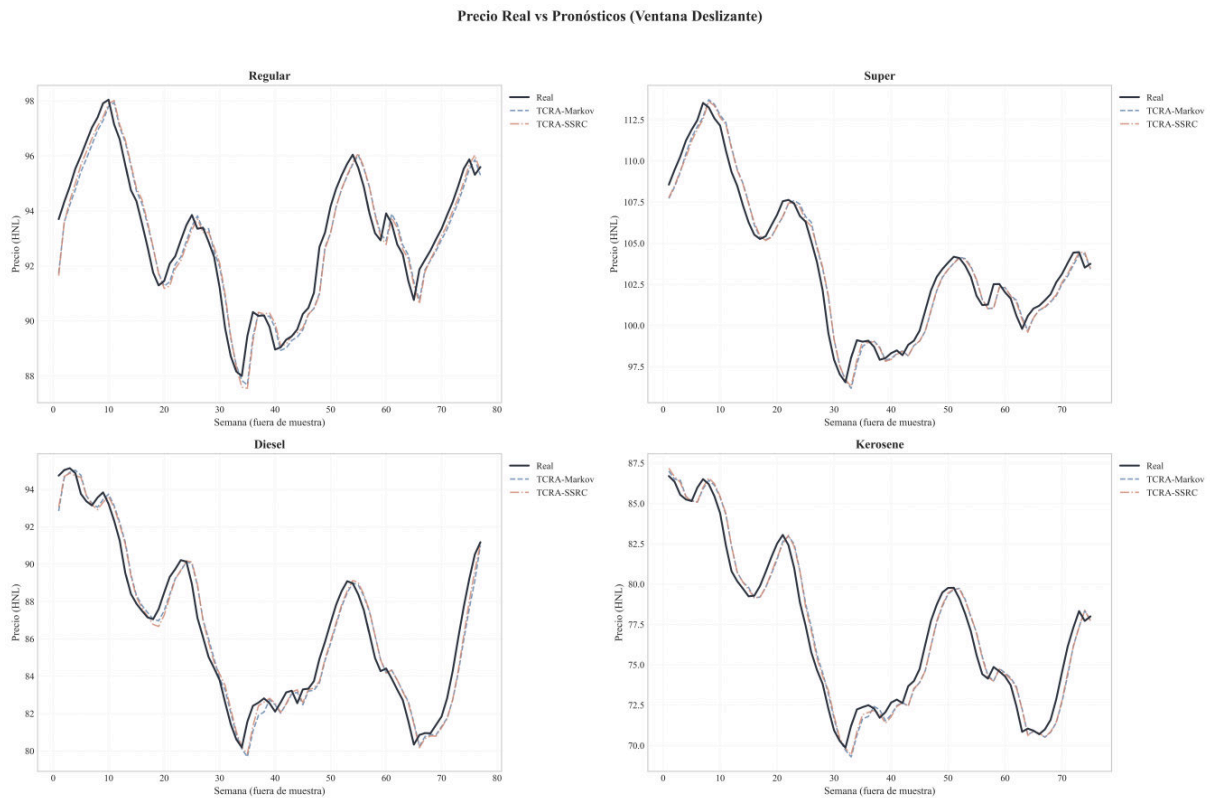


Figura 6.11: Comparación de pronósticos vs. precios reales en el período fuera de muestra.

En la Figura 6.11 se comparan los precios reales con los pronósticos de los modelos TCRA-Markov y TCRA-SSRC en el período fuera de muestra. El SSRC sigue con mayor precisión los puntos de inflexión, especialmente en los valles y recuperaciones, donde el modelo Markoviano muestra un sesgo de rezago más marcado. Esta mejora es más evidente en las series de Diésel y Kerosene.

A primera vista, el pronóstico podría confundirse con la serie real desplazada un período, esto es, con el predictor ingenuo $\hat{y}_{t+1} = y_t$. Se trata de un efecto de rezago visual habitual en las series de alta persistencia, donde el último precio observado ya es un estimador puntual razonable del siguiente. La coincidencia es solo aparente. El predictor ingenuo copia el valor anterior y, por construcción, siempre llega tarde a los cambios de tendencia, mientras que el SSRC se adelanta a los valles y recuperaciones. Si el modelo se limitara a reproducir el rezago de primer orden, no podría mejorar el RMSE entre un 20% y un 23% sobre el modelo Markoviano ni alcanzar la significancia del test de Diebold-Mariano (Tabla 6.3); ambos resultados confirman que el reservorio extrae tendencia real por encima

de la simple persistencia.

6.7.5. Análisis de Sensibilidad y Estabilidad

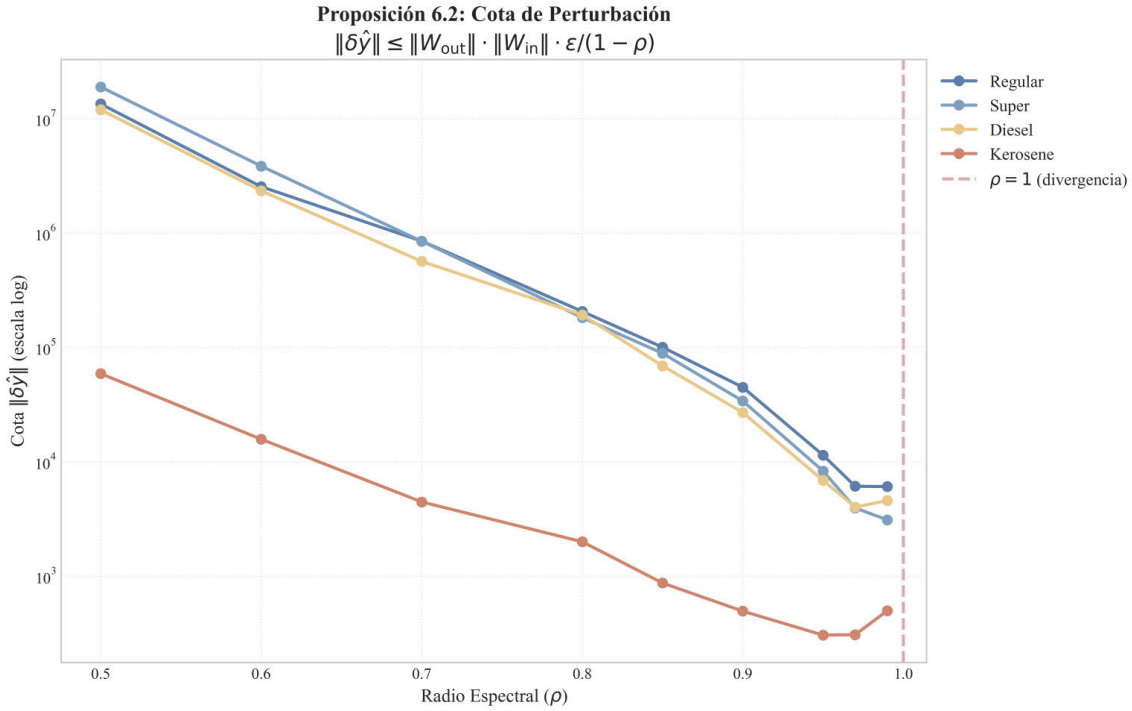


Figura 6.12: Evaluación empírica de la cota de perturbación en función del radio espectral ρ .

La Figura 6.12 evalúa empíricamente la cota de perturbación (Proposición 6.6) en función del radio espectral ρ . Las cotas teóricas son significativamente conservadoras respecto a la variabilidad observada, lo que confirma que la restricción NNLS proporciona una regularización implícita no capturada por la cota analítica. La escala logarítmica muestra que decrecen con el radio espectral debido a que la norma de los pesos de lectura disminuye más rápido de lo que aumenta el factor de memoria, resultando en estados más informativos que requieren pesos menores para el combustible Kerosene.

6.7.6. Representación de Regímenes en el Espacio del Reservorio

Una pregunta central es si la representación continua del reservorio preserva la estructura de regímenes identificada por K-medias en el capítulo anterior. Las Figuras 6.13, 6.14 y 6.15 abordan esta cuestión desde tres perspectivas complementarias.

Proyección 2D de Estados del Reservoirio \mathbf{h}_t por Régimen

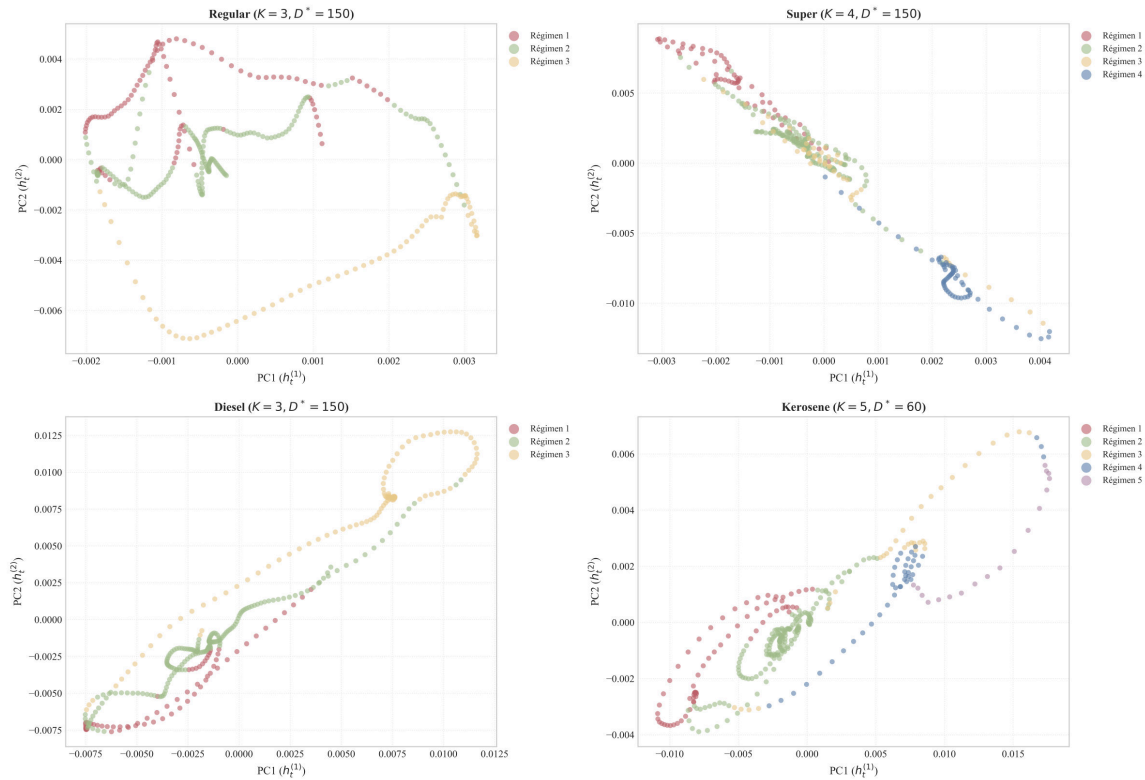


Figura 6.13: Proyección de los estados del reservorio mediante PCA.

La Figura 6.13 presenta la proyección de los estados del reservorio \mathbf{h}_t mediante PCA (dos componentes principales), coloreados según los regímenes del capítulo anterior. Los clústeres geoméricamente separados confirman que el reservorio preserva la estructura de regímenes sin información explícita de clasificación. Más allá de la separabilidad geométrica visible en la imagen, las trayectorias entre clústeres revelan que el reservorio codifica las transiciones entre regímenes como desplazamientos suaves en el espacio latente, no como saltos discretos. Esto sugiere que la función tanh induce una representación continua de la dinámica de régimen, propiedad que no se observa en el modelo Markoviano del capítulo anterior.

