

# **Métodos y Técnicas de Detección de Caos en Series Temporales**

TRABAJO FIN DE GRADO

Curso 2019/2020



**UNIVERSIDAD COMPLUTENSE  
MADRID**

FACULTAD DE CIENCIAS MATEMÁTICAS

DOBLE GRADO EN ECONOMÍA – MATEMÁTICAS Y ESTADÍSTICA

Alumno: Iván Rendo Barreiro

Tutora: Lucía Inglada Pérez

Madrid, 17 de julio del 2020

## RESUMEN.

Este trabajo revisa formalmente algunos de los fundamentos de la teoría del caos y los métodos más relevantes e influyentes de la literatura a la hora de detectar procesos caóticos y no linealidades en series temporales, distinguiendo así sistemas deterministas y estocásticos y procesos lineales y no-lineales. Los métodos expuestos son computados en R para aplicarse al caso de una serie bursátil (el histórico de la cotización de IBM en la bolsa neoyorquina). Estos métodos se aplican directamente a los retornos de la serie y a sus modelizaciones ARMA, GARCH y EGARCH. Los resultados obtenidos parecen indicar que la serie estudiada no presenta un componente determinista / caótico y que la modelización más adecuada de la misma es del tipo EGARCH, debido a que su estructura no lineal puede ser recogida a través de la modelización de la varianza.

*This research formally reviews some of the foundations of chaos theory and the most relevant and influential methods in the literature for detecting chaotic and non-linear processes in time series, distinguishing between deterministic and stochastic systems and linear and non-linear processes. The methods presented are computed in R and applied to a stock series (the IBM's historical quotation in the New York Stock Exchange). These methods are applied directly to the returns of the series and to its ARMA, GARCH and EGARCH modelizations. The results obtained seem to indicate that the series studied does not present a deterministic / chaotic component and that the most appropriate modelling of the series is of the EGARCH type, due to the fact that its non-linear structure can be collected through variance modelling.*

Detección de caos, Determinismo, No-linealidad, Series temporales. *Chaos Detection, Determinism, Non-linear, Time Series.*

**Tabla de contenidos**

1. Introducción	7
2. Introducción a las series temporales	10
2.1. No-linealidad y determinismo	10
2.2. Definición de caos	11
2.3. Otros conceptos previos	13
2.4. Algoritmo básico de predicción	16
2.5. Control de sistemas dinámicos	16
3. Métodos de detección de no-linealidad	18
3.1. Test de la dimensión de correlación	18
3.2. Test BDS	19
3.3. Test de White (o NN)	20
3.4. Test de Teräsvirta	21
3.5. Test RESET	22
3.6. Test de Keenan	23
3.7. Test de Tsay	24
4. Métodos de detección de caos	25
4.1. Métodos con muestras subrogadas	25
4.2. Test de Shintani-Linton	26
4.3. Test 0-1	28
4.4. Test de Matilla-García y Ruiz Marín	29
4.5. Algoritmo MESA	31
4.6. El test de los tres estados	31
4.7. La técnica de la valoración del ruido	33
4.8. El árbol de decisión para procesos caóticos	34
4.9. Test NEGM	37
4.10. Gráficos de recurrencia	38
5. Aplicación de los métodos	39
5.1. Datos y metodología	39
5.2. Resultados empíricos	40
6. Conclusión	43
Bibliografía	44
Anexos	49

**Listado de acrónimos**

<b>AADC</b>	Algoritmo del Árbol de Decisión para detección de Caos
<b>AAFT</b>	Transformada de Fourier de Amplitud Ajustada
<b>ARCH</b>	Modelo Autorregresivo Condicional Heterocedástico
<b>ARMA</b>	Modelo Autorregresivo de Media Móvil
<b>BDS</b>	Brock, Dechert y Scheinkman
<b>BIC</b>	Criterio de Información Bayesiano
<b>DT</b>	Desviación Típica
<b>EGARCH</b>	Modelo General Exponencial Autorregresivo Condicional Heterocedástico
<b>FFT</b>	Transformada Rápida de Fourier
<b>FNN</b>	Falso Vecino más Cercano
<b>GAN</b>	Red Generativa Antagónica
<b>GARCH</b>	Modelo General Autorregresivo Condicional Heterocedástico
<b>GPH</b>	Geweke, Porter-Hudak
<b>IBM</b>	<i>International Bussiness Machines</i>
<b>IID</b>	Independientes e Idénticamente Distribuidas
<b>KIC</b>	<i>Kepler Input Catalog</i>
<b>MCO</b>	Mínimos Cuadrados Ordinarios
<b>MESAH</b>	<i>Mean Squared error Histogram</i>
<b>NEGM</b>	Nychka, Ellner, Gallant, MacCaffrey
<b>NN</b>	Red Neuronal
<b>OCDE</b>	Organización para la Cooperación y el Desarrollo Económicos
<b>OGY</b>	Ott, Gerbogi y Yorke
<b>PIB</b>	Producto Interior Bruto
<b>PNB</b>	Producto Nacional Bruto
<b>RESET</b>	Test del Error de la Especificación de la Ecuación de Regresión
<b>RQA</b>	Recurrence Quantification Analysis
<b>R/S</b>	<i>Rescaled Range Analysis</i>
<b>SSR</b>	Suma de los Cuadrados de los Residuos
<b>TSC</b>	Clasificación de Series Temporales
<b>UPO</b>	Unstable Periodic Orbits
<b>3ST</b>	Test de los Tres Estados

