
**TÉCNICAS ESTADÍSTICAS PARA EL ANÁLISIS
DE COMPLEJIDAD DE SERIES DE TIEMPO**

Tesis que para obtener el grado de:
Doctor en Ciencias (Ingeniería Química)

Presenta:
M. en C. Jesús Alejandro Ortíz Cruz



Universidad Autónoma Metropolitana Unidad Iztapalapa

División de Ciencias Básicas e Ingeniería

Departamento de Ingeniería de Procesos e Hidráulica

Posgrado en Ingeniería Química

**TÉCNICAS ESTADÍSTICAS PARA EL ANÁLISIS
DE COMPLEJIDAD DE SERIES DE TIEMPO**

Tesis que para obtener el grado de:

Doctor en Ciencias (Ingeniería Química)

Presenta:

M. en C. Jesús Alejandro Ortíz Cruz

Asesor:

Dr. José de Jesús Álvarez Ramírez

México, D.F., a 31 de Enero de 2013

Agradecimientos

A mi asesor, el Dr. José de Jesús Álvarez Ramírez, al cual estoy profundamente agradecido por darme la oportunidad de trabajar en este proyecto y por guiarme apropiadamente durante toda esta etapa. También por el inigualable apoyo durante los difíciles momentos que acontecí junto con mi familia durante mi estancia en el posgrado.

A los doctores Eduardo Rodríguez Flores, Juan Carlos Echeverría Arjonilla, Rafael Escalera Pérez, Jesús Isidro González Trejo, Ricardo Aguilar López quienes formaron parte del comité revisor de la presente tesis. Gracias por sus comentarios y el tiempo dedicado a que este trabajo se enriqueciera más.

A mi esposa Ernestina quien siempre me ha apoyado en todas mis decisiones y se ha aventurado conmigo en este camino de la investigación.

A mi hijo Alejandro Santiago quien me brindó la fuerza para continuar esforzándome para concluir esta etapa de mi vida y sigue alentando a que cada día de lo mejor de mí.

A mi bebe Ricardo Emiliano quien se integra a la familia Ortiz Moreno y ahora junto con su hermano son mis razones de que siga en búsqueda de la felicidad y el éxito para mi familia.

A mi familia, quienes siempre estuvieron al pendiente y nos brindaron su apoyo a pesar de la distancia.

A todos mis amigos que siempre estuvieron en las buenas y en las malas conmigo y mi familia.

Finalmente, al Consejo Nacional de Ciencia y Tecnología por el apoyo económico brindado durante todo este tiempo para que pudiera concluir satisfactoriamente el presente trabajo de investigación.

Resumen

El registro de datos de variables de fenómenos natural ó procesos químicos en forma de series de tiempo se ha comenzado a recopilar desde hace varias décadas. El objetivo del registro es estudiar las series de tiempo para visualizar, entender, comprender y revelar comportamientos del sistema en estudio. Para lograr este fin se han desarrollado técnicas estadísticas que permiten analizar las series de tiempo; sin embargo, la principal consideración detrás de estas técnicas es que fueron diseñadas para trabajar con series de tiempo que satisfagan el criterio del estacionario, algo que comúnmente violan los sistemas complejos reales.

El presente trabajo de investigación desarrolla e implementa algoritmos estadísticos que permiten analizar series de tiempo de sistemas reales complejos, este tipo de sistemas se caracterizan por presentar un comportamiento fuertemente no lineal y además tienen la particularidad de contener una gran cantidad de ruido asociado.

La teoría básica para el correcto entendimiento del trabajo se presenta en el Capítulo 3. Las técnicas de análisis de complejidad desarrolladas tienen su fundamento en los conceptos de la entropía aproximada y del exponente de Hurst. El primero de estos conceptos nos proporciona una medida de la cantidad de configuraciones y/o patrones existentes en la serie de tiempo, mientras que el segundo nos proporciona una medida de la existencia de correlaciones de corto y largo alcance en la serie de tiempo y/o la falta de estas correlaciones.

Las metodologías propuestas y desarrolladas del presente trabajo de investigación se presentan en el Capítulo 4. Los casos de estudio se seleccionaron en el siguiente orden: primero se analizaron señales sintéticas del tipo $1/f^\beta$, posteriormente se validó la robustez de los métodos al analizar series de tiempo de sistemas complejos de la literatura especializada. Por ejemplo, análisis de complejidad de series de tiempo para fluctuaciones de ritmos cardíacos, la aplicación final de los métodos de análisis de complejidad se llevó a cabo al estudiar series de tiempo de variables sísmicas, de precios del petróleo y de datos de difracción de rayos X.

ÍNDICE GENERAL

Índice general	I
Índice de figuras	IV
Índice de tablas	VI
1. Introducción	1
1.1. Motivación	2
1.2. Antecedentes	3
1.3. Justificación	7
1.4. Objetivos	8
1.4.1. Objetivo general	8
1.4.2. Objetivos específicos	8
1.5. Estructura de la tesis	9
1.6. Artículos derivados de la tesis	9
2. Marco Teórico	10
2.1. Series de tiempo	10
2.1.1. Ejemplos representativos	10
2.1.2. Terminología	14

2.1.3.	Objetivos del análisis	15
2.1.4.	Métodos para analizar series de tiempo	17
2.1.5.	Técnicas descriptivas simples	18
2.2.	Gráfica de tiempo	21
2.2.1.	Transformaciones	22
2.3.	Analizando series que contienen tendencia	23
2.3.1.	Ajuste de curva	24
2.3.2.	Filtrado	25
2.3.3.	Filtros en series	28
2.3.4.	Diferenciación	29
2.4.	Analizando series que contienen variación estacional	29
3.	Métodos de estudio	32
3.1.	Entropía aproximada	32
3.2.	Entropía aproximada con remoción de tendencia	34
3.3.	Entropía aproximada multiescala	36
3.4.	Entropía asimetría temporal	37
3.5.	Análisis de rango reescalado	39
4.	Casos de estudio	43
4.1.	Señales sintéticas tipo $1/f^\beta$	43
4.2.	Series de tiempo de fluctuaciones de ritmos cardíacos	45
4.3.	Señales de intensidad de difracción de rayos X	50
4.3.1.	Introducción	51
4.3.2.	Análisis de complejidad	53
4.3.3.	Análisis fractal	61
4.4.	Eficiencia de mercados de petróleo crudo	67
4.4.1.	Introducción	67
4.4.2.	Metodología	71

4.4.3. Resultados	72
4.4.4. Conclusiones	81
4.5. Variaciones temporales de eventos sísmicos	82
4.5.1. Introducción	83
4.5.2. Metodología	87
4.5.3. Resultados	88
4.5.4. Conclusiones	89
5. Conclusiones	90
5.1. Entropía Aproximada	91
5.2. Exponente de Hurst	93
Bibliografía	95

ÍNDICE DE FIGURAS

2.1. Serie de tiempo del índice de precios anual del trigo Beveridge de los años 1500 a 1869.	11
2.2. Promedio de temperaturas mensuales alrededor del mundo de 1940-2011. Estación meteorológica en Recife, Brazil.	11
2.3. Ventas mensuales de automóviles en Quebec, Canada en el período de 1960-1968.	12
2.4. Total de nacimientos femeninos diarios en California, 1959.	13
2.5. Gráfico de control de proceso.	13
2.6. Representación de un proceso binario.	13
2.7. Representación de un proceso puntual.	14
2.8. Representación esquemática de un proceso lineal.	16
2.9. Representación esquemática de un proceso lineal.	28
4.1. Análisis <i>DAE</i> para señales sintéticas con ruido tipo $1/f^\beta$	44
4.2. Entropía aproximada multiescala de señales sintéticas $1/f^\beta$	45
4.3. Análisis <i>DAE</i> para las señales del movimiento Browniano y su correspondiente señal desordenada aleatoriamente.	46
4.4. Series de tiempo características de los grupos de individuos (a) sanos, (b) con insuficiencia cardíaca y (c) con fibrilación auricular respectivamente.	47

4.5. Análisis de <i>DAE</i> representativo de series de tiempo de ritmos cardíacos.	48
4.6. Promedio del análisis <i>DAE</i> para cada uno de los tres casos de estudio.	48
4.7. <i>DAE</i> para ruido blanco con un componente de retardo Kx_{i-1} , $0 < K < 1$	49
4.8. <i>DAE</i> para el ruido blanco con un componente de retraso Kx_{i-5}	50
4.9. Intensidad de patrones $I(2\theta)$ de rayos X para alúmina calcinada a 200 y 500 °C.	53
4.10. (a) Intensidad de patrones $I(2\theta)$ de rayos X , (b) Patrones de entropía $E(2\theta)$ de la señal de intensidad $I(2\theta)$ de alúmina calcinada a 200°C.	57
4.11. (a) Intensidad de patrones $I(2\theta)$ de rayos X , (b) Patrones de entropía $E(2\theta)$ de la señal de intensidad $I(2\theta)$ de alúmina calcinada a 500°C.	58
4.12. Patrones de entropía $E(2\theta)$ como función de la temperatura de calcinación para cuatro diferentes regiones angulares.	60
4.13. (a) Intensidad de patrones $I(2\theta)$ de rayos X , (b) Patrón del exponente de Hurst para la intensidad de patrones $I(2\theta)$ de rayos X para alúmina calcinada a 200°C.	64
4.14. (a) Intensidad de patrones $I(2\theta)$ de rayos X , (b) Patrones del exponente de Hurst de la señal de intensidad $I(2\theta)$ de alúmina calcinada a 500°C.	65
4.15. Exponente de Hurst como función de la temperatura de calcinación para cuatro diferentes regiones angulares.	66
4.16. (a) Precios del WTI para el período del 1 de Enero de 1986 al 15 de Marzo del 2011. (b) Diferencias logarítmicas de precios para el mismo período del panel (a).	72
4.17. Secuencias de diferencia de precios logarítmicas utilizando un filtro pasa bajas para escalas (a) diaria ($n = 1$), (b) semanal ($n = 5$), (c) mensual ($n = 20$) y (c) trimestral ($n = 60$).	73
4.18. Patrones de entropía multiescala $AE(\tau)$ para el subperíodo 1986-1989 de acuerdo con las secuencias filtradas pasa bajas de la Figura 4.17.	73
4.19. Comportamiento de la entropía como función del tamaño de ventanas N_s para escalas temporales diarias, semanales, mensuales y trimestrales.	74

4.20. Comportamiento de la entropía con respecto a la escala de tiempo τ , para cinco subperíodos seleccionados de tres años.	76
4.21. (a) Límites de referencia de los patrones de entropía para 10,000 muestras de secuencias gaussianas de ruidos no correlacionados de longitud $N_s = 600$ observaciones. (b) Patrón de entropía en comparación con el punto de referencia para la aleatoriedad.	77
4.22. Índice de eficiencia informativa del mercado $I_{IME}(\tau)$ con respecto al tiempo t y escala τ	80
4.23. (a) Mapa esquemático y sectores de la deformación tectónica activa a lo largo de la costa del Sur del margen activo mexicano. (b) Sismicidad del período 1998-2011 de acuerdo con el catálogo del Servicio Sismológico Nacional, México.	83
4.24. Diagrama de frecuencia-magnitud de la sismicidad de la Fig. 4.23.b.	88
4.25. Comportamiento del exponente de Hurst como función del tiempo t y escala N_s	88
4.26. Detalles de los patrones del exponente de Hurst mostrados en la Fig.4.25. . .	89

ÍNDICE DE TABLAS

- 1.1. Tipo de correlación en una serie de tiempo con base en el exponente de Hurst. 7

CAPÍTULO 1

INTRODUCCIÓN

El análisis de series de tiempo es una herramienta que ha sido empleada con gran intensidad en las últimas décadas con el propósito de estudiar y entender el comportamiento de series de tiempo de sistemas complejos de interés. En estos estudios, en general, se analizan los principales componentes de las series como son: tendencias, fluctuaciones cíclicas y/o estacionales, irregularidades y/o patrones presentes. Estudiar estas componentes no es algo trivial, por lo que se han desarrollado métodos que tienen en cuenta estas componentes; sin embargo, la mayoría de los métodos tienen una orientación del tipo lineal que en varios casos no proporcionan resultados satisfactorios cuando los sistemas de estudio tienen comportamientos no lineales complejos. Por lo tanto, ni los métodos estadísticos clásicos ni los métodos actuales para el estudio de series de tiempo proporcionan resultados satisfactorios cuando son aplicados al estudio de sistemas complejos; por ejemplo, Guillen y col. (2005) demostraron que la transformada rápida de Fourier (*FFT*) aplicada al monitoreo de sistemas industriales no es adecuada para analizar fallas presentes en la instrumentación de un proceso, debido a la presencia de componentes no estacionarios (fenómenos oscilantes localizados en el tiempo), transitorios y/o cambios abruptos que exhiben patrones de frecuencia variable en el tiempo. Jáuregui-Correa y Gonzalez-Brambila (2010) aplicaron la *FFT* en la identificación de discontinuidades en las propiedades de rigidez de sistemas mecánicos encontrando que el

método solo es adecuado para señales lineales y estacionarias, con resultados aceptables en la práctica.

El presente trabajo de investigación enfoca su estudio al desarrollo de técnicas estadísticas que permitan analizar y estudiar series de tiempo generadas por sistemas complejos. Sistemas que, en general, presentan fuertes comportamientos no lineales y además poseen la característica de tener gran contenido de ruido.

1.1. Motivación

El análisis de series de tiempo ha sido parte importante en los estudios relacionados con econometría, que es la ciencia que unifica la estadística, la teoría económica y las matemáticas para una comprensión real de las relaciones cuantitativas de la vida económica moderna. Es la unión de estas tres disciplinas lo que constituye una herramienta de análisis potente, constituyendo así lo que es la econometría (Frisch, 1930). Sin embargo, la aplicación del análisis de series de tiempo ha trascendido a otras disciplinas del conocimiento como en ingeniería biomédica donde se han estudiado las fluctuaciones del ritmo cardíaco (Costa y col., 2008; Goldberger y col., 2002; Pincus, 2006), mientras que en geología ha sido útil al estudiar series de tiempo de registros sísmicos (Astiz y Kanamori, 1984; Balankin y col., 2009; Day-Lewis y col., 2010). En la ingeniería mecánica el análisis de series de tiempo de registros de maquinarias ha revelado daños asociados con aspectos de vibración (Yan y Gao, 2007; Pérez-Canales y col., 2011) y en la física se ha estudiado el caos presente en las series de tiempo de registros pertenecientes a equipos láser NMR (por sus siglas en inglés, Nuclear Magnetic Resonance) (Elder y col., 2006), estos son sólo algunos ejemplos. Dada la diversidad de disciplinas en donde se está empleando con éxito el análisis de series de tiempo se esperaría que en la ingeniería química sucediera lo mismo, pero este no es el caso, basta con realizar una búsqueda sobre trabajos de investigación en esta área del conocimiento y es fácil darse cuenta que son escasos los trabajos donde se ha explotado su uso hasta el momento. Por lo que surge en forma natural la siguiente pregunta, ¿Por qué no se ha explotado la aplicación del análisis de series de tiempo en la ingeniería química?. Probablemente por la

complejidad de las señales que generan los procesos químicos de interés y porque además los métodos actuales no han dado resultados satisfactorios. El presente trabajo motivado por esta interrogante enfoca su estudio al desarrollo y/o modificación de técnicas estadísticas que permitan resolver este problema.

1.2. Antecedentes

El estudio de series de tiempo se lleva a cabo mediante dos categorías de estudio principal, los métodos que se basan en el dominio de la frecuencia y los basados en el dominio del tiempo. Los métodos basados en el dominio de la frecuencia incluyen el análisis espectral y recientemente el análisis de wavelets; mientras que por los métodos basados en el dominio del tiempo podemos mencionar los análisis de auto correlación y correlación cruzada.

Aunque los métodos antes mencionados funcionan bien en la mayoría de los casos, existen dificultades cuando se emplean en series de tiempo generadas por sistemas complejos. La razón se debe al comportamiento altamente no lineal de los sistemas complejos y a la presencia de ruido en estas señales. Estas características se ven reflejadas en comportamientos anormales de los componentes principales de las series de tiempo como son tendencia, estacionalidad y aleatoriedad.

El estudio de estos sistemas complejos ha generado nuevos desafíos para poder estudiar las series de tiempo que pertenecen a estos sistemas. Se han propuesto diversas metodologías; por ejemplo, el análisis de rango reescalado R/S (Hurst y col., 1965), el análisis de fluctuación sin tendencia (Peng y col., 1994) para afrontar el problema, la entropía aproximada (AE). Donde el concepto de entropía tiene la ventaja de robustez frente al ruido implícito en las series de tiempo por analizar. Si bien el concepto es ampliamente conocido por su estudio en la termodinámica, su empleo en la presente investigación es más bien desde el punto de vista de la teoría de información. En este mismo campo, Shannon (1948) en sus trabajos sobre teoría de la información acuñó el término entropía con el fin de cuantificar la incertidumbre que se transmite por un canal de información. El término entropía de información ha sido asociado a la irregularidad o complejidad en series de tiempo y la relevancia de su empleo

consiste en su fácil interpretación física. Por ejemplo, la entropía de Kolmogorov-Sinai mide la velocidad exponencial a la cual la información es obtenida y la entropía de Shannon mide la incertidumbre de una fuente de información.

Ahora bien, el inconveniente de los conceptos tradicionales de la entropía, como la entropía de Shannon o la de Kolmogorov-Sinai es la necesidad de una gran cantidad de datos libres de ruido con una alta resolución. Razón por la cual es impráctica su cuantificación con métodos computacionales tradicionales (Dorfman, 1999). Por otra parte, las señales pertenecientes a sistemas reales complejos son escasas en datos y generalmente contienen ruido. Un ejemplo de éstas son las señales experimentales, escasas en datos y con ruido asociado a equipos obsoletos o mal calibrados y/o a falta de experiencia por parte del técnico quien llevo a cabo el experimento y también a factores externos imprevistos o a la misma naturaleza del sistema en estudio.

Para resolver este problema Pincus (1991) introdujo el estadístico denominado entropía aproximada, AE , el cual tiene su fundamento en la entropía de Shannon. Este estadístico permite cuantificar el grado de irregularidad o complejidad en series de tiempo de sistemas reales complejos. Pincus (1991) empleó este estadístico al estudio de series de tiempo fisiológicas encontrando que es un discriminante entre series de tiempo pertenecientes a sujetos sanos y enfermos, algo que métodos tradicionales no pudieron discriminar.

Su aplicación se ha extendido a otras disciplinas distintas de la física biológica. Pueden verse ejemplos de su aplicación en el análisis de datos financieros con datos empíricos y modelos, incluidos los índices compuestos e individuales; la hipótesis del caminante aleatorio y del movimiento Browniano fraccional (Pincus y Kalman, 2004), el monitoreo de máquinas de cortado (Pérez-Canales y col., 2011) y el análisis de señales fisiológicas (Richman y Moorman, 2000) por mencionar algunos.

Costa y col. (2002) extendieron el análisis de la entropía aproximada de Pincus (1991) a una metodología multiescala, encontrando que un análisis multiescala puede extraer mayor información de la serie de tiempo. El nuevo estadístico desarrollado recibió el nombre de *entropía aproximada multiescala* y tiene fundamento en la AE . La principal consideración

propuesta es que tanto las señales completamente ordenadas como las completamente aleatorias no son del todo complejas. En un estudio posterior, Costa y col. (2005) emplearon el análisis de entropía multiescala en series de tiempo de ritmos cardíacos humanos de pacientes en condiciones fisiológicas y patológicas graves, encontrando que existe una pérdida de complejidad del sistema cardíaco con la edad (casos de arritmias cardíacas erráticas e insuficiencia cardíaca).

Esta nueva metodología de análisis de series de tiempo ha tomado gran interés en investigaciones recientes relacionadas con señales de sistemas complejos; por ejemplo, Martina y col. (2011) aplicaron el análisis de entropía multiescala al monitoreo de la dinámica del precio del petróleo. Los autores encontraron que altos valores de entropía corresponden a una dinámica más compleja y por lo tanto a un mercado menos predecible. Además el estudio reveló que, en los recientes 25 años las recesiones americanas coinciden con períodos de bajos niveles de entropía. En otro estudio realizado por Alvarez-Ramirez y Rodriguez (2011) emplearon el análisis de entropía multiescala al estudio de los registros del índice Dow Jones de 1928-2010 y encontraron la existencia de períodos dominantes de 4.5 y 22 años en la dinámica. Los autores sugieren que la crisis actual no es singular, sino que sus orígenes pueden explicarse en términos de patrones cíclicos de largo alcance. Para mayores referencias sobre el tema se puede consultar la investigación más reciente de Alvarez-Ramirez y col. (2012).

Otro enfoque en el análisis de complejidad de series de tiempo de sistemas complejos es el análisis fractal. El concepto de fractal está asociado con objetos geométricos que satisfacen dos criterios: *auto-similitud* y *dimensión fractal*. La *auto-similitud* significa que un objeto está compuesto de subunidades y éstas a su vez en sub-subunidades, en un proceso iterativo, de tal forma que siempre se tiene la forma original del objeto completo (Feder, 1988). Matemáticamente esta propiedad debe aplicarse a todas las escalas; sin embargo, en la realidad esto no siempre sucede ya que existen límites en las cuales es aplicable el concepto de auto-similitud. El segundo criterio, la *dimensión fractal* significa que un objeto es fractal si tiene una dimensión fraccionaria. Este requisito distingue los objetos fractales de los objetos

que cumplen la geometría de Euclides, los cuales tienen dimensión entera.

Para estimar la *dimensión fractal* puede emplearse el concepto del *exponente de Hurst*, nombrado así en honor de Harold Edwin Hurst (1951), hidrólogo inglés que estudió los registros egipcios de los desbordamientos del río Nilo para encontrar si el fenómeno era puramente aleatorio o tenía alguna relación con otros fenómenos naturales (por ejemplo, registros de precipitación pluvial, manchas solares, etc.). En su estudio Hurst (1951) desarrolló el método del (*rango reescalado R/S*) con el que pudo concluir que el fenómeno de estudio no era del todo aleatorio, sino que existían correlaciones en los registros.

Actualmente, los estudios realizados por Hurst (1951) encuentran aplicación en el análisis fractal propuesto por Mandelbrot (1985), quien estudió la dimensión de objetos que no corresponden a las clásicas geometrías de Euclides. Se interesaba por responder a preguntas del tipo *¿Cuál es la dimensión de la costa de Gran Bretaña?*. Mandelbrot (1985) se dió cuenta que este tipo de objetos tenían dimensión fraccionaria mayor que uno y que el análisis fractal se ajustaba a una ley de potencia, este tipo de ley también se encontró en los estudios de Hurst (1951). Esta reciprocidad de resultados entre los estudios de Hurst (1951) y Mandelbrot (1985), permitió a Mandelbrot (1985) descubrir que en efecto la dimensión fractal está relacionada con el exponente de Hurst. Estos resultados permitieron descartar la idea de modelos estocásticos para el comportamiento del río Nilo.

Actualmente, el análisis de fractalidad tienen dos objetivos principales de cálculo:

- La dimensión fractal
- El exponente de reescalamiento, H

Este último parámetro, (H), también es conocido como *exponente de Hurst* en honor a Hurst (1951). El valor que adquiera el exponente de Hurst permite saber si existen correlaciones en una serie de tiempo. La Tabla 1.1 resume el tipo de correlación que puede presentar una serie de tiempo.

El análisis de fractalidad ha resultado muy útil para encontrar correlaciones en series de tiempo de fenómenos complejos en diversas disciplinas que van desde finanzas (Alvarez-Ramirez y col., 2008b), biológicas (Goldberger y col., 2002) e ingeniería (Vela-Martínez

Tabla 1.1. Tipo de correlación en una serie de tiempo con base en el exponente de Hurst.

H	Tipo de correlación
$H > 0.5$	Correlación persistente
$H = 0.5$	No existe correlación alguna
$H < 0.5$	Correlación antipersistente

y col., 2009) y ha permitido estudiar y entender mejor estos fenómenos, mientras que estadísticos descriptivos (por ejemplo, media aritmética, desviación estándar, etc.) no han tenido resultados satisfactorios.

En el campo de la Ingeniería Química también ha demostrado ser útil el análisis de Hurst, por ejemplo, Fan y col. (1993) realizaron un análisis de fractalidad a las fluctuaciones de presión de un reactor de cama fluidizada. Sus resultados revelaron que las fluctuaciones de presión en la cama exhiben correlaciones de largo alcance.

Actualmente existen diversos métodos para el cálculo del exponente de Hurst en los análisis de fractalidad; sin embargo, los más empleados son:

- El análisis de reescalamiento (R/S), y
- El análisis de fluctuaciones sin tendencia (DFA).

Un inconveniente del análisis de reescalamiento (R/S) es la necesidad de señales estacionarias, algo que comúnmente violan los sistemas complejos reales ya que generalmente presentan tendencias no lineales. Mientras que el análisis DFA es más robusto en este sentido.

1.3. Justificación

Con estos antecedentes es evidente que el uso de los análisis de complejidad mediante los conceptos de entropía y fractalidad no han sido explotados en temas relacionados con la Ingeniería Química. Por lo tanto, en el presente trabajo se emplean éstos conceptos para el

análisis de series de tiempo de sistemas complejos de sistemas relacionados con la Ingeniería Química. Para esto, se tomarán conceptos existentes y se crearán nuevas metodologías de análisis de complejidad con el fin de sobrellevar los problemas actuales (ruido, series no estacionarias, etc.) al analizar series de tiempo de sistemas complejos.

1.4. Objetivos

1.4.1. Objetivo general

Desarrollar e implementar algoritmos estadísticos que puedan ser empleados en metodologías generadas para el análisis de complejidad de series de tiempo de sistemas reales con dinámicas complejas en temas de la Ingeniería Química.

1.4.2. Objetivos específicos

- Desarrollar algoritmos de análisis de complejidad basados en la entropía aproximada (AE) propuesta por Pincus y análisis de fractalidad basados en la metodología del rango reescalado (R/S) propuesto por Hurst.
- Sintonizar los parámetros de los algoritmos desarrollados para ser empleados en análisis de complejidad de series de tiempo de sistemas reales con dinámicas complejas.
- Desarrollar una metodología que remueva la tendencia en las series de tiempo ha estudiar con el propósito de resolver el problema de la no estacionalidad.
- Desarrollar metodologías de análisis de complejidad multiescala.
- Verificar los métodos desarrollados con ejemplos propuestos en la literatura especializada incluyendo el estudio de señales sintéticas.
- Aplicar las metodologías propuestas al estudio de series de tiempo de sistemas reales con dinámicas complejas relacionados con temas de la Ingeniería Química.

1.5. Estructura de la tesis

En el Capítulo 2 se presentan los conceptos básicos relacionados con el trabajo de investigación. Se mencionan las definiciones básicas y los temas de mayor relevancia sobre series de tiempo. En el Capítulo 3 se presentan los métodos desarrollados para el análisis de complejidad de series de tiempo desarrollados en este trabajo de investigación. Posteriormente en el Capítulo 4 se muestran algunos ejemplos de aplicación de los métodos desarrollados. Finalmente en el Capítulo 5 se presentan las conclusiones generales del presente trabajo de investigación y se plantean algunas sugerencias para trabajos futuros.

1.6. Artículos derivados de la tesis

El reporte de los Capítulos 3, 4 y 5 se basan en versiones modificadas de los artículos publicados, los cuales se mencionan a continuación:

- Ortiz-Cruz, A.; Rodriguez, E.; Ibarra-Valdez, C.; Alvarez-Ramirez, J. Efficiency of crude oil markets: Evidences from informational entropy analysis. *Energy Policy*, 41, 365-373 (2012).
- Ortiz-Cruz, A.; Santolalla, C.; Moreno, E.; De los Reyes-Heredia, J.A.; Alvarez-Ramirez, J. Fractal analysis of powder X-ray diffraction patterns. *Physica A: Statistical Mechanics and its Applications*, 391(4), 1642-1651 (2012).
- Alvarez-Ramirez, J.; Echeverria, J.C.; Ortiz-Cruz, A.; Hernandez, E. Temporal and spatial variations of seismicity scaling behavior in Southern México. *Journal of Geodynamics*, 54, 1-12 (2012).

CAPÍTULO 2

MARCO TEÓRICO

En este capítulo se presentan los conceptos básicos para comprender el estudio del análisis de series de tiempo. Donde se presentan definiciones básicas y se muestran algunos ejemplos de aplicación de series de tiempo en distintos campos de estudio.

2.1. Series de tiempo

Una serie de tiempo es una colección de observaciones hechas secuencialmente a través del tiempo. Los ejemplos ocurren en una variedad de campos, desde economía hasta ingeniería, y los métodos de análisis de series de tiempo constituyen una importante área de la estadística.

2.1.1. Ejemplos representativos

A continuación se presentan algunos ejemplos de series de tiempo, y su aplicación en distintas disciplinas del conocimiento.

Series de tiempo económicas y financieras

Muchas series de tiempo son almacenadas y empleadas rutinariamente en economía y finanzas. Algunos ejemplos incluyen precios de las acciones a diferentes escalas, totales de

exportación en meses sucesivos, ingresos medios de meses sucesivos, beneficios de las empresas en los años sucesivos, etc.

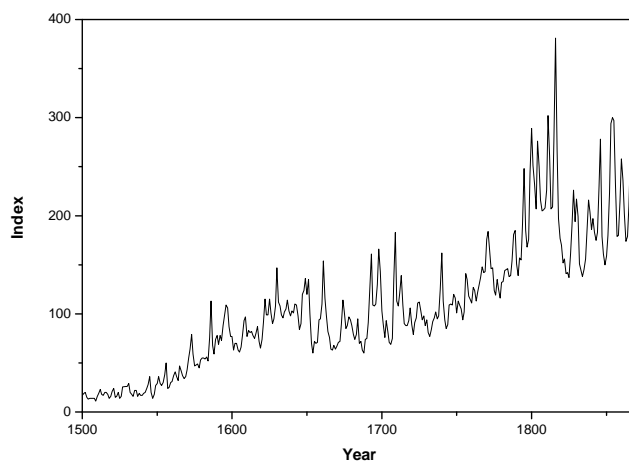


Figura 2.1. Serie de tiempo del índice de precios anual del trigo Beveridge de los años 1500 a 1869.

En la Figura 2.1 puede observarse una serie clásica del índice de precios del trigo Beveridge, la cual consiste de los precios de trigo promedio de cerca de 50 lugares en varios países, medido en años sucesivos desde 1500 hasta 1869. En esta serie se pueden observar algunos comportamientos cíclicos, y es de particular interés para historiadores económicos.

Series de tiempo físicas

Muchos tipos de series de tiempo ocurren en las ciencias físicas, particularmente en meteorología, ciencias marinas y geofísicas. Algunos ejemplos son las precipitaciones de días sucesivos y temperaturas del aire medidas en horas sucesivas, días o meses. La Figura 2.2 muestra las temperaturas mensuales promedio alrededor del mundo para el período de 1940 – 2011 correspondiente a 71 años, donde pueden observarse claramente las fluctuaciones estacionales.