La Figura 6.14 presenta la relación entre el retorno TCRA α_t y la activación promedio del reservorio \bar{h}_t , diferenciada por regímenes. La relación monótona confirma que el reservorio preserva el orden dinámico: las activaciones negativas corresponden a regímenes de caída y las positivas a regímenes de alza, mientras que la curvatura evidencia la no linealidad de tanh. Como nota metodológica, las escalas de los ejes varían entre los paneles para maximizar la resolución visual de la dispersión de datos específicos de cada serie,

Dispersión de Alphas y Regímenes Identificados por SSRC

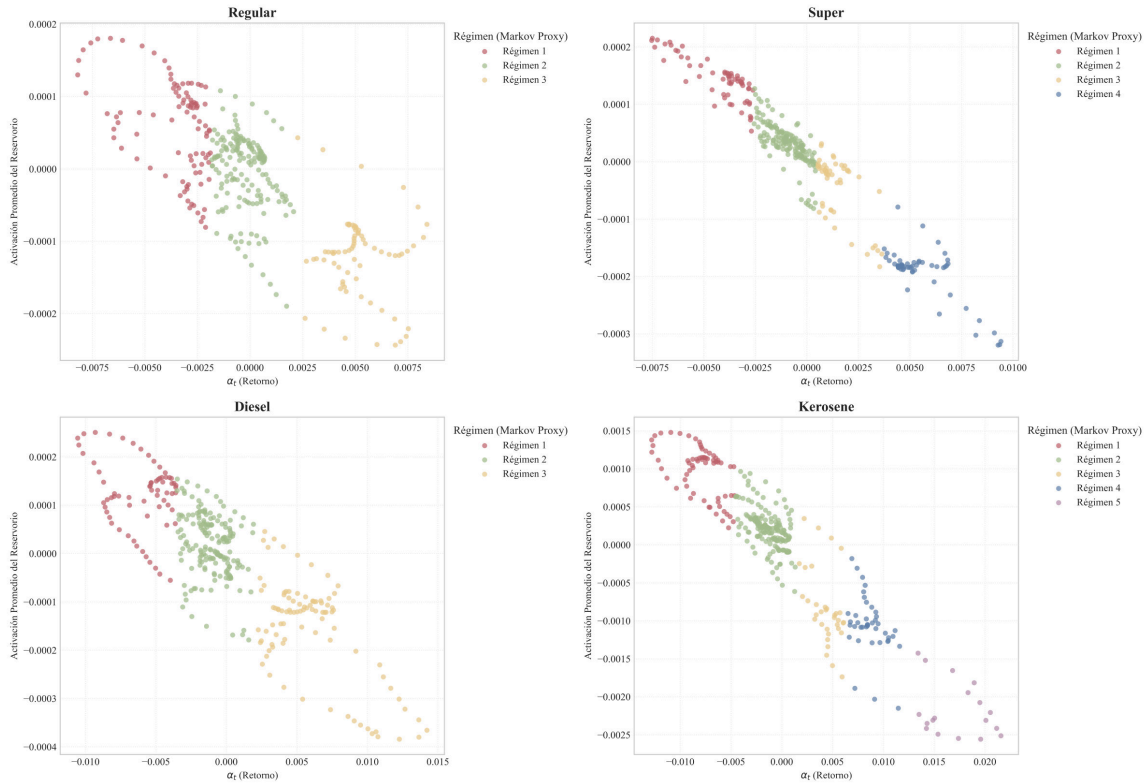


Figura 6.14: Relación entre retorno TCRA α_t y activación promedio del reservorio \bar{h}_t .

debido a las marcadas diferencias en la magnitud de los retornos históricos absolutos de cada combustible. En series como Diésel, la amplificación de señales débiles cerca del origen y la compresión de las extremas revelan que la saturación de tanh actúa como regulador implícito: el reservorio no solo mapea α_t a un espacio de mayor dimensión, sino que redistribuye la masa de probabilidad de forma asimétrica según el régimen dominante.

La Figura 6.15 muestra la identificación de regímenes de mercado mediante SSRC, superpuestos a la serie de precios. La coincidencia entre transiciones de color y puntos de inflexión de la serie (como la marcada caída de 2020) confirma que el SSRC mantiene la interpretabilidad económica y operacional en un espacio de estados continuo de alta dimensión sin necesidad de variables exógenas, lo que lo hace directamente aplicable a contextos de planificación energética con disponibilidad de datos limitada.

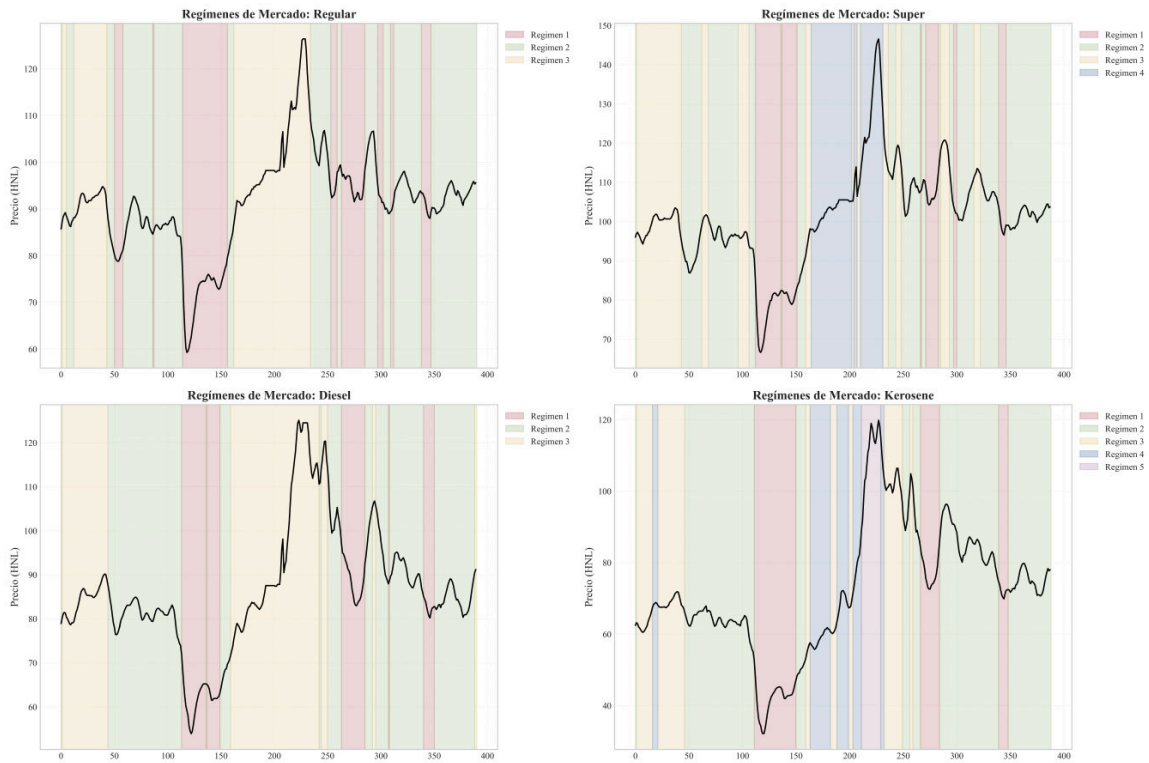


Figura 6.15: Identificación de regímenes de mercado mediante SSRC.

6.7.7. Discusión

Los resultados empíricos permiten responder a la pregunta central del capítulo: ¿aporta la extensión no lineal una ganancia predictiva sobre el modelo Markoviano lineal?

La respuesta es afirmativa y sustancial. El SSRC reduce el RMSE entre un 20% y un 23% en las cuatro series, con significancia estadística confirmada por el test de Diebold-Mariano en tres de cuatro casos. Este resultado valida empíricamente la predicción de la Observación 6.3.2: la generalización estrictamente más expresiva del SSRC efectivamente captura estructura no lineal que el modelo Markoviano no puede representar.

Los hallazgos centrales se detallan a continuación. Respecto a la relevancia del factor de fuga a . Los valores óptimos $a^* = 0.1$ para Regular y Diésel indican que la clave de la ganancia no es la no linealidad *per se*, sino la *memoria temporal adaptativa*: el Leaky-ESN con $a = 0.1$ retiene el 90% del estado previo, actuando como un filtro de suavizado que complementa la ventana $W \approx 50$ del operador TCRA. En contraste, Kerosene con $a^* = 1.0$ se beneficia de la memoria recurrente pura, consistente con su dinámica más volátil y su mayor número de regímenes ($K = 5$).

En cuanto a la saturación dimensional, la Figura 6.7 muestra que la ganancia se satura rápidamente: la diferencia entre $D = 20$ y $D = 150$ es inferior al 0.5 % en RMSE. Esto indica que la complejidad no lineal intrínseca del proceso generador de precios de combustibles hondureños es baja, lo cual es consistente con un mercado regulado donde los ajustes semanales son pequeños y predecibles. El reservorio no necesita alta dimensionalidad para capturar esta estructura.

Sobre la robustez del ensemble, la Figura 6.10 muestra que la totalidad de las 30 realizaciones produce RMSE inferior al modelo de referencia markoviano, con rangos intercuartílicos estrechos. Esto confirma que la ganancia no es un artefacto de una inicialización afortunada, sino una propiedad estructural del modelo.

Queda la pregunta de qué estructura captura el reservorio que el modelo Markoviano no puede representar. La ganancia se atribuye a la capacidad del SSRC de codificar *asimetrías temporales* en la dinámica de precios: las caídas y las recuperaciones no son simétricas ni en magnitud ni en velocidad. El modelo Markoviano, al operar sobre estados discretos de un paso, trata las transiciones bajistas→estables y alcistas→estables como eventos de naturaleza equivalente, gobernados por probabilidades estáticas en $\hat{\mathbf{P}}$. El reservorio, al retener memoria continua de las últimas $\tau = 1/a$ semanas en su vector de estado \mathbf{h}_t , puede distinguir la *trayectoria* que condujo al estado actual, y no únicamente el estado en sí mismo. Esta capacidad es particularmente relevante para el pico de precios de 2022 y la recuperación posterior (visible en la Figura 6.11, semanas $\sim 5-20$), donde el SSRC anticipa mejor los puntos de inflexión que el modelo Markoviano.

Queda por sopesar la parsimonia, porque un reservorio de dimensión alta suma parámetros que conviene justificar. La saturación dimensional ya descrita acota ese costo: como la diferencia entre $D = 20$ y $D = 150$ es inferior al 0.5 % en RMSE, el modelo no necesita muchas neuronas y una dimensión modesta basta para capturar la estructura. Sobre esa base, la ganancia del 20 % al 23 % en RMSE compensa con holgura los parámetros adicionales frente al modelo Markoviano. La restricción de no negatividad del NNLS refuerza el balance. Poda las direcciones de lectura que solo ajustarían ruido y, al mantener fijo el reservorio en lugar de entrenarlo por completo, evita el sobreajuste que provocaría un ajuste libre de todos los pesos sobre muestras de apenas unos cientos de observaciones.

6.8. Limitaciones

Sensibilidad a la inicialización. A diferencia del marco TCRA-Markov, que es determinista dados los hiperparámetros, el SSRC introduce variabilidad estocástica a través de las matrices aleatorias \mathbf{W}_{in} y \mathbf{W}_{res} . Aunque la ESP garantiza que la condición inicial \mathbf{h}_0 no afecta el resultado, la *elección* de las matrices de reservorio sí influye en el desempeño. El promediado sobre 30 realizaciones (*ensemble*) mitiga este efecto, como lo evidencia la baja variabilidad observada en la Figura 6.10, pero incrementa el costo computacional por un factor de 30.

Condicionamiento numérico. Los números de condición $\kappa(\mathbf{H}) \sim 10^9\text{--}10^{10}$ observados en la Tabla 6.1 indican una sensibilidad numérica potencialmente alta. Aunque la estabilidad práctica está asegurada por la regularización implícita del NNLS y el ensemble, este condicionamiento impone un límite superior a la dimensión útil del reservorio: aumentar D más allá de ~ 150 degradaría $\kappa(\mathbf{H})$ sin ganancia predictiva.