## Índice de figuras

<b>Figura 5.1:</b>	Gráfico de la evolución de la cotización de IBM	39
<b>Figura 6.1:</b>	Autocorrelograma de la serie de retornos	55
<b>Figura 6.2:</b>	Función de autocorrelación parcial de la serie de retornos	55
<b>Figura 6.3:</b>	Autocorrelograma de los residuos del modelo ARMA(1,1)	56
<b>Figura 6.4:</b>	Función de autocorrelación parcial de los residuos del modelo ARMA(1,1)	56
<b>Figura 6.5:</b>	Histograma de los retornos	57
<b>Figura 6.6:</b>	Gráfico Qq de los retornos con los cuantiles teóricos de una normal	57
<b>Figura 6.7:</b>	Histograma de los residuos del ARMA(1,1)	58
<b>Figura 6.8:</b>	Gráfico Qq de los residuos del ARMA(1,1) con los cuantiles teóricos de una normal	58
<b>Figura 6.9:</b>	Histograma de los residuos del GARCH(1,1)	59
<b>Figura 6.10:</b>	Gráfico Qq de los residuos del GARCH(1,1) con los cuantiles teóricos de una normal	59
<b>Figura 6.11:</b>	Histograma de los residuos del EGARCH(2,1)	60
<b>Figura 6.12:</b>	Gráfico Qq de los residuos del EGARCH(2,1) con los cuantiles teóricos de una normal	60
<b>Figura 6.13:</b>	Gráfico de recurrencia de los retornos (fNonlinear, m=1, d=1, eps=0,01)	64
<b>Figura 6.14:</b>	Gráfico de recurrencia del modelo ARMA(1,1) (fNonlinear, m=1, d=1, eps=0,01)	64
<b>Figura 6.15:</b>	Gráfico de recurrencia del modelo GARCH(1,1) (fNonlinear, m=1, d=1, eps=0,7)	65
<b>Figura 6.16:</b>	Gráfico de recurrencia del modelo EGARCH(2,1) (fNonlinear, m=1, d=1, eps=0,7)	65
<b>Figura 6.17:</b>	Gráfico de recurrencia de los retornos (fNonlinear, m=2, d=2, eps=0,01)	66
<b>Figura 6.18:</b>	Gráfico de recurrencia del modelo ARMA(1,1) (fNonlinear, m=2, d=2, eps=0,01)	66
<b>Figura 6.19:</b>	Gráfico de recurrencia del modelo GARCH(1,1) (fNonlinear, m=2, d=2, eps=0,7)	67
<b>Figura 6.20:</b>	Gráfico de recurrencia del modelo EGARCH(2,1) (fNonlinear, m=2, d=2, eps=0,7)	67
<b>Figura 6.21:</b>	Gráfico de recurrencia de los retornos (tseriesChaos, m=2, d=1)	68
<b>Figura 6.22:</b>	Gráfico de recurrencia del modelo ARMA(1,1) (tseriesChaos, m=2, d=1)	68
<b>Figura 6.23:</b>	Gráfico de recurrencia del modelo GARCH(1,1) (tseriesChaos, m=2, d=1)	69
<b>Figura 6.24:</b>	Gráfico de recurrencia del modelo EGARCH(2,1) (tseriesChaos, m=2, d=1)	69
<b>Figura 6.25:</b>	Gráfico de recurrencia caótico. Sistema de ecuaciones de Lorenz. (tseriesChaos, m=2, d=1)	70
<b>Figura 6.26:</b>	Gráfico de recurrencia de los retornos (propio, m=1, eps=0,5)	70
<b>Figura 6.27:</b>	Gráfico de recurrencia del ARMA(1,1) (propio, m=1, eps=0,5)	71
<b>Figura 6.28:</b>	Gráfico de recurrencia del GARCH(1,1) (propio, m=1, eps=0,5)	71
<b>Figura 6.29:</b>	Gráfico de recurrencia del EGARCH(2,1) (propio, m=1, eps=0,5)	72
<b>Figura 6.30:</b>	Histograma MESAHA de los retornos	74
<b>Figura 6.31:</b>	Histograma MESAHA del modelo ARMA(1,1)	74
<b>Figura 6.32:</b>	Histograma MESAHA del modelo GARCH(1,1)	75
<b>Figura 6.33:</b>	Histograma MESAHA del modelo EGARCH(2,1)	75
<b>Figura 6.34:</b>	Mapa del espacio de fases de los retornos (d=1, m=2)	76
<b>Figura 6.35:</b>	Mapa del espacio de fases del modelo ARMA(1,1) (d=1, m=2)	76
<b>Figura 6.36:</b>	Mapa del espacio de fases del modelo GARCH(1,1) (d=1, m=2)	77
<b>Figura 6.37:</b>	Mapa del espacio de fases del modelo EGARCH(2,1) (d=1, m=2)	77
<b>Figura 6.38:</b>	Mapa de Poincaré de los retornos (nonlinearTseries)	78
<b>Figura 6.39:</b>	Mapa de Poincaré del modelo ARMA(1,1) (nonlinearTseries)	78
<b>Figura 6.40:</b>	Mapa de Poincaré del modelo GARCH(1,1) (nonlinearTseries)	79
<b>Figura 6.41:</b>	Mapa de Poincaré del modelo EGARCH(2,1) (nonlinearTseries)	79
<b>Figura 6.42:</b>	Convergencia a K en el test de los tres estados para los retornos	80

<b>Figura 6.43:</b>	Convergencia a K en el test de los tres estados para el modelo ARMA(1,1)	81
<b>Figura 6.44:</b>	Convergencia a K en el test de los tres estados para el modelo GARCH(1,1)	81
<b>Figura 6.45:</b>	Convergencia a K en el test de los tres estados para el modelo EGARCH(2,1)	82
<b>Figura 6.46:</b>	Regresión del test de Matilla y Ruiz para los retornos	83
<b>Figura 6.47:</b>	Regresión del test de Matilla y Ruiz para el modelo ARMA(1,1)	83
<b>Figura 6.48:</b>	Regresión del test de Matilla y Ruiz para el modelo GARCH(1,1)	83
<b>Figura 6.49:</b>	Regresión del test de Matilla y Ruiz para el modelo EGARCH(2,1)	84
<b>Figura 6.50:</b>	Exponente de Lyapunov estimados para los retornos (DChaos)	84
<b>Figura 6.51:</b>	Exponente de Lyapunov estimados para el modelo ARMA(1,1) (DChaos)	85
<b>Figura 6.52:</b>	Exponente de Lyapunov estimados para el modelo GARCH(1,1) (DChaos)	85
<b>Figura 6.53:</b>	Exponente de Lyapunov estimados para el modelo EGARCH(2,1) (DChaos)	86

## Índice de tablas

<b>Tabla 1:</b>	Estadísticos y test básicos de las series	40
<b>Tabla 2:</b>	Resultados de los contrastes de estacionariedad	40
<b>Tabla 3:</b>	Resultados de los contrastes de linealidad	42
<b>Tabla 4:</b>	Resultados de los métodos de detección de caos	42
<b>Tabla 5:</b>	Resultados de la estimación ARMA	53
<b>Tabla 6:</b>	Resultados de la estimación GARCH	54
<b>Tabla 7:</b>	Resultados de la estimación EGARCH	54
<b>Tabla 8:</b>	Resultados del contraste Breusch-Godfrey	61
<b>Tabla 9:</b>	Test BDS para los retornos	61
<b>Tabla 10:</b>	Test BDS para el modelo ARMA(1,1)	62
<b>Tabla 11:</b>	Test BDS para el modelo GARCH(1,1)	62
<b>Tabla 12:</b>	Test BDS para el modelo EGARCH(2,1)	63
<b>Tabla 13:</b>	Contrastes de linealidad excepto BDS	63
<b>Tabla 14:</b>	Resultados Test 0-1	72
<b>Tabla 15:</b>	Exponentes de Hurst (pracma)	72
<b>Tabla 16:</b>	Exponentes de Hurst e intervalos de confianza (liftLRD)	73
<b>Tabla 17:</b>	Dimensión de la correlación (nonlinearTseries)	73
<b>Tabla 18:</b>	Resultados de la técnica de valoración del ruido	80
<b>Tabla 19:</b>	Resultados del test de Matilla y Ruiz	82
<b>Tabla 20:</b>	Resultados del árbol de decisión de procesos caóticos	86

## 1. INTRODUCCIÓN

“La predicción en economía y en meteorología tienen mucho que ver. Cuando la gente no está hablando del tiempo, lo hace de la economía.” Estas son las primeras palabras de David Kelsey en su artículo *The economics of chaos or the chaos of economics* [69]. Tras esta primera comparación, el autor explica el desarrollo de los métodos utilizados en la meteorología, ciencia que puede considerarse la vanguardia de la teoría del caos. Debe entenderse que el tiempo atmosférico es la síntesis de ciertos fenómenos atmosféricos que interactúan físicamente, sin aleatoriedad. Mediante la observación parcial de estos fenómenos, se realiza una predicción acerca de ciertas variables que conforman el tiempo meteorológico (precipitación, humedad, nubosidad) y generalmente se resume a través de un dibujo sobre nuestra ciudad. Este proceso es lógicamente determinista, es decir, es perfectamente predecible: dadas unas condiciones iniciales, puede conocerse el estado del sistema cierto tiempo después. Sin embargo, el hombre o mujer del tiempo se equivoca. Esto es debido, en general, a la complejidad del sistema. En particular, se desconocen tanto las propias condiciones iniciales (sería necesario conocer -como mínimo- las condiciones de cada elemento de la troposfera), como la dinámica exacta según la que evoluciona el sistema. Por ello, para su tratamiento, un enfoque puede ser considerar este proceso determinista como un proceso estocástico. Es decir, aceptar nuestro desconocimiento. Por ejemplo, Maxwell, a través de su distribución Maxwell-Boltzmann [90], basa en la estadística su descripción de la cinética de los gases pues es el único acercamiento posible, aun siendo plenamente consciente de su naturaleza determinista.