Algunos dispositivos mecánicos toman medidas continuamente y producen un trazado continuo en vez de observaciones a intervalos discretos de tiempo. Por ejemplo, en algunos laboratorios es importante mantener la temperatura y la humedad tan constantes como sea

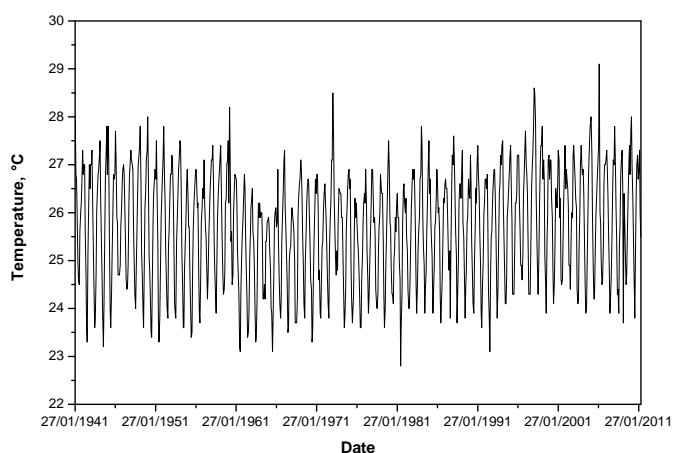


Figura 2.2. Promedio de temperaturas mensuales alrededor del mundo de 1940-2011. Estación meteorológica en Recife, Brazil.

posible y para ello se instalan dispositivos que miden continuamente estas variables. Se toman acciones cuando los valores de las variables se salen de límites pre-especificados. Exámenes visuales del trazado son adecuados para muchos propósitos; mientras que, para un análisis detallado es frecuente convertir el trazado continuo a series de tiempo discretas mediante un muestreo del trazado a intervalos equidistantes apropiados de tiempo. El análisis resultante es más sistemático, puede ser leído y manipulado eficientemente por programas estándares de series de tiempo.

Series de tiempo de mercadeo

El análisis de series de tiempo que surge en el mercadeo es un importante problema en comercio. Las variables observadas pueden incluir gráficas de ventas de semanas o meses sucesivos, ingresos monetarios, costos de publicidad y muchos más. Como un ejemplo, la Figura 2.3 muestra las ventas mensuales de automóviles de una cierta armadora en Quebec, Canada en un período de 8 años. Es importante mencionar que el análisis de este tipo de series ayuda a predecir las ventas futuras, donde se tienen que considerar efectos de tendencia y variación estacional dentro del plan de producción. También es de interés examinar la relación entre las ventas y otras series de tiempo tales como gastos en publicidad.

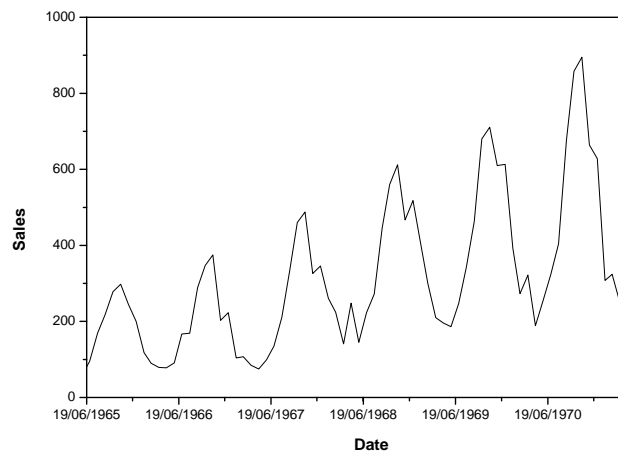


Figura 2.3. Ventas mensuales de automóviles en Quebec, Canada en el período de 1960-1968.

Series de tiempo demográficas

Varias series de tiempo ocurren en el estudio de los cambios de población. Algunos ejemplos incluyen medidas anuales de población de Canadá, nacimientos mensuales en California (Figura 2.4), entre otros. Los demógrafos quieren predecir cambios en la población tan lejanos como 10 ó 20 años en el futuro, se basan en el hecho de que la estructura de cambio es lenta para la población humana. Los métodos estándares de análisis de series de tiempo son usualmente inapropiados para manejar este tipo de problema, ya que ignoran los efectos de interdependencia entre variables, correlación espacial. Además se desprecian las relaciones de una variable medida en distintos instantes de tiempos, correlación temporal. Debido a estas razones estas técnicas son muy poco robustas en la detección de fallas y es aconsejable buscar otros métodos en la literatura especializada..

Datos de control de procesos

En control de procesos, el problema es detectar cambios en la operación de un proceso a través de la medición de una variable, la cual muestra características del proceso como son la calidad, el rendimiento, la controlabilidad, entre otras. Las mediciones pueden ser graficadas en un gráfico de control (Figura 2.5), donde se tienen límites de referencia y cuando

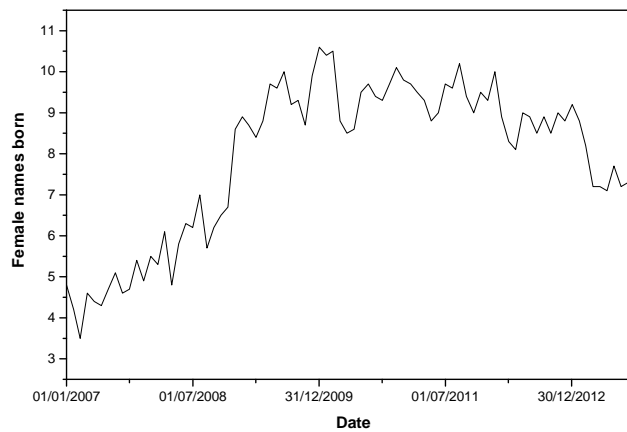


Figura 2.4. Total de nacimientos femeninos diarios en California, 1959.

las mediciones se alejan de un valor de referencia, se deben de tomar acciones correctivas apropiadas para controlar el proceso. Se han desarrollado técnicas especiales para este tipo de problemas en series de tiempo, y pueden encontrarse en el clásico libro sobre control de calidad estadístico de Montgomery (1996).

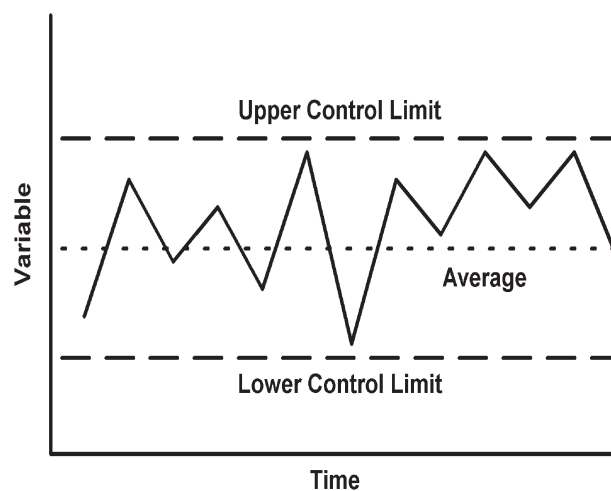


Figura 2.5. Gráfico de control de proceso.

Procesos binarios

Un tipo especial de series de tiempo surge cuando las observaciones pueden tomar uno de sólo dos valores, usualmente denotados por 0 y 1 (Figura 2.6). Por ejemplo, en las ciencias de la computación, la posición de un interruptor, sea “encendido” o “apagado”, puede ser registrada como 1 o 0, respectivamente. Las series de tiempo de este tipo son llamados procesos binarios y ocurren en muchas situaciones, incluyendo el estudio de la teoría de comunicaciones.

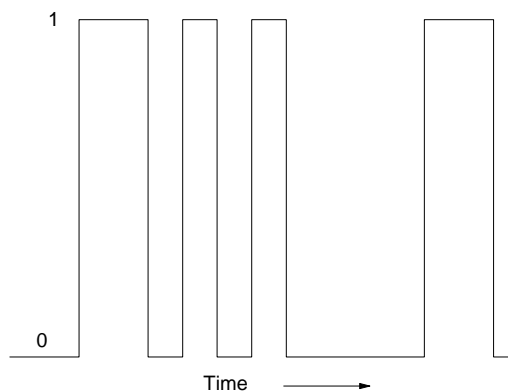


Figura 2.6. Representación de un proceso binario.

Procesos puntuales

Un tipo completamente diferente de series de tiempo ocurre cuando consideramos series de eventos que ocurren “aleatoriamente” a través del tiempo. Por ejemplo, podemos registrar los datos de desastres mayores en autopistas. Una serie de eventos de este tipo es usualmente llamado proceso puntual (Figura 2.7). Las observaciones de este tipo incluyen la distribución del número de eventos que ocurren en un período de tiempo dado y la distribución de los intervalos de tiempo entre los eventos. Los métodos de análisis de datos de procesos puntuales son generalmente muy diferentes a aquellos usados para analizar datos de series de tiempo (Cox y Isham, 1980).

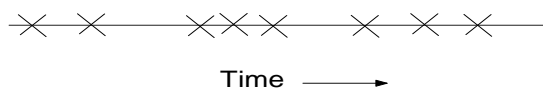


Figura 2.7. Representación de un proceso puntual.

2.1.2. Terminología

Una serie de tiempo se dice que es continua cuando las observaciones son hechas continuamente a través del tiempo, como en la Figura 2.6. El adjetivo “continuo” es usado en series de este tipo incluso cuando la variable medida sólo puede tomar un conjunto discreto de valores. Una serie de tiempo se dice que es discreta cuando las observaciones son tomadas sólo a tiempos específicos, usualmente equidistantes. El término “discreto” es usado para series de este tipo incluso cuando la variable medida es una variable continua.

Las series de tiempo discretas pueden surgir de varias maneras. Dada una serie de tiempo continua, podemos leer (o digitalizar) los valores a intervalos iguales de tiempo para obtener una serie de tiempo discreta, algunas veces llamada serie muestreada. El intervalo de muestreo entre lecturas sucesivas debe ser escogido cuidadosamente de manera que exista poca pérdida de información. Un tipo diferente de series discretas surge cuando la variable no tiene un valor instantáneo pero podemos agregar (o acumular) los valores sobre intervalos iguales de tiempo (Wikipedia). Ejemplos de este tipo son exportaciones mensuales y precipitaciones diarias. Finalmente, algunas series de tiempo son inherentemente discretas, un ejemplo es el dividendo pagado por una sociedad a los accionistas en años sucesivos.

Mucha de la teoría estadística es definida para muestras aleatorias de observaciones independientes. La característica especial del análisis de series de tiempo es el hecho de que las observaciones sucesivas son usualmente no independientes y el análisis deberá tener

en cuenta el orden del tiempo de las observaciones. Cuando las observaciones sucesivas son dependientes, los valores futuros pueden ser predichos a partir de observaciones pasadas. Si una serie de tiempo puede ser predicha exactamente, se dice que es determinística. Sin embargo, muchas series de tiempo son estocásticas ya que el futuro es sólo determinado parcialmente por valores pasados, así que predicciones exactas son imposibles y debe ser reemplazada por la idea de que valores futuros tienen una distribución de probabilidad, la cual es condicionada por un conocimiento de valores pasados.

2.1.3. Objetivos del análisis

Existen diversos objetivos al analizar una serie de tiempo. Estos objetivos pueden ser clasificados como de descripción, explicación, predicción y control.

Descripción

Cuando se tiene una serie de tiempo, el primer paso en el análisis es usualmente graficar las observaciones contra el tiempo para dar lo que es llamado una gráfica de tiempo, y luego obtener medidas descriptivas simples de las propiedades principales de la serie. El poder de la gráfica de tiempo como una herramienta descriptiva es ilustrado en la Figura 2.3, la cual muestra claramente que existe un efecto estacional regular, con “altas” ventas a principios de cada año y “bajas” ventas a mediados del mismo. La gráfica de tiempo también muestra que las ventas anuales se incrementan, es decir, existe una tendencia positiva. La descripción de tales series se considera a través de la interpretación de tendencias y variaciones estacionales. En contraste, para series más complejas se requerirán de técnicas sofisticadas para proveer un análisis adecuado, por ejemplo la construcción de modelos más complejos como los basados en procesos estocásticos.

Es cierto que existen técnicas de análisis más avanzadas, pero esto no significa que las técnicas descriptivas elementales sean menos importantes. Para analizar una serie de tiempo es necesario graficarla en primera instancia para visualizar su complejidad. Una gráfica no sólo mostrará tendencias y variaciones estacionales, sino que también revelará cualquier

observación “atípica” o “outliers” que no parecen ser consistentes con el resto de los datos. El tratamiento de los outliers es una materia compleja en la cual el sentido común es tan importante como la teoría. Un outlier puede ser una observación perfectamente válida, pero extrema, la cual puede, por ejemplo, indicar que el dato no está normalmente distribuido. Alternativamente, un outlier puede ser una observación anómala que se produce, por ejemplo, cuando el dispositivo de registro funciona mal o cuando un paro afecta severamente las ventas. En el último caso, el outlier necesita ser ajustado de alguna forma antes del análisis de los datos. Otras características que buscar en una gráfica de tiempo incluyen cambios repentinos o graduales en las propiedades de las series. Por ejemplo, se deben identificar cambios escalones en el nivel de las series y/o cambios en el patrón estacional. Si existe alguna fuente de discontinuidad en las series, entonces se necesitarán diferentes modelos para ajustar las diferentes secciones de las series.

Explicación

Cuando las observaciones se toman sobre dos o más variables, es posible usar la variación en una serie de tiempo para explicar la variación en otras series de tiempo. Esto puede conducir a un mayor entendimiento de las series de tiempo dadas.

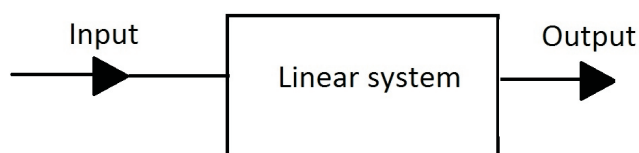


Figura 2.8. Representación esquemática de un proceso lineal.

Aunque los modelos de regresión múltiple ocasionalmente ayudan en estos casos, estos no están diseñados para manejar datos de series de tiempo. Un sistema lineal convierte una serie de entrada en una serie de salida mediante un operador lineal. Dada las observaciones en la entrada y salida para un sistema lineal (Figura 2.8), el análisis evaluará las propiedades del sistema lineal. Por ejemplo, si se desea evaluar el efecto de la temperatura y presión sobre el nivel del mar, condiciones económicas de un país, o bien los efectos de las ventas sobre los precios, entre otras. Una clase de modelos, llamados modelos de funciones de transferencia, nos permiten modelar datos de series de tiempo de una manera apropiada.

Predicción

Dada una serie de tiempo observada, se desea predecir los valores futuros de la serie. Esta es una tarea importante, por ejemplo, en el pronóstico de ventas, y en el análisis de series de tiempo económicas e industriales.

Control

Algunas veces las series de tiempo son recolectadas o analizadas para mejorar el control sobre un sistema físico o económico. Por ejemplo, cuando una serie de tiempo es generada para medir la “calidad ”de un proceso de manufactura, el objetivo del análisis puede ser mantener el proceso operando en un nivel óptimo. Los problemas de control están muy relacionados a predicciones en muchas situaciones, por ejemplo, si se puede predecir que un proceso de manufactura se está alejando del valor de referencia, entonces se deben realizar acciones correctivas apropiadas.

Los procedimientos de control varían considerablemente en estilo y sofisticación. En el control de calidad estadístico, las observaciones son trasladadas sobre gráficas de control y los operadores toman acciones como resultado del estudio de estas gráficas. Un procedimiento más complicado se basa en modelar los datos y usar un modelo para trabajar en una estrategia de control óptima, observe por ejemplo Box y col. (1994). En este método, se ajusta un modelo estocástico a la serie y los valores futuros de la serie son predichos. Posteriormente

las variables de entrada del proceso se ajustan para mantener el proceso en especificación.

2.1.4. Métodos para analizar series de tiempo

Entre los métodos más empleados para el análisis de series de tiempo se encuentran, las técnicas descriptivas simples, la cuales incluyen graficar las observaciones y buscar tendencias, fluctuaciones estacionales y demás. Una herramienta de diagnóstico importante es una función llamada función de autocorrelación, la cual ayuda para describir la evolución de un proceso a través del tiempo. Inferencias basadas en esta función frecuentemente reciben el nombre de análisis en el dominio del tiempo.

Otra función llamada función de densidad espectral, describe como la variación en una serie de tiempo surge por componentes cíclicos a diferentes frecuencias. Inferencias basadas en la función de densidad espectral se llaman análisis en el dominio de la frecuencia.

Análisis avanzados para series de tiempo incluyen los llamados modelos de espacio-estado, filtros de Kalman (el cual es un método general para generar el mejor estimado de la “señal” en una serie de tiempo con presencia de ruido), modelos no lineales y multivariantes.

2.1.5. Técnicas descriptivas simples

Las técnicas estadísticas para analizar series de tiempo varían desde métodos estadísticos directos hasta técnicas inferenciales sofisticadas. Los métodos descriptivos deberían emplearse generalmente antes de intentar procedimientos más complicados, porque ellos pueden ser vitales para “limpiar” los datos, y luego obtener “sentido” de ellos, antes de tratar de generar ideas respecto a un modelo adecuado.

Antes que nada, el analista debe estar seguro de que el problema práctico a analizar se entiende correctamente. En otras palabras, el contexto de un problema dado es crucial en el análisis de series de tiempo, como en todas las áreas de la estadística. Si es necesario, el analista deberá hacerse preguntas para obtener la información teórica necesaria y clarificar sus objetivos. En particular, debemos asegurarnos que los datos apropiados han sido, o serán, recolectados. Si la serie es muy corta, o se ha medido la variable equivocada, no será posible

resolver el problema. En muchas áreas de la estadística, un análisis típico comienza por calcular la media aritmética (o mediana o moda) y la desviación estándar (o varianza) para medir “ubicación” y “dispersión”. Sin embargo, el análisis de series de tiempo es diferente. Si una serie de tiempo contiene tendencia, estacionalidad o algún otro componente sistemático, los estadísticos de resumen usuales pueden estar seriamente desubicados y no serán representativos. Más aún, incluso cuando una serie de tiempo no contiene ningún componente sistemático, los estadísticos de resumen no tienen sus propiedades usuales. Así que, a continuación se mencionan los efectos típicos en series de tiempo, tal como tendencia, estacionalidad y correlación entre observaciones sucesivas para entender mejor el estudio de series de tiempo.

Tipos de variación

Los métodos tradicionales de análisis de series de tiempo se enfocan principalmente en descomponer las variaciones de una serie de tiempo en componentes que representan tendencia, variación estacional y otros cambios cíclicos. Cualquier variación remanente es tratada como una fluctuación “irregular”. Este enfoque no es siempre el mejor pero es particularmente valioso cuando la variación es dominada por la tendencia y la estacionalidad. Sin embargo, debe notarse que una descomposición en tendencia y variación estacional en general sólo se da si se hacen ciertas suposiciones. Por lo tanto algún tipo de modelado, ya sea explícito o implícito, podría estar involucrado en la realización de estas técnicas descriptivas, y esto demuestra la frontera borrosa que siempre existe entre las técnicas descriptivas e inferenciales en estadística. Las diferentes fuentes de variación se presentan a continuación y se describen con mayor detalle.

Variación estacional

Muchas series de tiempo, como gráficas de ventas y lecturas de temperaturas, exhiben variaciones periódicas anuales. Por ejemplo, el desempleo es típicamente “alto” en invierno pero “bajo” en verano. Esta variación anual es fácil de entender, y puede estimarse

directamente si la componente estacional es de interés directo. Alternativamente, la variación estacional puede ser removida de los datos, para dar datos con remoción de la variación estacional, si la componente estacional no es de interés directo.

Otras variaciones cíclicas

Aparte de los efectos estacionales, algunas series de tiempo exhiben variaciones en períodos fijos debido a otras causas físicas, un ejemplo es la variación diaria de la temperatura. Además algunas series de tiempo exhiben oscilaciones, las cuales no tienen un período fijo pero son predecibles de alguna manera. Por ejemplo, los datos económicos algunas veces son afectados por ciclos de negocios con períodos que varían de 3 a 4 años o hasta más de 10 años, dependiendo de la variable medida. Sin embargo, la existencia de tales ciclos de negocios es la materia de algunas controversias, y existe evidencia de que ninguno de tales ciclos es no simétrico. Además, en el caso de las series de tiempo económicas se pueden comportar diferente cuando entra o sale de una recesión.

Tendencia

Esta puede ser vagamente definida como “cambios a largo plazo en el nivel medio”. Una dificultad con esta definición es decidir que se entiende por “a largo plazo”. Por ejemplo, las variables climáticas algunas veces exhiben variaciones cíclicas sobre períodos de tiempo muy largos como 50 años. Si uno sólo tiene 20 años de datos, esta oscilación a largo plazo puede ser vista como una tendencia, pero si estuvieran disponibles cientos de años de datos, entonces la variación cíclica de largo plazo sería visible. No obstante en el corto plazo es más conveniente pensar en una oscilación de largo plazo como una tendencia. Así al hablar de una “tendencia”, debemos tomar en cuenta el número de observaciones disponibles y hacer una evaluación subjetiva de cuál es el significado para la frase “a largo plazo”. Al igual que para la estacionalidad, existen métodos disponibles tanto para estimar la tendencia, o para removerla para que el analista pueda enfocarse con mayor detalle en otras fuentes de variación.

Otras fluctuaciones irregulares

Una vez que las variaciones cíclicas y de tendencia han sido removidas del conjunto de datos, se tiene una serie de residuos que pueden o no ser “aleatorios”.

Series de tiempo estacionarias

Se dice que una serie de tiempo es estacionaria si no hay un cambio sistemático en la media (no la tendencia), si no hay un cambio sistemático en la varianza, y además han sido removidas las variaciones periódicas. En otras palabras, las propiedades de una sección de los datos son iguales que aquellas de cualquier otra sección. Hablando estrictamente, no existe algo como “una serie de tiempo estacionaria”, dado que la propiedad estacionaria esta definida para un modelo. Sin embargo, la frase es usualmente empleada para datos de series de tiempo que exhiben características que un modelo estacionario puede ajustar.

Mucho de la teoría de probabilidad de series de tiempo está relacionada con series de tiempo estacionarias, y es por esta razón que el análisis de series de tiempo frecuentemente requiere que se transforme una serie no estacionaria a una serie estacionaria para poder emplear la teoría. Por ejemplo, será de interés remover las variaciones de tendencia y de estacionalidad de un conjunto de datos y luego tratar de modelar la variación en los residuos por medio de un proceso estocástico estacionario. Sin embargo, también vale la pena destacar que componentes no estacionarios, tal como la tendencia, pueden ser de mayor interés que los residuos estacionarios.

2.2. Gráfica de tiempo

El primer, y más importante, paso en cualquier análisis de series de tiempo es graficar las observaciones contra el tiempo. Este gráfico se llama gráfica del tiempo, y mostrará importantes características de las series tal como tendencia, estacionalidad, outliers y discontinuidades. La gráfica es vital tanto para describir los datos como para ayudar a formular un modelo adecuado. Graficar una serie de tiempo no es tan sencillo. Debido

a que la elección de las escalas, el tamaño de las intersecciones y la manera en que los puntos son presentados (como una línea continua o puntos separados o cruces) puede afectar sustancialmente la forma en que “luce” la gráfica, así que el análisis debe ejercitarse atenta y objetivamente.

Hoy en día, las gráficas son realizadas usualmente por computadoras. Algunas son bien graficadas pero algunos software producen gráficas pobres y el usuario deberá estar preparado para modificarlas si es necesario o, mejor, dar a la computadora instrucciones apropiadas para producir una determinada gráfica.

2.2.1. Transformaciones

Graficar los datos puede sugerir que es adecuado considerar su transformación, por ejemplo, tomando logaritmos o raíces cuadradas. Las tres razones principales para hacer una transformación son las siguientes.

Para estabilizar la varianza

Si hay una tendencia en la serie y la varianza parece incrementar con la media, entonces será recomendable transformar los datos. En particular, si la desviación estándar es directamente proporcional a la media, una transformación logarítmica es indicada. O en otras palabras, si la varianza cambia a través del tiempo sin una tendencia presente, entonces una transformación no ayudara. En lugar de esto, se deberá considerar un modelo que permita un cambio de varianza.

Para hacer los efectos estacionales aditivos

Si hay una tendencia en la serie y el tamaño del efecto estacional parece incrementar con la media, entonces sería recomendable transformar los datos para hacer al efecto estacional constante año con año. El efecto estacional se dice entonces que es aditivo. En particular, si el tamaño del efecto estacional es directamente proporcional a la media, entonces el efecto estacional se dice que es multiplicativo y una transformación logarítmica es apropiada para

hacer el efecto aditivo. Sin embargo, esta transformación solo estabilizará la varianza si el término del error también se piensa que es multiplicativo, un requisito que algunas veces se pasa por alto.

Para hacer los datos normalmente distribuidos

El modelamiento y pronóstico son usualmente llevados a cabo con la presunción de que los datos son normalmente distribuidos. En la práctica esto no es necesariamente el caso; podría, por ejemplo, haber evidencias de sesgos en los que la tendencia serían picos en la gráfica de tiempo que estén todos en la misma dirección (todos arriba o abajo). Este efecto puede ser difícil de eliminar con una transformación y sería necesario modelar los datos usando una distribución del error diferente. Las transformaciones logarítmicas y de raíz cuadrada, antes mencionadas, son casos especiales de una clase general de transformaciones llamadas transformaciones de Box-Cox. Dada una serie de tiempo observada x_t y un parámetro de transformación, λ , la serie transformada esta dada por:

$$y_t = \begin{cases} (x_t^\lambda - 1)/\lambda & \lambda \neq 0 \\ \log x_t & \lambda = 0 \end{cases}$$

Esta es efectivamente sólo una transformación de potencia cuando $\lambda \neq 0$, dado que las constantes son introducidas para hacer y_t una función continua de λ en el valor $\lambda = 0$. El mejor valor de λ puede ser estimado, o estimado alternativamente por un procedimiento inferencial apropiado, tal como el de máxima verosimilitud.

2.3. Analizando series que contienen tendencia

Es muy difícil dar una definición precisa de tendencia y diferentes autores pueden usar el término en diferentes maneras. El tipo más simple de tendencia es la familiar “línea de tendencia + ruido”, para el cual la observación al tiempo t es una variable aleatoria X_t , dada por

$$X_t = \alpha + \beta t + \varepsilon_t \quad (2.1)$$

Donde α, β son constantes y ε_t denota un término de error aleatorio con media cero. El nivel medio al tiempo t esta dado por $m_t = (\alpha + \beta t)$; y es llamado algunas veces “término de la tendencia”. Otros autores prefieren describir la pendiente β como la tendencia, así la tendencia es el cambio en el nivel medio por unidad de tiempo.

La tendencia en la Ecuación 2.1 es una función determinística de tiempo y es llamada algunas veces tendencia lineal global. En la práctica, generalmente proporciona un modelo poco realista, y hoy en día hay más énfasis en modelos que permiten una tendencia lineal local. Una posibilidad es ajustar un modelo lineal por segmentos donde la línea de tendencia es lineal localmente pero con puntos cambiantes donde la pendiente y la intercepción cambian abruptamente. Es común que las líneas se unan en los puntos cambiantes, pero, aún incluso, los cambios repentinos en la pendiente frecuentemente parecen no naturales. Entonces, a veces parece más sensato buscar modelos que permitan una transición suave entre los diferentes submodelos. Extendiendo esta idea, incluso parece aún más natural permitir que los parámetros α y β en la Ecuación 2.1 evolucionen en el tiempo. Esto puede ser un hecho determinístico, pero es más común asumir que α y β evolucionan estocásticamente dando lugar a lo que es llamado una tendencia estocástica. Otra posibilidad, dependiendo de como luzcan los datos, es que la tendencia tenga una forma no lineal, tal como un crecimiento cuadrático. Un crecimiento exponencial puede ser particularmente difícil de manejar, incluso si se toman logaritmos para transformar la tendencia a una forma lineal. Incluso con la ayuda computacional actual, aún es difícil decidir que forma de tendencia es apropiada en un contexto dado (Ball y Wood, 1996; Ching-Chih y Tin-Chia, 2011; Iparraquirre-D’elia, 2011).

El análisis de una serie de tiempo que exhibe tendencia depende de, si se desea (1) medir la tendencia y/o (2) remover la tendencia de forma que se puedan analizar las fluctuaciones locales. También depende de si los datos exhiben la componente estacional. Con datos estacionales, es una buena idea empezar por calcular los promedios anuales sucesivos, y

esto dará una descripción simple de la tendencia. Un enfoque de este tipo usualmente es adecuado, particularmente si la tendencia es bastante pequeña, pero algunas veces se desea un enfoque más sofisticado.

2.3.1. Ajuste de curva

Un método tradicional de manejar datos no estacionales que contienen tendencia, particularmente datos anuales, es ajustar una función simple de tiempo tal como una curva polinomial (lineal, cuadrática, etc.), una curva de Gompertz o una curva logística (Meade, 1984; Franses, 1998). La tendencia lineal global en la Ecuación 2.1 es el tipo más simple de curva polinomial. La curva de Gompertz puede ser escrita en la forma

$$\log x_t = a + br^t$$

Donde a , b , r son parámetros con $0 < r < 1$, o en forma alternativa como

$$x_t = \alpha \exp[\beta \exp(-\gamma t)]$$

La cual luce muy distinta, pero de hecho es equivalente, donde $\gamma > 0$. La curva logística está dada por

$$x_t = a/(1 + be^{-ct})$$

Ambas curvas tienen forma de S y se aproximan a un valor asintótico cuando $t \rightarrow \infty$, con la curva de Gompertz generalmente convergiendo más lento que la logística. Ajustar la curva con datos debe conducir a ecuaciones simultáneas no lineales. Para todas las curvas de este tipo, la función de ajuste proporciona una medida de la tendencia, y los residuos proporcionan un estimado de las fluctuaciones locales, donde los residuos son la diferencia entre las observaciones y los valores correspondientes a la curva de ajuste.

2.3.2. Filtrado

Un segundo procedimiento para manejar una tendencia es usar un filtro lineal, el cual convierte una serie de tiempo x_t , en otra y_t , mediante el operador lineal

$$y_t = \sum_{r=-q}^{+s} a_r x_{t+r}$$

Donde a_r es un conjunto de pesos. Para suavizar fluctuaciones locales y estimar la media local, deben escogerse adecuadamente los pesos de forma que $\sum a_r = 1$, y entonces la operación se refiere frecuentemente como un promedio móvil. Los promedios móviles son discutidos en detalle por Kendall y col. (1983), y aquí solo se proporciona una breve introducción. Los promedios móviles son simétricos frecuentemente con $s = q$ y $a_j = a_{-j}$. El ejemplo más simple de un filtro de suavizado simétrico es el promedio móvil simple, para el cual $a_r = 1/(2q + 1)$ para $r = -q, \dots, +q$, y el valor suavizado de x_t esta dado por

$$Sm(x_t) = \frac{1}{2q + 1} \sum_{r=-q}^{+q} x_{t+r}$$

El promedio móvil simple no es generalmente recomendado para medir la tendencia, aunque puede ser útil para remover variaciones estacionales. Otro ejemplo simétrico es proporcionado por el caso cuando a_r son términos sucesivos en la expansión de $(1/2 + 1/2)^{2q}$. Así cuando $q = 1$, los pesos son $a_{-1} = a_{+1} = 1/4$, $a_0 = 1/2$. Conforme q aumente, los pesos aproximan a una curva normal.

Un tercer ejemplo es el promedio móvil de 15 puntos de Spencer, el cual es usado para suavizar estadísticas de mortalidad para obtener tablas de vida. Cubre 15 puntos consecutivos con $q = 7$, y los pesos simétricos son

$$\frac{1}{320}[-3, -6, -5, +3, +21, +46, +67, +74, \dots]$$

Un cuarto ejemplo, llamado promedio móvil de Henderson, es descrito por Kenny y Durbin (1982) y es ampliamente usado, en los paquetes estacionales $X - 11$ y $X - 12$. Este promedio móvil ayuda a seguir una tendencia polinomial cúbica sin distorsión, y la elección

de q depende del grado de irregularidad. El promedio móvil de 9 términos simétricos, por ejemplo, esta dado por

$$[-0.041, -0.010, +0.119, +0.267, +0.330, \dots]$$

La idea general es ajustar una curva polinomial, no el total de la serie, pero si un conjunto local de puntos. Por ejemplo, un ajuste polinomial para los primeros $(2q+1)$ puntos de datos puede ser usado para determinar el valor interpolado en la mitad del rango donde $t = (q+1)$, y el procedimiento luego puede ser repetido usando los datos desde $t = 2$ hasta $t = (2q+2)$ y así sucesivamente.