Interpretabilidad reducida. Los estados del reservorio $\mathbf{h}_t \in \mathbb{R}^D$ carecen de la interpretación directa que poseen los estados discretos $S_t \in \{s_1, \dots, s_K\}$ del modelo Markoviano. Sin embargo, las Figuras 6.13–6.15 demuestran que la estructura de regímenes se preserva implícitamente en el espacio del reservorio, lo que permite recuperar interpretabilidad parcial mediante técnicas de reducción de dimensionalidad.

Selección de hiperparámetros del reservorio. Los parámetros $(D, \rho_{\text{res}}, a)$ carecen de una teoría de selección óptima análoga al AIC del modelo Markoviano. La búsqueda en grilla de 240 configuraciones es costosa, y si se escala a más series convendría recurrir a Optimización Bayesiana. También, la discretización univariada podría beneficiarse de algoritmos como Fisher-Jenks, que garantizan particiones óptimas globales en una dimensión.

Conservadurismo de la cota de perturbación. La Proposición 6.6 proporciona una cota en el peor caso que, como se evidencia en la Figura 6.12, sobreestima la inestabilidad real por varios órdenes de magnitud. Una cota más ajustada requeriría análisis probabilístico que incorpore la distribución de α_t y la estructura de \mathbf{W}_{res} , lo cual es un problema abierto para trabajo futuro. En particular, la aplicación del Teorema del Círculo de Gerschgorin

sobre la matriz dispersa \mathbf{W}_{res} podría producir cotas bastante más ajustadas.

6.9. Resumen del Capítulo

Este capítulo ha establecido la relación formal entre el marco TCRA-Markov desarrollado a lo largo de la tesis y la computación de reservorio no lineal. Los resultados principales son:

Sobre la conexión teórica. Se demostró (Teorema 6.3) que TCRA-Markov es un caso particular del SSRC, obtenido al fijar $\mathbf{W}_{\text{res}} = \mathbf{0}$, $\sigma = \text{id}$ y $a = 1$. Esta relación de inclusión establece una **jerarquía de modelos**: TCRA-Markov (lineal, interpretable, estados discretos) \subset TCRA-SSRC (no lineal, aproximador universal, estados continuos). La Observación 6.3.2 garantiza que la extensión es estrictamente más expresiva, y la verificación numérica (Figura 6.2) confirma la identidad a nivel de precisión de máquina.

Sobre la estimación. La capa de lectura del SSRC se estima mediante el mismo procedimiento NNLS de los capítulos anteriores. El Teorema 6.4 establece que las propiedades de existencia y unicidad se heredan directamente. La restricción de no negatividad, además de preservar la interpretación como combinación positiva de filtros, aporta regularización implícita que mitiga el condicionamiento numérico adverso ($\kappa \sim 10^9$).

Sobre el factor de fuga. La variante Leaky-ESN (Definición 6.3) con $a < 1$ resultó esencial para tres de cuatro series, indicando que la principal ganancia del SSRC sobre TCRA-Markov reside en la *memoria temporal adaptativa* más que en la no linealidad pura. Los valores $a^* = 0.1$ para Regular y Diésel implican una constante de tiempo efectiva de $\tau = 10$ semanas, complementaria a la ventana TCRA de 50–52 semanas.

Sobre la aplicación empírica. El SSRC reduce el RMSE entre un 20 % y un 23 % respecto al modelo Markoviano, con significancia estadística confirmada por el test de Diebold-Mariano para Superior ($p = 0.002$), Diésel ($p = 0.033$) y Kerosene ($p = 0.006$). La totalidad de las 30 realizaciones aleatorias produce RMSE inferior al modelo de referencia markoviano, y la estructura de regímenes económicos se preserva en el espacio latente del reservorio.

Perspectiva. La arquitectura TCRA-SSRC aquí desarrollada establece que el paradig-

ma de reservorio aporta ganancia predictiva en mercados de combustibles regulados, a costa de interpretabilidad. La extensión natural es hacia sistemas multivariados y acoplados, donde la capacidad de aproximación no lineal del reservorio puede explotarse más plenamente en contextos con mayor riqueza estructural.

Capítulo 7

Conclusiones Generales y Trabajo Futuro

7.1. Síntesis del Trabajo

Esta tesis construyó un marco matemático para analizar regímenes económicos a partir de la *Tasa de Cambio Relativa Autorregresiva* (TCRA). El recorrido va desde los fundamentos algebraicos del operador hasta su extensión no lineal con computación de reservorio. Cada capítulo se apoyó en los resultados del anterior, de modo que la jerarquía de modelos creció en complejidad pero siempre con garantías teóricas bajo condiciones explícitas y validación sobre datos reales del mercado hondureño de combustibles.

El punto de partida fue formalizar la familia de operadores TCRA y sus variantes (TCRAM, ETCRA, ETCRAM) como transformaciones parametrizadas sobre series temporales. Se demostraron las relaciones de anidamiento entre variantes y se estableció que la señal α_t admite interpretación como coeficiente de un modelo AR(1) local estimado por mínimos cuadrados ponderados. Esta interpretación resultó ser más relevante de lo que aparenta: al darle a α_t significado como *tasa de cambio local*, se abrió la puerta a usarla tanto en modelos estocásticos discretos (cadenas de Markov) como en modelos dinámicos continuos (reservorios recurrentes).

La extensión no lineal del operador incluyó teoremas de existencia y unicidad del estimador bajo condiciones de no degeneración de los datos en la ventana. Estas garantías analíticas aseguran que la señal α_t esté bien definida en cada instante, condición necesaria para la modelación Markoviana posterior.

La integración con cadenas de Markov de tiempo discreto formalizó las propiedades de ϕ -mixing, la convergencia asintótica y la conexión entre la brecha espectral de la matriz de transición y la velocidad de convergencia al equilibrio estacionario. El resultado central aquí es que la señal α_t discretizada hereda propiedades de mezcla bajo condiciones de irreducibilidad, aperiodicidad y estabilidad de la partición, lo que permite hacer inferencia

estadística válida en los capítulos aplicados.

La estimación de la matriz de transición $\hat{\mathbf{P}}$ se resolvió mediante mínimos cuadrados no negativos (NNLS), incluyendo una prueba constructiva de existencia y unicidad. La formulación transforma el problema matricial en un sistema lineal aumentado mediante el producto de Kronecker. La ventaja práctica es directa: evita la optimización no convexa de la verosimilitud que hace al MS-AR vulnerable a mínimos locales, y garantiza convergencia global. Las probabilidades estocásticas ($\sum_i \hat{p}_{ij} = 1$ para cada columna j) se aseguran mediante normalización columna por columna, preservando la convención de matrices estocásticas por columnas del Capítulo 4.

La validación empírica aplicó el marco TCRA-Markov a series semanales de cuatro combustibles hondureños (Regular, Superior, Diésel y Kerosene) durante 2017–2025. El modelo fue superior a la discretización por cuantiles fijos ($p < 0.05$ en exactitud para las cuatro series) y, dato revelador, el modelo de referencia MS-AR no alcanzó convergencia bajo el protocolo de validación y conjunto de datos de este estudio (statsmodels v0.14, 50 reinicios aleatorios). El análisis espectral mostró un mercado con alta inercia (brechas espectrales $\gamma \in [0.021, 0.040]$ y tiempos de mezcla de 25 a 47 semanas). También se identificó que para ventanas W demasiado grandes el modelo se degenera, lo que llevó a establecer la restricción $W \leq 52$ como condición necesaria para que los regímenes sean identificables.

El apartado teórico final cerró el arco analítico de la tesis al demostrar que el marco TCRA-Markov es un caso particular de una arquitectura más general: el Computador de Reservorio Estructurado Estocástico (SSRC). El Teorema de Inclusión (Teorema 6.3) estableció formalmente que toda predicción del modelo Markoviano es reproducible por un reservorio con $\mathbf{W}_{\text{res}} = \mathbf{0}$, $\sigma = \text{id}$ y $a = 1$, pero no a la inversa. La extensión no lineal, validada empíricamente con mejoras del 20–23% en RMSE respecto al modelo Markoviano (estadísticamente significativas para tres de cuatro series, $p < 0.05$ vía test de Diebold-Mariano), demostró que la representación continua de estados mediante reservorios recurrentes captura estructura que la discretización Markoviana no puede representar, particularmente las asimetrías temporales entre caídas y recuperaciones de precios.

7.2. Resumen Cuantitativo de Resultados

La Tabla 7.1 resume las métricas de rendimiento del modelo TCRA-Markov para las cuatro series de combustibles analizadas, junto con los métodos de referencia. Se reportan la exactitud de predicción de régimen, el número de estados K , la ventana óptima W y la significancia estadística frente al modelo de referencia de cuantiles fijos. A nivel de desempeño, se observa que el modelo TCRA-Markov (K-medias) supera de forma consistente en exactitud a la referencia de cuantiles en todas las series de combustible estudiadas. Es de particular relevancia destacar que el modelo autorregresivo con cambio de régimen de Hamilton (MS-AR) no logró alcanzar convergencia bajo el protocolo de validación y optimización seleccionado (statsmodels v0.14, 50 reinicios aleatorios) debido a la naturaleza altamente no lineal y discontinua de las transiciones locales de precios regulados en Honduras. Esto confirma empíricamente la robustez y la ventaja computacional del método de mínimos cuadrados no negativos (NNLS) de formulación convexa frente al algoritmo de optimización iterativa EM empleado en el MS-AR tradicional.

Tabla 7.1: Resumen consolidado de métricas por serie y método.

| Serie | K | W | Exactitud | vs. Cuantiles | MS-AR |
|-------------------|---|----|-----------|---------------|----------|
| Gasolina Regular | 3 | 50 | 85.10% | $p < 0.01$ | No conv. |
| Gasolina Superior | 4 | 52 | 62.78% | $p < 0.05$ | No conv. |
| Diésel | 3 | 50 | 71.56% | $p < 0.05$ | No conv. |
| Kerosene | 5 | 52 | 68.20% | $p < 0.05$ | No conv. |

En todos los casos, el modelo TCRA-Markov superó significativamente al modelo de referencia de cuantiles fijos. El modelo MS-AR no alcanzó convergencia bajo el protocolo de validación y conjunto de datos de este estudio (statsmodels v0.14, 50 reinicios), lo que evidencia la ventaja práctica de la formulación NNLS en este contexto empírico frente a la optimización por máxima verosimilitud para estas series reguladas.

Esta divergencia plantea una interrogante fundamental sobre si es preferible prescindir de supuestos distribucionales en el modelado de este tipo de series temporales. El enfoque no paramétrico y libre de distribución (*distribution-free*) adoptado en el marco TCRA-Markov proporciona una notable robustez frente a colas pesadas, saltos discretos y asimetrías severas que suelen invalidar los supuestos de normalidad o linealidad de los modelos

paramétricos tradicionales. No obstante, la ventaja clave de la metodología propuesta no radica únicamente en la ausencia de supuestos distribucionales exógenos, sino en el desacoplamiento algorítmico que permite reformular un problema de optimización no convexo e inestable (el algoritmo de Esperanza-Maximización en MS-AR) en un problema de optimización cuadrática estrictamente convexo y determinista (NNLS). Esto garantiza teóricamente la convergencia global a un óptimo único, evitando las limitaciones de convergencia numérica y la inestabilidad de parámetros características de los métodos iterativos clásicos ante datos con alta persistencia y controles administrativos.

7.3. Contribuciones

Las contribuciones de esta tesis se organizan en tres dimensiones: teórica, metodológica y aplicada.

