

2017

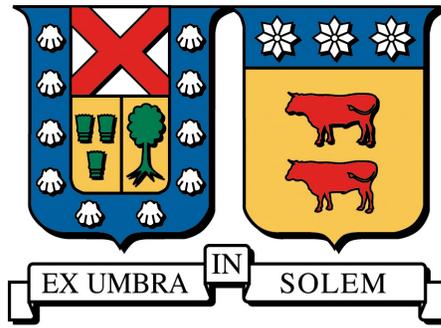
# ESTUDIO DE LAS CARACTERÍSTICAS MULTIFRACTALES EN LA CORRELACIÓN ENTRE ÍNDICES BURSÁTILES LATINOAMERICANOS Y EL PRECIO DEL CRUDO

GAJARDO CEA, GABRIEL ENRIQUE

---

<http://hdl.handle.net/11673/23610>

*Repositorio Digital USM, UNIVERSIDAD TECNICA FEDERICO SANTA MARIA*



**ESTUDIO DE LAS CARACTERÍSTICAS MULTIFRACTALES EN LA  
CORRELACIÓN ENTRE ÍNDICES BURSÁTILES LATINOAMERICANOS Y EL  
PRECIO DEL CRUDO**

Tesis de Grado presentado por

**Gabriel Enrique Gajardo Cea**

como requisito parcial para optar al título de  
**Ingeniero Civil Industrial**

y al grado de

**Magíster en Ciencias de la Ingeniería Industrial**

Profesor Referente:

Dr. Werner Kristjanpoller Rodríguez

Profesor Co-referente Interno:

Dr. Javier Scavia

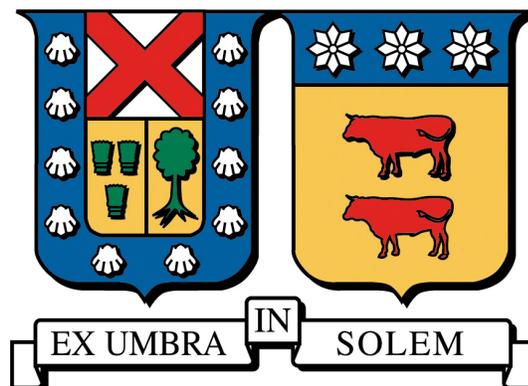
Profesor Co-referente Externo:

Dr. Marcelo Villena

AGOSTO 2017



UNIVERSIDAD TÉCNICA FEDERICO SANTA MARÍA  
DEPARTAMENTO DE INDUSTRIAS



**ESTUDIO DE LAS CARACTERÍSTICAS MULTIFRACTALES EN LA  
CORRELACIÓN ENTRE ÍNDICES BURSÁTILES LATINOAMERICANOS Y EL  
PRECIO DEL CRUDO**

Tesis de Grado presentado por

**Gabriel Enrique Gajardo Cea**

como requisito parcial para optar al título de  
**Ingeniero Civil Industrial**

y al grado de

**Magíster en Ciencias de la Ingeniería Industrial**

Profesor Referente:

Dr. Werner Kristjanpoller Rodríguez

Profesor Co-referente Interno:

Dr. Javier Scavia

Profesor Co-referente Externo:

Dr. Marcelo Villena

VALPARAÍSO, AGOSTO 2017



TÍTULO DE LA TESIS:

**ESTUDIO DE LAS CARACTERÍSTICAS MULTIFRACTALES EN LA CORRELACIÓN ENTRE ÍNDICES BURSÁTILES LATINOAMERICANOS Y EL PRECIO DEL CRUDO**

AUTOR:

**Gabriel Enrique Gajardo Cea**

TRABAJO DE TESIS, presentado en cumplimiento parcial de los requisitos para el Grado de Magíster en Ciencias de la Ingeniería Industrial y de Ingeniero Civil Industrial de la Universidad Técnica Federico Santa María.

Dr. Werner Kristjanpoller Rodríguez .....

Dr. Javier Scavia .....

Dr. Marcelo Villena .....

VALPARAÍSO, Chile. AGOSTO 2017







# AGRADECIMIENTOS

Al Vlado, Thelmo, Cristian, Canelo, Simi, Miguel Palacios, Miguel Valenzuela, Parot, Gabriela, José, Jerez, Tatán, Pancho, Cata, Cote, Kevin, Seba, Peirano, Jona y los cabros de energía y finanzas por toda la buena onda, enseñanzas y buenas experiencias durante los años de U. Sin sus distracciones esta tesis habría salido en 'dos meses full' :)

A los profes Werner y Andrés, quienes fueron la guía en este largo y horrible proceso de tesis.

A la Kettu por aguantarme tanto y ser una fuente constante de amor y cariño ♡.

Y agradezco especialmente a mi familia, Eileen, Enrique y Alejandro, por su amor incondicional, paciencia, apoyo y guía en todos estos años de crecimiento y desarrollo ♡. Sin ellos lograr esto hubiera sido infinitamente más complicado y frustrante.



---

## RESUMEN EJECUTIVO

En esta tesis se realiza un amplio análisis de las bondades del uso de la teoría fractal en el área financiera. El objetivo es demostrar las limitaciones de las técnicas y supuestos lineales clásicos y superar parcialmente estas limitaciones mediante el uso de técnicas asociadas a la geometría fractal desarrollada por Benoit Mandelbrot.

Para este objetivo se realiza una extensiva introducción a la geometría fractal, las ventajas que representa la introducción de dimensiones fraccionales en el análisis y como este puede ser extendida para el análisis de una propiedad llamada dependencia de largo plazo, cuya interpretación más extendida es la de decaimiento de la función de autocorrelación, o correlación cruzada, de forma más lenta que la asumida exponencial. Esto posee una serie de implicancias en la dependencia de las series en si mismas y en otras ante pequeños cambios, los cuales pueden poseer gran influencia en períodos posteriores.

Con ello en mente, se cuestionan los fundamentos clásicos de la Hipótesis del Mercado Eficiente y se plantean las características de la Hipótesis del Mercado Fractal que aportan realismo al análisis de series de precios. Se repasan una serie de métodos de estimación de las características de dependencia de largo plazo presentes en las series, con las cuales es posible aplicar métricas de eficiencia de mercado, abordándolas desde una perspectiva monofractal. El supuesto de monofractalidad es posteriormente cuestionado ante el aumento de las bondades del análisis fractal cuando se considera un sistema multifractal. Esto es extendido desde el marco de trabajo de la autocorrelación hacia su utilización en el estudio de la correlación cruzada.

Además, se realiza una amplia introducción al análisis de los procesos estocásticos y sus características. Se introducen además los modelos clásicos utilizados para la descripción de series financieras y como éstos han sido ampliamente estudiados para la caracterización de las series reales de precios.

Finalmente se realiza un repaso de los principales modelos econométricos para el modelamiento de series temporales, que si bien presentan gran utilidad en el estudio de las series reales, poseen supuestos poco realistas.

Con el objetivo de poner estos conocimientos en práctica, se analizan las autocorrelaciones

---

e índices de eficiencia de mercado de seis índices Latinoamericanos y dos de Estados Unidos, utilizando el método Multifractal Detrended Fluctuation Analysis (MF-DFA).

Luego de ello se investiga la correlación entre los índices en cuestión con el precio del crudo dado por la corriente del West Texas Intermediate (WTI), utilizando el método llamado Multifractal Detrended Cross-Correlation Analysis (MF-DCCA).

Con el objetivo de extender el análisis y ahondar en las características de la correlación, se analiza la presencia de asimetría en la correlación cruzada, mediante el estudio del Multifractal Detrended Asymmetric Cross-Correlation Analysis (MF-ADCCA), que permite demostrar que la correlación cambia según la tendencia de cada serie.

El set de datos utilizados cubre el período comprendido entre el 28 de Febrero de 2005 y el 29 de Diciembre de 2016, con lo cual es posible observar como la crisis subprime de 2008 influyó en las características de la correlación y autocorrelación. Los resultados indican un aumento en la ineficiencia de mercado luego de la crisis para la mayoría de países Latinoamericanos. Además se encuentra evidencia de correlación con dependencia de largo plazo para todos los índices, junto con altos grados de multifractalidad en ellas. Para analizar la asimetría se consideran por separado las tendencias de cada serie, siendo éstas de vital relevancia en el grado de correlación.

Finalmente, mediante el uso de ventana móvil, se examinó el cambio temporal de los exponentes de autocorrelación y de correlación cruzada y sus cambios según tendencia de las series, tanto para el caso clásico, y al considerar la asimetría, encontrándose interesantes patrones en su comportamiento.

**Keywords:** Correlación, Multifractalidad, Correlación asimétrica, Análisis de series de tiempo, Mercados bursátiles latinoamericanos, Mercado del crudo

---

## ABSTRACT

In this thesis a broad analysis of the benefits of the use of fractal theory in the financial area is carried out. The objective is to demonstrate the limitations of classical linear techniques and assumptions, while partially overcoming these limitations by using techniques associated with fractal geometry developed by Benoit Mandelbrot.

For this purpose we present an extensive introduction to the fractal geometry, and analyze the advantages of the use of fractional dimensions. We also show how this can be extended for the analysis of a property called long-term dependency, whose more popular interpretation is as a rate of decay of the autocorrelation function, or cross-correlation, slower than the assumed exponential. This has several implications in the dependence of the series on themselves and in others to small changes, which can have great influence in later periods. With this in mind, the classic foundations of the Efficient Market Hypothesis are questioned and the characteristics of the Fractal Market Hypothesis are presented, which bring realism to the analysis of price series. We review a series of methods for estimating the long-term dependency characteristics present in the series, with which it is possible to apply market efficiency metrics, addressing them from a monofractal perspective. The assumption of monofractality is later questioned in light of the increase in the goodness of fractal analysis when considering a multifractal system. This is extended from the framework of autocorrelation to its use in the study of cross-correlation.

In addition, a broad introduction is made to the analysis of stochastic processes and their characteristics. We also introduce the classic models used for the description of financial series and how they have been extensively studied for the acquisition of characteristics of the real series of prices. Finally, a review of the main econometric models for the modeling of time series is carried out, which, while very useful in the study of the real series, have unrealistic assumptions.

In order to put this knowledge into practice, the autocorrelations and market efficiency indices of six Latin American and two US indices are analyzed using the Multifractal Detrended Fluctuation Analysis (MF-DFA) method. After that, the correlation between the indices in question and the West Texas Intermediate (WTI) crude oil price is investigated,

---

using the method called Multifractal Detrended Cross-Correlation Analysis (MF-DCCA). With the aim of extending the analysis and delving into the characteristics of the correlation, we analyze the presence of asymmetry in cross-correlations, by studying the Multifractal Detrended Asymmetric Cross-Correlation Analysis (MF-ADCCA), which allows to demonstrate that the correlation changes according to the trend of each series.

The data set used covers the period from February 28, 2005 to December 29, 2016, making it possible to observe how the subprime crisis of 2008 influenced the characteristics of correlations and autocorrelations. The results indicate an increase in market inefficiency after the crisis for most Latin-American countries. There is also evidence of correlation with long-term dependence for all indices, along with high degrees of multifractality in them. To analyze the asymmetry, the trends of each series are considered separately, being these of vital relevance in the asymmetry of the correlations. By using the rolling window method, it is possible to examine the temporal change of the autocorrelation and cross-correlation exponents, under the normal case, and when considering the asymmetry, which allowed to find interesting patterns in its behavior.

**Keywords:** Cross-correlations, Multifractality, Asymmetric Cross-correlations, Time-Series Analysis, Latin-American stock markets, Crude oil market

# ÍNDICE DE CONTENIDOS

<b>1. Introducción</b>	<b>1</b>
1.1. Motivación . . . . .	1
1.2. Objetivo General . . . . .	5
1.3. Objetivos Específicos . . . . .	5
1.4. Alcance . . . . .	6
<b>2. Fractalidad en series financieras</b>	<b>7</b>
2.1. Fractales . . . . .	7
2.1.1. Propiedades de los fractales . . . . .	9
2.1.1.1. Dimensión . . . . .	9
2.1.1.2. Autoafinidad y Autosimilitud . . . . .	11
2.1.2. Fractales en finanzas . . . . .	13
2.1.3. Hipótesis los Mercados Eficientes (HME) . . . . .	14
2.1.4. Hipótesis de los Mercados Fractales (HMF) . . . . .	15
2.2. Tipos de correlación . . . . .	17
2.2.1. Autocorrelación con memoria larga . . . . .	17
2.2.2. Correlación de corto plazo . . . . .	18
2.2.3. Correlación con memoria larga . . . . .	18
2.3. Multifractalidad en series financieras . . . . .	20
2.3.1. Definición matemática de la multifractalidad . . . . .	22
2.3.2. Procesos Multifractales . . . . .	24
2.4. Obtención del índice de Hurst . . . . .	26
2.4.1. Análisis del Rango Reescalado . . . . .	28
2.5. MF-DFA . . . . .	30
2.5.1. Grado de eficiencia del mercado . . . . .	34
2.6. MF-DCCA . . . . .	37
2.7. MF-ADCCA . . . . .	40
<b>3. Procesos Estocásticos</b>	<b>43</b>
3.1. Series de Tiempo . . . . .	43
3.1.1. Autocovarianza . . . . .	46
3.1.2. Estacionariedad estricta . . . . .	47
3.1.3. Estacionariedad débil . . . . .	48
3.1.4. Ruido blanco . . . . .	49
3.1.5. Raíz unitaria . . . . .	50
3.2. Algunos tipos de procesos estocásticos . . . . .	51
3.2.1. Martingala . . . . .	51
3.2.2. Movimiento Browniano . . . . .	52
3.2.2.1. Propiedades . . . . .	55

3.2.3.	Movimiento Browniano Aritmético . . . . .	57
3.2.4.	Movimiento Browniano Geométrico . . . . .	59
3.2.5.	Movimiento Browniano Fraccional . . . . .	59
3.3.	Pruebas estadísticas . . . . .	62
3.3.1.	Test de Ljung-Box . . . . .	62
3.3.2.	Prueba de Jarque-Bera . . . . .	62
3.3.3.	Prueba de heterocedasticidad condicional autorregresiva . . . . .	64
3.3.4.	Prueba de Dickey-Fuller Aumentada . . . . .	65
3.3.5.	Prueba de memoria larga en la correlación . . . . .	66
3.3.6.	Coefficiente de correlación DCCA . . . . .	66
<b>4.</b>	<b>Modelos de series de tiempo financieras</b>	<b>69</b>
4.1.	Modelos lineales . . . . .	69
4.1.1.	Modelos autorregresivos y de media móvil . . . . .	70
4.1.2.	Vector autorregresivo . . . . .	71
4.2.	Modelos heterocedásticos . . . . .	72
4.2.1.	Modelo de Heterocedasticidad Condicional Autorregresiva (ARCH)	73
4.2.1.1.	Construcción de un modelo ARCH . . . . .	74
4.2.2.	Modelo Generalizado de Heterocedasticidad Condicional Autorre-	75
4.2.2.1.	gresiva (GARCH) . . . . .	76
4.2.2.1.	Construcción de un modelo GARCH . . . . .	76
4.2.3.	Variaciones de los modelos GARCH . . . . .	77
4.2.3.1.	EGARCH . . . . .	77
4.2.3.2.	GJR-GARCH . . . . .	77
4.3.	Medidas de calidad relativa de modelos . . . . .	77
4.3.1.	Criterio de información de Akaike . . . . .	77
4.3.2.	Criterio de información Bayesiano . . . . .	78
4.4.	Redes Neuronales Artificiales . . . . .	79
4.4.1.	Motivación . . . . .	79
4.4.2.	Red Neuronal Prealimentada . . . . .	79
4.4.3.	Estimación de modelos y pronósticos . . . . .	82
4.4.4.	Algoritmo de propagación hacia atrás ( <i>Backpropagation</i> ) . . . . .	83
4.4.4.1.	Algoritmo de Levenberg-Marquardt para la propagación	85
4.4.4.1.	hacia atrás ( <i>Backpropagation</i> ) . . . . .	85
<b>5.</b>	<b>Presentación y análisis de resultados</b>	<b>87</b>
5.1.	Contexto del petróleo . . . . .	87
5.2.	Contexto Latinoamericano . . . . .	89
5.3.	Análisis estadístico de las series . . . . .	96
5.3.1.	Estadísticas descriptivas . . . . .	96
5.3.2.	Prueba de memoria larga en la correlación . . . . .	97
5.3.3.	Coefficiente de correlación DCCA . . . . .	97
5.4.	Resultados MF-DFA . . . . .	99
5.4.1.	Cambios temporales en la autocorrelación . . . . .	102
5.4.2.	Ventana móvil de 4 años . . . . .	104
5.5.	Resultados MF-DCCA . . . . .	106
5.5.1.	Cambios temporales en la correlación . . . . .	108
5.6.	Resultados MF-ADCCA . . . . .	111
5.6.1.	Asimetría según tendencia del crudo . . . . .	111
5.6.2.	Asimetría según tendencia del índice . . . . .	112

5.6.3. Cambios temporales en la asimetría . . . . .	115
5.6.3.1. Ventana de 4 años . . . . .	115
5.7. Conclusiones generales . . . . .	118
<b>6. Glosario</b>	<b>121</b>
<b>Bibliografía</b>	<b>127</b>
<b>ANEXO</b>	<b>133</b>





# ÍNDICE DE TABLAS

5.1. Estadísticas . . . . .	96
5.2. DCCA para diferentes $s$ . . . . .	99
5.3. Grado de eficiencia según MF-DFA. . . . .	101
5.4. Multifractalidad según MF-DFA. . . . .	101
5.5. Ranking de ineficiencia. . . . .	101
5.6. Resumen $H(2)$ móvil, 4 años. . . . .	103
5.7. Exponente de correlación $H_{xy}(q)$ y su dependencia del tamaño de las fluctuaciones. . . . .	108
5.8. Estadísticas del $\alpha$ . . . . .	108
5.9. Resumen $H_{xy}(2)$ móvil, 4 años . . . . .	109
5.10. Asimetría en el $H(q)$ según WTI . . . . .	112
5.11. Asimetría en el $H(q)$ según el índice . . . . .	114
5.12. Estadísticos $\Delta H_{xy}(2)$ móvil para ventana de 4 años, WTI . . . . .	116
5.13. Estadísticos $\Delta H_{xy}(2)$ móvil para ventana de 4 años, índice . . . . .	116



# ÍNDICE DE FIGURAS

2.1. Copo de Nieve de Koch . . . . .	10
2.2. Simulación GARCH . . . . .	22
2.3. Orden $q$ . . . . .	33
2.4. Ejemplos MF-DFA . . . . .	33
2.5. Espectro multifractal . . . . .	34
3.1. Movimiento Browniano Aritmético . . . . .	58
3.2. Movimiento Browniano Fraccional . . . . .	61
a. Serie antipersistente. . . . .	61
b. Serie antipersistente. . . . .	61
c. Serie aleatoria. mBa clásico. . . . .	61
d. Serie persistente. . . . .	61
e. Serie persistente. . . . .	61
4.1. Neurona . . . . .	79
4.2. Red Neuronal . . . . .	80
5.1. Brent y WTI . . . . .	89
5.2. Países afectados por caídas del crudo . . . . .	90
5.3. $Q_{cc}(m)$ . . . . .	97
5.4. Coeficiente de correlación DCCA . . . . .	98
5.5. MF-DFA $f(\alpha)$ vs. $\alpha$ . . . . .	100
5.6. MF-DFA $H(q)$ vs. $q$ . . . . .	100
5.7. Ventana móvil, 4 años, $H(2), 1$ . . . . .	104
a. Ventana móvil, 4 años WTI. . . . .	104
b. Ventana móvil, 4 años Ibovespa. . . . .	104
c. Ventana móvil, 4 años IPSA. . . . .	104
d. Ventana móvil, 4 años IPC. . . . .	104
e. Ventana móvil, 4 años Merval. . . . .	104
5.8. Ventana móvil, 4 años, $H(2), 2$ . . . . .	105
a. Ventana móvil, 4 años SPBVL. . . . .	105
b. Ventana móvil, 4 años COLCAP. . . . .	105
c. Ventana móvil, 4 años DJIA. . . . .	105
d. Ventana móvil, 4 años SP500. . . . .	105
5.9. MF-DCCA $f(\alpha)$ vs. $\alpha$ . . . . .	107
5.10. MF-DCCA $H(q)$ vs. $q$ . . . . .	107
5.11. Ventana móvil de 4 años para $H_{xy}(2)$ . . . . .	110
5.12. Asimetría según WTI . . . . .	112
5.13. Asimetría según índice . . . . .	114
5.14. Ventana móvil de 4 años para $\Delta H_{xy}(2)$ . . . . .	117

1.	Relación $H_{xy}(q)$ con $\gamma(q)$ . . . . .	134
2.	Gráficos de precios 1 . . . . .	139
	a. Precios del WTI. . . . .	139
	b. Precios del Ibovespa. . . . .	139
	c. Precios del IPSA. . . . .	139
	d. Precios del IPC. . . . .	139
	e. Precios del Merval. . . . .	139
3.	Gráficos de precios 2 . . . . .	140
	a. Precios del SPBVL. . . . .	140
	b. Precios del COLCAP. . . . .	140
	c. Precios del DJIA. . . . .	140
	d. Precios del SP500. . . . .	140
4.	Gráficos de retornos 1 . . . . .	141
	a. Retornos del WTI. . . . .	141
	b. Retornos del Ibovespa. . . . .	141
	c. Retornos del IPSA. . . . .	141
	d. Retornos del IPC. . . . .	141
	e. Retornos del Merval. . . . .	141
5.	Gráficos de retornos 2 . . . . .	142
	a. Retornos del SPBVL. . . . .	142
	b. Retornos del COLCAP. . . . .	142
	c. Retornos del DJIA. . . . .	142
	d. Retornos del SP500. . . . .	142
6.	Comportamiento de escala 1 . . . . .	143
	a. DFA WTI. . . . .	143
	b. DFA Ibovespa. . . . .	143
	c. DFA IPSA. . . . .	143
	d. DFA IPC. . . . .	143
	e. DFA Merval. . . . .	143
7.	Comportamiento de escala 2 . . . . .	144
	a. DFA SPBVL. . . . .	144
	b. DFA COLCAP. . . . .	144
	c. DFA DJIA. . . . .	144
	d. DFA SP500. . . . .	144
8.	Comportamiento de escala según MF-DCCA . . . . .	145
9.	Comportamiento de escala según WTI . . . . .	146
10.	Comportamiento de escala según índice . . . . .	147
11.	Asimetría en $H(q)$ según WTI . . . . .	148
12.	Asimetría en $H(q)$ según índice . . . . .	149

# CAPÍTULO 1

## INTRODUCCIÓN

### 1.1. Motivación

Muchos sistemas poseen complejas relaciones en sus constituyentes, quienes interactúan en formas que alteran continuamente las características de la dependencia entre ellos. Dicha complejidad es especialmente alta en sistemas económicos donde existe una inmensurable cantidad de participantes, que no solo se ven afectados por su pasado remoto, sino que por el de los otros constituyentes. Usualmente los participantes pueden ver su actividad agregada en series de precios, ya sea de commodities, futuros, precios de acciones o índices bursátiles. Al menos dos medidas permiten cuantificar estas relaciones: la correlación y la covarianza cruzada. Ambas medidas están relacionadas, pero la primera además permite cuantificar si la correlación es poderosa o no, siendo usualmente utilizado el coeficiente de Pearson para ello. Sin embargo, este coeficiente no es robusto y puede entregar resultados engañosos ante la presencia de *outliers*, abundantes en la data real, especialmente en el tipo de datos aquí estudiados.

La Hipótesis de los Mercados Eficientes (HME), sugiere que los precios futuros son impredecibles, debido a la capacidad de incorporación de la información que poseen las series de precios. Esto se expresa matemáticamente en series que siguen una caminata aleatoria, mientras que los retornos se asemejan a un proceso de ruido blanco. Sin embargo, la evidencia ha demostrado que esto no ocurre en la práctica, donde los retornos muestran distribuciones de colas pesadas, en vez de las Gaussianas, y el precio posee algún nivel

de predictibilidad, dada su persistencia. Además de dichos problemas, es posible observar asimetrías, clústers de volatilidad, asimetría en las escalas de tiempo, entre muchas otras que hacen que la hipótesis falle al tratar de dar respuesta a los problemas del mundo real [1].

En muchos casos, se suele asumir que los retornos financieros son estacionarios, lo cual posee una serie de implicancias estadísticas para el modelamiento y análisis. Esta característica suele determinarse mediante estadísticos como el de Dickey Fuller o Phillips–Perron. Sin embargo, en muchos casos dicho estadístico puede entregar valores sesgados debido a quiebres estructurales, cúmulos de volatilidad y un sinnúmero de características que pueden hacer que los test incurran en algún tipo de error. Ante ello, resulta útil determinar las características de las correlaciones con métodos que no requieren tantos supuestos sobre las series.

Numerosos otros supuestos se hacen al momento de analizar riesgo y evaluar portafolios, así como al construir modelos. Sin embargo, se ha demostrado extensivamente que los supuestos sobre los que se realiza la mayoría de los análisis clásicos no corresponden a lo que es posible observar en la realidad, por lo que extensiones o nuevas formas de análisis son requeridas. Uno de estos principales supuestos es el de comportamiento Gaussiano de los retornos, que si bien simplifica enormemente el estudio, es posible observar que en la realidad se presentan sesgos, diferencias de curtosis, asimetrías y distribuciones bastante diferentes a la Normal, observándose colas pesadas y valores extremos.

Si bien los modelos clásicos han entregado buenas herramientas de análisis a lo largo de los años a pesar de sus numerosos supuestos de dudosa veracidad, el problema recae en que muchas series temporales exhiben comportamiento muy complejo, como la auto-afinidad o auto-similaridad, mediante la cual es posible caracterizar la autocorrelación o correlación cruzada mediante leyes de potencia. Esto se basa en la teoría de los fractales, los cuales son figuras geométricas en que pueden observarse partes que replican la forma del todo. Esto resulta de gran interés, ya que así como los fractales se repiten a través de las dimensiones, los precios de acciones parecieran moverse replicando patrones geométricos a través del tiempo. Ello permite encontrar universalidad en una gran cantidad de problemas. Hurst fue el primero en identificar estas características estudiando el río Nilo, mediante el análisis

del Rango Reescalado, el cual indica que los registros están bien descritos por la relación  $R/S \sim n^H$ , donde  $n$  es la escala de tiempo y  $H$  es el llamado Coeficiente o Exponente de Hurst, cuyos valores van del 0 al 1, donde un  $H = 0,5$  indica que no existe autocorrelación o solo hay en el corto plazo. Por otro lado si  $0,5 < H < 1$  se dice que existen autocorrelaciones con memoria larga de comportamiento persistente, mientras que si  $0 < H < 0,5$  se habla de autocorrelación de memoria corta o de comportamiento antipersistente.

El desarrollo de las matemáticas de Ito para procesos estocásticos y el movimiento Browniano rigen las bases sobre las cuales se trabajan los conceptos utilizados en finanzas y economía, por lo que una extensiva introducción se realiza para cuestionar su relación con la realidad. En este trabajo se cuestiona principalmente el supuesto de independencia de las variables financieras, que se analiza mediante tres técnicas de caracterización multifractal para el estudio de dicha dependencia, tanto a nivel de autocorrelación como de correlación con un commodity, específicamente, el precio del crudo.

En años recientes, los países industrializados han sufrido importantes impactos económicos derivados de las variaciones de dicho commodity, dando origen a un gran número de investigaciones en los efectos cuantitativos y cualitativos de estas variaciones en el crecimiento, tasas de cambio y mercados financieros, haciendo esencial el entendimiento de la dinámica del movimiento e influencia de dicho commodity. Sin embargo, considerando lo anteriormente mencionado, es necesario extender el análisis y buscar respuestas de otro tipo, que logren ajustarse de mejor forma al real comportamiento de las series financieras. Desde que los trabajos de Hurst [2, 3] y Mandelbrot [4] fueron publicados, se ha llevado a cabo mucha investigación con el objeto de identificar las características de memoria larga o dependencia de largo plazo en series del mundo real, inicialmente por medio del Análisis del Rango Reescalado (R/S), en algunas áreas tan disímiles como la minería [5], tráfico [6], contaminantes del aire [7], y economía [8, 9, 10]. Si bien éste método fue pionero en la determinación de las características antes mencionadas, presentaba sensibilidad a las autocorrelaciones de corto plazo y no estacionariedad de la serie, la cual podría llevar a estimadores sesgados en la estimación de los parámetros que rigen la dependencia de largo plazo [11]. Otros métodos surgieron para enfrentar este problema, como el método *Wavelet Transform Modulus Maxima* (WTMM) [12], el *Fluctuation Analysis* (FA) [13], el

*Detrended Fluctuation Analysis* (DFA) [14], el *Detrended Moving Average analysis* (DMA) [15], su extensión multifractal (MF-DMA) [16] y el *Multifractal Detrended Fluctuation Analysis* (MF-DFA) [17], el cual ha sido extensamente utilizado en el análisis de series de tiempo [18, 19, 20, 21, 22, 23], y ha entregado una nueva forma de investigar la ineficiencia de mercado [24, 25].

Con la identificación de las correlaciones con memoria larga en muchas series financieras [26, 27, 28, 29, 30, 31], se desarrollaron otros métodos, tales como el *Detrended Cross-Correlation Analysis* (DCCA) [32], el *Multifractal Detrending Moving Average Cross-Correlation Analysis* (MF-X-DMA) [33], y el *Multifractal Detrended Cross-Correlation Analysis* [34] (MF-DCCA), el cual integra el DCCA en el marco de trabajo del MF-DFA. A partir de entonces, el MF-DCCA ha sido ampliamente utilizado en el análisis de series financiera [35, 36, 37, 38, 39].

Sin embargo, estos modelos no consideran las respuestas asimétricas de los mercados ante diferentes noticias económicas [40, 41, 42, 43]. En este estudio se pretende analizar la correlación ante diferentes tendencias de uno de los mercados. Para este fin se aplica el *Multifractal Asymmetric Detrended Cross-Correlation Analysis* (MF-ADCCA) propuesto por Cao et al. [44]. Este modelo es una combinación del MF-DCCA y la versión asimétrica del DFA, el A-DFA propuesto por Alvarez-Ramirez et al. [45]. Al usar este modelo, es posible estudiar las propiedades multifractales de la correlación, sin caer en la división subjetiva entre mercados *bull* y *bear*.

Con el objetivo de estudiar las propiedades multifractales en la autocorrelación con memoria larga de los índices bursátiles y en su correlación con memoria larga con el precio del crudo, se aplica el MF-DFA y el MF-DCCA respectivamente. Para este fin, las teorías y modelos basados en la HME ya no son apropiadas para el análisis.

Este estudio se enfoca en seis índices bursátiles Latinoamericanos, en los cuales depende la mayoría del comercio de la región, y que han sido los principales receptores de inversión extranjera en las últimas décadas. De los seis países, Mexico, Brasil, Argentina y Colombia son productores de crudo, haciendo a Latinoamérica un exportador global en el mercado mundial de crudo. Por otro lado, Chile y Perú son importadores netos, por lo que en el análisis se encuentran países con diferentes relaciones con el mercado del crudo. La com-

preensión de la dinámica de la relación es de suma relevancia para gobiernos, organizaciones multinacionales e inversores en la región para prepararse ante potenciales movimientos adversos en la economía, por lo que cualquier información que permita enfrentar los aspectos no considerados por el análisis clásico será de suma utilidad.

Las contribuciones de este trabajo son las siguientes. Primero, según la información recopilada, este es el primer estudio de las características multifractales en la correlación con el petróleo de los índices bursátiles Latinoamericanos. Esto también aplica para la medición del grado de asimetría en la correlación. Además, se estudian las correlaciones de forma relativamente cualitativa con los estadísticos propuestos por Podobnik et al. [32] y por Zebende [46], y de forma cuantitativa mediante el MF-DCCA y MF-ADCCA. Finalmente, se utiliza la metodología de ventana móvil para estudiar la dinámica temporal de la correlación y de su asimetría.

## 1.2. Objetivo General

- Demostrar cómo el análisis fractal permite obtener información adicional a la que es posible obtener mediante modelos econométricos clásicos y ahondar en las bases de su aplicación teórica y práctica en los mercados Latinoamericanos.

## 1.3. Objetivos Específicos

- Aportar a la literatura económica y financiera de países Latinoamericanos con importantes conclusiones sobre su actividad económica agregada y su relación con el precio del crudo.
- Estudiar las limitaciones de los modelos econométricos clásicos y el poco realismo de sus supuestos.
- Estudiar métricas de eficiencia de mercados bursátiles y su evolución en el tiempo.

- Desarrollar las bases matemáticas del análisis fractal y su potencial de aplicación a las finanzas y economía.
- Aplicar métodos recientes de estimación de características multifractales estudiando sus bases matemáticas e implicancia práctica.

## 1.4. Alcance

El análisis que aquí se presenta pretende explorar la multifractalidad presente tanto en la autocorrelación de índices bursátiles como en la correlación con el precio del crudo. Para ello, el análisis se limita a Chile, Perú, Colombia, México, Argentina y Brasil, países que se considera poseen la mayor parte de la actividad económica de la región. Si bien se incurre en un análisis bastante agregado que no considera las diferentes relaciones con el petróleo de los componentes de dichos índices bursátiles, se plantea una metodología bastante apropiada y extensible a dicho análisis.

El estudio pretende dilucidar como es que cambios en una serie impactan en la otra en el largo plazo, demostrando primero si efectivamente dicha influencia existe.

Para justificar el análisis se realiza una extensiva revisión bibliográfica y matemática sobre los métodos utilizados, con el objetivo de introducir al lector en la teoría aplicada.

El período de análisis se limita desde febrero de 2005 hasta diciembre de 2016, por lo cual las conclusiones derivadas del análisis solo deben ceñirse a dicho período, ya que como se observa en el análisis de ventana móvil, los métodos utilizados son totalmente dependientes de la ventana estudiada.

## CAPÍTULO 2

# FRACTALIDAD EN SERIES FINANCIERAS

### 2.1. Fractales

Los fractales [47] son un objeto abstracto que presenta diversas características de complejidad mayor a las de la geometría tradicional o Euclidiana, cuyas formas limitadas a menudo poseen problemas al momento de modelar fenómenos naturales. Por ejemplo, un pino no puede ser adecuadamente replicado por la geometría Euclidiana, tan solo es posible realizar una aproximación en forma de cono, quizás junto a un cilindro. Un pino real posee una estructura similar a un cono, pero bastante más rugosa y áspera que lo que es posible modelar mediante dicha estructura. Específicamente, se encuentra que posee un tronco rugoso, con patrones que se repiten a diversas escalas, así como también ramas sucesivamente más pequeñas de forma similar.

Esta diferencia, la forma en que se escalan, es decir, como se altera la forma al aumentar o disminuir uniformemente un objeto, es la principal diferencia con la geometría Euclidiana. Por ejemplo, un polígono, el cual se dice que reside en un plano, es decir, en dos dimensiones  $d = 2$ , al ver aumentado el largo de sus límites por  $n = 2$ , verá multiplicada su área por 4. Es decir  $n^d = 2^2 = 4$ . A su vez, si el radio de una esfera es duplicada  $n = 2$ , como reside en un espacio tridimensional  $d = 3$ , ésta verá su volumen incrementado por 8, ya que  $n^d = 2^3 = 8$ . Ahora bien, en el caso de los objetos conocidos como fractales, el valor resultante no es necesariamente un entero, siendo la potencia, en este caso  $d$  no necesariamente un entero para el fractal, siendo denominada usualmente como dimensión

fractal, la cual por lo general será mayor a la dimensión topológica del objeto.

Este concepto, si bien complejo, puede verse más claro tomando las considerando que es una forma de cuantificar la rugosidad de la curva. En geometría usualmente cuantificamos de dimensión 0 a puntos, 1 a líneas, 2 a superficies, y 3 a volúmenes, las cuales están ligadas a los grados de libertad. Es posible considerar la rugosidad como aumento en dimensión, por ejemplo una curva rugosa tiene dimensión entre 1 y 2, y una superficie rugosa entre 2 y 3.

Por ello cuantificar fractales se convierte en definir la proporción de espacio físico llenado por ellos. Por ejemplo, una hoja arrugada que no ocupa el 100 % de su espacio físico, tendría dimensión entre 2 y 3, siendo considerados fractales los objetos que poseen dimensión de Hausdorff–Besicovitch superior a su dimensión topológica. La geometría fractal permite analizar 3 puntos importantes:

- Provee dimensiones adicionales y más cercanas a la realidad que la geometría Euclidiana
- La mayoría de sistemas complejos son caóticos, y exhiben conductas extrañas que no pueden ser asociadas con límites o campos que no se pueden representar en dimensiones enteras.
- Los fractales son escalables, es decir se puede reducir o ampliar su análisis para observar detalles, mientras que las formas básicas es posible observarlas en cada escala.

El concepto de fractal proviene del latín *fractus*, que significa 'roto' o 'fracturado', y Mandlebrot, considerado el padre de la teoría fractal, los definió como una 'formas geométricas rugosas que pueden ser reducidos en partes, donde cada una de ellas es aproximadamente una copia reducida de la forma completa'. Sin embargo, su definición aún no posee un consenso absoluto, por lo que si bien son fácilmente reconocibles, es mejor analizarlos en términos de sus propiedades y características, las cuales son principalmente la autosimilaridad y la dimensión fractal.

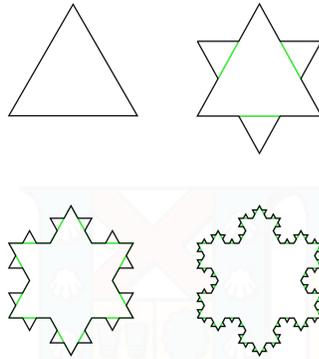
## 2.1.1. Propiedades de los fractales

### 2.1.1.1. Dimensión

La dimensión de Hausdorff–Besicovitch es un concepto matemático desarrollado por Felix Hausdorff [48] y estudiado extensivamente por Abram Samoilovitch Besicovitch [49, 50], que permite medir el tamaño local de cierto espacio que toma en consideración la distancia entre los puntos en él. En términos simples, es una métrica que permite definir la dimensión no entera asociada a un objeto fractal, no siendo la única, pero si la más utilizada.

Esta dimensión es equivalente a la clásica dimensión Euclidiana para los objetos que pertenecen a dicha rama de estudio, que son los que poseen un número pequeño de esquinas. Sin embargo, para objetos que presentan propiedades de autosimilitud y escala, permite entregar dimensiones fraccionales.

Imaginando uno de los casos que popularizó el estudio fractal, es posible observar, para las costas de Inglaterra, que esta debería tener una dimensión de 1, entregando una medida de  $X$  kilómetros. Al tratar de aproximar una estructura tan irregular como una costa, una buena idea sería hacerlo mediante muchos segmentos de líneas rectas, con lo que sería posible obtener, por ejemplo  $Y > X$  kilómetros. El problema que surge es que dependiendo del largo de los segmentos, algunas partes podrían aproximarse como rectas sin serlo, ante lo que resultaría normal intentar hacerlo con segmentos más pequeños, con lo cual se obtendría  $Z > Y$  kilómetros. Es posible repetir este proceso incontables veces, obteniendo siempre una longitud mayor de costa. Esto se debe a que la costa tiene una estructura fractal, infinitamente delgada al considerar las entradas a la tierra, ante lo que se considera que posee estructura fractal, siempre es posible encontrar complejidad ante menores escalas. Otro ejemplo, es el Copo de Nieve de Koch [51], ilustrado en la siguiente figura



**Figura 2.1:** Copo de Nieve de Koch

(Fuente: Addison, Paul S. (1997). *Fractals and Chaos: An Illustrated Course.*)

Claramente, esta figura posee una dimensión topológica de 1, pero el largo de la curva entre dos puntos es infinita, debido a que no hay línea recta, todo esta compuesto de un número infinito de segmentos unidos en diferentes ángulos. Intuitivamente, estos objetos fractales se pueden considerar demasiado complejos para tener 1 dimensión, pero demasiado simples para tener 2.

Si bien es la medida de dimensión fractal más utilizada, no es la única, pero si cumple la propiedad de ser menor al resto se las aproximaciones, solo cumpliéndose que es mayor o igual a la dimensión topológica  $D_T \leq D_H$ . En la práctica estas mediciones están afectadas por las aproximaciones metodológicas, y presentan sensibilidad ante ruido numérico y metodológico.

Su definición matemática es la siguiente

$$\dim_H(F) := \sup\{s : \mathcal{H}(F) = \infty\} := \inf\{s : \mathcal{H}(F) = 0\} \quad (2.1)$$

En el caso de series temporales autoafines la dimensión fractal está directamente relacionada con el índice de Hurst. Estando la primera, evidentemente limitada por  $1 < D < 2$ , y pudiendo ser calculada como

$$D = 2 - H \quad (2.2)$$

o generalizable para un objeto que vive en una dimensión topológica de  $n$

$$D = n + 1 - H \quad (2.3)$$

Si bien esta propiedad puede no cumplirse de forma estricta debido a los problemas en la estimación de la dimensión o el coeficiente de Hurst [52], es una buena estimación.

La idea detrás de esta expresión radica en que la dimensión fractal es una propiedad local, mientras que la dependencia de largo plazo, o memoria larga es una característica global. Ello implica que para procesos autoafines, las propiedades locales están reflejadas en las propiedades globales, resultando en la expresión (2.3).

El método Hausdorff-Besicovitch simplificado, para determinar la dimensión fractal puede resumirse como

- S: cantidad de segmentos
- L: escala de medición
- D: dimensión

$$S = L^D \quad (2.4)$$

$$\ln(S) = \ln(L^D) \quad (2.5)$$

$$\ln(S) = D \cdot \ln(L) \quad (2.6)$$

$$D = \frac{\ln(S)}{\ln(L)} \quad (2.7)$$

### 2.1.1.2. Autoafinidad y Autosimilitud

Un objeto autosimilar o autosemejante es un objeto que es exacta o aproximadamente similar a una parte de sí misma. Esto se da principalmente en fractales artificiales. Para el caso de series financieras, esto puede observarse en el hecho de que sin leyenda, no es posible determinar la periodicidad de la serie, no se sabrá si son series mensuales, anuales, diarias, etc.