Donde quiera que se escoja un filtro simétrico, existirá algo como un problema de efectos finales (Kendall y col., 1983), dado que $Sm(x_t)$ sólo puede ser calculado para $t = (q+1)$ hasta $t = N - q$. En algunas situaciones esto no será importante, como por ejemplo, en el manejo de análisis retrospectivos. Sin embargo, en otras situaciones, como en pronósticos, es particularmente importante obtener datos suavizados arriba de $t = N$. El analista puede proyectar los valores suavizados al ojo o, alternativamente, puede usar un filtro asimétrico que sólo involucre valores presentes y pasados de x_t . Por ejemplo, la técnica popular conocida como ajuste exponencial asume efectivamente que

$$Sm(x_t) = \sum_{j=0}^{\infty} \alpha(1-\alpha)^j x_{t-j}$$

Donde α es una constante tal que $0 < \alpha < 1$. Aquí observamos que los pesos $a_j = \alpha(1-\alpha)^j$ disminuyen geométricamente con j . Una vez estimada la tendencia, podemos observar las fluctuaciones locales mediante

$Res(x_t)$ = residuos de los valores ajustados

$$Res(x_t) = x_t - Sm(x_t)$$

$$Res(x_t) = \sum_{r=-q}^{+s} b_r x_{t+r}$$

Este también es un filtro lineal con $b_0 = 1 - a_0$, y $b_r = -a_r$ para $r \neq 0$. Si $\sum a_r = 1$,

entonces $\sum b_r = 0$ y el filtro es un removedor de la tendencia. ¿Cómo escogemos el filtro apropiado?, la respuesta a esta pregunta realmente requiere de una experiencia considerable más un conocimiento de los aspectos de frecuencia del análisis de series de tiempo. Como el nombre lo implica, los filtros son usualmente diseñados para producir una salida con énfasis en la variación a frecuencias particulares. Por ejemplo, para conseguir valores suavizados queremos remover las fluctuaciones locales que constituyen lo que es llamado la variación de frecuencia alta. En otras palabras queremos lo que es llamado un filtro pasa bajas. Para obtener $Res(x_t)$, debemos remover las fluctuaciones de largo alcance o las variaciones de frecuencia baja. En otras palabras queremos lo que es llamado un filtro pasa altas. El Slutsky o efecto Slutsky-Yule esta relacionado con este problema. Slutsky mostró que operando sobre una serie completamente aleatoria tanto con procedimientos de promedio y diferenciación uno puede inducir variaciones sinusoidales en los datos. Slutsky sugirió que un comportamiento periódico en algunas series de tiempo económicas aparentemente debe ser contado por los procedimientos de suavizado usados para formar los datos.

2.3.3. Filtros en series

Un procedimiento de suavizado puede ser manejado en una o más etapas. Como un ejemplo, dos filtros en serie pueden ser representados como en la Figura 2.9. Es fácil mostrar que una serie de operaciones lineales es todavía un filtro lineal total. Suponga que el filtro I, con pesos $\{a_r\}$, actúa sobre $\{x_t\}$ para producir $\{y_t\}$. Luego el filtro II con pesos $\{b_j\}$ actúa sobre $\{y_t\}$ para producir $\{z_t\}$. Ahora

$$z_t = \sum_j b_j y_{t+j}$$

$$z_t = \sum_j b_j \sum_r a_r x_{t+j+r}$$

$$z_t = \sum_k c_k x_{t+k}$$

Donde $c_k = \sum_r a_r b_{(k-r)}$ son los pesos para el filtro total, los pesos $\{c_k\}$ son obtenidos por un procedimiento llamado convolución, y entonces podemos escribir

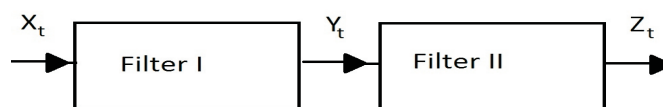


Figura 2.9. Representación esquemática de un proceso lineal.

$$\{c_k\} = \{a_r\} * \{b_j\}$$

Donde el símbolo $*$ representa al operador de convolución. Por ejemplo, el filtro $(\frac{1}{4}, \frac{1}{2}, \frac{1}{4})$ puede ser rescrito como

$$\left(\frac{1}{4}, \frac{1}{2}, \frac{1}{4}\right) = \left(\frac{1}{2}, \frac{1}{2}\right) * \left(\frac{1}{2}, \frac{1}{2}\right)$$

El promedio móvil de 15 puntos de Spencer es de hecho una convolución de cuatro filtros, nombrados

$$\left(\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4}\right) * \left(\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4}\right) * \left(\frac{1}{5}, \frac{1}{5}, \frac{1}{5}, \frac{1}{5}\right) * \left(-\frac{3}{4}, \frac{3}{4}, 1, \frac{3}{4}, -\frac{3}{4}\right)$$

Y este debe ser la mejor manera de calcularlo.

2.3.4. Diferenciación

Un tipo especial de filtro, el cual es particularmente útil para remover la tendencia, es simplemente diferenciar una serie de tiempo dada hasta que llegue a ser estacionaria. Este procedimiento es una parte integral de lo que es llamado procedimiento de Box-Jenkins. Para datos no estacionales, la diferenciación de primer orden es usualmente suficiente para alcanzar estacionalidad aparente. Aquí una nueva serie, por decir $\{y_2, \dots, y_N\}$, se forma a partir de una serie de observaciones originales, por decir $\{x_1, \dots, x_N\}$, mediante $y_t = x_t - x_{t-1} = \nabla x_t$,

para $t = 2, 3, \dots, N$. Ocasionalmente se puede llegar a requerir la diferenciación de segundo orden empleando el operador ∇^2 , donde

$$\nabla^2 x_t = \nabla x_t - \nabla x_{t-1} = x_t - 2x_{t-1} + x_{t-2}$$

La primera diferenciación se usa ampliamente y frecuentemente trabaja bien. Por ejemplo, Franses y Kleibergen (1996) muestran que los mejores pronósticos suelen ser obtenidos con datos económicos por el uso de la primera diferenciación en vez de ajustar una tendencia determinística (o global). Existe un procedimiento llamado diferenciación estacional, la cual remueve la variación estacional.

2.4. Analizando series que contienen variación estacional

En la sección 2.1.5 se comentó que la variación estacional tiene en general un período anual, mientras que en la sección 2.2.1 se distinguió entre la estacionalidad aditiva, la cual es contante año con año, y la estacionalidad multiplicativa. Los tres modelos de estacionalidad comúnmente empleados son,

$$A \quad X_t = m_t + S_t + \varepsilon_t$$

$$B \quad X_t = m_t S_t + \varepsilon_t$$

$$C \quad X_t = m_t S_t \varepsilon_t$$

Donde m_t es la componente que elimina la estacionalidad al tiempo t , S_t es el efecto estacional al tiempo t , y ε_t es el error aleatorio. El modelo A describe el caso aditivo, mientras que los modelos B y C involucran la estacionalidad multiplicativa. En el modelo C el término del error es también multiplicativo, y una transformación logarítmica convertirá a esta en un modelo aditivo (lineal), el cual será más fácil de manejar. Se deberá de examinar la gráfica de tiempo para ver que modelo puede dar la mejor descripción. Los índices estacionales $\{S_t\}$ son usualmente asumidos que cambian lentamente con el tiempo, así que $S_t \cong S_{t-s}$, donde s es el número de observaciones por año. Los índices usualmente son normalizados para que su suma de cero en el caso aditivo, o promediados a uno en el caso multiplicativo. En la

práctica se pueden presentar dificultades si los términos estacional y/o multiplicativo no son exactamente multiplicativos o aditivos.

El análisis de series de tiempo que exhiban variaciones estacionales depende de que se desee (a) medir el efecto estacional y/o (b) eliminar la estacionalidad. Para series que muestren una pequeña tendencia, es usualmente adecuado estimar el efecto estacional para un período particular (por ejemplo en el mes de Enero), encontrando el promedio de las observaciones de cada Enero menos el promedio anual correspondiente en el caso aditivo, o la observación de Enero dividida por el promedio anual en el caso multiplicativo. Para series que contengan una tendencia substancial, se requerirá de un enfoque más sofisticado. Con datos mensuales, la manera más común de eliminar el efecto estacional es calcular

$$Sm(x_t) = \frac{\frac{1}{2}x_{t-6} + x_{t-5} + x_{t-4} + \cdots + x_{t+5} + \frac{1}{2}x_{t+6}}{12}$$

Obsérvese que la suma de los coeficientes es 1. Un promedio móvil simple no puede ser utilizado ya que se extendería 12 meses y no se centraría en un valor entero de t . Un promedio móvil sobre 13 meses no se puede usar, ya que daría demasiado peso a los meses que estén en los dos extremos. Para datos de cuartos de mes, el efecto estacional puede ser eliminado calculando

$$Sm(x_t) = \frac{\frac{1}{2}x_{t-2} + x_{t-1} + x_t + x_{t+1} + \frac{1}{2}x_{t+2}}{4}$$

Para datos de 4 semanas, se puede emplear un promedio móvil simple sobre las 13 observaciones sucesivas. El efecto estacional puede ser estimado al calcular $x_t - Sm(x_t)$ o $x_t/Sm(x_t)$ dependiendo si el efecto estacional es aditivo o multiplicativo. Debe revisarse que los efectos estacionales sean razonablemente estables, y luego se calculan los efectos de promedio mensuales (o cuartiles, etc.). El efecto estacional también puede ser eliminado por diferenciación (Box y Jenkins, 1976).

En este capítulo se exponen las metodologías empleadas en los algoritmos computacionales como son la entropía aproximada, la entropía aproximada con remoción de tendencia, la entropía aproximada multiescala y finalmente el análisis de rango reescalado (R/S). Cuya finalidad es emplear las metodologías anteriormente descritas en el análisis de complejidad de series de tiempo.

3.1. Entropía aproximada

El concepto de entropía es empleado en el campo de la termodinámica, mecánica estadística y teoría de la información. Con respecto al campo de la termodinámica se relaciona desde un punto de vista macroscópico, donde la entropía puede ser relacionada con el número y diversidad de patrones y variaciones que un sistema puede mostrar. Y en el campo de la teoría de la información, la entropía se relaciona con la complejidad o incertidumbre que existe en cualquier señal aleatoria. En el caso de la entropía de Shannon o la entropía de Kolmogorov-Sinai caracterizan la ganancia de información a través de un índice del contenido de desorden e incertidumbre. La relevancia de la entropía consiste en su fácil interpretación física, por ejemplo, la entropía de Kolmogorov-Sinai mide la velocidad exponencial a la cual

la información es obtenida cuando el sistema es dinámicamente perturbado.

Una aplicación directa de los conceptos de entropía requiere la disponibilidad de series de datos infinitos con precisión y resolución infinita. Esto no es posible ya que las mediciones de sistemas reales son muestreados con una resolución limitada ε y con velocidad finita de muestreo $1/T_s$. Para resolver este problema, se introdujo el estadístico entropía aproximada (AE , por su acrónimo en inglés) para cuantificar la irregularidad y complejidad en series de tiempo reales (Pincus, 1991). La cuantificación de la entropía aproximada se basa en la probabilidad de que patrones en las series de tiempo que son similares permanecen similares en el siguiente incremento de comparación. Por lo tanto, las series de tiempo con altos valores de AE están asociadas a altas irregularidades (mayor complejidad) o mayores fluctuaciones, y de manera contraria series de tiempo con bajos valores de AE se asocian a menores irregularidades (menor complejidad) o menores fluctuaciones.

El algoritmo de AE puede ser descrito como sigue. Consideremos una serie de tiempo de longitud N muestreada a intervalos de tiempo T_s , $\{X_i\} = \{x_1, x_2, \dots, x_N\}$, donde $x_i = x(t_0 + iT_s)$. Observemos que la longitud N puede ser relacionada a una escala de tiempo $\tau = NT_s$. Dos secuencias de vectores de m -dimensión $u^{(m)}(i) = \{x_i, x_{i+1}, \dots, x_{i+m-1}\}$ y $v^{(m)}(j) = \{x_j, x_{j+1}, \dots, x_{j+m-1}\}$, $i \neq j$, $1 \leq i, j \leq N - m + 1$. Estos vectores $u^{(m)}(i)$ y $v^{(m)}(j)$ son llamados similares si su distancia $d_{u,v}(i, j) = \max\{ju(i+k) - v(j+k) : 0 \leq k \leq m-1\}$, es menor a una tolerancia ε establecida. Para cada uno de los $N - m + 1$ vectores $u^{(m)}(i)$, el número de vectores similares $v^{(m)}(j)$ está dado midiendo sus respectivas distancias. Si $n_i^{(m)}$ es el número de vectores $v^{(m)}(j)$ similares a $u^{(m)}(i)$, la frecuencia relativa de encontrar un vector $v^{(m)}(j)$ similar a $u^{(m)}(i)$ con un nivel de tolerancia ε está dado por:

$$C_i(m, \varepsilon, \tau) = \frac{n_i^{(m)}}{N - m + 1} \quad (3.1)$$

Donde $N - m + 1$ es el número de vectores $v^{(m)}(j) \neq u^{(m)}(i)$ que son potencialmente similares a $u^{(m)}(i)$. Bajo la hipótesis de estacionalidad, lo siguiente es buscar la frecuencia relativa del logaritmo de $C_i(m, \varepsilon, \tau)$, por ejemplo

$$\Phi(m, \varepsilon, \tau) = \frac{1}{N - m + 1} \sum_{i=1}^{N-m+1} \ln C_i(m, \varepsilon, \tau) \quad (3.2)$$

Para N finita, la entropía aproximada es estimada por el estadístico

$$AE(m, \varepsilon, \tau) = \frac{1}{T_s} [\Phi(m, \varepsilon, \tau) - \Phi(m + 1, \varepsilon, \tau)] \quad (3.3)$$

De esta manera, valores bajos de $AE(m, \varepsilon, \tau)$ reflejan series de tiempo más regulares, mientras que valores altos están asociados a una menor predecibilidad (mayor complejidad) de las series de tiempo en la escala de tiempo (τ). Los efectos de tolerancia (ε), longitud de datos (N) y dimensión (m) en el desempeño computacional de la AE fueron explorados por Pincus (1991), encontrando estimadores estadísticos estables para $N > 500$. Por otra parte, los parámetros $m = 2$ y $\varepsilon = 0.15\sigma$, donde σ es la desviación estándar de la serie de tiempo, se utilizan comúnmente en aplicaciones del método. Estos valores de los parámetros son ahora estándares en muchas aplicaciones de los algoritmos para el cálculo de la entropía aproximada, los cuales fueron empleados en este trabajo de investigación.

3.2. Entropía aproximada con remoción de tendencia

El análisis de entropía aproximada es un estadístico introducido por Pincus (1991) para analizar la complejidad de datos experimentales con pocos datos y con ruido. El algoritmo de entropía aproximada (AE) se describe en la sección 3.1.

La hipótesis de estacionalidad es una definición que debe aplicarse en los cálculos computacionales de la AE . Sin embargo, los datos reales complejos no exhiben un estado estacionario; y con frecuencia pueden llevar a interpretaciones erróneas en la estimación de la AE . En muchos casos, las no estacionalidades son inducidas por fuerzas exógenas que actúan en la dinámica del sistema, provocando largas tendencias en el comportamiento de las series de tiempo. En este trabajo hemos introducido un análisis de entropía aproximada sin tendencia DAE para remover los efectos de tendencia en series de tiempo reales complejas. Para conseguir lo anterior, la serie de tiempo original, $\{X_i\} = \{x_1, x_2, \dots, x_N\}$ de longitud

N , se dividió en $M = [N/L]$ cajas que no se sobreponen de muestras de longitud L . Si Δt es el tiempo de muestreo, $\tau = L\Delta t$ representa la escala de tiempo. En cada ventana de tiempo, los datos se ajustan con un ajuste lineal de mínimos cuadrados, la cual representa la tendencia en esa ventana. La coordenada vertical del segmento de línea es $x_i^{(\tau)}$. Al remover la tendencia en la serie de tiempo original se obtiene una serie de tiempo sin tendencia $\{Z_i^{(\tau)}\}$ correspondiente a una escala de tiempo τ :

$$z_i^{(\tau)} = x_i - x_i^{(\tau)}, \quad i = 1, 2, \dots$$

Los siguientes pasos corresponden a la aplicación del algoritmo para calcular la AE mediante las ecuaciones (3.1-3.3) para las series de tiempo sin tendencia $\{Z_i^{(\tau)}\}$. El procedimiento se repite para todas las escalas de tiempo (tamaños de caja) para caracterizar las variaciones de la AE con la escala de tiempo. Es decir, el proceso que elimina la tendencia en conjunto con el procedimiento estándar de calcular computacionalmente la AE conduce naturalmente a la caracterización multiescala de la complejidad de la serie de tiempo cuantificada en términos de la funcionalidad de la entropía multiescala $AE(\tau)$.

Al aplicar el procedimiento del DAE descrito anteriormente e interpretar los resultados, los siguientes comentarios deben tomarse en consideración:

El paso de eliminación de la tendencia actúa como un filtro pasa altas al remover componentes de frecuencia menores que τ^{-1} . De esta manera, $DAE(\tau)$ debe reflejar la regularidad de la serie de tiempo para frecuencias no menores que τ^{-1} .

Debe observarse que la tolerancia está dada en términos de la desviación estándar de la serie de tiempo sin tendencia en el sentido que $r(\tau) = \rho\sigma(\tau)$, para un valor fijo de $\rho(0, 1)$. Sin embargo, $\sigma(\tau)$ no depende de la varianza de la serie de tiempo original, por ejemplo, el valor absoluto de los elementos de los datos. En cambio si se usa una tolerancia fija, se estaría midiendo también la desviación estándar y no sólo la entropía.

Se sugiere limitar los cálculos de la DAE para cajas de tamaño $L \geq 10$. Cálculos computacionales de la AE inadecuados se pueden obtener con cajas de tamaño menores debido a efectos del muestreo de tamaño finito.

El análisis de entropía multiescala ha sido considerado por Costa y col. (2002), pero desde otro punto de vista. Ellos aplican el algoritmo estándar de la *AE* dado por las ecuaciones (3.1-3.3) a series de tiempo de grano grueso, $\{W_i^{(\tau)}\}$, de acuerdo a la ecuación

$$W_i^{(\tau)} = \frac{1}{L} \sum_{i=(j-1)L+1}^{jL} x_i, \quad 1 \leq j \leq N/L \quad (3.4)$$

Para $L = 1$, la serie de tiempo $\{W_i^{(\tau)}\}$ es simplemente la serie de tiempo original. Mientras que la propuesta en este trabajo involucra el empleo de un filtro pasa altas, el análisis de entropía multiescala de Costa y col. (2002) involucra el empleo de un filtro pasa bajas al promediar la ecuación (3.4). En este sentido, el método de eliminación de la tendencia en este trabajo debe proveer información acerca de la dependencia de la entropía con la escala para componentes de frecuencias altas. En contraste, la propuesta de Costa y col. (2002) se enfoca en los patrones de entropía multiescala para componentes de frecuencia baja. Por ejemplo, con el método de multiescala basado en la ecuación (3.3) la entropía del ruido blanco disminuye con la escala de tiempo, mientras que con el enfoque de eliminación de la tendencia predice un valor de entropía constante para todas las escalas de tiempo. En este sentido, ambos enfoques de multiescala son complementarios y deben proveer diferente información del cambio de entropía con la escala de tiempo.

3.3. Entropía aproximada multiescala

En general, la complejidad de sistemas reales no está restringida a una sola escala. De esta forma, se espera que la entropía sea dependiente de la escala, lo que significa que una señal puede ser más incierta para determinadas escalas de tiempo pero más regular en otras. Por lo cual, una caracterización completa de entropía para series de tiempo debe considerar la variabilidad a través de rangos de escalas no triviales. La idea fundamental de un enfoque de entropía multiescala deberá consistir en proporcionar un índice que refleje la velocidad media de creación de información a una escala dada. Como resultado, el grado total de predecibilidad e incertidumbre de una señal es evaluado considerando los valores de entropía

estimados para un rango predefinido de escalas de tiempo. El método de entropía aproximada multiescala involucra dos pasos:

1. Un método que permite visualizar la representación de un sistema dinámico a diferentes escalas de tiempo. Dada una serie de tiempo $X = \{x_1, x_2, \dots, x_N\}$, y un procedimiento de filtro pasa-bajas $LP(f)$ con una frecuencia de corte f , se obtienen series de tiempo filtradas como $Y_f = LP(f) \cdot X$, donde $Y_f = \{y_{f,1}, y_{f,2}, \dots, y_{f,N}\}$. De esta forma, las nuevas series de tiempo Y_f conservan la complejidad de la señal X para frecuencias menores que f , o escalas de tiempo mayores que $\tau = \frac{1}{f}$. Diferentes esquemas de filtros pasa-bajas $LP(f)$ están disponibles en paquetes comerciales. En este trabajo, al igual que en los análisis técnicos de mercados para obtener tendencias a largo plazo de señales financieras, el filtro pasa-bajas se realiza con filtros de promedio móvil. De esta forma, la señal filtrada es obtenida como sigue

$$Y_{f,i} = \frac{1}{n} \sum_{j=1}^n x_{i+j-1}$$

La escala de tiempo está dada por $\tau = \Delta tn$ y $f = \frac{1}{\tau}$. Donde, Δt es el período de muestreo.

2. La cuantificación del grado de irregularidad de cada una de las series de tiempo Y_f es realizado usando el procedimiento de entropía aproximada descrito en la sección 3.1.

3.4. Entropía asimetría temporal

A lo largo del tiempo, la entropía puede cambiar debido a perturbaciones, bifurcaciones, etc. Para algunos períodos, el sistema entrópico puede incrementar o disminuir por cambios en la configuración del sistema. Por ejemplo, para sistemas cerrados que convergen hacia un atractor estacionario, la entropía debería mostrar una disminución a medida que el número de posibles configuraciones se reducen asintóticamente. El incremento o disminución en la entropía debería ser detectado como una asimetría temporal, donde la entropía hacia atrás es

diferente de la entropía hacia al frente. El algoritmo de Pincus (1991) puede adaptarse para estimar la asimetría temporal al introducir la direccionalidad del tiempo en las comparaciones de los vectores $y^{(m)}(i)$ y $y^{(m)}(j)$. Para lograr esto, se fija el vector $y^{(m)}(i)$ y se realizan las comparaciones con sus vectores relativos hacia atrás y hacia al frente $y^{(m)}(j)$. Haciendo $n_{B,i}^{(m)}$ y $n_{F,i}^{(m)}$ el número de vectores $y^{(m)}(j)$ hacia atrás (ejem., $j < i+m$) y hacia al frente (ejem., $j > i$) similares a $y^{(m)}(i)$, respectivamente. Observe que $n_{B,i}^{(m)} + n_{F,i}^{(m)} = n_i^{(m)}$. Las frecuencias relativas de encontrar vectores $y^{(m)}(j)$ hacia atrás y hacia el frente similar al vector $y^{(m)}(i)$ son

$$C_{B,i}(m, \varepsilon, \tau) = \frac{n_{B,i}^{(m)}}{i-1}$$

$$C_{F,i}(m, \varepsilon, \tau) = \frac{n_{F,i}^{(m)}}{N-i-m+1}$$

Donde $i-1$ y $N-i-m+1$ son el número de vectores $y^{(m)}(j)$ que están hacia atrás y al frente de $y^{(m)}(i)$, respectivamente. Al igual que en la ecuación (3.1), se busca la frecuencia relativa del logaritmo de $C_{B,i}(m, \varepsilon, \tau)$ y $C_{F,i}(m, \varepsilon, \tau)$ como sigue:

$$\Phi_B(m, \varepsilon, \tau) = \frac{1}{N-m+1} \sum_{i=1}^{N-m+1} \ln C_{B,i}(m, \varepsilon, \tau)$$

$$\Phi_F(m, \varepsilon, \tau) = \frac{1}{N-m+1} \sum_{i=1}^{N-m+1} \ln C_{F,i}(m, \varepsilon, \tau)$$

Para una longitud finita N , la entropía aproximada hacia atrás y al frente es

$$AE_B(m, \varepsilon, \tau) = \frac{1}{T_s} [\Phi_B(m, \varepsilon, \tau) - \Phi_B(m, \varepsilon, \tau)]$$

$$AE_F(m, \varepsilon, \tau) = \frac{1}{T_s} [\Phi_F(m, \varepsilon, \tau) - \Phi_F(m, \varepsilon, \tau)]$$

La asimétrica temporal se encuentra cuando $AE_B(m, \varepsilon, \tau) \neq AE_F(m, \varepsilon, \tau)$. La diferencia entre AE_B y AE_F caracteriza la asimetría temporal de la trayectoria del sistema x_i . Fuera del equilibrio, sus diferencias permanecen finitas. En el equilibrio, la simetría temporal $AE_B(m, \varepsilon, \tau) = AE_F(m, \varepsilon, \tau)$ se restablece, así que desaparecen sus diferencias con la

producción de entropía. De esta manera, uno puede introducir el índice de asimetría temporal como

$$\alpha_{AE}(m, \varepsilon, \tau) = AE_F(m, \varepsilon, \tau) - AE_B(m, \varepsilon, \tau)$$

En la práctica, debido a limitantes en la precisión numérica, es muy difícil encontrar que $\alpha_{AE}(m, \varepsilon, \tau) = 0$. En este sentido, uno puede considerar la existencia de la asimetría temporal cuando $|\alpha_{AE}(m, \varepsilon, \tau)| < \rho$ para una tolerancia especificada ρ .

3.5. Análisis de rango reescalado

El análisis de rango reescalado (R/S) es un método empleado para caracterizar las dependencias internas de secuencias. El análisis de rango reescalado (R/S) fue desarrollado por Hurst (1951), un hidrólogo inglés quien trabajó en el problema del control de la represa del proyecto del río Nilo en 1907. Él había estudiado los registros de 8 siglos y medio que los egipcios habían mantenido de los desbordamientos del río Nilo, los cuales no parecían ser ruido blanco. Hurst desarrolló el análisis de reescalamiento (R/S) para distinguir entre series de tiempo totalmente aleatorias y las totalmente correlacionadas. La idea principal detrás del análisis R/S es observar el comportamiento del reescalamiento de las derivaciones acumulativas reescaladas de la media. Para sistemas independientes, las distancias cubiertas incrementan, en promedio, a la raíz cuadrada del número de observaciones. Si el sistema cubre una distancia larga, esto es el sistema no puede ser independiente por definición, y los cambios deben ser influenciados unos con otros, así que estos deben ser correlacionados. El método R/S para análisis fractales de secuencias ha encontrado aplicaciones en diversas campos de las ciencias e ingenierías, entre ellas finanzas (Lo, 1991), geofísica (Chamoli y col., 2007), análisis de fracturas (Varotsos y col., 2002), fisiología (Martini y col., 2004), ingeniería mecánica (Vela-Martínez y col., 2009), ingeniería minera (Zhongying y col., 2011), ecología (Wang y col., 2011) por mencionar algunos ejemplos.

Para una secuencia de longitud N dada por $X = \{x_1, \dots, x_N\}$, donde $x_i = x(t_i)$ y

$\Delta t = t_i - t_{i-1}$ es el período de muestreo, el análisis R/S puede ser llevado a cabo mediante los siguientes pasos:

Paso 1. Dividir la serie en D subseries de longitud M ; llamados, $I_{M,d} = \{r_{d,1}, \dots, r_{d,M}\}$, $d = 1, \dots, D$, con $\bigcup_{d=1}^D I_{M,d} = X$. Haciendo $\tau = M\Delta t$ que sea la escala correspondiente de las subseries de longitud M . Calcular la media $E_{M,d}$ para cada subseries $I_{M,d}$ como $E_{M,d} = \frac{1}{M} \sum_{k=1}^M r_{d,k}$. Calcular las subseries $Y_{M,d} = \{y_{d,1}, \dots, y_{d,M}\}$ de las salidas acumuladas a partir de la media como sigue: $y_{d,j} = \sum_{k=1}^j (r_{k,d} - E_{M,d})$

Paso 2. El rango de las secuencias de subseries que $Y_{M,d}$ es calculado como

$$R_{M,d} = \max \{y_{d,1}, \dots, y_{d,M}\} - \min \{y_{d,1}, \dots, y_{d,M}\}, \quad d = 1, \dots, D$$

Este rango es reescalado por la correspondiente desviación estándar, $S_{M,d} = \sqrt{\frac{1}{M} \sum_{k=1}^M (r_{k,d} - E_{M,d})^2}$, así que el rango reescalado es dado por $(R/S)_{M,d} = R_{M,d}/S_{M,d}$.

El rango promedio es dado por $\overline{(R/S)}_M = \frac{1}{D} \sum_{d=1}^D (R/S)_{M,d}$.

Paso 3. Para mostrar explícitamente la escala $\tau = M\Delta t$ asociada a las subseries de longitud M , hacemos uso de la notación $\overline{(R/S)}_\tau$ (esto es, $\overline{(R/S)}_\tau \equiv \overline{(R/S)}_M$). Hurst encontró que el rango reescalado, $\overline{(R/S)}_\tau$, para diversos registros muestra un comportamiento de reescalamiento tipo ley de potencia de la siguiente manera

$$\overline{(R/S)}_\tau \propto \tau^H$$

donde H es el reescalamiento positivo o exponente de Hurst. Por lo tanto, después de realizar el análisis para todas las secciones de N , se puede graficar $\log \left(\overline{(R/S)}_\tau \right)$ contra $\log(\tau)$. Y realizando una regresión mediante mínimos cuadrados, se encuentra que la pendiente de la regresión que corresponde a la estimación del exponente de Hurst H .

La relación entre la dimensión fractal D_f de las secuencias X y el exponente de Hurst H puede ser expresado como $D_f = 2 - H$. De esta manera, el análisis de reescalamiento R/S provee una caracterización de la fractalidad de una secuencia dada. El impacto de la

observación punto a punto en direcciones hacia atrás y hacia adelante se puede expresar en términos de la función de correlación $C(\tau) = 2^{(2H-1)} - 1$. Si la secuencia es independiente (por ejemplo, procesos Gaussianos, t-Student o gamma), la gráfica es aproximadamente una línea recta con pendiente $H = 0.5$ y por lo tanto $C(\tau) = 0$. Si $H > 0.5$ y $C(\tau) > 0$ lo cual significa que la secuencia es persistente con memoria de largo alcance afectando estructuras en la secuencia en todas las escalas. La persistencia implica que si la señal ha estado incrementando o decreciendo, los cambios serán tales que la señal continuará incrementando o decreciendo, respectivamente.

La intensidad del comportamiento con tendencia o persistencia, incrementa conforme H incrementa. Debe notarse que para $H = 1.0$ se tiene que $C(\tau) = 1$, lo cual indica correlaciones positivas perfectas. A la inversa, cuando $H < 0.5$ se tienen antipersistencias con correlaciones negativas (por ejemplo, $C(\tau) < 0$). Esto significa que, cuando la secuencia haya estado incrementando, es más probable que la secuencia decremente para observaciones cercanas. En este sentido, un proceso antipersistente mostrará más oscilaciones bruscas que un proceso como de la caminata aleatoria pura. Es importante notar que los procesos estocásticos persistentes tienen poco ruido mientras que los procesos antipersistentes muestran la presencia de ruido de alta frecuencia. La caminata aleatoria con $H \neq 0.5$ es referida como Movimiento Browniano Fraccional (FBM).