7.3.1. Contribuciones Teóricas

1. Se formalizó la estructura de retículo (*lattice*) de variantes de la TCRA, demostrando las relaciones de anidamiento donde TCRAM y ETCRA representan ramas paralelas que confluyen en la variante general ETCRAM (Capítulo 2).
2. El operador TCRA polinómico admite un único estimador bajo condiciones de no degeneración. La demostración, basada en la convexidad estricta de la funcional de costo, asegura que la señal α_t queda bien definida.
3. La secuencia de estados $\{S_t\}$ obtenida por discretización de α_t hereda propiedades de ϕ -mixing, lo que habilita la inferencia estadística asintótica sobre la cadena de Markov resultante.
4. Para la estimación de $\hat{\mathbf{P}}$ mediante NNLS aumentado se probó existencia y unicidad de la solución cuando la matriz de diseño tiene rango completo, y se caracterizaron las condiciones de rango en función de la representación de todos los estados en los datos.

5. El Teorema de Inclusión TCRA-Markov \subset SSRC establece que el modelo Markoviano lineal es un caso degenerado del computador de reservorio estructurado, fijando una jerarquía estricta de expresividad entre ambos marcos.
6. Se extendió el teorema de la Propiedad de Estado de Eco al caso Leaky-ESN: la contracción del mapa de reservorio se preserva para todo $a \in (0, 1]$ siempre que $\rho(\mathbf{W}_{\text{res}}) < 1$.
7. Finalmente, se caracterizó analíticamente el fenómeno de degeneración cuando $W \rightarrow T$: $\text{Var}(\alpha_t) \rightarrow 0$ y K-medias converge a $K_{\text{eff}} = 1$ independientemente de K nominal.

7.3.2. Contribuciones Metodológicas

El marco TCRA-Markov de dos etapas integra el operador TCRA con cadenas de Markov estimadas vía NNLS, evitando la optimización no convexa del MS-AR y asegurando convergencia global. Este marco se aceptó para publicación en IEEE CONCAPAN XLII (2025). La extensión no lineal (la arquitectura TCRA-SSRC) reemplaza la discretización K-medias por un reservorio recurrente Leaky-ESN, preservando la estimación NNLS para la capa de lectura. Es clave señalar que la observabilidad directa de los estados de transición, provista originalmente por la formulación del operador TCRA, es el elemento habilitador que permite prescindir de algoritmos iterativos no convexos como Esperanza-Maximización (EM) y aplicar optimización cuadrática convexa (NNLS) en ambas etapas.

También se diseñó un protocolo de optimización jerárquico en tres niveles (RMSE \rightarrow Exactitud \rightarrow AIC) con búsqueda exhaustiva en grilla y restricción fundamentada de $W \leq 52$. La validación se completó con el test de Diebold-Mariano, combinando ventanas deslizantes, ensemble de 30 realizaciones y pruebas de significancia.

7.3.3. Contribuciones Aplicadas

En el plano empírico, la tesis aporta una caracterización cuantitativa del mercado hondureño de combustibles: se identificaron de 3 a 5 regímenes por serie, con centroides interpretables y distribuciones estacionarias que formalizan la inercia del mercado regulado.

Las pruebas estadísticas confirman que TCRA-Markov supera a la discretización por cuantiles ($p < 0.05$) y que el modelo de referencia MS-AR no alcanza la convergencia bajo el protocolo y conjunto de datos de este estudio, mientras que el modelo TCRA-SSRC mejora sobre TCRA-Markov en 20–23 % ($p < 0.05$ para tres de cuatro series). Un resultado no anticipado es la jerarquía de complejidad del mercado: Regular y Diésel ($K^* = 3$) exhiben una dinámica más simple que Superior ($K^* = 4$) y Kerosene ($K^* = 5$), y la complejidad estructural se correlaciona positivamente con la dificultad de predicción.

7.4. Limitaciones Generales

Las siguientes limitaciones delimitan el alcance de los resultados obtenidos en esta investigación:

1. **Alcance y representatividad de los datos:** Los datos empíricos provienen exclusivamente del mercado hondureño de combustibles (un mercado regulado administrativamente, con ajustes semanales de precios y baja volatilidad en comparación con mercados libres) y todos los pronósticos se realizan a un solo paso adelante ($h = 1$ semana). La generalización a mercados desregulados, a otras clases de activos financieros o a horizontes múltiples de tiempo requiere investigación adicional.
2. **Condicionamiento numérico del reservorio:** Los números de condición de la matriz de estados alcanzan valores elevados ($\kappa(\mathbf{H}) \sim 10^9$ – 10^{10}). Aunque en la práctica la restricción de no negatividad de NNLS y el promediado de un ensemble de 30 realizaciones mitigan la sensibilidad numérica del sistema, una solución matemática formal (como la regularización de Tikhonov con selección adaptativa) queda pendiente de mayor desarrollo teórico.
3. **Asunción de homogeneidad y estabilidad temporal:** El modelo Markoviano asume probabilidades de transición constantes en el tiempo, y el SSRC asume pesos de lectura fijos tras su entrenamiento. Sin embargo, el período analizado 2017–2025 incluye choques estructurales severos como la pandemia de COVID-19 y el pico global de precios energéticos de 2022, eventos que potencialmente violan el supuesto de

estacionariedad. No se realizó una prueba formal para evaluar el cambio estructural temporal de los parámetros.

4. **Supuesto de Markov de primer orden sin validación de orden:** La estructura del modelo Markoviano se construyó bajo la asunción de primer orden (Suposición 3.1), lo que significa que el estado futuro depende únicamente del estado actual. Esta hipótesis se adoptó por parsimonia metodológica y no fue precedida de pruebas de orden estadísticas formales (e.g., pruebas Chi-cuadrado para transiciones de segundo orden o superiores), omitiendo posibles dependencias de mayor alcance.
5. **Sensibilidad a la inicialización y selección heurística:** La estimación del espacio discretizado mediante K-medias exhibe sensibilidad a la semilla de inicialización aleatoria de los centroides. Adicionalmente, el protocolo jerárquico seleccionó los hiperparámetros óptimos sobre el dataset completo en lugar de implementar una validación cruzada anidada rigurosa, lo que introduce un sesgo de selección potencial en los resultados predictivos reportados.
6. **Enfoque univariante y exclusión de variables exógenas:** El modelado se restringe a una estructura puramente autorregresiva (endógena), donde la dinámica de cada combustible se predice basándose únicamente en sus precios históricos. Esto impide capturar de forma directa el impacto de choques exógenos macroeconómicos globales (como el precio internacional del crudo WTI) o variables de política local (como variaciones discrecionales en los subsidios).

7.5. Líneas de Trabajo Futuro

Los resultados y limitaciones de esta tesis sugieren varias direcciones de investigación futura, organizadas por horizonte temporal.

7.5.1. Extensiones Inmediatas

Sistemas multivariados acoplados. La extensión natural del marco es pasar de series univariantes a sistemas de series temporales interrelacionadas. En el contexto

centroamericano, los precios de combustibles de las seis economías del istmo están acoplados por proximidad geográfica, acuerdos comerciales y dependencia compartida del crudo WTI. Un modelo TCRA-Markov multivariado estimaría una matriz de transición conjunta $\hat{\mathbf{P}} \in \mathbb{R}^{K^m \times K^m}$, donde m es el número de series acopladas, capturando las dependencias cruzadas entre mercados. Sin embargo, esto induce una explosión dimensional exponencial (con $m = 6$ países y $K = 4$ regímenes por país, el sistema requiere estimar una matriz de 4096×4096). Para mitigar este problema, se propone el uso de estructuras de tipo *Hub-and-Spoke* (donde un nodo central conecta y coordina los mercados satélites, reduciendo la complejidad a mK^2), factorizaciones matriciales de bajo rango o modelos gráficos probabilísticos. Esta línea de investigación se encuentra actualmente en desarrollo como artículo de investigación independiente.

Comparación de dinámicas regionales centroamericanas. Aunque esta tesis se enfoca en el mercado regulado hondureño, una línea de investigación prioritaria consiste en extender el análisis comparativo a nivel de Centroamérica para disponer de un escenario de validación más amplio. Esta comparativa regional es fundamental para evaluar el comportamiento de los operadores bajo marcos regulatorios divergentes. Por ejemplo, permitiría contrastar la inercia temporal y la exactitud predictiva en países con control administrativo de precios (como Honduras y Costa Rica) frente a economías con regímenes de libre mercado (como El Salvador y Guatemala). Implementar este análisis regional enriquecerá la robustez del modelo y ayudará a dimensionar el impacto real de las políticas públicas sobre la dinámica local de precios.

Pruebas de estabilidad temporal. La partición de la muestra en subperíodos (prepandemia 2017–2019, pandemia 2020–2021, pospandemia 2022–2025) y la comparación de las matrices $\hat{\mathbf{P}}$ estimadas en cada subperíodo, mediante pruebas de homogeneidad basadas en distancia entre matrices estocásticas, permitirían evaluar formalmente el supuesto de estabilidad temporal. Para mitigar el problema de muestras pequeñas y celdas con pocas observaciones en subperíodos cortos (que incrementan la variabilidad de la estimación), se propone el uso de métodos de bootstrap paramétrico para estimar la incertidumbre y obtener intervalos de confianza robustos para cada celda de la matriz de transición.

Regularización adaptativa del SSRC. La sustitución de NNLS por un esquema de regularización de Tikhonov con selección del parámetro de regularización η_R (utilizando esta notación para evitar confusión con el factor de olvido λ del operador TCRA) mediante validación cruzada generalizada (GCV) podría mejorar el condicionamiento numérico sin sacrificar la interpretabilidad de los pesos no negativos, mediante una formulación de NNLS regularizada

$$\hat{\mathbf{w}} = \arg \min_{\mathbf{w} \geq 0} \|\mathbf{H}\mathbf{w} - \mathbf{y}\|^2 + \eta_R \|\mathbf{w}\|^2.$$

7.5.2. Extensiones de Mediano Plazo

Topología de reservorio estructurada. En el SSRC actual, la matriz \mathbf{W}_{res} es aleatoria y dispersa. Trabajos recientes en la literatura de computación de reservorio sugieren que topologías deterministas (e.g., anillo, ciclo con saltos) pueden reducir la varianza entre realizaciones y mejorar la reproducibilidad. Explorar estas topologías en el contexto TCRA es una extensión natural.

Incorporación de variables exógenas. El marco actual es puramente endógeno: α_t se calcula a partir de los precios pasados. La incorporación de variables exógenas excede el análisis univariante al incorporar el precio del crudo WTI, el tipo de cambio o índices macroeconómicos como entradas adicionales. Sin embargo, la inyección directa de múltiples variables en un único reservorio masivo puede violar la propiedad de estado de eco (ESP) al alterar el radio espectral efectivo y la contracción del sistema. Como alternativa viable, se propone una arquitectura de reservorio profundo modular (*Deep ESN modular*), en la cual cada variable de entrada es procesada de forma independiente por su propio operador y sub-reservorio, garantizando de manera aislada la contracción, para posteriormente concatenar las capas de salida antes de realizar la estimación por NNLS.

Pronóstico multi-horizonte. La extensión del horizonte de predicción de $h = 1$ a $h \in \{1, 4, 8, 12\}$ semanas, ya sea mediante recursión del reservorio o mediante formulación directa con múltiples salidas, permitiría evaluar la degradación de la capacidad predictiva con el horizonte y determinar el alcance práctico del marco para la planificación económica.

7.5.3. Visión de Largo Plazo

Generalización a otras clases de activos. El marco TCRA-Markov/SSRC no asume nada específico sobre la naturaleza de la serie temporal subyacente. Su aplicación a tasas de interés, índices bursátiles, tipos de cambio o indicadores macroeconómicos abre un programa de investigación que podría validar la generalidad del enfoque más allá del mercado de combustibles.