En este sentido, Poincaré afirma en 1914 que “El azar es solo la medida de nuestra ignorancia. Los fenómenos fortuitos son, por definición, aquellos cuyas leyes ignoramos.”<sup>1</sup> [110]. El propio Poincaré identifica en ciertos sistemas (como el clima o la bolsa) una propiedad fundamental de los procesos caóticos: una ligera variación en las condiciones iniciales conlleva una gran divergencia en su evolución, haciendo imposible la predicción a largo plazo. Volviendo al artículo inicial, David Kelsey identifica que lo concerniente a la meteorología es totalmente atribuible a la economía, y que la nueva (por aquel entonces) teoría del caos podría mejorar la modelización y la predicción en ambas ciencias.

Los sistemas caóticos son, por tanto, impredecibles por naturaleza, pero no aleatorios. Así, la escuela económica austriaca, muchas veces tildada de *antimatemática*, encuentra en la teoría del caos un salvoconducto, un nuevo argumento en su crítica usual, como resume Murray N. Rothbard en el título de un artículo [118]: “La teoría del caos: destruyendo la economía matemática, ¿desde dentro?” En él, el autor establece que, asumiendo a partir de la teoría de caos que la economía es un sistema demasiado complejo y dependiente de las condiciones iniciales, la economía matemática carece de sentido porque “¿Quién, en el mundo, puede predecir la llegada de una mariposa que aletea?”<sup>2</sup> haciendo alusión al famoso *efecto mariposa* que da nombre y es puesto de moda -entre otros- por una película de 2004 protagonizada por Ashton Kutcher. Realmente, el objetivo último del autor es defender el individualismo metodológico, pilar fundamental de la Escuela de Viena desde el *Methodenstreit*<sup>3</sup>; los fenómenos sociales resultan de las motivaciones y acciones de los individuos.

En este punto cabe preguntarse si el hecho de que tanto la economía como la meteorología estén formados por sistemas deterministas, aunque caóticos, es trasladable al resto de ámbitos existentes en la naturaleza. Es esta una pregunta cuya respuesta parece más alcanzable desde la filosofía. Leucipo de Mileto adelanta en el s. V. a. C. [9] que “Nada sucede porque sí, sino que todo sucede con razón y por necesidad” hasta llegar a Laplace, quien responde a la pregunta que le lanza Napoléon sobre el papel de Dios<sup>4</sup> en el universo, “Señor, no he necesitado esa hipótesis” [119]. Pues, Laplace, padre del determinismo científico cree que, dadas todas las condiciones en un cierto instante, pasado y

<sup>1</sup>“Chance is only the measure of our ignorance. Fortuitous phenomena are, by definition, those whose laws we are ignorant of.”

<sup>2</sup>“Who in the world can predict the arrival of a flapping butterfly?”

<sup>3</sup>El *Methodenstreit* (en alemán “disputa de métodos”) fue una controversia económica iniciada en la década de 1880 y que persistió durante más de una década, entre la Escuela Austriaca de ese campo y la Escuela Histórica (alemana). El debate se refería al lugar que ocupaba la teoría general en las ciencias sociales y al uso de la historia para explicar la dinámica de la acción humana. También abordó temas de política, incluyendo los papeles del individuo y del Estado.

<sup>4</sup>Si bien la figura de Dios es recurrente en las diferentes teorías acerca del determinismo en los individuos, normalmente llegando a posturas conciliadoras (e.g. Escrivá de Balaguer argumenta que “Dios quiere correr el riesgo de nuestra libertad” [36]), en el ámbito científico siempre tuvo un papel marginal.

futuro pueden ser determinados por un conjunto completo de leyes. Antes de discutir la existencia del determinismo y sus consecuencias desde el punto de vista del estudio de las series temporales, se expone un fragmento del filósofo Gustavo Bueno [19] que introduce un matiz interesante sobre la finalidad del sistema, y que trata de aunar elementos del estudio de sistemas complejos y de la metafísica.

*“Un «sistema dinámico» determinista es un sistema de-finido (sic) (es decir, determinado según un cierto modo de finalidad); aunque también un sistema «caótico determinista» puede –por su determinismo, más que por su caoticidad– considerarse de-finido siendo ahora los fines los llamados «atractores» (por ejemplo, el «punto fijo») susceptibles de ser dibujados en el espacio de fases del sistema. También para Aristóteles las causas finales se caracterizaban por su capacidad «atractiva» –a diferencia de la capacidad impulsiva de las causas eficientes– (cabría eliminar las connotaciones animistas de la idea aristotélica de fin teológico redefiniendo al Acto Puro como el atractor que se dibuja en el espacio de fases de los astros que se mueven eterna y circularmente).”*

A día de hoy, aceptando que objeto y observación son independientes, se discute si existe aleatoriedad (o, usualmente, *indeterminismo*) en algunos procesos estudiados desde la mecánica cuántica, como el colapso de la función de onda<sup>5</sup>. El filósofo Karl Popper, además de aceptar este argumento, rebate el determinismo científico a través del principio de responsabilidad<sup>6</sup> [112]. El físico-químico Prigogine, por su parte, cree que “*cuanto más sabemos sobre el universo, más difícil es creer en el determinismo*” y que el determinismo pierde su poder explicativo en situaciones donde están presentes la irreversibilidad y la inestabilidad [115].

Para solventar este problema en la sección 2 se define formalmente lo que se entiende por un proceso determinista, estocástico y caótico.

El objetivo del presente trabajo es revisar los resultados más importantes que existen en la literatura acerca del problema de establecer si una serie temporal observada viene generada por un proceso lineal o no-lineal y, en el segundo caso, si es caótico-determinista o, por el contrario, se trata de un proceso estocástico y, así, tratarla adecuadamente con distintos objetivos que se desarrollan después con más detalle: modelización, predicción o control.

El interés del trabajo sigue la línea de expuesta por Inglada-Pérez [60] que, de manera resumida, abarca los siguientes puntos. Primero, consiste en dar una respuesta a si los modelos lineales, frecuentemente utilizados en la economía, describen con precisión el mundo real o si, por el contrario, no están correctamente especificados y un modelo no lineal mejoraría su desempeño, por ejemplo, a través de métodos no paramétricos, tratando de mejorar los resultados predictivos del modelo. Segundo, en el ámbito de la predicción, la existencia del régimen caótico ayuda a la obtención de buenos resultados a través de un algoritmo de predicción no-lineal. Debido a la naturaleza de los procesos caóticos, la predicción es válida en el corto plazo, pues se realiza con base en la estructura ordenada del sistema. Por el contrario, en el largo plazo, no se obtendrán -en general- buenos resultados, debido a que la divergencia que conlleva una mínima desviación en las condiciones iniciales es mayor cuanto más evolucione el sistema. Esto es, al aumentar el tiempo hasta el momento de la predicción, la varianza de esa predicción crece de manera exponencial. Tercero, en lo concerniente al diseño de la política económica, la existencia de una dinámica caótica ayuda al control de cierta variable de referencia, como se explica más adelante. Por último, en el caso concreto de las series del mercado de valores, el análisis de la dinámica caótica -de existir- puede ayudar a la comprensión y predicción de dichas series, siendo esto objeto de gran interés por parte de inversores o economistas.

Tras lo expuesto, y en relación con el problema del determinismo, puede considerarse que, dada una serie temporal, si esta presenta un comportamiento irregular pero dentro de una estructura global ordenada<sup>7</sup>, debe tratarse a través de las herramientas de la ciencia de la complejidad con el fin de aprovechar la información obtenible de ese orden subyacente. Sin embargo, si la dinámica del sistema es tan compleja que no se encuentran siquiera indicios de dicha estructura ordenada, el tratamiento aconsejable será a través de su consideración como proceso estocástico. *Quod*

<sup>5</sup>El propio Einstein, sobre esto, pronuncia su famosa frase “*Dios no juega a los dados con el universo*”. [35]

<sup>6</sup>Dada la tarea de predecir el futuro con cierta precisión, una teoría determinista debería ser capaz de conocer la precisión necesaria para las condiciones iniciales para llevar a cabo esa tarea antes de que se pueda hacer cualquier cálculo.