El análisis R/S proporciona una cuantificación, en términos del exponente de Hurst, de las correlaciones contenidas en una secuencia. La desventaja principal del método R/S es que es sensible a la memoria de corto alcance. Por lo tanto, con el fin de corregir dependencias de corto alcance se ha aplicado el análisis R/S a bloques con datos desordenados, por ejemplo, se toma una permutación aleatoria de la serie de datos con bloques de tamaño predefinidos (en general, bloques de tamaño pequeño) y se lleva a cabo el análisis R/S a estos datos desordenados. Esto es justificado debido a la crítica de Lo (2004), la cual menciona que el estadístico R/S es sensible a la presencia de dependencias de corto alcance y los efectos de las permutaciones aleatorias en esos bloques pequeños son necesarios para destruir cualquier estructura particular de autocorrelación con estos bloques. En este trabajo de investigación

se usaron bloques de tamaño 15 para llevar a cabo las secuencias de desorden.

CAPÍTULO 4

CASOS DE ESTUDIO

Este capítulo está dedicado a la aplicación de los métodos desarrollados en el Capítulo 3. Antes de emplear las metodologías del análisis de complejidad directamente a los casos de estudio, primero se estudiaron los resultados al analizar series de tiempo sintéticas del tipo $1/f^\beta$. Lo anterior permitió entender y sintonizar los parámetros para aplicar los métodos al estudio de series de tiempo de sistemas reales. Los casos de estudio seleccionados varían desde la ingeniería biomédica, sismología, economía e ingeniería química. Se emplea la entropía aproximada para el estudio de la complejidad en series de tiempo como: sintéticas del tipo $1/f^\beta$, fluctuaciones de ritmos cardíacos, intensidad de difracción de rayos X , precios de petróleo, así como de datos sísmicos. Por otra parte, se emplea el exponente de Hurst para el análisis de fractalidad en señales de difracción de rayos X .

4.1. Señales sintéticas tipo $1/f^\beta$

Se emplearon señales con ruido tipo $1/f^\beta$ con longitudes de 30,000 datos. Las señales fueron tomadas de la base de datos de PhysioNet. La Figura 4.1 exhibe la entropía aproximada sin tendencia como una función de la escala de tiempo para distintos valores de β . Como era de esperarse, para el caso cuando $\beta = 0$ el análisis de complejidad *DAE*

es invariante con la escala de tiempo. Lo que indica una carencia de cualquier orden o patrones de estructura para cualquier escala de tiempo. Para el caso de $\beta > 0$ el análisis *DAE* es una función que decae monótonicamente, con la mayor rapidez de disminución para el movimiento Browniano ($\beta = 2$). Esto significa que la microestructura (esto es, patrones a escalas menores) es menos regular que la macroestructura (es decir, patrones a escalas mayores).

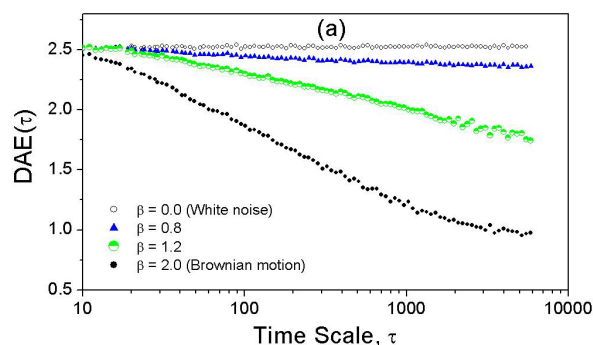
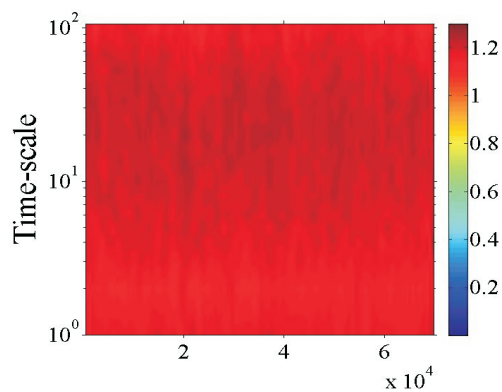
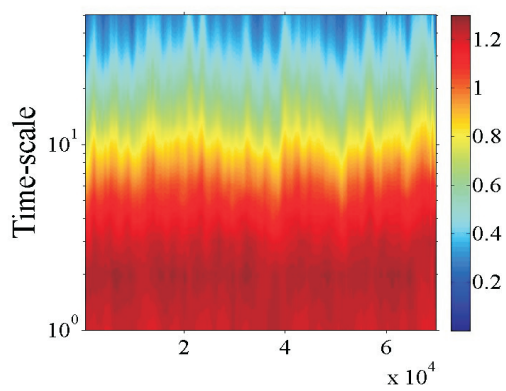
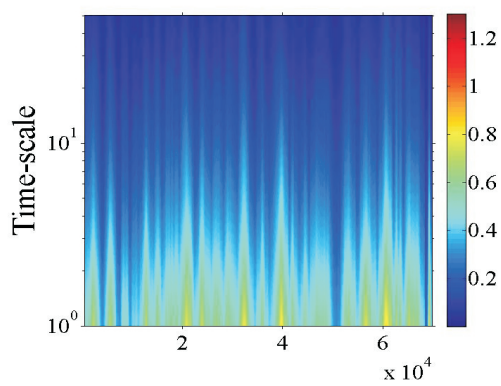


Figura 4.1. Análisis *DAE* para señales sintéticas con ruido tipo $1/f^\beta$.

De igual manera, se analizaron las señales sintéticas con la entropía multiescala, para revelar información de las señales a distintas escalas de tiempo. La Figura 4.2 muestra los resultados del análisis a multiescala. Donde puede observarse que, para el ruido blanco a cualquier escala de tiempo presenta valores altos de entropía relacionados con una falta de orden o patrones estructurados. En el caso de la señal sintética $\beta = 0.8$, a escalas bajas de tiempo presenta una falta de estructura en la señal. Sin embargo, a escalas de tiempo mayores el nivel de entropía disminuye. Lo anterior indica que la señal en la microestructura carece de patrones estructurados, pero en la macroestructura posee una estructuración en la señal. Y finalmente, en el caso de la señal de movimiento browniano se obtienen los mismos resultados que en el análisis *DAE*, donde a cualquier escala existen patrones estructurados.

Para demostrar que el análisis *DAE* disminuye como consecuencia de la estructura intrínseca de la señal, se generó una nueva serie de tiempo a partir de la original pero el orden en los datos fue tomado en forma aleatoria para destruir correlaciones y degradar el contenido de información presente en la serie de tiempo original. La Figura 4.3 muestra el

(a) Señal de ruido blanco ($\beta = 0$)(b) Señal sintética ($\beta = 0.8$)(c) Señal de movimiento browniano ($\beta = 2$)**Figura 4.2.** Entropía aproximada multiescala de señales sintéticas $1/f^\beta$.

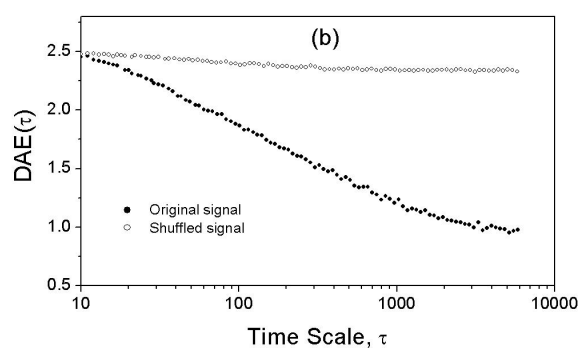


Figura 4.3. Análisis DAE para las señales del movimiento Browniano y su correspondiente señal desordenada aleatoriamente.

análisis DAE como una función de la escala de tiempo para la señal de movimiento Browniano original ($\beta = 2$) y el correspondiente a la señal desordenada aleatoriamente. Encontrándose mayores valores de entropía para la serie de tiempo desordenada aleatoriamente que para la señal de movimiento Browniano original. Comportamientos similares se encontraron para los demás casos de la Figura 4.1. Para escalas de tiempo mayores a 300 aproximadamente, la función $DAE(\tau)$ es constante, parecido al comportamiento del ruido blanco (Figura 4.1). En cambio para escalas de tiempo pequeñas, la función $DAE(\tau)$ presenta una pequeña pendiente negativa.

4.2. Series de tiempo de fluctuaciones de ritmos cardíacos

Se analizaron series de ritmos cardíacos, con el fin de verificar los resultados del análisis DAE contra casos reportados en la literatura especializada, los datos empleados están disponibles en PhysioNet. Las series de tiempo estudiadas son descritas por Mietus y col. (2002) como series de tiempo de intervalos de ritmo cardíaco (RR) derivadas de 24 h de grabación del monitoreo Holter del electrocardiograma (ECG) para sujetos sanos, sujetos con insuficiencia cardíaca, que limita considerablemente la actividad de los sujetos, y sujetos con fibrilación auricular, una arritmia cardíaca seria. Los datos para el grupo de control

normal fueron obtenidos de 24 h de grabación del monitoreo Holter para 72 sujetos sanos, 35 hombres y 37 mujeres, edades 54.6 ± 16.2 (media \pm SD), rango 20-78 años. Los datos del ECG fueron muestreados a 128 Hz. Los datos para el grupo de insuficiencia cardíaca fueron obtenidos en 24 h de grabación Holter de 43 sujetos (28 hombres y 15 mujeres) con edades de 55.5 ± 11.4 (media \pm SD), rango 22-78 años. Los datos del ECG fueron muestreados a 128 Hz. Los datos para el grupo de fibrilación auricular se obtuvieron de 10 h de grabación Holter muestreados a 250 Hz para 9 sujetos. El conjunto de datos fue filtrado para excluir artefactos, contracciones pre ventriculares y detección de latidos faltantes. Las series de tiempo de los intervalos de ritmos cardíacos representativas de los grupos de sujetos sanos, insuficiencia cardíaca y fibrilación auricular se presentan en la Figura 4.4. Obsérvese que las series de tiempo de los sujetos sanos y con insuficiencia cardíaca muestran dinámicas con tendencia, las cuales pueden ser causadas por distintas condiciones operacionales del sistema cardiovascular.

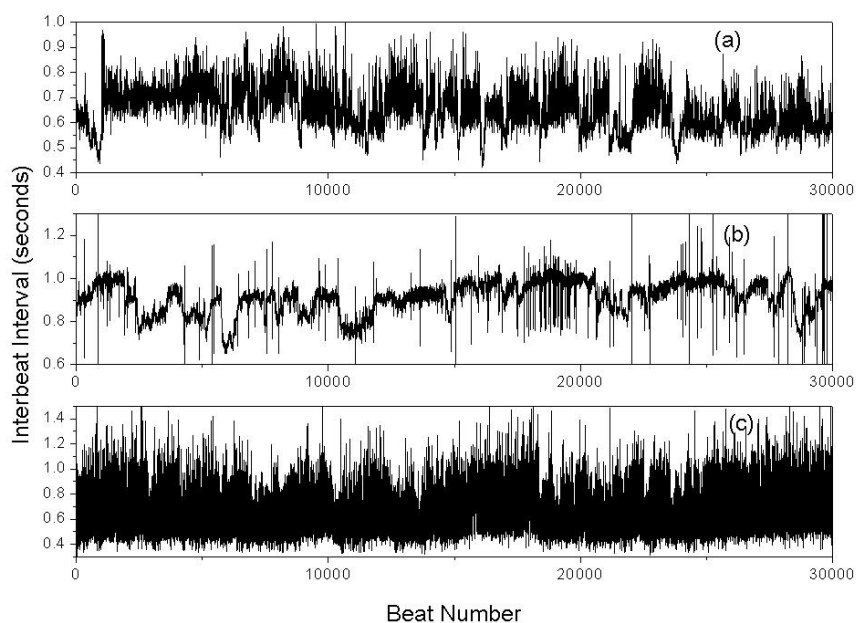


Figura 4.4. Series de tiempo características de los grupos de individuos (a) sanos, (b) con insuficiencia cardíaca y (c) con fibrilación auricular respectivamente.

La Figura 4.5 presenta los resultados típicos del análisis de entropía sin tendencia para

las series de tiempo de intervalos RR para el conjunto de personas sanas, con insuficiencia cardíaca y con fibrilación auricular. Se puede observar un comportamiento muy distinto entre los 3 grupos: (a) La entropía para los casos sanos es una función decreciente de la escala de tiempo. (b) Las entropías de los casos con insuficiencia cardíaca son menores que los casos sanos. En general, esta entropía mide un incremento con la escala de tiempo, aunque puede encontrarse un valor máximo para algunos casos. (c) La entropía para los casos con fibrilación auricular es constante, similar al ruido blanco (Figura 4.1). Estos resultados indican que las dinámicas de insuficiencia cardíaca son menos complejas y mas regulares que las de condiciones sanas. Los casos con fibrilación auricular exhiben altos valores de entropía, sugiriendo que los intervalos de los latidos no contienen estructuras regulares, pero actúan como un proceso puramente aleatorio.

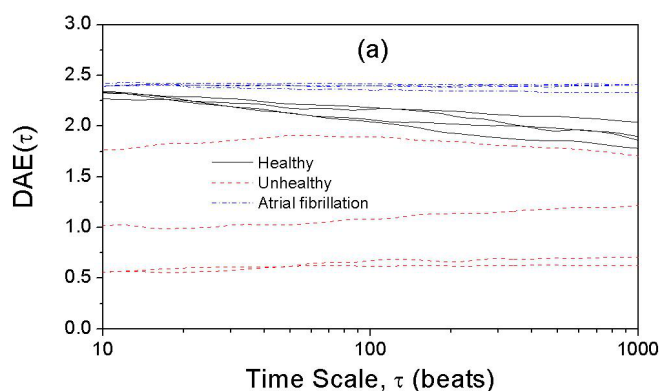


Figura 4.5. Análisis de DAE representativo de series de tiempo de ritmos cardíacos.

La Figura 4.6 presenta el resultado promedio de todos los sujetos de cada caso de estudio. Se observa una clara discriminación entre las distintas condiciones cardíacas y que es posible discriminarlas mediante el empleo del análisis de entropía sin tendencia sobre un amplio rango de escalas de tiempo. De hecho, la condición de insuficiencia cardíaca es caracterizada por menores valores de entropía que los sujetos sanos. En contraste, la condición de fibrilación auricular puede discriminarse al detectar una invariancia en la medición de la entropía con las escalas de tiempo.

Por otra parte, se consideró que el sistema cardiovascular está equipado con un

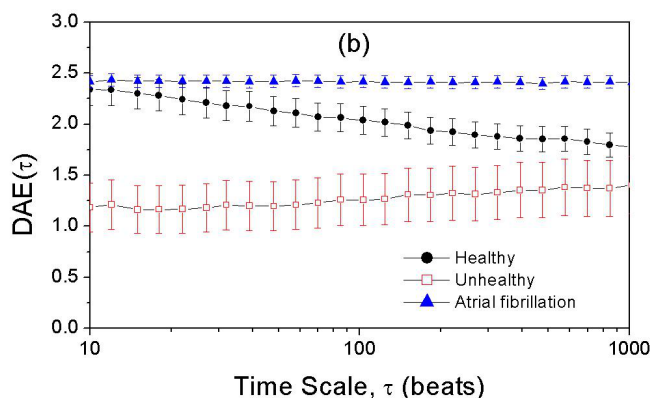


Figura 4.6. Promedio del análisis DAE para cada uno de los tres casos de estudio.

controlador retroalimentado o baroreflejo con el propósito de regular el ritmo cardíaco al enfrentarse a cambios en las condiciones de operación (por ejemplo, cambios en el ritmo cardíaco requerido) y disturbios (por ejemplo, cambios de postura y temperatura ambiental). El control retroalimentado se compone básicamente de dos elementos individuales; llamados, las actividades autonómicas y el baroreflejo (Costa y col., 2002). La actividad simpática incrementa al mismo tiempo que la actividad parasimpática disminuye el ritmo cardíaco. Las acciones de estos dos componentes resultan en un control regulado de la dinámica cardiovascular al enfrentar disturbios complejos.

Para ganar habilidad y conocimiento del papel de la actividad de retroalimentación del sistema cardiovascular, se consideró el modelo simple de regresión-media $x_i = x_r + K(x_{i-1} - x_r) + w_{i-1}$, donde w_{i-1} representa un disturbio tipo ruido blanco, x_r es una referencia de un ritmo cardíaco, $K(x_{i-1} - x_r)$ es la señal de control retroalimentado y K es la ganancia del control retroalimentado. En ausencia de disturbios, la señal x_i converge al valor requerido x_r para $-1 < K < 1$. La Figura 4.7 muestra el análisis DAE para tres valores distintos de la ganancia K . Observe la similitud con el análisis DAE para condiciones sanas, sugiriendo que la reducción de la entropía para escalas de tiempo altas está relacionada a la acción del baroreflejo que regula la dinámica del ritmo cardíaco. El efecto del componente de retardo es reducir la incertidumbre de la señal para escalas de tiempo grandes, lo cual sugiere que el comportamiento del DAE en los sujetos sanos se debe a un buen sintonizado del regulador

del ritmo cardíaco.

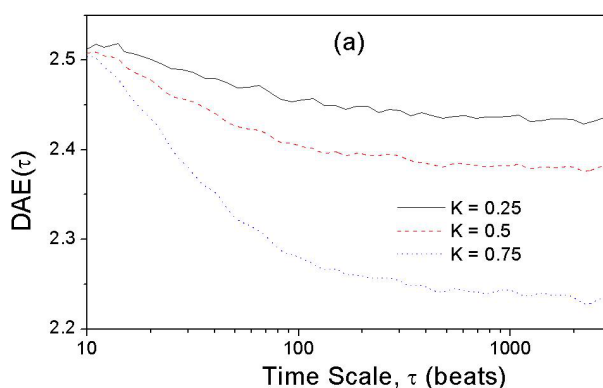


Figura 4.7. DAE para ruido blanco con un componente de retardo Kx_{i-1} , $0 < K < 1$.

En un trabajo reciente Alvarez-Ramirez y col. (2009), mostraron que la condición de insuficiencia cardíaca puede estar asociada con retrasos en el sistema baroreflejo, el cual degrada la eficiencia de la dinámica del sistema cardiovascular. Acciones de retraso en controles retroalimentados pueden limitar seriamente la capacidad del sistema para reducir efectos adversos de disturbios (Morari y Zafiriou, 1989). Para mostrar que el comportamiento del análisis DAE para condiciones no sanas puede relacionarse con retrasos en el control retroalimentado, consideramos el control retroalimentado $K(x_{i-d} - x_r)$, donde $d > 1$ representa el retraso. La Figura 4.8 muestra el análisis DAE para $K = 0.5$ y $d = 6$ latidos. Observe la similitud con algunas de las curvas del análisis DAE para condiciones no sanas de la Figura 4.5. El retraso en la retroalimentación es incapaz de reducir la incertidumbre para escalas de tiempo altas, así que el ritmo cardíaco no está regulado por el valor medio x_r . De esta forma, el sistema cardiovascular es operado en condiciones inestables donde el ritmo cardíaco deseable no puede alcanzarse. Finalmente, es interesante observar que las curvas del análisis DAE para el caso de fibrilación auricular son similares a las del ruido blanco, la cual corresponde al caso con $K = 0$. Esto sugiere que la fibrilación auricular es una condición donde la función del sistema baroreflejo no participa tal como se esperaba por la irregular actividad auricular, indicando que en la condición de insuficiencia cardíaca no se manifiesta el mecanismo regulador del ritmo cardíaco. En resumen, los resultados antes mencionados

indican que el análisis de entropía sin tendencia, DAE , proporciona una importante capacidad de entender como el sistema baroreflejo se encarga de regular las dinámicas cardíacas.

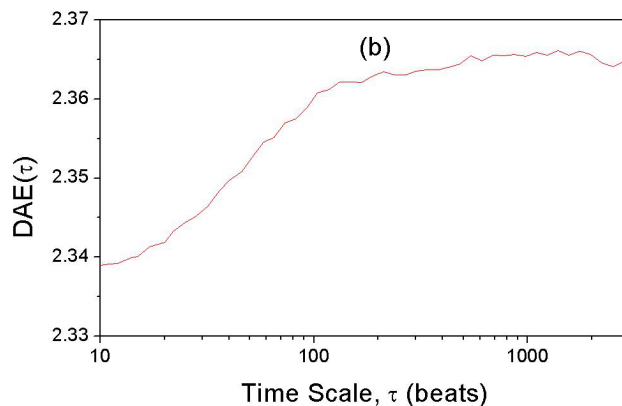


Figura 4.8. DAE para el ruido blanco con un componente de retraso Kx_{i-5} .

4.3. Señales de intensidad de difracción de rayos X

Las señales de intensidad de difracción de rayos X con ruido intenso son comunes en la caracterización de materiales con cierto grado de amorficidad; sin embargo, la identificación de picos asociados a ciertas fases del material no está bien definida. En esta sección se muestran los resultados de los análisis de entropía aproximada multiescala y del rango reescalado para caracterizar los patrones de intensidad de difracción de rayos X con ruido intenso. La idea es que existen patrones de regularidad en las señales de intensidad que pueden estar asociados con las fases del material pero debido al intenso ruido en la misma se encuentran ocultos.

Como casos de estudio se emplearon distintas muestras de alúmina calcinada a diferentes temperaturas debido a la gran información existente sobre la estabilidad de fases y transiciones para ilustrar la aplicabilidad de los métodos. Se demuestra que los análisis tanto de entropía aproximada multiescala como de rango escalado son capaces de detectar regiones angulares en las señales de intensidad asociadas a fases de alúmina correspondientes a fichas de difracción de polvos (PDF).

4.3.1. Introducción

La difracción de rayos X en polvos (XRD) es tal vez la técnica con mayor aceptación y confiabilidad en la caracterización de materiales. En general, la muestra a estudiar debe encontrarse en forma de granos finos del material cristalino siendo el comportamiento de los patrones de XRD consecuencia de la simetría y orientación de los cristales en la muestra de estudio. Por lo tanto, cuando los patrones de difracción $2-D$ son obtenidos estos muestran aros concéntricos de picos dispersos correspondientes a diferentes espacios- d en la estructura cristalina. Donde la posición e intensidad de los picos se emplean para identificar la estructura o fase principal del material.

Una ventaja adicional de la XRD es que ésta implica una preparación no complicada del material y una caracterización no destructiva del mismo, lo cual es conveniente en la caracterización de materiales desconocidos. La caracterización de materiales a partir de datos de XRD se consigue por comparación de los patrones de difracción contra un estándar conocido o datos proporcionados por *The International Center for Diffraction Data*.

Los sólidos cristalinos exhiben patrones de difracción distintivos con fuertes picos de intensidad que son fácilmente observados. La posición que corresponde a la configuración de espacios y la intensidad relativa de las líneas son indicadores de una fase particular del material, los cuales se emplean como patrones de comparación. En contraste a los sólidos cristalinos donde los patrones de XRD exhiben marcados picos, los materiales no cristalinos pueden conducir a señales con ruido intenso. Como ejemplos de materiales con algún grado de amorficidad están varios polímeros y semiconductores, donde la cristalinidad está limitada a una parte de la estructura molecular. Un calcinado incompleto a bajas temperaturas de materiales cerámicos puede producir estructuras amorfas donde las fases cristalinas pueden estar mezcladas con material no cristalino dando configuraciones geométricas complejas con patrones no regulares. Las Figuras 4.9.a y 4.9.b muestran la intensidad de patrones $I(2\theta)$ de rayos X correspondientes a las temperaturas de calcinación de 200 y 500°C para alúmina preparada por el método sol-gel (Wang y col., 1999). De acuerdo a la ruta de transformación de fases más probable para las condiciones de preparación de la alúmina (Gates y col., 1979),

la señal de intensidad correspondiente a 200°C muestra tres picos prominentes que sugieren la presencia de γ^* -bohemita a ~ 10 grados (Wang y col., 1999), grupos de óxidos a ~ 22 grados (Liu y col., 2008) y bayerita a ~ 40 grados (*PDF Card* 20 – 0011). Sin embargo, a pesar de la presencia de estos picos, los patrones de *XRD* presentan ruido intenso para los cuales las regiones angulares de intensidad regular no pueden ser fácilmente distinguidas.

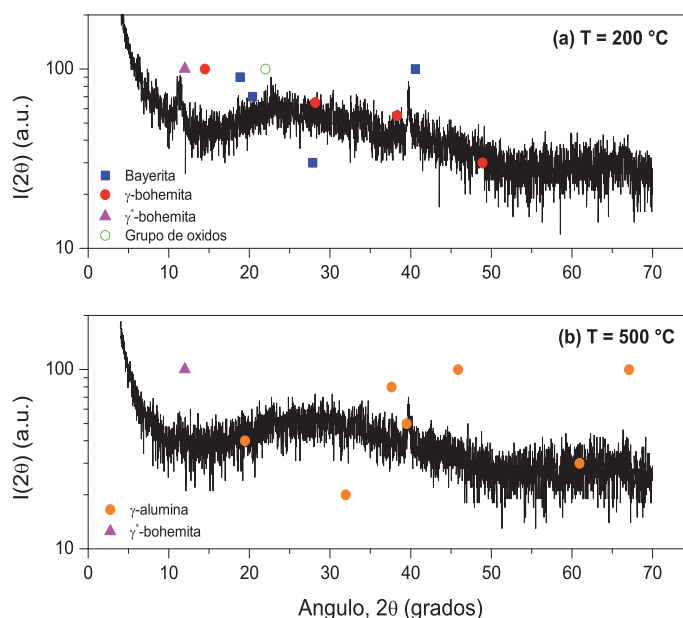


Figura 4.9. Intensidad de patrones $I(2\theta)$ de rayos X para alúmina calcinada a 200 y 500 °C.

Para el caso de la temperatura de calcinación a 500°C se presenta un análisis de mayor reto donde sólo un pico prominente puede observarse. Para estas condiciones de calcinación, los grupos de bayerita y bohemita se convierten en las fases η - y γ -alúmina. De hecho, el pico prominente a 40 grados (50% de intensidad, *PDF Card* 10 – 0425) indica la presencia de γ -alúmina. Sin embargo, otro importante pico de γ -alúmina (por ejemplo, la línea de intensidad de 100% a 45.901 grados) no se muestra en los patrones de *XRD*. En general, los patrones de ruido en estas señales de *XRD* reflejan la presencia de estructuras espaciales desordenadas. Como consecuencia, la interpretación de los patrones de *XRD* llega a ser, en muchos casos, un reto para los ojos profesionales no especializados.

Los avances continuos en hardware y software, en particular mejoras ópticas y detectores rápidos, han mejorado drásticamente la capacidad analítica de los análisis de difracción de

rayos X (por ejemplo, la velocidad de los análisis). Sin embargo, la técnica aún requiere futuras mejoras para refinar e interpretar los patrones de XRD con ruido intenso como aquellos mostrados en la Figura 4.9. Las mejoras en estos temas pueden permitir a los usuarios de XRD una mayor precisión en la caracterización del material y sus fases.

4.3.2. Análisis de complejidad

Se empleó el enfoque del análisis de entropía aproximada para analizar la complejidad de los patrones de XRD con ruido intenso. Esta técnica ha sido empleada en diversas disciplinas desde economía (Pincus, 2008; Martina y col., 2011; Alvarez-Ramirez y Rodriguez, 2011), ciencias biomédicas (Pincus, 2006; Pan y col., 2011; Hassan y col., 2011) hasta ciencia de materiales (Yan y Gao, 2007; Pérez-Canales y col., 2011) con éxito en la caracterización de señales complejas.

Metodología

La motivación de este trabajo de investigación radica en que el ruido contenido en la señal puede proporcionar información de las fases del material. Es decir, las señales no son aleatorias del todo. Existen regularidades, menores fluctuaciones, en la señal que pueden detectarse con el análisis de entropía aproximada sugiriendo la existencia de estructuras más ordenadas, fases cristalinas, en las señales de intensidad. De esta forma, un fuerte pico de intensidad en un ángulo dado se manifiesta porque la radiación dispersa de rayos X se concentra en el detector por geometrías regulares que corresponden a estructuras cristalinas. En contraste, para materiales amorfos no se consigue concentrar la radiación, debido a la ausencia de regularidad.

La idea fundamental es que el patrón de ruido intenso no es tan complejo del todo, sino que contiene patrones internos que pueden asociarse a la estructura del material. Estos patrones pueden ser comparados con fichas de difracción de polvos (PDF) para concluir la presencia o no de probables fases del material. Se seleccionó como caso de estudio a la alúmina calcinada a diferentes temperaturas por la gran cantidad de información con respecto a la

estabilidad de fases y transiciones disponibles en la literatura científica.

La metodología usa el análisis de entropía aproximada multiescala para detectar regiones angulares en los patrones de *XRD* con ruido intenso, $I(2\theta)$, que contienen patrones asociados a la regularidad en la estructura del material. La idea fundamental es que patrones regulares de $I(2\theta)$ reflejan un cierto grado de orden en la geometría del material. En esta forma, un fuerte pico en el patrón de intensidad $I(2\theta)$ que refleja estructuras de un material altamente organizado corresponde a patrones con mayor regularidad asociados a una menor complejidad. Por otra parte, geometrías altamente desorganizadas corresponden a un comportamiento aleatorio en las fluctuaciones de intensidad de la difracción de rayos X .

Basado en estas ideas, la metodología propuesta para analizar los patrones de *XRD* con ruido intenso consiste en transformar la señal de intensidad $I(2\theta)$ en un patrón de entropía $E(2\theta)$ mediante el análisis de entropía aproximada multiescala; es decir,

$$I(2\theta) \rightarrow E(2\theta)$$

De este manera, el análisis para caracterizar la presencia de fases del material se lleva a cabo en los patrones de entropía $E(2\theta)$ en vez de en la señal de intensidad $I(2\theta)$. Así, bajos valores de $E(2\theta)$ están asociados con estructuras más regulares, cristalinas, y posibles fases del material. Mientras que, altos valores de $E(2\theta)$ se asocian con estructuras menos regulares, amorfas.

Resultados

Para cada corrida, los datos experimentales de *XRD* consistieron de 3301 observaciones. El tamaño de la ventana deslizante fue seleccionado como 250 observaciones, fijándose 21 escalas para realizar el análisis.

La Figura 4.10.a muestra la intensidad $I(2\theta)$ de patrones de *XRD* para la temperatura de calcinación de 200°C. Para esta temperatura de calcinación, la alúmina está compuesta probablemente por bayerita y ciertas formas de bohemita, como γ^* -bohemita que tiene la misma estructura cristalina que la γ -alúmina Wang y col. (1999). La figura 4.10.a también

muestra la localización, extraídas de tarjetas *PDF*, para bayerita (*Card* 20–0011) y bohemita (*Card* 21 – 1307), indicando que los picos a ~ 40 grados pueden ser relacionados a bayerita. Siguiendo el estudio realizado por Wang y col. (1999), los picos a ~ 10 grados pueden ser relacionados a la presencia de γ^* –bohemita. Se ha sugerido que los picos a ~ 22 grados pueden estar relacionados a la presencia de grupos de óxidos complejos (Liu y col., 2008). A pesar de que los fuertes picos en los patrones de *XRD* pueden ser explicados por la presencia de algunas fases de alúmina, la localización de otras fases probables, como bayerita, es incierta debido al intenso ruido que afecta la interpretación de los resultados experimentales.