La invarianza a escala es una forma exacta de autosimilitud, donde a cualquier aumento existe una parte que es similar al entero. Esto puede ser observado en el Copo de Nieve de Koch en la Figura (2.1), que puede ver aumentada parte de su estructura y encontrar el mismo patrón de forma indefinida. Esta similitud puede no ser trivial, y encontrarse más detalle en escalas arbitrariamente pequeñas.

Existen varios tipo de autosimilitud, entre las que destacan la exacta, la aproximada, la estadística y la cualitativa. La exacta es la del caso del Copo de Nieve de Koch, una o más partes de la estructura completa ven repetido su patrón de forma perfecta. La invarianza a escala explicitada anteriormente posee la característica que puede multiplicarse continuamente por tres sin cambiar de forma.

La aproximada es la más frecuentemente encontrada en la naturaleza, presentándose diversas diferencias entre las partes y el todo, y usualmente generándose solo dentro de una porción limitada del total. Es posible interpretarla como una autosimilitud exacta con una perturbación aleatoria.

La estadística corresponde al mantenimiento de ciertas propiedades estadísticas ante el cambio de escala. Este es el caso de las costas geográficas.

Un objeto autoafín es un objeto que permite un escalamiento diferente para las distintas direcciones de una base ortonormal. Esto es una diferencia importante con la autosimilitud, que se refiere a escalamientos isotrópicos, mientras que la autoafinidad se refiere a distintos valores del escalamiento. Ello implica que los objetos bajo dicha transformación son invariantes tanto ante dilataciones como rotaciones.

Matemáticamente, es posible considerar que un proceso aleatorio  $\{X(t)\}$ , dado  $X(0) = 0$ , será autoafín si cumple

$$\{X(ct_1), \dots, X(ct_2)\} \stackrel{d}{=} \{c^H X(t_1), \dots, c^H X(t_2)\} \quad (2.8)$$

para algún  $H > 0$ , llamado el índice de autoafinidad, o exponente de escala de  $X(t)$ .

Algunos autores entregan la misma definición para la autosimilitud [53], pero dicho término

es más estricto.

Existen dos tipos de procesos autoafines utilizados en finanzas. Los primeros son los descritos arriba que asumen incrementos independientes y estacionarios, correspondientes a procesos L-estables, como el movimiento Browniano, que puede ser visto como la caminata aleatoria continua. Los segundos son los correspondientes a los movimientos Brownianos fraccionales (mBf), procesos autoafines  $B_H(t)$  de características continuas. El exponente  $H$ , el exponente de autoafinidad, satisface que  $0 < H < 1$ . En el caso particular de un mBf si toma un valor de 0,5, resulta equivalente al movimiento Browniano clásico. Sin embargo, para otros valores de  $H$ , el mBf posee incrementos dependientes. Se hablará de autocorrelación negativa (no en el sentido clásico lineal), o antipersistencia cuando  $0 < H < 0,5$ , y autocorrelación positiva, o persistencia, cuando  $0,5 < H < 1$ . Se dice que en este último caso estamos en presencia del fenómeno denominado como memoria larga o dependencia de largo plazo.

Los conceptos de movimiento Browniano y movimiento Browniano fraccional se estudian en detalle en el capítulo de Procesos Estocásticos.

### 2.1.2. Fractales en finanzas

Se sugiere que el mercado exhibe correlación temporal, es decir, fluctuaciones volátiles que ocurren con determinada tendencia, en que es posible observar distribuciones de probabilidad de colas pesadas. Es fácil observar la estructura fractal en gráficos de precios de, por ejemplo, acciones o índices, ya que poseen ciertas características:

- Son autosimilares, sin leyenda no es posible identificar tiempo de análisis.
- Poseen alta sensibilidad a las condiciones iniciales (relación con teoría del caos).
- Dimensión fraccional entre 1 y 2.
- No se cumplen los supuestos de normalidad, ni independencia de rendimientos bursátiles (Hipótesis de Mercados Eficientes). Esto ha originado que se utilice el

Coefficiente de Hurst para determinar la correlación entre precios sucesivos, el cual esta relacionado con la dimensión fractal.

Se debe realizar un procedimiento de análisis estadístico y fractal, que se basan en las pruebas de normalidad y colas pesadas (Pareto-Lévy). Todo esto para obtener el coeficiente de Hurst.

### 2.1.3. Hipótesis los Mercados Eficientes (HME)

Los mercados eficientes se definen como aquellos que 'reflejan toda la información disponible y se ajustan automáticamente a la nueva'. Es decir, la competencia entre los participantes lleva a un equilibrio donde el precio de mercado es buen reflejo de su valor, no existiendo sobre o subvaloración de títulos.

La naturaleza de esto radica en que al existir un número alto de compradores y vendedores tratando de obtener rentabilidad, toda potencial información que permita predecir lo que puede ocurrir en el futuro es tomada en cuenta, incorporándose de forma prácticamente inmediata en los precios de los activos. A diferencia de lo que podría pensarse, ello no implica que todos los involucrados sean racionales, pueden ser irracionales, pero su comportamiento agregado se considerará aleatorio. Es por ello que no existirá información posible a ser utilizada para generar rentabilidad, no existiendo ningún tipo de 'memoria', ni posibilidad de vencer de forma sostenida al mercado.

La implicancia de aleatoriedad es fundamental, ya que el mejor modelo de los precios será el de caminata aleatoria, con todas las implicancias que ello posee estadísticamente, y cuyas características pueden ser estudiadas en el capítulo de Procesos Estocásticos.

Existen 3 formas en que se manifiesta esta teoría

- **Forma débil**

Los precios reflejan completamente la información pasada, por lo que todos manejan la misma información y pueden aprovecharla de la misma manera para superar al mercado. Se supone que existe información desconocida que influirán los precios ante su llegada determinando su movimiento. El análisis técnico no tendrá éxito bajo

este esquema.

- **Forma semifuerte**

Los precios reflejan la información pasada y reaccionan de forma instantánea para reflejar la nueva información, por lo que no es posible utilizarla para generar retornos, resultando de poca utilidad tanto el análisis técnico como el fundamental.

- **Forma fuerte**

Los precios reflejan absolutamente toda la información, pública y privada, no pudiéndose obtener rendimientos superiores al mercado. Esta forma requiere que no existan barreras legales para que se conozca la información, como las usuales relacionadas con información privilegiada. A pesar de que existan algunos que vencen consistentemente al mercado, éstos caen bajo lo predicho de acuerdo a los supuestos Gaussianos.

Para el correcto funcionamiento de esta teoría debe existir plena liquidez, de forma que inversionistas con diferentes características de riesgo puedan negociar de forma expedita y sean capaces de tomar precios justos. Además, considerando los supuestos de aleatoriedad, los precios deberían verse bien reflejados por una distribución normal, lo cual no suele cumplirse, como se puede observar en las usuales colas pesadas de la data financiera, y que puede ser comprobado por el estadístico Jarque-Bera, introducido también en el Capítulo de Procesos Estocásticos.

#### **2.1.4. Hipótesis de los Mercados Fractales (HMF)**

La hipótesis de los mercados eficientes posee una serie de críticas tanto teóricas como empíricas. Las desviaciones se pueden atribuir en parte a imperfecciones de mercado dadas por sesgos cognitivos estudiados por economistas conductuales, que observan sobre-reacción, excesos de confianza, sesgos de información y otros numerosos errores humanos. Con ello en mente, la Hipótesis de los Mercados Fractales [8] surge como una alternativa más realista, no exenta de críticas, al comportamiento real de los mercados de capitales.

Bajo esta teoría, la liquidez toma importancia fundamental, y no se realizan supuestos

estadísticos sobre los procesos.

La diferencia más relevante radica en que la información por si misma no impactará de forma uniforme a los precios, sino que será asimilada de forma diferente, de acuerdo a los diferentes horizontes de inversión.

Es posible resumir la hipótesis en los siguientes puntos:

- Se habla de un mercado estable si hay varios inversores y diferentes horizontes temporales, lo cual implica liquidez del mercado. Desajustes de corto plazo (CP), son ajustables por inversores de largo plazo (LP). Esto implica que los inversores deben compartir los mismos niveles de riesgo, pero este debe estar ajustado para la escala del horizonte de inversión.
- El conjunto de información está más relacionado con la sensibilidad del mercado, actividad de trading, liquidez y factores técnicos de CP, más que a LP. A medida que aumenta el horizonte de inversión, el análisis fundamental es el que domina. Los cambios de precios pueden reflejar información importante solo para ese horizonte de inversión.
- Se rompe la estabilidad cuando inversores de LP dejan de participar o pasan a comerciar con el conjunto de información de CP. Esto ocurre cuando inversores creen que la información fundamental que usan para invertir ya no es fiable.  
En caso en que el horizonte se hace uniforme, el mercado entra en caída libre debido a la gran falta de liquidez, y aparecen discontinuidades en la secuencia de precios, las que generan grandes cambios y aparecen colas pesadas en la distribución de frecuencia de rentabilidades.
- Como los precios reflejan una combinación de comercio técnico a CP y valuación fundamental de LP, los cambios en precios de CP son probablemente más volátiles o ruidosos que los de LP. Las tendencias de CP son resultado del comportamiento colectivo, sin que estas sean, necesariamente, tendencias económicas de LP.
- Si un activo no tiene vínculo al ciclo económico entonces no habrá tendencia de LP

(horizonte de evaluación). Comercio, liquidez e información dominan en el CP.

## 2.2. Tipos de correlación

### 2.2.1. Autocorrelación con memoria larga

En la literatura se suele llamar Memoria larga, Dependencia de largo plazo o Persistencia de largo plazo a la idea de que el mayor ciclo aparente en una muestra será proporcional al número total de observaciones. Esta idea se ha formalizado de muchas formas, pero la más común ha sido la de considerar un decaimiento más lento que la exponencial en la función de autocorrelación.

Tasas de decaimiento hiperbólico se reportaron por primera vez por Taylor [54] y son ahora un hecho bien documentado en las series de tiempo financieras.

Suponiendo que  $\{X(t), t \in T\}$  es un proceso estacionario, se dirá que posee dependencia de largo plazo si existe un número real  $\alpha \in (0, 1)$  y una constante  $c \geq 0$ , tal que  $\lim_{k \rightarrow \infty} \frac{\rho(k)}{ck^{-\alpha}} = 1$ , o de forma alternativa  $\rho(k) \sim ck^{-\alpha}$  mientras  $k \rightarrow \infty$ .

Es posible relacionarlo con fractales al reemplazar  $\alpha$ , y considerar que el proceso estocástico  $X(t)$  es estacionario, si existe un número real  $H$  tal que la función de autocorrelación posee la siguiente tasa de decaimiento:

$$\rho(k) \sim ck^{2H-2}, k \rightarrow \infty \quad (2.9)$$

Al aumentar  $k$ , el ratio entre ambos se hace 1.  $H$  representa el coeficiente de Hurst, el cual representa la memoria de largo plazo de la serie de tiempo, pudiendo expresarse la autocovarianza de memoria larga en términos del coeficiente de Hurst como

$$\mathbb{E}(X_1(t)X_1(t+k)) \sim \frac{c}{k^{2-2H}}, k \rightarrow \infty \quad (2.10)$$

Si  $H > 1$ , el valor  $d = 2H - 2$  tomará un valor mayor a 0,5 y el proceso no es estacionario. Este comportamiento en la función de autocovarianza implica que  $\sum_{k=0}^{\infty} \rho(k) = \infty$ , es decir, que la correlación de los rezagos altos puede ser pequeña, pero su efecto acumulado es

significativo.

Es importante tener claro que resulta imposible teóricamente concluir que existe dependencia de largo plazo en muestras finitas, debido a que pueden existir rezagos en que la covarianza sea muy grande, influyendo en la tasa de decaimiento de la correlación.

### 2.2.2. Correlación de corto plazo

Se dice que dos procesos estacionarios  $\{x_t\}$  y  $\{y_t\}$  poseen correlación de corto plazo si para  $n > 0$  y/o para  $n < 0$ , la función de correlación se comporta como

$$\rho_{xy}(n) \sim \exp(-n/\delta) \quad (2.11)$$

donde  $\delta$  es el decaimiento característico  $0 \leq \delta < +\infty$ .

### 2.2.3. Correlación con memoria larga

Para entender la correlación con memoria larga y su relevancia, es necesario primero definir el significado e implicancias de este tipo de correlación. Asumiendo los procesos  $X_1(t)$  y  $X_2(t)$ , de media cero, con dependencia de largo plazo en su autocorrelación. Esto es para  $H \in (0,5, 1)$ [55]

$$\mathbb{E}(X_1(t)X_1(t+s)) \sim \frac{1}{s^{2-2H}} \quad (2.12)$$

$$\mathbb{E}(X_2(t)X_2(t+s)) \sim \frac{1}{s^{2-2G}} \quad (2.13)$$

donde  $H$  y  $G$  son los exponentes de Hurst de  $X_1$  y  $X_2$ .

La covarianza se define por la siguiente relación, donde a priori se asume  $\gamma \in (0,5, 1)$ , y  $A$  y  $B$  son constantes

$$\mathbb{E}(X_1(t)X_2(t+s)) \sim \frac{A}{s^{2-2\gamma}} \quad (2.14)$$

$$\mathbb{E}(X_2(t)X_1(t+s)) \sim \frac{B}{s^{2-2\gamma}} \quad (2.15)$$

En general, para ruido Gaussiano fraccional, se debería cumplir que  $\gamma = \frac{H+G}{2}$ , pero en general,  $\gamma \leq \frac{H+G}{2}$ . En el anexo se demuestra que esto no se cumple, dado que las series poseen naturaleza multifractal. De todas formas, sí se encuentra que la relación entre el coeficiente de correlación calculado  $H_{xy}(q)$  y el estimado por la relación anterior  $\gamma$  resulta más parecido ante pequeñas fluctuaciones.

Existen modelos que se enfocan en determinar los exponentes  $H$ ,  $G$  y  $\gamma$ . Como en muchas aplicaciones  $X_1$  y  $X_2$  pueden estar contaminadas con tendencias no estacionarias, los métodos sin tendencia, presentados en secciones posteriores, han recibido mucha atención. Mediante el *Detrended Fluctuation Analysis* (DFA) es posible estimar el coeficiente de Hurst, el cual mejora los problemas relacionados con la no estacionariedad de las series y sensibilidad a correlación de corto plazo, que llevan a estimadores sesgados.

En el DFA, primero se calcula el perfil, como  $x(t) = \sum_{i=1}^t X(i)$ , con lo cual es posible calcular los coeficientes del DFA, es decir, la varianza sin tendencia de  $x(t)$  sobre una ventana de largo  $n$ .

$$F_{DFA}^2(n) = \frac{1}{\lfloor T/n \rfloor \cdot n} \sum_j (x_n^{(j)} - P_d(x_n^{(j)}))^T (x_n^{(j)} - P_d(x_n^{(j)})) \quad (2.16)$$

$T$  es el número total de datos,  $x_n^{(j)}$  es el vector de elementos de  $x(t)$  que caen dentro de la  $j$ -ésima ventana de largo  $n$ , y  $P_d(x_n^{(j)})$  es ajuste polinomial de mínimos cuadrados de orden  $d$  a  $x_n^{(j)}$ .  $H$  está dado por el ajuste lineal de  $\log(n)$  vs.  $\log(\sqrt{F_{DFA}^2(n)})$ . El grado según el cual  $X(t)$  está autocorrelacionado según una ley de potencia se cuantifica por la calidad del ajuste lineal.

El Detrended Cross-Correlation Analysis (DCCA) generaliza el método DFA a pares de series de tiempo, estimando  $\gamma$ , y cuantificando el escalamiento de las correlaciones. Dichos coeficientes están dados por

$$F_{DCCA}^2(n) = \frac{1}{\lfloor N/n \rfloor \cdot n} \sum_j ((x_1)_n^{(j)} - P_d((x_1)_n^{(j)}))^T ((x_2)_n^{(j)} - P_d((x_2)_n^{(j)})) \quad (2.17)$$

$$:= \frac{1}{\lfloor N/n \rfloor \cdot n} \sum_j F_{j,DCCA}(n) \quad (2.18)$$

Al igual que en el caso anterior,  $\gamma$  está dado por el ajuste de mínimos cuadrados de  $\log(n)$  vs.  $\log(\sqrt{F_{DCCA}^2(n)})$ .

Si bien existen otros métodos para estimar  $\gamma$ , como la estimación del periodograma promedio [56], el análisis *height cross-correlation* [57] y enfoques basados en promedios móviles [58], el enfoque de esta tesis se basa en el DCCA debido a la extensión de su uso y su demostrada efectividad y complementariedad con pruebas estadísticas.

En secciones posteriores se explicará a cabalidad el método de estimación de  $\gamma$ , extendiendo el análisis al caso de correlación multifractal.

La separación entre los dos tipos de correlación mencionados anteriormente es estricta. La correlación cruzada de corto plazo posee una función de correlación con decaimiento exponencial que se va a cero relativamente rápido. En cambio la correlación con memoria larga se caracteriza por un decaimiento asintótico según una ley potencial de la función de correlación.

El decaimiento exponencial es caracterizado por correlaciones significativas en rezagos pequeños, pero que rápidamente se van a cero. Por otro lado, el decaimiento según ley potencial presenta valores altos en la función de correlación incluso para rezagos bastante grandes.

Si la correlación es positiva, se llama al proceso positivamente correlacionado con memoria larga o positivamente persistente.

## 2.3. Multifractalidad en series financieras

La descripción probabilística de precios financieros, propuesto por Bachelier (1900) [59], inicialmente se enfocó en cambios en precios independientes y normalmente distribuidos. Economistas financieros han reconocido grandes discrepancias entre el modelo de Bachelier y la data financiera real. Primero, las series financieras muestran dependencia temporal en la alternación de períodos de grandes cambios de precios, con períodos de menores cambios. Segundo, las colas de los histogramas de la data observada son típicamente mucho mas gordas que las predichas por la distribución Gaussiana.

Las extensiones ARCH y GARCH (explicados en profundidad en secciones posteriores), se han vuelto el modo de pensamiento predominante en la representación estadística de las series de precios. Para un  $GARCH(p, q)$  original, las innovaciones en los retornos se especifican como

$$\varepsilon_t = u_t h_t^{1/2} \quad (2.19)$$

donde  $\{u_t\}$  es i.i.d. y

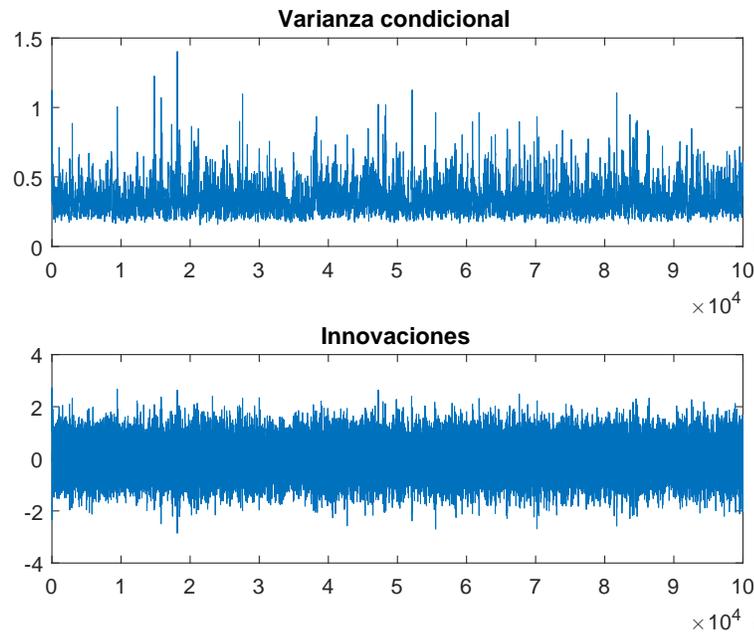
$$h_t = \alpha_0 + \sum_{j=1}^q \beta_j h_{t-j} + \sum_{i=1}^p \alpha_i \varepsilon_{t-i}^2 \quad (2.20)$$

Entre las extensiones a esta especificación, Nelson [60] introduce una respuesta asimétrica a los shocks positivos y negativos. Otros trabajos agregan una componente estocástica a la volatilidad. Además, algunos modelos relajan el supuesto de i.i.d para  $\{u_t\}$ , y solo asumen estacionariedad.

Usualmente para los modelos GARCH, la distribución condicional de los retornos tiene segundo momento finito y variante con el tiempo. Esto tiene directa relación con los clústers de volatilidad y permite abordarlo de forma tal que puede mitigar el problema de colas pesadas.

Como es posible observar en la Figura (2.2), al realizar simulaciones con modelos GARCH, para el ojo desnudo los procesos que no consideran memoria larga parecen ser ruidos blancos observados a distancia.

En teoría, los modelos GARCH permiten agregar parámetros de orden superior para capturar ciclos de corta frecuencia. En la práctica, representar fenómenos de memoria larga con procesos no integrados lleva a representaciones poco robustas, cuyo número de parámetros crece con el número de observaciones, por lo que resulta poco atractivo.



**Figura 2.2:** Simulación de proceso GARCH con  $c = 0,01$ ,  $\alpha_1 = 0,05$  y  $\beta_1 = 0,92$ . El código de Matlab puede ser encontrado en el Anexo.

(Fuente: Elaboración Propia)

### 2.3.1. Definición matemática de la multifractalidad

Los procesos usuales de incrementos independientes no reflejan una característica principal de los mercados financieros: la alternación entre períodos de altos cambios y pequeños cambios [61]. En cambio el mBf resulta útil para modelar la tendencia de los cambios de precios para ser seguidos por cambios de la misma (u opuesta) dirección. Sin embargo, estos modelos no capturan ni las colas pesadas ni las fluctuaciones en volatilidad. Ambos modelos presentan una forma bastante fuerte de invarianza a escala, donde la distribución de los retornos en diferentes submuestras son idénticos en general. Ello contrasta con las observaciones empíricas. En particular, muchos conjuntos de datos se vuelven más delgadas en sus colas a medida que aumenta el intervalo de muestreo, pero sin necesariamente volverse Gaussianas con intervalos largos de muestreo. De hecho, los procesos multifractales no suelen convertirse nunca en Gaussianos.

Como fue definido previamente, los procesos autoafines son los que satisfacen la siguiente regla

$$X(ct) \stackrel{d}{=} c^H X(t) \quad (2.21)$$

Ante lo cual la teoría multifractal se presenta con la relación mas general

$$X(ct) \stackrel{d}{=} M(c)X(t) \quad (2.22)$$

donde  $X$  y  $M$  son funciones independientes y aleatorias. Bajo estacionariedad estricta, las translaciones arbitrarias a través del eje del tiempo permiten la extensión de la ecuación anterior a reglas de escala locales

$$X(t + c\Delta t) - X(t) \stackrel{d}{=} M(c)[X(t + \Delta t) - X(t)] \quad (2.23)$$

para cualquier  $c > 0$ . El factor de escala  $M(c)$  es una variable aleatoria, cuya distribución no depende el instante particular  $t$ . Los procesos autosimilares satisfacen (2.22), con  $M(c) = c^H$ . Siguiendo con la analogía, definimos el índice generalizado  $H(c) = \log_c M(c)$ , y se reescribe la relación anterior

$$X(ct) \stackrel{d}{=} c^{H(c)}X(t) \quad (2.24)$$

En contraste con los procesos auto-similares, el índice  $H(c)$  es una función aleatoria de  $c$ . Esto implica que la multifractalidad permite una mayor variedad de comportamientos que es posible bajo autoafinidad. Ello tambien impone fuertes restricciones en la distribución del proceso. Por ejemplo, si  $c_2/c_1 = c_3/c_2$ , y la condición (2.22) se sostiene, entonces

$$\frac{X(c_2t)}{X(c_1t)} \stackrel{d}{=} \frac{X(c_3t)}{X(c_2t)} \quad (2.25)$$

dado que ambos ratios se distribuyen como  $M(c_2/c_1)$ .

Además se requiere que los factores de escala aleatorios satisfagan la propiedad de  $M(ab) \stackrel{d}{=} M_1(a)M_2(b)$ , donde  $M_1$  y  $M_2$  son copias independientes de  $M$ . Esta condición, implica la siguiente regla de escalamiento

$$\mathbb{E}(|X(t)|^q) = c(q)t^{\tau(q)+1} \quad (2.26)$$

donde  $\tau(q)$  y  $c(q)$  son funciones deterministas de  $q$ .

La regla de escalamiento en (2.26) es la propiedad definitoria de los procesos multifractales. En este contexto, la condición (2.21) solo caracteriza una clase particular de multifractales. Esto lleva a la principal característica del análisis, **la Multifractalidad se define como una propiedad global de los momentos del proceso.**

La mayoría del trabajo multifractal se concentra en el análisis de la función de escala, o  $\tau(q)$ . Con un valor de  $q = 0$ , el intercepto de dicha función estará dado por  $\tau(0) = -1$ .

Un proceso autoafín con índice  $H$  se considera multifractal con función de escala  $\tau(q) = qH - 1$ , que dada su linealidad, la función de escala estará dada por un solo coeficiente, su pendiente. Esto implica un proceso uniescala, o monofractal. Los procesos Multifractales (o multiescala), permiten funciones cóncavas de escala más generales.

### 2.3.2. Procesos Multifractales

Resulta conveniente definir los procesos multifractales en términos de los momentos, debido a que ello presenta implicaciones gráficas y demostrables. La idea más simple, e intuitiva es la concentración en la escala de los momentos de los incrementos.

Se define un proceso estocástico multifractal  $\{X(t)\}$  como aquel que posee incrementos estacionarios y satisface

$$\mathbb{E}(|X(t)|^q) = c(q)t^{\tau(q)+1}, \quad \forall t \in T, q \in Q \quad (2.27)$$

donde  $T$  y  $Q$  son intervalos en el eje real.  $\tau(q)$  y  $c(q)$  son funciones de dominio  $Q$ . Además,  $T$  y  $Q$  tienen largos positivos, y  $0 \in T, [0, 1] \subseteq Q$ .

Por ende, un proceso multifractal es globalmente escalable, en el sentido de que sus momentos satisfacen la relación de escala. La función  $\tau(q)$  es llamada función de escala del proceso multifractal.

Los procesos autoafines son multifractales. Teniendo un proceso autoafín  $\{X(t), t \geq 0\}$ , con

índice de autoafinidad  $H$ , satisface  $X(t) \stackrel{d}{=} t^H X(1)$ , y por ello satisface

$$\mathbb{E}(|X(t)|^q) = t^{Hq} \mathbb{E}(|X(1)|^q) \quad (2.28)$$

Con lo que la condición (2.27) se sostiene si se cumple

$$\tau(q) = Hq - 1 \quad (2.29)$$

$$c(q) = \mathbb{E}(|X(1)|^q) \quad (2.30)$$

Como se mencionó antes, para procesos monofractales la función de escala  $\tau(q)$  es lineal y completamente determinada por su índice  $H$ , limitándose su comportamiento por un único parámetro de pendiente. Al ser  $\tau(q)$  no lineal, es posible atribuir dicho comportamiento al valor de  $H$ , quien también dependerá del orden  $q$ , por lo que se suele reescribir la relación como

$$\tau(q) = qH(q) - 1 \quad (2.31)$$

$H(q)$  recibe el nombre de exponente de Hurst generalizado. Y en lo que sigue, el enfoque estará en la determinación de este exponente de Hurst generalizado. Sin embargo, si se busca ampliar el análisis, es posible definir un exponente de Hurst local, directamente de los valores de la función de fluctuación. Esto se debe a que el exponente de Hurst tendrá cierta variación en el tiempo para los procesos multifractales. Para lo anterior se utiliza el espectro multifractal, el cual refleja la variación temporal del exponente de Hurst local, es decir, la variación temporal de estructura invariante de cierta escala de la serie temporal. Si bien esto puede realizarse cambiando el análisis desde el estudio de segmentos sin superposición (lo cual se estudia en secciones posteriores) a uno con superposición, es posible realizar una buena aproximación mediante el cálculo del exponente de Holder.

Mientras mayores sean los cambios en los períodos de pequeñas y grandes fluctuaciones, mayor será el espectro. En términos simples, el exponente de Holder describe el nivel de continuidad de una curva. Si  $\alpha > 0$  la serie es continua, y si  $\alpha = 1$  la serie cumple la condición de Lipschitz, que la hace más cercana a una función continuamente diferenciable. El espectro, por tanto, describirá de otra forma la rugosidad de la serie.

La determinación del exponente de Holder y de la función que lo describe, se puede realizar mediante la transformada de Legendre

$$\alpha = \frac{d\tau(q)}{dq} = H(q) + qH'(q) \quad (2.32)$$

$$f(\alpha) = q(\alpha - H(q)) + 1 \quad (2.33)$$

El exponente de Holder se utiliza para caracterizar la fuerza de la singularidad en una serie de tiempo, y de las expresiones anteriores es posible observar que para series monofractales, el ancho del espectro será cercano a cero debido a que  $H(q)$  será constante.

Para una serie multifractal, la parábola será inversa, aproximándose a una distribución de probabilidad.

## 2.4. Obtención del índice de Hurst

Como se mencionó anteriormente, el primero que estudió series fractales fue Harold Edwin Hurst (1880-1978), quien era constructor de represas en los primeros años del siglo 20, y cuyo trabajo durante un buen tiempo se desarrolló en el río Nilo. Al momento de diseñar la represa, se enfrentó a un gran problema, recurrente en hidrología, que consistía en determinar la capacidad de almacenamiento dependiente del flujo que entraba al río a causa de lluvias o riachuelos, así como el flujo de salida del río, que estaba principalmente enfocado al riego.

Usualmente se asumía que el flujo era un proceso aleatorio, que dada la complejidad y gran número de componentes parecía ser razonable. Pero para Hurst al parecer esto no era explicación suficiente, por lo que pidió registros históricos del Nilo a lo largo de cerca de mil años. Al hacer esto observó que flujos grandes, eran por lo general seguidos de flujos grandes, y a su vez, flujos pequeños eran seguidos por flujos incluso más pequeños. Si bien podía pensarse que esto ocurría por ciclos, no fue posible demostrar una longitud periódica de los mismos. De hecho los análisis estándar no revelaban correlación significativa, lo que llevó a Hurst a desarrollar su propia metodología. H. E. Hurst, desarrolló el estadístico de Rango Reescalado, o estadístico R/S. Su razonamiento fue que si se tiene un proceso

$X_i, i = 1, \dots, n$ , i.i.d. normal, entonces el estadístico  $R/S$  debería crecer en el orden de  $\sqrt{n}$ , donde  $n$  era la escala de análisis, relacionando este fenómeno con la geometría fractal. Sin embargo, Hurst encontró que para la data del río Nilo, el estadístico  $R/S$  se incrementaba no en el orden de  $\sqrt{n}$ , sino que en el orden  $n^H$ , donde  $H$  se denominó como el índice de Hurst, y para la data del río Nilo, tomaba un valor entre 0,75 y 0,8, H.E. Hurst descubrió que este fenómeno aparecía en numerosos otros conjuntos de datos, lo cual se denominó el *fenómeno de Hurst*.

La aproximación se realizaba mediante la siguiente regresión lineal

$$\ln(R/S)_n = \ln(c) + H \cdot \ln(n) \quad (2.34)$$

Es posible generar 3 grandes casos para la interpretación del Hurst

- Si  $H = 0,5$  se tiene un proceso independiente. El análisis  $R/S$  no requiere que el proceso estudiado distribuya de forma normal, solo que sea independiente, y el análisis no es paramétrico. La hipótesis nula es que el comportamiento del fenómeno es una caminata aleatoria o movimiento Browniano discretizado.
- Si  $0,5 < H \leq 1$ , se tiene una serie de tiempo persistente, es decir hay efectos de memoria de largo plazo. Si por ejemplo  $H = 0,7$ , se tiene un 70 % de probabilidades de que el siguiente período siga la misma tendencia que el anterior. Teóricamente lo que suceda en  $t$  afectará en un horizonte de tiempo que tiende a infinito, debido a las características caóticas que generan las condiciones iniciales. Usualmente, las series tienden a ser persistentes en la naturaleza, cosa que también ocurre en las series económicas.
- Si  $0 \leq H < 0,5$ , se tiene anti-persistencia. Un sistema así cubre menos distancia que uno aleatorio, en el caso de una partícula errática. Se suele asemejar a un sistema de reversión a la media.
- Es posible determinar la dimensión fractal  $D$  mediante la ecuación  $D = 2 - H$ .

El método consiste, a grandes rasgos, en dividir el conjunto o la serie original en subconjuntos más pequeños de igual tamaño que se analizarán por separado. Se calcula el

valor esperado para cada submuestra y se calcula la diferencia de cada elemento con la media, para posteriormente construir un perfil que vaya sumando dichas diferencias. En las submuestras se calculan los valores máximos y mínimos, y luego la diferencia entre ambos, es decir, el rango. Esta diferencia se compara con la desviación estándar muestral para estimar el rango reescalado  $(R/S)_n$  para diferentes tamaños de subconjuntos de  $n$ , y mediante la regresión descrita anteriormente se puede obtener el coeficiente de Hurst.

El estudio realizado por H.E. Hurst [2, 3] pretendía determinar la capacidad de reserva que se necesitaría para desarrollar el riego a través del Nilo a capacidad máxima. Básicamente se dedicó a resolver el siguiente problema.

En años de alto flujo, el agua del Nilo no se utilizaba de forma óptima, pero en los años de bajo flujo existía una baja en la cantidad de agua. Por ello, para el uso eficiente de los recursos hídricos, se pretendía buscar un flujo anual de salida constante. La pregunta a responder fue: ¿Cuál es el óptimo flujo anual de salida y cual es la capacidad de almacenamiento necesario para mantenerlo?. La solución implicó el desarrollo del Rango Reescalado.

### 2.4.1. Análisis del Rango Reescalado

Suponiendo que  $X_i$  es el flujo de salida total de agua desde la fuente del Nilo en el año  $i$ ,  $i = 1, \dots, N$ . Para estimar el exponente de Hurst, primero se debe estimar la dependencia del Rango Reescalado en una escala de tiempo  $s$  de la observación. Una serie con largo  $N$  se divide en subseries de tiempo sin superposición más cortas de largos  $s$ , con  $s = \{N, N/2, N/4, \dots\}$ , y para cada subserie se calcula el estadístico  $R/S(s)$ , que depende del tamaño de la subserie, y es el cociente entre el rango de variación ( $R$ ) y su desviación típica ( $S$ ).

#### Procedimiento para cada valor de $s$ :

1. Con las variaciones se genera serie de tiempo  $x_t$ ,  $t = 1, \dots, s$
2. Se calcula el perfil como sigue para cada subserie de largo  $s$

$$Y(t) = \sum_{j=1}^t (x_j - \bar{x}) \quad (2.35)$$

3. Se calcula el rango de cada submuestra:

$$R(s) = \max\{Y(t), t = 1, \dots, s\} - \min\{Y(t), t = 1, \dots, s\} \quad (2.36)$$

4. Se calculan las desviaciones estándar

$$S(s) = \sqrt{\frac{1}{s} \sum_{j=1}^s (x_j - \bar{x})^2} \quad (2.37)$$

5. Se calcula el promedio del rango reescalado  $R(s)/S(s)$  para las subseries de tamaño  $s$ .

6. Se ajusta una regresión lineal log log a la expresión

$$\mathbb{E}(R(s)/S(s)) = (R/S)_s = c \cdot s^H \quad (2.38)$$

quedando

$$\log((R/S)_s) \text{ vs. } \log(s) \quad (2.39)$$

La pendiente del ajuste lineal corresponde al coeficiente de Hurst.

## 2.5. Multifractal Detrended Fluctuation Analysis

El MF-DFA [17] es la extensión multifractal del Detrended Fluctuation Analysis [14], un método introducido en 1994 como una alternativa al Rango Reescalado, y que ha obtenido gran aceptación gracias a la característica de ser aplicable a series no estacionarias. Asumiendo una serie de tiempo  $x_i$ ,  $i = 1, \dots, N$ , donde  $N$  es el largo de las series, el método consiste los siguientes pasos

- **Paso 1:** Se construye el perfil

$$X(i) = \sum_{t=1}^i (x_t - \bar{x}), \quad i = 1, \dots, N \quad (2.40)$$

Donde  $\bar{x}$  representa el promedio de la serie para todo el período.

- **Paso 2:** Se divide el perfil  $X(i)$  en  $N_s \equiv [N/s]$  ventanas sin superposición de igual largo  $s$ . Dado que el largo de la longitud de la serie  $N$  no es necesariamente un múltiplo de la escala de tiempo  $s$ , una parte del perfil puede quedar al final. Para no descartar esta parte, el mismo procedimiento se aplica empezando desde el final de la serie. Esto implica que se obtienen  $2N_s$  segmentos.
- **Step 3:** La tendencia,  $\hat{X}_v(i)$  para cada uno de los  $2N_s$  segmentos son estimados por medio de una regresión lineal o polinomial. Esto precede la determinación de la varianza sin tendencia, la cual se calcula como

$$F^2(v, s) = \frac{1}{s} \sum_{i=1}^s (X((v-1)s + i) - \hat{X}_v(i))^2 \quad (2.41)$$

para cada segmento  $v$ ,  $v = 1, \dots, N_s$ , y

$$F^2(v, s) = \frac{1}{s} \sum_{i=1}^s (X(N - (v - N_s)s + i) - \hat{X}_v(i))^2 \quad (2.42)$$

para cada segmento  $v$ ,  $v = N_s + 1, \dots, 2N_s$ .

Dependiendo del orden de este polinomio es que se denominan DFA1, DFA2, DFA3, etc.

Considerando que se elimina la tendencia mediante la sustracción del ajuste polinomial en el perfil, los diferentes órdenes del polinomio difieren en su capacidad de eliminar las tendencias en la serie. En un ajuste MF-DFA $m$  (m-ésimo orden del MF-DFA) las tendencias de orden  $m$  en el perfil (o de orden  $m - 1$  en la serie original) son eliminadas. Por ello es que una comparación de resultados para diferentes órdenes del MF-DFA permiten estimar el tipo de tendencia polinomial en la serie de tiempo analizada.

- **Paso 4:** Por medio del promedio sobre todos los segmentos, la función del  $q$ -ésimo orden de fluctuación se puede obtener para  $q \neq 0$

$$F_q(s) = \left( \frac{1}{2N_s} \sum_{v=1}^{2N_s} (F^2(v, s))^{q/2} \right)^{1/q} \quad (2.43)$$

El valor de  $H(0)$ , el cual corresponde al límite  $H(q)$  para  $q \rightarrow 0$ , no puede ser determinado usando directamente el procedimiento de la ecuación (2.43) debido al exponente divergente. Para solucionar esto es necesario emplear un procedimiento logarítmico

$$F_q(s) = \exp \left( \frac{1}{4N_s} \sum_{v=1}^{2N_s} \ln[F^2(v, s)] \right) \quad (2.44)$$

Donde, en general,  $q$  puede tomar cualquier valor real (para  $q = 0$ , ver el **Paso 5**). Para  $q = 2$  se tiene el DFA clásico.

En este paso el interés radica en cómo las funciones de fluctuación  $F_q(s)$  dependen de la escala de tiempo  $s$  para los distintos valores de  $q$ . Esto implica que es necesario repetir los pasos 2 a 4 para distintas escalas de tiempo. Resulta aparente que  $F_q(s)$  aumenta con incrementos de  $s$ . Evidentemente,  $F_q(s)$  depende del orden  $m$  del DFA, y sólo está definida para  $s \geq m + 2$ .

- **Paso 5:** El comportamiento de escala de las fluctuaciones se analiza observando los gráficos log-log de  $F_q(s)$  versus  $s$  para cada valor de  $q$ . En el caso donde la series posea autocorrelación de memoria larga,  $F_q(s)$  se incrementará con  $s$ , de acuerdo a

una ley de potencia.

$$F_q(s) \sim s^{H(q)} \quad (2.45)$$

Para series monofractales,  $H(q)$  no depende de  $q$ , ya que el comportamiento de escala  $F_s^2(v, s)$  es idéntico para todos los segmentos y la ecuación (2.43) entrega el mismo comportamiento para todos los valores de  $q$ . Solo si pequeñas o grandes fluctuaciones se escalan de forma diferente existirá dependencia significativa de  $H(q)$  en  $q$ .

Para escalas grandes,  $s > N/4$ ,  $F_q(s)$  se vuelve estadísticamente poco confiable debido a que el número de segmentos  $N_s$  para el procedimiento en el **Paso 4** se vuelve muy pequeño. Por ello se suelen excluir escalas donde  $s > N/4$  en el procedimiento de ajuste para determinar  $H(q)$ .

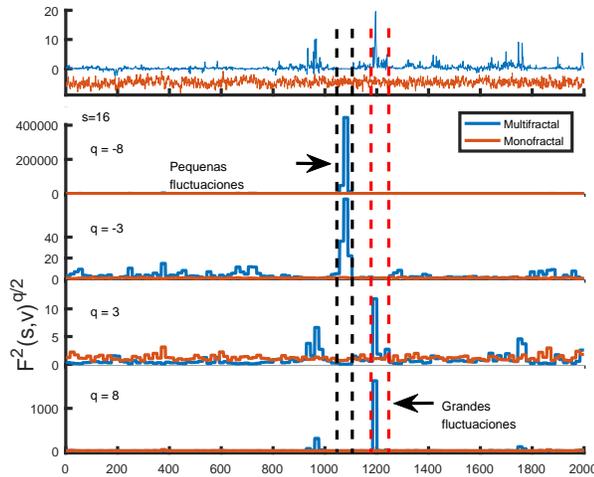
Aparte de eso, se presentan desviaciones sistemáticas (pero corregibles) del comportamiento de escala en la ecuación (2.45) para escalas muy pequeñas ( $s \approx 10$ ), por lo que para evitar problemas se utiliza una escala mínima de 16. En general, el exponente  $H(q)$  en la ecuación (2.45) puede depender de  $q$ . Para series estacionarias,  $H(2)$  es idéntico al exponente de Hurst, por ello es que se denomina a la función  $H(q)$  como exponente de Hurst generalizado.

Para valores positivos de  $q$ , los segmentos  $v$  con alta varianza  $F^2(v, s)$  (grandes desviaciones del ajuste correspondiente) dominarán el promedio de  $F_q(s)$ , ello implica que  $H(q)$  describe el comportamiento de escala de los segmentos con fluctuaciones altas.

Para valores negativos de  $q$ ,  $H(q)$  los segmentos  $v$  con varianza  $F^2(v, s)$  pequeña dominarán el promedio de  $F_q(s)$ , es decir,  $H(q)$  describirá el comportamiento de escala de segmentos con fluctuaciones pequeñas.