Fundamentación teórica y cotas de convergencia. Las cotas de perturbación derivadas en esta tesis para el SSRC son conservadoras por varios órdenes de magnitud. Una línea de largo plazo consiste en desarrollar cotas más ajustadas utilizando la localización de valores propios vía círculos de Gerschgorin sobre la estructura dispersa de \mathbf{W}_{res} . Asimismo, se propone formalizar las tasas explícitas de convergencia para la propiedad de estado de eco (ESP) como función conjunta del factor de fuga a y el radio espectral ρ , así como estimar analíticamente la velocidad de convergencia del Aproximador Universal en función de la dimensión del reservorio D .

7.6. Recomendaciones

Las conclusiones y resultados de esta investigación permiten formular las siguientes recomendaciones específicas, orientadas a la toma de decisiones y a la investigación futura:

7.6.1. Recomendaciones para el Mercado de Combustibles de Honduras

1. **Dirigido a la Secretaría de Energía (SEN) y a la Comisión Administradora de Petróleo (CAP):** Se recomienda la adopción del operador de Tasa de Cambio Relativa Autorregresiva (TCRA) en conjunto con cadenas de Markov para establecer un sistema de monitoreo sistemático de las transiciones de precios de combustibles a nivel nacional. La capacidad de estimar la matriz de transición $\hat{\mathbf{P}}$ con garantías de convergencia global (mediante la formulación NNLS) ofrece un instrumento cuantitativo robusto para predecir si el mercado entrará en regímenes de alza fuerte

o descenso fuerte, sirviendo como insumo técnico para el diseño óptimo de subsidios gubernamentales y mitigación de choques de precios externos.

2. **Dirigido a la Secretaría de Desarrollo Económico (SDE):** Se recomienda utilizar las estimaciones de la brecha espectral y los tiempos de mezcla (que varían de 25 a 47 semanas en el mercado hondureño) como métricas oficiales para evaluar el grado de inercia y persistencia en los precios de los energéticos. Esto permitirá fundamentar técnicamente el análisis de impacto inflacionario por rezago de distribución en la cadena de transporte de mercancías.

7.6.2. Recomendaciones para el Desarrollo de Modelos Estocásticos de Regímenes

1. **Dirigido a los investigadores en matemáticas aplicadas y econometría de la Universidad Nacional Autónoma de Honduras (UNAH):** Se recomienda el uso de la Computación de Reservorio Estocásticamente Estructurada (SSRC) como una extensión no lineal para el modelado de series de tiempo con cambios bruscos de dinámica. Se sugiere explotar las propiedades de estabilidad y el Teorema de Inclusión demostrados en esta tesis para estructurar reservorios con restricciones físicas o de signo, garantizando la interpretabilidad de los coeficientes de salida \mathbf{W}_{out} obtenidos a través de optimización convexa.
2. **Dirigido a la comunidad científica internacional en aprendizaje automático:** Se recomienda profundizar en la regularización adaptativa de Tikhonov acoplada a la restricción de no negatividad para matrices de reservorios de alto condicionamiento numérico ($\kappa(\mathbf{H}) \sim 10^9-10^{10}$), con el fin de robustecer el modelado dinámico continuo de series altamente volátiles.

7.7. Reflexiones Finales

El recorrido de esta tesis va desde la formalización de un operador de tasa de cambio local hasta la demostración de que ese operador, integrado con cadenas de Markov, es un

caso particular de un computador de reservorio no lineal. Existe una dicotomía fundamental en matemáticas aplicadas entre simplicidad interpretable y complejidad expresiva, y este trabajo la recorre en toda su extensión.

El modelo TCRA-Markov, con sus K estados discretos y su matriz de transición de $K(K - 1)$ parámetros libres, es transparente: un economista puede leer $\hat{\mathbf{P}}$ y comprender la dinámica del mercado. El modelo TCRA-SSRC, con su reservorio de D neuronas y su representación continua de estados, es más expresivo (entre un 20 y un 23% más preciso en el pronóstico), aunque menos inmediatamente interpretable. Ambos tienen su lugar.

Lo que unifica ambos extremos es la formulación NNLS, que garantiza soluciones no negativas, convergencia global y, sobre todo, la posibilidad de interpretar los pesos de salida como contribuciones no negativas, no necesariamente porcentuales, de cada componente al pronóstico. Esta propiedad, que podría parecer un detalle técnico, es lo que hace que el marco completo sea útil más allá de la academia: un modelo que mejora la predicción pero no puede explicar por qué tiene valor limitado para la toma de decisiones en política económica.

La aplicación a combustibles hondureños no fue elegida al azar. En una economía donde el precio del galón de diésel afecta directamente el costo del transporte de alimentos, la capacidad de anticipar regímenes de precios y de cuantificar la incertidumbre asociada a esa anticipación, tiene consecuencias prácticas tangibles. Si el marco desarrollado en estas páginas contribuye, aunque sea marginalmente, a mejorar la planificación en ese contexto, habrá cumplido su propósito.

Bibliografía

- [1] A. C. Aitken. On least squares and linear combination of observations. *Proceedings of the Royal Society of Edinburgh*, 55:42–48, 1936.
- [2] T. W. Anderson. *The Statistical Analysis of Time Series*. John Wiley and Sons, New York, 1971.
- [3] T. W. Anderson and L. A. Goodman. Statistical inference about markov chains. *The Annals of Mathematical Statistics*, 28(1):89–110, 1957.
- [4] A. Ang and G. Bekaert. Regime switches in interest rates. *Journal of Business & Economic Statistics*, 20(2):163–182, 2002.
- [5] Banco Central de Honduras. Informe de la economía hondureña. Technical report, Banco Central de Honduras, Tegucigalpa, Honduras, 2022.
- [6] Banco Central de Honduras. Memoria anual 2023. Technical report, Banco Central de Honduras, Tegucigalpa, Honduras, 2023.
- [7] L. Banegas and F. Vides. Stochastically structured reservoir computers for financial and economic system identification, 2025.
- [8] E. Bolthausen. The berry-esseen theorem for functionals of discrete markov chains. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, 54(1):59–73, 1980.
- [9] G. E. P. Box, G. M. Jenkins, G. C. Reinsel, and G. M. Ljung. *Time Series Analysis: Forecasting and Control*. John Wiley & Sons, Hoboken, NJ, 5th edition, 2015.
- [10] R. C. Bradley. Basic properties of strong mixing conditions: A survey and some open questions. *Probability Surveys*, 2:107–144, 2005.
- [11] R. Bro and S. De Jong. A fast non-negativity-constrained least squares algorithm. *Journal of Chemometrics*, 11(5):393–401, 1997.
- [12] CEPAL. Estudio económico de América Latina y el Caribe 2024: Trampa de bajo crecimiento, cambio climático y dinámica del empleo. Technical report, Comisión Económica para América Latina y el Caribe, Santiago de Chile, 2024.

- [13] F. X. Diebold and R. S. Mariano. Comparing predictive accuracy. *Journal of Business & Economic Statistics*, 13(3):253–263, 1995.
- [14] L. Gonon and J.-P. Ortega. Fading memory echo state networks are universal. *Neural Networks*, 138:10–13, 2021.
- [15] W. H. Greene. *Econometric Analysis*. Pearson, New York, NY, 8th edition, 2018.
- [16] L. Grigoryeva and J.-P. Ortega. Echo state networks are universal. *Neural Networks*, 108:495–508, 2018.
- [17] M. Guidolin and A. Timmermann. Forecasts of US short-term interest rates: A flexible forecast combination approach. *Journal of Econometrics*, 150(2):297–311, 2009.
- [18] J. D. Hamilton. A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica*, 57(2):357–384, 1989.
- [19] J. D. Hamilton. *Time Series Analysis*. Princeton University Press, Princeton, NJ, 1994.
- [20] S. Holm. A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics*, 6(2):65–70, 1979.
- [21] R. J. Hyndman and G. Athanasopoulos. *Forecasting: Principles and Practice*. OTexts, Melbourne, Australia, 3rd edition, 2021.
- [22] R. J. Hyndman and A. B. Koehler. Another look at measures of forecast accuracy. *International Journal of Forecasting*, 22(4):679–688, 2006.
- [23] P. A. Imam and J. R. W. Temple. State capacity and growth regimes. IMF Working Paper WP/25/14, International Monetary Fund, 2025. Clasifica regímenes de crecimiento para 108 países en desarrollo mediante cadenas de Markov de estados finitos.
- [24] International Monetary Fund. IMF DataMapper: Real GDP Growth, Annual Percent Change. https://www.imf.org/external/datamapper/NGDP_RPCH@WEO,

2025. World Economic Outlook, October 2025. Clasificación visual por rangos: $\geq 6\%$, $3-6\%$, $0-3\%$, $-3-0\%$, $< -3\%$. Consultado: marzo 2025.
- [25] H. Jaeger. The “echo state” approach to analysing and training recurrent neural networks. *GMD Report*, 148(34):13, 2001.
- [26] J. G. Kemeny and J. L. Snell. *Finite Markov Chains*. D. Van Nostrand Company, Princeton, NJ, 1960.
- [27] C.-J. Kim and C. R. Nelson. *State-Space Models with Regime Switching: Classical and Gibbs-Sampling Approaches with Applications*. MIT Press, 1999.
- [28] S. Kolassa. Why the “best” point forecast depends on the error or accuracy measure. *International Journal of Forecasting*, 36(1):208–211, 2020.
- [29] H.-M. Krolzig. *Markov-Switching Vector Autoregressions: Modelling, Statistical Inference, and Application to Business Cycle Analysis*. Springer, Berlin, 1997.
- [30] C. L. Lawson and R. J. Hanson. *Solving Least Squares Problems*. SIAM, Philadelphia, PA, USA, 1995.
- [31] D. A. Levin, Y. Peres, and E. L. Wilmer. *Markov Chains and Mixing Times*. American Mathematical Society, Providence, RI, 2009.
- [32] M. Lukoševičius and H. Jaeger. Reservoir computing approaches to recurrent neural network training. *Computer Science Review*, 3(3):127–149, 2009.
- [33] H. Lütkepohl. *New Introduction to Multiple Time Series Analysis*. Springer, Berlin, 2005.
- [34] J. MacQueen. Some methods for classification and analysis of multivariate observations. In *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability*, volume 1, pages 281–297. University of California Press, 1967.
- [35] S. Makridakis, E. Spiliotis, and V. Assimakopoulos. The M4 competition: Results, findings, conclusion and way forward. *International Journal of Forecasting*, 34(4):802–808, 2018.

- [36] J. R. Norris. *Markov Chains*. Cambridge University Press, 1998.
- [37] Proceso Digital. Tabla de precios de los combustibles, 2025. Último acceso: Julio de 2025.
- [38] L. R. Rabiner. A tutorial on hidden markov models and selected applications in speech recognition. *Proceedings of the IEEE*, 77(2):257–286, 1989.
- [39] T. Teräsvirta. Specification, estimation, and evaluation of smooth transition autoregressive models. *Journal of the American Statistical Association*, 89(425):208–218, 1994.
- [40] H. Tong. *Non-linear Time Series: A Dynamical System Approach*. Oxford University Press, Oxford, 1990.
- [41] R. S. Tsay. *Analysis of Financial Time Series*. John Wiley & Sons, Hoboken, NJ, 2005.
- [42] R. S. Tsay. *Analysis of Financial Time Series*. John Wiley and Sons, Hoboken, NJ, third edition, 2010.
- [43] F. Vides and G. Lopez Gutierrez. Estimación de matrices de transición de procesos de 1-markov: Aplicación de mínimos cuadrados no negativos (nnls). Technical report, Departamento de Matemática, Universidad Nacional Autónoma de Honduras, July 2025. Documento técnico interno. Contiene la implementación de referencia de la función SRep para estimación NNLS de matrices de transición.
- [44] World Bank. Honduras overview, 2023. Accessed: 2025.
- [45] World Bank. GDP (current US\$) — Honduras, 2024. World Development Indicators. Accessed: 2025.
- [46] World Bank. GDP growth (annual%) — Honduras, 2024. World Development Indicators. Accessed: 2025.
- [47] World Bank. GDP per capita (constant LCU) — Honduras, 2024. World Development Indicators. Accessed: 2025.