La Figura 4.10.b muestra los correspondientes patrones de entropía $E(2\theta)$ obtenidos por el análisis de entropía aproximada multiescala. Las siguientes características pueden ser observadas:

1. Se planteó la hipótesis que fuertes picos de intensidad deberían corresponder a mayores regularidades, menor complejidad en la estructura del material, en términos del análisis de entropía aproximada multiescala, $E(2\theta)$. Los resultados en la Figura 4.10.b respaldan esta idea mostrando que se detectan las fases correspondientes a tres picos prominentes en la señal original y otras fases que no son evidentes por simple inspección. La marcada intensidad en los picos ~ 10 , 22 y 40 grados corresponden a los menores valores de entropía y por lo tanto a una menor complejidad en la estructura del material que puede asociarse a fases cristalinas.
2. El patrón de entropía $E(2\theta)$ no muestra marcados picos como en los patrones de intensidad $I(2\theta)$ de *XRD*. En lugar, el comportamiento del patrón $E(2\theta)$ obtenido del análisis de entropía multiescala muestra zonas de mínimos y máximos. Esto sugiere que las fases cristalinas en la estructura de la alúmina no se encuentran aisladas sino que coexisten en una distribución geométrica compleja, reflejada en el ruido intenso de los patrones de intensidad $I(2\theta)$ de *XRD*. Por ejemplo, la región angular en la vecindad de 20 grados sugiere un patrón de cristalinidad conformado por bayerita y grupos óxidos. Similarmente, la vecindad de 40 grados pueden ser relacionada a un patrón de cristalinidad complejo conformado por bayerita y bohemita.

3. El análisis de entropía aproximada multiescala puede revelar la existencia de regiones asociadas a regularidades no triviales que no son exhibidas por los patrones originales de *XRD*. De hecho, el análisis de entropía aproximada multiescala reveló la existencia de patrones regulares en la vecindad de 50 grados, con valores de entropía similar a los patrones en la vecindad de 40 grados. En contraste al patrón de intensidad original $I(2\theta)$, el análisis de entropía aproximada multiescala sugiere que la región de 50 grados puede estar relacionado a la fase de bohemita.
4. El menor valor de entropía está localizado en la vecindad de 10 grados, lo cual puede ser atribuido a la presencia de la fase γ^* -bohemita, como ya se ha sugerido por Wang y col. (1999).

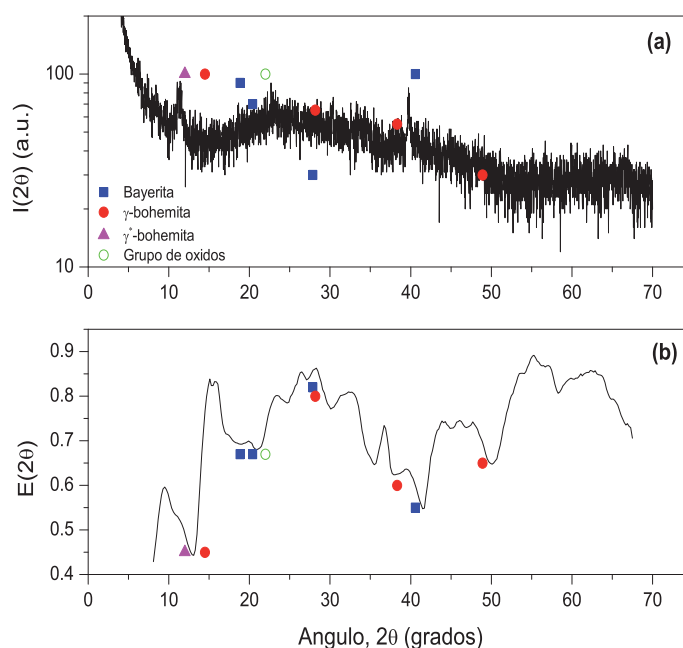


Figura 4.10. (a) Intensidad de patrones $I(2\theta)$ de rayos X , (b) Patrones de entropía $E(2\theta)$ de la señal de intensidad $I(2\theta)$ de alúmina calcinada a 200°C.

Los resultados anteriores muestran la habilidad del análisis de entropía aproximada multiescala para detectar regiones angulares en señales de intensidad con ruido intenso, que pueden ser atribuidas a las fases del material. El caso de la temperatura de calcinación a

500°C es más desafiante ya que los patrones de *XRD* en la Figura 4.11.a muestran un sólo pico fuerte cerca de los 40 grados sugiriendo que la fase γ -alúmina se ha formado. La presencia de γ^* -bohemia cerca de los 10 grados se observa ligeramente como un pico pequeño en tal región angular. Sin embargo, la presencia de otras fases γ -alúmina no puede localizarse directamente de la señal de intensidad original $I(2\theta)$. Por ejemplo, la fase localizada a 20 grados no exhibe ningún pico prominente en la señal original. De esta forma, la presencia de la fase γ -alúmina no puede ser afirmada desde una directa inspección de la intensidad de patrones en la región cerca de 20 grados.

La Figura 4.11.b muestra el comportamiento de los patrones de entropía $E(2\theta)$ obtenidos por el análisis de entropía aproximada multiescala de los patrones *XRD* de la Figura 4.11.a. Donde puede observarse regiones de mínimos, las cuales pueden estar asociadas a la presencia γ -alúmina en el material calcinado, tanto para la región a 40 y 20 grados; y otras regiones angulares reportadas en tarjetas *PDF* de alúmina. También se observa la presencia de γ^* -bohemia en el patrón de entropía $E(2\theta)$ cerca de los 10 grados. Interesantemente, la entropía de la γ^* -bohemia aumentó de 0.44 a 0.75 en comparación con la temperatura de calcinación a 200°C. Esto sugiere, como se esperaba, que fases de bohemia son transformadas a γ -alúmina conforme la temperatura de calcinación se incrementa dentro del rango de 100 – 500°C.

Conforme la temperatura de calcinación se cambió, nuevas fases fueron obtenidas por la transformación a otras fases. De esta forma, la entropía de ciertas fases fue reducida como efecto de la transformación en la estructura de la materia. La Figura 4.12 presenta la evolución de la entropía para regiones en la vecindad de cuatro diferentes ángulos. De acuerdo a los resultados obtenidos, se puede interpretar lo siguiente:

1. La entropía para la región cerca de 12 grados aumenta después de 200°C, indicando una gradual destrucción de la γ^* -bohemia para producir γ -alúmina conforme la temperatura se incrementa. Es evidente que γ^* -bohemia es el principal precursor para la formación de γ -alúmina ya que estas fases tienen una estructura cristalina semejante.

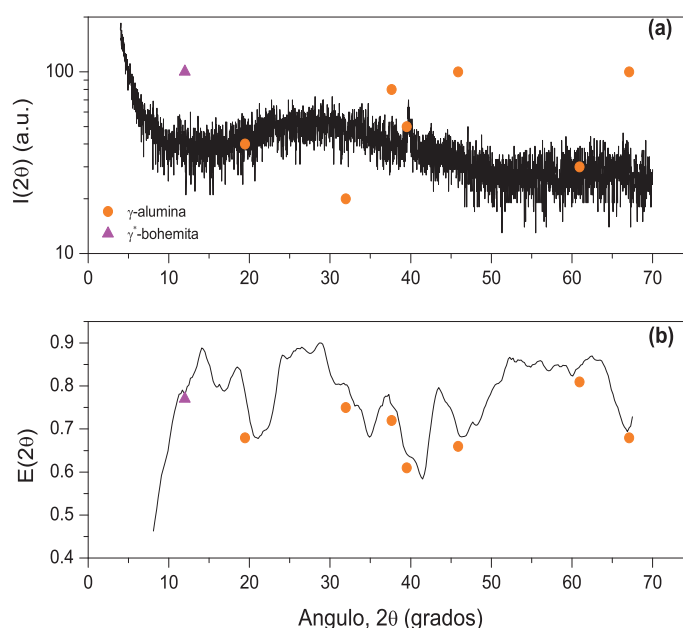


Figura 4.11. (a) Intensidad de patrones $I(2\theta)$ de rayos X , (b) Patrones de entropía $E(2\theta)$ de la señal de intensidad $I(2\theta)$ de alúmina calcinada a 500°C .

2. La entropía de los grupos óxidos, localizado cerca de 22 grados, muestra un comportamiento no monótono, el cual podría reflejar transiciones complejas de estas fases a la fase γ -alúmina.
3. La entropía cerca de los 40 grados muestra un comportamiento no monótono, el cual refleja la transición gradual de bayerita y γ -bayerita a la fase γ -alúmina. Se observa que la entropía sufre un fuerte aumento cuando la temperatura de calcinación está por arriba de 500°C , lo cual sugiere una extensa destrucción de la geometría cristalina. Esta observación apoya la práctica de usar temperaturas de calcinación no mayores a 500°C para preparar catalizadores con base alúmina Gates y col. (1979).
4. La entropía para la región cerca de los 50 grados muestra un incremento hasta un máximo en la temperatura de 200°C y posteriormente una disminución hasta un mínimo para la temperatura de 500°C y después un fuerte incremento en su comportamiento, indicando la gradual destrucción de γ -bohemita a estructuras de fases γ conforme la temperatura de calcinación se incrementa hasta los 500°C .

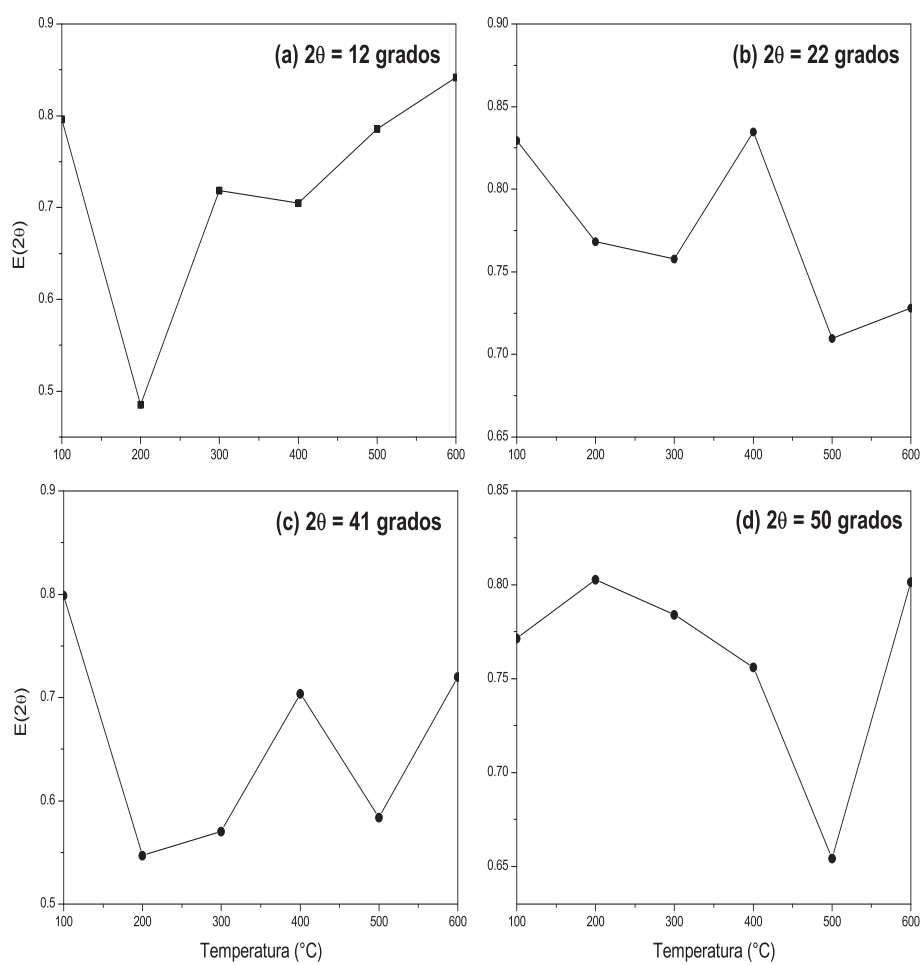


Figura 4.12. Patrones de entropía $E(2\theta)$ como función de la temperatura de calcinación para cuatro diferentes regiones angulares.

Los resultados anteriores muestran que el análisis de entropía aproximada multiescala puede ser usado como una herramienta para monitorear los cambios de fases de diferentes materiales durante el proceso de calcinación.

Conclusiones

En esta sección se ha presentado una metodología basada en el análisis de entropía aproximada multiescala para caracterizar los patrones de difracción de rayos X con ruido intenso. La idea es detectar regiones angulares donde las fluctuaciones de intensidad de difracción de rayos X no son aleatorias del todo, sino que exhiben cierto grado de regularidad. Para lograrlo, regularidades en las señales de intensidad se cuantifican mediante un análisis de entropía aproximada multiescala que pueden ser asociados con patrones geométricos contenidos en la compleja estructura del material. El enfoque del análisis fue ilustrado con alúmina calcinada a diferentes temperaturas, mostrando que el método puede proveer información sobre la localización de fases de alúmina corroborados contra tarjetas *PDF*.

4.3.3. Análisis fractal

En esta sección se propone el enfoque del análisis del rango reescalado para analizar los patrones de *XRD* con ruido intenso. Esta técnica ha sido empleada en diversas disciplinas desde finanzas (Opong y col., 1999), geofísica (Chamoli y col., 2007), análisis de fracturas (Varotsos y col., 2002), fisiología (Martini y col., 2004), ingeniería mecánica (Vela-Martínez y col., 2009), minería de datos (Zhongying y col., 2011), entre otras.

Metodología

En esta sección se presenta la idea de que el patrón de ruido intenso no es tan aleatorio del todo, y que puede contener patrones que pueden asociarse a fases del material. La metodología usa el análisis del rango reescalado para detectar regiones angulares de patrones de *XRD* con ruido intenso, $I(2\theta)$. La idea es que las correlaciones intrínsecas $I(2\theta)$ reflejan un cierto grado de orden en la geometría del material complejo. En esta forma, un fuerte pico

en el patrón de intensidad $I(2\theta)$ que refleja estructuras de un material altamente organizado corresponde a una alta correlación. Por otra parte, geometrías altamente desorganizadas corresponden a un comportamiento aleatorio en las fluctuaciones de intensidad de la difracción de rayos X , las cuales son cuantificadas por el exponente de Hurst cercanas a 0.5.

Con base en estas ideas, la metodología propuesta para analizar los patrones de XRD con ruido intenso consiste en transformar la señal de intensidad $I(2\theta)$ en un patrón de exponentes de Hurst $H(2\theta)$ mediante el análisis del rango reescalado R/S ;

$$I(2\theta) \rightarrow H(2\theta)$$

La transformación antes mencionada se logra al implementar el análisis de rango reescalado R/S (Capítulo 3.5) mediante un esquema de ventana deslizante. De esta manera, el análisis para caracterizar la presencia de fases del material se lleva a cabo en los patrones de correlación $H(2\theta)$ en vez de en la señal de intensidad original $I(2\theta)$. En cierto sentido, el análisis R/S puede ser visto como un operador no lineal que filtra la señal de intensidad para remover las componentes de alta frecuencia a un dominio de señal transformada. Recordando que $H = 0.5$ corresponde a señales independientes no correlacionadas (ejemplos, procesos de ruido blanco Gaussianos, t-Student o gamma) y que un incremento del exponente de Hurst refleja fuertes correlaciones (por ejemplo, un orden geométrico dominante), es conveniente introducir un índice de cristalinidad fractal como sigue:

$$C(2\theta) = H(2\theta) - 0.5$$

De esta manera, $C(2\theta) = 0$ deberá corresponder a un material amorfo total, para el cual ninguna regularidad puede encontrarse en su estructura geométrica. Mientras mayor sea el valor de $C(2\theta)$ mayor será su orden molecular. Los máximos valores $C(2\theta)$ deben corresponder a estructuras cristalinas totales que inducen a fuertes picos en la señal.

Resultados

Para cada corrida, los datos experimentales de *XRD* consistieron de 3301 observaciones. Para balancear la localidad y precisión de los computos de correlación (por ejemplo, el exponente de Hurst), la ventana deslizante fue seleccionada como 300 observaciones. Ventanas más pequeñas conducen a inestabilidades numéricas del exponente de Hurst por la pendiente del $\log(R/S)_\tau$ contra $\log t$. Por otra parte, una ventana excesivamente grande puede conducir a locaciones inciertas de regiones angulares con alta correlación, que pueden ser relacionadas con las fases del material.

La Figura 4.13.a muestra la intensidad $I(2\theta)$ de patrones de *XRD* para la temperatura de calcinación de 200°C. Para esta temperatura de calcinación, la alúmina está compuesta probablemente por bayerita y ciertas formas de bohemita, específicamente γ^* -bohemita que tiene la misma estructura cristalina que la γ -alúmina Wang y col. (1999). La figura 4.10.a también muestra la localización, extraídas de tarjetas *PDF*, para bayerita (*Card* 20 – 0011) y bohemita (*Card* 21 – 1307), indicando que los picos a ~ 40 grados pueden ser relacionados a bayerita. Siguiendo el estudio realizado por Wang y col. (1999), los picos a ~ 10 grados pueden ser relacionados a la presencia de γ^* -bohemita. Se ha sugerido que los picos a ~ 22 grados pueden estar relacionados a la presencia de grupos de óxidos complejos (Liu y col., 2008). A pesar de que los fuertes picos en los patrones de *XRD* pueden ser explicados por la presencia de algunas fases de alúmina, la localización de otras fases probables, como bayerita, es incierta debido a la presencia de fluctuaciones de intensidad de alta frecuencia que afectan la interpretación de los resultados experimentales.

La Figura 4.10.b muestra el correspondiente patrón de cristalinidad $C(2\theta)$ obtenidos por el análisis fractal R/S . Las siguientes características pueden ser observadas:

1. Se planteó la hipótesis que los fuertes picos de intensidad deben corresponder patrones de correlación altos medidos en términos del exponente de Hurst $H(2\theta)$. Los resultados en la Figura 4.10.b respaldan esta idea mostrando que la intensidad marcada en los picos a ~ 10 , 22 y 40 grados corresponden a los máximos del índice de cristalinidad $C(2\theta) = H(2\theta) - 0.5$.

2. El índice de cristalinidad $C(2\theta)$ no muestra fuertes picos como el patrón de intensidad XRD $I(2\theta)$. En lugar, el índice de cristalinidad $C(2\theta)$ obtenido del análisis fractal R/S está distribuida sobre valores máximos. Esto sugiere que las fases cristalinas en la estructura de la alúmina no se encuentran aisladas sino que coexisten en una distribución geométrica compleja, probablemente fractal, que se refleja en intenso ruido de los patrones de intensidad $I(2\theta)$ de XRD. El análisis R/S tiene la habilidad de "filtrar" los patrones de intensidad de ruido y revelar la presencia de regiones de alta correlación que pueden ser atribuidas a una o más fases de material. Por ejemplo, la región angular cerca de los 20 grados sugiere un patrón de cristalinidad conformado por bayerita y grupos óxidos. Similarmente, cerca de los 40 grados se puede relacionar un patrón complejo de cristalinidad conformado por bayerita y bohemita.
3. El análisis R/S puede revelar la existencia de regiones angulares asociadas con correlaciones no triviales que no son exhibidas por los patrones originales de XRD. De hecho, el análisis R/S reveló la existencia de patrones no aleatorios cerca de los 50 grados, con un índice de cristalinidad similar al de la cercanía de los 40 grados. Al contrario del patrón de intensidad original $I(2\theta)$, el índice de cristalinidad fractal sugiere que la región de 50 grados puede estar relacionada con la fase bohemita.
4. En contraste con el patrón de intensidad original en la Figura 4.13.a, el exponente de Hurst muestra claramente la ubicación de picos que pueden relacionarse con diferentes fases de alúmina. El mayor valor del índice de cristalinidad fractal es de aproximadamente 1.1 a ~ 10 grados, lo cual puede ser atribuido a la presencia de la fase γ^* -bohemita como ya se ha sugerido por Wang y col. (1999).

Los resultados anteriores muestran la habilidad del análisis R/S para detectar regiones angulares en el ruido intenso y que pueden ser atribuidas a las fases del material. En el caso de la temperatura de calcinación a 500°C es más retador ya que la presencia de otras fases γ -alúmina no pueden ser localizadas directamente de la señal de intensidad original $I(2\theta)$. Por ejemplo, en la Figura 4.14.a se observa solo un pico prominente a 40 grados, lo

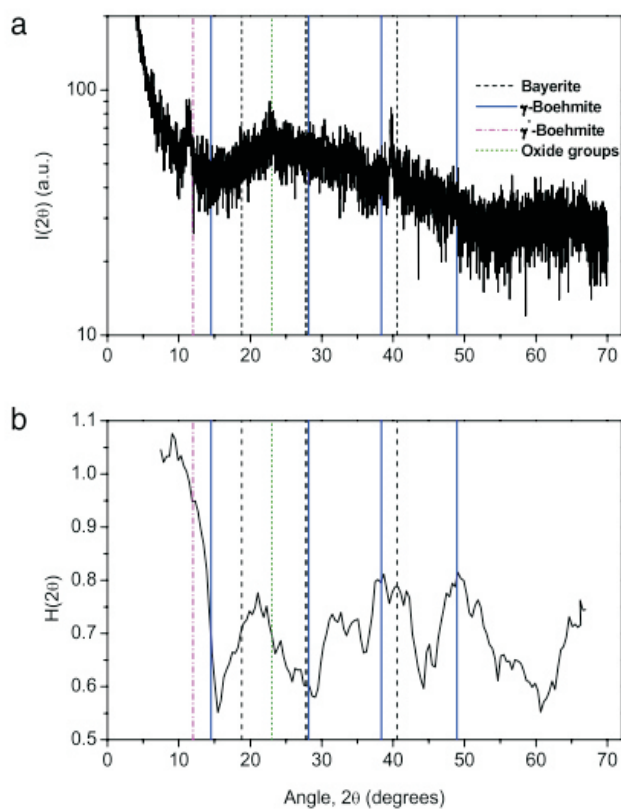


Figura 4.13. (a) Intensidad de patrones $I(2\theta)$ de rayos X, (b) Patrón del exponente de Hurst para la intensidad de patrones $I(2\theta)$ de rayos X para alúmina calcinada a 200°C.

que sugiere que la fase γ -alúmina se ha formado. Sin embargo, la presencia de otras fases γ -alúmina no puede discernirse a partir de la ubicación de la tarjeta *PDF* 10-0425. Por ejemplo, la fase localizada a 20 grados no exhibe ningún pico prominente en la señal original. La presencia de γ^* -bohemitita cerca de los 10 grados se observa ligeramente como un pico pequeño en tal región angular. De esta forma, la presencia de la fase γ -alúmina no puede ser afirmada desde una directa inspección de la intensidad de patrones en la región cerca de los 20 grados.

La Figura 4.14.b muestra el comportamiento de la cristalinidad fractal $C(2\theta)$ correspondiente al patrón *XRD* de la Figura 4.14.a. En donde se muestran varias regiones de correlación que son reflejadas como índices de cristalinidad positiva. Interesantemente, todas

estas regiones de correlación pueden estar asociadas a la presencia γ -alúmina en el material calcinado. El máximo índice de cristalinidad está localizado en la región de los 40 grados donde se encuentra el fuerte pico mostrado en la señal de intensidad de *XRD*. Se observa que se detectan las fases γ -alúmina tanto para la región de los 40 y 20 grados; y otras regiones angulares reportadas en tarjetas *PDF* de alúmina. También se observa la presencia de γ^* -bohemita en el índice de cristalinidad $C(2\theta)$ cerca de los 10 grados. Interesantemente, el índice de cristalinidad de la γ^* -bohemita disminuyó desde 0.6 a 0.3 en comparación con la temperatura de calcinación a 200°C. Esto sugiere, como se esperaba, que fases de bohemita son transformadas a γ -alúmina conforme la temperatura de calcinación se incrementa dentro del rango de 100 – 500°C.

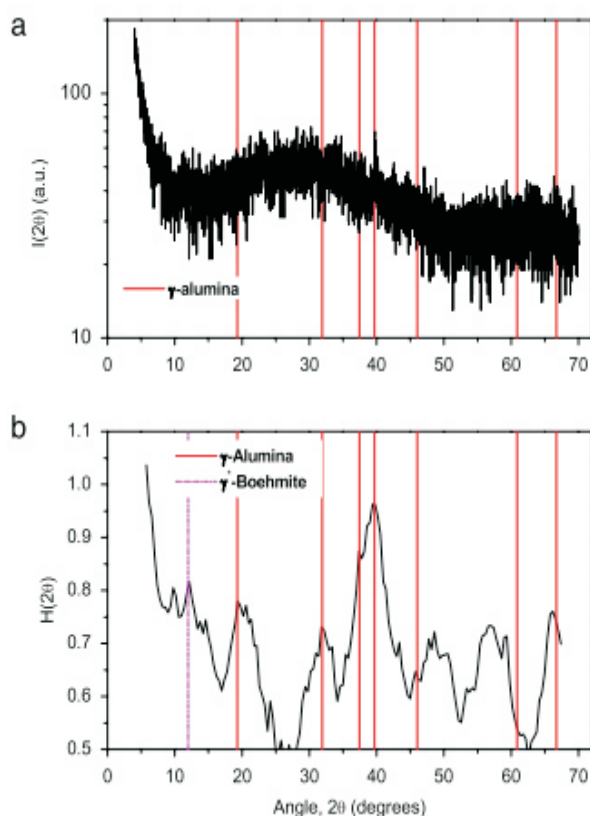


Figura 4.14. (a) Intensidad de patrones $I(2\theta)$ de rayos X, (b) Patrones del exponente de Hurst de la señal de intensidad $I(2\theta)$ de alúmina calcinada a 500°C.

Conforme la temperatura de calcinación se cambió, nuevas fases fueron obtenidas por la transformación a otras fases. De esta forma, el exponente de Hurst de ciertas fases fue reducido como efecto de la transformación en la estructura del material. La Figura 4.15 presenta la evolución del exponente de Hurst para regiones en la vecindad de cuatro diferentes ángulos. Lo siguiente puede ser comentado:

1. El exponente de Hurst para la región cerca de los 12 grados disminuye después de 200°C, indicando una gradual destrucción de la γ^* -bohemita para producir γ -alúmina conforme la temperatura se incrementa. Es evidente que γ^* -bohemita es el principal precursor para la formación de γ -alúmina ya que estas fases tienen una estructura cristalina semejante.
2. El exponente de Hurst de los grupos óxidos, localizado cerca de 22 grados, muestra un comportamiento no monótono, el cual podría reflejar transiciones complejas de estas fases a la fase γ -alúmina.
3. El exponente de Hurst cerca de los 40 grados muestra un comportamiento monótono creciente, el cual refleja la transición gradual de bayerita y γ -boehmita a la fase γ -alúmina. Se observa que el exponente de Hurst sufre un fuerte decremento cuando la temperatura de calcinación está por arriba de 500°C, lo cual sugiere una extensa destrucción de la geometría cristalina. Esta observación apoya la práctica de usar temperaturas de calcinación no mayores a 500°C para preparar catalizadores con base alúmina Gates y col. (1979).
4. El exponente de Hurst para la región cerca de los 50 grados muestra un comportamiento monótono decreciente, indicando la gradual destrucción de cristales de γ -bohemita a estructuras de fases γ conforme la temperatura de calcinación se incrementa hasta los 500°C.

Los resultados anteriores muestran que el exponente de Hurst puede ser usado como una herramienta para monitorear cambios de fases de diferentes materiales durante el proceso de calcinación.

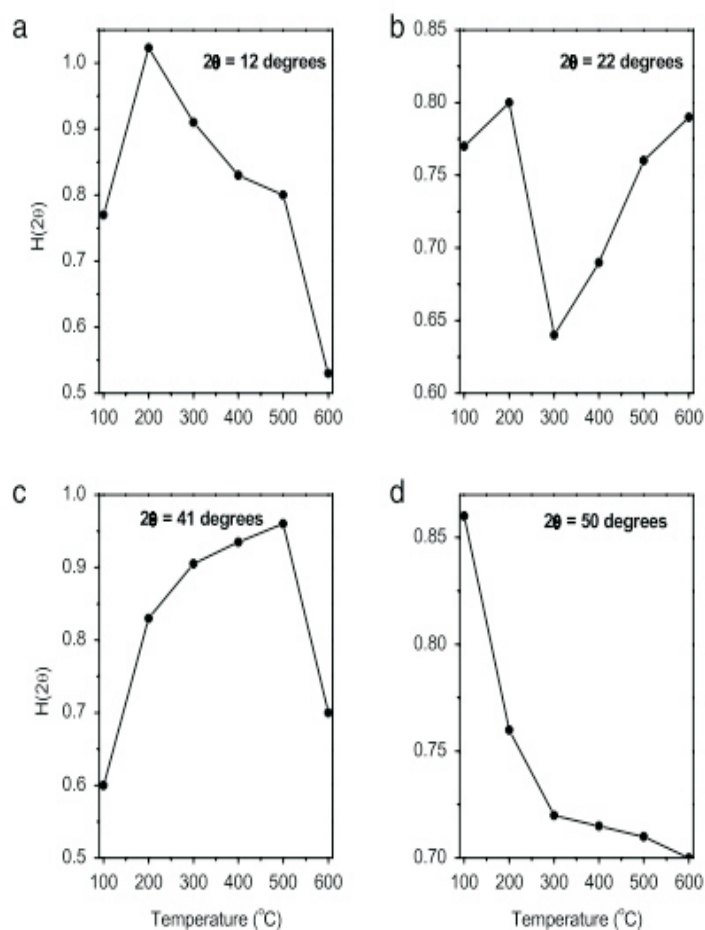


Figura 4.15. Exponente de Hurst como función de la temperatura de calcinación para cuatro diferentes regiones angulares.

Conclusiones

En esta sección se ha presentado una metodología basada en el análisis fractal para caracterizar los patrones de difracción de rayos X con ruido intenso. La idea es detectar regiones angulares donde las fluctuaciones de intensidad de difracción de rayos X no son aleatorias del todo, sino que exhiben cierto grado de regularidad reflejadas como correlaciones internas. Para lograrlo, regularidades en las señales de intensidad se cuantifican mediante el análisis R/S por medio del llamado exponente de reescalamiento de Hurst que revela la regularidad de los patrones geométricos contenidos en la compleja, probablemente fractal,

estructura del material. El enfoque del análisis fue ilustrado con alúmina calcinada a diferentes temperaturas, mostrando que el método puede proveer información sobre la localización angular de picos regulares asociados a líneas prominentes (tarjetas *PDF*) de fases de la alúmina. Un índice de cristalinidad fue introducido en términos de una distancia de aleatoriedad, así la regularidad de una cierta fase puede ser cuantificada incluso cuando el material no es cristalino.

4.4. Eficiencia de mercados de petróleo crudo

Debido a la importancia fundamental del petróleo, su incremento de precio puede ser responsable de grandes variaciones en las economías locales, provocando así un fuerte impacto en la actividad económica mundial. En esta sección se estudió la eficiencia informativa de los mercados de petróleo crudo por medio de métodos de análisis de entropía multiescala. Se utilizaron series de datos al cierre de *WTI* (por sus siglas en inglés, West Texas Intermediate), donde se analizó la evolución de la complejidad informática y eficiencia del mercado de petróleo crudo. Los resultados indicaron que el mercado del petróleo crudo es, en términos de la complejidad informática, eficiente durante el período analizado a excepción de dos períodos que corresponden a la década de 1990 y finales del 2000 (períodos de recesión en la economía americana). En general, los resultados mostraron que la desregulación ha mejorado el funcionamiento del mercado en el sentido de hacerlo menos predecible.