Lo anterior indica que mientras mayor es el rango del exponente de Hurst generalizado  $H(q)$  más complicada es la estructura de la serie.

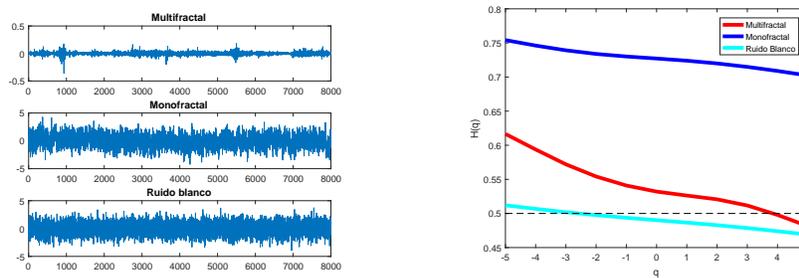
Para ejemplificar lo anterior, se grafica en la Figura (2.3) el valor de  $F^2(16, v)^{q/2}$ , para  $q = \{-1, -3, 1, 3\}$ .



**Figura 2.3:** Influencia del orden  $q$  de fluctuación. (Fuente: Elaboración Propia)

Como es posible observar, los períodos de pequeñas fluctuaciones serán ponderados de forma diferente por valores negativos de  $q$ , y para valores positivos de  $q$ . Ello permite estimar el comportamiento de escala de los diferentes niveles de fluctuación en la serie.

Para complementar el análisis, a continuación se presentan ejemplos de lo que es posible obtener con el algoritmo al considerar series conocida de 3 tipos: multifractal, mono fractal y ruido blanco.

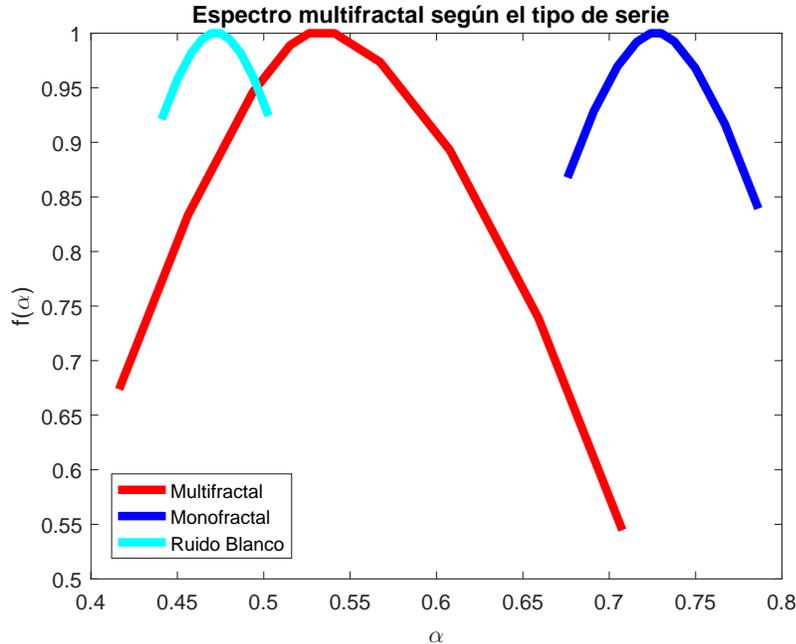


**Figura 2.4:** Tipo de serie y sus índices de Hurst generalizados.

(Fuente: Elaboración Propia)

Como se observa en la Figura (2.4), el ruido blanco, al ser completamente aleatorio, presenta  $H(q) \approx 0,5, \forall q$ , mientras que la serie monofractal presenta un comportamiento similar de  $H \approx 0,75, \forall q$ . Por otro lado, la serie multifractal presenta valores del Hurst generalizado bastante diferentes según el orden de la fluctuación,  $H(q_{min}) = 0,62$  y  $H(q_{max}) = 0,48$ . Esto indica el comportamiento multifractal de la serie.

Al estudiar el espectro multifractal, dado por los exponentes de Holder, es posible observar el grado de multifractalidad de otro modo



**Figura 2.5:** Espectro multifractal de series conocidas.

(Fuente: Elaboración Propia)

Como se observa, el espectro de la serie ruido blanco y el de la serie monofractal presenta pequeñas fluctuaciones alrededor de cierto exponente de Holder  $\alpha$ , indicando un comportamiento más homogéneo de la curva. Por otro lado la serie multifractal presenta una mayor gama de los mismos.

### 2.5.1. Grado de eficiencia del mercado

Para que se considere 'eficiente' un índice o precio de una acción, todos los tipos de fluctuaciones deben seguir un comportamiento de caminata aleatoria, lo cual se traduce en que los  $H(q)$ s relacionados con los diferentes  $q$  tiendan a un valor de 0,5. El análisis de este tipo se suele concentrar tanto en fluctuaciones pequeñas como grandes, pudiéndose definir las medidas que se presentan a continuación.

La primera es la más intuitiva, en un mercado perfectamente eficiente se obtendrá que

$ED = 0$ .

$$ED = |H(q = 2) - 0,5| \quad (2.46)$$

Sin embargo, esta medida es bastante ineficiente para series multifractales, debido a que no cubrirá las complejas relaciones de la autocorrelación. Con ello en mente, se plantea la siguiente medida, dada por la ecuación

$$DME = \frac{1}{q_{max} - q_{min} + 1} \sum_{q=q_{min}}^{q_{max}} |H(q) - 0,5| \quad (2.47)$$

Y se hablará de un mercado eficiente mientras más cercano a cero sea este valor.

Si el objetivo es enfocarse en la diferencia entre grandes y pequeñas fluctuaciones, se plantea la siguiente medida como opción. Esto debido a que cuando  $q < 0$  las pequeñas fluctuaciones se ven aumentadas en influencia, mientras que para  $q > 0$ , las grandes se ven aumentadas [62].

$$ME = \frac{1}{2}(|H(q_{min}) - 0,5| + |H(q_{max}) - 0,5|) \quad (2.48)$$

En dicha ecuación, los exponentes de escalamiento  $H(q_{min})$  y  $H(q_{max})$  son usados para denotar las características de persistencia de fluctuaciones pequeñas y altas respectivamente. Para hablar de 'eficiencia' de la acción o índice el valor  $ME$  debe ser cercano a cero.

### Causas de la multifractalidad en autocorrelación

Además de lo anterior es posible analizar el efecto del orden de las series y las colas pesadas en el nivel de multifractalidad. Ello se hace por separado de dos formas:

- Para el análisis del efecto de la correlación de largo plazo, se genera una serie aleatoria (*shuffled*) y se calcula  $h_{shuf}(q)$ . Aleatorizar los datos permite eliminar la influencia del tiempo sin destruir la distribución de la serie.

La influencia de la correlación se medirá entonces como:

$$h_{cor} = h(q) - h_{shuf}(q) \quad (2.49)$$

- Para el análisis del efecto de las colas pesadas, se genera una serie sustituta (*surrogate*) al aleatorizar sus fases en el espacio de Fourier, mediante el *Amplitude Adjusted Fourier Transform* (IAAFT) o el *Iterative Amplitude Adjusted Fourier Transform* (AAFT). Las nuevas series son Gaussianas y mantienen la autocorrelación que pueda existir, por lo que si la no-normalidad era la fuente de multifractalidad, estas series serían monofractales y sus exponentes de Hurst generalizados  $h_{sur}(q)$  serían constantes.

La influencia de la no normalidad de la distribución será medida como:

$$h_{PDF} = h(q) - h_{sur}(q) \quad (2.50)$$

Este análisis no se lleva a cabo en esta tesis.

## 2.6. Multifractal Detrended Cross-Correlation Analysis

El método *Multifractal Detrended Cross-Correlation Analysis* es la extensión multifractal del DCCA mencionado en secciones anteriores [34]. Hereda el gran desempeño incluso con series altamente no estacionarias, y sus pasos son similares a los del MF-DFA. Asumiendo dos series de tiempo  $x_i$  e  $y_i$ ,  $i = 1, \dots, N$ , donde  $N$  es el largo de las series, el método sigue los siguientes pasos

- **Paso 1:** Se construye el perfil

$$X(i) = \sum_{t=1}^i (x_t - \bar{x}), \quad Y(i) = \sum_{t=1}^i (y_t - \bar{y}), \quad i = 1, \dots, N \quad (2.51)$$

Donde  $\bar{x}$  y  $\bar{y}$  representan el promedio de las series para todo el período.

- **Paso 2:** Se divide el perfil  $X(i)$  y  $Y(i)$  en  $N_s \equiv [N/s]$  ventanas sin superposición de igual largo  $s$ . Dado que el largo de la longitud de la serie  $N$  no es necesariamente un múltiplo de la escala de tiempo  $s$ , una parte del perfil puede quedar al final. Para no descartar esta parte, el mismo procedimiento se aplica empezando desde el final de la serie. Esto implica que se obtienen  $2N_s$  segmentos.
- **Step 3:** Las tendencias locales,  $\hat{X}_v(i)$  y  $\hat{Y}_v(i)$  para cada uno de los  $2N_s$  segmentos son estimados por medio de una regresión lineal o polinomial. Esto precede la determinación de la covarianza sin tendencia, la cual se calcula como

$$F(v, s) = \frac{1}{s} \sum_{i=1}^s |X((v-1)s+i) - \hat{X}_v(i)| \cdot |Y((v-1)s+i) - \hat{Y}_v(i)| \quad (2.52)$$

para cada segmento  $v$ ,  $v = 1, \dots, N_s$ , y

$$F(v, s) = \frac{1}{s} \sum_{i=1}^s |X(N - (v - N_s)s + i) - \hat{X}_v(i)| \cdot |Y(N - (v - N_s)s + i) - \hat{Y}_v(i)| \quad (2.53)$$

para cada segmento  $v$ ,  $v = N_s + 1, \dots, 2N_s$ .

- **Paso 4:** Por medio del promedio sobre todos los segmentos, la función del  $q$ -ésimo orden de fluctuación se puede obtener como sigue para  $q \neq 0$

$$F_q(s) = \left( \frac{1}{2N_s} \sum_{v=1}^{2N_s} [F(v, s)]^{q/2} \right)^{1/q} \quad (2.54)$$

y como sigue para  $q = 0$

$$F_q(s) = \exp \left( \frac{1}{4N_s} \sum_{v=1}^{2N_s} \ln[F(v, s)] \right) \quad (2.55)$$

Para el caso de  $q = 2$ ,  $F_q(s)$  equivale al  $F_{DCCA}(s)$  utilizado en los estadísticos de prueba anteriores.

- **Paso 5:** El comportamiento de escala de las fluctuaciones se analiza observando los gráficos log-log de  $F_q(s)$  versus  $s$  para cada valor de  $q$ . En el caso donde las dos series poseen correlación de memoria larga,  $F_q(s)$  se incrementará con  $s$ , de acuerdo a una ley de potencia.

$$F_q(s) \sim s^{H_{xy}(q)} \quad (2.56)$$

El exponente de escala  $H_{xy}(q)$  es conocido como el exponente de correlación generalizado y describe la relación de acuerdo a una ley de potencia entre las dos series. Puede ser obtenida calculando la pendiente de los graficos log-log de  $F_q(s)$  versus  $s$ , a través de Mínimos Cuadrados Ordinarios (MCO).

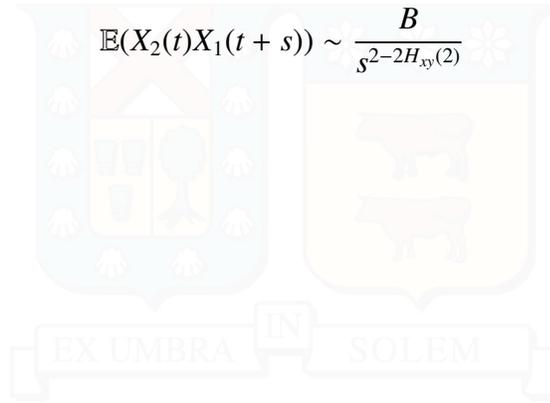
En el caso de  $q = 2$ , el exponente de correlación generalizado tiene propiedades e interpretación similar a la del Coeficiente de Hurst univariado que es posible calcular mediante el Rango Reescalado o el DFA.

Si  $H_{xy}(2) > 0,5$ , la serie posee correlación persistente, es decir, un cambio positivo (negativo) en una serie, será más probablemente seguido por un valor positivo (negativo). En el caso en que  $H_{xy}(2) < 0,5$  las series poseen correlación antipersistente, lo cual significa que un cambio positivo (negativo) será más probablemente seguido por un cambio negativo (positivo) en el otro precio. Para  $H_{xy}(2) = 0,5$  solo correlaciones de corto plazo (o ningún tipo de correlación) están presentes en la relación entre las series.

Esto queda más claro cuando se considera la covarianza mencionada al analizar la correlación con memoria larga, por ejemplo

$$\mathbb{E}(X_1(t)X_2(t+s)) \sim \frac{A}{s^{2-2H_{xy}(2)}} \quad (2.57)$$

$$\mathbb{E}(X_2(t)X_1(t+s)) \sim \frac{B}{s^{2-2H_{xy}(2)}} \quad (2.58)$$



## 2.7. Multifractal Asymmetric Detrended Cross-Correlation Analysis

El método *Multifractal Asymmetric Detrended Cross-Correlation Analysis* [44] se basa en el A-DFA y el MF-DCCA, y permite examinar las características asimétricas de dos series de tiempo correlacionadas. Este método logra gran desempeño incluso con series altamente no estacionarias, y sus pasos son similares a los pasos en el MF-DCCA.

Asumiendo dos series de tiempo  $x_i$  e  $y_i$ ,  $i = 1, \dots, N$ , donde  $N$  es el largo de las series, el método sigue los siguientes pasos

- **Paso 1:** Se construye el perfil

$$X(i) = \sum_{t=1}^i (x_t - \bar{x}), \quad Y(i) = \sum_{t=1}^i (y_t - \bar{y}), \quad i = 1, \dots, N \quad (2.59)$$

Donde  $\bar{x}$  y  $\bar{y}$  representan el promedio de la serie para todo el período.

- **Paso 2:** Se divide el perfil  $X(i)$  y  $Y(i)$  en  $N_s \equiv [N/s]$  ventanas sin overlapping de igual largo  $s$ . Dado que el largo de la longitud de la serie  $N$  no es necesariamente un múltiplo de la escala de tiempo  $s$ , una parte del perfil puede quedar al final. Para no descartar esta parte, el mismo procedimiento se aplica empezando desde el final de la serie. Esto implica que se obtienen  $2N_s$  segmentos.
- **Step 3:** Las tendencias,  $X^v(i)$  y  $Y^v(i)$  para cada uno de los  $2N_s$  segmentos son estimados por medio de una regresión lineal como  $X^v(i) = a_{X^v} + b_{X^v} \cdot i$  y  $Y^v(i) = a_{Y^v} + b_{Y^v} \cdot i$ . Esto precede la determinación de la covarianza sin tendencia, la cual se calcula como

$$F(v, s) = \frac{1}{s} \sum_{i=1}^s |X((v-1)s+i) - X^v(i)| \cdot |Y((v-1)s+i) - Y^v(i)| \quad (2.60)$$

para cada segmento  $v$ ,  $v = 1, \dots, N_s$ , y

$$F(v, s) = \frac{1}{s} \sum_{i=1}^s |X(N - (v - N_s)s + i) - X^v(i)| \cdot |Y(N - (v - N_s)s + i) - Y^v(i)| \quad (2.61)$$

para cada segmento  $v$ ,  $v = N_s + 1, \dots, 2N_s$ .

- **Paso 4:** Por medio del promedio sobre todos los segmentos, la función del  $q$ -ésimo orden de fluctuación se puede obtener como sigue para diferentes comportamientos de la tendencia de la serie de tiempo  $x_t$

$$F_q^+(s) = \left( \frac{1}{M^+} \sum_{v=1}^{2N_s} \frac{\text{sign}(b_{X^v}) + 1}{2} [F(v, s)]^{q/2} \right)^{1/q} \quad (2.62)$$

$$F_q^-(s) = \left( \frac{1}{M^-} \sum_{v=1}^{2N_s} \frac{-[\text{sign}(b_{X^v}) - 1]}{2} [F(v, s)]^{q/2} \right)^{1/q} \quad (2.63)$$

cuando  $q \neq 0$ , y

$$F_0^+(s) = \exp \left( \frac{1}{2M^+} \sum_{v=1}^{2N_s} \frac{\text{sign}(b_{X^v}) + 1}{2} [F(v, s)]^{q/2} \right)^{1/q} \quad (2.64)$$

$$F_0^-(s) = \exp \left( \frac{1}{2M^-} \sum_{v=1}^{2N_s} \frac{-[\text{sign}(b_{X^v}) - 1]}{2} [F(v, s)]^{q/2} \right)^{1/q} \quad (2.65)$$

para  $q = 0$ .  $M^+ = \sum_{v=1}^{2N_s} \frac{\text{sign}(b_{X^v}) + 1}{2}$  y  $M^- = \sum_{v=1}^{2N_s} \frac{-[\text{sign}(b_{X^v}) - 1]}{2}$  son el número de series con tendencias positivas y negativas, respectivamente. Se asume que  $b_{X^v} \neq 0$  para  $v = 1, \dots, 2N_s$ , de forma tal que  $M^+ + M^- = 2N_s$ .

EL MF-DCCA tradicional se implementa calculando la función del  $q$ -ésimo orden de fluctuación para  $q \neq 0$

$$F_q(s) = \left( \frac{1}{2N_s} \sum_{v=1}^{2N_s} [F(v, s)]^{q/2} \right)^{1/q} \quad (2.66)$$

y como sigue para  $q = 0$

$$F_q(s) = \exp \left( \frac{1}{4N_s} \sum_{v=1}^{2N_s} \ln[F(v, s)] \right) \quad (2.67)$$

- **Paso 5:** El comportamiento de escala de las fluctuaciones se analiza observando los gráficos log-log de  $F_q(s)$  versus  $s$  para cada valor de  $q$ . En el caso donde las dos series poseen correlación de memoria larga,  $F_q(s)$  se incrementará con  $s$ , de acuerdo a una ley de potencia.

$$F_q(s) \sim s^{H_{xy}(q)} \quad (2.68)$$

$$F_q^+(s) \sim s^{H_{xy}^+(q)} \quad (2.69)$$

$$F_q^-(s) \sim s^{H_{xy}^-(q)} \quad (2.70)$$

El coeficiente  $H_{xy}$  tiene las mismas interpretaciones que las mencionadas en el MF-DCCA. La diferencia radica en que es posible obtener diferentes exponentes ante diferentes tendencias.

Para medir el grado de asimetría de la correlación se puede plantear, para cada  $q$ , la siguiente métrica.

$$\Delta H_{xy}(q) = H_{xy}^+(q) - H_{xy}^-(q) \quad (2.71)$$

Mientras mayor sea el valor absoluto, mayor es el comportamiento asimétrico. Si  $\Delta H_{xy}(q)$  es igual o cercano a cero, entonces la correlación es simétrica en el orden  $q$  para diferentes tendencias de la serie de tiempo  $x_t$ . Si el valor de  $\Delta H_{xy}(q)$  es positivo, significa que el exponente de correlación es mayor cuando la serie temporal  $x_t$  tiene una tendencia positiva que cuando es negativa. Por otro lado, si es negativa, el exponente de correlación es menor cuando la serie de tiempo  $x_t$  tiene tendencia positiva que cuando es negativa.

Si el valor del exponente de correlación generalizado  $H_{xy}(q)$  depende del valor de  $q$ , la correlación entre las dos series de tiempo es multifractal. Al igual que en el MF-DCCA, para valores  $q > 0$ ,  $H_{xy}(q)$ ,  $H_{xy}^+(q)$  y  $H_{xy}^-(q)$  describen el comportamiento de escala de las grandes fluctuaciones, mientras que para  $q < 0$ , describe el comportamiento de escala para pequeñas fluctuaciones. En este trabajo, se estudia el comportamiento multifractal para diferentes tendencias, tanto del precio del crudo como de los índices bursátiles.

## CAPÍTULO 3

# PROCESOS ESTOCÁSTICOS

Un proceso estocástico es un objeto matemático definido como un conjunto de variables aleatorias. Cada una de las variables que conforman el proceso estocástico posee su propia función de distribución y comparten un único parámetro en su indexación que determina su evolución, siendo usualmente el tiempo.

Es posible encontrar este tipo de procesos en numerosos fenómenos, donde la aleatoriedad toma protagonismo. Algunos ejemplos son el crecimiento poblacional de bacterias, accidentes en una carretera, transmisión de enfermedades, ventas de retail, juegos de azar, cambios meteorológicos, estadísticas de crímenes en una ciudad, productos en mal estado y series financieras.

### 3.1. Series de Tiempo

Suponiendo que se observa una muestra de tamaño  $T$  de una variable aleatoria  $X_t$  [63]

$$\{x_1, x_2, \dots, x_T\} \quad (3.1)$$

Por ejemplo, considerando una colección de  $T$  variables  $\varepsilon_t$  idénticamente distribuidas (i.i.d.)

$$\{\varepsilon_1, \varepsilon_2, \dots, \varepsilon_T\} \quad (3.2)$$

con  $\varepsilon_t \sim N(0, \sigma^2)$ . Esto se refiere a una muestra de tamaño  $T$  de un proceso Gaussiano de ruido blanco.

La muestra observada en (3.1) representa  $T$  números particulares, pero este conjunto de  $T$  números es solo un posible resultado del proceso subyacente generador de la data.

Suponiendo que fuera posible observar el proceso por un período infinito de tiempo

$$\{x_t\}_{t=-\infty}^{\infty} = \{\dots, x_{-1}, x_0, x_1, x_2, \dots, x_T, x_{T+1}, x_{T+2}, \dots\} \quad (3.3)$$

la secuencia infinita  $\{x_t\}_{t=-\infty}^{\infty}$  aún sería una única realización del proceso de serie temporal. Por ejemplo, podría programarse un computador para trabajar generando una secuencia infinita de variables i.i.d.  $N(0, \sigma^2)$ ,  $\{\varepsilon_t^{(1)}\}_{t=-\infty}^{\infty}$ , y un segundo computador generando una secuencia separada  $\{\varepsilon_t^{(2)}\}_{t=-\infty}^{\infty}$ . Se puede considerar que son dos realizaciones independientes de un proceso Gaussiano de ruido blanco.

Imaginando un conjunto de  $I$  computadores generando secuencias  $\{\varepsilon_t^{(1)}\}_{t=-\infty}^{\infty}$ ,  $\{\varepsilon_t^{(2)}\}_{t=-\infty}^{\infty}$ ,  $\dots$ ,  $\{\varepsilon_t^{(I)}\}_{t=-\infty}^{\infty}$ , y considerando las observaciones asociadas a  $t$  de cada secuencia, es posible obtener

$$\{x_t^{(1)}, x_t^{(2)}, \dots, x_t^{(I)}\} \quad (3.4)$$

Esto puede ser descrito como una muestra de  $I$  realizaciones de la variable aleatoria  $X_t$ . Es posible ajustar alguna densidad, denotada por  $f_{X_t}(x_t)$ , la cual es llamada la densidad incondicional de  $X_t$ . Por ejemplo, para el proceso Gaussiano de ruido blanco, esta densidad está dada por

$$f_{X_t}(x_t) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(\frac{-y_t^2}{2\sigma^2}\right) \quad (3.5)$$

La esperanza de la observación  $t$ -ésima se refiere a la media de la función de distribución, asumiendo que existe

$$\mathbb{E}(X_t) = \int_{-\infty}^{\infty} x_t f_{X_t}(x_t) dy_t \quad (3.6)$$

Es posible interpretar esto como la probabilidad límite del promedio del conjunto

$$\mathbb{E}(Y_t) = \text{plim}_{I \rightarrow \infty} (1/I) \sum_{i=1}^I X_t^{(i)} \quad (3.7)$$

Para el caso en que  $\{X_t\}_{t=-\infty}^{\infty}$  representa la suma de una constante  $\mu$  más un proceso de ruido blanco Gaussiano  $\{\varepsilon_t\}_{t=-\infty}^{\infty}$ ,

$$X_t = \mu + \varepsilon_t \quad (3.8)$$

su media toma la forma

$$\mathbb{E}(Y_t) = \mu + \mathbb{E}(\varepsilon_t) = \mu \quad (3.9)$$

Si  $X_t$  es una tendencia temporal más un ruido blanco Gaussiano

$$X_t = \beta t + \varepsilon_t \quad (3.10)$$

entonces su media toma la forma

$$\mathbb{E}(X_t) = \beta t \quad (3.11)$$

La esperanza  $\mathbb{E}(X_t)$  a menudo se llama *media incondicionada* de  $X_t$ , la cual es denotada por  $\mu_t$

$$\mathbb{E}(X_t) = \mu_t \quad (3.12)$$

Esta notación permite que la media pueda ser una función de la fecha de la observación  $t$ . Para el proceso denotado por (3.8), su media (3.9), no resulta una función del tiempo. Por otro lado, para el caso del proceso denotado por (3.10), su media (3.11) si resulta ser una función del tiempo.

La varianza de la variable aleatoria  $X_t$ , denotada por  $\gamma_{0t}$ , es definida de forma similar como

$$\gamma_{0t} \equiv \mathbb{E}(X_t - \mu_t)^2 = \int_{-\infty}^{\infty} (x_t - \mu_t)^2 f_{Y_{X_t}}(x_t) dx_t \quad (3.13)$$

En el caso del proceso con tendencia temporal 3.10, la varianza está dada por

$$\gamma_{0t} = \mathbb{E}(X_t - \beta t)^2 = \mathbb{E}(\varepsilon_t^2) = \sigma^2 \quad (3.14)$$

### 3.1.1. Autocovarianza

Dada una realización particular como  $\{x_t^{(1)}\}_{t=-\infty}^{\infty}$  en un proceso de serie temporal, considerar un vector  $y_t^{(1)}$  asociado a la fecha  $t$ . Este vector consiste en las  $[j + 1]$  observaciones más recientes en  $x$  hasta la fecha  $t$

$$x_t^{(1)} \equiv \begin{bmatrix} y_t^{(1)} \\ y_{t-1}^{(1)} \\ \vdots \\ y_{t-j}^{(1)} \end{bmatrix} \quad (3.15)$$

Cada realización puede generar un valor particular del vector  $y_t$ , y se requiere calcular la distribución de probabilidad de este vector  $x_t^{(i)}$  para todas las realizaciones  $i$ . Esta distribución es la distribución conjunta de  $(X_t, X_{t-1}, \dots, X_{t-j})$ . De esta distribución es posible calcular la  $j$ -ésima autocovarianza de  $X_t$ , denotada por  $\gamma_{jt}$

$$\gamma_{jt} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} (x_t - \mu_t)(x_{t-j} - \mu_{t-j}) \quad (3.16)$$

$$\cdot f_{X_t, X_{t-1}, \dots, X_{t-j}}(x_t, x_{t-1}, \dots, x_{t-j}) dx_t dx_{t-1} \cdots dx_{t-j} \\ = \mathbb{E}(X_t - \mu_t)(Y_{t-j} - \mu_{t-j}) \quad (3.17)$$

Esta ecuación tiene la forma de una covarianza entre dos variables  $X$  e  $Y$

$$Cov(X, Y) = \mathbb{E}((X - \mu_X)(Y - \mu_Y)) \quad (3.18)$$

Por ello, la ecuación (3.16) puede ser descrita como la covarianza de  $X_t$  con su propio valor rezagado, de ahí el término *autocovarianza*. Es necesario señalar que en dicha ecuación, el rezago cero corresponde a la varianza de  $X_t$ , como se señaló anteriormente en (3.14) mediante  $\gamma_{0t}$ .

La autocovarianza  $\gamma_{jt}$  puede ser vista como el  $(1, j + 1)$  elemento de la matriz de varianza-covarianza del vector  $y_t$ . Por esta razón, las autocovarianzas se describen como los segundos momentos del proceso para  $X_t$ .

Como ejemplo de cálculo de autocovarianzas, notar que para el proceso (3.8) las autocova-

rianzas son todas cero para  $j \neq 0$

$$\gamma_{jt} = \mathbb{E}((X_t - \mu)(X_{t-j} - \mu)) = \mathbb{E}(\varepsilon_t \varepsilon_{t-j}) = 0, \quad j \neq 0 \quad (3.19)$$

### 3.1.2. Estacionariedad estricta

Aún cuando se podría considerar en el análisis estadístico de instrumentos financieros que el desempeño de los retornos en  $t$  no son un reflejo de su comportamiento pasado, es posible creer que existen ciertas propiedades estadísticas que permaneces estables a través del tiempo. De otro modo, no existe posibilidad de realizar un análisis estadístico significativo de los retornos, ni considerar modelos que puedan resultar de interés para representar su distribución. Esta invarianza de los momentos de la distribución de un proceso aleatorio es lo que se conoce como la hipótesis de estacionariedad.

La estacionariedad estricta de un proceso  $\{X_t\}$  implica que para cualquier conjunto finito de instantes de tiempo  $\{t_1, \dots, t_k\}$  y para cualquier período de tiempo dado  $\tau$ , la distribución conjunta de  $\{X_{t_1}, \dots, X_{t_k}\}$ , es la misma que la distribución conjunta de  $\{X_{t_1+\tau}, \dots, X_{t_k+\tau}\}$  [64], es decir

$$F_{X_{t_1}, \dots, X_{t_k}}(x_1, \dots, x_k) = F_{X_{t_1+\tau}, \dots, X_{t_k+\tau}}(x_1, \dots, x_k) \quad (3.20)$$

Como  $\tau$  no afecta  $F_X(\cdot)$ ,  $F_X$  no es una función del tiempo.

Una propiedad interesante de los procesos estacionarios estrictos es que una vez que se encuentra uno, es posible producir otros al aplicar una operación 'regular', como por ejemplo, un promedio móvil.

No siempre resulta evidente que los retornos puedan ser descritos como un proceso estrictamente estacionario. Sin embargo, es un supuesto conveniente para asegurar que uno puede estimar los momentos de los retornos al tomar muestras de cualquier intervalo. Al observar los gráficos usuales de retornos, se encuentra que la media suele ser constante y estar muy cerca del cero, mientras que la varianza esta acotada y describe patrones que se repiten a través de diferentes períodos de tiempo. Por ello, resulta más razonable que la invariabilidad de los todos momentos, asumir que el primer momento es constante, pero que el segundo tiene relación con su pasado, lo cual implica tener nociones del concepto de

autocovarianza, equivalente a la covarianza entre la serie, y el rezago de la misma.

### 3.1.3. Estacionariedad débil

Si ni la media  $\mu_t$ , ni la autocovarianza  $\gamma_{jt}$  dependen del tiempo  $t$ , entonces el proceso  $X_t$  es llamado *estacionario en covarianza* o *débilmente estacionario*. Esto resulta muy útil debido a que encontrar procesos estacionarios estrictos en la práctica es muy complejo, e incluso si se obtiene uno, su demostración también puede serlo.

$$\mathbb{E}(X_t) = \mu \quad \forall t \quad (3.21)$$

$$\mathbb{E}((X_t - \mu)(X_{t-j} - \mu)) = \gamma_j \quad \forall t, j \quad (3.22)$$

$$(3.23)$$

Por ejemplo, el proceso (3.8) es estacionario en covarianza

$$\mathbb{E}(X_t) = \mu \quad (3.24)$$

$$\mathbb{E}((X_t - \mu)(X_{t-j} - \mu)) = \begin{cases} \sigma^2 & j = 0 \\ 0 & j \neq 0 \end{cases} \quad (3.25)$$

Por otro lado, el proceso de (3.10) no es estacionario en covarianza, ya que su media,  $\beta t$ , es una función del tiempo.

Notar que si un proceso es estacionario en covarianza, la covarianza entre  $X_t$  y  $X_{t-j}$  depende solo del valor de  $j$ , la distancia separando las observaciones, y no en el valor de  $t$ , la fecha de la observación. De esto es posible derivar que para un proceso estacionario en covarianza,  $\gamma_j$  y  $\gamma_{-j}$  representan la misma magnitud. Es posible mostrar esto brevemente recordando la definición

$$\gamma_j = \mathbb{E}((X_t - \mu)(X_{t-j} - \mu)) \quad (3.26)$$

Si el proceso es estacionario en covarianza, entonces esta magnitud es la misma para cualquier valor de  $t$  que puede escogerse. Por ejemplo, si se reemplaza  $t$  con  $t + j$

$$\gamma_j = \mathbb{E}((X_{t+j} - \mu)(X_{[t+j]-j} - \mu)) = \mathbb{E}((X_{t+j} - \mu)(X_t - \mu)) = \mathbb{E}((X_t - \mu)(X_{t+j} - \mu)) \quad (3.27)$$

esta expresión es simplemente la definición de  $\gamma_{-j}$ . Por ello, para cualquier proceso estacionario en covarianza

$$\gamma_j = \gamma_{-j}, \quad \forall j \in \mathbb{Z} \quad (3.28)$$

Es necesario aclarar que un proceso estrictamente estacionario con segundo momento finito es débilmente estacionario, pero un proceso débilmente estacionario no necesariamente es estrictamente estacionario. Una excepción es un proceso Gaussiano débilmente estacionario, como una normal multivariada, que si presenta también estacionariedad estricta.

### 3.1.4. Ruido blanco

El bloque básico para todos los procesos considerados en el análisis de modelos lineales es la secuencia  $\{\varepsilon_t\}_{t=-\infty}^{\infty}$ , cuyos elementos tienen media cero y varianza  $\sigma^2$

$$\mathbb{E}(\varepsilon_t) = 0 \quad (3.29)$$

$$\mathbb{E}(\varepsilon_t^2) = \sigma^2 \quad (3.30)$$

para los cuales, las series  $\varepsilon$  no están correlacionadas a través del tiempo

$$\mathbb{E}(\varepsilon_t \varepsilon_\tau) = 0, \quad t \neq \tau \quad (3.31)$$

Un proceso que satisfaga estas últimas tres ecuaciones es descrito como un proceso de ruido blanco.

En ciertos casos, se suele reemplazar la última ecuación con la condición, ligeramente más exigente, de que  $\varepsilon_t, \varepsilon_\tau$  son independientes para  $t \neq \tau$ .

Notar que esta última condición implica (3.31), pero el recíproco no es cierto.

Un proceso que cumpla con las cuatro condiciones es llamado un *proceso de ruido blanco independiente*.

Finalmente, si las cuatro se sostienen junto a

$$\varepsilon_t \sim N(0, \sigma^2) \quad (3.32)$$

entonces se denomina *Proceso de ruido blanco Gaussiano*.

### 3.1.5. Raíz unitaria

La característica de raíz unitaria suele aparecer en procesos que evolucionan a través del tiempo. Esta característica puede causar problemas en inferencia enfocada a modelos de series de tiempo.

El nombre 'raíz unitaria' deriva del hecho que un proceso estocástico que posea raíz unitaria tendrá al 1 como raíz de la ecuación característica del proceso. Si las otras raíces caen dentro del círculo unitario, es decir, tienen un módulo menor que uno, entonces la primera diferencia del proceso será estacionaria. En caso contrario, el proceso deberá ser diferenciado más de una vez para volverlo estacionario.

Si se considera un proceso estocástico de tiempo discreto definido como  $\{Y_t, t = 1, \dots, \infty\}$  y que puede ser escrito como un proceso autorregresivo de orden  $p$

$$Y_t = \alpha_1 Y_{t-1} + \alpha_2 Y_{t-2} + \dots + \alpha_p Y_{t-p} + \varepsilon_t \quad (3.33)$$

Se considera que  $\varepsilon_t$  es ruido blanco y por simplicidad se asume que  $Y_0 = 0$ . La existencia de raíz unitaria se estudiará buscando las raíces de la ecuación característica

$$m^p - m^{p-1}\alpha_1 - m^{p-2}\alpha_2 - \dots - \alpha_p = 0 \quad (3.34)$$

En el caso en que  $m = 1$  es una raíz con multiplicidad<sup>1</sup> 1 de dicha ecuación, el proceso tiene raíz unitaria, o dicho de otra forma, es integrado de orden uno ( $I(1)$ ). En el caso en que  $m = 1$  es una solución, pero de multiplicidad  $k$ , el proceso es integrado de orden  $k$

<sup>1</sup>Número de veces que  $m = 1$  es solución del polinomio característico

$(I(k))$ , siendo su  $k$ -ésima diferencia estacionaria.

Por ejemplo, teniendo un proceso  $Y_t = \alpha_1 Y_{t-1} + \varepsilon_t$ , sabemos que su ecuación característica es  $m - \alpha_1 = 0$ , por lo cual en el caso en que  $\alpha_1 = 1$ ,  $m$  tomará el valor de 1, haciendo que el proceso tenga raíz unitaria y la serie no sea estacionaria, siendo el proceso dependiente del tiempo.

Para el caso de  $\alpha_1 = 1$ , el proceso corresponde a una caminata aleatoria. Si consideramos nuevamente que  $Y_0 = 0$ , podemos expresar el proceso como

$$Y_t = Y_0 + \sum_{j=1}^t \varepsilon_j \quad (3.35)$$

Por lo que la varianza estará dada por

$$\mathbb{V}(Y_t) = \sum_{j=1}^t \sigma^2 = t\sigma^2 \quad (3.36)$$

Es decir, la varianza depende de  $t$  violando uno de los principales requisitos en la estacionariedad, y volviéndose cada vez más grande a medida que  $t$  crece.

Es necesario hacer la diferencia con procesos de tendencia estacionaria, donde la tendencia puede ser removida y hacer estacionario al proceso, sin necesidad de integrarlo. Dicha tendencia no necesariamente debe ser lineal.

La diferencia entre ambos radica en que para la tendencia estacionaria, al sufrir una desviación el proceso, este volver a la tendencia de largo plazo por muy grande que pueda ser el ruido transitorio. En cambio en los procesos con raíz unitaria un cambio importante afectará de forma permanente el comportamiento de la serie.

## 3.2. Algunos tipos de procesos estocásticos

### 3.2.1. Martingala

Proceso estocástico en que el conocimiento de eventos pasados nunca ayuda a predecir la media de futuros aciertos o reducir su incerteza. La esperanza condicional del valor de

la siguiente observación es igual a la última observación. Representa el típico caso de la caminata aleatoria.

Matemáticamente esto es, dado un espacio de probabilidad definido por  $(\Omega, \mathcal{F}, P)$ , donde  $\Omega$  es el conjunto de todos los posibles resultados del experimento aleatorio,  $\mathcal{F}$  es la  $\sigma$ -álgebra asociada a  $\Omega$  y  $P$  es la medida de probabilidad.

Sea  $\mathbb{F}$  una filtración de  $\sigma$ -álgebras,  $\mathcal{F}_1 \subset \mathcal{F}_2 \subset \dots \subset \mathcal{F}_T \subseteq \mathcal{F}$ . Si  $\{X(t)\} = X_1, X_2, \dots, X_n$  es una sucesión de las variables aleatorias que forman el proceso estocástico, cuando este se adapta a la filtración  $\mathbb{F}$  recibirá el nombre de Martingala si

$$\mathbb{E}(X(t)|\mathcal{F}) = X(s) \quad (3.37)$$

Es decir, el proceso estocástico  $\{X(t), t \geq 0\}$  es una Martingala si su esperanza en tiempo  $t$ , con  $t > s$ , con la información conocida del pasado hasta  $s$ , dada por  $\mathcal{F}_s$ , es el valor que el proceso tomo en  $s$ , es decir, su esperanza en tiempo futuro es el valor que tiene en tiempo presente. A esto se refiere la ausencia de deriva estadística.

Otros tipos corresponden a una Submartingala si

$$\mathbb{E}(X(t)|\mathcal{F}) \geq X(s) \quad (3.38)$$

y a una Supermartingala si

$$\mathbb{E}(X(t)|\mathcal{F}) \leq X(s) \quad (3.39)$$

### 3.2.2. Movimiento Browniano

Considerando una caminata aleatoria

$$y_t = y_{t-1} + \varepsilon_t \quad (3.40)$$

donde las innovaciones son variables aleatorias de distribución estándar

$$\varepsilon_t \sim i.i.d N(0, 1) \quad (3.41)$$

Si el proceso se empieza con  $y_0$ , entonces

$$y_t = \varepsilon_1 + \varepsilon_2 + \dots + \varepsilon_t \quad (3.42)$$

$$y_t \sim N(0, t) \quad (3.43)$$

Además, el cambio en el valor de  $y$  entre las fechas  $t$  y  $s$

$$y_s - y_t = \varepsilon_{t+1} + \varepsilon_{t+2} + \dots + \varepsilon_s \quad (3.44)$$

sigue una distribución  $N(0, (s - t))$  y es independiente del cambio entre las fechas  $r$  y  $q$  para cualquier fecha  $t < s < r < q$ .

Considerando el cambio entre  $y_{t-1}$  y  $y_t$ , y asumiendo que las innovaciones  $\varepsilon_t$  pueden ser descritas como la suma de dos variables normales independientes

$$\varepsilon_t = e_{1t} + e_{2t} \quad (3.45)$$

con  $e_{it} \sim i.i.dN\left(0, \frac{1}{2}\right)$ , es posible asociar  $e_{1t}$  con el cambio entre  $y_{t-1}$  y el valor de  $y$  en algún punto interior, por ejemplo,  $y_{t-(1/2)}$ ,

$$y_{t-(1/2)} - y_{t-1} = e_{1t} \quad (3.46)$$

y con  $e_{2t}$  con los cambios entre  $y_{t-(1/2)}$  y  $y_t$

$$y_t - y_{t-(1/2)} = e_{2t} \quad (3.47)$$

Si se muestrea el proceso en fechas enteras  $t = 1, 2, \dots$  el proceso de (3.46) y (3.47), tendrá las mismas propiedades que (3.40), ya que

$$y_t - y_{t-1} = e_{1t} + e_{2t} \sim i.i.d N(0, 1) \quad (3.48)$$

Adicionalmente, el proceso de (3.46) y (3.47) se define también en las fechas no enteras  $\{t + \frac{1}{2}\}_{t=0}^{\infty}$  y conserva la propiedad, tanto para fechas enteras y no enteras, de que  $y_s -$

$y_t \sim N(0, s - t)$  con  $y_s - y_t$  independiente del cambio sobre cualquier otro intervalo sin superposición.