- [48] World Bank. GDP per capita (current US\$) — Honduras, 2024. World Development Indicators. Accessed: 2025.
- [49] I. B. Yildiz, H. Jaeger, and S. J. Kiebel. Re-visiting the echo state property. *Neural Networks*, 35:1–9, 2012.

A.1. Demostración Extendida del Teorema de ϕ -Mixing

Sea $\{S_t\}$ una cadena de Markov homogénea con espacio de estados finito $S = \{s_1, \dots, s_N\}$, matriz de transición P , e irreducible y aperiódica.

Teorema .1 ([10]). *Una cadena de Markov $\{S_t\}$ es ϕ -mixing con tasa geométrica si existe $\delta > 0$ tal que para todo par de estados $s_i, s_j \in S$, existe $t \in \mathbb{N}$ con:*

$$P^t(s_j|s_i) > \delta.$$

Idea de la demostración. El resultado se deduce a partir del hecho de que las cadenas de Markov irreducibles, aperiódicas y con espacio de estados finito son uniformemente ergódicas. Tal propiedad implica mezcla fuerte ϕ -mixing con decaimiento exponencial. \square

Observación práctica: En el modelo TCRA-Markov, la irreducibilidad puede garantizarse si el rango de $\alpha_{t,\lambda,W}$ permite transiciones entre todos los estados definidos por π .

A.2. Estabilidad Discreta de la Función de Cuantización

π

La función de cuantización $\pi: \mathbb{R} \rightarrow \mathcal{S}$ definida en la Definición 3.1 es una función escalonada y, por tanto, discontinua en los umbrales $\{\theta_j\}$. En sentido estricto, π no es Lipschitz continua. Sin embargo, la estabilidad del modelo se garantiza mediante la separación entre umbrales, como se demuestra en el Lema 3.2.

Si se desea una versión suavizada de π para análisis asintótico, se puede considerar la aproximación sigmoide:

$$\pi^*(x) = \sum_{j=1}^K s_j \cdot h_j(x), \quad h_j(x) = \frac{1}{1 + \exp(-a(x - \theta_j))},$$

donde $a > 0$ controla la pendiente de la transición. La constante de Lipschitz de π^* satisface $L_{\pi^*} \leq a/4$, alcanzada en los umbrales ($x = \theta_j$). Cuando $a \rightarrow \infty$, $\pi^* \rightarrow \pi$ puntualmente,

pero $L_{\pi^*} \rightarrow \infty$, lo cual es consistente con la discontinuidad de π . Esta suavización es útil para demostraciones técnicas, pero en la práctica del modelo TCRA-Markov se trabaja directamente con la versión discreta.

A.3. Trayectorias y Transiciones del Proceso $\{S_t\}$ con Datos Reales

Las siguientes figuras y tablas presentan las trayectorias de estados y los diagramas de transición obtenidos directamente de los datos de combustibles (Superior y Regular), utilizando las matrices $\hat{\mathbf{P}}$ estimadas por el modelo TCRA-Markov (ene 2017 – jul 2025). Para el análisis de las trayectorias temporales, se contrasta la secuencia observada (asignada por el retorno local α_t y los centroides de K-medias con $K = 4$ estados) con una trayectoria simulada de Monte Carlo sobre una ventana de 20 periodos seleccionada por exhibir la máxima variabilidad de regímenes. Es importante recalcar que la trayectoria simulada corresponde a una realización estocástica particular libre y no pretende realizar una predicción determinista paso a paso ni coincidir con la serie real observada. La discrepancia paso a paso en ciertos tramos es inherente al carácter estocástico del método de Monte Carlo; su propósito es ilustrar la consistencia dinámica del modelo al reproducir patrones equivalentes de persistencia y volatilidad. Asimismo, los diagramas de transición grafican la estructura Markoviana de probabilidades estimadas con datos reales, mostrando únicamente aquellas transiciones con una probabilidad de ocurrencia marginal ≥ 0.02 para facilitar la lectura.

Combustible Superior ($K = 4$ estados)

Tabla A.1: Matriz de transición estimada $\hat{\mathbf{P}}$ para combustible **Superior** ($K = 4$).

| | s_1 | s_2 | s_3 | s_4 |
|-------|-------|--------------|--------------|--------------|
| s_1 | 0.500 | 0.011 | – | – |
| s_2 | 0.250 | 0.711 | 0.097 | 0.011 |
| s_3 | – | 0.267 | 0.819 | 0.232 |
| s_4 | 0.250 | 0.011 | 0.085 | 0.758 |

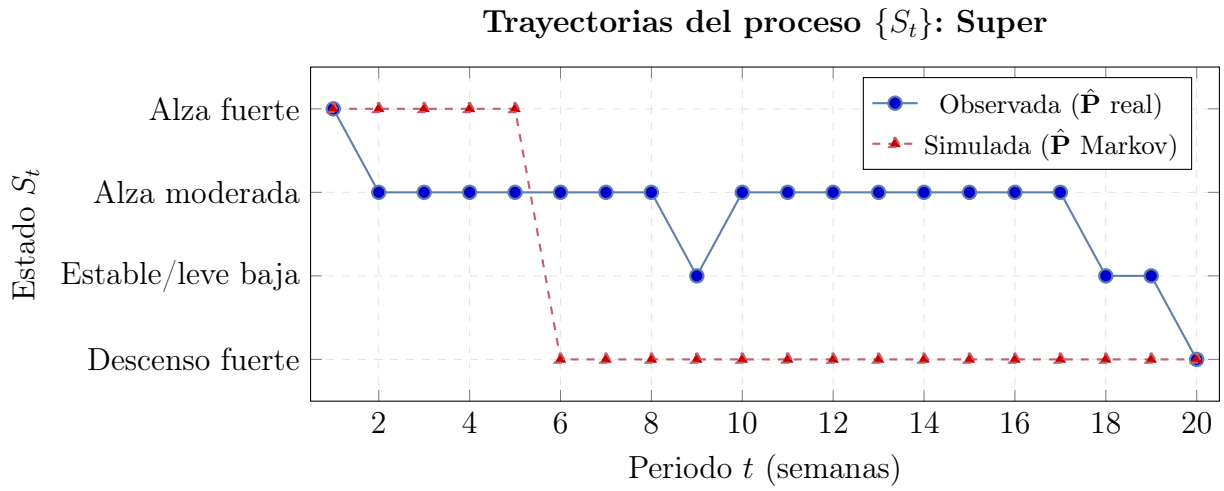


Figura A.1: Trayectorias del proceso $\{S_t\}$ para combustible **Superior** (secuencia observada vs. simulada).

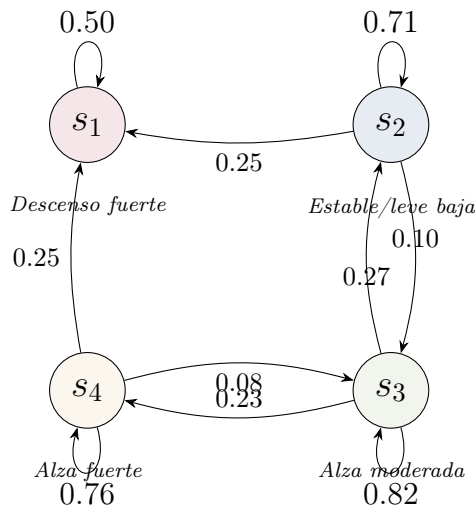


Figura A.2: Diagrama de transición entre regímenes para combustible **Superior** ($K = 4$).

Combustible Regular ($K = 4$ estados)

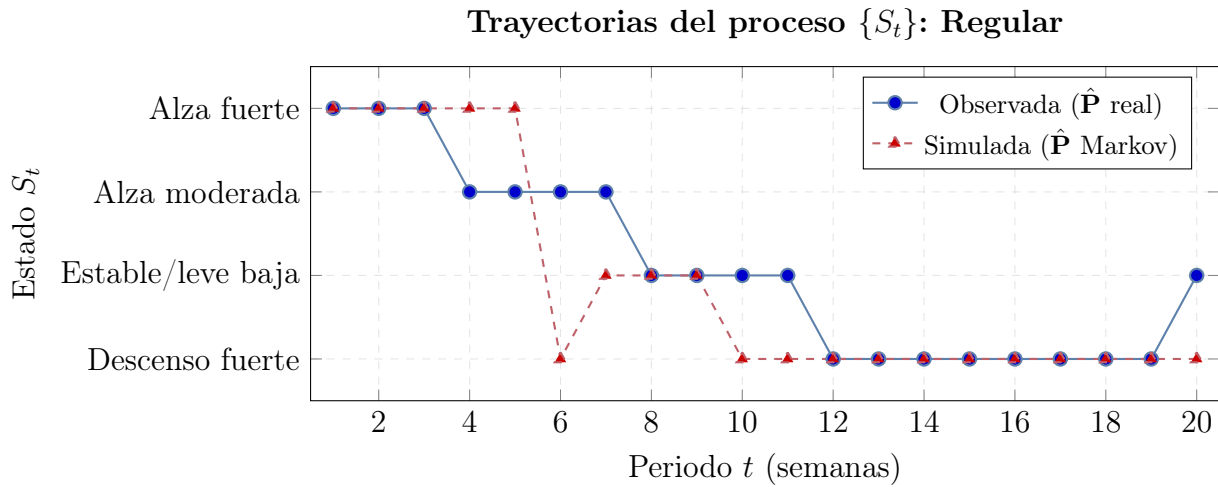


Figura A.3: Trayectorias del proceso $\{S_t\}$ para combustible **Regular** (secuencia observada vs. simulada).

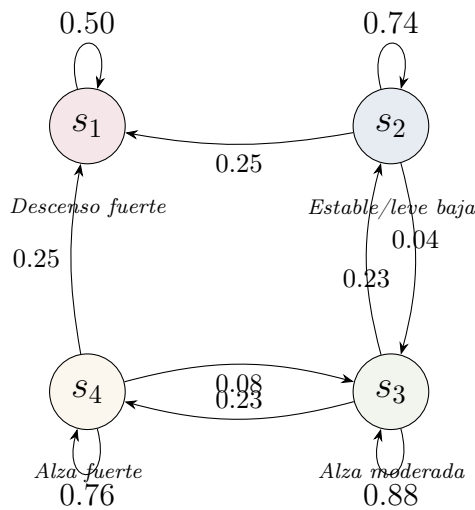


Figura A.4: Diagrama de transición entre regímenes para combustible **Regular** ($K = 4$).

B.1. Implementación Unificada del Operador TCRA

La siguiente función en Python calcula cualquier variante de la familia TCRA según los parámetros proporcionados (Sección 2.7).

```

1 import numpy as np
2
3 def compute_tcra(v, W=None, lambd=1.0):

```

Tabla A.2: Matriz de transición estimada $\hat{\mathbf{P}}$ para combustible **Regular** ($K = 4$).

| | s_1 | s_2 | s_3 | s_4 |
|-------|-------|--------------|--------------|--------------|
| s_1 | 0.500 | 0.019 | — | — |
| s_2 | 0.250 | 0.736 | 0.043 | 0.010 |
| s_3 | — | 0.226 | 0.878 | 0.225 |
| s_4 | 0.250 | 0.019 | 0.079 | 0.765 |

```

4      """
5      Operador unificado para la familia TCRA.
6      Calcula beta y alpha bajo el marco L^2 ponderado.
7
8      Parametros
9      -----
10     v      : array 1-D, serie temporal (longitud T+1).
11     W      : int o None. Ventana. Si None, W = T (datos completos).
12     lambd  : float en (0,1]. Factor de decaimiento. Default: 1.0.
13
14     Retorna
15     -----
16     alpha  : float, tasa de cambio relativa.
17     beta   : float, factor multiplicador óptimo.
18     """
19     v = np.asarray(v, dtype=float)
20     T = len(v) - 1
21
22     if W is None or W > T:
23         W = T
24     elif W < 1:
25         raise ValueError("W debe ser >= 1.")
26
27     start = T - W + 1
28     v_target = v[start : T+1]
29     v_lag     = v[start-1 : T]
30
31     weights = lambd ** np.arange(W-1, -1, -1) if lambd != 1.0 else 1.0
32

```

```
33     num = np.sum(weights * v_target * v_lag)
34     den = np.sum(weights * (v_lag ** 2))
35
36     beta = num / den if den != 0 else 1.0
37     alpha = beta - 1.0
38     return alpha, beta
```

Código 1: Función unificada para calcular las variantes de la TCRA.