4.4.1. Introducción

Las relaciones entre los precios del petróleo crudo y la actividad económica involucran complejos mecanismos que intervienen durante la formación de precios en los mercados del petróleo. Lo anterior, ha sido ampliamente investigado por Rasche y Tatom (1977); Hamilton (1983); Santini (1985); Gisser y Goodwin (1986); Rotemberg y Woodford (1996); Carruth y col. (1998); Hamilton (2003); Barsky y Kilian (2004); Oladosu (2009). Las coincidencias entre los altos incrementos en el precio del petróleo crudo y el estallido de las recesiones

económicas han sido documentados (Hamilton, 1983; Mork, 1989), lo que sugiere que la dinámica del precio del petróleo crudo puede ser utilizada como indicador de la actividad económica mundial. Otros estudios apuntan hacia efectos permanentes de los precios del petróleo crudo sobre la inflación y, a corto plazo, con efectos asimétricos sobre las tasas de crecimiento de la producción (Cuñado y Perez de Gracia, 2003; Oladosu, 2009). En resumen, los resultados de diversas investigaciones han indicado que los mercados del petróleo crudo tienen un impacto importante en el desempeño de las economías regionales y mundiales, y que esta influencia tiene un carácter complejo y multifactorial.

La importancia de la hipótesis del mercado eficiente (*EMH*, Efficient Market Hypothesis) se basa en el hecho de que en un mercado eficiente toda la información disponible y relevante se refleja completamente e inmediatamente en el precio de un valor de mercado a fin de que nadie pueda tomar ventaja de esta información (Fama, 1970, 1991). De esta manera, no hay activos ni subvaluados ni sobrevaluados en un mercado eficiente, y el precio de mercado de los activos financieros constituye una guía adecuada para el presupuesto de capital y su asignación.

Dependiendo de la naturaleza y la fuente de la información, tres formas de eficiencia de mercado se puede establecer: eficiencia de forma sólida, semisólida y débil (Fama, 1970, 1991). En la eficiencia de forma sólida, el conjunto de información disponible puede ser a la vez información pública y privada. Se limita para la forma semi-sólida a toda la información pública y a sólo los movimientos de precios pasados para la forma débil. La mayoría de las pruebas empíricas para el *EMH* se enfocan en la forma débil. De esta manera, la eficiencia del mercado está relacionada con la ausencia de condiciones de arbitraje ya que el comportamiento del mercado de petróleo crudo no puede predecirse utilizando la dinámica de rendimientos de precios pasados. En resumen, los resultados en relación con el cumplimiento de la *EMH* deben proporcionar pistas importantes sobre la dinámica de los mercados del petróleo crudo y de su impacto en las economías mundiales y regionales.

Arouri y col. (2010) han señalado que la eficiencia del mercado es deseable para modelos de valoración de activos y el proceso de inversión en la toma de decisiones; mientras tanto se

apoya en supuestos fuertes tales como, mercados sin fricciones, disponibilidad de información y transparencia, racionalidad de los inversores y el arbitraje. Recientemente, Alvarez-Ramirez y col. (2008a) utilizaron métodos basados en el análisis de fluctuación sin tendencia para mostrar que los mercados de petróleo crudo son consistentes con la *EMH* en horizontes largos de tiempo, aunque variaciones auto-correlacionadas pueden ser exhibidas en escala de tiempo cortas. Charles y Darné (2009) utilizaron pruebas de relaciones de variancia para mostrar que el mercado del petróleo crudo Brent es eficientemente débil mientras que el mercado del petróleo crudo *WTI* parece que fue ineficiente en el sub-período 1994-2008. En contraste con el reporte de investigación de Tabak y Cajueiro (2007) este resultado sugiere que la desregulación no ha mejorado la eficiencia del mercado del petróleo crudo *WTI* en el sentido de hacer que los rendimientos sean menos predecibles. Por medio del análisis de fluctuación multiescala, Wang y Liu (2010) sugieren que los comportamientos a corto, mediano y largo plazo fueron, generalmente, transformados en comportamientos eficientes a través del tiempo. Lean y col. (2010) no encontraron ninguna evidencia de la varianza media y del dominio estocástico en precios de crudo al contado y a futuro, dando a entender que no hay oportunidad de arbitraje entre estos dos mercados y que los mercados al contado y a futuros son eficientes y racionales. Alvarez-Ramirez y col. (2010) estudiaron auto-correlaciones fractales retardadas de los precios al contado del *WTI* para encontrar que la presencia de auto-correlaciones pueden estar enmascaradas por efectos de retardo. Arouri y col. (2010) utilizaron modelos de espacio de estado para encontrar evidencias de la predictibilidad a corto plazo en los precios del petróleo crudo a través del tiempo, por lo que la hipótesis de convergencia hacia la eficiencia informativa débil debería ser rechazada. Los resultados de los enfoques no lineales han sugerido que el mercado del petróleo crudo no es eficiente ya que la dinámica de precios se puede predecir hasta cierto punto mediante el uso de modelos no lineales (Wang y Yang, 2010), algoritmos genéticos (Fan y col., 2008) y la descomposición de ondas (tipo wavelet) (Jammazi y Aloui, 2012).

Por otra parte, el análisis de entropía (en términos informativos) de series financieras se ha visto restringido a un número limitado de estudios de investigación. Por ejemplo,

Gulko (1999) propuso primeramente el uso de los conceptos de entropía para estudiar series de tiempo financieras, mostrando que el formalismo de máxima entropía, también llamado eficiencia informativa, haciendo la hipótesis del mercado eficiente operacional y medible. Este formalismo se utiliza para establecer que mercados entrópicos no admiten el arbitraje y se basan en la teoría de arbitraje de precios Ross y en el modelo de Black-Scholes. Darbellay y Wuertz (2000) demostraron la utilidad de los conceptos de entropía para caracterizar series de tiempo financieras, mostrando que la principal ventaja del enfoque de entropía reside en su capacidad para tener en cuenta dependencias no lineales en la estructura de autocorrelación de los sistemas dinámicos subyacentes.

Mientras que algunas otras investigaciones, en términos de entropía, se enfocan en la cuantificación de la eficiencia informativa del mercado en los mercados de petróleo crudo. Kaffashi y col. (2008); Hassan y col. (2011); Pincus y Kalman (2004) sugirieron que el algoritmo de entropía aproximada es adecuado para el análisis de series de tiempo financieras, ya que puede ser aplicado a secuencias muy cortas y además puede ser utilizado como un marcador de la estabilidad del mercado. Recientemente, los conceptos de entropía se han utilizado para cuantificar la eficiencia del mercado de divisas y los mercados de valores. Dado que la entropía es un índice de la cantidad de información (medida en términos de riqueza de patrones) contenida en una serie de tiempo, la alta entropía puede estar relacionada con la baja previsibilidad de la dinámica del mercado y, por lo tanto, a la alta eficiencia del mercado.

La idea de partida es que los precios en los mercados eficientes no se pueden predecir a causa de la falta de correlaciones intrínsecas y patrones regulares. Es decir, los rendimientos de la trayectoria de precios para un mercado de forma débil en la *EMH* deben corresponder al ruido estocástico no correlacionado. A su vez, los rendimientos deberían mostrar un máximo en el contenido de entropía. En relación con los resultados existentes en la literatura sobre eficiencia informativa de los mercados de petróleo crudo (Serletis y Andreadis, 2004; Tabak y Cajueiro, 2007; Maslyuk y Smyth, 2008; Alvarez-Ramirez y col., 2010; Arouri y col., 2010), el aporte basado en el análisis de entropía se puede resumir como sigue:

- Un índice de eficiencia de mercado informacional se introduce en términos de una distancia entrópica a la aleatoriedad. Trabajos anteriores se han centrado en el estudio de eficiencia de mercado informacional como una cuestión de todo o nada. En el enfoque propuesto, el índice de eficiencia informacional proporciona un medio para cuantificar una distancia desde la forma de eficiencia informativa débil.
- En concordancia con los conceptos similares de hipótesis de mercados adaptativos (Lo, 2004), el análisis empírico mostró que la información de eficiencia de mercado informacional para el mercado del petróleo crudo exhibe cambios temporales importantes que dependen del horizonte de tiempo. De esta forma, implica que los participantes en el mercado de petróleo crudo (ejem., inversionistas, gobiernos, productores, etc.) se adaptan a condiciones cambiantes endógenas y exógenas, (ejem., presupuesto público ajustable o la maximización de las ganancias).
- Se muestra un movimiento conjunto entre la eficiencia de mercado del petróleo crudo y recesiones económicas de los EE.UU. Esto sugiere que la eficiencia de mercados de petróleo crudo puede ser utilizada como un indicador de la evolución de las recesiones económicas.

En resumen, se reconoce la complejidad del problema, para lo cual se propuso un índice cuantitativo de la eficiencia informativa del mercado del petróleo crudo, y además se evaluaron los cambios temporales que pudieran estar relacionados con eventos socio-económicos.

4.4.2. Metodología

Se analizaron los datos de precios del petróleo crudo *WTI* de cierre diarios durante el período de desregulación del mercado de los EE.UU. El *WTI* (por sus siglas en inglés, West Texas Intermediate) se emplea como precio de referencia para fijar el precio de otros petróleos crudos producidos en medio oriente o el mar del Norte. Aunque intermediarios del petróleo crudo Brent y West Texas (*WTI*) tienen diferentes precios en dólares US/barriles, algunos

estudios (Fattouh, 2010) han mostrado solo ligeras diferencias en la dinámica de precios. De esta manera, solo se enfocó en el análisis al precio *WTI*, el cual puede ser libremente obtenido de la Administración de Información de Energía (*EIA*) en el departamento de energía de los EE.UU. (Administration).

Los datos abarcan desde el 1 de Enero de 1986 al 15 de Marzo del 2011 (6520 observaciones). Las Figuras 4.16.a y 4.16.b muestran la representación gráfica de los precios y de las diferencias logarítmicas de la serie de tiempo, respectivamente. Las ocurrencias de eventos extremos son resaltadas, los cuales presuntamente tuvieron un impacto importante en la dinámica del mercado. En la década del 1998 al 2008, el precio del petróleo crudo se incrementó más del 700% en promedio.

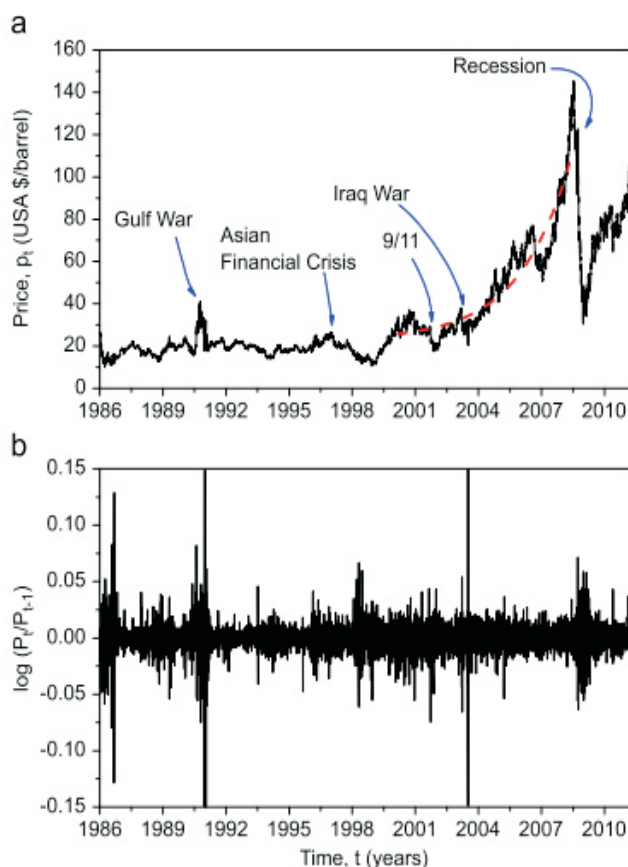


Figura 4.16. (a) Precios del WTI para el período del 1 de Enero de 1986 al 15 de Marzo del 2011. (b) Diferencias logarítmicas de precios para el mismo período del panel (a).

Se empleó el análisis de entropía debido a que esta técnica muestra mejor rendimiento que otras (ejem., análisis de Hurst) para series de tiempo de pequeña longitud y permite la introducción de un índice de eficiencia de mercado como una medida de aleatoriedad y detecta cambios bruscos en la eficiencia sin especificar un modelo de evolución. Durante el análisis se emplearon las diferencias logarítmicas de los precios del petróleo crudo. La media y la desviación estándar de estas diferencias fueron 0.024% y 2.32% diariamente, respectivamente.

4.4.3. Resultados

Para demostrar que el análisis de entropía depende de la escala de tiempo τ se analizaron las secuencias de diferencias de precios logarítmicas para el período de 1986–1989. La Figura 4.17 muestra las señales obtenida después de emplear un filtro pasa bajas para escalas de tiempo $n=1$ (diaria, señal original), $n=5$ (semanal), $n=20$ (mensual) y $n=60$ (trimestral). Cabe resaltar que las señales fueron normalizadas por su desviación estándar.

Los resultados del análisis de entropía multiescala para el período 1986-1989 se muestran en la Figura 4.18. Puede observarse que la entropía depende de la escala de tiempo $AE(\tau)$, y que además exhibe una tendencia decreciente ya que el contenido de información es reducido con respecto al incremento en la longitud promedio de n . Sin embargo, se observa un máximo local cerca de 72 días hábiles, lo cual sugiere que la información del mercado es agregada cerca de escalas de tiempo trimestrales para el período 1986-1989.

La metodología de entropía multiescala se empleó con el objetivo de estudiar algunos hechos a cerca de la complejidad y eficiencia de mercados del petróleo crudo en los recientes 25 años. El método se basa en monitorear las variaciones de entropía con respecto a la escala y al tiempo en una ventana deslizante de tamaño N_s . De esta forma, como un pre requisito para el cálculo computacional de la entropía se debe seleccionar un tamaño adecuado de ventana N_s . La entropía es una función creciente del tamaño de ventana N_s , donde los patrones existentes no son destruidos a medida que aumenta N_s .

Para evaluar los efectos del tamaño de ventana, la entropía fue computarizada a través

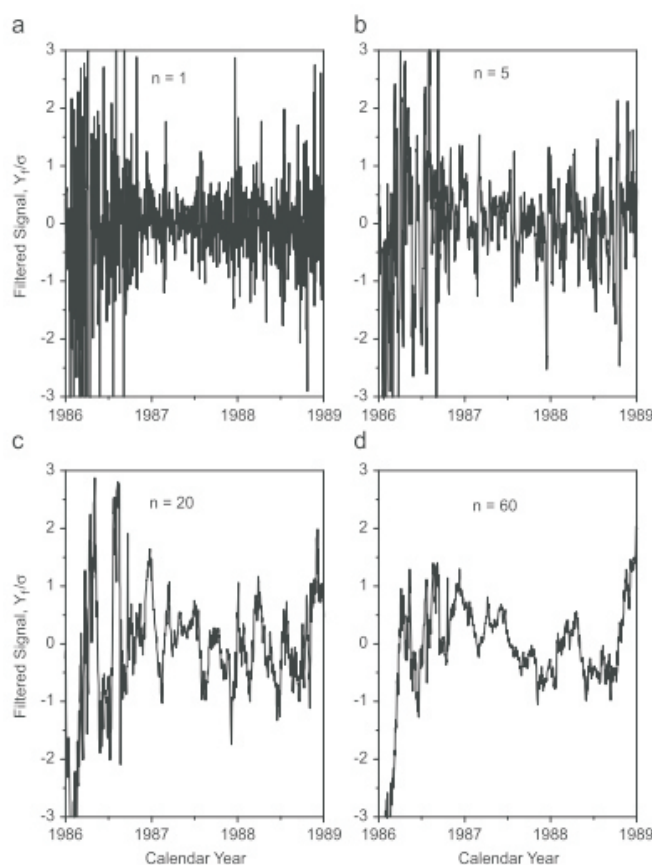


Figura 4.17. Secuencias de diferencia de precios logarítmicas utilizando un filtro pasa bajas para escalas (a) diaria ($n=1$), (b) semanal ($n=5$), (c) mensual ($n=20$) y (d) trimestral ($n=60$).

de todos los tamaños de ventana N_s contenidas en la serie de tiempo, posteriormente se utilizó por un filtro pasa bajas para una escala de tiempo $\tau = n$ días hábiles. Después se realizó un promediado sobre todas las posibles ventanas contenidas en la serie de tiempo Y_f . La Figura 4.19 muestra el comportamiento de la entropía como una función del tamaño de ventana N_s para escalas de tiempo diarias, semanales, mensuales y trimestrales. Para estas escalas de tiempo seleccionadas se observa que cuando $N_s > 500$, la entropía exhibe un comportamiento creciente bien definido con respecto a la escala de tiempo τ . Además puede observarse que, por ejemplo, para $N_s < 500$, la entropía promedio semanal es mayor que la entropía promedio diaria. Se seleccionó un tamaño de ventana de $N_s = 600$ días hábiles (cerca de 2.5 años) para el cálculo computacional de la entropía, lo cual garantizo que la

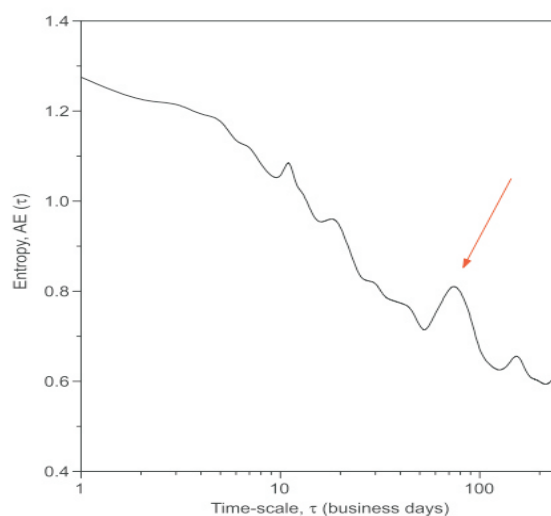


Figura 4.18. Patrones de entropía multiescala $AE(\tau)$ para el subperíodo 1986-1989 de acuerdo con las secuencias filtradas pasa bajas de la Figura 4.17.

entropía promedio es creciente con respecto a τ . Se podría haber seleccionado un mayor valor de N_s , aunque con esto se reduciría la localidad de las fluctuaciones de entropía con respecto al tiempo.

Complejidad del mercado informacional

En la Figura 4.20 se muestra el comportamiento de la entropía con respecto a la escala de tiempo τ , para cinco subperíodos seleccionados de tres años. Para escalas de hasta un trimestre, la entropía es una función decreciente de la escala de tiempo, lo cual indica que el contenido de información es reducido conforme el horizonte de tiempo se incrementa. En algunos casos, como en los subperíodos 1986-1989 y 1993-1996, se observan máximos prominentes locales a 80 y 200 días hábiles, respectivamente. Lo que sugiere que, la formación de precios a largo plazo se vio afectada en gran medida por los flujos de información anuales agregados a escalas temporales anuales. En contraste, muchos subperíodos muestran un decremento de entropía a través de todo un rango de escalas.

Gulko (1999) ha señalado que cuanto mayor es el número de patrones en los precios (máxima entropía), más difícil es la predicción de la dinámica de los precios en curso.

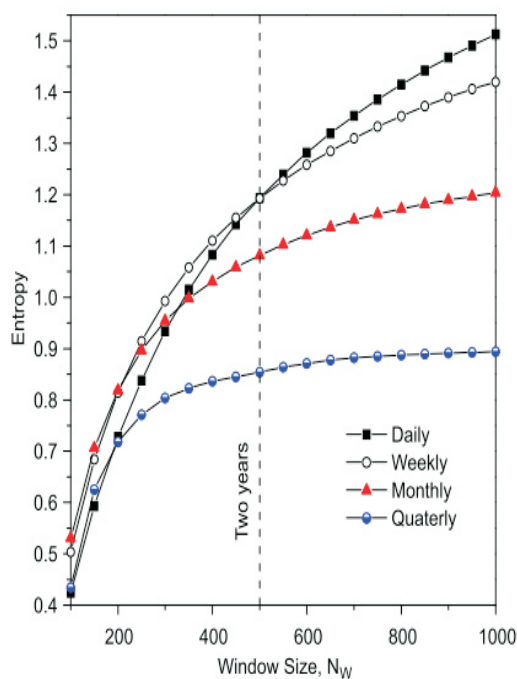


Figura 4.19. Comportamiento de la entropía como función del tamaño de ventanas N_s para escalas temporales diarias, semanales, mensuales y trimestrales.

Consecuentemente, mercados con bajos niveles de entropía son relativamente fáciles de predecir, en contraste con mercados con alta entropía para los cuales la predicción debe de requerir métodos y algoritmos más complejos. En este sentido, mercados con altos niveles de entropía son mucho más complejos que los mercados con bajos niveles de entropía. De acuerdo con esta idea, los resultados en la Figura 4.20 indican las siguientes características para la complejidad, en términos del contenido de entropía, de mercado del petróleo:

- La Figura 4.20 muestra que la complejidad de las diferencias logarítmicas de precios no-triviales es dependiente de la escala de tiempo τ . El patrón general es que la entropía es alta para pequeñas escalas de tiempo mientras que muestra decrementos para altas escalas de tiempo. Este comportamiento sugiere que la predictibilidad del precio del petróleo crudo es relativamente mayor en el largo plazo así como la entropía muestra valores relativamente pequeños para escalas de tiempo largas. También se considera que la entropía también evoluciona con el tiempo. Por ejemplo, el subperíodo 2004-2007

exhibe niveles altos de entropía a través de todo el rango de escala a diferencia de los otros cuatro subperíodos. Puede observarse que el mercado del petróleo crudo muestra un incremento en la diversidad de patrones de precios durante el subperíodo 2004-2007, el cual fue causado por muchos factores que incluyeron el crecimiento robusto de la economía de los EE.UU. seguido de la crisis del 2001, el incremento en la demanda sostenida del petróleo crudo por países asiáticos del Pacífico, el uso del petróleo crudo como un activo para la especulación financiera, etc. Por el contrario, el subperíodo 2007-2010 corresponde a la reciente recesión económica donde se tiene un severo decremento de entropía a través de todo el rango de escala, el cual puede corresponder a una importante reducción de la eficiencia del mercado donde la dinámica de precios se vuelven más predecibles.

- La dependencia de la escala en los patrones de entropía sugieren la siguiente estructura en el mercado del petróleo crudo. En la operación a corto plazo (de días a semanas), el mercado muestra la mayor complejidad (medida en términos de diversidad de patrones) así como mayores valores de entropía indicando que los movimientos de precios son poco predecibles. La diversidad del mercado a corto plazo en la dinámica de precios surge de los efectos combinados de la especulación (posiciones no comerciales), niveles de inventarios (Merino y Ortiz, 2005), eventos meteorológicos (ejem., huracán Katrina), etc. En contraste, a largo plazo, para escalas de tiempo altas mayores de un cuarto, el mercado es relativamente más predecible cuando la dinámica de precios a largo plazo depende de factores estructurales, tales como la macroeconomía y ambientes políticos, planificación de la producción a largo plazo (es decir, número y capacidad de los campos petrolíferos de crudo Saudíes y Venezolanos), la volatilidad del dólar, etc. De esta forma, la eficiencia a corto plazo debería contribuir a la estabilidad del mercado del petróleo crudo promoviendo la asignación eficiente de recursos, que se refleja con altos valores de entropía y reducida habilidad de especulación para explotar condiciones de arbitraje.

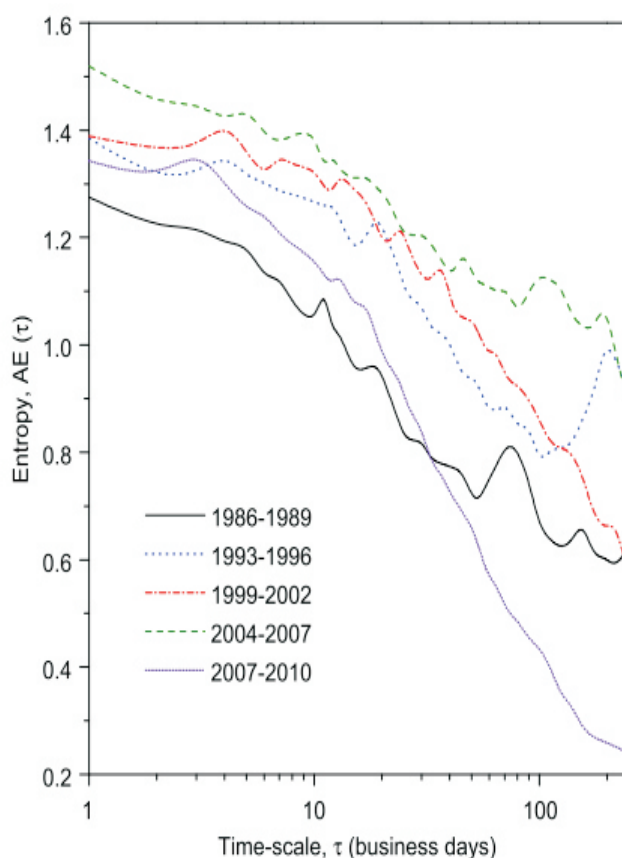


Figura 4.20. Comportamiento de la entropía con respecto a la escala de tiempo τ , para cinco subperíodos seleccionados de tres años.

Los resultados de la Figura 4.20 sugieren que en el corto plazo la dinámica del mercado es muy incierta, inducido por las dinámicas a corto plazo relacionadas con perturbaciones diarias y semanales, con convergencia estocástica hacia una tendencia de equilibrio en el precio impuesta por mecanismos de oferta-demanda a mediano y largo plazo (Wang y Liu, 2010). En cuanto a la entropía en altas escalas de tiempo, es evidente que la disminución de valores podría ser inducido por mercados futuros que tiran de los precios hacia precios de contratos con fecha de vencimiento en plazos trimestrales. La dependencia de la complejidad del mercado de petróleo crudo con el tiempo y la escala sugiere que los mercados participantes adaptan el procesamiento de la información a horizontes de tiempo específicos para realizar tareas de planificación energética y ganancias. Esto está en concordancia con el enfoque

evolutivo (Lo, 2004), que establece que la complejidad del mercado es una característica que varía continuamente en el tiempo y en todos los mercados.

Eficiencia del mercado informativo

Algunas investigaciones se han enfocado en la cuantificación de la eficiencia del mercado del petróleo crudo, en términos de entropía (Oh y col., 2007; Risso, 2008, 2009). El fundamento de la relación entre la eficiencia de mercado débil y entropía proviene del hecho de que los movimientos en los precios no se pueden predecir mediante el procesamiento de los movimientos de precios, por ejemplo, análisis técnicos. Las pruebas empíricas para demostrar si un mercado no es eficiente se han centrado en el rechazo de la hipótesis de que las diferencias logarítmicas de precios reflejan dinámicas estocásticas no correlacionadas (proceso de paso aleatorio). A su vez, esta hipótesis implica que los precios son impredecibles ya que no hay patrones regulares que puedan ser encontrados en la evolución del mercado. Campbell y col. (1997) argumentan que este enfoque es una idealización que es inalcanzable ya que el mercado es operado por agentes heterogéneos con diferentes puntos de vista y capacidades en el procesamiento de información (Giglio y col., 2008). En cambio, el concepto de eficiencia de mercado relativa puede ser un concepto más útil que el enfoque adoptado por la literatura tradicional. En otras palabras, un índice de cuantificación de eficiencia de los mercados debe ser relativo, proporcionando una medida del nivel de predictibilidad de la dinámica de mercados. Una alternativa para probar la *EMH* se puede dar desde el punto de vista de entropía si se considera que la dificultad de pronosticar un mercado está directamente relacionada con el número y la diversidad en el movimiento del precio (contenido de entropía).

Para construir el índice de referencia (benchmark) para la estimación de un grado de eficiencia del mercado se consideraron muestras de secuencias de ruido gaussiano no correlacionadas de tamaño $N_s = 600$ observaciones. La Figura 4.21.a muestra los límites del índice de referencia de los patrones de entropía para 10,000 muestras. Los límites superior e inferior representan, respectivamente, los valores de entropía máximo y mínimo para el conjunto de secuencias. De esta manera, el patrón de entropía multiescala para cualquier

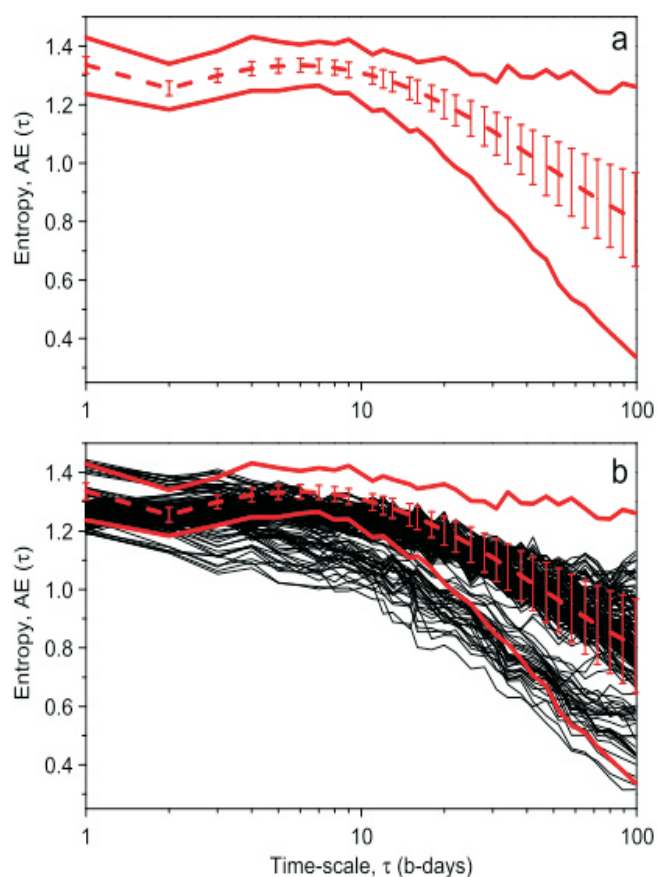


Figura 4.21. (a) Límites de referencia de los patrones de entropía para 10,000 muestras de secuencias gaussianas de ruidos no correlacionados de longitud $N_s = 600$ observaciones. (b) Patrón de entropía en comparación con el punto de referencia para la aleatoriedad.

secuencia de ruido gaussiano no correlacionado de tamaño $N_s = 600$ observaciones debe estar contenida en la envolvente, tal y cómo se muestra en la Figura 4.21.a. Se observa que, la geometría fractal es similar en la envolvente entre las fronteras, lo cual refleja el hecho de que incluso el ruido no correlacionado es un sistema dinámico fraccional. La idea subyacente es que la definición de la eficiencia del mercado informativo es que una secuencia real de tamaño N_s que debe contener la mayor diversidad de patrones si su patrón de entropía multiescala se encuentra dentro de la envolvente (Figura 4.21.a). Es decir, el mercado del petróleo crudo será completamente eficiente informativamente (es decir, 100 %) si la secuencia de diferencias logarítmicas está contenida dentro de la envolvente. En contraste, si el patrón de entropía

está por debajo del límite inferior, el mercado del petróleo crudo será sólo parcialmente eficiente.

Para establecer una medida de la eficiencia del mercado se estableció un índice de eficiencia del mercado informativo $I_{IME}(\tau)$ (Ecuación 4.1) para una escala de tiempo τ determinada. Con la introducción de la definición de $I_{IME}(\tau)$, se evita el enfoque de todo o nada permitiendo más flexibilidad al proporcionar una medida de cuantificación para la eficiencia del mercado de petróleo crudo.

$$I_{IME}(\tau) = \frac{AE(\tau)}{B_{min}(\tau)} \times 100 \quad (4.1)$$

Donde $AE(\tau)$ es la secuencia de entropía y $B_{min}(\tau)$ es el límite inferior de la envolvente del benchmark (Figura 4.21.a). De esta forma, $I_{IME}(\tau) = 100\%$ si $AE(\tau) \geq B_{min}(\tau)$. En contraste, $I_{IME}(\tau) < 100\%$ si el patrón de entropía no está contenido en el benchmark de entropía.