Bajo el mismo razonamiento, se puede imaginar particionar los cambios (innovaciones) entre  $t - 1$  y  $t$  en  $N$  subperíodos separados

$$y_t - y_{t-1} = e_{1t} + e_{2t} + \dots + e_{Nt} \quad (3.49)$$

con  $e_{it} \sim i.i.dN(0, 1/N)$ . El resultado sería un proceso con las mismas propiedades de (3.40), definida en una cuadrícula de fechas cada vez más finas a medida que aumentamos  $N$ .

El límite cuando  $N \rightarrow \infty$  es un proceso de tiempo continuo conocido como el Movimiento Browniano Estándar. El valor de este proceso en una fecha  $t$  se denota como  $W(t)^2$ .

Un proceso continuo es una variable aleatoria que toma un valor para cada número  $t$  real no negativo, a diferencia de lo que ocurre en un proceso de tiempo discreto, el cual solo está definido para valores enteros de  $t$ . Para enfatizar esta distinción, se denota la fecha entre paréntesis al describir un proceso de tiempo continuo en la fecha  $t$ , por ejemplo  $W(t)$ , y se utilizaran subíndices para variables de tiempo discreto, por ejemplo  $y_t$ .

Un proceso discreto se representa como una secuencia de variables aleatoria contables, denotado como  $\{y_t\}_{t=1}^{\infty}$ . Una realización de un proceso de tiempo continuo puede verse como una función estocástica, denotada como  $W(\cdot)$ , donde  $W : t \in [0, \infty) \rightarrow \mathbb{R}^1$ .

Una realización particular de movimiento Browniano resulta ser una función continua de  $t$ . Para demostrar esto se tiene que tener en consideración el cambio entre  $t$  y  $t + \Delta$  esta distribuida por  $N(0, \Delta)$ . Tal cambio es esencialmente arbitrariamente pequeño cuando el intervalo  $\Delta$  va a cero.

**Definición:** El movimiento Browniano Estándar  $W(\cdot)$  es un proceso estocástico de tiempo continuo, asociando cada fecha  $t \in [0, 1]$  con el escalar  $W(t)$ , tal que

- $W(0) = 0$
- Para cualquier fecha  $0 \leq t_1 < t_2 < \dots < t_k \leq 1$ , los cambios  $[W(t_2) - W(t_1)]$ ,  $[W(t_3) - W(t_2)]$ , ...,  $[W(t_k) - W(t_{k-1})]$  están distribuidos independientemente según una Normal

multivariada, con  $[W(s) - W(t)] \sim N(0, s - t)$ .

- $W(t)$  tiene incrementos independientes, es decir, si  $0 \leq s_1 < t_1 \leq s_2 < t_2$ , entonces  $W_{t_1} - W_{s_1}$  y  $W_{t_2} - W_{s_2}$  son valores independientes aleatorios.
- Para cualquier realización dada,  $W(t)$  es continuo en  $t$  con probabilidad 1.

Existen ventajas en restringir el análisis a fechas  $t$  de un intervalo cerrado, por que siempre se asumirá de aquí en adelante. A pesar de ser un proceso continuo en  $t$ , no puede ser diferenciado utilizando cálculo estándar, ya que la dirección del cambio en  $t$  puede ser completamente diferente a la de  $t + \Delta$ , sin importar que tan pequeña sea  $\Delta$ .

### 3.2.2.1. Propiedades

- **Media:** La distribución de probabilidad incondicional corresponde a una normal con media 0 y varianza  $t$ , en algún tiempo  $t$

$$f_{W(t)}(x) = \frac{1}{\sqrt{2\pi t}} \exp(-x^2/(2t)) \quad (3.50)$$

Por lo que la esperanza es cero:  $\mathbb{E}(W(t)) = 0$ , y la varianza es calculada como

$$\mathbb{V}(W(t)) = \mathbb{E}(W(t)^2) - \mathbb{E}^2(W(t)) = \mathbb{E}(W(t)^2) - 0 = \mathbb{E}(W(t)^2) = t \quad (3.51)$$

- **Covarianza:** Esperanza y varianza derivan del hecho que los incrementos tienen distribución normal, es decir

$$W(t) = W(t) - W(0) \sim N(0, t) \quad (3.52)$$

Los resultados de la covarianza y la correlación siguen la definición de que los incrementos no superpuestos son independientes, donde la propiedad de ausencia de

correlación es utilizada. Si se asume que  $t_1 < t_2$

$$\text{Cov}(W(t_1), W(t_2)) = \mathbb{E}([W(t_1) - \mathbb{E}(W(t_1))] \cdot [W(t_2) - \mathbb{E}(W(t_2))]) = \mathbb{E}(W(t_1) \cdot W(t_2)) \quad (3.53)$$

y considerando que

$$W(t_2) = (W(t_2) - W(t_1)) + W(t_1) \quad (3.54)$$

$$\mathbb{E}(W(t_1) \cdot W(t_2)) = \mathbb{E}(W(t_1) \cdot ((W(t_2) - W(t_1)) + W(t_1))) \quad (3.55)$$

$$= \mathbb{E}(W(t_1) \cdot (W(t_2) - W(t_1))) + \mathbb{E}(W(t_1)^2) \quad (3.56)$$

y como  $W(t_1) = W(t_1) - W(t_0)$  y  $W(t_2) - W(t_1)$  son independientes

$$\mathbb{E}(W(t_1) \cdot (W(t_2) - W(t_1))) = \mathbb{E}(W(t_1))\mathbb{E}(W(t_2) - W(t_1)) = 0 \quad (3.57)$$

Finalmente

$$\text{Cov}(W(t_1), W(t_2)) = \mathbb{E}(W(t_1)^2) = t_1 = \min\{t_1, t_2\} \quad (3.58)$$

- **Correlación:** Con los resultados anteriores, es fácil ver que corresponde simplemente a

$$\text{Corr}(W(s), W(t)) = \frac{\text{Cov}(W(s), W(t))}{\sigma_{W(s)}\sigma_{W(t)}} = \frac{\min(s, t)}{\sqrt{st}} = \sqrt{\frac{\min(s, t)}{\max(s, t)}} \quad (3.59)$$

- **Varianza de los Incrementos:**

$$\mathbb{E}(W(t) - W(s))^2 = |t - s|, \quad s, t \in \mathbb{R} \quad (3.60)$$

- **Distribución de los incrementos:**

$$W(t) - W(s) \sim N(0, t - s) \quad (3.61)$$

Es posible simular un movimiento Browniano en un intervalo de tiempo  $[0, T]$  particio-

nando el intervalo total en intervalos finitos,  $0 = t_0 < t_1 < \dots < t_n = T$ , y calcular de forma recursiva la siguiente expresion [64]

$$\Delta W(t_i) = W(t_i) - W(t_{i-1}) = \varepsilon_i \sqrt{\Delta t_i} \quad (3.62)$$

con  $\varepsilon_i \sim N(0, 1)$  como variable aleatoria i.i.d. (ruido blanco) y  $\Delta t_i = t_i - t_{i-1}$ .

Para un proceso de tiempo discreto  $\{B_t = B(t) : t = t_0, \dots, t_n\}$  se verifican todas las propiedades del mBe.

### 3.2.3. Movimiento Browniano Aritmético

El mBe se puede extender a los procesos estocásticos, mediante una ecuación diferencial estocástica, donde se define al Movimiento Browniano Aritmético (mBa) como el proceso que la satisface [64]

$$dX(t) = \alpha dt + \sigma dW(t) \quad (3.63)$$

donde  $\alpha$  y  $\sigma$  son constantes y  $W(t)$  es un movimiento Browniano. Al resolver la ecuación diferencial estocástica se obtiene la siguiente expresión del proceso

$$X(t) = X(0) + \alpha t + \sigma W(t) \quad (3.64)$$

Dicho proceso posee las siguientes propiedades heredadas del mBe.

- Para cualquier  $s$  y  $t$  donde  $s < t$

$$X(t) - X(s) \sim N(\alpha(t - s), \sigma^2(t - s)) \quad (3.65)$$

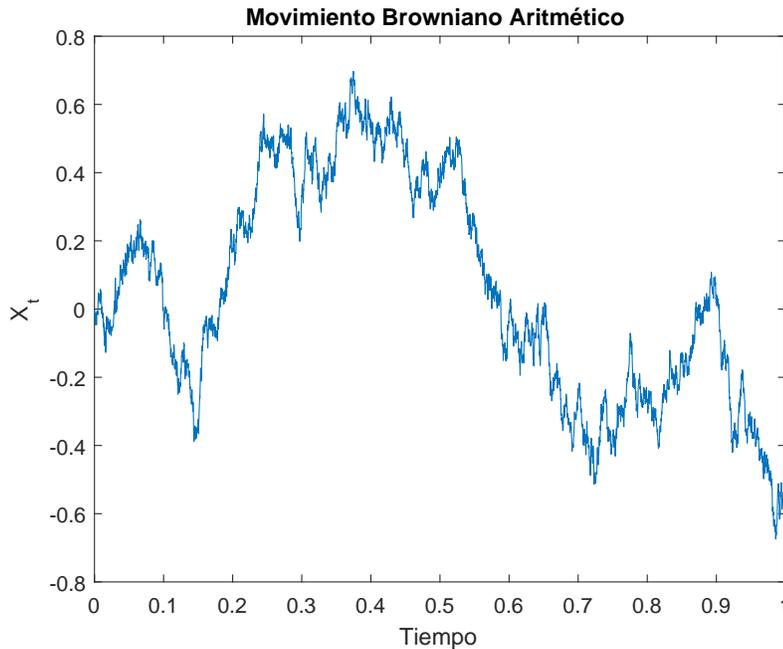
- Los incrementos  $\{X(t) - X(s) : s < t\}$  son estacionarios y estadísticamente independientes.
- Los caminos seguidos por  $X(t)$  son siempre continuos, sin ser diferenciables, con probabilidad 1.

$\alpha$  corresponde a la tasa esperada de deriva, mientras que  $\sigma^2$  corresponde a la varianza de los incrementos de  $X(t)$  por paso de tiempo. Como se observa en las propiedades, al tener incrementos independientes, uno no puede conocer un incremento futuro a partir de incrementos anteriores, lo cual lo hace un gran candidato para el modelamiento de series financieras.

Es posible simular una realización del mBa en un intervalo  $[0, T]$  con determinados  $\alpha$  y  $\sigma$ , utilizando el modelo discreto señalado anteriormente. Esto se realiza generando  $n > 0$  particiones de igual tamaño en el intervalo, e iniciando con  $X_0 = 0$ , iterando para los  $i = 1, \dots, n$  pasos la fórmula

$$X_i = X_{i-1} + \alpha \frac{T}{n} + \sigma \varepsilon_i \sqrt{\frac{T}{n}} \quad (3.66)$$

Asumiendo  $T = 1$ ,  $n = 2^{12}$ ,  $X_0 = 0$ ,  $\alpha = 0$  y  $\varepsilon_i \sim N(0, 1)$ , es posible obtener una realización del mBa, cuyo código de Matlab se encuentra en el Anexo.



**Figura 3.1:** Movimiento Browniano Aritmético con  $T = 1$ ,  $n = 2^{12}$ ,  $X_0 = 0$ ,  $\alpha = 0$  y  $\varepsilon_i \sim N(0, 1)$   
(Fuente: Elaboración Propia)

Mientras mayor sea el valor de  $n$ , más cercano estará el camino del Movimiento Browniano.

El proceso de tiempo discreto  $X_{t_i} = X(i)$  para  $t_1 = iT/n$ ,  $i = 0, 1, \dots, n$  definido por la ecuación (3.66) es una caminata aleatoria con tasa de deriva constante  $\alpha$  y varianza constante  $\sigma^2$  por unidad de tiempo, y distribución estándar. Claramente no es estacionaria, pero sus incrementos son estacionarios e independientes. Estos hechos implican que cualquier información pasada es irrelevante para los valores futuros del proceso, por lo que este proceso satisface la forma débil de la Hipótesis de los Mercados Eficientes.

### 3.2.4. Movimiento Browniano Geométrico

Una limitación importante del Movimiento Browniano aritmético como modelo de serie financiera es que en este modelo, la variación de precios sobre cualquier período de tiempo esta normalmente distribuido, lo cual implica que los precios pueden tomar valores negativos con probabilidad positiva. Esto puede solucionarse considerando una distribución log-Normal. Es decir, si  $\{P(t)\}_{t \geq 0}$  es un proceso continuo derivado de las observaciones discretas del precio de una serie, se busca modelar  $P(t)$  como  $e^{X(t)}$ , con  $X(t)$  como movimiento Browniano aritmético. Esto entrega que  $X(t) = \ln P(t)$ , haciendo que el precio esté log-normalmente distribuido, y con ello, los retornos logarítmicos

$$r_t = \ln \left( \frac{P(t)}{P(t-1)} \right) = X(t) - X(t-1) \quad (3.67)$$

son estacionarios e independientes.

El modelo continuo  $P(t) = e^{X(t)}$  es llamado Movimiento Browniano Geométrico (mBg)

### 3.2.5. Movimiento Browniano Fraccional

Esta generalización del Movimiento Browniano posee la característica de que los incrementos no requieren ser independientes. A pesar de ello, sigue siendo un proceso Gaussiano de tiempo continuo, cuya notación suele ser  $B_H(t)$  en  $[0, T]$ , donde  $H$  corresponde al índice o coeficiente de Hurst asociado al Movimiento Browniano Fraccional. Cumple con las siguientes propiedades

$$B_H(0) = 0 \quad (3.68)$$

$$\mathbb{E}(B_H(t)) = 0, \quad \forall t \quad (3.69)$$

$$\mathbb{E}(B_H(t)B_H(s)) = C_H(s, t) = \frac{1}{2} (|t|^{2H} + |s|^{2H} - |t - s|^{2H}), \quad \forall s, t \in \mathbb{R} \quad (3.70)$$

Es fácil ver que cuando el índice de Hurst es igual a  $1/2$  el proceso es un Movimiento Browniano clásico, ya que  $\mathbb{E}(B_{0,5}(t)B_{0,5}(s)) = \frac{1}{2}(s + t - |t - s|) = \min\{s, t\}$ .

El valor del coeficiente de Hurst determina el tipo de proceso del Movimiento Browniano Fraccional.

- $H > 1/2, C_H(s, t) > 0$ : Los incrementos del proceso están positivamente correlacionados. Existe dependencia de largo plazo.
- $H = 1/2, C_H(s, t) = 0$ : El proceso corresponde a un movimiento Browniano o proceso Wiener.
- $H < 1/2, C_H(s, t) < 0$ : Los incrementos del proceso están negativamente correlacionados.

Los incrementos del proceso  $X(t) = B_H(t + 1) - B_H(t)$  se conocen como Ruido Fraccional Gaussiano.

Existe una generalización del Movimiento Browniano Fraccional, denominado como  $n$ -mBf, el cual es Gaussiano, autosimilar y no estacionario, cuyos incrementos de orden  $n$  son estacionarios. Para  $n = 1$ , corresponde al mBf clásico.

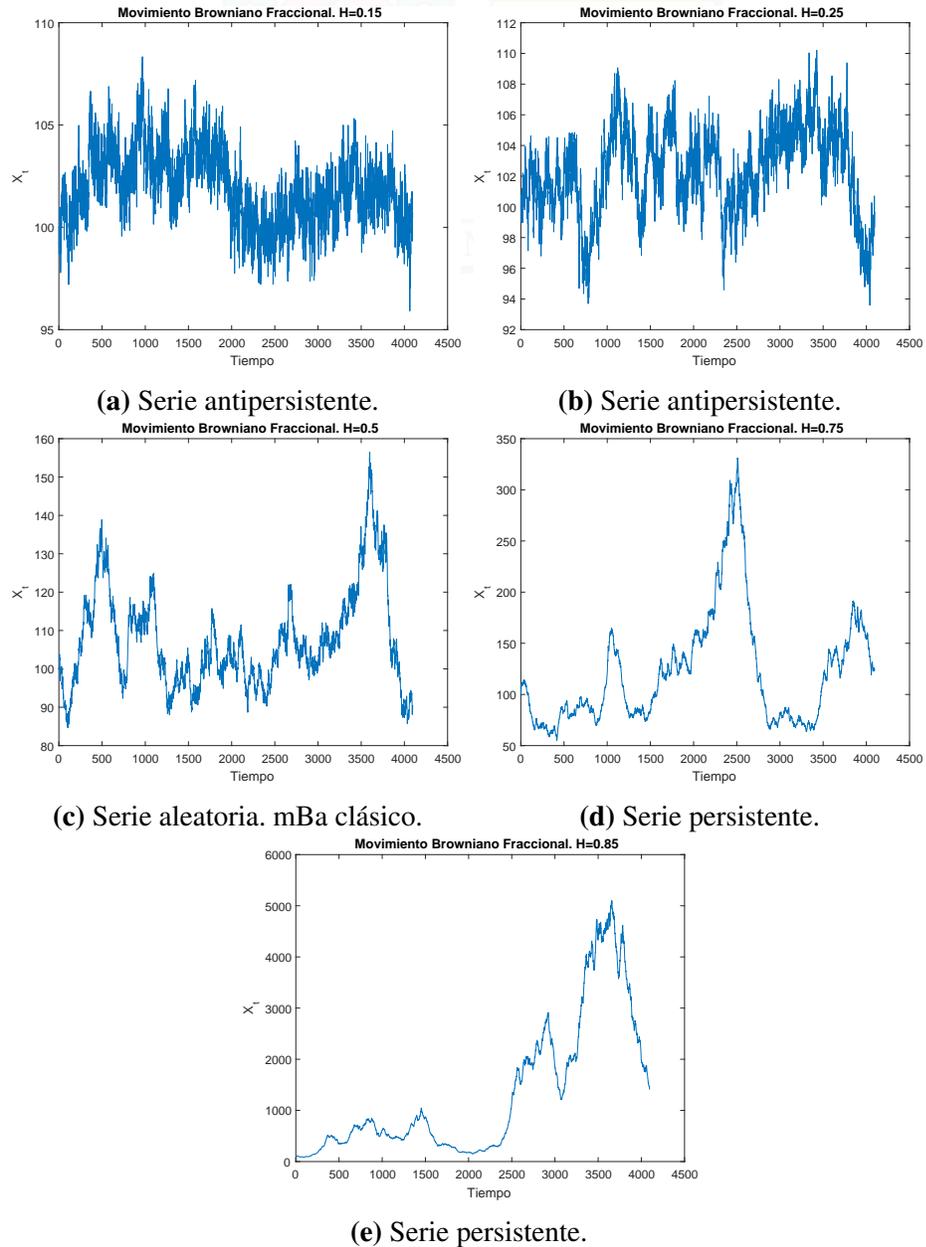
Al igual que el Movimiento Browniano, el mBf posee la propiedad de autosimilaridad, haciendo de éste un sistema fractal. Esto es, para cualquier  $H \in (0, 1)$  y  $\alpha > 0$

$$B_H(\alpha t) = \alpha^H B_H(t), \quad \forall t \quad (3.71)$$

Es necesario notar que cuando  $H \neq 0,5$ , no es posible considerar al proceso como Markoviano o como Martingala, donde en el primero se asume que la probabilidad no es afectada

por el proceso, y en el segundo que la esperanza para el siguiente resultado está dada por lo que ya ha ocurrido. Esto implica que los cálculos de Ito presentarán problemas.

En las siguientes figuras es posible observar simulaciones de un Movimiento Browniano Fraccional. El código para generarlo se encuentra en el Anexo.



**Figura 3.2:** Movimiento Browniano Fraccional para distintos valores del coeficiente de Hurst.

(Fuente: Elaboración Propia)

### 3.3. Pruebas estadísticas

#### 3.3.1. Test de Ljung-Box

Para determinar la potencial dependencia de la varianza con su pasado es posible utilizar la función de autocorrelación (ACF) para el cuadrado de los residuales. Y para efectivamente probar la heterocedasticidad condicional, o lo que es también conocido como el efecto *ARCH*, se puede utilizar el estadístico de Ljung-Box,  $Q(h)$ , para la serie  $\{a_t^2\}$ , en el cual la hipótesis nula a probar es si los primeros  $h$  rezagos de la ACF de  $\{a_t^2\}$  son cero. Si esta hipótesis nula es rechazada, es decir, el p-value es menor al 5 % de significancia usualmente utilizado, tenemos evidencia de efectos *ARCH*, y se recomienda el uso de modelos de volatilidad.

Escrito de manera formal, esto es

$H_0$  Los datos se distribuyen de forma independiente

$H_a$  Los datos no se distribuyen de forma independiente.

Con el estadístico de prueba

$$Q = n(n + 2) \sum_{k=1}^h \frac{\hat{\rho}_k^2}{n - k} \quad (3.72)$$

donde  $n$  es el tamaño de la muestra,  $\hat{\rho}_k^2$  es la autocorrelación de la muestra en el rezago  $k$ , y  $h$  es el número de rezagos que se está probando. La prueba se rechazará si se cumple que

$$Q > \chi_{1-\alpha, h}^2 \quad (3.73)$$

donde  $\chi_{1-\alpha, h}^2$  es una chi-cuadrado, con  $h$  grados de libertad, al  $\alpha$  nivel de significancia.

#### 3.3.2. Prueba de Jarque-Bera

La prueba de Jarque-Bera es una prueba de bondad de ajuste que busca entregar evidencia sobre si el tercer y cuarto momento de una serie temporal, la simetría y la curtosis, corresponden a la de una distribución normal [65]. Este test es uno de los más conocidos para la tarea de determinar si los datos siguen una distribución normal, y presenta muy

buenos resultados con series de tiempo largas ( $n > 2000$ ), a diferencia de lo que ocurre en otras pruebas como la de Shapiro-Wilk, la cual compara la distribución acumulada real con la ideal.

La simetría y curtosis muestral pueden ser calculadas, respectivamente como

$$S = \frac{\hat{\mu}_3}{\hat{\sigma}^3} = \frac{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^3}{\left(\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2\right)^{3/2}} \quad (3.74)$$

$$C = \frac{\hat{\mu}_4}{\hat{\sigma}^4} = \frac{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^4}{\left(\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2\right)^2} \quad (3.75)$$

donde  $\hat{\mu}_3$  y  $\hat{\mu}_4$  son los estimadores del 3er y 4to momento,  $\bar{x}$  es la media y  $\hat{\sigma}^2$  el estimador de la varianza.

Para  $S$ , se considera que

- Si  $S = 0$  la distribución es simétrica
- Si  $S < 0$  la distribución es sesgada a la izquierda, asimétrica negativa, o de colas pesadas a la izquierda
- Si  $S > 0$  la distribución es sesgada a la derecha, asimétrica positiva o de colas pesadas a la derecha

mientras que para  $C$  se considera que

- Si  $C = 3$ , la estructura de variabilidad es igual a la Gaussiana. Mesocúrtica.
- Si  $C < 3$ , la estructura de variabilidad es mayor a la ideal. Platicúrtica.
- Si  $C > 3$ , la estructura de variabilidad es menor a la ideal. Leptocúrtica.

El estadístico a utilizar para la prueba de hipótesis corresponde a

$$JB = \frac{n}{6} \left( S^2 + \frac{(C - 3)^2}{4} \right) \quad (3.76)$$

La idea de la prueba es verificar de forma conjunta si el indicador de simetría esperado es cero y el de curtosis es 3. Específicamente, el test prueba las siguientes hipótesis

$H_0$  Los datos se distribuyen según una Normal.

$H_a$  Los datos no distribuyen según una Normal.

El estadístico de prueba, JB, distribuye de acuerdo a una Chi-Cuadrado con dos grados de libertad ( $\chi^2_{1-\alpha,2}$ ).

A modo de extensión, se dice que una distribución posee colas pesadas si el área fuera del  $I=[\mu - 3\sigma, \mu + 3\sigma]$  es mayor al 1 %.

### 3.3.3. Prueba de heterocedasticidad condicional autorregresiva

La prueba ARCH [66], desarrollada por Engle en 1982, fue desarrollada para estudiar series temporales con heterocedasticidad condicional, o dicho de otra forma, autocorrelación en el cuadrado de los residuales. Este estadístico se justifica debido al hecho de que una serie puede no presentar autocorrelación pero ser de todas formas serialmente dependiente debido a un proceso de varianza condicional.

Suponiendo una serie de tiempo  $Y_t = \mu_t + \varepsilon_t$ , donde  $\mu_t$  es la media y  $\varepsilon_t$  son las innovaciones de media cero. Se asume además que las innovaciones son generadas por  $\varepsilon_t = \sigma_t z_t$ , donde  $z_t$  es un proceso i.i.d de media cero y varianza unitaria que cumple  $\mathbb{E}(\varepsilon_t \varepsilon_{t+h}) = 0$  para todos los rezagos donde  $h \neq 0$ .

Si  $H_t$  denota la historia del proceso disponible hasta el tiempo  $t$ , la varianza condicional de  $Y_t$  puede ser expresada como

$$\mathbb{V}(Y_t|H_{t-1}) = \mathbb{V}(\varepsilon_t|H_{t-1}) = \mathbb{V}(\varepsilon_t^2|H_{t-1}) = \sigma_t^2 \quad (3.77)$$

Por lo tanto, la heterocedasticidad condicional en el proceso de varianza es equivalente a la autocorrelación en el cuadrado del proceso de innovación.

Definiendo la serie de residuales  $e_t = Y_t - \hat{\mu}_t$ . Si toda la autocorrelación en la serie original  $Y_t$  es capturada en la media condicional, los residuales no estarán correlacionados y tendrán media cero. Sin embargo, los residuales pueden seguir siendo dependientes serialmente.

Para ello se prueba la autocorrelación pero para el cuadrado de los residuales, analizando la significancia de los parámetros del modelo

$$e_t^2 = \alpha_0 + \alpha_1 e_{t-1}^2 + \dots + \alpha_m e_{t-m}^2 + \mu_t \quad (3.78)$$

Donde las hipótesis nula y alternativa serán

$H_0$  No existe autocorrelación en el cuadrado de los residuales.

$H_a$  Existe evidencia de presencia de efectos ARCH en los residuales.

Dicho de otra forma, la hipótesis nula plantea que  $\alpha_0 = \alpha_1 = \alpha_2 = \dots = \alpha_m = 0$ . Para ello, la prueba utiliza el estadístico F de significancia global, que bajo la hipótesis nula sigue una Chi-Cuadrado con  $m$  grados de libertad ( $\chi_{1-\alpha, m}^2$ )

### 3.3.4. Prueba de Dickey-Fuller Aumentada

Esta prueba contrasta la hipótesis nula de presencia de raíz unitaria versus diferentes tipos de no estacionariedad [67]. El modelo a considerar para la elaboración del test es el siguiente

$$\Delta Y_t = c + \delta t + \phi Y_{t-1} + \beta_1 \Delta Y_{t-1} + \dots + \beta_{p-1} \Delta Y_{t-p} + \varepsilon_t \quad (3.79)$$

$c$  es constante,  $\delta$  es el coeficiente de la tendencia temporal, y  $p$  el número de rezagos del proceso autorregresivo.  $c = \delta = 0$  equivale a una caminata aleatoria sin deriva ni constante, y si solo  $\delta = 0$  corresponde a una caminata aleatoria sin deriva.

Hacer la prueba de si  $\phi = 0$ , es equivalente a probar si existe raíz unitaria. Es decir, las hipótesis corresponden a

$H_0$  Existe raíz unitaria.

$H_a$  No existe evidencia de raíz unitaria.

Los software que permiten realizar estas pruebas permiten hacer las pruebas con o sin deriva y tendencia temporal.

El estadístico de prueba dependerá del caso estudiado, en algunos casos se utilizan los clásicos estadísticos  $t$  con distribución definida y en otros es necesario realizar la interpolación en tablas estimadas mediante simulaciones de Montecarlo.

### 3.3.5. Prueba de memoria larga en la correlación

Previo a la aplicación del análisis principal por medio de MF-DCCA y MF-ADCCA, es necesario cuantificar el grado de correlación de memoria larga entre el WTI y los índices bursátiles. Ellos se realiza mediante el estadístico desarrollado por Podobnik [68], el cual tiene la siguiente estructura

$$C_i = \frac{\sum_{k=i+1}^N x_k y_{k-i}}{\sqrt{\sum_{k=1}^N x_k^2 \sum_{k=1}^N y_k^2}} \quad (3.80)$$

$$Q_{cc}(m) = N^2 \sum_{i=1}^m \frac{C_i^2}{N-i} \quad (3.81)$$

El estadístico sigue, aproximadamente, una distribución Chi-Cuadrado con  $m$  grados de libertad, es decir  $\chi^2(m)$ . Si el estadístico de prueba excede el valor crítico de la  $\chi^2(m)$ , con 5% de significancia en este estudio, se dice que existe evidencia de memoria larga en la correlación entre las series.

Es necesario destacar que este estadístico debe ser usado para probar la existencia de correlación de memoria larga cuando cada una de las series presenta memoria larga en su correlación. Las hipótesis, por ende, son

$H_0$  No hay correlación con memoria larga.

$H_a$  Existe evidencia de memoria larga en la correlación.

### 3.3.6. Coeficiente de correlación DCCA

Para verificar y agregar robustez al enfoque, se aplica otro método propuesto por Zebende . Esta estadístico se define como el ratio entre la función de covarianza sin tendencia  $F_{DCCA}^2$  y las varianzas sin tendencia  $F_{DFA}^2$  de cada serie.

$$\rho_{DCCA} = \frac{F_{DCCA}^2(s)}{F_{DFA1}(s) F_{DFA2}(s)} \quad (3.82)$$

Al igual que el coeficiente de Pearson, el  $\rho_{DCCA}$  es un coeficiente sin dimensión de rango  $-1 \leq \rho_{DCCA} \leq 1$ . En el caso donde las series no poseen correlación, el valor de  $\rho_{DCCA}$  será aproximadamente cero. En caso contrario, existirá evidencia de correlación. Ahora bien, series Gaussianas sin aparente correlación pueden mostrar valores distintos de cero,

pero gracias a la elaboración de diversas simulaciones con distintas series Gaussianas y no Gaussianas, de características monofractales y sin dependencia de largo plazo, se elaboran tablas tanto en el estudio por Podobnik antes citado y el trabajo de Blythe [55] con el objetivo de evitar correlaciones espurias. Estos valores dependen de diversos parámetros, pero en este estudio se utilizó el máximo valor crítico considerado para evitar dicha clase de resultados, el cual corresponde a un valor de 0,2, conforme a los datos utilizados.

Además, para develar la influencia de la escala, se calcula el valor del coeficiente para diferentes tamaños de ventana.

De acuerdo al trabajo de Podobnik, se recomienda utilizar este estadístico para cuantificar si es que la correlación entre las variables existe, pero complementarlo con la prueba de memoria larga en la correlación de la sección (3.3.5), y posteriormente con el análisis de escala de la función de varianza sin tendencia  $F_{DCCA}^2(s)$  para determinar si efectivamente existe memoria larga en la correlación.

Esto se debe a que el estadístico puede comportarse de igual forma en dos casos muy diferentes. El primero cuando no existe memoria larga ni en la autocorrelación, ni en la correlación cruzada, donde el coeficiente será relativamente constante con  $n$  debido a que  $2H_{xy} - H_x - H_y$  será muy pequeño. Pero esto también ocurrirá en el caso en que  $H_{xy} = (H_x + H_y)/2$ , donde no necesariamente existe solo dependencia de corto plazo. Por ello, para añadir robustez a las conclusiones, este test debe complementarse con los otros.

El valor de este estadístico, cuando es bien utilizado, es que ayuda a cuantificar el nivel de la correlación. Si bien el valor de  $H_{xy}$  cuantifica la correlación de memoria larga correspondiente a una ley de potencia, esto no es equivalente a cuantificar el nivel de correlación, ante lo cual este estadístico permite dar respuesta [46].



## CAPÍTULO 4

# MODELOS DE SERIES DE TIEMPO FINANCIERAS

### 4.1. Modelos lineales

Una serie de tiempo  $\{X_t\}$  es un proceso lineal si tiene la siguiente forma

$$X_t = \sum_{k=-\infty}^{\infty} \psi_k W_{t-k}, \quad \forall t \quad (4.1)$$

donde  $\{\psi_k\}$  es una secuencia de constantes con  $\sum_{k=-\infty}^{\infty} |\psi_k| < \infty$ , y  $\{W_t\}$  es un ruido blanco (débil) con media cero y varianza  $\omega^2$ :  $W_t \sim WN(0, \sigma^2)$ . La condición  $\sum_{k=-\infty}^{\infty} |\psi_k| < \infty$ , y  $W_t$  sirve para asegurar que la serie en 4.1 converja. Un caso particular del teorema de descomposición de Wold indica que cualquier proceso débilmente estacionario  $X_t$  tiene una representación lineal en la forma

$$X_t = \sum_{k=0}^{\infty} \psi_k W_{t-k} + V_t \quad (4.2)$$

donde  $\psi_0 = 1$ ,  $\sum_{k=0}^{\infty} \psi_k^2 < \infty$ ,  $W_t \sim WN(0, \sigma^2)$ ,  $Cov(W_s, V_t) = 0$  para todo  $s$  y  $t$ , y donde  $V_t$  es determinista. En el proceso lineal de la ecuación (4.2) tiene coeficientes nulos para  $k < 0$ , siendo un caso de media móvil infinita, por lo que según la descomposición de Wold de  $X_t$  en la ecuación (4.2), es posible obtener la estructura de las autocovarianzas de  $X_t$  en

términos de los coeficientes  $\psi_k$ , es decir, los pesos de medias móviles.

$$\gamma(0) = \mathbb{V}(X_t) = \sigma^2 \sum_{k=0}^{\infty} \psi_k^2 \quad (4.3)$$

$$\gamma(h) = \text{Cov}(X_{t+h}, X_t) = \sigma^2 \sum_{k=0}^{\infty} \psi_k \psi_{k+h} \quad (4.4)$$

De dichas ecuaciones uno puede obtener las autocorrelaciones

$$\rho(h) = \frac{\gamma(h)}{\gamma(0)} = \frac{\sum_{k=0}^{\infty} \psi_k \psi_{k+h}}{\sum_{k=0}^{\infty} \psi_k^2} \quad (4.5)$$

con las que luego es posible realizar análisis.

La representación de Wold aquí mostrada es conveniente para deducir propiedades de un proceso estacionario, pero no resulta de mucha utilidad para modelar el comportamiento del proceso, ya que la estimación de  $\psi_k$  para sumas infinitas en las ecuaciones 4.3 no es posible. Para ello surgen como aproximadores los modelos autorregresivos y de medias móviles (ARMA), donde la componente determinista  $V_t$  equivale a la media  $\mathbb{E}(X_t)$  multiplicado por alguna constante.

#### 4.1.1. Modelos autorregresivos y de media móvil

Considerando un ruido blanco débil  $W_t \sim W N(0, \sigma^2)$ , y siendo  $p \geq 1$  y  $q \geq 1$ . Una serie de tiempo  $X_t$  es autorregresiva de orden  $p$  ( $AR(p)$ ) si

$$X_t = W_t + \phi_1 X_{t-1} + \phi_2 X_{t-2} + \dots + \phi_p X_{t-p} \quad (4.6)$$

o de media móvil de orden  $q$  si

$$X_t = W_t + \theta_1 W_{t-1} + \dots + \theta_q W_{t-q} \quad (4.7)$$

o autorregresivo y media móvil de orden  $p$ ,  $q$  si

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \dots + \phi_p X_{t-p} + W_t + \theta_1 W_{t-1} + \dots + \theta_q W_{t-q} \quad (4.8)$$

donde  $\phi_1, \dots, \phi_p, \theta_1, \dots, \theta_q$  son números reales, en las tres ecuaciones.

Hasta ahora se ha considerado series con media cero para simplificar la notación, pero ello no implica una pérdida de generalidad, puesto que si un proceso dado,  $Y_t$  por ejemplo, tiene media  $\mathbb{E}(Y_t) = \mu$ , es posible ajustar el modelo ARMA a  $X_t = Y_t - \mu$ , el cual posee media cero. Esto simplemente añade una constante al modelo. Por ejemplo, suponiendo que  $Y_t$  sigue un proceso  $AR(1)$ , tal que  $Y_t - \mu = W_t + \phi_1(Y_{t-1} - \mu)$ , el cual es equivalente a  $Y_t = W_t + \phi_1 Y_{t-1} + \phi_0$ , donde  $\phi_0 = (1 - \phi_1)\mu$ . Si, por otro lado,  $Y_t$  sigue un proceso  $MA(1)$ , entonces  $Y_t = \mu + W_t + \phi_1 W_{t-1}$ .

### 4.1.2. Vector autorregresivo

Un modelo Vector Autorregresivo (VAR) [63] describe la evolución de un conjunto de series de tiempo. La idea del modelo es capturar la interdependencia lineal de las variables. Una vez que se entiende la especificación del modelo AR es fácil generalizarlo para  $k$  series de tiempo o variables endógenas, donde los rezagos de una variable afectan a las otras, y a su vez dicha variable se ve afectada por los rezagos de la otra.

Asumiendo un periodo de muestra finito  $t = 1, \dots, T$ , es posible agrupar las variables en un vector columna  $Y_t$  de largo  $k$ , cuyo  $i$ -ésimo elemento, es decir  $y_{it}$ , corresponde a la variable  $i$  en el tiempo  $t$ . Es posible escribir un VAR de orden  $p$ , denotado en adelante como  $VAR(p)$ , como

$$Y_t = c + A_1 Y_{t-1} + A_2 Y_{t-2} + \dots + A_p Y_{t-p} + \varepsilon_t \quad (4.9)$$

Esta ecuación incluye los rezagos de las  $k$  variables en el vector  $Y_{t-l}$ , que corresponde a los rezagos de orden  $l = 1, \dots, p$ , un vector  $c$  de intercepto, cada  $A_i$  será una matriz  $k \times k$

invariante en el tiempo, y  $\varepsilon_t$  será un vector  $k \times 1$  de términos de error que satisface

$$\mathbb{E}(\varepsilon_t) = 0 \quad (4.10)$$

$$\mathbb{E}(\varepsilon_t \varepsilon_t') = \Omega \quad (4.11)$$

$$\mathbb{E}(\varepsilon_t \varepsilon_{t-k}') = 0 \quad (4.12)$$

donde la primera ecuación determina la media cero de cada término del error, la segunda la matriz de covarianza de los términos del error de tamaño  $k \times k$  y la tercera la ausencia de correlación para un rezago distinto de cero  $k \neq 0$ , es decir ausencia de correlación serial de los vectores de errores.

En ocasiones resulta más fácil ver el modelo escrito en forma matricial. Por ejemplo, para un VAR(1) será

$$\begin{bmatrix} y_{1t} \\ y_{2t} \end{bmatrix} = \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} + \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \cdot \begin{bmatrix} y_{1,t-1} \\ y_{2,t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix} \quad (4.13)$$

O descrito en ecuaciones separadas

$$y_{1t} = c_1 + A_{11}y_{1,t-1} + A_{12}y_{2,t-1} + \varepsilon_{1t} \quad (4.14)$$

$$y_{2t} = c_2 + A_{21}y_{1,t-1} + A_{22}y_{2,t-1} + \varepsilon_{2t} \quad (4.15)$$

## 4.2. Modelos heterocedásticos

Las series financieras usualmente exhiben retornos diarios con varianza alta por cierto intervalo de tiempo, y similares en baja varianza por otro. Es posible observar esta dependencia del tiempo en la varianza de los retornos con su pasado al analizar la función de autocorrelación del cuadrado de los retornos, obteniéndose valores significativos en la misma.

Estas observaciones motivan modelos de retornos financieros donde su varianza evoluciona en el tiempo, los cuales se denominan *Modelos de Heterocedasticidad Condicional*. En términos formales, se busca que  $\mathbb{V}((X_t)|F_{t-1})$  varíe con  $F_{t-1}$ .

### 4.2.1. Modelo de Heterocedasticidad Condicional Autorregresiva (ARCH)

El modelo de Heterocedasticidad Condicional Autorregresiva (ARCH, por sus siglas en inglés), fue concebido por Robert Engle [66] bajo el supuesto de que la varianza condicional es una función lineal de las  $p$  innovaciones pasadas. Por ello, es posible definir un modelo  $ARCH(p)$  como

$$Y_t = \mu_t + a_t \quad (4.16)$$

$$\sigma_t^2 = \omega + \alpha_1 a_{t-1}^2 + \dots + \alpha_p a_{t-p}^2 \quad (4.17)$$

Para que este modelo esté bien definido y la varianza condicional sea positiva, los parámetros deben cumplir que  $\omega > 0$  y  $\alpha_1 \geq 0, \dots, \alpha_p \geq 0$ . Definiendo  $W_t = a_t^2 - \sigma_t^2$ , podemos escribir la ecuación (4.17) como

$$a_t^2 = \omega + \sum_{k=1}^p \alpha_k a_{t-k}^2 + W_t \quad (4.18)$$

y como  $\mathbb{E}(W_t | F_{t-1}) = 0$ , la ecuación (4.18) dice que el modelo  $ARCH(p)$  corresponde a un  $AR(p)$  para el cuadrado de las innovaciones,  $a_t^2$ . Por ello, el proceso es débilmente estacionario si, y solo si,  $|\alpha_1 + \dots + \alpha_p| < 1$ , en cuyo caso la varianza incondicional (en adelante, varianza constante) corresponde a

$$\sigma^2 = \mathbb{V}(a_t) = \frac{\omega}{1 - \alpha_1 - \dots - \alpha_p} \quad (4.19)$$

Para realizar inferencia con modelos ARCH usualmente se asume que  $a_t = \sigma_t \varepsilon_t$ , donde  $\{\varepsilon_t\}$  es una secuencia de variables aleatorias i.i.d. con  $\varepsilon_t \sim N(0, 1)$ .

A pesar de que esta realización de  $a_t$  no está serialmente correlacionada, por la ecuación (4.17) es claramente dependiente a través del tiempo. Se tiene que grandes (pequeños) valores de volatilidad deben ser seguidos por otros grandes (pequeños) valores de volatilidad. Por ello, el modelo  $ARCH$  permite trabajar con clústering de volatilidad.

### 4.2.1.1. Construcción de un modelo ARCH

Dada una serie de retornos  $\{r_t : t = 1, \dots, T\}$  a la que se pretende ajustar un modelo ARCH, lo primero es determinar el orden del modelo. Para ello se consideran los residuales  $a_t = r_t - \mu_t$  de la media en (4.17), y dado que el modelo ARCH es una autorregresión del cuadrado de los residuales (4.18), es posible usar la función de autocorrelación (PACF) para  $\{a_t^2\}$  con el objetivo de determinar el orden de  $p$ . El siguiente paso es estimar los parámetros  $\omega, \alpha_1, \dots, \alpha_p$  del modelo. Esto se hace mediante el Método de Estimación de Máxima Verosimilitud.