B.2. Código Python para Análisis de Sensibilidad

A continuación, se presenta el código completo en Python utilizado para realizar el análisis de sensibilidad de los parámetros de discretización W y λ , como se describe en la subsección de “Análisis de Sensibilidad de los Parámetros de Discretización”.

```
1 import numpy as np
2 import pandas as pd
3
4 # --- Funciones Auxiliares ---
5
6 def calculate_alpha(v, W, lam):
7     """Calcula la serie de tasas relativas alpha_t usando la fórmula de
8         TCRA."""
9     alphas = []
10    T = len(v)
11    weights = lam ** np.arange(W - 1, -1, -1)
12
13    for t in range(W, T):
14        v_current_window = v[t-W+1 : t+1]
15        v_previous_window = v[t-W : t]
16
17        numerator = np.sum(weights * v_current_window * v_previous_window)
18        denominator = np.sum(weights * v_previous_window**2)
19
20        if denominator == 0:
21            beta_t = 1.0
```

```
21     else:
22         beta_t = numerator / denominator
23
24         alphas.append(beta_t - 1)
25
26     return np.array(alphas)
27
28 def assign_states(alphas):
29     """Asigna estados discretos segun los umbrales definidos."""
30     conditions = [
31         alphas < -0.01,
32         (alphas >= -0.01) & (alphas <= 0.02),
33         (alphas > 0.02) & (alphas <= 0.04),
34         alphas > 0.04
35     ]
36     choices = [1, 2, 3, 4]
37
38     return np.select(conditions, choices, default=0)
39
40 # --- Script Principal del Analisis de Sensibilidad ---
41
42 # 1. DATOS Y PARAMETROS
43 datos_reales_pib = np.array([
44     16214.3, 16211.7, 16381.9, 16692.9, 17296.5, 17903.1, 18634.1,
45     19337.8, 19708.9, 18813.3, 19104.7, 19431.7, 19828.9, 19983.0,
46     20199.1, 20579.2, 20982.7, 21595.6, 22019.3, 22177.7, 19838.3,
47     21961.6, 22491.3, 22900.9
48 ])
49
50 ws_to_test = [2, 3, 4, 5]
51 lambdas_to_test = np.arange(1.0, 0.79, -0.01)
52 params_to_test = [(w, lam) for w in ws_to_test for lam in lambdas_to_test]
53
54 # 2. REALIZAR EL ANALISIS
55 results = []
56 for W, lam in params_to_test:
```

```
57     alphas = calculate_alpha(datos_reales_pib, W, lam)
58     states = assign_states(alphas)
59
60     unique_states, counts = np.unique(states, return_counts=True)
61     state_counts = dict(zip(unique_states, counts))
62
63     total_states = len(states)
64     if total_states > 0:
65         s1_perc = (state_counts.get(1, 0) / total_states) * 100
66         s2_perc = (state_counts.get(2, 0) / total_states) * 100
67         s3_perc = (state_counts.get(3, 0) / total_states) * 100
68         s4_perc = (state_counts.get(4, 0) / total_states) * 100
69     else:
70         s1_perc, s2_perc, s3_perc, s4_perc = 0, 0, 0, 0
71
72     results.append({
73         'W': W,
74         'lambda': lam,
75         '% s1': s1_perc,
76         '% s2': s2_perc,
77         '% s3': s3_perc,
78         '% s4': s4_perc
79     })
80
81 # 3. MOSTRAR RESULTADOS
82 df_results = pd.DataFrame(results)
83 pd.set_option('display.max_rows', None)
84 print(df_results.round(2))
```

Código 2: Script de Python para el cálculo de distribución de estados en el análisis de sensibilidad.

B.3. Código Python para Validación con Múltiples Particiones

A continuación, se presenta el código completo en Python utilizado para realizar el análisis de robustez de la validación predictiva, como se resume en la Tabla 4.2. El script itera sobre múltiples particiones de entrenamiento/prueba para evaluar la estabilidad del rendimiento del modelo.

```
1 import numpy as np
2 import pandas as pd
3 from scipy.optimize import nnls
4
5 # --- Funciones Auxiliares ---
6
7 def calculate_alpha(v, W, lam):
8     """Calcula la serie de tasas relativas alpha_t usando la fórmula de
9         TCRA."""
10    alphas = []
11    T = len(v)
12    weights = lam ** np.arange(W - 1, -1, -1)
13    for t in range(W, T):
14        v_current_window = v[t-W+1 : t+1]
15        v_previous_window = v[t-W : t]
16        numerator = np.sum(weights * v_current_window * v_previous_window)
17        denominator = np.sum(weights * v_previous_window**2)
18        if denominator == 0:
19            beta_t = 1.0
20        else:
21            beta_t = numerator / denominator
22        alphas.append(beta_t - 1)
23    return np.array(alphas)
24
25 def assign_states(alphas):
26     """Asigna estados discretos segun los umbrales definidos."""
27     conditions = [
```

```

27     alphas < -0.01,
28     (alphas >= -0.01) & (alphas <= 0.02),
29     (alphas > 0.02) & (alphas <= 0.04),
30     alphas > 0.04
31 ]
32 choices = [1, 2, 3, 4]
33 return np.select(conditions, choices, default=0)
34
35 def SRep(datos, ss):
36     """Estimador NNLS de matriz de transicion estocastica por columnas."""
37     n0 = datos.shape[0]
38     if ss >= datos.shape[1]:
39         ss = datos.shape[1] - 1
40
41     # 1. Matriz de diseno S0 y vector objetivo S1
42     S0_data = datos[:, :ss]
43     S0 = np.kron(S0_data, np.identity(n0)).T
44     S1 = S0.T @ (datos[:, 1:(1 + ss)].T).reshape(ss * n0)
45
46     # 2. Restriccion de suma unitaria
47     C = np.kron(np.identity(n0), np.ones((1, n0)))
48
49     # 3. Sistema aumentado
50     Mr = np.zeros((n0**2 + n0, n0**2))
51     Mr[:n0**2, :] = S0.T @ S0
52     Mr[n0**2:, :] = C
53
54     rhs = np.zeros((n0**2 + n0))
55     rhs[:n0**2] = S1
56     rhs[n0**2:] = 1
57
58     # 4. Resolver NNLS
59     p_flat = nnls(Mr, rhs)[0]
60
61     # 5. Reconponer y normalizar
62     Pr = p_flat.reshape(n0, n0).T

```

```
63     col_sums = Pr.sum(axis=0)
64     col_sums[col_sums == 0] = 1 # Proteccion contra division por cero
65     Pr = Pr / col_sums
66
67     return Pr
68
69 def calculate_count_matrix(states_sequence, num_states=4):
70     """Calcula la matriz de conteo de transiciones C."""
71     count_matrix = np.zeros((num_states, num_states), dtype=int)
72     for i in range(len(states_sequence) - 1):
73         from_state = states_sequence[i] - 1
74         to_state = states_sequence[i+1] - 1
75         if from_state >= 0 and to_state >= 0:
76             count_matrix[to_state, from_state] += 1
77     return count_matrix
78
79 # --- Script Principal con Multiples Divisiones ---
80
81 # 1. DATOS Y PARAMETROS
82 datos_reales_pib = np.array([
83     16214.3, 16211.7, 16381.9, 16692.9, 17296.5, 17903.1, 18634.1,
84     19337.8, 19708.9, 18813.3, 19104.7, 19431.7, 19828.9, 19983.0,
85     20199.1, 20579.2, 20982.7, 21595.6, 22019.3, 22177.7, 19838.3,
86     21961.6, 22491.3, 22900.9
87 ])
88 W_model = 2
89 lambda_model = 1.0
90 NUM_STATES = 4
91
92 # 2. GENERACION DE DATOS BASE
93 alphas_full = calculate_alpha(datos_reales_pib, W_model, lambda_model)
94 states_full = assign_states(alphas_full)
95 num_samples = len(states_full)
96 X_full = np.zeros((NUM_STATES, num_samples))
97 for i, state in enumerate(states_full):
98     if state > 0:
```

```
99         X_full[state - 1, i] = 1
100
101 C_full = calculate_count_matrix(states_full, num_states=NUM_STATES)
102 P_hat_full = SRep(X_full, ss=X_full.shape[1] - 1)
103
104 # 3. BUCLE PARA PROBAR DIFERENTES DIVISIONES
105 split_ratios_to_test = np.arange(0.60, 1.0, 0.05)
106
107 for ratio in split_ratios_to_test:
108     train_perc = int(ratio * 100)
109     test_perc = 100 - train_perc
110
111     print("\n" + "#" * 70)
112     print(f"###   RESULTADOS PARA DIVISION {train_perc}% / {test_perc}%
113           ###")
114     print("#" * 70 + "\n")
115
116     # DIVISION EN ENTRENAMIENTO Y PRUEBA
117     split_index = int(num_samples * ratio)
118     states_train = states_full[:split_index]
119     states_test = states_full[split_index:]
120     X_train = X_full[:, :split_index]
121
122     # CALCULO DE MATRICES
123     C_train = calculate_count_matrix(states_train, num_states=NUM_STATES)
124     C_test = calculate_count_matrix(states_test, num_states=NUM_STATES)
125     P_hat_trained = SRep(X_train, ss=X_train.shape[1] - 1)
126
127     # EVALUACION PREDICTIVA
128     predictions = []
129     actual_outcomes = []
130     for i in range(len(states_test) - 1):
131         current_state_idx = states_test[i] - 1
132         prob_next_state = P_hat_trained[:, current_state_idx]
133         predicted_state = np.argmax(prob_next_state) + 1
134         predictions.append(predicted_state)
```

```
134     actual_next_state = states_test[i+1]
135     actual_outcomes.append(actual_next_state)
136
137     # SALIDA PARA ESTA DIVISION
138     idx = [f'a s{i+1}' for i in range(4)]
139     cols = [f'de s{i+1}' for i in range(4)]
140
141     print("--- Matriz de Conteo C (Entrenamiento) ---")
142     print(pd.DataFrame(C_train, index=idx, columns=cols))
143
144     print("\n--- Matriz de Conteo C (Prueba) ---")
145     print(pd.DataFrame(C_test, index=idx, columns=cols))
146
147     print("\n--- Matriz de Conteo C (Completa) ---")
148     print(pd.DataFrame(C_full, index=idx, columns=cols))
149
150     print("\n--- Matriz P_hat (Entrenamiento) ---")
151     print(pd.DataFrame(P_hat_trained, index=idx, columns=cols).round(4))
152
153     print("\n--- Matriz P_hat (Completa) ---")
154     print(pd.DataFrame(P_hat_full, index=idx, columns=cols).round(4))
155     print("\n" + "="*50 + "\n")
156
157     print("--- Desglose de Predicciones en Prueba ---")
158     prediction_df = pd.DataFrame({
159         'Estado Actual (t)': states_test[:-1],
160         'Real (t+1)': actual_outcomes,
161         'Predicho (t+1)': predictions
162     })
163     if not prediction_df.empty:
164         correct = (prediction_df['Real (t+1)'] ==
165                  prediction_df['Predicho (t+1)'])
166         prediction_df['Correcto?'] = correct
167     print(prediction_df)
168
169     # METRICAS RESUMEN
```

```
170     correct_predictions = prediction_df['Correcto?'].sum() if 'Correcto?'
      in prediction_df else 0
171 total_predictions = len(prediction_df)
172 accuracy = correct_predictions / total_predictions if total_predictions
      > 0 else 0
173
174 print("\n--- Metricas Resumen ---")
175 print(f"Predicciones correctas: {correct_predictions} de {
      total_predictions}")
176 print(f"Precision (Accuracy): {accuracy:.3f}")
```

Código 3: Script de Python para validación con múltiples particiones.