<sup>7</sup>Este hecho se denomina la “paradoja del caos”.



*natura non dat, Salmantica non præstat*.<sup>8</sup> Esta idea está recogida incluso en la teoría clásica de series temporales, donde suele decirse que los residuos de un modelo (lineal), usualmente denotados por  $\hat{\epsilon}_t$ , recogen *el resto de variables* no introducidas en el modelo (y sus relaciones no-lineales).

Se ha tratado de que los métodos escogidos fueran aquellos más representativos, por su uso actual o pasado o su importancia para el desarrollo de la ciencia, tratando de ser objetivo. Sin embargo, hay un componente subjetivo muy relevante. Sobre este aspecto, el economista Myrdal niega la posibilidad de neutralidad en las ciencias sociales [99] -y es extensible a este caso- al estar siempre sometido a algún tipo de subjetividad aun en la elección del tema de investigación. Por ello, propone -para mantener la objetividad- explicitar las preferencias o juicios de valor que tome el investigador, aunque cree imposible la neutralidad. Revelar o especificar las preferencias propias es complicado, más aún cuando no existe una respuesta general de por qué un método parece más seleccionable que otro. Por ello, me limitaré a aceptar que varios métodos que aquí se presentan *no deberían* aparecer si la decisión se tomara exclusivamente en base a la importancia y tampoco si se tuvieran en cuenta los resultados que de ellos se obtienen, sino que tienen otras características que los hacen interesantes desde mi punto de vista (e.g. simplicidad, novedad, teorías en la que se basan...).

Como es esperable en lo que a series temporales se refiere, la mayoría de las aplicaciones en la práctica tienen como objetivo la predicción. Con esa meta, se han publicado diversos artículos de relevancia recientemente (desde 2019) en distintos ámbitos, como ejemplo: Lahmiri y Bekiros utilizan en *deep learning* y teoría de caos para predecir la evolución del precio de las criptomonedas [78]; Meseguer-Ruiz, Osborn *et al.* definen un índice para la distribución de las precipitaciones en la España peninsular basado en la dimensión fractal [96]; y, en economía, Wen y Yang caracterizan a través del caos la evolución de los sistemas financieros [145]. Además, la investigación teórica también está vigente y pueden incorporarse herramientas de este campo a otros: Naruse *et al.* definen un nuevo tipo de red antagonista<sup>9</sup> basada en series temporales caóticas [100].

El resto de secciones del presente trabajo describirán los siguientes puntos: en la sección 2, se tratan las nociones más básicas que conciernen al estudio de la no-linealidad de las series temporales, así como algunos resultados y conceptos más avanzados que son de interés a la hora entender la motivación y los mecanismos de la detección de caos; en la sección 3, se explican distintos métodos de detección de no linealidad; en la sección 4, se desarrollan algunas técnicas de detección de caos; en la sección 5, se aplican las técnicas descritas en las secciones 3 y 4 al histórico del valor de cotización de IBM<sup>10</sup> en la bolsa neoyorquina; y, para finalizar, en la sección 6, se presenta una breve conclusión. La elección de una serie financiera se debe simplemente a una inclinación personal hacia la ciencia económica. Y, en concreto, se opta por la serie de IBM por tratarse de un valor con un histórico suficientemente extenso en la bolsa de Nueva York y por una cuestión de gustos: puede ser la empresa de estas características que más apueste por el estudio de los sistemas complejos [128, 144]. Sin embargo, cualquier serie temporal es válida para realizar este mismo estudio, aunque los resultados indudablemente difiriesen.

<sup>8</sup>Lo que la naturaleza no da, Salamanca no lo otorga. Con esto se pretende hacer referencia a que la naturaleza de la serie temporal viene dada, y el hecho de que esa estructura global ordenada sea o no *aprovechable* en su modelización depende sobre todo de su complejidad, más que de los métodos utilizados para detectarla o reconstruirla. Como ejemplo, supóngase la economía de una isla donde existe un volcán activo. Lógicamente, el hecho de que el volcán erupcione o no en un determinado momento  $t$  es un proceso indudablemente determinista, pero demasiado complejo y, a día de hoy, impredecible -hasta cierto punto-. Por tanto, en cualquier modelización del estado de dicha economía, todo lo que puede hacerse es tratar la erupción del volcán como una variable aleatoria exógena.

<sup>9</sup>Las famosas GAN, por sus siglas en inglés.

<sup>10</sup>International Business Machines.

## 2. INTRODUCCIÓN A LAS SERIES TEMPORALES

### 2.1 No-linealidad y determinismo

#### 2.1.1 Definición de serie temporal

**Definición 1.** Una **serie temporal** es un conjunto de observaciones de una característica cuantitativa ordenadas sucesivamente en el tiempo.

Algunos autores señalan que los intervalos entre una observación y otra deben ser equidistantes, pero no lo requiere la definición general. Las series temporales, a su vez, pueden venir generadas por un sistema determinista o estocástico, y que este sea lineal o no, como se describe a continuación.

#### 2.1.2 Definición de determinismo y estocasticidad

**Definición 2.** Se dice que un proceso  $\mathcal{X} = \{X_n : n \geq 0\}$  es **determinista** si, dadas ciertas condiciones iniciales, siempre sigue la misma realización. Esto es,  $\forall n, x_n = f(x_0, n)$ , siendo  $x_0$  las condiciones iniciales.

O, lo que es lo mismo, que no existe aleatoriedad en su trayectoria.

**Definición 3.** Un proceso **estocástico**  $\mathcal{X} = \{X_n : n \geq 0\}$  es una sucesión de variables aleatorias definidas sobre un espacio de probabilidad  $(\Omega, \mathcal{A}, P)$  donde, independientemente de sus condiciones iniciales, la realización de la serie es en alguna medida aleatoria [47].

Entre estos procesos están aquellos que generan series temporales, que serán los de nuestro estudio. El índice que ordena dichas series es el tiempo,  $t$ .

#### 2.1.3 Definición de linealidad frente a no-linealidad

Tanto un proceso estocástico como uno determinista pueden ser lineales o no lineales, teniendo especial interés estos últimos en el objeto de nuestro estudio.

**Definición 4.** Decimos que un proceso o un sistema **determinista** es **lineal** cuando el operador  $H$  que lo define en función de una entrada y un tiempo dado  $t$ , satisface el principio de superposición: dados dos inputs  $x_1(t)$  y  $x_2(t)$  y sus respectivos outputs  $y_1(t) = H(x_1(t))$  y  $y_2(t) = H(x_2(t))$  se tiene que:

$$\alpha y_1(t) + \beta y_2(t) = H(\alpha x_1(t) + \beta x_2(t))$$

donde  $\alpha, \beta \in \mathbb{R}$ .