Asumiendo que la distribución de probabilidad de  $a_t$  condicionada a la varianza es normal, es decir,  $a_t \sim N(0, \sigma_t^2)$ , la función de máxima verosimilitud es

$$L(\omega, \alpha_1, \dots, \alpha_p) = \prod_{t=p+1}^T \left( \frac{1}{\sqrt{2\pi\sigma_t^2}} \exp\left(\frac{-a_t^2}{2\sigma_t^2}\right) \right) \quad (4.20)$$

Maximizar esa función es lo mismo que maximizar el logaritmo

$$-\frac{1}{2} \sum_{t=p+1}^T \left( \ln(2\pi) + \ln(\sigma_t) + \frac{a_t^2}{\sigma_t^2} \right) \quad (4.21)$$

Los valores que maximizan la ecuación anterior son los mismos que para la fórmula sin constantes aditivas. Por ello, la log-verosimilitud a considerar es

$$l(\omega, \alpha_1, \dots, \alpha_p) = -\frac{1}{2} \sum_{t=p+1}^T \left( \ln(\sigma_t^2) + \frac{a_t^2}{\sigma_t^2} \right) \quad (4.22)$$

Se procede a buscar iterativamente los valores de  $\omega, \alpha_1, \dots, \alpha_p$  que cuando se sustituyen en el modelo (4.17) para que  $\sigma_t^2$  maximice el valor de  $l(\omega, \alpha_1, \dots, \alpha_p)$ , comenzando con la varianza en el tiempo  $t = p + 1$  estimado con los  $p$  residuales pasados  $a_t, \dots, a_p$ , que significa

$$\sigma_{p+1}^2 = \omega + \alpha_1 a_p^2 + \alpha_2 a_{p-1}^2 + \dots + \alpha_p a_1^2 \quad (4.23)$$

donde  $a_i^2 = (r_i - \hat{\mu})^2$ , con  $\hat{\mu}$  siendo la media simple.

### 4.2.2. Modelo Generalizado de Heterocedasticidad Condicional Autorregresiva (GARCh)

El modelo Generalizado de Heterocedasticidad Condicional Autorregresiva (GARCh, por sus siglas en inglés) fue propuesto por Bollerslev en 1986 [69], como una extensión de los modelos  $ARCH(p)$  para valores grandes de  $p$ , los cuales a menudo son observados en la práctica.

La idea de Bollerslev consistió en hacer que la varianza también dependa de sus  $q$  valores pasados, obteniéndose la siguiente fórmula recursiva para  $\sigma_t^2$

$$X_t = \mu_t + a_t \quad (4.24)$$

$$\sigma_t^2 = \omega + \sum_{i=1}^p \alpha_i a_{t-i}^2 + \sum_{j=1}^q \beta_j \sigma_{t-j}^2 \quad (4.25)$$

Para que la varianza condicional de este  $GARCH(p, q)$  esté bien definido, sea positivo y que el proceso sea débilmente estacionario, los coeficientes deben satisfacer  $\omega > 0$ ,  $\alpha_i \geq 0$ ,  $i = 1, \dots, p$ ,  $\beta_j \geq 0$ ,  $j = 1, \dots, q$ , y que  $\sum_{i=1}^{\max(p,q)} (\alpha_i + \beta_i) < 1$ .

Estas condiciones del modelo se deducen de un razonamiento similar al del  $ARCH(p)$ , reescribiendo la ecuación (4.25) como sigue

$$a_t^2 = \omega + \sum_{k=1}^{\max(p,q)} (\alpha_k + \beta_k) a_{t-k}^2 - \sum_{k=1}^q \beta_k W_{t-k} + W_t \quad (4.26)$$

con  $W_{t-k} = a_{t-k}^2 - \sigma_{t-k}^2$ ,  $k = 0, 1, \dots, q$ . Esta ecuación define un modelo  $ARMA(\max(p, q), q)$  para  $a_t^2$ . De esto sigue que el proceso es débilmente estacionario si, y solo si,  $\sum_{k=1}^{\max(p,q)} (\alpha_k + \beta_k) < 1$ . Además, la varianza incondicionada para el  $GARCH(p, q)$  es obtenida tomando (4.25) como

$$\sigma^2 = \mathbb{V}(a_t) = \frac{\omega}{1 - \sum_{k=1}^{\max(p,q)} (\alpha_k + \beta_k)} \quad (4.27)$$

Esta varianza incondicionada puede ser interpretada como la varianza predicha de largo plazo. Para ver esto, se puede considerar la predicción de un paso simplemente con el modelo  $GARCH(p, q)$ . Esto es una extensión directa de un pronóstico  $ARCH(1)$ , y tiene la

forma

$$\sigma_{t+1|t}^2 = \omega + \alpha\sigma_{t+h-1|t}^2 + \beta\sigma_{t+h-1|t}^2 \quad (4.28)$$

Tomando el límite como  $h \rightarrow \infty$ , obtenemos la predicción de la varianza de largo plazo

$$V_L = \lim_{h \rightarrow \infty} \sigma_{t+h|t}^2 = \frac{\omega}{1 - \alpha - \beta} \quad (4.29)$$

Esta es la varianza media por período de tiempo (en este caso diario) considerada implícita en el modelo GARCH. La volatilidad implícita diaria es  $\sqrt{V_L}$ . Esto también provee otra interpretación para los modelos GARCH, como modelos para una aproximación a la varianza condicional por una varianza de largo plazo,  $V_L$ , más el cuadrado de las innovaciones y varianzas pasadas recientes. Para el modelo  $GARCH(1, 1)$ , este modelo es

$$X_t = \mu_t + a_t \quad (4.30)$$

$$\sigma_t^2 = \gamma V_L + \alpha a_{t-1}^2 + \beta \sigma_{t-1}^2 \quad (4.31)$$

con  $\gamma = 1 - \alpha - \beta$  y  $\alpha + \beta < 1$ .

#### 4.2.2.1. Construcción de un modelo GARCH

Los pasos para la construcción de un GARCH son similares a los de la construcción del ARCH para una serie de retornos dada  $\{r_t\}$ . Sin embargo, determinar el orden  $(p, q)$  del modelo no es fácil. Se ha observado que el cuadrado de los residuales  $a_t = r_t - \mu_t$  verifica un modelo  $ARMA(\max(p, q), q)$ , por lo que es posible aplicar un test de autocorrelación parcial para estimar un número  $p'$  para la parte AR del modelo, pero este se obtiene como  $p' = \max(p, q)$ , y se deben probar diferentes combinaciones de los parámetros  $p$  y  $q$ . Ahora bien, en la práctica, modelos de bajo orden son suficientes, y de hecho el modelo  $GARCH(1, 1)$  con  $\alpha_1 + \beta_1$  cercano a 1 han demostrado ser suficientemente buenos, debido a la gran cantidad de información contenida en el autorregresivo de la varianza condicional. Posteriormente, los parámetros  $\omega, \alpha, \beta$  y  $\mu$  son estimados con el método de máxima verosimilitud, donde la función de log-verosimilitud tiene la misma forma que para el modelo ARCH, dada por (4.22).

### 4.2.3. Variaciones de los modelos GARCH

#### 4.2.3.1. EGARCH

Con el objetivo de lidiar con la asimetría usualmente encontrada en series de retornos, el GARCH Exponencial (EGARCH) es desarrollado por Nelson en 1991. Permite modelar el logaritmo de la varianza condicional como

$$\log(\sigma_t^2) = \alpha_0 + \sum_{i=1}^p \gamma_i \log(\sigma_{t-i}^2) + \sum_{j=1}^q \alpha_j \left[ \frac{|\varepsilon_{t-j}|}{\sigma_{t-j}} - E \left( \frac{|\varepsilon_{t-j}|}{\sigma_{t-j}} \right) \right] + \sum_{j=1}^q \xi_j \left( \frac{\varepsilon_{t-j}}{\sigma_{t-j}} \right) \quad (4.32)$$

sin restricciones de parámetros.

#### 4.2.3.2. GJR-GARCH

Glosten, Jagannathan and Runkle proponen un GARCH modificado que toma en consideración la diferencia en la volatilidad condicional dependiendo del signo de los shocks. Se define como sigue

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^p \beta_j \sigma_{t-j}^2 + \sum_{j=1}^p \xi_j \mathbb{I}[\varepsilon_{t-j} < 0] \varepsilon_{t-j}^2 \quad (4.33)$$

Como se observa en el último término, existe un componente *ARCH* para el caso en que el retorno de los rezados considerados sea negativo. Con ello, existe la posibilidad de tener diferentes influencias en la varianza respecto a noticias 'buenas' y 'malas'.

## 4.3. Medidas de calidad relativa de modelos

### 4.3.1. Criterio de información de Akaike

El AIC (por sus siglas en inglés) es una medida de la calidad relativa de un modelo estadístico ajustado a un determinado set de datos [70]. El modelo tiene una serie de

características que se deben tener en consideración en su uso

- La medida entregada por el AIC sirve para cierto set de datos y debe ser usada para comparar modelos ajustados a dicho set. Esto implica que no entrega información sobre la calidad de ajuste para un modelo en particular, sino que un medio para la selección de modelos. Puede darse el caso de tener solo modelos malos y esta medida no indicará nada al respecto.
- La inspiración de dicho modelo viene de la rama de la Teoría de la Información, es decir, es una medida de cuantificar la información perdida al estimar un determinado modelo. La medida permite tener en consideración tanto la bondad de ajuste del modelo como su complejidad. Si bien un mayor número de parámetros mejora el ajuste, también puede provocar problemas de *overfitting*.

Si se define  $k$  como el número de parámetros del modelo, y  $L$  es el valor máximo de la función de verosimilitud del modelo, el AIC estará dado por la siguiente expresión

$$AIC = 2k - 2\ln(L) \quad (4.34)$$

El valor de mejor ajuste corresponderá al menor valor de este criterio.

### 4.3.2. Criterio de información Bayesiano

Este criterio, BIC por sus siglas en inglés, es muy similar al AIC. La principal diferencia radica en que impone una mayor penalización a la inclusión [71].

Si nuevamente se considera  $k$  como el número de parámetros,  $L$  como el valor máximo de la función de verosimilitud del modelo y  $n$  el número de datos observados, es posible definir el BIC como

$$BIC = \ln(n)k - 2\ln(L) \quad (4.35)$$

Una limitación del BIC es que  $n \gg k$ . Además, entre AIC y BIC se suele utilizar AIC por poseer ventajas de desempeño y ser más práctico [72].

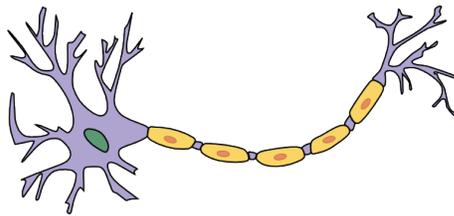
## 4.4. Redes Neuronales Artificiales

### 4.4.1. Motivación

El cerebro humano posee gran capacidad de procesamiento de información y generalización incluso bajo condiciones de datos incompletos, incertidumbre y pocas repeticiones. Esto ha inspirado a investigadores a utilizar como punto de partida las neuronas para el desarrollo de modelos matemáticos de aprendizaje.

Existen diferencias sustanciales entre el cerebro humano y computadores digitales en el procesamiento de la información. Mientras que los computadores digitales permiten resolver problemas mediante el uso de algoritmos secuenciales, el cerebro humano utiliza una gran red de células llamadas neuronas que permiten realizar trabajo en paralelo, las cuales poseen numerosas conexiones que entregan al humano la capacidad de aprendizaje y generalización para la resolución de problemas.

Motivados por la gran eficiencia que presenta el cerebro en sus tareas, en las últimas décadas investigadores han intentado replicar estos modelos biológicos para la construcción de modelos computacionales inteligentes, con el objetivo de procesar gran cantidad de información de manera eficiente. Estos sistemas son llamado Redes Neuronales Artificiales (en adelante RNA), que poseen numerosos elementos de procesamiento interconectados de forma tal que permiten realizar las tareas de aprendizaje.



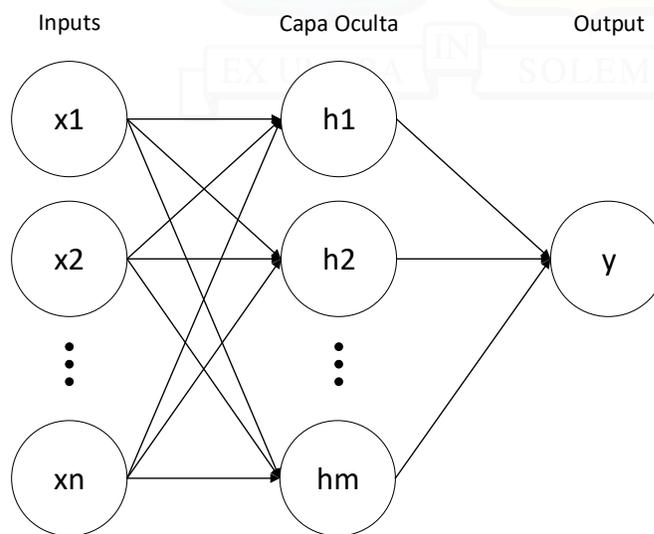
**Figura 4.1:** Representación de una neurona natural.

(Fuente: *Principles of neural science* [73])

### 4.4.2. Red Neuronal Prealimentada

La red neuronal *feed-forward*, en muchos casos llamada 'prealimentada', es una de los métodos mas populares para aproximar una función no lineal multivariada. Su característica

distintiva son las capas ocultas, donde las variables de input activan alguna función no lineal de transformación, generando una señal continua a los nodos de salidas. El esquema de capas ocultas representa un modelo altamente eficiente de un proceso estadístico no lineal paralelamente distribuido. De hecho, las redes neuronales han sido capaces de aproximar cualquier función continua uniformemente en conjuntos compactos, para un número apropiado de nodos en la capa oculta, y una muestra de entrenamiento suficientemente grande [74].



**Figura 4.2:** Ilustración de una red neuronal

(Fuente: Elaboración Propia)

La figura (4.2) muestra una red neuronal prealimentada con  $n$  nodos de entradas (*inputs*),  $x_1, x_2, \dots, x_n$ , un nodo de salida (*output*), y  $m$  nodos o neuronas ocultas,  $h_1, h_2, \dots, h_m$ , que constituyen la capa oculta.

Cada neurona  $h_i$  recibe la información del input y se activa, transmitiendo una señal al output. Si una cierta combinación lineal de los inputs,  $\sum_{i \rightarrow j} \omega_{ij} x_i$ , es mayor que algún valor límite  $\alpha_j$ , es decir,  $h_j$  transmite un valor no negativo si

$$z_j = \sum_{i \rightarrow j} \omega_{ij} x_i - \alpha_j > 0 \quad (4.36)$$

donde la suma se lleva a cabo sobre todos los nodos de inputs  $x_i$  que alimentan a la neurona  $h_j$ , y  $\omega_{ij}$  son los pesos. La señal se produce por una función de activación aplicada a  $z_j$ , la cual suele ser una sigmoide o logística, usada también en este estudio, y que tiene la siguiente forma

$$P(z) = \frac{\exp(z)}{1 + \exp(z)} = \frac{1}{1 + \exp(-z)} \quad (4.37)$$

Si no hay activación ( $z_j \leq 0$ ) la señal es cero, y cuando si existe activación de la señal, el valor cumple  $0 < P(z) < 1$ . Componiendo una combinación lineal de los posibles valores de las señales transmitidas por la capa oculta, más un algún valor  $d$  para tener en cuenta el sesgo en la conexión, al considerar una función de activación  $f$  para el nodo de salida, es posible obtener una expresión matemática para la red neuronal prealimentada con conexiones desde inputs a nodos ocultos, y de estos al output

$$y = f \left( d + \sum_{j=1}^m \theta_j P \left( \alpha_j + \sum_{i \rightarrow j} \omega_{ij} x_i \right) \right) \quad (4.38)$$

Para una red neuronal mas compleja, se puede permitir conexiones directas desde nodos de input a nodos de output, por lo que tenemos una expresión más general para estas redes neuronales

$$y = f \left( d + \sum_{i=1}^n \phi_i x_i + \sum_{j=1}^m \theta_j P \left( \alpha_j + \sum_{i \rightarrow j} \omega_{ij} x_i \right) \right) \quad (4.39)$$

donde  $x_1, \dots, x_n$  son los valores de input, e  $y$  es el output,  $P(\cdot)$  es la función de activación para cada uno de los  $m$  nodos ocultos,  $\alpha_j$  es el valor límite y  $\omega_{ij}$  son los pesos a la través de la conexión para el  $j$ -ésimo nodo oculto,  $\phi_i$  y  $\theta_j$  son los pesos para las combinaciones lineales de orden  $n$  y  $m$ ,  $d$  es el sesgo de la conexión y  $f$  es la función de activación para el nodo de output.  $f$  puede ser logística, lineal o umbral. En este estudio se utiliza una función lineal.

Notar que para el caso de  $f$  siendo lineal, la ecuación (4.39) toma la forma de un modelo ARMA para una serie de tiempo  $\{y_t\}$ . De hecho tomando  $x_1 = y_{t-1}, \dots, y_m$ , donde  $p$  representa diferentes rezagos de la serie, y el output siendo  $y = y_t$ , como la serie de tiempo

a pronosticar y  $f$  una función lineal en (4.39), se obtiene

$$y_t = d + \sum_{i=1}^n \phi_i y_{t-i} + \sum_{j=1}^m \theta_j P \left( \alpha_j + \sum_{i \rightarrow j} \omega_{ij} y_{t-i} \right) \quad (4.40)$$

Esto demuestra que la red neuronal prealimentada es una generalización de un modelo  $ARMA(n, m)$ .

El caso de  $f$  como función umbral también resulta de gran interés en las aplicaciones de series de tiempo. Si  $y = r_t$  es una serie de retornos y  $x_1, x_2, x_p$  son diferentes rezagos de esta serie de retornos  $\{r_t\}$ , o de otro estadístico como su volatilidad histórica, entonces la ecuación (4.39) con función umbral puede ser interpretada como un modelo para pronosticar la dirección de los retornos, donde  $f(z) = 1$  implica retornos positivos y  $f(z) = 0$  implica retornos negativos.

### 4.4.3. Estimación de modelos y pronósticos

El interés del modelamiento mediante redes neuronales radica en ajustarla a una serie de retornos. Sea  $\{(x_t, r_t) : t = 1, \dots, T\}$  sea la data disponible donde  $r_t$  son los retornos de un activo financiero y  $x_t$  sea el vector de inputs (también llamado características) que pueden contener varios rezagos de la serie, el volumen, la varianza o cualquier indicador fundamental de la serie. El ajuste del modelo a una red neuronal requiere dividir los datos en dos submuestras, de forma que una sea usada en el entrenamiento del modelo y la otra para su prueba. En la fase de entrenamiento de los parámetros (pesos, límites y sesgos) en la ecuación (4.39) se escogen de forma tal que el error de pronóstico es minimizado. Usualmente se requiere que los parámetros escogidos minimicen el Error Cuadrático Medio (MSE, por sus siglas en inglés, *Mean Squared Error*).

$$MSE(w) = \frac{1}{T} \sum_{t=1}^T (r_t - y_t)^2 \quad (4.41)$$

donde  $w$  es un vector de los parámetros requeridos e  $y_t$  es la salida de la ecuación (4.39) en los inputs  $x_t$  y  $r_t$  los retornos reales.

Este es un problema de estimación no lineal que puede ser resuelto por un método iterativo,

que usualmente es el algoritmo de propagación hacia atrás, o *Back Propagation*. Este algoritmo, como su nombre lo indica, trabaja en empezando por el output y utilizando reglas de gradiente para ajustar los pesos de forma iterativa en base a los errores obtenidos. En el paso de prueba de la red neuronal en el set de aprendizaje, se usa la segunda submuestra para predecir algunos valores y se compara dichas estimaciones con el valor real de dicha muestra. Los paso se repiten varias veces hasta ajustar los parámetros, con la posible aplicacion de validación cruzada para seleccionar el mejor modelo con respecto al desempeño de los pronósticos.

A pesar de su popularidad, el método de validación puede fallar en encontrar los mejores parámetros en la presencia de mínimos locales, y puede requerir mucho tiempo computacional en el procesamiento de la red neuronal.

Además de estos métodos, se pueden utilizar métodos basados en la máxima verosimilitud o Bayes para la selección de redes apropiadas.

#### 4.4.4. Algoritmo de propagación hacia atrás (*Backpropagation*)

Sea  $N$  una red neuronal con  $v$  conexiones,  $n$  inputs,  $m$  outputs,  $L$  capas y  $K$  ejemplos de entrenamiento. Nuevamente se denotará como  $x_1, x_2, \dots, x_k$  los vectores en  $\mathbb{R}^n$ .  $y, y', y_1, \dots, y_k$  vectores en  $\mathbb{R}^m$ .  $\omega, \omega_1, \omega_2, \dots$  en  $\mathbb{R}^v$ . Estos serán los inputs, outputs y pesos, respectivamente. La red neuronal corresponde a una función  $y = f_N(\omega, x)$ , la cual, dada un peso  $\omega$  mapea un input  $x$  a un input  $y$ . El algoritmo de propagación hacia atrás toma una secuencia de ejemplos de entrenamiento  $(x_1, y_1), \dots, (x_K, y_K)$  y produce una secuencia de pesos  $\omega_0, \omega_1, \dots, \omega_k$ , empezando de alguna secuencia  $\omega_0$  [75],

1. Determinar secuencia  $\omega_0$ . Puede ser aleatoriamente determinada o según algún método apropiado.
2. Escoger una señal  $k$ -ésima señal a propagar hacia adelante a través de la red, calculando sus valores de salida para las  $i$  neuronas en cada capa  $l$ , utilizando

$$o_i^l(k) = f \left( \sum_{p=0}^{n-l} \omega_{ip}^l o_p^{l-1} \right) \quad (4.42)$$

3. Calcular el valor del error total  $E = E(k) + E$  y la señal del error  $\delta_i^{(L)}$  mediante

$$\delta_i^{(L)} = [y_i - o_i^L][f'(tot_i^L)] \quad (4.43)$$

la señal  $tot_i^{(l)}$  representa la suma de todas las señales llegando al nodo  $i$  en la capa oculta  $l$ , desde la capa previa  $l - 1$ .

El error total toma la forma

$$E_c = \sum_{k=1}^K E(k) = \frac{1}{2} \sum_{k=1}^K \sum_{i=1}^q [y_i(k) - o_i(k)]^2 \quad (4.44)$$

La formulación del problema de optimización se puede plantear como el set de pesos que minimiza  $E_c$  o  $E(k)$ , dependiendo de si los ejemplos de entrenamiento se presentan al mismo tiempo o de a uno.

Para este trabajo se utiliza el segundo esquema, formulándose el problema de minimización como

$$\min_{\omega} E(k) = \min_{\omega} \frac{1}{2} \sum_{i=1}^q [y_i(k) - o_i(l)]^2 \quad (4.45)$$

4. Actualizar los pesos de acuerdo a

$$\Delta\omega_{ij}^{(l)} = -\eta\delta_i^{(l)}o_j^{(l-1)} \quad (4.46)$$

para  $l = L, \dots, 1$ , utilizando

$$\delta_i^{(L)} = [y_i - o_i^L][f'(tot_i^L)] \quad (4.47)$$

y dirigiéndose hacia atrás en la red  $N$  utilizando

$$\delta_i^{(l)} = o_i^{(l)}(1 - o_i^{(l)}) \sum_{p=1}^{n_l} \delta_p^{(l+1)}\omega_{pi}^{(l+1)} \quad (4.48)$$

para  $l < L$ .

5. Repetir el proceso empezando desde el Paso 2 utilizando otro ejemplo de entrena-

miento. El número de veces que se realice este procedimiento con todos los ejemplos de entrenamiento se denomina como número de *epochs*.

6. Chequear si el error acumulado  $E$  en la capa de output se ha vuelto menor que un valor predeterminado. Si es así, se concluye que la red ha sido exitosamente entrenada. Si no, se repite el proceso para un nuevo *epoch*.

#### 4.4.4.1. Algoritmo de Levenberg-Marquardt para la propagación hacia atrás (*Back-propagation*)

Este algoritmo fue diseñado para aumentar la velocidad de entrenamiento sin tener que calcular la matrix Hessiana [76]. Cuando la función de desempeño toma la forma de suma de cuadrados, como es el caso, la matriz Hessiana puede ser aproximada como

$$H = J^T J \quad (4.49)$$

y el gradiente se puede calcular como

$$g = J^T e \quad (4.50)$$

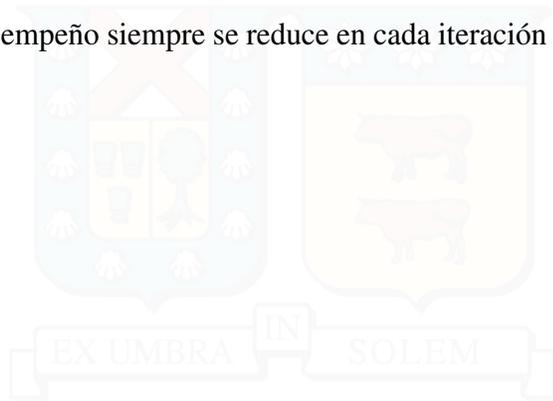
donde  $J$  es la matriz Jacobiana que contiene las primeras derivadas de los errores de la red con respecto a los pesos y sesgos, y  $e$  es un vector de errores de la red. La matriz Jacobiana puede ser calculada a través de la técnica de propagación hacia atrás, que es mucho menos complejo que calcular la matriz Hessiana.

El algoritmo de Levenberg-Marquardt utiliza dicha aproximación a la matriz Hessiana en la actualización de tipo Newton

$$x_{k+1} = x_k - [J^T J + \mu I]^{(-1)} J^T e \quad (4.51)$$

Cuando el escalar  $\mu$  es cero, esto corresponde al método de Newton utilizando la matriz Hessiana aproximada. Cuando  $\mu$  es grande, esto se convierte en el gradiente descendiente con paso pequeño.

El método de Newton es más rápido y preciso cerca de un error mínimo, por lo que el enfoque es apuntar hacia un método de Newton lo más rápido que se pueda. Por ello, se disminuye  $\mu$  luego de cada paso exitoso, es decir, donde se reduce la función de desempeño. Por otro lado, resulta útil aumentarlo cuando un paso incrementa la función de desempeño. De esta forma, la función de desempeño siempre se reduce en cada iteración del algoritmo.



## CAPÍTULO 5

# PRESENTACIÓN Y ANÁLISIS DE RESULTADOS

### 5.1. Contexto del petróleo

El comportamiento del precio del petróleo durante el período analizado, observable en la Figura (5.1), puede dividirse en 4 períodos de características bastante interesantes.

Durante el primer período, 2005 a 2007, experimentó un nivel relativamente estable en torno a los 60 USD/B (dólares por barril), para luego, durante el segundo periodo, experimentar una increíble subida hasta 2008, donde alcanza más de 140 USD/B, antes de desplomarse hasta en torno a los 30 USD/B al inicio de 2009. El tercer período, 2009 a 2014 se caracteriza por cierta estabilidad y movimiento al alza del petróleo. Dicho período también presenta gran interés debido a la diferencia de precios entre Brent y WTI, causados principalmente por hechos como la primavera árabe en 2011, reducción de la oferta de crudo fuera de EEUU dado por cortes de la OPEC y acompañados por el aumento de la oferta estadounidense.

Tanto el WTI como el Brent se consideran 'light' o 'livianos', es decir flotan en agua. Además son 'sweet' o 'dulces', es decir, poseen un contenido de sulfuros menor al 0,5 %. En términos de gravedad API, el Brent presenta un valor de 38,06, mientras que el WTI posee 39,6, lo que significa que es más liviano. Además, el WTI es relativamente más dulce.

Sin embargo, la diferencia más relevante radica en el lugar en que son extraídos. El Brent se obtiene de más de una docena de campos petrolíferos en el Mar del Norte, que se sitúa entre las costas de Noruega y Dinamarca en el este, Reino Unido en el oeste, y Alemania,

Bélgica y Francia al sur. Por otro lado, el WTI está ubicado en Texas, EEUU, por lo que sus costos de transporte tienden a ser mayores.

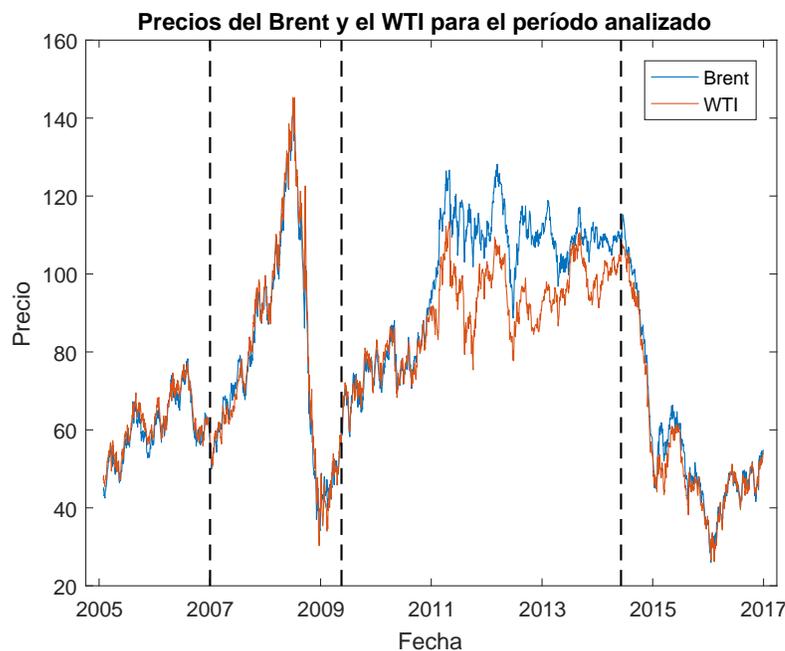
A pesar de ello, las diferencias entre ambos se han mantenido relativamente bajas, salvo para el período comprendido entre finales de 2010 hasta mediados de 2015. Como se mencionó antes, las diferencias entre ambos pueden deberse principalmente a factores geopolíticos, los que afectan menos al WTI por encontrarse en tierra de EEUU. Cuando el WTI gana fuerza se suele atribuir a factores de oferta y demanda domésticos. A pesar de lo anterior y del hecho que la mayoría del Brent se destina a mercados Europeos, el Brent ha adquirido cada vez más importancia como benchmark en los últimos años, ya que ha demostrado ser un indicador de precios globales, y es extraído de más de una docena de campos en el Mar del Norte. Su uso cada vez más extendido tiene el efecto de volverlo aún más poderoso, ya que está siendo utilizado como benchmark en África, el Mediterráneo y el sudeste asiático, ligándolo a un mercado más grande. Los saudíes lo cambiaron en 2009 como benchmark por el Argus Sour Crude Index (ASCI).

Según Bloomberg<sup>2</sup>, las diferencias ocurren por escasez de capacidad de refinería. Y se postula, que las diferencias que aparecen luego de 2011 se deben a un exceso de producción en EEUU, mientras que el Brent subió a causa de los movimientos árabes. Como los barriles con precio de WTI no pudieron ser fácilmente transportados al golfo, no fue posible sacar provecho de estos problemas, generando una diferencia de precios sostenida.

Esta tesis fue realizada con el WTI en mente, debido a que los precios de EEUU tienden a ser más coherentes con lo que se esperaría pagar dada la proximidad geográfica. Ahora bien, en términos prácticos, la relevancia mas importante recae en los retornos, los cuales, como se muestra en las gráficas de precios, deberían ser similares, dados los movimientos de ambas series.

---

<sup>2</sup>'WTI Prices Don't Reflect International Oil Market, Study Says'. Bloomberg.com. 13 April 2007. Retrieved 16 April 2013.



**Figura 5.1:** Precios del Brent y el WTI para el período estudiado.

(Fuente: Elaboración Propia)

## 5.2. Contexto Latinoamericano y relación con el crudo

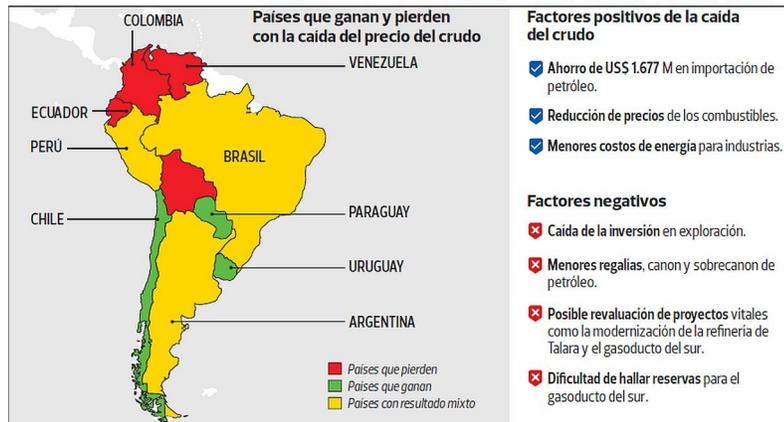
La realidad de la región es bastante compleja. Por un lado hay países productores, entre los que se encuentran Venezuela, México, Brasil, Colombia y Ecuador. Países de exportaciones modestas como Perú, Bolivia y Argentina. Y países importadores netos de energía, como es el caso de Chile.

Como la región está compuesta por diferentes países con diferente relación con el petróleo, existirá diferente influencia de los precios del mismo entre ellos. Mientras los países exportadores netos sentirán cierto nivel de estrés ante la reducción de los ingresos, los importadores observarán beneficios causados por los ahorros públicos y de los consumidores finales de combustible, es decir, personas y empresas.

La caída de los precios del petróleo desde 2015, observable en la Figura (5.1), posee impactos diferentes según la estructura descrita anteriormente, los cuales pueden verse resumidos en la Figura (5.2). El impacto, evidentemente, no estará igualmente repartido y serán los más dependientes de las exportaciones quienes se verán más afectados.

Dadas las características de Latinoamérica, se suele dar un círculo vicioso en que los precios

de la energía poseen gran influencia. El descenso de los ingresos petroleros repercute en las potenciales políticas sociales y proyectos de inversión, lo cual reduce el respaldo popular en los gobiernos, y con ello surgen populismos y sus problemas relacionados. Por ello, estudiar la influencia de la tendencia del petróleo resulta fundamental.



**Figura 5.2:** Países afectados por caídas del precio del crudo.

(Fuente: El Comercio Perú)

A continuación se resumen algunos de los elementos más relevantes de los países en estudio. La mayoría de datos están sacados de **The CIA World Factbook**.

## Brasil

- Economía más grande de Latinoamérica. Su Producto Interior Bruto se ha visto incrementado en un 284 % en el período que comprende los años del 2000 al 2011.
- Posee una producción de 2,255,000 barriles diarios, haciéndolo el décimotercero mayor productor del mundo, y el tercero de Latinoamérica. Esto le permite ser autosuficiente desde 2006, siendo las exportaciones de un 6,1 % del total.
- En los últimos años ha sufrido una gran crisis política dada por un escándalo de corrupción centrado en Petrobras.
- Posee el 73 % de sus reservas en aguas profundas, por lo que ha sido pionero en dicho tipo de extracción.

- El índice estudiado corresponde al Ibovespa (Bolsa de Valores do Estado de São Paulo). Este índice lo componen cerca de 50 empresas que cotizan en la Bolsa de São Paulo.

## Chile

- Importador neto de petróleo, solo produce unos 6.600 barriles al día, focalizados en Magallanes, siendo el 81vo productor del mundo y el octavo de Latinoamérica.
- Su falta de yacimientos de petróleo lo hace muy dependiente de los precios internacionales.
- El índice bursátil a analizar es el Índice de Precio Selectivo de Acciones (IPSA), elaborado por la Bolsa de Comercio de Santiago. Está compuesto, de forma ponderada, por las 40 empresas de mayor capitalización bursátil. Fue el más rápido en recuperarse luego de la crisis de 2008.

## México

- Tercera economía más grande de Latinoamérica.
- Produce 2,459,000 barriles de crudo diarios, lo que lo convierte en el 11vo mayor productor del mundo y segundo de Latinoamérica.
- Ha experimentado grandes cambios relacionados con los hidrocarburos, pasando de ser el 80 % de sus exportaciones en 1980 a solo el 6 % en 2016.
- Tras tres años consecutivos de pérdidas, producto de una gestión deficiente, se realizó un rescate público de Pemex.
- En éste análisis se considera el Índice de Precios y Cotizaciones (IPC), el principal índice bursátil de la Bolsa Mexicana de Valores. Este índice considera las 37 empresas con mayor liquidez del mercado.

## Argentina

- Bajo el gobierno de los Kirchner, en 2012, YPF pasó a manos del gobierno argentino. El país pasó de ser exportador de hidrocarburos a importador neto.
- Posee una considerable riqueza petrolera y gasífera, que le permite organizar una cadena de producción petroquímica. Dichos productos corresponden a cerca del 20 % de las exportaciones.
- El país produce 532,100 barriles diarios, convirtiéndolo en el 28vo productor del mundo, y el sexto de Latinoamérica.
- Si bien no se estudia en éste análisis, en el primer semestre de 2017 la producción de petróleo y gas en Argentina fue el peor en 25 años y estuvo apenas por encima de 1981, retrocediendo su nivel de producción 36 años.
- Ha experimentado varias recesiones en los años considerados en el análisis.
  - 2009: Afectada por la crisis financiera mundial.
  - 2012: Considera una baja de 1,1 % en el PIB.
  - 2013: Devaluación de la moneda e inflación de 38,5 %
  - 2015: Aumentan tasas de interés para contener inflación, pero aumenta al 39,2 % y pobreza aumenta al 33 %.
- A pesar de lo anterior, entre 2003 y 2008, su PIB crece un 43,6 %.
- El Merval (por MERcado de VALores) es el principal índice del Mercado de Valores de Buenos Aires. Las acciones se seleccionan en base a capitalización y número de transacciones.

## Perú

- Importador de crudo. A 2014 la exportación de crudo correspondía a un 1,2 % del total de exportaciones.

- Producción de 69,300 barriles diarios (2014) lo convierte en el séptimo exportador de crudo de LA y 51vo del mundo.
- Economía dependiente de fluctuaciones de precios globales. Debido a gran crecimiento económico, su dependencia del crudo se ha hecho más fuerte.
- En la última década ha intentado mejorar su capacidad exportadora de petróleo, lo que ha recibido numerosas críticas por potenciales daños al medio ambiente, ya que la dependencia de la economía en precios del petróleo causa problemas cuando está a la baja. Se postula que deberían concentrar sus esfuerzos en la exportación de productos con alto desarrollo tecnológico, en vez de depender de los recursos naturales. Varios proyectos han quedado detenidos debido a las bajas del último par de años.
- Petrobras y Talisman Energy Inc. abandonan producción de petróleo en 2012.
- El índice bursátil analizado es el S&P/BVL Peru General, que sigue la historia del IGBVL. Está diseñado para ser el benchmark del mercado peruano reflejando la tendencia promedio de las principales acciones cotizadas en Bolsa. Asume la reinversión de los dividendos y agrega las acciones de la Bolsa de Valores de Lima que cumplen ciertos requisitos de capitalización, liquidez y días de trading.

## Colombia

- Cuarta economía más grande de Latinoamérica.
- Exportador neto de crudo, alcanzando el 42,41 % de sus exportaciones en 2012. El refinado alcanza el 7,3 % el mismo año.  
Otras exportaciones importantes corresponden a briquetas de carbón, café y oro.
- Cuarto productor de Latinoamérica, y vigésimo del mundo.
- 989,900 barriles producidos por día al 2014. Alcanza un máximo en 2013 con un millón diario.

- Para el análisis se utiliza el COLCAP, uno de los índices bursátiles de la Bolsa de Valores de Colombia. Refleja la ponderación de los precios de las 20 acciones más líquidas donde el valor de Capitalización Bursátil ajustada de cada compañía determina su participación dentro de dicho índice. Debido a sus características se ha convertido en el principal índice de referencia del mercado de renta variable colombiano.

## Estados Unidos

- Mayor productor de petróleo de América, con 8,653,000 barriles diarios. Tercero del mundo. A su vez, es el mayor importador en el mundo, con un 8,2 % de las importaciones en crudo.
- Segunda economía más grande del mundo, luego de China. Se caracteriza por abundantes recursos naturales, en que se encuentra el petróleo.
- El West Texas Intermediate (WTI) es una corriente de crudo producida en Cushing, Oklahoma (Texas), y sirve como benchmark del precio mundial de crudo, aunque, como se mencionó anteriormente, está siendo paulatinamente reemplazado en influencia por el Brent.
- A modo de benchmark, se utilizan los dos índices bursátiles más importantes.
  - Standard & Poor's 500 (SP500): Se basa en la capitalización bursátil de las 500 empresas más grandes que cotizan en la bolsa de New York o NASDAQ. Si diferencia de el Dow Jones Industrial Average en la diversidad de rubros que lo conforman. Se considera el más representativo del mercado de acciones en EEUU y el mejor marcador de tendencias en la economía Norteamericana. Para la inclusión se ponderan los siguientes elementos: capitalización bursátil, liquidez, domicilio, capital flotante, clasificación del sector, viabilidad financiera, periodo de tiempo durante el cual ha cotizado en bolsa y ser negociada en la bolsa de valores.

- Dow Jones Industrial Average (DJIA): Refleja el comportamiento del precio de la acción de las 30 compañías industriales más importantes y representativas de Estados Unidos. El índice no solo se ve afectado por el desempeño del sector industrial, sino que eventos políticos externos como guerra y terrorismo, así como desastres naturales.



## 5.3. Análisis estadístico de las series

### 5.3.1. Estadísticas descriptivas

En la tabla (5.1) se presentan los estadísticos más importantes para caracterizar las series de retornos logarítmicos y las pruebas estadísticas mencionadas en el capítulo de procesos estocásticos.