La Figura 4.21.b muestra el patrón de entropía multiescala $AE(\tau)$ para 100 submuestras de diferencias logarítmicas de precios del petróleo crudo. Se observa que el patrón de la entropía no está necesariamente contenido en la región de referencia (benchmark), por lo que la eficiencia del mercado depende de la escala de tiempo τ . Es decir, el mercado del petróleo crudo puede ser completamente eficiente para determinadas escalas de tiempo, y parcialmente eficiente para otras. También, la eficiencia del mercado depende del tiempo t . La Figura 4.22 muestra el índice de eficiencia del mercado informativo $I_{IME}(\tau)$ con respecto al tiempo t y escala τ . La reducción de la eficiencia del mercado (es decir, menos del 100%) se puede observar para finales de los 80 cuando el mercado del petróleo crudo fue adaptado a condiciones de desregulación. Se observaron otros dos períodos importantes, a principios de 1990 y finales del 2000, donde se observó una disminución en la eficiencia del mercado. De cualquier manera, en general, el mercado del petróleo crudo ha sido informativamente eficiente a lo largo de las últimas dos décadas, lo cual concuerda con los resultados obtenidos con el análisis de Hurst, que indica que el mercado del petróleo crudo es consistente con la hipótesis del mercado eficiente (Alvarez-Ramirez y col., 2008a). Al igual que en la conclusión

obtenida por Tabak y Cajueiro (2007), estos resultados sugieren que la desregulación del mercado en 1986 ha mejorado la eficiencia en el mercado del petróleo crudo en el sentido de hacer los rendimientos menos predecibles.

Los resultados anteriores confirman que el mercado del petróleo crudo es un sistema adaptativo evolutivo donde la eficiencia no es constante, sino que presenta variaciones importantes en el tiempo. Como ha sido señalado por Lim y Brooks (2011), las características de la microestructura del mercado, límites en el arbitraje, ruido en el comercio y la existencia de imperfecciones en el mercado son potenciales factores que pueden ser incrementados para períodos de salida en la eficiencia del mercado. Además, los resultados confirman los hallazgos previos desde la prueba de raíz unitaria los cuales demuestran que la dinámica de los precios del petróleo crudo puede ser caracterizada como un proceso de paso aleatorio, que se aleja de la eficiencia, siendo importantes y significativos en términos de grandes eventos que tienen impacto en la economía mundial (Maslyuk y Smyth, 2008).

Eficiencia de los mercados del petróleo crudo durante las recesiones económicas en Estados Unidos

En los Estados Unidos, las recesiones económicas han significado un declive en la actividad económica, estos descensos son impulsados por cambios en la reglamentación del gobierno y en las políticas fiscales, comerciales y monetarias. Los ciclos en la agricultura, el consumo y la inversión, así como la salud de la industria bancaria también contribuyen con estos descensos. Las recesiones en los EE.UU. han afectado crecientemente a economías en una escala mundial, especialmente debido a que la globalización ha entrelazado a las economías. Risso (2008) encontró evidencias de que la probabilidad de sufrir un colapso financiero en los mercados de valores aumenta a medida que disminuye la eficiencia informativa.

La Figura 4.22 muestra que el mercado del petróleo crudo ha sido informativamente eficiente a excepción de dos períodos que corresponden a dos grandes recesiones económicas en los EE.UU. De hecho, se pueden observar decrementos importantes en la eficiencia informativa del mercado $I_{IME}(\tau)$ para los subperíodos 1990-1993 y 2008-2012, que

corresponden a las dos crisis económicas más graves en los EE.UU. en los últimos 25 años. A principios de 1990, la recesión se prolongó desde Julio de 1990 a Marzo de 1991 y tenía un descenso de -1.4% del producto interno bruto (máximo al mínimo). La reciente recesión (llamada la Gran Recesión) se inició en Diciembre de 2007 y NBER (por sus siglas en ingles, The National Bureau of Economic Research) declaró oficialmente su término en Junio de 2009. En este caso, el deterioro del producto interno bruto fue más pronunciado con una disminución del 4.1% .

Al parecer estas dos recesiones han tenido un efecto importante en la eficiencia del mercado de petróleo crudo, lo que se vio reflejado en fuertes descensos de entropía sobre prácticamente todo el rango de escala de tiempo. Durante los períodos de recesión, la dinámica del mercado fue dominado por patrones regulares de precios así como de la diversidad de participantes y de las expectativas que son reducidas por efectos de deficiencias económicas. Se reconoció una baja económica breve en la década del 2000, la cual fue explicada por la caída de la especulativa burbuja (.com), la caída del gasto e inversión de las empresas, y los ataques del 11 de Septiembre. La caída del producto interno bruto fue corta, -0.3% , y la aplicación de políticas monetarias eficientes permitieron una agitación breve y superficial. La Figura 4.22 muestra que esta recesión no indujo un cambio significativo en la dinámica de la entropía del mercado de petróleo crudo. A su vez, esto sugiere que, dado que la recesión no introdujo un efecto importante en los mercados de energía, el origen de la recesión a principios del 2000 no fue sino estructural, inducida por perturbaciones exógenas que fueron en consecuencia absorbidas.

Los resultados de la Figura 4.22 están de acuerdo con hallazgos recientes (Risso, 2008) que muestran que para los mercados de valores diferentes, la probabilidad de tener una quiebra aumenta a medida que disminuye la eficiencia informativa. De hecho, las recesiones económicas graves han provocado importantes disminuciones de la eficiencia en diferentes escalas de tiempo. A su vez, estos decrementos de eficiencia podrían estar relacionados con disminuciones en la entropía del precio ya que el mercado se compromete con la direccionalidad del precio preferencial debido a, por ejemplo, la reducción de la demanda y las

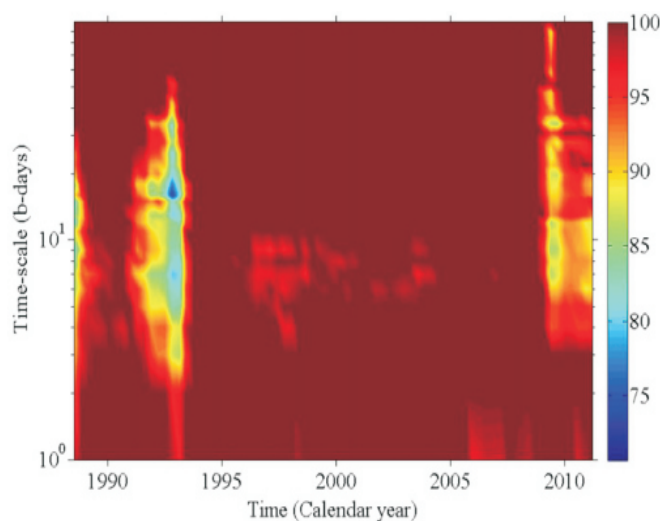


Figura 4.22. Índice de eficiencia informativa del mercado $I_{IME}(\tau)$ con respecto al tiempo t y escala τ .

expectativas de inversión. Por ejemplo, en la reciente crisis económica en el 2008, el petróleo crudo sufrió una fuerte disminución de alrededor de 160 dls/barril a unos 60 dls/barril inducido principalmente por el hundimiento de las expectativas de crecimiento económico. La dinámica de los precios del petróleo crudo fueron gobernados en su mayoría por ajustes negativos en el consumo, ruido e inversionistas del mercado. Como consecuencia, el contenido de aleatoriedad, y por lo tanto la diversidad de patrones, en la dinámica del precio del petróleo crudo se reduce en una gran cantidad. Si la eficiencia informativa del mercado de petróleo crudo, medido en términos de entropía, es visto como un indicador de duración de la recesión de los EE.UU., la Figura 4.22 indica que se logró una recuperación completa a principios de 1990 cuando la eficiencia del mercado de petróleo crudo se recuperó completamente y el crecimiento económico se restableció en 1993 (es decir, tres años después de que comenzara la recesión).

4.4.4. Conclusiones

En esta sección se utilizó el análisis de entropía multiescala bajo un esquema de ventana deslizante para estudiar la complejidad y eficiencia del mercado de petróleo crudo durante

el período 1986-2011. La idea subyacente es que cuanto mayor es la entropía, mayor será la diversidad de patrones en las fluctuaciones de precios. A su vez, esto implica que un mercado con altos valores de entropía es más complejo que aquellos con bajos valores de entropía. Los resultados obtenidos están en concordancia con los resultados anteriores (Alvarez-Ramirez y col., 2008a; Elder y Serletis, 2008; Arouri y col., 2010) en que se muestra evidencia de que la eficiencia del mercado varía en el tiempo para rendimientos diarios de petróleo crudo durante el período de desregulación que comenzó en 1986.

Además se introdujo un índice dependiente de la escala para cuantificar el grado de eficiencia del mercado. Con excepción de dos períodos (a principios de los 90's y a finales del 2000, correspondientes a recesiones económicas en los EE.UU.), el índice de eficiencia es del 100 % ya que el patrón de entropía multiescala muestra un comportamiento similar al ruido no correlacionado en un conjunto de escalas temporales (τ).

Los resultados obtenidos en esta sección muestran que la complejidad y la eficiencia del mercado del petróleo crudo son dependientes de la escala de tiempo (τ). Lo anterior debe ser considerado para los diferentes participantes del mercado, desde inversionistas hasta los que buscan oportunidades de beneficio dentro de estrategias de inversión activas, así también para los gobiernos. Por ejemplo, durante el inicio de una recesión económica, el precio del petróleo crudo se vuelve más predecible ya que las empresas posponen inversiones a largo plazo en su búsqueda de una cierta direccionalidad de recuperación en la recesión (Bernanke y col., 1997). A su vez, la complejidad del mercado del petróleo crudo se incrementa y las fluctuaciones en los precios son cada vez más inciertas, lo que se refleja como incrementos en los valores de entropía a largo plazo.

4.5. Variaciones temporales de eventos sísmicos

En esta sección se utilizó el análisis R/S para estudiar las propiedades de reescalamiento fractal en términos del exponente de Hurst para datos sísmicos del Sur de México del período 1998-2011. El estudio se centra en esta región geográfica, ya que es una de las zonas sísmicas más activas en México. Los resultados obtenidos indican que al considerar sólo los eventos

sísmicos que cumplen con el requisito de la ley de Gutenberg-Richter ($b = 0.97$, $M_{GR} = 3.6$), se encuentran agrupaciones en el tiempo para escalas de alrededor de 100 y 135 eventos. De esta forma, los resultados revelan que el análisis de la actividad sísmica a través del análisis R/S puede proveer de conocimientos adicionales sobre la actividad relacionada con la aparición de grandes terremotos.

4.5.1. Introducción

Cuando se habla de terremotos, la importancia de la planificación de riesgos no sólo se debe considerar la magnitud e intensidad o grado de devastación de los eventos, sino también normas de edificación, planes de evacuación, estrategias de predicción, políticas de prevención y estándares urbanos. El objetivo del monitoreo de eventos sísmicos es entender los mecanismos geodinámicos, por ejemplo tectónicos relacionados con la subducción que son potencialmente desencadenantes de eventos sísmicos extremos.

El monitoreo de señales en regiones sísmicas incluyen anomalías electromagnéticas, señales de auto-potencial y actividad sísmica para amplios rangos de magnitud. En este contexto, el análisis de señales monitoreadas puede proporcionar información valiosa de características dinámicas de procesos sísmicos y mecanismos geofísicos. A su vez, los resultados derivados de análisis estadísticos confiables pueden mejorar el diseño de políticas de prevención, y de esta forma, mitigar los efectos adversos de eventos sísmicos.

Ambiente tectónico

La placa de Cocos es la formación tectónica principal en la región Sur de México que es subducida a lo largo de la fosa Mesoamericana (MAT , Middle America Trench) (Figura 4.23.a). En la Figura 4.23.b puede observarse una alta fracción de eventos sísmicos ocurridos en la región entre el MAT y la línea costera.

El movimiento de la placa de Cocos con respecto a la placa de América del Norte se dirige desde el norte hacia el noreste, un poco hacia la izquierda con una línea normal a la MAT . La tasa de convergencia aumenta hacia el este desde 4.8 cm/año en 104.5 W hasta 7.5

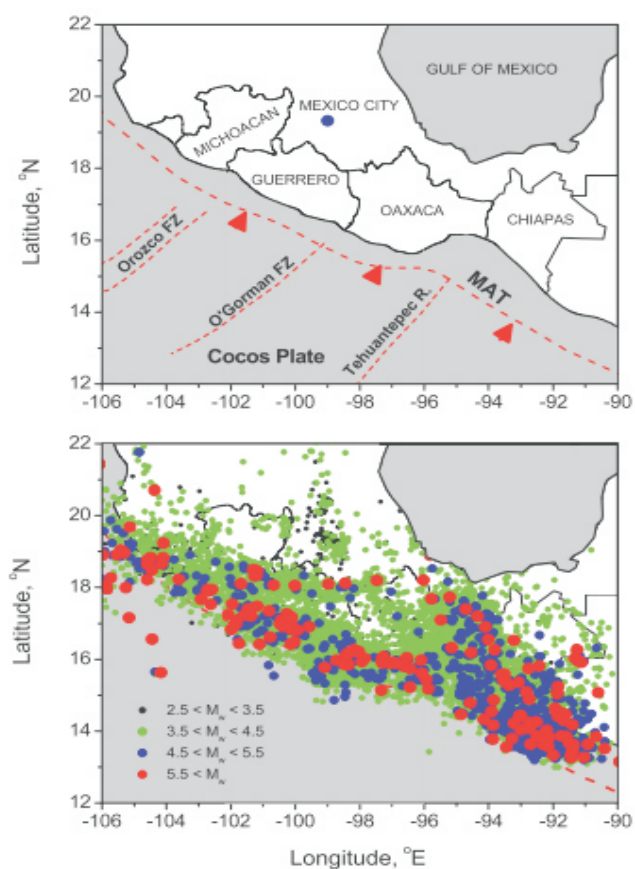


Figura 4.23. (a) Mapa esquemático y sectores de la deformación tectónica activa a lo largo de la costa del Sur del margen activo mexicano. (b) Sismicidad del período 1998-2011 de acuerdo con el catálogo del Servicio Sismológico Nacional, México.

cm/año en 94 W (DeMets y col., 1994). La edad de la placa de Cocos bajo subducción varía a lo largo de la *MAT*, con algunas discontinuidades a través de las zonas de fractura que se extienden hacia el este desde el flanco dorsal del Pacífico Oriental. Las más prominentes son las zonas de fracturas de Orozco y O'Gorman.

Sismicidad

En los últimos dos siglos a lo largo de la parte mexicana de la *MAT* se han producido grandes terremotos poco profundos ($M_w = 7.5$) con intervalos de recurrencia de 30-50 años en segmentos discretos de 100-200 km (Singh y col., 1981). Se han identificado brechas

sísmicas en donde grandes terremotos no se producen en largos períodos de tiempo (Astiz y Kanamori, 1984). La brecha sísmica de Guerrero, entre Zihuatanejo y Acapulco ($\sim 101^\circ$ Oeste), es una región bien conocida de alto riesgo sísmico debido a grandes eventos sísmicos que ocurrieron allí en el comienzo del siglo XX ($M_w = 7.5$, en 1911). Posteriormente, sólo se registró un terremoto importante ($M_w = 6.7$, 28 de Abril de 2002). En 1957 se produjeron terremotos significativos en la costa de Guerrero ($M_w = 7.0$ y 7.1 , Acapulco) y en 1962 en Costa Rica ($M_w = 7.8$). Esto sugiere que en la zona de Guerrero la subducción podría haber acumulado una gran cantidad de tensión en los últimos 50 años.

Un sector en la costa de Michoacán, donde la zona de fractura de Orozco se cruza con la *MAT* al sureste fue considerada en gran medida como una brecha sísmica hasta la ocurrencia de un gran terremoto ($M_w = 8.1$) ocurrido el 19 de Septiembre de 1985 llenó el vacío. La región está flanqueada en ambos lados por segmentos sísmicos activos, con la intersección entre la zona de fractura Orozco y la *MAT* al sureste de la laguna. Existe evidencia de que terremotos periódicos de gran magnitud forman elevaciones en las terrazas marinas (Bodin y Klinger, 1986).

El sector de Tehuantepec en la zona de subducción es también muy interesante porque ningún terremoto importante con empuje superficial ha sido registrado desde hace más de dos siglos. Se ha argumentado que esta zona es asísmica o tiene intervalos de recurrencia anormalmente largos para grandes terremotos (Astiz y Kanamori, 1984). La posibilidad de que la brecha de Tehuantepec puede ser asísmica para grandes terremotos se explica por la influencia de la subducción de la cresta de Tehuantepec, debido a que la brecha se encuentra cerca de la triple unión, convirtiéndose así en una zona de transición con respecto a la geometría en la zona de Benioff (Manea y col., 2005).

Ya que, la Ciudad de México fue afectada por el terremoto del 19 de septiembre de 1985 ($M_w = 8.1$), causando la muerte de aproximadamente 10,000 personas, y miles de millones de dólares en daños estructurales. Por consiguiente, esta región es altamente monitoreada y se han realizado muchos estudios en relación con sus características geodinámicas (Kostoglodov y Ponce, 1994; Zhang y col., 2009).

Antecedentes

Se han realizado esfuerzos importantes para la obtención de una caracterización sistemática de los procesos sísmicos en el siglo reciente. La clásica ley de frecuencia-magnitud de Gutenberg-Richter (Gutenberg y Richter, 1954) y la ley Omori para la descomposición temporal de la actividad de réplica (Omori, 1984; Utsu y col., 1995) se utilizan comúnmente para evaluar la distribución espacial de la tensión en un fallo o para estimar riesgos sísmicos después de fuertes sismos (Reasenberg y Jones, 1989; Gerstenberger y col., 2005). Un supuesto en las leyes de Gutenberg-Richter y Omori es que los procesos sísmicos muestran un comportamiento auto-similar descrito por las leyes de escala de potencia. Es decir, la sismicidad no tiene escala típica o dominante temporal o espacial, sino que exhibe una gran diversidad de escalas donde se expresa el fenómeno. En cierto sentido, una ley de escala de potencia de la sismicidad se convierte en un índice del grado de heterogeneidad tanto de las propiedades del proceso así como de la agrupación de la actividad sísmica. Importantes esfuerzos de investigación se han dedicado en los últimos años para hacer frente a las propiedades de reescalamiento de eventos sísmicos, dando lugar a una amplia aceptación de que la actividad sísmica no es en absoluto al azar, pero presenta fenómenos de agrupamiento en el tiempo y en el espacio (Hainzl, 2003). La idea que subyace detrás de los estudios de escala es que la actividad sísmica en alguna escala temporal mayor puede determinarse a partir de la que se produce a una escala menor. Las leyes de escala para la variabilidad temporal y espacial de los terremotos han sido obtenidas de diversas regiones sísmicas con diferentes propiedades tectónicas (Bak y col., 2002; Corral, 2004; Balankin y col., 2009) y los resultados han indicado que los períodos inter-eventos no siguen una distribución exponencial como en el proceso de Poisson, sino más bien una distribución de ley de potencia. A su vez, esto sugiere que los registros de recurrencia en eventos sísmicos no son independientes, sino que contienen correlaciones seriales a largo plazo.

Sin embargo, se debe considerar que la ley de escala de las funciones de distribución de probabilidad (*PDF*) sólo es estadísticamente de primer orden y no revela nada acerca de la estructura subyacente de correlaciones seriales. Es de particular importancia para la

evaluación de los patrones de sismicidad, la derivación de información sobre la estructura de correlación entre los bloques de períodos inter-eventos, que pueden ser relacionadas a agrupaciones temporales de eventos sísmicos que ocurren en una región geográfica dada. Dirigidos a abordar esta cuestión, se ha demostrado que los métodos tomados de la estadística no lineal, tales como el análisis R/S (Hurst y col., 1965) y el análisis de fluctuación sin tendencia (Peng y col., 1994) pueden proporcionar información importante acerca de la dinámica en eventos sísmicos (Telesca y col., 2001, 2007; Telesca y Lovallo, 2009).

Los resultados empíricos muestran que las secuencias de períodos inter-eventos son persistentes, lo que implica que la ocurrencia de eventos sísmicos no es aleatoria, sino regidos por mecanismos estocásticos con efectos en la memoria a largo plazo (Xu y Burton, 2006; Telesca y col., 2007). Otros estudios basados en el catálogo de sismicidad de California han indicado que la secuencia de eventos sísmicos no puede ser descrito por un modelo estadístico simple o único ya que el fenómeno muestra una variedad de comportamientos complejos (Telesca y col., 2004; Jiménez, 2011). Se ha explorado a su vez la estructura multifractal entre períodos inter-eventos y terremotos sucesivos, mostrando un incremento gradual de la multifractalidad ante actividades sísmicas importantes (Dimitriu y col., 2000); así como de una pérdida de multifractalidad durante réplicas (Telesca y Lapenna, 2006; Zamani y Agh-Atabai, 2009). Cabe destacar que las estimaciones de dimensiones generalizadas de las características de los períodos multifractales inter-eventos permiten una interpretación tectónica (Molchan y Kronrod, 2009). También se ha sugerido que la caracterización combinada de propiedades fractales de secuencias de terremotos y fallas permite una mejor detección del riesgo sísmico en una región geográfica determinada (Henares-Romero y col., 2010).

4.5.2. Metodología

Para el análisis R/S se obtuvieron datos de sismicidad del catálogo del Servicio Sismológico Nacional Mexicano para el período Enero de 1998 a Marzo de 2011. Las fronteras de la región examinada se delinear en la Figura 4.23.a ($\sim 90^\circ - 106^\circ$ Oeste y $\sim 12^\circ - 22^\circ$

Norte), que contiene los estados mexicanos a lo largo de la trinchera mesoamericana la cual forma parte de la placa de Cocos (Michoacán, Guerrero, Oaxaca y Chiapas). El catálogo de terremotos incluye más de 16,300 eventos y contiene información sobre la fecha, hora, ubicación (latitud y longitud), profundidad y magnitud para cada evento individual. Los eventos clasificados en términos de valores de magnitud se muestran en la Figura 4.23.b. Puede observarse que una fracción grande de eventos sísmicos se encuentran localizados dentro de la *MAT* y el borde costero. Observe la distribución espacial no-regular con una geometría de tipo fractal.

Para los estudios de sismicidad es necesario un catálogo de terremotos completo y homogéneo. Balankin y col. (2009) realizaron el cálculo de la magnitud mínima para que el catalogo sea considerado completo (M_{GR}). Se encontró que el catálogo sísmico obedece a la ley de Gutenberg-Richter $\log_{10}N(M_w > m) = a - bm$ ($b = 0.97$) para eventos sísmicos de magnitud $M_w > M_{GR} = 3.6$ (Figura 4.24). Donde puede observarse que un ajuste lineal es la mejor opción para la ley de Gutenberg-Richter ($b = 0.97$, $M_{GR} = 3.6$). De esta manera, sólo sismicidades con $M_w \geq M_{GR}$ se consideraron para el análisis R/S. Esto sugiere que algunos pequeños eventos con magnitud de $M_w < 3.6$ no fueron completamente registrados en el catálogo. Por lo que análisis de correlaciones de período inter-eventos se limitó a eventos sísmicos de magnitud $M_w > 3.6$. Para los casos en estudio, la eliminación de eventos sísmicos con $M_w = 3.6$ implica la eliminación de aproximadamente 15 % de los eventos sísmicos en promedio.

4.5.3. Resultados

Después de quitar los eventos sísmicos con $M_w = 3.6$, el exponente de Hurst fue calculado para ventanas deslizantes de tamaño $N_{rw} = 800$ eventos y con un deslizamiento $N_{rol} = 800$ eventos. Los resultados se muestran en la Figura 4.25 para el exponente de Hurst como una función del tiempo t (expresado en días naturales) y escala N_s (expresado en número de eventos). Un análisis más detallado de las distribuciones de magnitud de frecuencia de la Fig. 4.24 muestra que el valor M_{GR} puede ser mayor que 3.6. Es evidente que también se

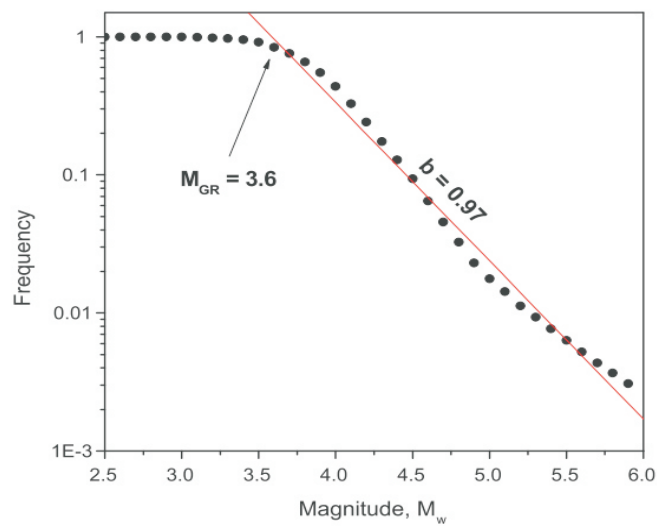


Figura 4.24. Diagrama de frecuencia-magnitud de la sismicidad de la Fig. 4.23.b.

pueden considerar los valores ≈ 3.7 ó 3.8 .

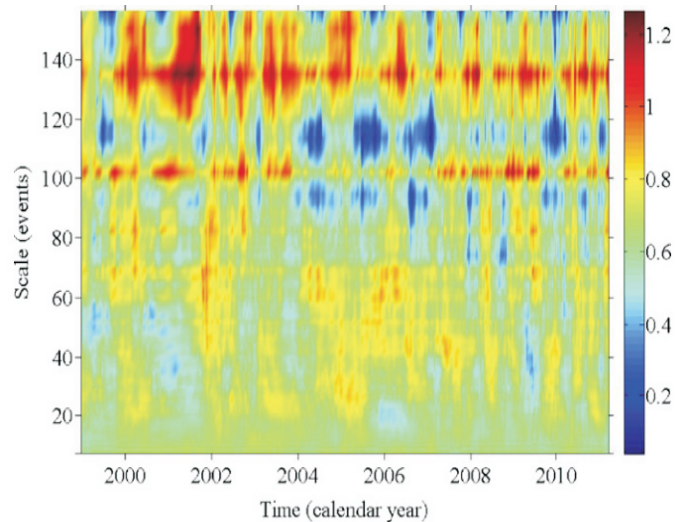


Figura 4.25. Comportamiento del exponente de Hurst como función del tiempo t y escala N_s .

En la Figura 4.25 se muestra la estructura de las variaciones temporales del exponente de Hurst, donde se pueden observar las siguientes características:

- En general, el exponente de Hurst es mayor a 0.5 para todos los tiempos y escalas. Esto indica que los eventos sísmicos en la región sur de México presentan dinámicas

persistentes. Esto es, como ocurre en distintas regiones sísmicas (Xu y Burton, 2006; Telesca y col., 2007), la secuencia de sismicidad no es aleatorio, sino que muestra agrupación en el tiempo. De este modo, un evento sísmico se acompaña de muchos otros eventos para los que el período inter-evento es similar ya sea más corto o más largo. La existencia de correlaciones fractales sugieren que toda la litosfera en la zona mexicana de subducción es un sistema auto-organizado en estado crítico, donde una fuerza está actuando simultáneamente en todas las placas para distribuir la energía sobre la carcasa de la litosfera.

- Se observan dos patrones coherentes para escalas de aproximadamente 100 y 135 eventos. Una vista detallada se muestra en la Figura 4.26, donde se observa que se alternan períodos de altos y bajos valores del exponente de Hurst. En estos períodos, el exponente de Hurst varía desde ~ 0.5 (fluctuaciones no correlacionadas) hasta ~ 1.0 (ruido $1/f$). Esto significa que las correlaciones de los eventos sísmicos no son uniformes, mostrando cambios significativos en el tiempo y escala. El origen de las bandas coherentes del exponente de Hurst alrededor de 100 y 153 eventos no está bien definido, pero bien podría estar relacionado con fenómenos de resonancia o con las ondas que viajan a lo largo de las inmediaciones del *MAT*.

4.5.4. Conclusiones

En esta sección se empleó el análisis R/S para cuantificar la fuerza de las correlaciones en las secuencias sísmicas de período inter-evento en el Sur de México. Los resultados confirman los hallazgos previos de otras regiones geográficas (Xu y Burton, 2006; Telesca y col., 2007) en el sentido de que la ocurrencia de eventos sísmicos no es aleatorio, sino que muestra el comportamiento de agrupación temporal. En particular, se encontró la existencia de dos escalas de resonancia en alrededor de 100 y 135 eventos donde la sismicidad exhibe elevadas correlaciones. Curiosamente, los períodos de bajas correlaciones de intereventos están ligados a la ausencia de terremotos importantes. Esto es, la mayor correlación de períodos de intereventos ocurrió con terremotos de mayor riesgo de alta magnitud. Los

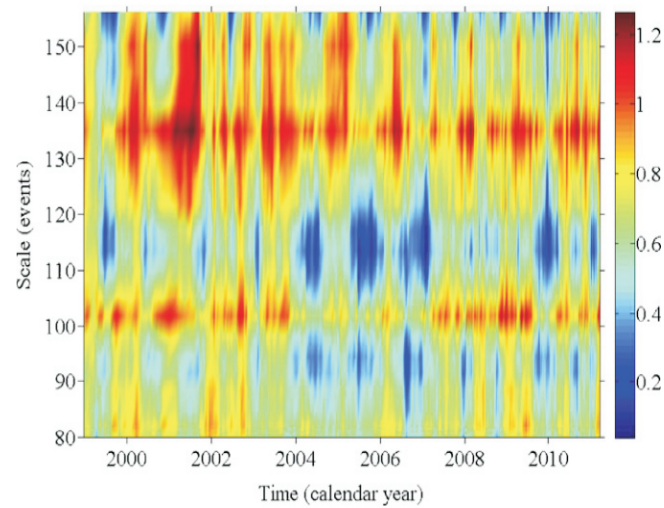


Figura 4.26. Detalles de los patrones del exponente de Hurst mostrados en la Fig.4.25.

resultados muestran la capacidad del análisis de Hurst R/S para revelar aspectos importantes sobre la dinámica de sismicidad que pueden ser útiles para el seguimiento de la probabilidad de ocurrencia de grandes terremotos en la región Sur de México.

CAPÍTULO 5

CONCLUSIONES

En este capítulo se presentan las conclusiones derivadas del presente trabajo de investigación. Para cumplir el objetivo principal del presente trabajo se tomo de la literatura especializada estimadores estadísticos que permitieran un análisis de complejidad de series de tiempo que fueran exitosos en temas relevantes a la ingeniería química. Estos estadísticos son la entropía aproximada desarrollada por Pincus (1991) y el exponente de Hurst nombrado así en honor de Hurst (1951). Estos estimadores estadísticos se seleccionaron debido a que los resultados obtenidos con ellos pueden interpretarse en forma física y por lo tanto existen límites naturales para poder interpretar los resultados.

A continuación se presentan las conclusiones derivadas de cada uno de estos. Para presentarlas en una forma más ordenada han sido separadas en dos secciones. La primera sección corresponde a las conclusiones derivadas de los casos de estudio en los que se emplearon metodologías de análisis mediante el estadístico entropía aproximada y la segunda sección para las conclusiones derivadas de los casos de estudio correspondientes al estadístico exponente de Hurst.

5.1. Entropía Aproximada

En el Capítulo 4 se presentan los resultados del análisis de complejidad relacionados con señales sintéticas del tipo $1/f^\beta$, estudio del cual podemos concluir que las señales que deben tener los máximos valores de entropía son aquellas pertenecientes a las de ruido blanco, lo anterior incluso para cualquier escala de tiempo; es decir, las señales de ruido blanco están relacionadas con una falta de orden o patrones de estructura en la señal. Por otra parte, las señales que presentan los mayores patrones de estructura a cualquier escala de tiempo son las pertenecientes a las del movimiento browniano. Estos resultados proporcionan límites teóricos reales que deben tenerse en cuenta al interpretar resultados derivados de los análisis de complejidad de series de tiempo mediante el estadístico entropía aproximada. Trabajos futuros deberán explorar mayores detalles sobre las señales sintéticas. Es decir, si bien se puede demostrar que los valores de entropía pertenecientes al movimiento browniano están relacionados a patrones de estructura existentes en la señal, también deberían estudiarse sus efectos para poder asegurar la aleatoriedad en las señales de ruido blanco.