**Ejemplo 5.** En un caso sencillo, la trayectoria descrita por un **movimiento rectilíneo uniforme** es un proceso determinista lineal:

$$x(t) = x_0 + v_0 t$$

donde  $x_0$  y  $v_0$  son las condiciones iniciales y  $t$  la variable temporal. En este caso:

$$\alpha H(x_1(t)) + \beta H(x_2(t)) = \alpha(x_1 + v_1 t) + \beta(x_2 + v_2 t) = (\alpha x_1 + \beta x_2) + (\alpha v_1 + \beta v_2)t = H(\alpha x_1(t) + \beta x_2(t))$$

**Definición 6.** Decimos que un proceso **estocástico** es **lineal** si cada observación  $x_n$  puede expresarse como:

$$x_n = \mu + u_n + \psi_1 u_{n-1} + \psi_2 u_{n-2} + \dots = \mu + u_n + \sum_{i=1}^{\infty} \psi_i u_{n-i}$$

donde  $\mu, \psi_1, \psi_2, \dots$  son parámetros desconocidos, mientras  $u_n, u_{n-1}, u_{n-2}, \dots$  es una secuencia de perturbaciones aleatorias distribuidas idéntica e independientemente con media cero y varianza  $\sigma_u^2$  (ruido blanco).

Definimos así, de manera negativa, un sistema determinista no-lineal como todo aquel proceso determinista que no satisface la definición 4. Análogamente, un sistema estocástico no-lineal es aquel proceso estocástico que no puede escribirse como indica la definición 6.

## 2.2 Definición de caos

Debido a que el objetivo principal del presente trabajo es detectar si un proceso es caótico o no, con los fines últimos mencionados en la introducción, se debe definir qué se entiende por caos. No existe una definición única de caos, por lo que se exponen varias [4, 7, 8, 10, 82, 94, 129, 153]. En general, suele contraponerse a los sistemas estables e inestables -en los que las condiciones iniciales determinan la convergencia o la no convergencia respectivamente- dado que un ligero cambio en las condiciones iniciales de un proceso caótico produce un gran cambio en la trayectoria del proceso. Se dan, por tanto, distintas definiciones de caos y se introducen los conceptos de periodicidad y cuasiperiodicidad, muy vinculados al problema.

### 2.2.1 El exponente (maximal) de Lyapunov y la definición de Caos

Todas las definiciones existentes para calificar a un sistema como caótico requieren que tres componentes estén presentes en el mismo: movimiento oscilante, determinista y sensibilidad a las condiciones. Una de las formalizaciones más extendidas parte del exponente maximal de Lyapunov. Los procesos caóticos son impredecibles (por lo menos, a largo plazo) porque pequeñas desviaciones en las condiciones iniciales (en  $t_0$ ) conllevan serias divergencias en el primer instante ( $t_0 + 1$ ) que, a su vez, suponen una mayor divergencia en el segundo instante ( $t_0 + 2$ ) y así sucesivamente, dando lugar a una divergencia exponencial. De manera contraria, en sistemas periódicos -aún por definir- esta divergencia es lenta.

La media ponderada de estos incrementos en la divergencia en la trayectoria del sistema entre los instantes  $t_0 + i$  y  $t_0 + i + 1$  se denomina *exponente de Lyapunov*. Así, cuanto mayor es, más caos presenta el sistema. Formalmente:

**Definición 7.** Sea  $\delta \mathbf{Z}(t)$  la separación en  $t$  de dos trayectorias del sistema y  $\delta \mathbf{Z}_0$  el vector de separación inicial ( $t = 0$ ) entre las mismas. Queremos formalizar la idea de que el exponente de Lyapunov (maximal),  $\lambda$ , es el ratio de divergencia entre las trayectorias, tal que  $|\delta \mathbf{Z}(t)| \approx e^{\lambda t} |\delta \mathbf{Z}_0|$ . Así, el exponente **maximal de Lyapunov** ( $\lambda$ ) se determina como:

$$\lambda = \lim_{t \rightarrow \infty} \lim_{|\delta \mathbf{Z}_0| \rightarrow 0} \frac{1}{t} \ln \frac{|\delta \mathbf{Z}(t)|}{|\delta \mathbf{Z}_0|}$$

En general, si dicho exponente es positivo las trayectorias divergen a un ratio exponencial y esto será la base para la definición de caos. Además, el término maximal aparece porque dicho valor puede ser distinto para distintas orientaciones del vector inicial  $\mathbf{Z}_0$ . Así, habrá tantos exponentes de Lyapunov como dimensiones tenga el espacio de fases<sup>11</sup>

**Definición 8.** Sea el sistema dinámico  $\dot{x}_i = f_i(x)$  presente en un espacio n-dimensional. Sea la matriz jacobiana  $J_{ij}(t) = \frac{df_i(x)}{dx_j} |_{x(t)}$  y la matriz  $Y$  la determinada por la ecuación  $\dot{Y} = JY$  con condiciones iniciales  $Y_{ij}(0) = \delta_{ij}$ <sup>12</sup>. El **espectro de exponentes de Lyapunov**  $\{\lambda_1, \lambda_2, \dots, \lambda_n\}$  se define como los autovalores de la matriz  $\Lambda$  determinada como:  $\Lambda = \lim_{t \rightarrow \infty} \frac{1}{2t} \log(Y(t)Y'(t))$ .

En general, los exponentes de Lyapunov no se pueden calcular analíticamente. Algunas formas de estimación del exponente maximal de Lyapunov se describen en el trabajo. Otros métodos ampliamente utilizados son los de Kantz [65], Rosenstein *et al.* [117], Kurths y Herzel [74] o Wolf *et al.* [149].

<sup>11</sup>Véase la definición 11.

<sup>12</sup>Deltas de Kronecker.

La definición de caos más usual en la práctica es la siguiente, basada en el exponente maximal de Lyapunov y utilizada en este trabajo.

**Definición 9.** Se dice que un sistema es **caótico** si es determinista, acotado, no-lineal y tiene un exponente maximal de Lyapunov positivo. Además, diremos que es **hipercaótico** si más de uno de sus exponentes de Lyapunov son positivos.

El objetivo del trabajo es distinguir qué procesos son caóticos, si bien los hipercaóticos<sup>13</sup> son -en general- difícilmente diferenciables del ruido. Igualmente, deben distinguirse los procesos caóticos de los periódicos y cuasiperiódicos, que también son deterministas y no-lineales.

**Definición 10.** Decimos que un sistema es **periódico** si es determinista, no-lineal y su exponente maximal de Lyapunov es negativo. Esto es, trayectorias inicialmente cercanas se mantienen próximas.

Será, sin embargo, **cuasiperiódico** si es determinista, no-lineal y con un exponente maximal de Lyapunov negativo y, además, presenta ciclos como en uno periódico, pero cuando revisita los estados en un nuevo ciclo no lo hace idénticamente en los anteriores, sino en unos parecidos.

Finalmente, pueden darse otras definiciones de caos -más teóricas y que sirven para una mayor comprensión del fenómeno- como las recogidas por Inglada-Pérez [60] a partir del trabajo de Gimeno *et al.* [46] que, en general, tratan de expresar una misma idea con ciertos matices. Son las siguientes:

- Barnett y Choi [4]: *medio para producir un origen determinista en un proceso estocástico. Por tanto, la literatura sobre el caos determinista puede ser una explicación teóricamente profunda de los fenómenos estocásticos.*
- Baumol y Benhabib [7]: *es un fenómeno en el que un mecanismo dinámico que es muy simple y determinista proporciona trayectorias a lo largo del tiempo tan complicadas que no detectarían la mayoría de los test de aleatoriedad usuales.*
- Stewart [129]: *comportamiento sin ley gobernado completamente por la ley.*
- Medio [94]: *comportamiento estocástico que ocurre en un sistema determinista.*
- Blank [10]: *la teoría del caos trabaja con procesos deterministas que parecen aleatorios (estocásticos) pero cuya dimensión es finita.*
- Prigoginen [114]: *el caos determinista se encuentra definido por ecuaciones que si bien engendran un comportamiento aleatorio, son tan deterministas como las leyes de Newton.*
- Baumol y Quandt [8]: *comportamiento totalmente determinista que es, al menos en algunos aspectos, indistinguible de un proceso aleatorio o, más bien, de un proceso perturbado por elementos sustancialmente aleatorios.*
- Yang y Brorsen [153]: *procesos deterministas que parecen estocásticos son definidos como caos determinista.*
- Lorenz [82]: *proceso que parece comportarse de acuerdo con el azar, aunque su desarrollo esté determinado por leyes precisas.*