B.4. Código Fuente: Estimación y Simulación

El siguiente listado contiene la implementación del algoritmo SRep para la estimación de matrices de transición mediante Mínimos Cuadrados No Negativos (NNLS), así como la simulación de la evolución de la cadena de Markov.

C.1. Código Fuente: Computación de Reservorio (SSRC)

Los siguientes listados implementan las funciones centrales del modelo SSRC descrito en el Capítulo 6. El código fuente completo está disponible en el repositorio del proyecto.

D.1. Código para Generación de Figuras

Los siguientes listados contienen los scripts de Python utilizados para generar las figuras principales de la tesis. Todos los gráficos usan la paleta de colores Nord y calidad IEEE (600 DPI).

Código 1 Implementación del algoritmo SRep para estimar matrices de transición estocásticas mediante NNLS.

```

1 import numpy as np
2 from scipy.optimize import nnls
3
4 def SRep(datos, ss):
5     """Estimador NNLS de matriz de transicion estocastica por columnas."""
6     n0 = datos.shape[0]
7
8     # 1. Matriz de diseno S0 y vector objetivo S1
9     S0_data = datos[:, :ss]
10    S0 = np.kron(S0_data, np.identity(n0)).T
11    S1 = S0.T @ (datos[:, 1:(1 + ss)].T).reshape(ss * n0)
12
13    # 2. Restriccion de suma unitaria
14    C = np.kron(np.identity(n0), np.ones((1, n0)))
15
16    # 3. Sistema aumentado
17    Mr = np.zeros((n0**2 + n0, n0**2))
18    Mr[:n0**2, :] = S0.T @ S0
19    Mr[n0**2:, :] = C
20    rhs = np.zeros((n0**2 + n0))
21    rhs[:n0**2] = S1
22    rhs[n0**2:] = 1
23
24    # 4. Resolver NNLS
25    p_flat = nnls(Mr, rhs)[0]
26
27    # 5. Recomponer y normalizar
28    Pr = p_flat.reshape(n0, n0).T
29    col_sums = Pr.sum(axis=0)
30    col_sums[col_sums == 0] = 1 # Proteccion contra division por cero
31    Pr = Pr / col_sums
32    return Pr

```

Código 2 Simulación de la evolución temporal de una cadena de Markov a partir de la matriz de transición \hat{P} .

```

1 # Simulacion del proceso de Markov
2 T = 20
3 p0 = np.zeros((4, T))
4 p0[0, 0] = 1 # Estado inicial: s1
5
6 for j in range(T - 1):
7     p0[:, j+1] = P_hat @ p0[:, j]

```

Código 3 Generación de las matrices fijas del reservorio (Definición 6.4). El reescalamiento espectral garantiza $\rho(\mathbf{W}_{\text{res}}) = \rho_{\text{res}} < 1$, condición suficiente para la propiedad ESP (Teorema 6.1).

```

1 import numpy as np
2 from typing import Tuple, Optional
3
4 def create_reservoir(
5     input_dim: int,
6     reservoir_dim: int,
7     spectral_radius: float,
8     input_scale: float = 1.0,
9     sparsity: float = 0.9,
10    seed: Optional[int] = None
11) -> Tuple[np.ndarray, np.ndarray]:
12    """
13    Genera las matrices fijas del reservorio (W_in, W_res).
14    Implementa la Definición 6.3 del capítulo SSRC.
15
16    Parametros
17    -----
18    input_dim : int      -- Dimension de la entrada (d=1 para TCRA).
19    reservoir_dim : int  -- Dimension del reservorio (D).
20    spectral_radius : float -- Radio espectral rho_res < 1 (ESP).
21    input_scale : float -- Escala omega de W_in.
22    sparsity : float    -- Fraccion de ceros en W_res.
23    seed : int         -- Semilla para reproducibilidad.
24
25    Retorna
26    -----
27    W_in : ndarray (D, d), W_res : ndarray (D, D)
28    """
29    rng = np.random.RandomState(seed)
30
31    # W_in: Uniforme en [-omega, omega]
32    W_in = rng.uniform(-input_scale, input_scale,
33                      size=(reservoir_dim, input_dim))
34
35    # W_res: Dispersa aleatoria, reescalada espectralmente
36    W_res_raw = rng.randn(reservoir_dim, reservoir_dim)
37    mask = rng.rand(reservoir_dim, reservoir_dim) > sparsity
38    W_res_raw *= mask
39
40    # Reescalar: W_res <- rho_res * W_res^(0) / rho(W_res^(0))
41    eigenvalues = np.linalg.eigvals(W_res_raw)
42    rho_current = np.max(np.abs(eigenvalues))
43
44    if rho_current > 0:
45        W_res = spectral_radius * W_res_raw / rho_current
46    else:
47        W_res = W_res_raw
48
49    return W_in, W_res

```

Código 4 Propagación de la señal α_t a través del reservorio Leaky-ESN. Los primeros washout estados se descartan para eliminar transitorios de la condición inicial $\mathbf{h}_0 = \mathbf{0}$.

```

1 import numpy as np
2
3 def propagate_reservoir(
4     alpha_series, W_in, W_res,
5     washout=50, leak_rate=1.0
6 ):
7     """
8     Propaga la senal TCRA alpha_t a traves del reservorio.
9     Implementa el Leaky-ESN (Definición 6.2):
10     h_t = (1-a)*h_{t-1} + a*tanh(W_in*u_t + W_res*h_{t-1})
11     Cuando a = 1.0, se reduce al ESN clasico.
12
13     Parametros
14     -----
15     alpha_series : ndarray (T,) -- Serie TCRA alpha_t.
16     W_in : ndarray (D, d)      -- Matriz de entrada.
17     W_res : ndarray (D, D)     -- Matriz de reservorio.
18     washout : int              -- Pasos transitorios.
19     leak_rate : float          -- Factor de fuga a in (0, 1].
20
21     Retorna
22     -----
23     H : ndarray (D, T_eff) -- Estados post-washout.
24     """
25     T = len(alpha_series)
26     D = W_res.shape[0]
27     U = alpha_series.reshape(-1, 1)
28
29     H_all = np.zeros((D, T))
30     h = np.zeros(D)
31     a = leak_rate
32
33     for t in range(T):
34         h = (1.0-a)*h + a*np.tanh(W_in @ U[t] + W_res @ h)
35         H_all[:, t] = h
36
37     H = H_all[:, washout:]
38     return H

```

Código 5 Script para generar la Figura 4.7: evolución temporal de las probabilidades de ocupación de estados bajo la dinámica de la matriz \hat{P} .

```

1  """
2  Genera la figura 'evolucion_probabilidades_markov.png'
3  Evolucion de probabilidades de la Matriz P estimada (K=4 regimenes)
4  con paleta Nord para la tesis.
5  """
6  import numpy as np
7  import matplotlib
8  matplotlib.use('Agg')
9  import matplotlib.pyplot as plt
10
11 # Paleta Nord
12 NORD_0 = "#2E3440"
13 NORD_3 = "#4C566A"
14 NORD_10 = "#5E81AC" # Frost Blue -- alza fuerte
15 NORD_11 = "#BF616A" # Aurora Red -- caida
16 NORD_13 = "#EBCB8B" # Aurora Yellow -- subida moderada
17 NORD_14 = "#A3BE8C" # Aurora Green -- estable
18
19 plt.style.use('default')
20 plt.rcParams.update({
21     'font.family': 'serif',
22     'font.serif': ['Times New Roman', 'DejaVu Serif'],
23     'font.size': 12,
24 })
25
26 # Matriz P estimada (ejemplo Super, K=4)
27 P = np.array([
28     [0.20, 0.30, 0.30, 0.05],
29     [0.35, 0.33, 0.43, 0.05],
30     [0.40, 0.32, 0.22, 0.05],
31     [0.05, 0.05, 0.05, 0.85],
32 ])
33 P = P / P.sum(axis=0, keepdims=True)
34
35 # Estado inicial: p0 = (1, 0, 0, 0)
36 K = P.shape[0]
37 p0 = np.zeros(K)
38 p0[0] = 1.0
39
40 # Simular evolucion
41 n_steps = 20
42 trajectory = np.zeros((K, n_steps))
43 trajectory[:, 0] = p0
44 for t in range(1, n_steps):
45     trajectory[:, t] = P @ trajectory[:, t-1]
46
47 colors = [NORD_11, NORD_14, NORD_13, NORD_10]
48 labels = ['caida', 'estable', 'subida', 'alza fuerte']
49
50 fig, ax = plt.subplots(figsize=(12, 7))
51 for k in range(K):
52     ax.plot(range(n_steps), trajectory[k, :],
53           marker='D', markersize=5, linewidth=2.5,
54           color=colors[k], label=labels[k])
55 ax.set_xlabel('Tiempo')
56 ax.set_ylabel('Probabilidad')
57 ax.set_title('Evolucion de Probabilidades (Matriz P Estimada)')
58 ax.legend(fontsize=12)
59 ax.grid(True, alpha=0.3, linestyle=':')
60 plt.savefig('evolucion_probabilidades_markov.png',
61           dpi=600, bbox_inches='tight')

```

Código 6 Script para generar la Figura 4.8: panel 2×2 con la evolución de probabilidades para las cuatro series de combustibles.

```
1 """
2 Genera figura panel 2x2 con evoluciones de probabilidades
3 para las 4 series de combustibles (Superior, Regular, Diésel, Kerosene).
4 """
5 import numpy as np
6 import matplotlib.pyplot as plt
7 from scipy.optimize import nnls
8
9 # Cargar datos de cada serie
10 def load_transition_matrix(fuel_name, output_dir):
11     """Carga la matriz de transicion desde archivos .npy generados
12     por el pipeline 03_modeling.py."""
13     P = np.load(f'{output_dir}/{fuel_name}_P_hat.npy')
14     return P
15
16 def simulate_evolution(P, n_steps=100, initial_state=0):
17     """Simula la evolucion de la distribucion de probabilidad
18     a partir de un estado inicial dado."""
19     K = P.shape[0]
20     p0 = np.zeros(K)
21     p0[initial_state] = 1.0
22
23     traj = np.zeros((K, n_steps))
24     traj[:, 0] = p0
25     for t in range(1, n_steps):
26         traj[:, t] = P @ traj[:, t-1]
27     return traj
28
29 # Configuracion de la grilla 2x2
30 fuels = ['Super', 'Regular', 'Diesel', 'Kerosene']
31 fig, axes = plt.subplots(2, 2, figsize=(14, 10))
32 for idx, (fuel, ax) in enumerate(zip(fuels, axes.flat)):
33     P = load_transition_matrix(fuel, 'outputs/models')
34     traj = simulate_evolution(P, n_steps=100)
35     K = P.shape[0]
36     for k in range(K):
37         ax.plot(range(100), traj[k, :], linewidth=2)
38         ax.set_title(fuel, fontweight='bold')
39         ax.set_xlabel('Tiempo')
40         ax.set_ylabel('Probabilidad')
41         ax.grid(True, alpha=0.3)
42 plt.suptitle('Evolucion de Probabilidades (Matriz P Estimada)',
43             fontweight='bold', fontsize=16)
44 plt.tight_layout()
45 plt.savefig('allseriecombustiblesNNLS.png', dpi=600)
```