Como podría esperarse, es posible observar cierta homogeneidad en las series de retornos. Las medias típicamente son muy cercanas al cero, con grandes fluctuaciones indicadas por los valores mínimos y máximos de los retornos. Tanto la curtosis y asimetría se alejan ampliamente de lo que podría esperarse de una distribución normal, lo cual es posible comprobar mediante el test Jarque-Bera, que indica rechazo de la hipótesis nula de normalidad al 1 % de significancia. En todas las series se rechaza la hipótesis de raíz unitaria al 1 % de significancia.

Además, según los estadísticos de Ljung-Box ( $Q(10)$ ) y  $ARCH(10)$ , para todos existe dependencia de la volatilidad de los rezagos, siendo la única excepción el Merval, que no pasa la prueba de Ljung-Box.

Estos últimos resultados pueden justificarse por la estructura de los gráficos de retorno, los cuales presentan cúmulos de volatilidad en períodos económicos turbulentos. Dichos gráficos se encuentran disponible en el Anexo.

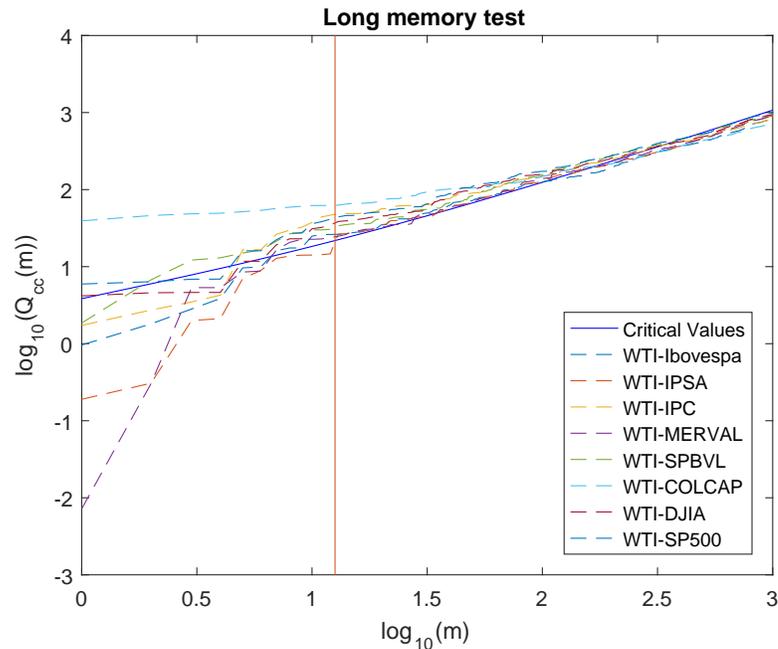
	WTI	IBOVESPA	IPSA	IPC	MERVAL	SPBVL	COLCAP	DOW JONES	SP500
Mean (%)	0.005	0.036	0.033	0.049	0.098	0.053	0.032	0.025	0.026
S.D (%)	2.688	1.904	1.131	1.362	2.185	1.882	1.748	1.203	1.301
Max. (%)	27.564	15.473	15.025	11.112	14.290	22.982	14.741	10.326	10.424
Min. (%)	-12.827	-12.096	-7.236	-7.266	-13.190	-12.653	-22.325	-9.637	-9.470
Kurt.	13.528	8.581	19.694	9.335	7.620	22.266	20.009	13.035	12.942
Skew.	0.674	0.092	0.302	0.139	-0.349	0.671	-1.019	-0.241	-0.418
J-B	11933.025***	3306.243***	29590.145***	4264.528***	2315.469***	39551.769***	31117.362***	10702.726***	10555.938***
ADF	-53.166***	-51.491***	-44.559***	-46.998***	-48.706***	-44.825***	-43.452***	-56.001***	-56.202***
Q(10)	32.235***	34.387***	59.469***	71.895***	12.499	56.228***	103.277***	65.070***	70.559***
ARCH(10)	408.419***	687.662***	298.446***	572.424***	339.299***	288.260***	570.148***	784.484***	803.852***
Obs.	2542	2542	2542	2542	2542	2542	2542	2542	2542

TABLA 5.1: Las estadísticas descriptivas de las series de retornos. J-B representa el estadístico Jarque-Bera de normalidad, ADF representa la prueba de Dickey-Fuller Aumentada,  $Q(10)$  representa el estadístico Ljung-Box-Pierce Q con 10 rezagos y  $ARCH(10)$  es el test de Engle con 10 rezagos.

\*\* Indica rechazo al 5 % de significancia.

\*\*\* Indica rechazo al 1 % de significancia.

### 5.3.2. Prueba de memoria larga en la correlación



**Figura 5.3:** Estadístico de memoria larga en la correlación entre los índices bursátiles y el WTI.

(Fuente: Elaboración Propia)

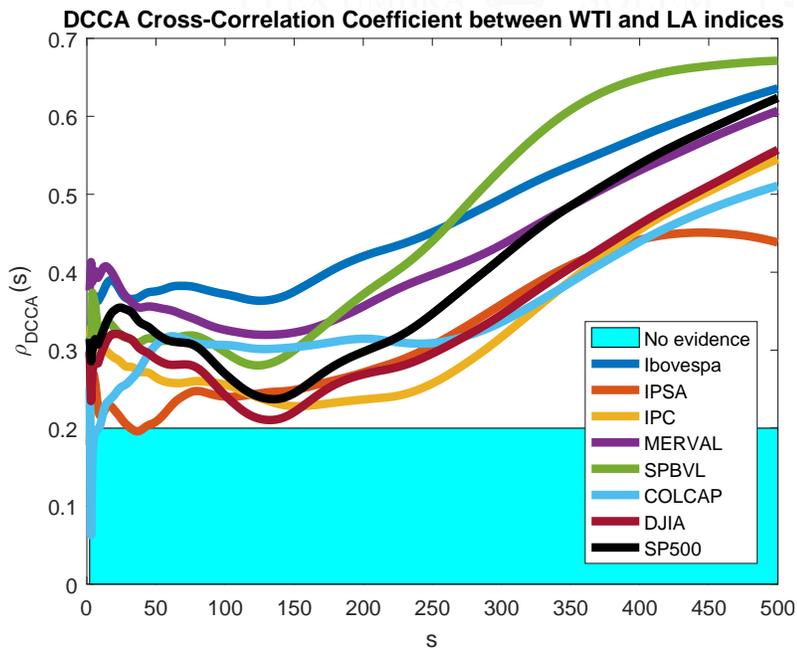
De acuerdo a la prueba de memoria larga presentada en forma logarítmica en la figura (5.3), es posible encontrar evidencia de dicha característica para todas las series cuando se analiza  $m \gtrsim 15$ , donde para todos los casos se supera el valor del estadístico de prueba, pudiendo rechazarse la hipótesis nula de ausencia de correlación de memoria larga para todos los pares.

### 5.3.3. Coeficiente de correlación DCCA

Con el objetivo de cuantificar la correlación se aplica la prueba de Zebende [46], la cual permite identificar la correlación entre series de comportamiento complejo. Esto es especialmente útil ante la característica de autoafinidad, donde se puede caracterizar la autocorrelación mediante leyes de potencia. Es necesario especificar que la prueba aquí planteada se basa en desarrollo empírico mediante simulaciones de procesos ARFIMA en que se puede variar la influencia que tiene una serie sobre otra, por lo que su aplicación solo permite obtener evidencia en el aspecto monofractal. Si bien resulta una limitación

importante, el uso de este estadístico ha presentado muy buenos resultados aun bajo un esquema multifractal, por lo que se utiliza como **uno** de los pasos para determinar correlación con memoria larga.

En la Figura (5.4) es posible observar como cambia la cuantificación de esta correlación con la ventana de tiempo estudiada. Se observa que ante ventanas pequeñas, si bien existe correlación fuera del área turquesa en que no es posible establecer significancia [55], esta experimenta un notable aumento al elevarse la ventana considerada, obteniéndose evidencia más clara de la presencia de la correlación al estudiar un mayor intervalo de tiempo de las series.



**Figura 5.4:** Coeficiente de correlación DCCA. La zona turquesa indica los valores en que no existe evidencia estadística de correlación.

(Fuente: Elaboración Propia)

Para mayor detalle en términos numéricos, se presenta en la Tabla (5.2) el valor del coeficiente  $\rho_{DCCA}$ . Como se mencionó antes, para ventanas grandes.

Con ambos estadísticos entregando información sobre la existencia de correlación con memoria larga, se puede dar inicio al análisis principal mediante el MF-DCCA y el MF-DFA. Sin embargo, primero se estudiará en detalle las características de la memoria larga en la autocorrelación.

s	s=16	s=32	s=64	s=128	s=256	s=512
WTI-Ibovespa	0.389	0.366	0.382	0.364	0.456	0.642
WTI-IPSA	0.229	0.199	0.234	0.245	0.312	0.432
WTI-IPC	0.296	0.279	0.258	0.237	0.263	0.554
WTI-MERVAL	0.405	0.362	0.349	0.320	0.401	0.615
WTI-SPBVL	0.335	0.313	0.315	0.281	0.450	0.672
WTI-COLCAP	0.235	0.262	0.318	0.301	0.310	0.517
WTI-DJIA	0.317	0.314	0.281	0.211	0.301	0.567
WTI-SP500	0.343	0.349	0.311	0.239	0.355	0.632

TABLA 5.2: Valor del  $\rho_{DCCA}$  para diferentes ventanas s.

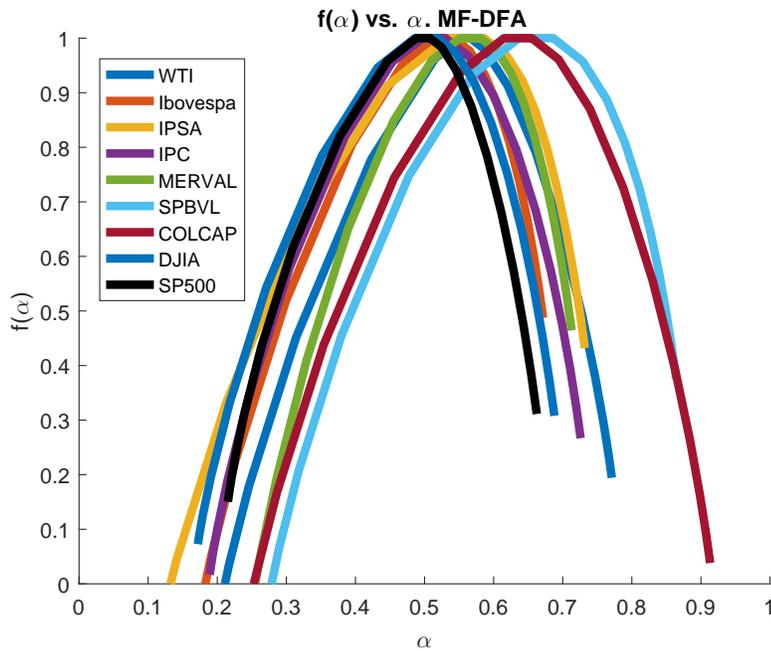
(Fuente: Elaboración Propia)

## 5.4. Resultados MF-DFA

Mediante la aplicación del MF-DFA es posible obtener mucha información en solo un par de gráficos. El primero, la Figura (5.5), confirma que existe multifractalidad en la correlación, la cual resulta ser muy alta. Es posible observar este alto nivel en las primeras columnas de la Tabla (5.4), donde se puede observar que el  $\alpha$  es bastante ancho, pudiendo abandonarse con total seguridad el supuesto de monofractalidad.

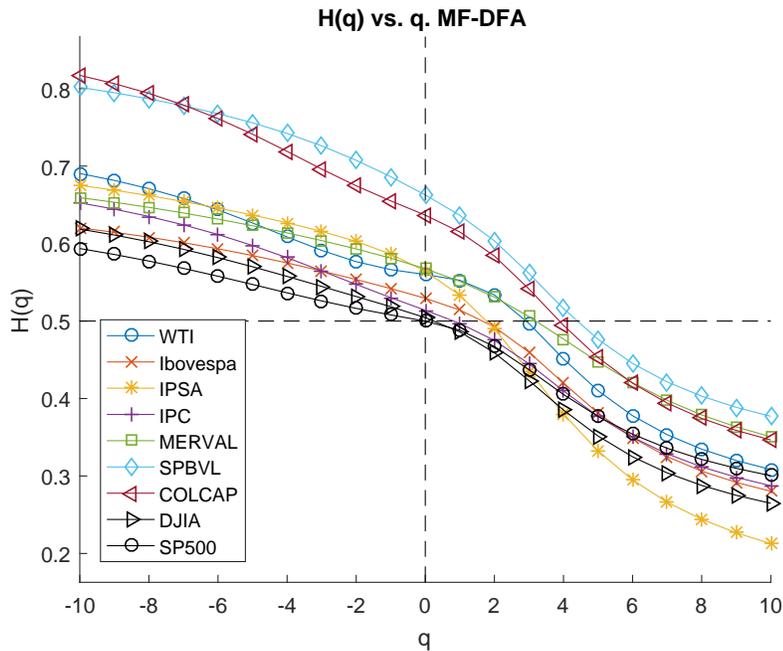
Esta conclusión toma fuerza al observar la Figura (5.6), el cual indica que existen diferentes valores del Hurst generalizado  $H(q)$  para los diferentes niveles de fluctuación. Al observar los valores de dichos exponentes en las últimas columnas de la Tabla (5.4), es posible concluir que el mayor grado de multifractalidad en la autocorrelación de las series está dado por el de las grandes fluctuaciones, indicadas por el orden  $q > 0$  de los momentos. Es posible observar que el comportamiento de la autocorrelación es bastante homogéneo para todas las series, destacando solo el SPBVL y el COLCAP, los que presentan un comportamiento extremadamente persistente para fluctuaciones pequeñas.

Al realizar el análisis de eficiencia, disponible en las Tablas (5.3) y (5.5), se encuentra que los dos indicadores multifractales arrojan resultados similares, solo cambiando el orden de algunas series, pero manteniendo al SP500 como el más eficiente. Ahora bien, estos indicadores evalúan eficiencia basándose en el supuesto, comprobado por este análisis, de que las series poseen autocorrelación multifractal. Si consideramos a las series como monofractales, el Ibovespa aparece primero en el ranking, pero como dicho supuesto es evidentemente erróneo, se desecha esta conclusión.



**Figura 5.5:** Comportamiento multifractal de las series en estudio.

(Fuente: Elaboración Propia)



**Figura 5.6:** Exponente de Hurst  $H(q)$  generalizado.

(Fuente: Elaboración Propia)

	$H(q)_{min}$	$H(q)_{max}$	$\Delta H$	ED	DME	ME
WTI	0.308	0.690	0.382	0.034	0.115	0.191
Ibovespa	0.280	0.620	0.340	0.008	0.100	0.170
IPSA	0.213	0.675	0.462	0.010	0.147	0.231
IPC	0.287	0.653	0.365	0.025	0.106	0.183
MERVAL	0.351	0.659	0.309	0.031	0.098	0.154
SPBVL	0.377	0.802	0.425	0.104	0.167	0.212
COLCAP	0.346	0.817	0.471	0.084	0.166	0.235
DJIA	0.265	0.619	0.354	0.041	0.103	0.177
SP500	0.300	0.594	0.293	0.033	0.082	0.147

TABLA 5.3: Resultados del análisis de eficiencia.

(Fuente: Elaboración Propia)

	$\alpha_{min}$	$\alpha_{max}$	$\Delta\alpha$	$\Delta H_{q>0}$	$\Delta H_{q<0}$	$\Delta H_{q>0}/\Delta H_{q<0}$
WTI	0.203	0.771	0.568	0.252	0.130	1.930
Ibovespa	0.175	0.672	0.495	0.249	0.091	2.740
IPSA	0.088	0.732	0.644	0.352	0.111	3.171
IPC	0.189	0.726	0.537	0.227	0.139	1.634
MERVAL	0.236	0.713	0.477	0.217	0.092	2.364
SPBVL	0.273	0.864	0.591	0.286	0.139	2.058
COLCAP	0.235	0.913	0.678	0.290	0.180	1.608
DJIA	0.172	0.688	0.516	0.240	0.114	2.096
SP500	0.215	0.662	0.447	0.199	0.094	2.127

TABLA 5.4: Multifractalidad de la serie e influencia de las pequeñas y grandes fluctuaciones en el espectro del Hurst generalizado.

(Fuente: Elaboración Propia)

Ranking	ED	DME	ME
1	Ibovespa	SP500	SP500
2	IPSA	MERVAL	MERVAL
3	IPC	Ibovespa	Ibovespa
4	MERVAL	DJIA	DJIA
5	SP500	IPC	IPC
6	DJIA	IPSA	SPBVL
7	COLCAP	COLCAP	IPSA
8	SPBVL	SPBVL	COLCAP

TABLA 5.5: Ranking de eficiencia.

(Fuente: Elaboración Propia)

### 5.4.1. Cambios temporales en la autocorrelación

Al analizar los cambios temporales en la autocorrelación monofractal para una ventana de 4 años, es posible observar que la mayoría de índices presenta un comportamiento bastante estable. Las excepciones corresponden al WTI y el Ibovespa, quienes presentan un fuerte cambio estructural a partir de mediados de 2013. Esto puede deberse al período que está capturando la ventana de 4 años a partir de esa fecha. Tanto el Ibovespa como el WTI presentan un comportamiento más estable en el período 2009-2010, lo cual evidentemente afecta las características de su autocorrelación.

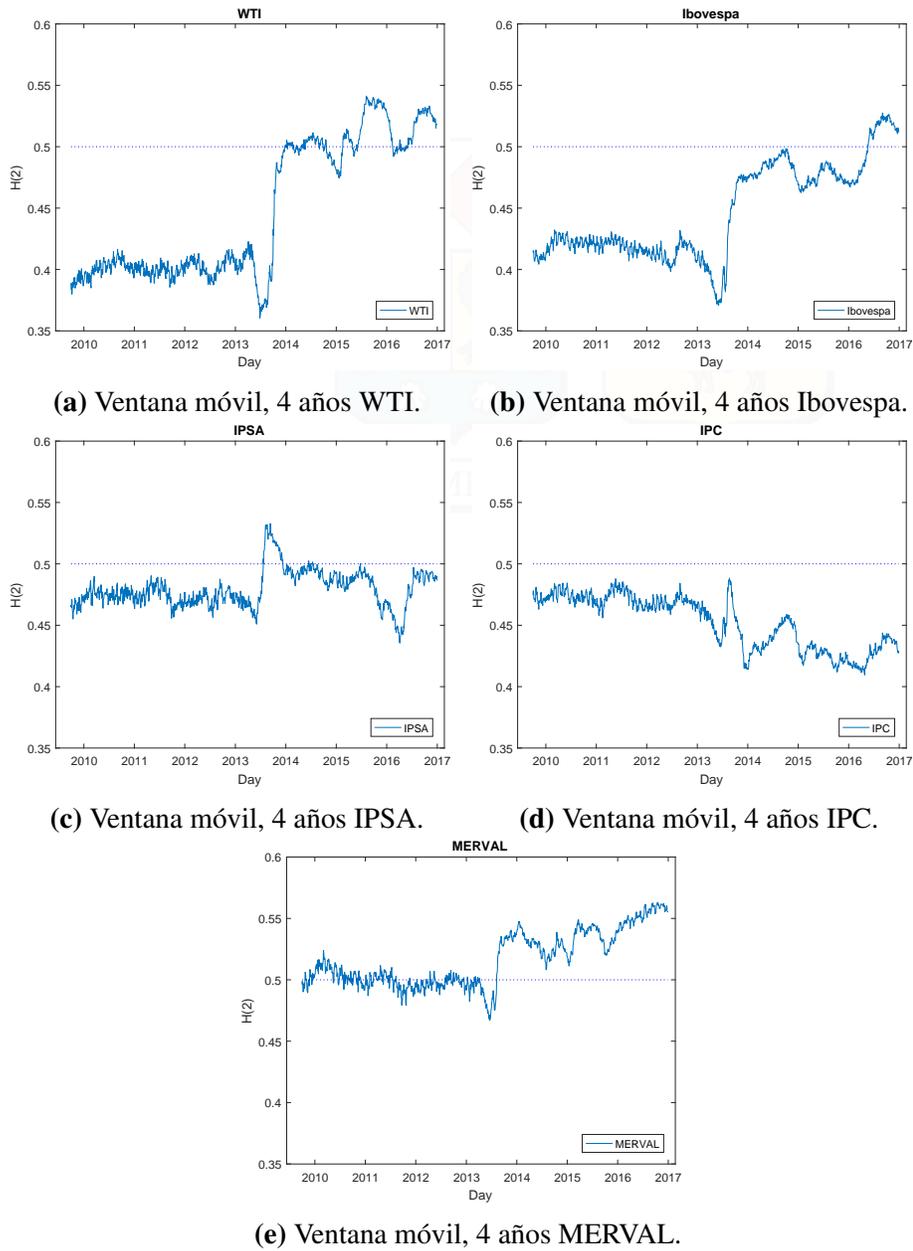
A pesar de lo anterior, el Ibovespa si mantiene un comportamiento antipersistente relativamente constante. A éste se suman el IPSA, IPC, DJIA y SP500, los que poseen un comportamiento mucho más estable y que puede ser comprobado al analizar la Tabla (5.6). Por otro lado, el Merval presenta un comportamiento aleatorio desde el punto de vista monofractal, que luego de 2013 se vuelve ligeramente persistente. A su vez, el COLCAP y el SPBVL muestran un comportamiento altamente persistente durante prácticamente todo el período, con leves indicios de comportamiento aleatorio a mediados de 2013, y en el caso del COLCAP, también a mediados de 2015.

Las diferencias de comportamiento con respecto al análisis del período completo pueden estar dadas debido a que el comportamiento de escala para ambos casos posee importantes diferencias. El valor máximo de la escala a analizar, corresponde, para el análisis completo a 634, mientras que para la ventana de 4 años, sólo a 250. Esto puede dar indicios de cambios estructurales en el comportamiento de escala, que deben ser analizados en profundidad en el futuro.

	Mean	Median	S.D.	Max.	Min.	Persistente (%)	Antipersistente (%)	$\bar{E}D$
WTI	0.448	0.410	0.056	0.541	0.360	29.4	70.6	0.062
Ibovespa	0.447	0.427	0.038	0.527	0.371	8.1	91.9	0.056
IPSA	0.479	0.477	0.015	0.533	0.436	5.5	94.5	0.023
IPC	0.452	0.457	0.022	0.488	0.410	0	100	0.048
MERVAL	0.517	0.511	0.022	0.563	0.467	69.1	30.9	0.021
SPBVL	0.544	0.541	0.019	0.601	0.485	98.6	1.4	0.044
COLCAP	0.539	0.547	0.029	0.596	0.467	88.3	11.7	0.043
DJIA	0.447	0.447	0.014	0.497	0.409	0	100	0.053
SP500	0.436	0.441	0.016	0.462	0.395	0	100	0.064

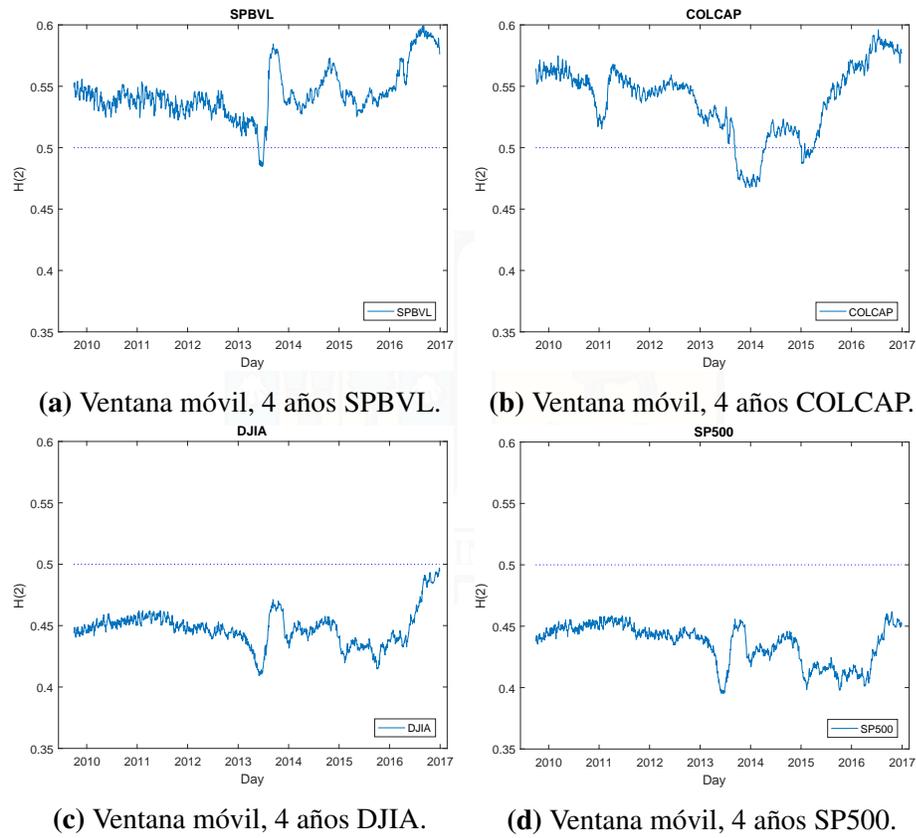
TABLA 5.6: Estadísticos del  $H(2)$  para una ventana de 4 años.

### 5.4.2. Ventana móvil de 4 años



**Figura 5.7:** Ventana móvil de 4 años para  $H(2)$  según MF-DFA.

(Fuente: Elaboración Propia)



**Figura 5.8:** Ventana móvil de 4 años para  $H(2)$  según MF-DFA..

(Fuente: Elaboración Propia)

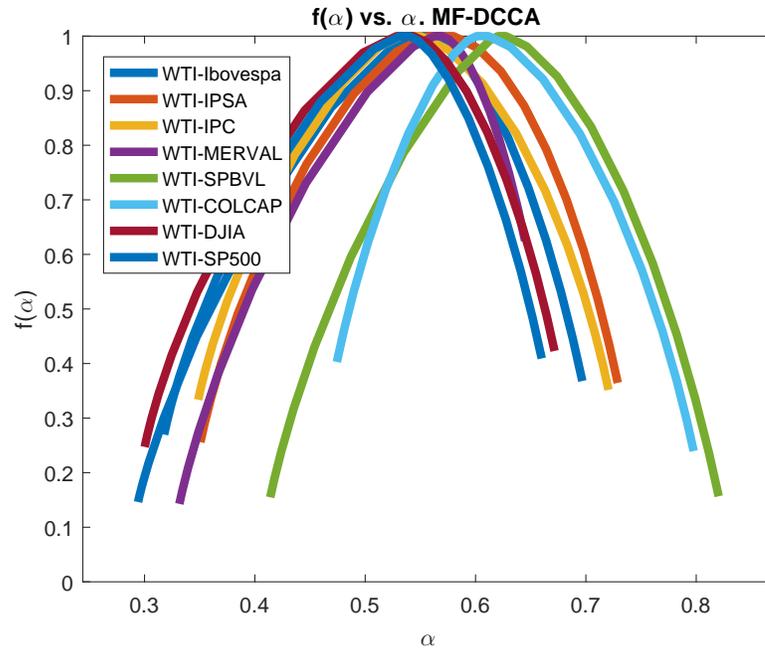
## 5.5. Resultados MF-DCCA

Nuevamente es posible encontrar multifractalidad en la correlación, como es posible observar en las curvas de la Figura (5.9). Como puede observarse en la Tabla (5.8), el comportamiento de  $\alpha$  no implica una multifractalidad tan pronunciada como en la autocorrelación, aunque sigue siendo lo suficientemente alta para desechar cualquier supuesto de monofractalidad.

Al analizar el coeficiente de correlación  $H_{xy}(q)$  en la Figura (5.10), se encuentra que se presenta un comportamiento mayoritariamente persistente para la mayoría de órdenes de fluctuación, destacándose los índices COLCAP y SPBVL, los cuales no presentan antipersistencia, ni siquiera ante grandes fluctuaciones. De hecho, su correlación deja de presentar dependencia de largo plazo.

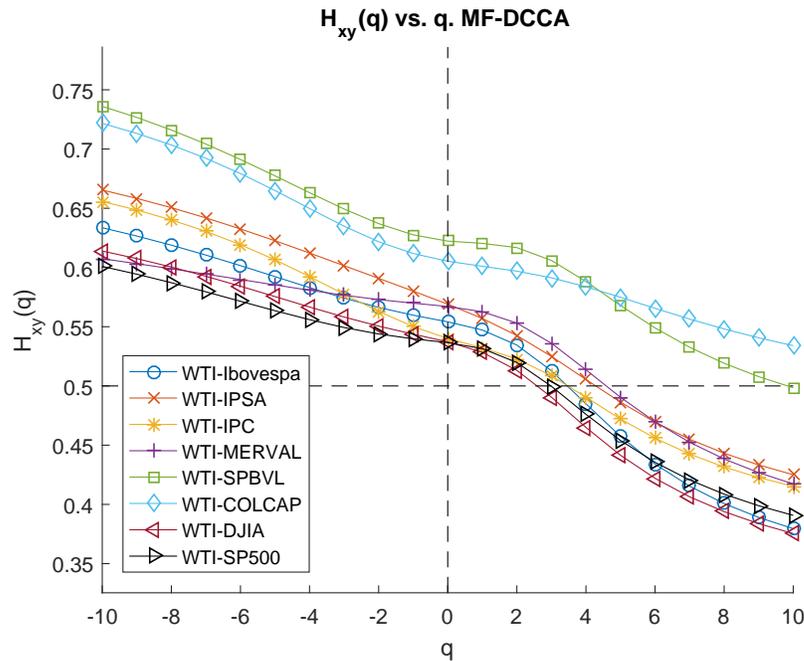
En la Tabla (5.7) se puede observar que el comportamiento de las grandes fluctuaciones presenta el mayor aporte a la naturaleza multifractal del coeficiente de correlación  $H_{xy}(q)$  para todos los países, excepto Colombia (COLCAP).

Para mayor información respecto al comportamiento de escala de las series, es posible encontrar los gráficos de las funciones de fluctuación monofractal ( $q = 2$ ) en el Anexo.



**Figura 5.9:** Comportamiento multifractal de las correlación.

(Fuente: Elaboración Propia)



**Figura 5.10:** Exponente de correlación  $H_{xy}(q)$ .

(Fuente: Elaboración Propia)

	$H_{xy}(q)_{min}$	$H_{xy}(q)_{max}$	$\Delta H_{xy}$	$\Delta H_{xy}(q > 0)$	$\Delta H_{xy}(q < 0)$	$\Delta H_{xy}(q > 0)/\Delta H_{xy}(q < 0)$
WTI-Ibovespa	0.380	0.634	0.254	0.175	0.079	2.198
WTI-IPSA	0.425	0.665	0.240	0.144	0.096	1.495
WTI-IPC	0.415	0.656	0.240	0.125	0.116	1.080
WTI-MERVAL	0.417	0.607	0.190	0.150	0.040	3.715
WTI-SPBVL	0.498	0.736	0.238	0.124	0.114	1.088
WTI-COLCAP	0.534	0.722	0.188	0.071	0.116	0.612
WTI-DJIA	0.375	0.614	0.239	0.162	0.077	2.110
WTI-SP500	0.391	0.601	0.210	0.146	0.064	2.266

TABLA 5.7: Exponente de correlación  $H_{xy}(q)$  y su dependencia del tamaño de las fluctuaciones.

	$\alpha_{min}$	$\alpha_{max}$	$\Delta\alpha$
WTI-Ibovespa	0.294	0.697	0.403
WTI-IPSA	0.351	0.729	0.378
WTI-IPC	0.349	0.721	0.372
WTI-MERVAL	0.332	0.645	0.313
WTI-SPBVL	0.414	0.820	0.406
WTI-COLCAP	0.474	0.798	0.323
WTI-DJIA	0.300	0.672	0.371
WTI-SP500	0.318	0.660	0.342

TABLA 5.8: Estadísticas del  $\alpha$ .

### 5.5.1. Cambios temporales en la correlación

Para estudiar los cambios temporales en la correlación, se procede a realizar los cálculos del coeficiente de correlación monofractal  $H_{xy}(2)$  para una ventana de 4 años. Si bien se demostró que las series poseen una estructura multifractal, el coeficiente monofractal permite estudiar cambios relativos a través del tiempo de una forma simple, y aproximable a comportamientos multifractales.

Al analizar la ventana de 4 años, cuyos gráficos se encuentran disponibles en la Figura (5.11), se pueden observar interesantes cambios en estructura. Los estadísticos de resumen se encuentran en la Tabla (5.9). Ibovespa, DJIA y SP500 presentan un comportamiento similar, presentando características antipersistentes hasta 2013, donde su comportamiento

se eleva notoriamente, pasando a ser aleatorio para el primer índice, y a obtenerse poca evidencia de antipersistencia para los estadounidenses. Esto puede ser explicado debido a que la ventana móvil de 4 años considera todos los cambios en estructura de correlación post crisis subprime, por lo que los coeficientes a partir de ese período no están contaminados por la turbulencia de la crisis.

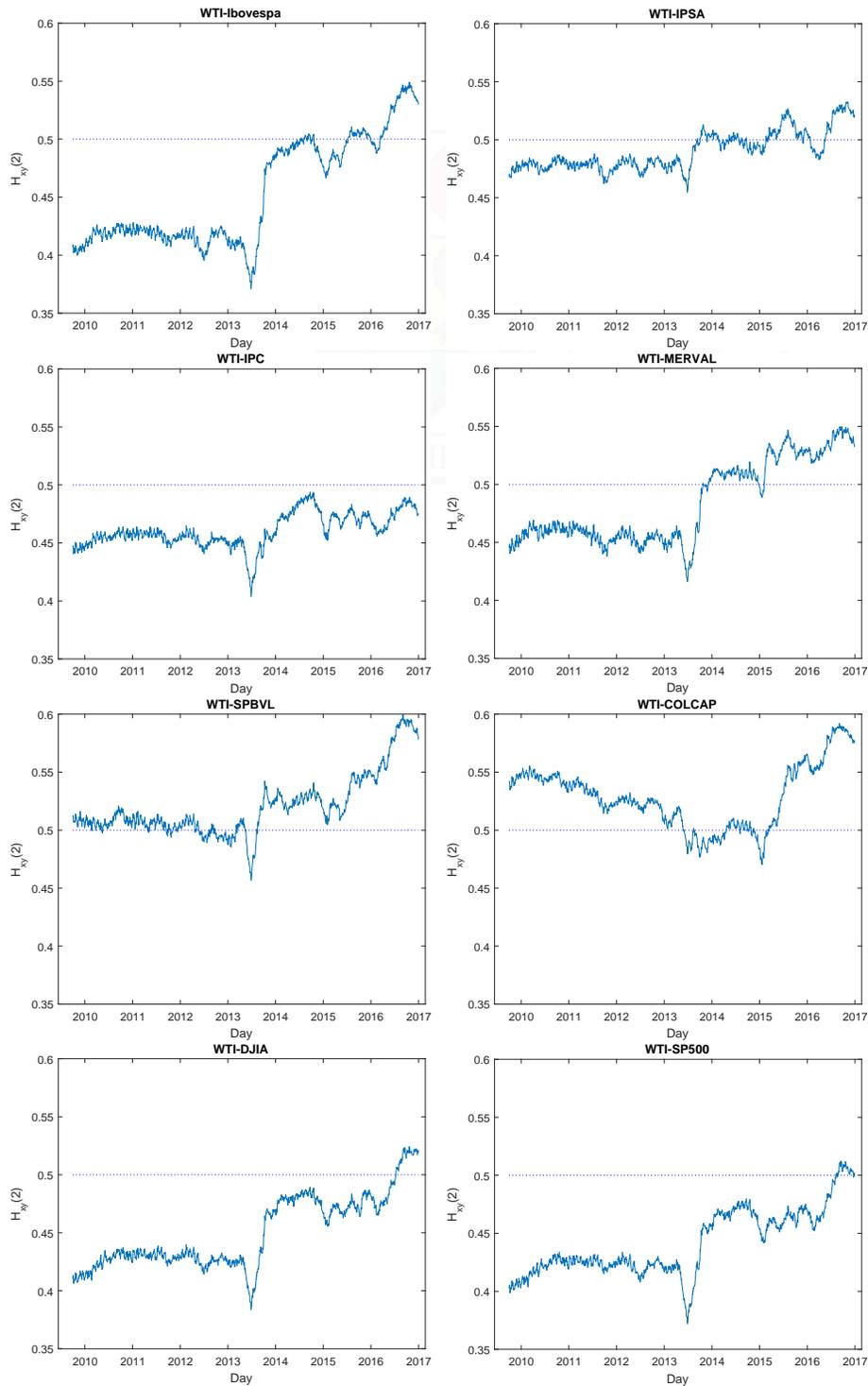
Algo similar ocurre en el caso de Merval, quien posee un comportamiento antipersistente en la correlación, pero que sin embargo aumenta a mediados de 2013, volviéndose persistente. El IPC presenta el comportamiento más homogéneo durante el período, observándose sólo correlación antipersistente. Por otro lado, el SPBVL, presenta solo correlación de corto plazo (o ausencia de correlación) hasta mediados de 2013, donde su comportamiento se vuelve cada vez más persistente. El Colcap posee un comportamiento persistente hasta mediados de 2013, donde deja de mostrar evidencia de correlación con memoria larga durante casi dos años, para posteriormente subir sostenidamente el nivel de persistencia en su correlación con el WTI. Finalmente, el IPSA es quien presenta la correlación con menor evidencia de memoria larga, siendo bastante estable para todo el período en estudio.

El notorio cambio en algunas de las correlaciones a mediados de 2013 está dado por la captura del tercer período de interés estudiado, el comprendido entre 2009 y 2014. En dicho período se encuentra que existe una tendencia al alza del precio del petróleo, y sigue una estructura bastante diferente a los períodos anteriores, especialmente el 2007-2009.

	Mean	Median	S.D.	Max.	Min.	Persistente (%)	Antipersistente (%)	$\bar{ED}$
WTI-Ibovespa	0.453	0.424	0.046	0.549	0.371	19.7	80.3	0.054
WTI-IPSA	0.490	0.484	0.016	0.533	0.455	29.6	70.4	0.016
WTI-IPC	0.461	0.458	0.014	0.494	0.404	0.0	100.0	0.039
WTI-Merval	0.485	0.464	0.036	0.550	0.416	41.6	58.4	0.035
WTI-SPBVL	0.520	0.511	0.027	0.600	0.457	84.7	15.3	0.023
WTI-Colcap	0.528	0.526	0.027	0.592	0.470	82.7	17.3	0.032
WTI-DJIA	0.450	0.434	0.031	0.524	0.384	6.3	93.7	0.052
WTI-SP500	0.442	0.429	0.028	0.512	0.372	4.5	95.5	0.059

Tabla 5.9: Estadísticos del  $H_{xy}(2)$  para una ventana de 4 años.

## Ventana móvil de 4 años



**Figura 5.11:** Ventana móvil de 4 años para  $H_{xy}(2)$  según MF-DCCA.

(Fuente: Elaboración Propia)

## 5.6. Resultados MF-ADCCA

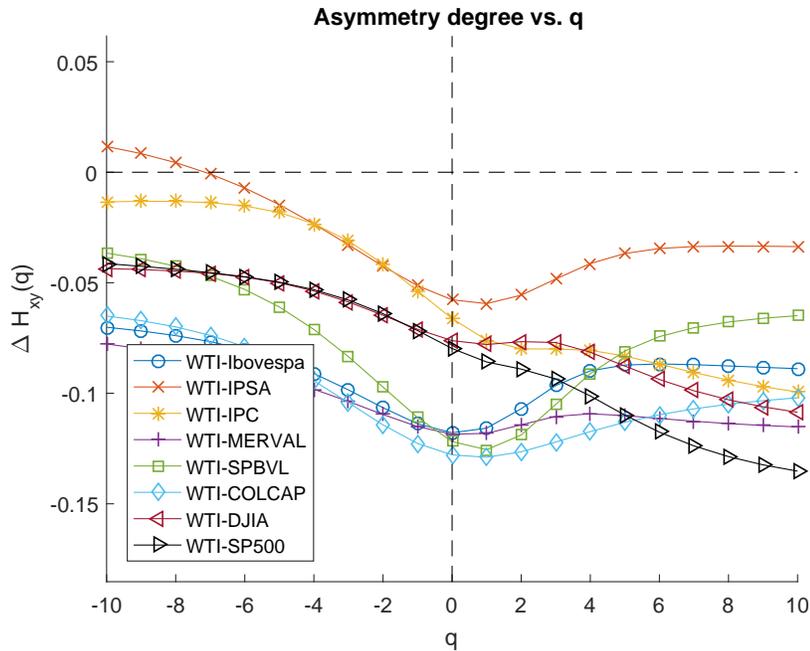
### 5.6.1. Asimetría según tendencia del crudo

Como es posible observar en la Figura (5.13), el coeficiente de correlación presenta grandes diferencias al considerar tendencias positivas o negativas del precio del crudo. Esta asimetría es negativa para todos los países excepto el IPSA, quien presenta asimetría positiva para muy pequeñas fluctuaciones. Es posible apreciar que la asimetría tiende a aumentar para grandes fluctuaciones para el IPC, DJIA, SP500 y Merval. Por otro lado, el Ibovespa, IPSA, SPBVL y COLCAP presentan un máximo de asimetría en torno al orden de fluctuación 0, siendo relativamente simétrico en torno a este valor. Ello implica que la asimetría es máxima para cierto nivel de fluctuaciones, pero similar ante fluctuaciones muy pequeñas y muy grandes.

La evidencia que esto entrega deja bastante clara la influencia que tiene la tendencia del petróleo en los índices bursátiles Latinoamericanos. Ante una tendencia negativa, existirá una correlación con memoria larga notoriamente persistente, llegando a diferencias absolutas sobre 0,13.

Cuando se analiza el caso monofractal en la Tabla (5.10), la asimetría es apreciable en términos más claros. Esto puede interpretarse como una prueba de la influencia que tiene el WTI en los índices, debido a que su tendencia será fundamental en su dependencia. Chile (IPSA) vuelve a adquirir protagonismo al ser el país con el menor grado de asimetría monofractal, con un valor absoluto de 0,056, mientras que Colombia (COLCAP) presenta el mayor con una magnitud de 0,126.

Resulta interesante considerar que en el caso monofractal todos los índices, exceptuando el SPBVL, presentan correlación con muy poca evidencia de memoria larga para tendencias positivas del petróleo. Destacando solo el SPBVL con un valor persistente de 0,54, y los índices estadounidenses con un valor antipersistente de 0,46. Esto permite concluir que la correlación de largo plazo se ve más influenciada por tendencias del petróleo a la baja.