### 2.2.2 Relación entre caos y no-linealidad

Para terminar con estas definiciones, se realizará una aclaración para tratar de comprender la diferencia práctica entre caos y no-linealidad. En general, dada una serie temporal, existen distintos métodos para detectar si viene generada por un proceso no-lineal o si, dentro de los no-lineales, presenta caos. Hay, sin embargo, otras complejidades, como la distinción primera entre un sistema estocástico y uno determinista caótico (similares en su realización) o la presencia de ruido (aleatorio) en sistemas deterministas, que se tratarán en el trabajo. Lo que ha de quedar claro es que un sistema caótico siempre es no-lineal y el contrario no tiene por qué darse (de hecho, hay varios métodos en el trabajo que disciernen, dentro del determinismo no-lineal, entre periódicos, cuasiperiódicos y caóticos), por lo que

<sup>13</sup>Se dice que un sistema es hipercaótico si presenta, al menos, dos exponentes de Lyapunov positivos.

contrastar positivamente que la serie procede de un sistema no-lineal puede ser un buen comienzo para detectar caos, pero no es un test de caos en sí mismo.

### 2.3 Otros conceptos previos

Existen algunos conceptos previos, definiciones, métricas y representaciones que puede ser útil conocer antes de comenzar el estudio de los métodos de detección de caos, debido a que son recurrentes y/o importantes para el mismo o como estudio preliminar.

#### 2.3.1 Otros conceptos

**Definición 11.** Definimos el **espacio de fases** de un sistema como el espacio que representa todos los estados posibles del sistema. Cada punto en el espacio de fases corresponde a un estado en el sistema.

**Definición 12.** Algunos sistemas deterministas presentan **ruido** en su dinámica. Partiendo del concepto de ruido del campo de las señales electrónicas, podemos definir el ruido como el error o la perturbación aleatoria que muestra la realización del sistema determinista.

Este ruido intrínseco puede ser despreciable y el sistema puede ser modelado, en teoría, por ecuaciones deterministas si lo eliminamos, al contrario de lo que ocurriría con la aleatoriedad de un proceso estocástico. Además, podemos diferenciar entre ruido **blanco** y **coloreado**. El primero es aquel cuyas perturbaciones son variables aleatorias, independientes unas de otras e idénticamente distribuidas, de media cero y varianza constante y finita. El ruido coloreado se refiere a aquel proceso estocástico que no mantiene el espectro de potencia.

**Definición 13.** Se define el **espectro de potencia** como

$$S_{xx}(\omega) = \lim_{T \rightarrow \infty} E(|\hat{x}(\omega)|^2)$$

donde  $\hat{x}(\omega) = \frac{1}{\sqrt{T}} \int_0^T x(t) e^{-i\omega t} dt$  es la transformada de Fourier truncada en  $[0, T]$  para una serie  $x_t$  e indica cómo se distribuye la potencia de una señal.

**Definición 14.** Decimos que un **embebimiento** es el proceso mediante el cual un objeto  $X$  es encajado en otro  $Y$  a través de un mapeo inyectivo que *preserva la estructura*  $f: X \hookrightarrow Y$ .

Dependiendo del campo en el que el embebimiento esté definido, *preservar la estructura* puede tener distintos significados. En el caso de las series temporales, debido a que el espacio de fases no se observa, sino que se tiene una secuencia de valores escalares, debemos reconstruir mediante un embebimiento el espacio de fases, convirtiendo las observaciones en vectores de estado. Esto se resuelve mediante el método de los retardos.

Normalmente, la serie temporal es una secuencia de escalares de una cantidad que depende del estado actual del sistema, esto es:

$$s_n = s(\mathbf{x}(n\Delta t)) + \eta_n$$

siendo  $s_n$  la serie observada,  $\mathbf{x}$  el estado en el tiempo  $\{n, n+t, n+2t, \dots\}$ ,  $s(x(n))$  la función que asocia un estado  $x$  en un tiempo  $n$  a su realización y  $\eta_n$  una perturbación aleatoria (ruido *de medición*) y que suponemos que no existe. Una reconstrucción de retardos de dimensión  $m$  está formada por los vectores  $\mathbf{s}_n$ , dados como:

$$\mathbf{s}_n = (s_{n-(m-1)\tau}, s_{n-(m-2)\tau}, \dots, s_{n-\tau}, s_n)$$

donde  $\tau \in \mathbb{N}$  es el retardo temporal. Encontrar la dimensión  $m$  y el retardo  $\tau$  adecuados que preserven la dinámica tras el embebimiento no es sencillo. El teorema de Takens, expuesto más adelante, señala que una condición suficiente es

$m \geq 2d + 1$ , siendo  $d$  la dimensión del atractor. En la práctica se utiliza el método del falso vecino más cercano (FNN<sup>14</sup>) para determinar una correcta dimensión  $m$ .

En cuanto a  $\tau$ , Frazer y Swinney en [42] desarrollaron una medida de información tal que, dada la serie temporal  $x_t$  con  $t = \{1, \dots, n\}$ :

$$I(\tau) = \sum_{x_t, x_{t+\tau}} P(x_t, x_{t+\tau}) \log \frac{P(x_t, x_{t+\tau})}{P(x_t)P(x_{t+\tau})}$$

y se escoge como  $\tau$  para el embebimiento aquel asociado al primer mínimo local que presente  $I(\tau)$ .

**Definición 15.** Decimos que una serie temporal presenta **memoria prolongada** (*long memory*) si existe una fuerte correlación entre dos observaciones muy distanciadas en el tiempo.

El hecho de que una serie presente memoria prolongada mejora su predictibilidad, por lo que queremos comprobar si efectivamente posee dicha característica. Uno de los contrastes más utilizados es el test de Gewek, Porter-Hudak (GPH) de 1987.

En teoría del caos, el objetivo muchas veces es determinar el atractor del sistema: el cuerpo geométrico al que un sistema tiende a evolucionar, sean cuales sean sus condiciones iniciales. Desde luego, dependerá de éstas que la trayectoria se aproxime o no al atractor. El cuerpo geométrico mencionado puede ser tanto un punto, como una curva, como una variedad o un conjunto de estructura fractal. En este último caso se habla de un atractor extraño, presentes en procesos caóticos y que se desarrollan más adelante. Primeramente, debe formalizarse la definición de atractor:

**Definición 16.** Sea un sistema cuya dinámica viene determinada por la función  $f(t, \cdot)$ . Esto es, dado  $a \in \mathbb{R}^n$ ,  $f(0, a) = a$  y, para cierto valor  $t > 0$ ,  $f(t, a)$  es el resultado de la evolución desde el estado  $a$  en  $t$  unidades de tiempo. Así, un **atractor** es un subconjunto  $A$  del espacio de fases caracterizado por las siguientes condiciones:

1.  $A$  es un subconjunto invariante bajo  $f$ . Esto es, para todo  $a \in A$ ,  $\forall t > 0$  se da  $f(t, a) \in A$ , es decir, la evolución desde un punto del atractor se mantiene en el atractor.
2. Existe un entorno de  $A$ , denominado la *cuenca de atracción* de  $A$  y que se denota por  $B(A)$  tal que  $B(A)$  es el conjunto de todos los puntos  $b$  en el espacio de fases que cumplen la siguiente propiedad: para cualquier entorno abierto  $N \subset A$ , existe una constante positiva  $T$  tal que  $f(t, b) \in N$ ,  $\forall t > T$ .
3. No existe ningún subconjunto no vacío de  $A$  que cumpla las propiedades anteriores.