En este mismo Capítulo 4 y derivado del estudio de las señales sintéticas y obtención de los resultados del análisis de complejidad de series de tiempo de fluctuaciones de ritmos cardíacos obtenidos por Costa y col. (2002), se concluye que los resultados más estables se obtienen al sintonizar los parámetros del estadístico AE en la forma propuesta por Pincus (1991). De esta forma es posible comparar resultados contra otros presentados en la literatura especializada, por lo que en la actualidad es algo normal emplear los parámetros según lo propone Pincus (1991); sin embargo, trabajos futuros deberán explorar el efecto de sintonizar los parámetros mediante diferentes esquemas. Este tipo de trabajos ya se ha empezado a estudiar, por ejemplo, trabajos actuales están interesados en establecer un equilibrio entre la robustez del método y la rapidez de cálculo al variar la forma de sintonizar los parámetros del estadístico AE .

En este mismo caso de estudio se observa la importancia de remover la tendencia presente

en la serie de tiempo previo al análisis de complejidad en la serie de tiempo, de otra forma los resultados varían y puede presentarse el caso de no estar comparando resultados en forma apropiada. También se puede observar la ventaja de extender el análisis a multiescala donde es posible observar que se revelan mayores detalles de la señal en otras escalas de tiempo. Por ejemplo, en el caso de estudio presente en el Capítulo 4 relacionado con el análisis de complejidad de series de tiempo de fluctuaciones de ritmos cardíacos, las distintas series de tiempo pueden ser discriminadas en forma más sencilla al analizar el comportamiento de la serie a diversas escalas de tiempo. Para el caso de estudio se observaron tres distintos comportamientos para las series de tiempo de acuerdo al padecimiento de cada paciente con una determinada condición de estado de salud.

Con estos resultados en consideración en el Capítulo 4 se presenta el estudio de señales de intensidad de difracción de rayos X , un tema relacionado con la ingeniería química. Y se observa que el estimador estadístico tiene éxito de aplicación. En este caso la metodología de estudio varía, el estudio no se lleva a cabo en la señal original sino en una señal de entropía generada a partir de la señal original. El análisis es multiescala y sólo se presentan los valores promedios para cada escala de tiempo con el fin de generar una sola señal de entropías para analizar. Es directamente sobre esta última señal donde se interpretan los resultados. Para este caso, es interesante observar que los valores máximos en la señal coinciden con regiones angulares asociadas con fases del material. Diversas metodologías pueden generarse según convenga, pero siempre debe tenerse en cuenta interpretar los resultados bajo las consideraciones resultantes de los casos de estudio previos en el Capítulo 4.

Finalmente se presenta la metodología del análisis de complejidad con el esquema de ventana deslizante, la cual resulta importante cuando se tenga interés con cambios repentinos o graduales en las series de tiempo por analizar. En el Capítulo 4 se presenta un caso de estudio bajo este esquema, el estudio de la eficiencia de mercados de petróleo crudo. En este caso se tiene interés en conocer la robustez del mercado del petróleo en términos de

previsibilidad, en este caso se desea saber si los cambios súbitos actuales o graduales en el tiempo han robustecido al mercado o lo han debilitado y si es posible predecir el mercado del petróleo. Para este caso una metodología con un esquema de una ventana deslizante resulta conveniente porque no se está interesado en un valor puntual de la señal sino en los cambios que ocurren en la series de tiempo a distintas escalas de tiempo.

5.2. Exponente de Hurst

Para este parámetro (o estimador) estadístico en el Capítulo 4 se presentan dos casos de estudio, uno que estudia las variaciones temporales de eventos sísmicos y otro que estudia las señales de intensidad de difracción de rayos X . El primero de estos casos de estudio se lleva a cabo con el fin de probar y sintonizar el método con resultados de la literatura especializada, mientras que el segundo ejemplifica su aplicación a un tema relacionado con la ingeniería química.

En el Capítulo 4 se presenta un caso de estudio relacionado con las variaciones temporales de eventos sísmicos, se demuestra que el análisis R/S resulta importante al cuantificar la fuerza de las correlaciones en las secuencias sísmicas de períodos interevento del Sur de México. Los resultados muestran la capacidad del análisis de Hurst (R/S) para revelar aspectos importantes sobre la dinámica de sismicidad que pueden ser útiles para el seguimiento de la probabilidad de ocurrencia de grandes terremotos en el Sur de la región de México. Resultados similares se han encontrado en otras regiones geográficas de la Tierra.

Finalmente en el Capítulo 4 se presenta el caso de estudio de las señales de intensidad de difracción de rayos X ahora analizado mediante el estimador estadístico del exponente de Hurst (R/S). Se demuestra que el análisis R/S puede ser empleado para encontrar regiones angulares asociadas con fases de un material para distintas señales de difracción de rayos X . Cuando se comparan los resultados mediante el exponente de Hurst (R/S) y el de entropía aproximada (AE) se puede observar que los resultados son complementarios y no excluyentes. Por lo tanto, al llevar a cabo un análisis de complejidad en series de tiempo de sistemas complejos no existe una metodología superior y se recomienda examinar

los resultados obtenidos por las dos metodologías y seleccionar la que mejor convenga para interpretar los resultados.

Se pueden desarrollar trabajos futuros para estudiar la robustez de la metodología del exponente de Hurst (R/S) frente a nuevos estimadores estadísticos que calculan el exponente, como lo es la reciente metodología del análisis de las fluctuaciones con remoción de la tendencia (DFA).

Administration, U.S. Energy Information. “www.eia.gov”.

Alvarez-Ramirez, J. y Rodriguez, E., 2011. “Long-term recurrence patterns in the late 2000 economic crisis: Evidences from entropy analysis of the Dow Jones index”, *Technological Forecasting and Social Change*, Vol. 78, No. 8, 1332–1344.

Alvarez-Ramirez, J., Alvarez, J., y Rodriguez, E., 2008a. “Short-term predictability of crude oil markets: A detrended fluctuation analysis approach”, *Energy Economics*, Vol. 30, 2645–2656.

Alvarez-Ramirez, J., Calderon, J., Rodriguez, E., y Fernandez-Anaya, G., 2008b. “Time-varying Hurst exponent for US stock markets”, *Physica A*, Vol. 387, 6159–6169.

Alvarez-Ramirez, J., Rodriguez, E., y Echeverria, J. C., 2009. “Delays in the human heartbeat dynamics”, *Chaos*, Vol. 19, 028502.

Alvarez-Ramirez, J., Alvarez, J., y Solis, R., 2010. “Crude oil market efficiency and modeling: Insights from the multiscaling autocorrelation pattern”, *Energy Economics*, Vol. 32, 993–1000.

Alvarez-Ramirez, J., Rodriguez, E., y Alvarez, J., 2012. “A multiscale entropy approach for market efficiency”, *International Review of Financial Analysis*, Vol. 21, 64–69.

-
- Arouri, E.H.M., Dinh, T.H., y Nguyen, D.K., 2010. “Time-varying predictability in crude-oil markets: the case of GCC countries”, *Energy Policy*, Vol. 38, 4371–4380.
- Astiz, L. y Kanamori, H., 1984. “An earthquake doublet in Ometepec, Guerrero, Mexico”, *Phys. Earth Planet. Int.*, Vol. 34, 24–45.
- Bak, P., Christensen, K., Danon, L., y Scanlon, T., 2002. “Unified scaling law for earthquakes”, *Phys. Rev. Lett.*, Vol. 88, 178501.
- Balankin, A.S., Morales-Matamoros, D., Patiño-Ortiz, J., Patiño-Ortiz, M., Pineda-León, E., y Samayoa-Ochoa, D., 2009. “Scaling dynamics of seismic activity fluctuations”, *Europhys Lett.*, Vol. 85, 39001.
- Ball, M. y Wood, A., 1996. “Trend growth in post-1850 british economic history: the kalman filter and historical judgment”, *Journal of the Royal Statistical Society. Series D (The Statistician)*, Vol. 45, No. 2, 143–152.
- Barsky, R.B. y Kilian, L., 2004. “Oil and the macroeconomy since the 1970s”, *Journal of Economic Perspectives*, Vol. 18, No. 4, 115–134.
- Bernanke, B.S., Gertler, M., y Watson, M., 1997. “Systematic monetary policy and the effects of oil price shocks”, *Brookings Papers on Economic Activity*, Vol. 1997, 91–157.
- Bodin, P. y Klinger, T., 1986. “Coastal uplift and mortality of intertidal organisms caused by the September 1985 Mexico earthquakes”, *Science*, Vol. 233, 1071–1073.
- Box, G. E. P. y Jenkins, G. M., 1976. *Time series analysis, forecasting and control*. San Francisco: Holden-day, 1st edición.
- Box, G. E. P., Jenkins, G. M., y Reinsel, G. C., 1994. *Time series analysis: forecasting and control*. Prentice Hall, Englewood Cliffs, NJ., 3rd edición.
- Campbell, J.Y., Lo, A.W., y MacKinlay, A.C., 1997. “The Econometrics of Financial Markets”, *Princeton University Press, Princeton*.

-
- Carruth, A.A., Hooker, M.A., y Oswald, A.J., 1998. “Unemployment equilibria and input prices: theory and evidence from the United States”, *The Review of Economics and Statistics*, Vol. 80, 621–628.
- Chamoli, Ashutosh, Bansal, Abhey Ram, y Dimri, V.P., 2007. “Wavelet and rescaled range approach for the Hurst coefficient for short and long time series”, *Computers & Geosciences*, Vol. 33, No. 1, 83–93.
- Charles, A. y Darné, O., 2009. “The efficiency of the crude oil markets: evidence from variance ratio tests”, *Energy Policy*, Vol. 37, 4267–4272.
- Ching-Chih, C. y Tin-Chia, L., 2011. “The nonlinear dynamic process of macroeconomic development by modelling dry bulk shipping market”, *Applied Economics Letters*, Vol. 18, No. 17, 1655–1663.
- Corral, A., 2004. “Long-term clustering, scaling, and universality in the temporal occurrence of earthquakes”, *Phys.Rev.Lett.*, Vol. 92, 108501.
- Costa, M., Goldberger, A. L., y Peng, C.-K., Jul 2002. “Multiscale Entropy Analysis of Complex Physiologic Time Series”, *Phys. Rev. Lett.*, Vol. 89, No. 6, 068102.
- Costa, M., Goldberger, A. L., y Peng, C.-K., Feb 2005. “Multiscale entropy analysis of biological signals”, *Phys. Rev. E*, Vol. 71, No. 2, 021906.
- Costa, M., Goldberger, A. L., y C.-K.Peng, January 2008. “Multiscale Analysis of Heart Rate Dynamics: Entropy and Time Irreversibility Measures”, *Cardiovasc Eng.*, Vol. 8, No. 2, 88–93.
- Cox, D. R. y Isham, V., 1980. *Point processes*. Chapman & Hall.
- Cuñado, J. y Perez deGracia, F., 2003. “Do oil price shocks matter? Evidence from some European countries”, *Energy Economics*, Vol. 25, 137–154.
- Darbellay, G.A. y Wuertz, D., 2000. “The entropy as a tool for analysing statistical dependences in financial time series”, *Physica A*, Vol. 287, 429–439.

-
- Day-Lewis, A., Zoback, M., y Hickman, S., 2010. “Scale-invariant stress orientations and seismicity rates near the San Andreas Fault”, *Geophys.Res.Lett.*, Vol. 37, L24304.
- DeMets, C., Gordon, R.G., Argus, D.F., y Stein, S., 1994. “Effect of recent revisions to the geomagnetic reversal time scale on estimates of current plate motions”, *Geophys.Res.Lett.*, Vol. 21, 2191–2194.
- Dimitriu, P.P., Scordilis, E.M., y Karacostas, V.G., 2000. “Multifractal analysis of the Arnea,Greece seismicity with potential implications for earthquake prediction”, *Nat. Hazards*, Vol. 21, 277–295.
- Dorfman, J. R., 1999. *An introduction to chaos in nonequilibrium statistical mechanics*. Cambridge University Press.
- Elder, J. y Serletis, A., 2008. “Long memory in energy futures prices”, *Review of Financial Economics*, Vol. 17, 146–155.
- Elder, T., Labbé, N., Harper, D., y Rials, T., 2006. “Time domain-nuclear magnetic resonance study of chars from southern hardwoods”, *Biomass and Bioenergy*, Vol. 30, No. 10, 855 – 862.
- Fama, E.F., 1970. “Efficient capital markets: a review of theory and empirical work”, *The Journal of Finance*, Vol. 25, 383–417.
- Fama, E.F., 1991. “Efficient capital markets II”, *Journal of Finance*, Vol. 46, 1575–1617.
- Fan, L. T., Kang, Y., Neogi, D., y Yashima, M., 1993. “Fractal analysis of fluidized particle behavior in liquid-solid fluidized beds”, *AIChE Journal*, Vol. 39, No. 3, 513–517.
- Fan, Y., Liang, Q., y Wei, Y.-M., 2008. “A generalized pattern matching approach for multistep prediction of crude oil prices”, *Energy Economics*, Vol. 30, 889–904.
- Fattouh, B., 2010. “The dynamics of crude oil price differentials”, *Energy Economics*, Vol. 32, 334–342.

-
- Feder, J., 1988. *Fractals*. New York: Plenum Press.
- Franses, P. H., 1998. *Time series models for business and economic forecasting*. Cambridge: Cambridge Univ. Press.
- Franses, P. H. y Kleibergen, F., 1996. “Unit roots in the Nelson-Plosser data: Do they matter for forecasting?”, *Int. J. Forecasting*, Vol. 12, 283–288.
- Frisch, R., 1930. “Necessary and sufficient conditions regarding the form of an index number which shall meet certain of Fisher’s test”, *Journal of the American Statistical Association*, Vol. 25, 397–406.
- Gates, B. C., Katzer, J. R., y Schuit, G.C.A., 1979. *Chemistry of catalytic processes*. McGraw-Hill, New York.
- Gerstenberger, M.C., Wiemer, S., Jones, L.M., y Reasenber, P.A., 2005. “Real-time forecasts of tomorrows earthquakes in California”, *Nature*, Vol. 435, 328–331.
- Giglio, R., Matsushita, R., y Da Silva, S., 2008. “The relative efficiency of stock markets”, *Economics Bulletin*, Vol. 7, 1–12.
- Gisser, M. y Goodwin, T.H., 1986. “Crude oil and the macroeconomy: tests of some popular notions”, *Journal of Money, Credit and Banking*, Vol. 18, 95–103.
- Goldberger, A. L., Amaral, L. A. N., Hausdorff, J. M., Ivanov, P. Ch., Peng, C.-K., y Stanley, H. E., 2002. “Fractal dynamics in physiology: Alterations with disease and aging”, *PNAS*, Vol. 99, 2466–2472.
- Guillen, P., Paredes, M., y Camacho, O., 2005. “A proposal method for fault detection and diagnosis in chemical processes instrumentation using wavelet transform”, *Revista Técnica de la Facultad de Ingeniería. Universidad del Zulia*, Vol. 28, No. 1, 68 – 77.
- Gulko, L., 1999. “The entropic market hypothesis”, *International Journal of Theoretical and Applied Finance*, Vol. 2, 293–329.

-
- Gutenberg, B. y Richter, C.F., 1954. *Seismicity of the Earth and Associated Phenomena*. Princeton University Press, Princeton, NJ., 2nd edición.
- Hainzl, S., 2003. “Self-organization of earthquake swarms”, *J. Geodynamics*, Vol. 35, 157–172.
- Hamilton, J.D., 1983. “Oil and macroeconomy since World War II”, *The Journal of Political Economy*, Vol. 91, No. 2, 228–248.
- Hamilton, J.D., 2003. “What is an oil shock?”, *Journal of Econometrics*, Vol. 113, No. 2, 363–398.
- Hassan, M., Terrien, J., Marque, C., y Karlsson, B., 2011. “Comparison between approximate entropy, correntropy and time reversibility: Application to uterine electromyogram signals”, *Medical Engineering & Physics*, Vol. 33, No. 980-986.
- Henares-Romero, J., López-Casado, C., Badal, J., y Peláez, J.A., 2010. “Seismicity pattern of the Betic Cordillera (Southern Spain) derived from the fractal properties of earthquakes and faults”, *Earthquake Sci.*, Vol. 23, 309–323.
- Hurst, H. E., 1951. “Long-term storage capacity of reservoirs”, *Trans. Amer. Soc. Civil Eng.*, Vol. 116, 770–808.
- Hurst, H.E., Black, R.P., y Simaika, Y.M., 1965. “Long-Term Storage: An Experimental Study”, *Constable, London*.
- Iparraguirre-D’elia, J.L., 2011. “Economic cycle fluctuations in Colombia. A comparative analysis according to univariate methods”, *Semest. Econ.*, Vol. 14, 61 – 85.
- Jammazi, R. y Aloui, Ch., 2012. “Crude oil price forecasting: Experimental evidence from wavelet decomposition and neural network modelling”, *Energy Economics*, Vol. 34, 828–841.
- Jiménez, A., 2011. “Comparison of the Hurst and DEA exponents between the catalogue and its clusters: the California case.”, *Physica A*, Vol. 390, 2146–2154.

-
- Jáuregui-Correa, J.C. y Gonzalez-Brambila, O.M., 2010. *Mechanical Vibrations of Discontinuous Systems*. Nova Science Publishers Inc., 1 edición.
- Kaffashi, F., Foglyano, R., Wilson, C.G., y Loparo, K.A., 2008. “The effect of time delay on approximate & sample entropy calculations”, *Physica D*, Vol. 237, 3069–3074.
- Kendall, M. G., Stuart, A., y Ord, J. K., 1983. *The advanced theory of statistics*, volumen 3. London: Griffin, 4th edición.
- Kenny, P. B. y Durbin, J., 1982. “Local trend estimation and seasonal adjustment of economic and social time series”, *Journal of the Royal Statistical Society. Series A*, Vol. 145, No. 1, 1–41.
- Kostoglodov, V. y Ponce, L., 1994. “Relationship between subduction and seismicity in the Mexican part of the Middle America trench”, *J. Geophys. Res.*, Vol. 99, 729–742.
- Lean, H.H., McAleer, M., y Wong, W.-K., 2010. “Market efficiency of oil spot and futures: A mean-variance and stochastic dominance approach”, *Energy Economic*, Vol. 32, 979–986.
- Lim, K.-P. y Brooks, R., 2011. “The evolution of stock market efficiency over time: A survey of the empirical literature”, *Journal of Economic Surveys*, Vol. 25, 69–108.
- Liu, Q., Wang, A., Wang, X., Gao, P., Wang, X., y Zhang, T., 2008. “Synthesis, characterization and catalytic applications of mesoporous gamma-alumina from boehmite sol.”, *Microporous and Mesoporous Materials*, Vol. 111, 323–333.
- Lo, A.W., 1991. “Long-term memory in stock market prices”, *Econometrica*, Vol. 59, 1279–1314.
- Lo, A.W., 2004. “The adaptive markets hypothesis: market efficiency from an evolutionary perspective”, *Journal of Portfolio Management*, Vol. 30, 15–29.
- Mandelbrot, B., 1985. “Self-Affine Fractals and Fractal Dimension”, *Physica Scripta*, Vol. 32, 257–260.

-
- Manea, M., Manea, V.C., Ferrari, L., Kostoglodov, V., y Bandy, W.L., 2005. “Tectonic evolution of the Tehuantepec ridge”, *Earth Planet.Sci.Lett.*, Vol. 238, 64–77.
- Martina, E., Rodriguez, E., Escarela-Perez, R., y Alvarez-Ramirez, J., 2011. “Multiscale entropy analysis of crude oil price dynamics”, *Energy Economics*, Vol. 33, 936–947.
- Martini, A.C., Stuz, G., Vincenti, L.M., Santillán, M.E., Carlini, V., Rubiales deBarioglio, S., y Fiol deCuneo, M., 2004. “Evaluación de parámetros morfométricos, reproductivos y conductuales en ratones macho hiponutridos”, *V Congreso de la FASEN. Neuroendocrinología*, Vol. 41, 66.
- Maslyuk, S. y Smyth, R., 2008. “Unit root properties of crude oil spot and futures prices”, *Energy Policy*, Vol. 36, 2591–2600.
- Meade, N., 1984. “The use of growth curves in forecasting market development a review and appraisal”, *Journal of Forecasting*, Vol. 3, No. 4, 429–451.
- Merino, A. y Ortiz, A., 2005. “Explaining the so-called “price premium” in oil markets”, *OPEC Review*, Vol. 29, 133–152.
- Mietus, J. E., Peng, C. K., Henry, I., Goldsmith, R.L., y Goldberger, A. L., 2002. “The pNNx files: re-examining a widely used heart rate variability measure”, *Heart*, Vol. 88, 378–380.
- Molchan, G. y Kronrod, T., 2009. “The fractal description of seismicity”, *Geophys.J.Int.*, Vol. 179, 1787–1799.
- Montgomery, D. F., 1996. *Introduction to Statistical Quality Control*. John Wiley and Sons.
- Morari, M. y Zafiriou, E., 1989. *Robust Process Control*. Prentice-Hall.
- Mork, K.A., 1989. “Oil and the macroeconomy when prices go up and down: an extension of Hamiltons results”, *The Journal of Political Economy*, Vol. 97, 740–744.

- Oh, G., Kim, S., y Eom, C., 2007. “Market efficiency in foreign exchange markets”, *Physica A*, Vol. 382, 209–212.
- Oladosu, G., 2009. “Identifying the oil price-macroeconomy relationship: an empirical mode decomposition analysis of US data”, *Energy Policy*, Vol. 37, No. 12, 5417–5426.
- Omori, F., 1984. “On aftershocks of earthquakes”, *J. College Sci. Imperial Univ. Tokyo*, Vol. 7, 111–200.
- Opong, K.K., Mulholland, G., Fox, A.F., y Farahmand, K., 1999. “The behaviour of some UK equity indices: An application of Hurst and BDS tests”, *Journal of Empirical Finance*, Vol. 6, No. 3, 267–282.
- Pan, Y.-H., Wang, Y.-H., Liang, S.-F., y Lee, K.-T., 2011. “Fast computation of sample entropy and approximate entropy in biomedicine”, *Computer Methods and Programs in Biomedicine*.
- Peng, C.-K., Buldyrev, S.V., Havlin, S., Simons, M., Stanley, H.E., y Goldberger, A.L., 1994. “Mosaic organization of DNA nucleotides”, *Phys. Rev. E*, Vol. 49, 1685–1689.
- PhysioNet. “www.physionet.org”.
- Pincus, S. M., 1991. “Approximate entropy as a measure of system complexity”, *Proc. Natl. Acad. Sci. USA*, Vol. 88, No. 6, 2297–2301.
- Pincus, S. M., October 2006. “Approximate entropy as a measure of irregularity for psychiatric serial metrics”, *Bipolar Disorders*, Vol. 8, No. 5, 430–440.
- Pincus, S. M., 2008. “Approximate Entropy as an Irregularity Measure for Financial Data”, *Econometric Reviews*, Vol. 27, No. 4-6, 329–362.
- Pincus, S. M. y Kalman, R. E., 2004. “Irregularity, volatility, risk, and financial market time series”, *Proc. Natl. Acad. Sci. USA*, Vol. 101, No. 38.

-
- Pérez-Canales, D., ÁlvarezRamírez, J., Jáuregui-Correa, J.C., Vela-Martínez, L., y Herrera-Ruiz, G., 2011. “Identification of dynamic instabilities in machining process using the approximate entropy method”, *International Journal of Machine Tools and Manufacture*, Vol. 51, No. 6, 556–564.
- Rasche, R.H. y Tatom, J., 1977. “Energy resources and potential GNP”, *Federal Reserve Bank of St. Louis Review*, Vol. 59, 10–24.
- Reasenber, P.A. y Jones, L.M., 1989. “Earthquake hazard after a mainshock in California”, *Science*, Vol. 243, 1173–1176.
- Richman, J. S. y Moorman, J. R., 2000. “Physiological timeseries analysis using approximate entropy and sample entropy”, *AJP-Heart*, Vol. 278, No. 6, H2039–H2049.
- Risso, W.A., 2008. “The informational efficiency and the financial crashes”, *Research in International Business and Finance*, Vol. 22, 396–408.
- Risso, W.A., 2009. “The informational efficiency: the emerging markets versus the developed markets”, *Applied Economics Letters*, Vol. 16, 485–487.
- Rotemberg, J.J. y Woodford, M., 1996. “Imperfect competition and the effects of energy price increases on economic activity”, *Journal of Money, Credit and Banking*, Vol. 28, 549–577.
- Santini, D.J., 1985. “The energy-squeeze model: energy price dynamics in US business cycles”, *International Journal of Energy Systems*, Vol. 5, 18–25.
- Serletis, A. y Andreadis, I., 2004. “Random fractal structures in North American energy markets”, *Energy Economics*, Vol. 26, 389–399.
- Shannon, C. E., 1948. “A Mathematical Theory of communication”, *The Bell System Technical Journal*, Vol. 27, 379–423.

- Singh, S.K., Astiz, L., y Havskov, J., 1981. “Seismic gaps and recurrence periods of large earthquakes along the Mexican subduction zone: a reexamination”, *Seismological Society of America*, Vol. 71, 827–843.
- Tabak, B.M. y Cajueiro, D.O., 2007. “Are the crude oil markets becoming weakly efficient over time? A test for time-varying long-range dependence in prices and volatility”, *Energy Economics*, Vol. 29, 28–36.
- Telesca, L. y Lapenna, V., 2006. “Measuring multifractality in seismic sequences”, *Tectonophysics*, Vol. 423, 115–123.
- Telesca, L. y Lovallo, M., 2009. “Non-uniform scaling features in central Italy seismicity: non-linear approach in investigating seismic patterns and detection of possible earthquake precursors”, *Geophys. Res. Lett.*, Vol. 36, L01308.
- Telesca, L., Cuomo, V., Lapenna, V., y Macchiato, M., 2001. “Identifying space-time clustering properties of the 1983-1997 Irpinia-Basilicata (Southern Italy) seismicity”, *Tectonophysics*, Vol. 330, 93–102.
- Telesca, L., Cuomo, V., Lapenna, V., y Macchiato, M., 2004. “Detrended fluctuation analysis of the spatial variability of the temporal distribution of Southern California seismicity”, *Chaos Solitons Fractals*, Vol. 21, 335–342.
- Telesca, L., Lovallo, M., Lapenna, V., y Macchiato, M., 2007. “Long-range correlations in two-dimensional spatio-temporal seismic fluctuations”, *Physica A*, Vol. 377, 279–284.
- Utsu, T., Ogata, Y., y Mutsaers, R.S., 1995. “The centenary of the Omori formula for a decay law of aftershock activity”, *Phys. Earth*, Vol. 43, 1–33.
- Varotsos, P. A., Sarlis, N. V., y Skordas, E. S., 2002. “Long-range correlations in the electric signals that precede rupture”, *Phys. Rev. E*, Vol. 66, No. 1, 011902.

-
- Vela-Martínez, L., Jáuregui-Correa, J.C., y ÁlvarezRamírez, J., 2009. “Characterization of machining chattering dynamics: An R/S scaling analysis approach”, *International Journal of Machine Tools and Manufacture*, Vol. 48, No. 11, 832–842.
- Wang, J. A., Bokhimi, X., Morales, A., Novaro, O., Lopez, T., y Gomez, R., 1999. “Aluminum local environment and defects in the crystalline structure of sol-gel alumina catalyst”, *The Journal of Physical Chemistry B*, Vol. 103, No. 2, 299–303.
- Wang, T. y Yang, J., 2010. “Nonlinearity and intraday efficiency tests on energy futures markets”, *Energy Economics*, Vol. 32, No. 2, 496–503.
- Wang, Y. y Liu, L., 2010. “Is WTI crude oil market becoming weakly efficient over time?: New evidence from multiscale analysis based on detrended fluctuation analysis”, *Energy Economics*, Vol. 32, 987–992.
- Wang, Y.Z., Li, B., Wang, R.Q., Su, J., y Rong, X.X., 2011. “Application of the Hurst exponent in ecology”, *Computers and Mathematics with Applications*, Vol. 61, 2129–2131.
- Wikipedia. “http://es.wikipedia.org/wiki/Teorema_de_muestreo_de_Nyquist-Shannon”.
- Xu, Y. y Burton, P.W., 2006. “Time varying seismicity in Greece: Hursts analysis and Monte Carlo simulation applied to a new earthquake catalogue for Greece”, *Tectonophysics*, Vol. 423, 125–136.
- Yan, R. y Gao, R. X., 2007. “Approximate Entropy as a diagnostic tool for machine health monitoring”, *Mechanical Systems and Signal Processing*, Vol. 21, No. 824-839.
- Zamani, A. y Agh-Atabai, M., 2009. “Temporal characteristics of seismicity in the Alborz and Zagros regions of Iran, using a multifractal approach”, *J. Geodynamics*, Vol. 47, 271–279.
- Zhang, C., Lomnitz, C., Shi, Y., y Ma, L., 2009. “Triggering of the 2001-2002 Mexico slow-slip event, M_w 7.5, by the shallow normal earthquake of 8 October 2001, M_w 5.8 and its aftershocks,”, *Seismol. Res. Lett.*, Vol. 80, 108–118.

Zhongying, Z., Yulin, W., Guangdi, L., y Xiang, S., 2011. “An rescaled range analysis on the characteristics of coal seam development in the Eastern depression of the Liaohe Basin”, *Mining Science and Technology (China)*, Vol. 21, No. 2, 223–227.



Casa abierta al tiempo

UNIVERSIDAD AUTÓNOMA METROPOLITANA

Fecha : 27/01/2014

Página : 1/1

CONSTANCIA DE PRESENTACION DE EXAMEN DE GRADO

La Universidad Autónoma Metropolitana extiende la presente CONSTANCIA DE PRESENTACION DE DISERTACIÓN PÚBLICA de DOCTOR EN CIENCIAS (INGENIERIA QUIMICA) del alumno JESUS ALEJANDRO ORTIZ CRUZ, matrícula 209382053, quien cumplió con los 348 créditos correspondientes a las unidades de enseñanza aprendizaje del plan de estudio. Con fecha treinta y uno de enero del 2014 presentó la DEFENSA de su DISERTACIÓN PÚBLICA cuya denominación es:

TECNICAS ESTADISTICAS PARA EL ANALISIS DE COMPLEJIDAD DE SERIES DE TIEMPO

Cabe mencionar que la aprobación tiene un valor de 180 créditos y el programa consta de 528 créditos.

El jurado del examen ha tenido a bien otorgarle la calificación de:

Aprubar

JURADO

Presidente

[Signature]

DR. JUAN CARLOS ECHEVERRIA ARJONILLA

Vocal

Secretario

[Signature]

DR. EDUARDO RODRIGUEZ FLORES

Vocal

[Signature]

DR. JESUS ISIDRO GONZALEZ TREJO

Vocal

[Signature]

DR. RICARDO AGUILAR LOPEZ

[Signature]

DR. RAFAEL ESCALERA PEREZ

UNIDAD IZTAPALAPA

Coordinación de Sistemas Escolares

Av. San Rafael Atlixco 186, Col. Vicentina, México, DF, CP 09340 Apdo. Postal 555-320-9000, Tels. 5804-4880 y 5804-4883 Fax: 5804-4876