**Figura 5.12:** Grado de asimetría para cambios en la tendencia del crudo.

(Fuente: Elaboración Propia)

	$H_{xy}(2)$	$H_{xy}^+(2)$	$H_{xy}^-(2)$	$\Delta H_{xy}(2)$
WTI-Ibovespa	0.535	0.469	0.576	-0.107
WTI-IPSA	0.543	0.510	0.566	-0.056
WTI-IPC	0.523	0.474	0.554	-0.080
WTI-MERVAL	0.553	0.482	0.596	-0.114
WTI-SPBVL	0.617	0.540	0.659	-0.119
WTI-COLCAP	0.597	0.525	0.651	-0.126
WTI-DJIA	0.513	0.464	0.541	-0.077
WTI-SP500	0.520	0.463	0.552	-0.089

TABLA 5.10:  $H_{xy}(2)$ ,  $H_{xy}^+(2)$  y  $H_{xy}^-(2)$  para las series en análisis para diferentes tendencias del WTI.

(Fuente: Elaboración Propia)

## 5.6.2. Asimetría según tendencia del índice

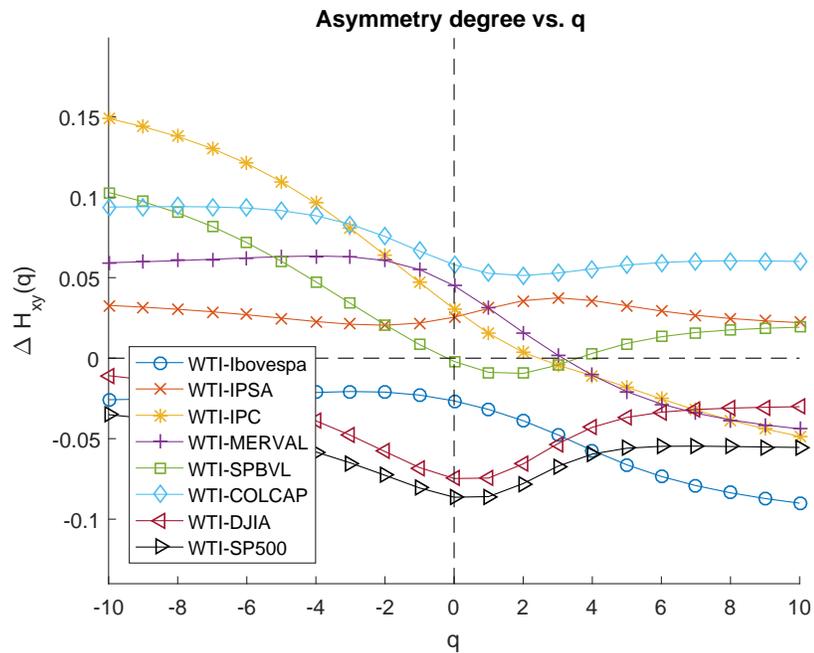
Al analizar los cambios en la correlación según la tendencia del índice, disponibles en la Figura (5.13), es posible apreciar que la asimetría no es tan homogénea. Para pequeñas

fluctuaciones, los índices Latinoamericanos presentan una asimetría positiva, es decir, cuando el índice está al alza, el exponente de correlación tiende a ser mayor. La única excepción para los países Latinoamericanos es Brasil (Ibovespa), quien presenta una muy pequeña asimetría negativa para pequeñas fluctuaciones.

Ahora bien, para el IPC y el Merval es posible observar el mismo comportamiento que en el caso anterior, donde para grandes fluctuaciones, el grado de asimetría se vuelve negativo. En contraste, el COLCAP y el IPSA presentan una asimetría relativamente homogénea. En este análisis destaca el nivel de asimetría de los índices de Estados Unidos, los cuales poseen una leve asimetría para pequeñas y grandes fluctuaciones, pero ésta se intensifica en fluctuaciones medianas. Esto queda mucho más claro al analizar el caso monofractal disponible en la Tabla (5.11), donde ambos índices presentan el mayor caso de asimetría. Siguiendo este análisis, se observa que el SPBVL y el COLCAP presentan gran nivel de persistencia en la correlación, tanto para alzas, como para bajas.

Lamentablemente, el análisis no resulta muy esclarecedor al considerar las diferencias entre las diferentes relaciones que poseen los países en estudio con el petróleo. Solo al considerar los índices de EEUU es posible observar diferencias más notorias.

Resulta interesante el hecho de que la mayoría de la correlación en el período estudiado resulta ser persistente, es decir, existe correlación positiva entre las variables petróleo e índices bursátiles, contrastando con lo que uno tendería a pensar. Si aumenta el precio del petróleo, el precio de la energía sube, por ende sería sensato suponer que los índices bursátiles bajen ante el mayor costo de operación. Esto no sucede así, y ha sido documentado por otros estudios. Las explicaciones de esto pueden relacionarse con la forma en que reaccionan los involucrados ante factores comunes globales. Al ser similar, ambos experimentarán movimientos similares, que pueden observarse como correlación. Se ha demostrado incluso que en sectores como transporte, quien se vería afectado de mayor forma teóricamente ante aumentos del petróleo, posee correlación directa.



**Figura 5.13:** Grado de asimetría para cambios en la tendencia del crudo.

(Fuente: Elaboración Propia)

	$H_{xy}(2)$	$H_{xy}^+(2)$	$H_{xy}^-(2)$	$\Delta H_{xy}(2)$
WTI-Ibovespa	0.535	0.497	0.535	-0.039
WTI-IPSA	0.543	0.546	0.511	0.036
WTI-IPC	0.523	0.506	0.502	0.004
WTI-MERVAL	0.553	0.556	0.540	0.016
WTI-SPBVL	0.617	0.593	0.602	-0.009
WTI-COLCAP	0.597	0.612	0.561	0.051
WTI-DJIA	0.513	0.481	0.547	-0.066
WTI-SP500	0.520	0.479	0.557	-0.078

**TABLA 5.11:**  $H_{xy}(2)$ ,  $H_{xy}^+(2)$  y  $H_{xy}^-(2)$  para las series en análisis para diferentes tendencias del índice.

(Fuente: Elaboración Propia)

### 5.6.3. Cambios temporales en la asimetría

#### 5.6.3.1. Ventana de 4 años

En la Figura (5.14) se presenta como cambia la asimetría a través del tiempo para una ventana móvil de 4 años. Lo primero que salta a la vista es el cambio que experimenta a mediados de 2013, donde todas las series experimentan un paso de una asimetría altamente negativa, a una positiva en el caso Latinoamericano, y a una bastante débil para el caso estadounidense. Esto puede deberse a que la ventana de 4 años esta capturando los efectos de la crisis subprime, y que el período considerado es bastante diferente a los dos anteriores. Posterior a ese período, se encuentra que vuelve a bajar paulatinamente o se estanca en un nivel mas simétrico.

Para todos los países, el período comprendido entre el 2008 y 2013 presenta gran similitud en el grado de asimetría independientemente de la tendencia considerada, excepto para Colombia. Sin embargo, posterior a 2013, Brasil y Argentina experimentan aumentos en la disparidad de la asimetría, y Colombia la mantiene, destacándose por la diferencia de asimetrías durante todo el período. El IPSA y los índices estadounidenses, experimentan un mayor grado de simetría luego de 2013.

Como se observa en las Tablas (5.12) y (5.13), la mayor parte del tiempo la asimetría es negativa para todos los índices, sea cual sea la tendencia considerada. Esto quiere decir que las tendencias a la baja de cualquiera de los dos elementos poseen mayor fuerza en la correlación.

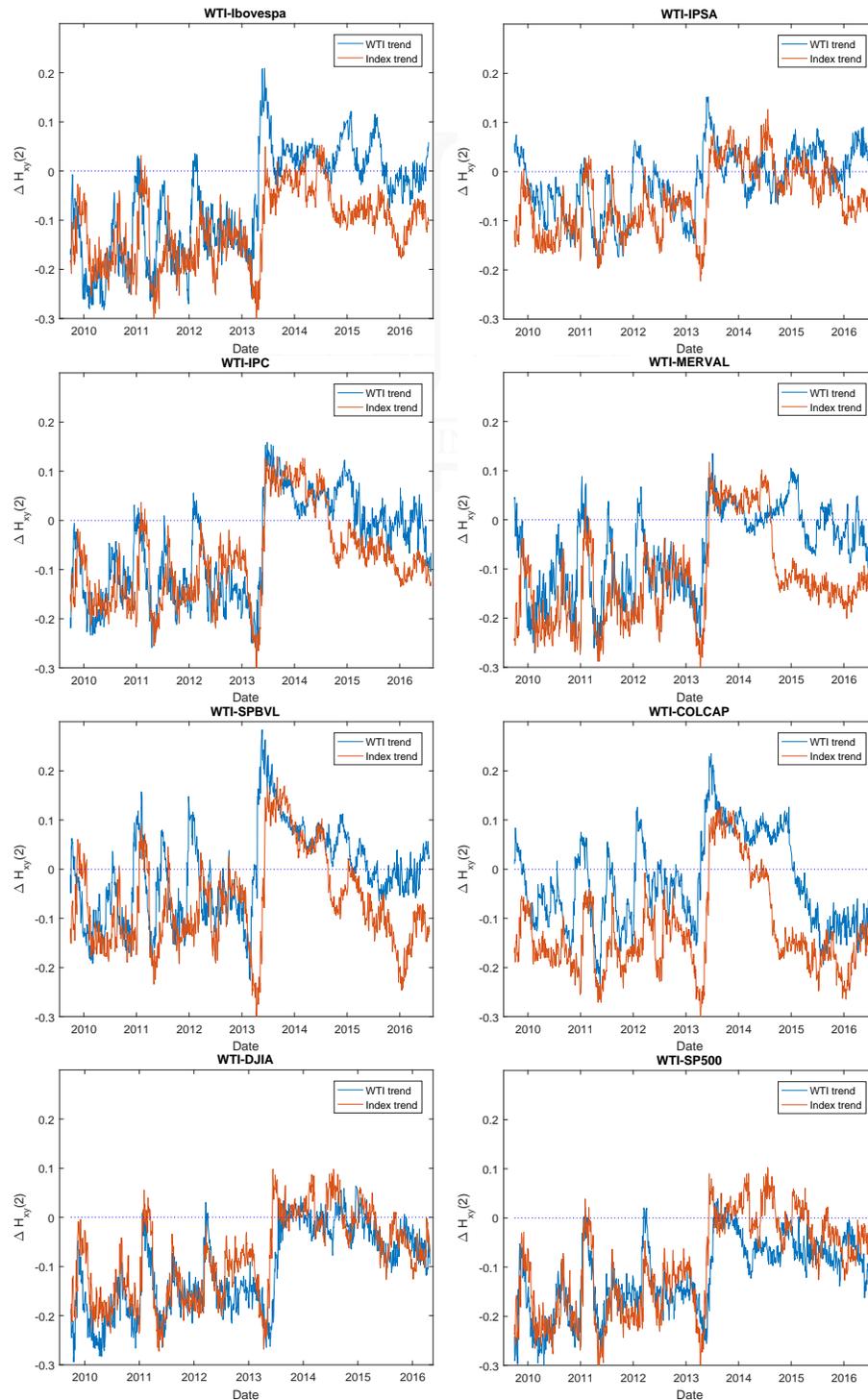
	Mean	Median	S.D.	Max.	Min.	$\Delta H_{xy}(2) \geq 0$ (%)	$\Delta H_{xy}(2) < 0$ (%)
WTI-Ibovespa	-0.058	-0.031	0.103	0.209	-0.283	37.451	62.549
WTI-IPSA	-0.017	-0.008	0.063	0.152	-0.195	45.111	54.889
WTI-IPC	-0.052	-0.036	0.096	0.159	-0.285	35.467	64.533
WTI-MERVAL	-0.064	-0.057	0.084	0.135	-0.277	27.818	72.182
WTI-SPBVL	-0.006	-0.011	0.089	0.284	-0.226	45.180	54.820
WTI-COLCAP	-0.028	-0.044	0.090	0.235	-0.233	35.179	64.821
WTI-DJIA	-0.101	-0.092	0.084	0.064	-0.294	12.822	87.178
WTI-SP500	-0.114	-0.101	0.073	0.039	-0.300	3.446	96.554

TABLA 5.12: Estadísticos del grado de asimetría  $\Delta H_{xy}(2)$  cuando se considera la tendencia del precio del crudo, para una ventana de 4 años.

	Mean	Median	S.D.	Max.	Min.	$\Delta H_{xy}(2) \geq 0$ (%)	$\Delta H_{xy}(2) < 0$ (%)
WTI-Ibovespa	-0.116	-0.110	0.072	0.053	-0.303	4.664	95.336
WTI-IPSA	-0.055	-0.058	0.067	0.127	-0.223	23.868	76.132
WTI-IPC	-0.077	-0.083	0.090	0.130	-0.300	18.107	81.893
WTI-MERVAL	-0.118	-0.133	0.092	0.117	-0.301	17.147	82.853
WTI-SPBVL	-0.067	-0.083	0.095	0.186	-0.300	23.388	76.612
WTI-COLCAP	-0.126	-0.152	0.091	0.127	-0.304	11.934	88.066
WTI-DJIA	-0.078	-0.072	0.083	0.098	-0.272	20.645	79.355
WTI-SP500	-0.095	-0.089	0.096	0.102	-0.309	19.479	80.521

TABLA 5.13: Estadísticos del grado de asimetría  $\Delta H_{xy}(2)$  cuando se considera la tendencia del índice, para una ventana de 4 años.

## Ventana móvil de 4 años



**Figura 5.14:** Ventana móvil de 4 años para el grado de asimetría  $\Delta H_{xy}(2)$  según MF-ADCCA.

(Fuente: Elaboración Propia)

## 5.7. Conclusiones generales

Luego de realizar una extensa introducción a los procesos multifractales, analizar sus bondades en el análisis y determinar como pueden superar los problemas de los modelos econométricos clásicos se inicia el análisis para 8 índices burstátiles.

Inicialmente, mediante pruebas estadísticas, se comprobó y cuantificó el nivel de correlación y de memoria larga en la misma.

Se realizó un análisis de la autocorrelación de las series mediante MF-DFA, de su correlación con el WTI mediante MF-DCCA y se analizó su asimetría mediante MF-ADCCA.

En el análisis de autocorrelación se encuentra que las series presentan alto nivel de ineficiencia, causado probablemente por las consecuencias de la crisis subprime. Se encuentra que el IPSA destaca en esta ineficiencia, y que la mayoría de las series presentan multifractalidad guiada principalmente por fluctuaciones grandes.

Se encuentra multifractalidad en la correlación, comprobándose las bondades de este análisis, para cada par de series en particular. El comportamiento resulta muy similar que en el caso bivariado, lo cual tiene sentido considerando la naturaleza de los métodos.

Se realiza la aplicación del MF-ADCCA en mercados Latinoamericanos, modelo que posee una serie de ventajas como que no es necesario especificar ningún tipo de modelo estadístico a la data, es fácil de implementar, los exponentes de escalamiento son calculados de forma directa y entregan numerosas propiedades, a la vez que entrega información sobre la multifractalidad de la correlación.

La utilidad de este análisis recae en que la asimetría en la correlación añade nuevas pistas respecto a la compleja relación entre los pares de series, sugiriendo que los mecanismos subyacentes entre sistemas complejos actúan de forma diferente dependiendo de la dirección de la dinámica. Para mercados financieros, este comportamiento de escala ayuda a revelar los riesgos del mercado financiero

Mediante el análisis de ventana móvil es posible observar que el análisis depende totalmente de la ventana analizada. Mediante ella es posible detectar los cambios en la dinámica de la autocorrelación y la correlación con el WTI. Es posible observar que

cualquier conclusión derivada de los métodos utilizados deben ceñirse al período utilizado. Cualquier intento de extrapolar los datos puede llevar a grandes errores ante potenciales períodos altos de volatilidad. Esto se debe a que probablemente exista un cambio estructural en el escalamiento de las fluctuaciones.

Uno de los problemas encontrados, es que si bien el método entrega gran cantidad de información respecto a las características de la correlación, no identifica muchas diferencias relevantes relacionados con la estructura de las importaciones/exportaciones de los países en estudio.

## Extensiones

Una de las principales extensiones sería realizar un análisis sectorial en Latinoamérica. Si bien el análisis agregado presentado en esta tesis permite obtener información y poner en práctica los métodos, se cae en una importante limitación metodológica al no considerar las diferencias entre la composición de los diferentes índices bursátiles, tanto entre ellos, como en la ponderación de las empresas de diferente sector productivo en cada uno.

Otra variación interesante podría ser respecto a la frecuencia de la data, es posible que diferentes resultados puedan obtenerse al aplicar la metodología semanal o mensualmente. Además, resultaría interesante desarrollar una forma de utilizar el método para generar estrategias de inversión, especialmente considerando las características de la autocorrelación que es posible obtener por medio del MF-DFA. También resultaría útil contrastar los presentes métodos con otros de similares objetivos para determinar potenciales debilidades. Es necesario ahondar en la posibilidad de crear coeficientes de correlación multifractal, tanto en el testeado como en la cuantificación, dado que los estadísticos aquí presentados se limitan al caso monofractal, y si bien existe evidencia clara de la multifractalidad y la memoria larga en las series, aún existe información que verificar.

Además, a modo de comparación, resultaría interesante estudiar otras zonas geográficas, especialmente Medio Oriente y Europa para identificar cambios en asimetría y la correlación.



## CAPÍTULO 6

## GLOSARIO

- **Dependencia de largo plazo:** También llamada memoria larga o persistencia de largo plazo. Se trata de un fenómeno estadístico que puede presentarse en el análisis de series temporales.

Se relaciona con la tasa de decaimiento de una función de autocorrelación con aumentos entre el intervalo de tiempo considerado. Se espera que en un proceso sin dependencia de largo plazo, dicha función de autocorrelación decaiga de forma exponencial. Si este decaimiento resulta más lento que el exponencial, se dice que el fenómeno presenta dependencia de largo plazo. Tiene relación directa con los procesos autosimilares.

- **Dimensión Fractal:** Número que cuantitativamente describe como un objeto llena su espacio. En un plano geométrico euclidiano, los objetos son sólidos y continuos, es decir, no tienen agujeros o saltos, por lo que sus dimensiones son enteras. Los fractales suelen ser rugosos y discontinuos, por lo que su concepto de dimensión cambia.
- **Distribución Fractal:** Una función de densidad de probabilidad estadísticamente autosimilar. Esto es, en diferentes incrementos de tiempo, sus características estadísticas permanecen iguales.
- **Exponente de Hurst:** A modo de aproximación, se puede considerar como una me-

dida del sesgo en el movimiento Browniano fraccional. Si es igual a 0,5 corresponde al tradicional. En caso de que  $0,5 < H \leq 1$  se presenta una serie en que se observa persistencia, y en el caso de que  $0 \leq H \leq 0,5$ , se tiene un sistema que presenta reversión a la media, o antipersistencia.

- **Filtración:** Teniendo una estructura algebraica  $\mathcal{F}$ , una filtración corresponde a un conjunto indexado  $\mathcal{F}_i$ , que cumple la siguiente condición

$$\forall i, j \in I : i \leq j \implies \mathcal{F}_i \subseteq \mathcal{F}_j \quad (6.1)$$

Es decir, asumiendo que  $i$  es el parámetro de tiempo del proceso estocástico, la filtración se puede interpretar como la representación de toda la información hasta el instante dado, sin incluir información del futuro. Es decir, a medida que  $i$  crece,  $\mathcal{F}$  contiene mas información, sin ser nunca anticipativo.

- **Fractal:** Objeto cuyas partes estan relacionadas, en alguna forma, al objeto completo. Posee dimensión no necesariamente entera y autosimilitud o autoafinidad.
- **Hipótesis de mercado fractal:** Una hipótesis de mercado que establece:
  - Mercado con muchos inversores con distintos horizontes de inversión.
  - La información que es relevante para cada horizonte de inversion es diferente.

Mientras el mercado mantenga esta estructura fractal, sin escala de tiempo característica, el mercado se mantiene estable. Cuando el horizonte de inversión del mercado se vuelve uniforme, el mercado se hace inestable ya que todos están utilizando la misma información al transar, lo cual se traduce en algunos de los problemas estadísticos de las series reales.

- **Innovaciones:** Diferencia entre el valor observado en  $t$  y el valor pronosticado en dicho  $t$  por algún modelo específico. Se suele denominar como  $\varepsilon_t$ .  
Se espera que la serie temporal de innovaciones corresponda a un ruido blanco, es decir, que posea media cero y sus valores no estén correlacionados. Resulta de interés

estudiarla sola ya que entrega información sobre el 'filtrado' o 'blanqueamiento' de una serie por un determinado modelo.

- **Integral de Ito:** Es una generalización de la integral de Riemman para procesos estocásticos, tales como el movimiento Browniano y otras martingalas. Esto resulta útil al considerar que dichos procesos no son diferenciables, por lo que el trabajo de Ito permite aplicar las técnicas clásicas del cálculo a procesos que no cumplen los requisitos. Integrando y integradores son procesos estocásticos

$$Y_t = \int_0^t H_s dX_s \quad (6.2)$$

$H_s$  es un proceso cuadráticamente integrable (integral del cuadrado del valor absoluto finita), adaptado a la filtración generada por  $X$ , que es un movimiento Browniano u otra semimartingala. El resultado de esta integración entrega otro proceso estocástico

- **Invarianza a escala.**

Dada una constante  $c$ , se tiene que:

$$f(x) = ax^k \quad (6.3)$$

$$f(cx) = c^k ax^k \quad (6.4)$$

Si  $\sigma^2$  representa la varianza, y se tiene que  $x = (x_i - \bar{x})$ ,  $k = 2$ ,  $a = 1/n$ :

$$g\left(\sum f(x)\right) = \sigma^2 \quad (6.5)$$

Si existe invarianza se cumplirá que:

$$g\left(\sum f(cx)\right) = c^2 \sigma^2 \quad (6.6)$$

- **Ley de Potencia:** Corresponde a una relacion entre dos cantidades, donde un cambio relativo en una cantidad resulta en un cambio proporcional relativo en la otra, por ejemplo,  $N = Cn^h$ , con  $C$  y  $n$  reales y constantes. Esto es independientemente del

tamaño inicial de dichas cantidades.

Un atributo importante de la Ley de Potencia es su invarianza a escala. Dada una relación  $f(x) = Cx^h$ , escalar el argumento  $x$  por un factor constante  $k$  solo causa un escalamiento proporcional de la función en sí. Es decir

$$f(kx) = C(kx)^h = k^h Cx^h = k^h f(x) \propto f(x) \quad (6.7)$$

para cualquier constante  $k$ . A esto se le llama invarianza a escala, al multiplicar el argumento  $x$  por  $k$ , se multiplica la ley de potencia por la constante  $k^h$ . Esto permite que una función quede determinada por su exponente, permitiendo calcular estadísticas sobre las diferentes escalas de observación y con ello estimar el exponente. Este comportamiento es lo que permite observar una relación lineal cuando se toman en consideración los logaritmos de  $f(x)$  y  $x$ .

Un ejemplo de esta ley se encuentra la estructura autosimilar de los fractales.

- **Ley de los grandes números:** Si bien se habla de una 'Ley', suele usarse este término para referirse a un conjunto de teoremas que describen la media de una sucesión de variables aleatorias cuando el número de ensayos aumenta. Tiene dos versiones. Si se considera una sucesión de variables aleatorias i.i.d de valor esperado  $\mu$  y varianza  $\sigma^2$ .

- Ley débil: Si la sucesión es finita, el promedio  $\bar{X}_n = \sum_{i=1}^n X_i/n$ , para cualquier número positivo  $\varepsilon$ , converge en probabilidad a  $\mu$  cuando  $n$  crece.

$$\lim_{n \rightarrow \infty} \mathbb{P}(|\bar{X}_n - \mu| < \varepsilon) = 1 \quad (6.8)$$

- Ley fuerte: Si la sucesión es infinita, el promedio  $\bar{X}_n = \sum_{i=1}^n X_i/n$ , para cualquier número positivo  $\varepsilon$ , converge en probabilidad a  $\mu$  cuando  $n$  crece.

$$\mathbb{P}\left(\lim_{n \rightarrow \infty} \bar{X}_n = \mu\right) = 1 \quad (6.9)$$

Mediante esta Ley se justifica el uso del promedio del muestreo repetitivo como el valor esperado de una variable aleatoria.

- **Martingala:** Proceso estocástico en que el conocimiento de eventos pasados nunca ayuda a predecir la media de futuros aciertos o reducir su incerteza. La esperanza condicional del valor de la siguiente observación es igual a la última observación. Representa el típico caso de la caminata aleatoria, la cual posee la propiedad de ser Markoviana.
- **Movimiento Browniano fraccional:** fBm, por sus siglas en inglés. Una caminata aleatoria continua con sesgo, comparable a un lanzamiento de dados que están cargados. Al contrario del movimiento browniano tradicional, las probabilidades están inclinadas en una dirección.
- **Proceso estocástico autosimilar:** Dado un proceso  $\{X(t), t \in T\}$  poseerá la propiedad de autosimilaridad si, para cualquier  $a$ , y dado un parámetro  $H$ , se cumple que  $\{X(at) \stackrel{d}{=} a^H \cdot X(t)\}$
- **Proceso Lévy:** Proceso estocástico con incrementos independientes y estacionarios. Es decir, los cambios sucesivos son aleatorios e independientes, además de idénticos sobre diferentes intervalos de tiempo, siempre que éstos sean de la misma longitud. Dentro de ésta categoría de procesos estocásticos se encuentra el movimiento Browniano, y el proceso Poisson. Se asemeja a una caminata aleatoria, pero en tiempo continuo.

Se dice que un proceso estocástico  $Z = \{Z(t)\}_{t \geq 0}$  es un proceso Lévy si se cumple

- Independencia de los incrementos. Es decir, si para  $0 \leq t_1 < t_2 < \dots < t_n < \infty$ , los valores de  $Z(t_2) - Z(t_1), \dots, Z(t_n) - Z(t_{n-1})$  son independientes.
- Estacionariedad de los incrementos. Para cualquier  $s < t$ ,  $Z(t) - Z(s) \stackrel{d}{=} Z(t - s)$ . Esto es, la distribución de probabilidad del incremento  $Z(t) - Z(s)$  solo depende de  $t - s$ , e incrementos del mismo intervalo de tiempo poseen la misma distribución.
- Continuidad en la probabilidad. Si para  $\varepsilon > 0$ , y para  $t \geq 0$ , se cumple que

$$\lim_{h \rightarrow 0} \mathbb{P}(|Z(t+h) - Z(t)| > \varepsilon) = 0 \quad (6.10)$$

- **Semi Martingala:** Proceso que puede ser descompuesto como la suma de un proceso Martingala local y un proceso adaptado de variación finita. Existen varias clases, siendo los mas conocidos el Movimiento Browniano y Procesos Poisson. Este tipo de comportamiento lo presentan los precios de acciones.

Un proceso semi martingala posee componentes que a su vez son semimartingalas.

- **Teorema del Límite Central:** Asumiendo que se tiene una suma de  $n$  variables aleatorias  $Y_n = X_1 + \dots + X_n$ , las que son i.i.d., independientes y con media  $\mu$  y varianza  $\sigma^2$  finitas. La media de  $Y_n$  será  $n \cdot \mu$  y la varianza  $n \cdot \sigma^2$ , dada su independencia, con lo cual se plantea, que cuando  $n \rightarrow \infty$ , la estandarización de  $Y_n$  converge a la Normal estándar, es decir

$$\lim_{n \rightarrow \infty} \mathbb{P} \left( \frac{Y_n - n\mu}{\sigma \sqrt{n}} \leq z \right) = \Phi(z) \quad (6.11)$$

Para el caso sin normalizar, se puede definir que la variable aleatoria

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i \quad (6.12)$$

se aproxima a una normal con  $\mu_{\bar{X}} = \mu$  y  $\sigma_{\bar{X}}^2 = \frac{\sigma^2}{n}$ , para  $n$  suficientemente grande.

Un aspecto muy interesante de este teorema es que solo pide existencia de media y varianza, no una distribución para  $X_i$ .

- **Variables aleatorias independientes e idénticamente distribuidas:** A lo largo de este documento se utilizará numerosamente la sigla i.i.d. para denotar la característica de una colección o secuencia de variables aleatorias con la misma distribución de probabilidad que las otras y que son mutuamente independientes. Este supuesto tiende a simplificar las matemáticas detrás de muchos métodos estadísticos, no siendo necesariamente completamente realista. Si bien implica que todas tienen la misma distribución, ello no implica que las potenciales realizaciones tengan la misma probabilidad de ocurrencia.

Este supuesto posee especial relevancia para la aplicación del teorema del límite central, que establece que la distribución de probabilidad de la suma o promedio de variables i.i.d. con varianza finita se aproxima a una distribución normal.

## Bibliografía

- [1] Rama Cont. Empirical properties of asset returns: stylized facts and statistical issues. 2001.
- [2] Harold Edwin Hurst. Long-term storage capacity of reservoirs. *Trans. Amer. Soc. Civil Eng.*, 116:770–808, 1951.
- [3] Harold Edwin Hurst. A suggested statistical model of some time series which occur in nature. *Nature*, 180:494, 1957.
- [4] J.W. Van Ness B.B. Mandelbrot. Fractional brownian motions, fractional brownian noises and applications. *SIAM Review*, 10:422—437, 1968.
- [5] Zhao Zhongying, Wang Yulin, Liu Guangdi, and Sun Xiang. An rescaled range analysis on the characteristics of coal seam development in the eastern depression of the liaohe basin. *Mining Science and Technology (China)*, 21(2):223–227, 2011.
- [6] Yu Xue, Li-Si Jia, Wei-Zhong Teng, and Wei-Zhen Lu. Long-range correlations in vehicular traffic flow studied in the framework of kerner’s three-phase theory based on rescaled range analysis. *Communications in Nonlinear Science and Numerical Simulation*, 22(1):285–296, 2015.
- [7] M Meraz, E Rodriguez, R Femat, JC Echeverria, and J Alvarez-Ramirez. Statistical persistence of air pollutants (o<sub>3</sub>, so<sub>2</sub>, no<sub>2</sub> and pm<sub>10</sub>) in mexico city. *Physica A: Statistical Mechanics and its Applications*, 427:202–217, 2015.
- [8] Edgar E Peters. *Fractal market analysis: applying chaos theory to investment and economics*, volume 24. John Wiley & Sons, 1994.
- [9] Daniel O Cajueiro and Benjamin M Tabak. Evidence of long range dependence in asian equity markets: the role of liquidity and market restrictions. *Physica A: Statistical Mechanics and its Applications*, 342(3):656–664, 2004.
- [10] John T Barkoulas, Christopher F Baum, and Nickolaos Travlos. Long memory in the greek stock market. *Applied Financial Economics*, 10(2):177–184, 2000.
- [11] Andrew W Lo. Long-term memory in stock market prices. Technical report, National Bureau of Economic Research, 1989.
- [12] Stephane Mallat and Wen Liang Hwang. Singularity detection and processing with wavelets. *IEEE transactions on information theory*, 38(2):617–643, 1992.

- [13] Chung Kang Peng, Sergej V Buldyrev, et al. Long-range correlations in nucleotide sequences. *Nature*, 356(6365):168, 1992.
- [14] C-K Peng, Sergey V Buldyrev, Shlomo Havlin, Michael Simons, H Eugene Stanley, and Ary L Goldberger. Mosaic organization of dna nucleotides. *Physical review E*, 49(2):1685, 1994.
- [15] E Alessio, A Carbone, G Castelli, and V Frappietro. Second-order moving average and scaling of stochastic time series. *The European Physical Journal B-Condensed Matter and Complex Systems*, 27(2):197–200, 2002.
- [16] Gao-Feng Gu, Wei-Xing Zhou, et al. Detrending moving average algorithm for multifractals. *Physical Review E*, 82(1):011136, 2010.
- [17] Jan W Kantelhardt, Stephan A Zschiegner, Eva Koscielny-Bunde, Shlomo Havlin, Armin Bunde, and H Eugene Stanley. Multifractal detrended fluctuation analysis of nonstationary time series. *Physica A: Statistical Mechanics and its Applications*, 316(1):87–114, 2002.
- [18] François G Schmitt, Li Ma, and Thierry Angounou. Multifractal analysis of the dollar–yuan and euro–yuan exchange rates before and after the reform of the peg. *Quantitative Finance*, 11(4):505–513, 2011.
- [19] Dong-Hua Wang, Xiao-Wen Yu, and Yuan-Yuan Suo. Statistical properties of the yuan exchange rate index. *Physica A: Statistical Mechanics and its Applications*, 391(12):3503–3512, 2012.
- [20] Yudong Wang, Yu Wei, and Chongfeng Wu. Analysis of the efficiency and multifractality of gold markets based on multifractal detrended fluctuation analysis. *Physica A: Statistical Mechanics and its Applications*, 390(5):817–827, 2011.
- [21] Darko Stošić, Dusan Stošić, Tatijana Stošić, and H Eugene Stanley. Multifractal analysis of managed and independent float exchange rates. *Physica A: Statistical Mechanics and its Applications*, 428:13–18, 2015.
- [22] Syed Aun R Rizvi, Ginanjar Dewandaru, Obiyathulla I Bacha, and Mansur Masih. An analysis of stock market efficiency: Developed vs islamic stock markets using mf-dfa. *Physica A: Statistical Mechanics and its Applications*, 407:86–99, 2014.
- [23] Luciano Zunino, Alejandra Figliola, Benjamin M Tabak, Darío G Pérez, Mario Garavaglia, and Osvaldo A Rosso. Multifractal structure in latin-american market indices. *Chaos, Solitons & Fractals*, 41(5):2331–2340, 2009.
- [24] Rongbao Gu, Hongtao Chen, and Yudong Wang. Multifractal analysis on international crude oil markets based on the multifractal detrended fluctuation analysis. *Physica A: Statistical Mechanics and its Applications*, 389(14):2805–2815, 2010.
- [25] Duan Wang, Boris Podobnik, Davor Horvatić, and H Eugene Stanley. Quantifying and modeling long-range cross correlations in multiple time series with applications to world stock indices. *Physical Review E*, 83(4):046121, 2011.

- [26] Blake LeBaron, W Brian Arthur, and Richard Palmer. Time series properties of an artificial stock market. *Journal of Economic Dynamics and control*, 23(9):1487–1516, 1999.
- [27] Sergio Arianos and Anna Carbone. Cross-correlation of long-range correlated series. *Journal of Statistical Mechanics: Theory and Experiment*, 2009(03):P03037, 2009.
- [28] Akihiko Utsugi, Kazusumi Ino, and Masaki Oshikawa. Random matrix theory analysis of cross correlations in financial markets. *Physical Review E*, 70(2):026110, 2004.
- [29] Boris Podobnik, Davor Horvatic, Alexander M Petersen, and H Eugene Stanley. Cross-correlations between volume change and price change. *Proceedings of the National Academy of Sciences*, 106(52):22079–22084, 2009.
- [30] Boris Podobnik, Duan Wang, Davor Horvatic, Ivo Grosse, and H Eugene Stanley. Time-lag cross-correlations in collective phenomena. *EPL (Europhysics Letters)*, 90(6):68001, 2010.
- [31] Boris Podobnik, DF Fu, H Eugene Stanley, and P Ch Ivanov. Power-law autocorrelated stochastic processes with long-range cross-correlations. *The European Physical Journal B-Condensed Matter and Complex Systems*, 56(1):47–52, 2007.
- [32] Boris Podobnik and H Eugene Stanley. Detrended cross-correlation analysis: a new method for analyzing two nonstationary time series. *Physical review letters*, 100(8):084102, 2008.
- [33] Zhi-Qiang Jiang, Wei-Xing Zhou, et al. Multifractal detrending moving-average cross-correlation analysis. *Physical Review E*, 84(1):016106, 2011.
- [34] Wei-Xing Zhou et al. Multifractal detrended cross-correlation analysis for two nonstationary signals. *Physical Review E*, 77(6):066211, 2008.
- [35] Liansheng Yang, Yingming Zhu, Yudong Wang, and Yiqi Wang. Multifractal detrended cross-correlations between crude oil market and chinese ten sector stock markets. *Physica A: Statistical Mechanics and its Applications*, 462:255–265, 2016.
- [36] Xincheng Lu, Xinxin Sun, and Jintian Ge. Dynamic relationship between japanese yen exchange rates and market anxiety: A new perspective based on mf-dcca. *Physica A: Statistical Mechanics and its Applications*, 474:144–161, 2017.
- [37] Ling-Yun He and Shu-Peng Chen. Multifractal detrended cross-correlation analysis of agricultural futures markets. *Chaos, Solitons & Fractals*, 44(6):355–361, 2011.
- [38] Xiaoyang Zhuang, Yu Wei, and Feng Ma. Multifractality, efficiency analysis of chinese stock market and its cross-correlation with wti crude oil price. *Physica A: Statistical Mechanics and its Applications*, 430:101–113, 2015.
- [39] Mayukha Pal, P Madhusudana Rao, and P Manimaran. Multifractal detrended cross-correlation analysis on gold, crude oil and foreign exchange rate time series. *Physica A: statistical mechanics and its applications*, 416:452–460, 2014.

- [40] Steven J Davis and John Haltiwanger. Sectoral job creation and destruction responses to oil price changes. *Journal of monetary economics*, 48(3):465–512, 2001.
- [41] Lance Bachmeier. Monetary policy and the transmission of oil shocks. *Journal of Macroeconomics*, 30(4):1738–1755, 2008.
- [42] Bebonchu Atem, Devin Kapper, and Eddery Lam. Do exchange rates respond asymmetrically to shocks in the crude oil market? *Energy Economics*, 49:227–238, 2015.
- [43] Jiaqi Jiang and Rongbao Gu. Asymmetrical long-run dependence between oil price and us dollar exchange rate—based on structural oil shocks. *Physica A: Statistical Mechanics and its Applications*, 456:75–89, 2016.
- [44] Guangxi Cao, Jie Cao, Longbing Xu, and LingYun He. Detrended cross-correlation analysis approach for assessing asymmetric multifractal detrended cross-correlations and their application to the chinese financial market. *Physica A: Statistical Mechanics and its Applications*, 393:460–469, 2014.
- [45] Jose Alvarez-Ramirez, Eduardo Rodriguez, and Juan Carlos Echeverria. A dfa approach for assessing asymmetric correlations. *Physica A: Statistical Mechanics and its Applications*, 388(12):2263–2270, 2009.
- [46] GF Zebende. Dcca cross-correlation coefficient: quantifying level of cross-correlation. *Physica A: Statistical Mechanics and its Applications*, 390(4):614–618, 2011.
- [47] Benoit B Mandelbrot and Roberto Pignoni. *The fractal geometry of nature*, volume 173. WH freeman New York, 1983.
- [48] Felix Hausdorff. Dimension und äußeres maß. *Mathematische Annalen*, 79(1):157–179, 1918.
- [49] AS Besicovitch. On linear sets of points of fractional dimension. *Mathematische Annalen*, 101(1):161–193, 1929.
- [50] Abram S Besicovitch. Sets of fractional dimensions (iv): on rational approximation to real numbers. *Journal of the London Mathematical Society*, 1(2):126–131, 1934.
- [51] Paul S Addison. *Fractals and chaos: an illustrated course*. CRC Press, 1997.
- [52] Tilmann Gneiting and Martin Schlather. Stochastic models that separate fractal dimension and the hurst effect. *SIAM review*, 46(2):269–282, 2004.
- [53] Gennady Samorodnitsky and Murad S Taquq. Stable non-gaussian random processes. *Econometric Theory*, 13:133–142, 1997.
- [54] Stephen J Taylor. *Modelling financial time series*. World Scientific, 2008.
- [55] Duncan AJ Blythe, Vadim V Nikulin, and Klaus-Robert Müller. Robust statistical detection of power-law cross-correlation. *Scientific reports*, 6:27089, 2016.