Nótese que la definición precisa de una métrica definida sobre el espacio de fases. En el caso de  $\mathbb{R}^n$  suele tomarse la métrica euclídea. Además, otras definiciones del atractor existen en la literatura. Por ejemplo, Milnor establece varias definiciones [98] en las que algunas el atractor debe tener volumen positivo (con el fin de evitar que un punto pueda considerarse como un atractor) y en otras se relaja el supuesto de que  $B(A)$  es un entorno de  $A$ .

Hay distintos tipos de atractores, pero pueden separarse en clásicos y extraños. Los segundos son los que pueden encontrarse en los procesos caóticos y están caracterizados por presentar estructura fractal<sup>15</sup>. Si bien casi todos los atractores extraños analizados son caóticos, esto no tiene por qué ser así. Un atractor extraño, para ser caótico, precisa exhibir una gran sensibilidad a las condiciones iniciales. Así, cualesquiera dos puntos arbitrariamente cercanos en el atractor conducirán a dos puntos suficientemente lejanos del atractor y, tras otro número determinado de iteraciones, volverán a puntos arbitrariamente cercanos. Por otra parte, se dice que un atractor es extraño no caótico si presenta un sistema con exponentes de Lyapunov no positivos (i.e. no es caótico), pero no es diferenciable a trozos. Un ejemplo de un proceso que presenta un atractor extraño no caótico sería el latido de la estrella KIC 5520878, el primer sistema dinámico extraño no caótico observado en la naturaleza [79].

<sup>14</sup>False Nearest Neighbour. Véase el trabajo de Kennel et al. [70].

<sup>15</sup>La definición de estructura fractal no es sencilla y difiere según distintos autores, incluyendo un pequeño grupo que no cree que una definición deba darse. Sin embargo, la definición más usual es la planteada por Mandelbrot, que describe un conjunto con estructura fractal como *una forma geométrica áspera o fragmentada que puede ser dividida en diferentes partes, cada una de las cuales es una copia (al menos aproximada) del conjunto* [85]. En general, los matices que se dan a esta definición suelen venir por redefinir la similaridad para entender cuándo se puede considerar al fragmento *una copia* del todo.

### 2.3.2 Definiciones de algunas medidas

Una de las métricas más utilizadas en teoría de caos es el *coeficiente de Hurst*. Este coeficiente también sirve como medida de la **memoria prolongada** de una serie.

**Definición 17.** Dada la serie temporal  $x_t$ , sea  $R(n)$  el rango<sup>16</sup> de la serie hasta el momento  $n$  y  $S(n)$  la desviación típica hasta ese instante, el **exponente de Hurst**,  $H$  es tal que:

$$\lim_{n \rightarrow \infty} E \left( \frac{R(n)}{S(n)} \right) = n^H$$

Podemos entender  $H$  como un índice de dependencia y cuantifica la tendencia relativa de una serie de regresar a la media o de agruparse en torno a una dirección, siendo  $H \in [0, 1]$ . La interpretación de los valores de  $H$  se muestran en la tabla a continuación.

$H$	Interpretación
0,5	No existe autocorrelación de largo plazo.
$> 0,5$	Existe autocorrelación positiva en el largo plazo.
$< 0,5$	Existen intercambios en la trayectoria en el sentido de que a un valor alto le suele seguir uno bajo y viceversa.

### 2.3.3 Algunas representaciones

Las representaciones nos ayudan a tener una noción básica del tipo de proceso al que nos enfrentamos, así como del ruido que en él esperamos. La más básica e importante es el **mapa del espacio de fases** que es, simplemente, la representación en el plano del valor de los puntos que forman los vectores  $(x_{t-\tau}, x_t)$  para todo  $t = \tau + 1, \dots, T$ . Takens proporciona las condiciones bajo las cuales un atractor diferenciable<sup>17</sup> puede ser reconstruido a partir de las observaciones y una función genérica [131], a través del denominado Teorema de Takens, que sigue a continuación de manera simplificada.

*Idea del teorema de Takens.* Supóngase que el vector de estado  $x_t$  de dimensión  $d$  evoluciona de acuerdo a una dinámica determinista y continua, pero desconocida. Supóngase también que la variable observable  $y$  es función (diferenciable) de  $x$  y acoplada a todos los componentes de  $x$ . Considérese entonces la variable  $y_t = y(t)$  y el pasado  $y_{t-\tau}, y_{t-2\tau}, \dots, y_{t-k\tau}$  para un periodo  $\tau \in \mathbb{N}$  dado. La idea es que, al aumentar el número de retardos  $k$ , la dinámica es cada vez más predecible, hasta el punto en que, cuando  $k \rightarrow \infty$ , dicha dinámica es determinista. De hecho, la dinámica es efectivamente determinista para una dimensión finita  $k' \leq 2d + 1$  y, además, esa dinámica determinista es completamente equivalente a aquella del espacio de estados original. Concretamente, ambas se relacionan por un difeomorfismo<sup>18</sup>

El mapeo de Poincaré se define para órbitas periódicas a través de la sección de Poincaré como sigue.

<sup>16</sup>Definido como la diferencia entre el mayor valor y el menor valor de la serie.

<sup>17</sup>Resultados posteriores generalizan el teorema a conjuntos de dimensión Minkowski–Bouligand, definida esta como:

$$\dim_{\text{box}}(S) = \lim_{\varepsilon \rightarrow 0} \frac{\ln N(\varepsilon)}{\ln(1/\varepsilon)}$$

siendo  $N(\varepsilon)$  el número de hipercubos de lado  $\varepsilon$  necesarios para recubrir el conjunto.

<sup>18</sup>Dadas dos variedades  $\mathcal{A}$  y  $\mathcal{B}$ , decimos que la aplicación  $f : \mathcal{A} \rightarrow \mathcal{B}$  es un **difeomorfismo** si:

1. Es un homeomorfismo diferenciable.
2. Tiene inversa diferenciable.



**Definición 18.** Sea  $(\mathbb{R}, M, \phi)$  un sistema dinámico global, con  $M$  el espacio de fases y  $\phi$  la función de evolución. Sea  $\gamma$  la órbita periódica que pasa a través de un punto  $p$  y  $S$  una sección transversal y localmente diferenciable de  $\phi$  a través de  $p$ , que llamamos la sección de Poincaré a través de  $p$ .

Dado un entorno  $U \subset S$  de  $p$ , la función  $P : U \rightarrow S$  es un **mapeo de Poincaré** para la órbita  $\gamma$  y la sección  $S$  si:

1.  $P(p) = p$
2.  $P(U)$  es un entorno de  $p$  y  $P : U \rightarrow P(U)$  es un difeomorfismo.
3.  $\forall p' \in U$  la semi-órbita positiva de  $p'$ <sup>19</sup> interseca  $S$  por primera vez en  $P(p)$ .

E.g., un mapa de Poincaré de la órbita de un planeta alrededor del sol solo representa la posición del mismo una vez cada año. Existen otras representaciones que pueden ayudar a la detección de procesos caóticos o a entender los mismos, como los *diagramas de bifurcación*.

## 2.4 Algoritmo básico de predicción

Como se ha mencionado, una de las aplicaciones básicas de la teoría de caos es la mejora de la predicción de las series temporales que lo presenten. En la sección 5 se discute cómo detectar que efectivamente son procesos caóticos. Sin embargo, para acabar de entender el objetivo del problema y la naturaleza de estos procesos deterministas, se expone a continuación uno de los algoritmos más simples de predicción en series caóticas que, debido a sus características particulares, será válido para el corto plazo exclusivamente (debido a la divergencia que se explica a través de los exponentes de Lyapunov).