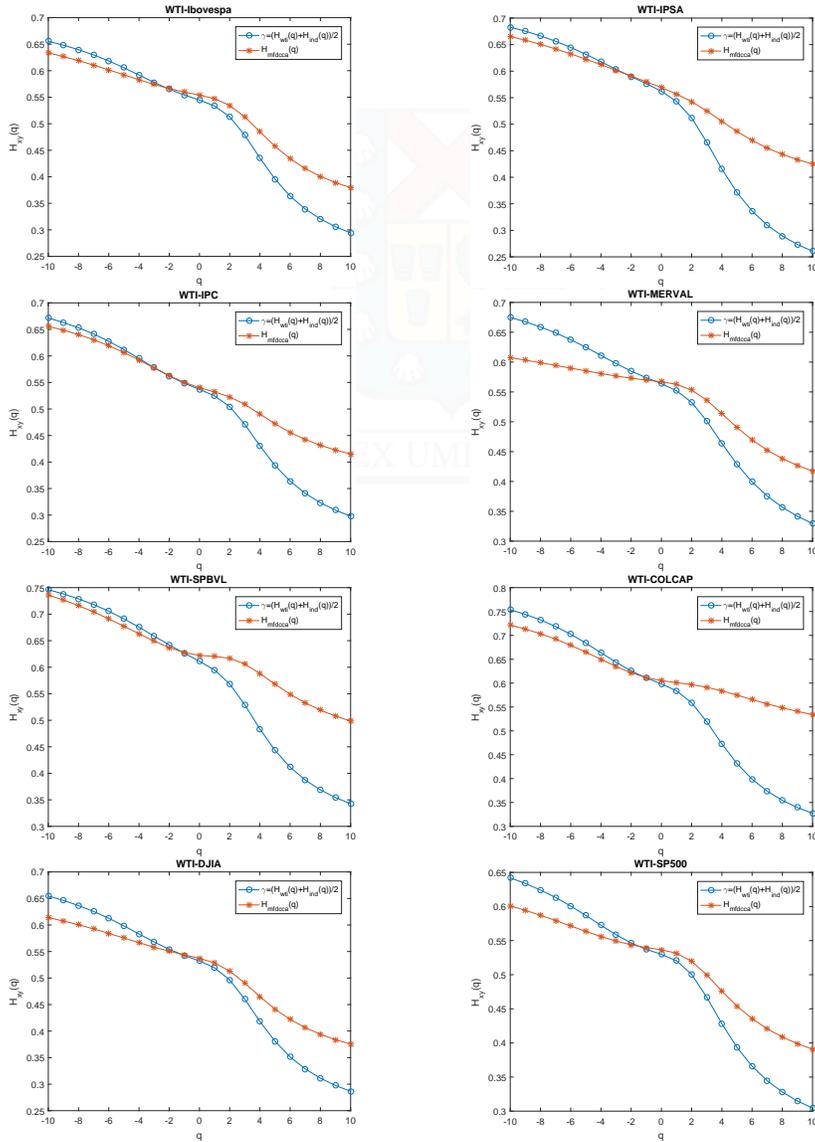
- [56] Rebecca J Sela and Clifford M Hurvich. The averaged periodogram estimator for a power law in coherency. *Journal of Time Series Analysis*, 33(2):340–363, 2012.
- [57] Ladislav Kristoufek. Multifractal height cross-correlation analysis: A new method for analyzing long-range cross-correlations. *EPL (Europhysics Letters)*, 95(6):68001, 2011.
- [58] Ladislav Kristoufek. Detrending moving-average cross-correlation coefficient: Measuring cross-correlations between non-stationary series. *Physica A: Statistical Mechanics and its Applications*, 406:169–175, 2014.
- [59] Louis Bachelier. *Théorie de la spéculation*. Gauthier-Villars, 1900.
- [60] Daniel B Nelson. Conditional heteroskedasticity in asset returns: A new approach. *Econometrica: Journal of the Econometric Society*, pages 347–370, 1991.
- [61] Benoit B Mandelbrot, Adlai J Fisher, and Laurent E Calvet. A multifractal model of asset returns. 1997.
- [62] SK Aggarwal, Michele Lovallo, PK Khan, BK Rastogi, and Luciano Telesca. Multifractal detrended fluctuation analysis of magnitude series of seismicity of kachchh region, western india. *Physica A: Statistical Mechanics and its Applications*, 426:56–62, 2015.
- [63] James Douglas Hamilton. *Time series analysis*, volume 2. Princeton university press Princeton, 1994.
- [64] Argimiro Arratia. *Computational Finance, An Introductory Course with R*, volume 1. Atlantis Publishing Corporation, 2014.
- [65] Carlos M Jarque and Anil K Bera. Efficient tests for normality, homoscedasticity and serial independence of regression residuals. *Economics letters*, 6(3):255–259, 1980.
- [66] Robert F Engle. Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. *Econometrica: Journal of the Econometric Society*, pages 987–1007, 1982.
- [67] Said E Said and David A Dickey. Testing for unit roots in autoregressive-moving average models of unknown order. *Biometrika*, 71(3):599–607, 1984.
- [68] Boris Podobnik, Ivo Grosse, Davor Horvatić, S Ilic, P Ch Ivanov, and H Eugene Stanley. Quantifying cross-correlations using local and global detrending approaches. *The European Physical Journal B-Condensed Matter and Complex Systems*, 71(2):243–250, 2009.
- [69] Tim Bollerslev. Generalized autoregressive conditional heteroskedasticity. *Journal of econometrics*, 31(3):307–327, 1986.
- [70] Hirotogu Akaike, BN Petrov, and F Csaki. Information theory and an extension of the maximum likelihood principle. 1973.

- [71] Gideon Schwarz et al. Estimating the dimension of a model. *The annals of statistics*, 6(2):461–464, 1978.
- [72] Kenneth P Burnham and David R Anderson. *Model selection and multimodel inference: a practical information-theoretic approach*. Springer Science & Business Media, 2003.
- [73] Eric R Kandel, James H Schwartz, Thomas M Jessell, Steven A Siegelbaum, A James Hudspeth, et al. *Principles of neural science*, volume 4. McGraw-hill New York, 2000.
- [74] Raúl Rojas. *Neural networks: a systematic introduction*. Springer Science & Business Media, 2013.
- [75] David E Rumelhart, Geoffrey E Hinton, Ronald J Williams, et al. Learning representations by back-propagating errors. *Cognitive modeling*, 5(3):1, 1988.
- [76] Donald W Marquardt. An algorithm for least-squares estimation of nonlinear parameters. *Journal of the society for Industrial and Applied Mathematics*, 11(2):431–441, 1963.

# ANEXOS



# Relación $H_{xy}(q)$ con $\gamma(q)$



**Figura 1:** Relación  $H_{xy}(q)$  con  $\gamma(q)$ .

(Fuente: Elaboración Propia)

## Códigos de simulaciones

### Simulación de Movimiento Browniano Aritmético

**Código 1:** Código de Matlab para simulación de mBa

```
1 %% Generacion de movimiento Browniano aritmetico
2 alfa=0; sigma=1; T=1; n=2^(12); X0=0;
3 dT=T/n;
4 t=0:dT:1;
5 R = normrnd(0,sigma,[1,n]);
6 X=[X0 alfa*dT+sigma*sqrt(dT).*R];
7 Xt=cumsum(X);
8
9 fil=['mBa'];
10 figure;
11 plot(t,Xt);
12 xlabel('Tiempo');
13 ylabel('X_t');
14 title('Movimiento Browniano Aritmetico');
15 saveas(gcf,fil,'epsc');
```

## Simulación de Movimiento Browniano Fraccional

**Código 2:** Código de Matlab para simulación de mBf

```

1 %% Generacion de movimiento Browniano Fraccional
2 clear;
3 m=0; %Estadísticas de la serie
4 s=0.01;
5 sk=0;
6 ku=3;
7 nestad=[m s sk ku];
8
9 p=1; %Simulaciones
10 q=4096; %Periodos
11 h=0.85;
12 n=1;
13 estad=zeros(p,5);
14
15 for k=1:p
16     clear covar;
17     clear e;
18     clear corra;
19     clear Lq;
20     clear Ln;
21     clear Lp;
22     clear eea;
23     clear Ya;
24
25     covar=zeros(q);
26     e=randn(1,q);
27
28     for i=1:q
29         for j=1:q
30             covar(i,j)=(s^2/2)*(abs(i-j+n)^(2*h)-2*abs(i-j)^(2*h)
                +abs(i-j-n)^(2*h));

```

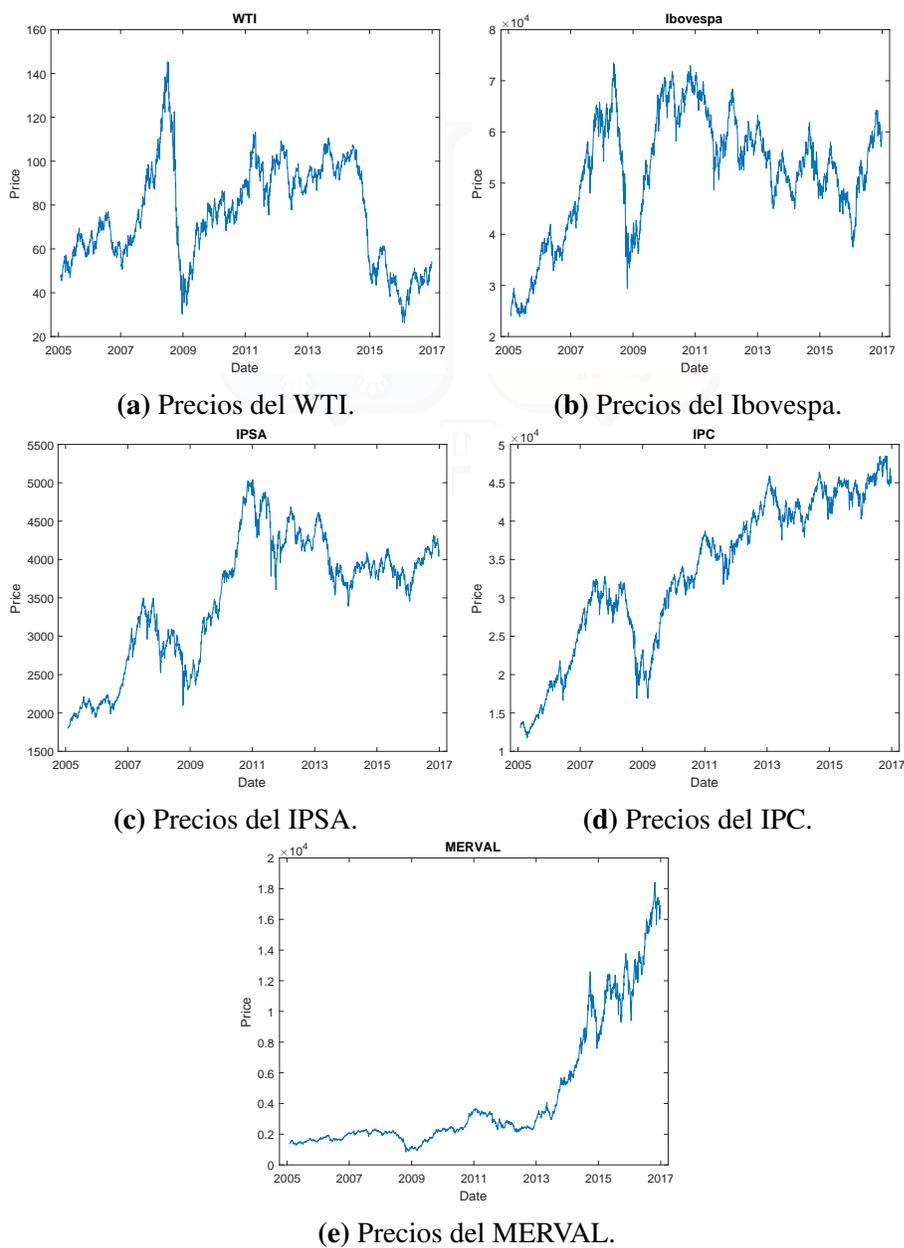
```
31     end
32 end
33
34 corra=cova/s^2;
35 La=chol(corra, 'lower');
36 eea=e*La;
37 Ya=m+s*eea;
38
39 estad(k,1)=floor(k);
40 estad(k,2)=mean(Ya);
41 estad(k,3)=sqrt(var(Ya));
42 estad(k,4)=skewness(Ya);
43 estad(k,5)=kurtosis(Ya);
44
45 meana=mean(estad(:,2:end),1);
46
47 for i=2:q
48     X(k,1)=100;
49     X(k,i)=X(k,i-1)*exp(Ya(i));
50 end
51
52 end
53 figure;
54 fil=['mBf_' num2str(h)];
55 plot(1:q,X);
56 xlabel('Tiempo');
57 ylabel('X_t');
58 title(['Movimiento Browniano Fraccional. H=' num2str(h)]);
59 saveas(gcf,fil,'eps');
```

## Simulación de Modelo GARCH

**Código 3:** Código de Matlab para simulación de GARCH

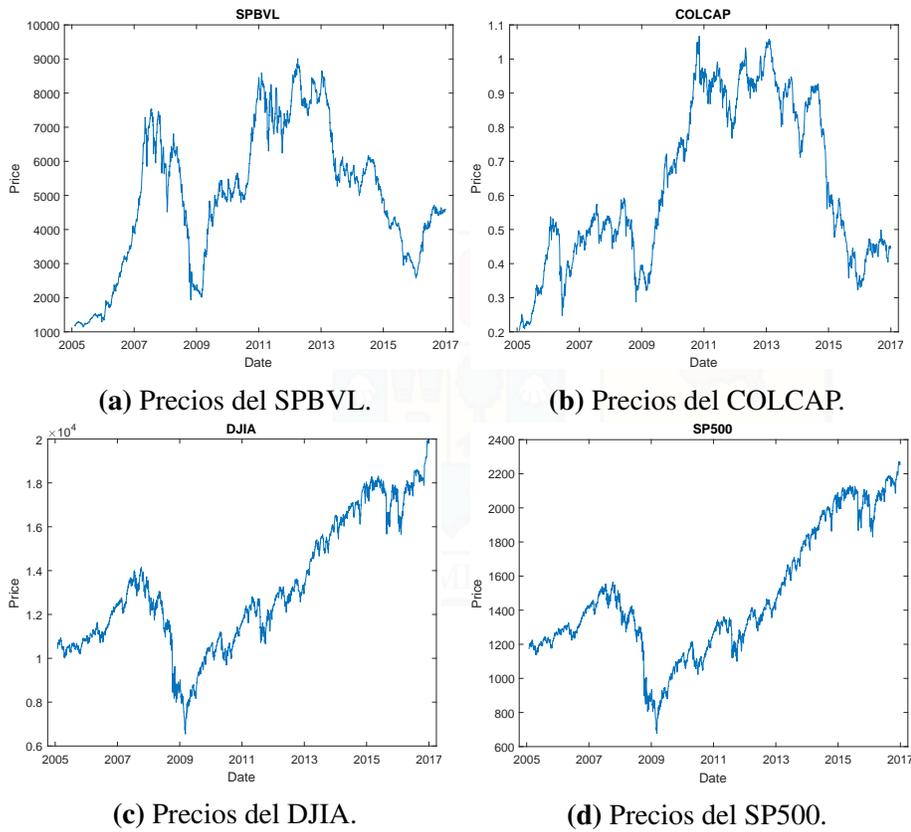
```
1 %% Simulaciones GARCH
2 Mdl = garch('Constant',0.01,'GARCH',0.92,'ARCH',0.05);
3 n=100000; % Cantidad de periodos a simular
4 rng default; % Para reproducir los resultados
5 [Vn,Yn] = simulate(Mdl,n,'NumPaths',1);
6
7 Vn(1,:); % Varianzas
8 fil=['figures/sim_GARCH'];
9 figure;
10 subplot(2,1,1);
11 plot(Vn);
12 xlim([0,n]);
13 title('Varianza condicional');
14 subplot(2,1,2);
15 plot(Yn);
16 xlim([0,n]);
17 title('Innovaciones');
18 saveas(gcf,fil,'eps');
```

## Gráficos de precios



**Figura 2:** Gráficos de precios de las series en análisis.

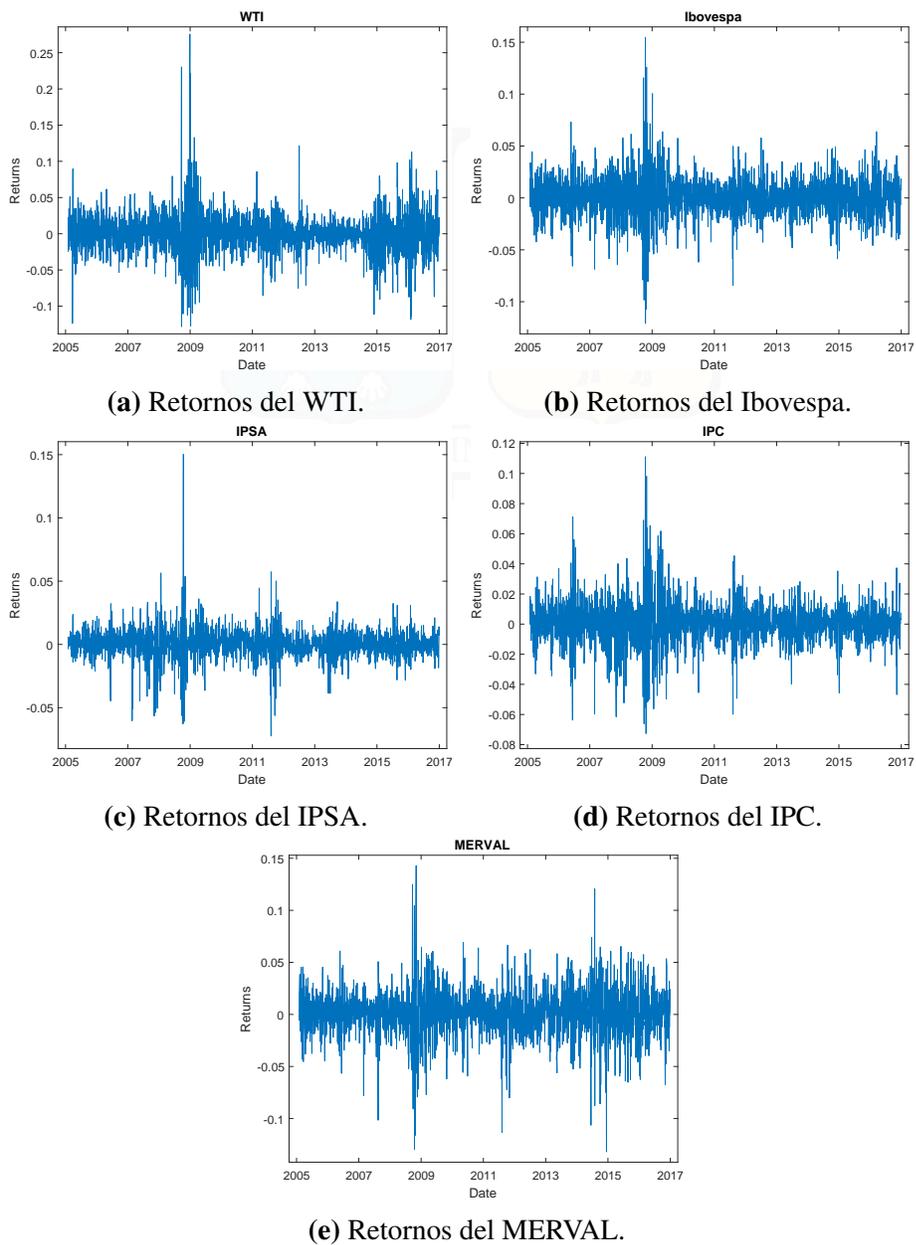
(Fuente: Elaboración Propia)



**Figura 3:** Gráficos de precios de las series en análisis.

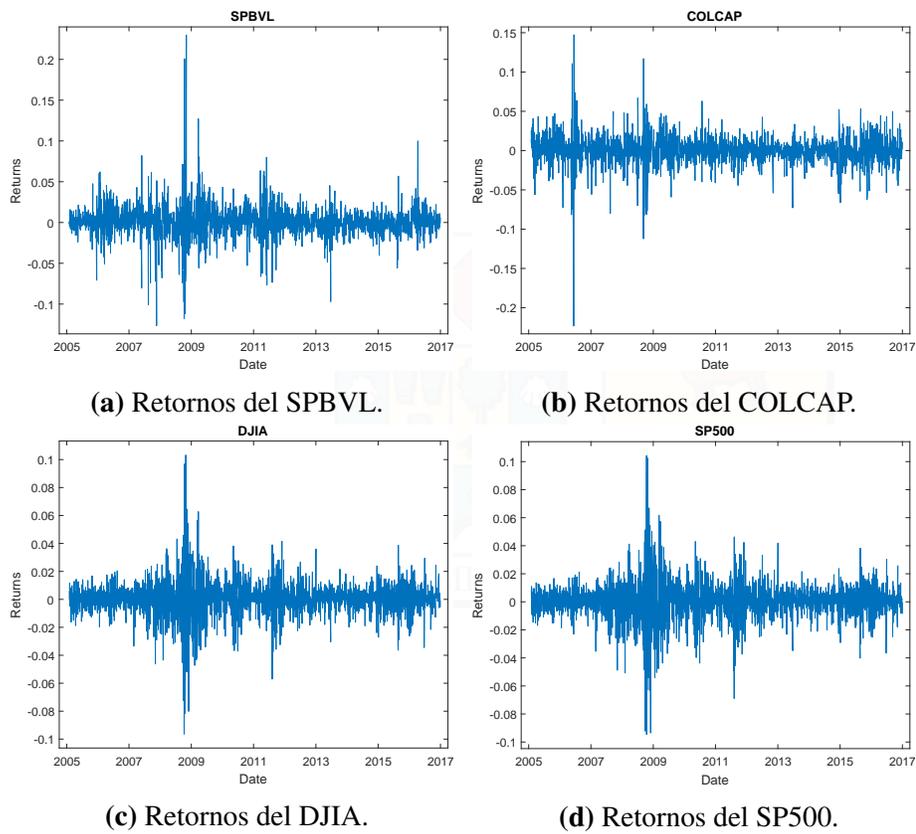
(Fuente: Elaboración Propia)

## Gráficos de retornos



**Figura 4:** Gráficos de retornos de las series en análisis.

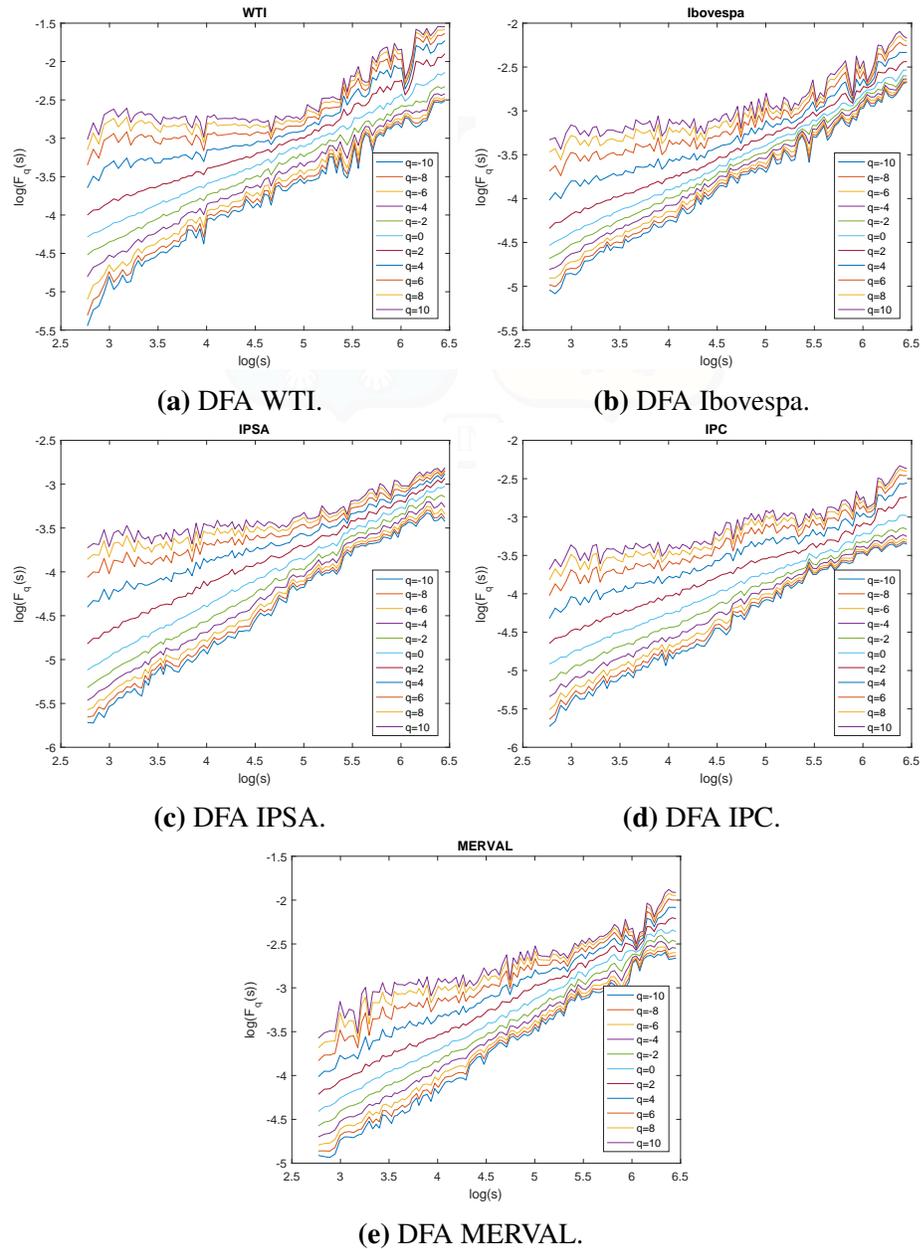
(Fuente: Elaboración Propia)



**Figura 5:** Gráficos de retornos de las series en análisis.

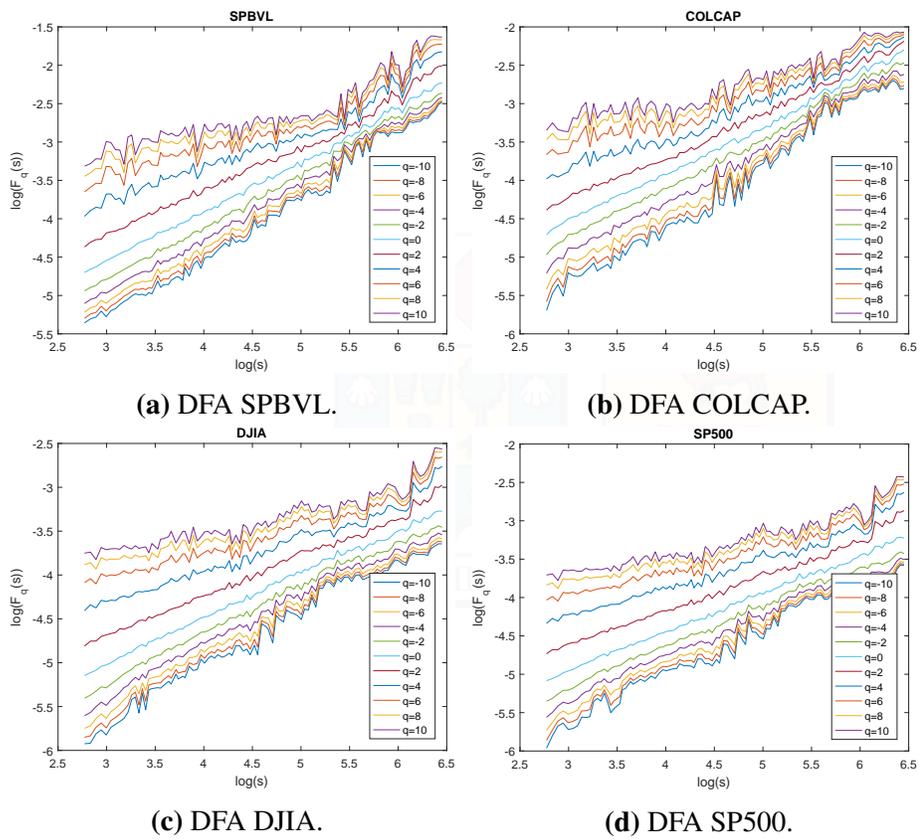
(Fuente: Elaboración Propia)

## Comportamiento de escala MF-DFA



**Figura 6:** Comportamiento de escala para distintos órdenes de fluctuación en el MF-DFA.

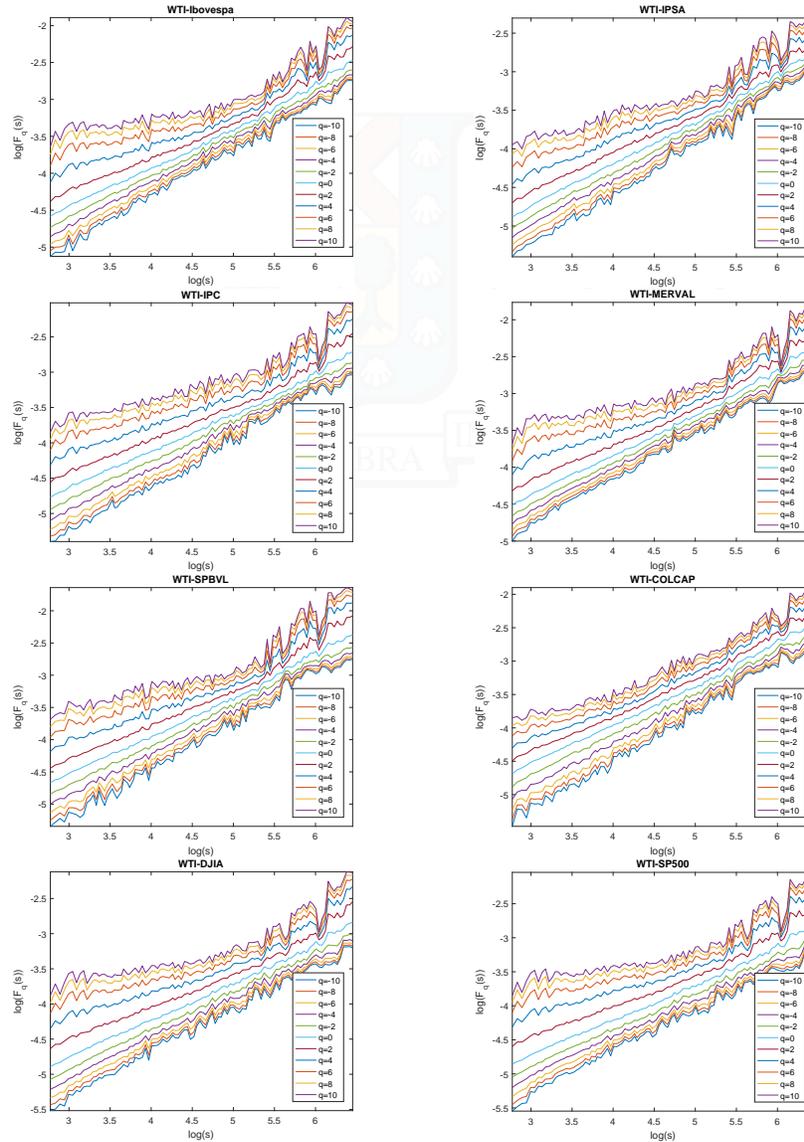
(Fuente: Elaboración Propia)



**Figura 7:** Comportamiento de escala para distintos órdenes de fluctuación en el MF-DFA.

(Fuente: Elaboración Propia)

## Comportamiento de escala según MF-DCCA

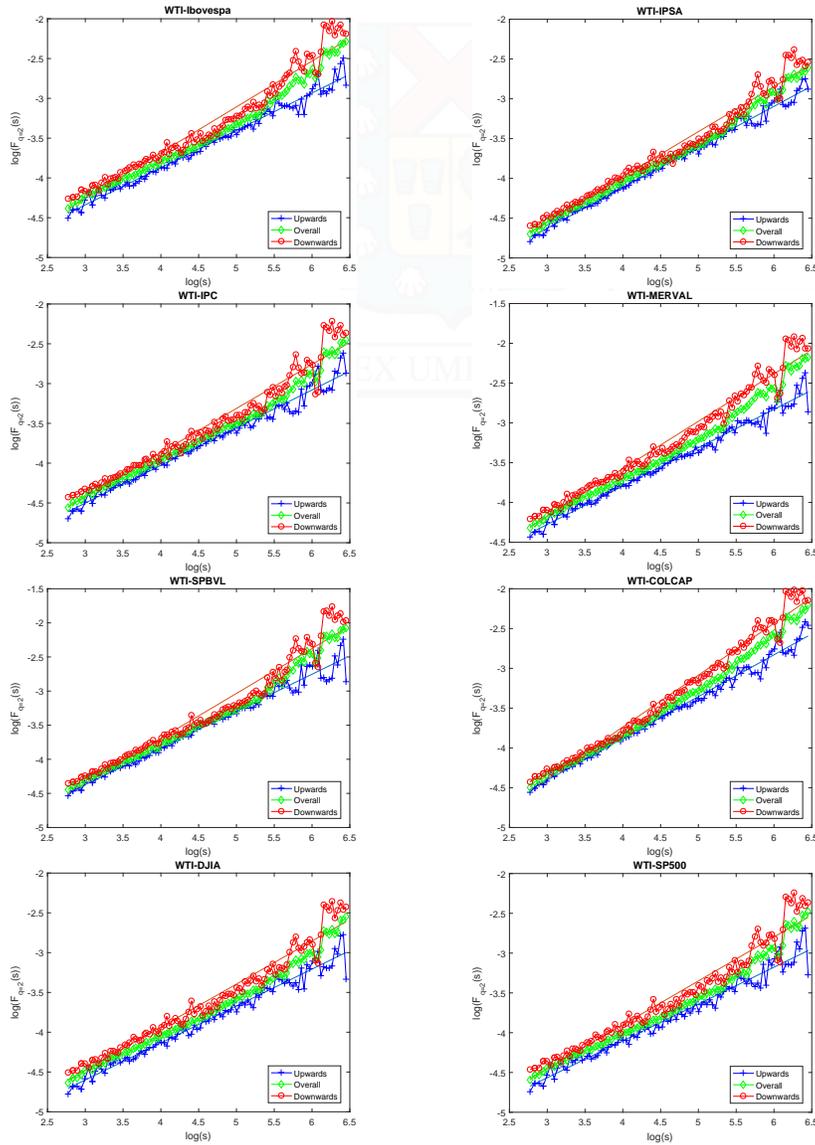


**Figura 8:** Gráficos de las funciones de fluctuación para diferentes órdenes  $q$  en el MF-DCCA.

(Fuente: Elaboración Propia)

# Comportamiento de escala según MF-ADCCA

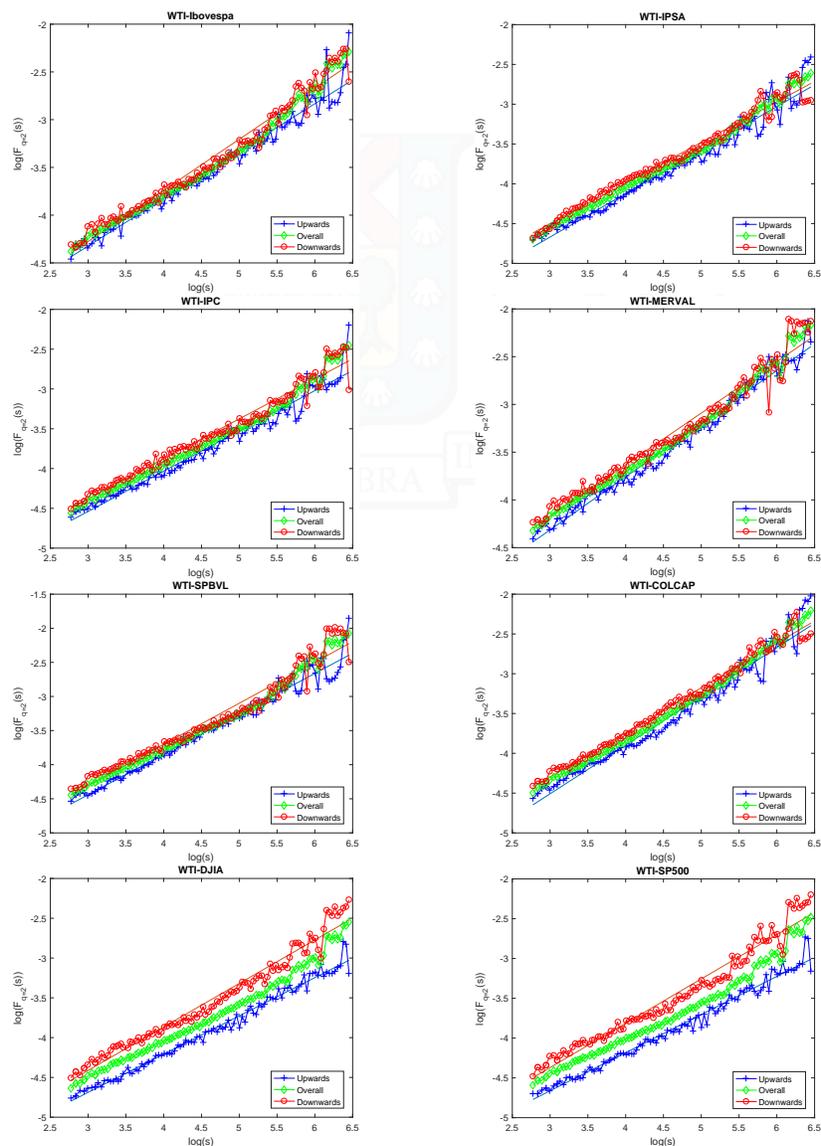
## Según tendencia del WTI



**Figura 9:** Gráficos de las funciones del segundo orden de fluctuación ( $q = 2$ ) de acuerdo a la tendencia del WTI en el MF-ADCCA.

(Fuente: Elaboración Propia)

## Según tendencia del índice bursátil

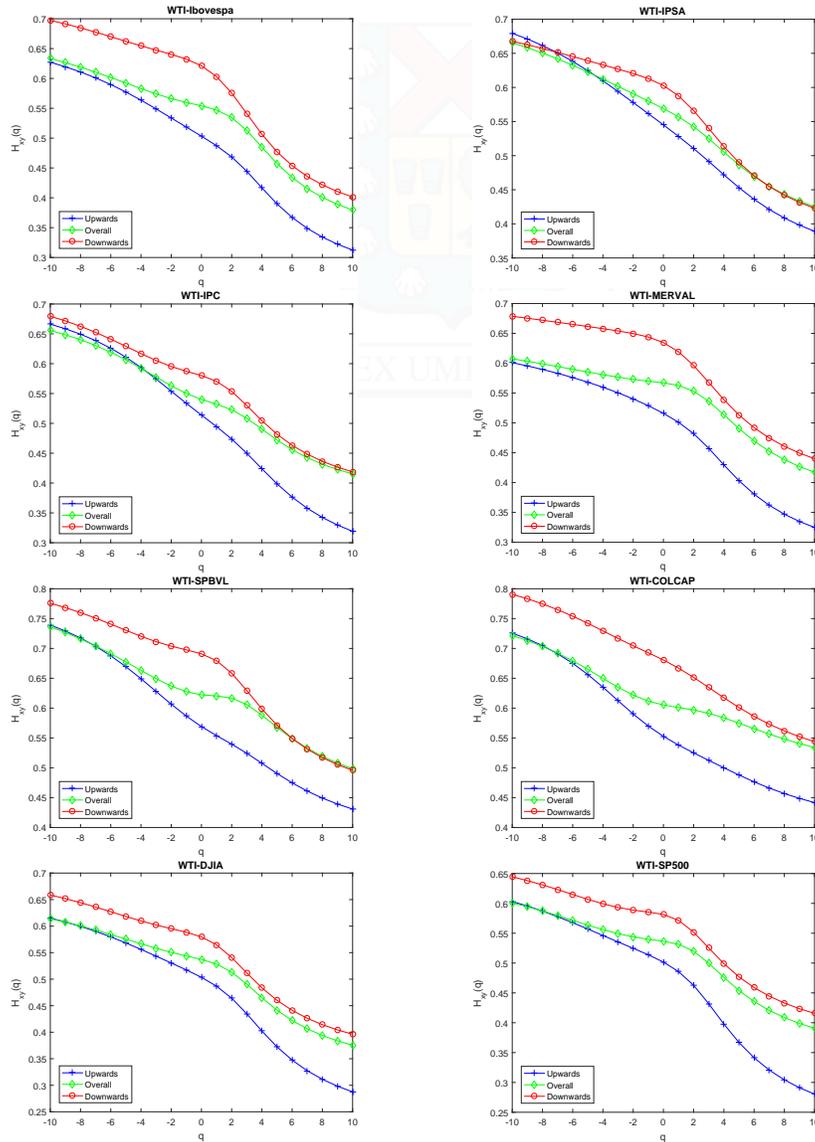


**Figura 10:** Gráficos de las funciones del segundo orden de fluctuación ( $q = 2$ ) de acuerdo a la tendencia del índice bursátil en el MF-ADCCA.

(Fuente: Elaboración Propia)

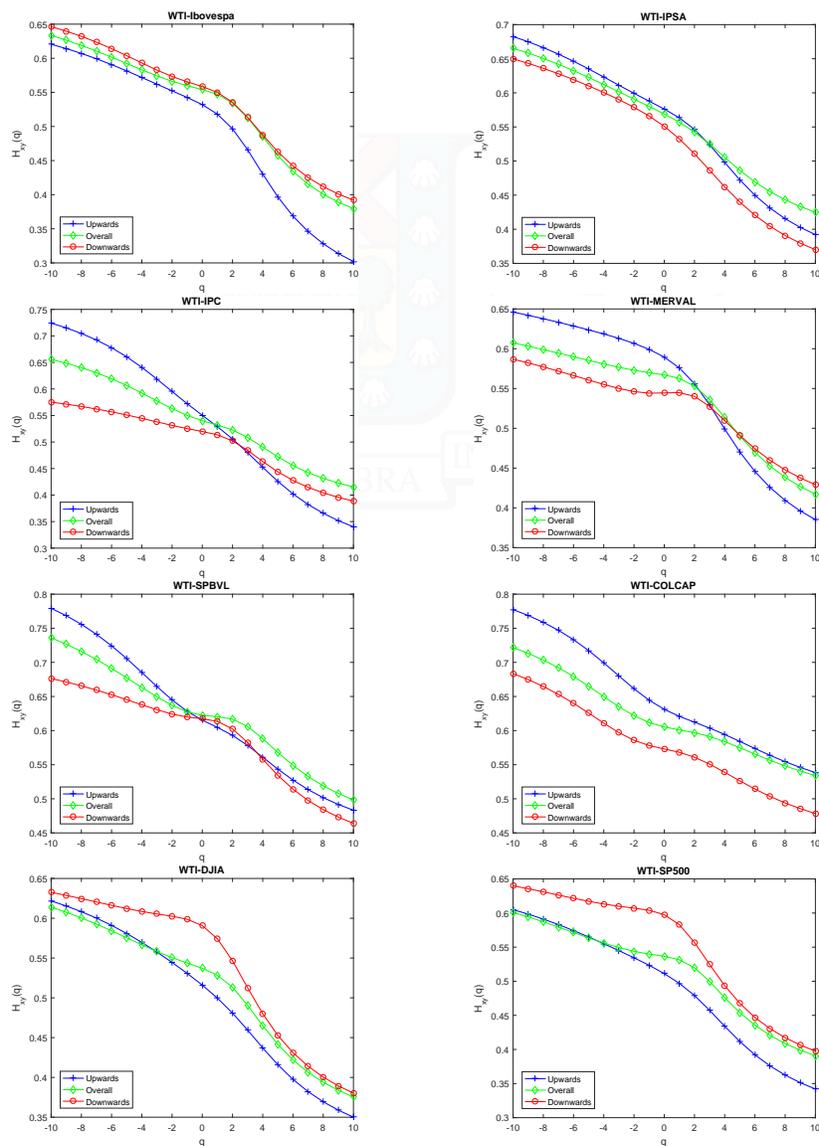
# Comportamiento de correlación según MF-ADCCA

## Según tendencia del WTI



**Figura 11:** Gráficos de los coeficientes de correlación  $H_{xy}(q)$  de acuerdo a la tendencia del WTI.  
(Fuente: Elaboración Propia)

## Según tendencia del índice bursátil



**Figura 12:** Gráficos de los coeficientes de correlación  $H_{xy}(q)$  de acuerdo a la tendencia del índice bursátil.

(Fuente: Elaboración Propia)