**Definición 19.** Sea  $x_1, \dots, x_T$  la serie temporal,  $m$  la dimensión del embebimiento y  $\tau$  el valor del retardo, tal que el embebimiento resulta en el conjunto de vectores  $\mathbf{s}_{(m-1)\tau+1}, \dots, \mathbf{s}_T$  de  $\mathbb{R}^m$ . La predicción mediante el **algoritmo simple** para un tiempo  $\Delta t$  después de  $T$ , dado un parámetro  $\varepsilon > 0$ , se define como:

$$\hat{s}_{T+\Delta t} = \frac{1}{|\mathcal{U}_\varepsilon(\mathbf{s}_T)|} \sum_{\mathbf{s}_t \in \mathcal{U}_\varepsilon(\mathbf{s}_T)} s_{t+\Delta t}$$

donde  $|\mathcal{U}_\varepsilon(\mathbf{s}_T)|$  denota el número de elementos en el entorno  $\mathcal{U}_\varepsilon(\mathbf{s}_T)$  de radio  $\varepsilon$  y centrado en  $\mathbf{s}_T$ .

## 2.5 Control de sistemas dinámicos

Otra aplicación de la teoría de caos, sobre todo en economía, es el de la política a través del control del sistema. En general, puede entenderse el control en un sistema dinámico como el proceso que, mediante la modificación de ciertas variables, resulta en un cambio deseado en el comportamiento de dicho sistema, como explica Chaparro [22]. A continuación se expone una introducción a su trabajo acerca de esta cuestión. La propia autora divide la teoría de control en dos: una metodología clásica, que se sirve de métodos como el lugar geométrico de las raíces o la respuesta de frecuencia; y una moderna, que se basa en la descripción del espacio de estados [32, 103].

Formalmente, considérese el siguiente sistema dinámico caracterizado a través de un sistema de ecuaciones diferenciales:

$$\begin{cases} \dot{x}_t = f(x_t, \mu_t, t) \\ y_t = g(x_t, \mu_t, t) \end{cases}$$

siendo  $x \in \mathbb{R}^n$  el vector de estado que determina la evolución del sistema;  $t \in \mathbb{R}$  la variable temporal;  $\mu_t \in \mathbb{R}^p$  el vector de variables exógenas, entre las que se encuentran las variables de control; y, finalmente,  $y \in \mathbb{R}^q$  el vector de salida del sistema. La idea es que, actuando sobre  $\mu$  pueden alcanzarse distintos objetivos en el sistema.

<sup>19</sup>La **semiórbita positiva** a través de  $x$  se define en un intervalo  $I(x) = (t_x^-, t_x^+)$  como

$$\gamma_x^+ = \{\phi(t, x) : t \in (0, t_x^+)\}$$



En concreto, la evolución de un sistema caótico puede ser modificada de dos maneras: realizando cambios en las condiciones iniciales, i.e. actuando sobre la estabilidad dinámica; o variando los valores de los parámetros, actuando así sobre la estabilidad estructural<sup>20</sup>. Los objetivos pueden ser los siguientes [22]:

1. *La estabilización de trayectorias caóticas.* Se trata de estabilizar trayectorias caóticas alrededor de órbitas inestables periódicas  $x_*(t)$  pertenecientes al atractor [11]. Considérese el sistema

$$\dot{x}_t = F(x_t, u_t) \quad (2.1)$$

donde  $u$  es el vector de variables de control de dimensión  $m$  y  $F$  diferenciable [40, 41]. El vector  $u$  representa la variación de los parámetros del sistema  $u_t = p_t - p_0$ . El objetivo, entonces, se define como la estabilización de  $x_{*0}$  con  $u = 0$  y  $F(x_{*0}, 0) = 0$ . Esto es:

$$\lim_{t \rightarrow \infty} x_t = x_*$$

2. *El seguimiento (o tracking) de las trayectorias.* Sea  $x_{*t}$  una órbita periódica del sistema (2.1). De ser inestable, el objetivo de seguimiento se define como

$$\lim_{t \rightarrow \infty} [x_t - x_{*t}] = 0$$

Esto es, hacer converger  $x_t$  a la órbita de  $x_{*t}$ .

3. *Generación de comportamiento caótico.* También denominado *anticontrol*. Dada una trayectoria  $x_{*t}$  estable, se define el objetivo de caotización como

$$\lim_{t \rightarrow \infty} G(x_t) \geq G_*$$

siendo  $G$  una función objetivo dada, usualmente  $G(x_t) = \lambda_1$ , donde  $\lambda_1$  representa el exponente maximal de Lyapunov del sistema.

4. *Sincronización controlada.* El último caso busca provocar un cambio en el estado de varios sistemas de manera coordinada. Por ejemplo, siendo  $x_1$  y  $x_2$  dos estados del sistema, se puede definir un objetivo como

$$\lim_{t \rightarrow \infty} [x_{1t} - x_{2t}] = 0$$

que conlleva la convergencia de la trayectoria  $x_t$  al conjunto  $\{x : x_1 = x_2\}$ .

Con tales fines, existen distintas metodologías de control. E.g., el método OGY<sup>21</sup> [104], o los desarrollados por Georgii *et al.* [57], Pettini [107], Brainman y Goldhirsch [13] o Hübler y Plapp [58]. Un ejemplo, también desarrollado por Chaparro [22], es el siguiente, donde se utiliza el método OGY de control en el modelo de Soliman de inflación y desempleo.

#### *Control en el modelo de Soliman de inflación y desempleo.*

El modelo de Soliman de 1996 [127] que relaciona inflación y desempleo parte del clásico desarrollado por Phillips [108], donde se especifica por primera vez la famosa curva de Phillips y es el primero de los numerosos modelos existentes hoy que relacionan desempleo e inflación.

El modelo genérico es un sistema en tiempo discreto caracterizado por las siguientes ecuaciones:

$$\begin{cases} \pi_t = f(u_t) + a\pi_t^e \\ \pi_{t+1}^e = \pi_t^e + c(\pi_t - \pi_t^e) \\ u_{t+1} = -b(m - \pi_t) + u_t \end{cases} \quad (2.2)$$

<sup>20</sup>El concepto de *estabilidad estructural* refiere a la estabilidad de los puntos de equilibrio, mientras que el de *estabilidad dinámica* a la respuesta del sistema frente a cambios en las condiciones iniciales.

<sup>21</sup>Ott, Gerbogi y Yorke.

donde  $\pi_t$  y  $\pi_t^e$  son las tasas de inflación real y esperada en el periodo  $t$ ,  $u_t$  el nivel de desempleo,  $a \in [0, 1]$  un parámetro que indica el grado en que la expectativa de inflación es incorporada a la inflación real. El parámetro  $c \in [0, 1]$ , por otro lado, representa el grado en que los agentes corrigen sus predicciones a partir de los errores de predicción pasados en cada periodo;  $f$  es la función que incluye el término no-lineal que relaciona  $u_t$  y  $\pi_t$ , con derivada negativa para cualquier nivel de desempleo. Finalmente,  $b > 0$  es la elasticidad del desempleo al crecimiento de la oferta monetaria real:  $m - \pi_t$ , siendo  $m$  la oferta monetaria nominal.

Soliman propone  $f(u_t) = \alpha_1 + \frac{\alpha_2}{u_t} + \frac{\alpha_3}{u_t^2}$  en lo que denomina el modelo empírico y es la que se considera en este ejemplo. Desde luego, la variable de control es la oferta monetaria nominal  $m$ . El objetivo será estabilizar la dinámica en el punto fijo inestable  $(\pi_E^e, u_E) = (16; 0,6960)$ . Sin entrar en el detalle de la obtención del resultado a través del método de OGY, el control óptimo del sistema a través de la oferta monetaria se consigue con la siguiente regla<sup>22</sup>:

$$m = \begin{cases} \bar{m} & |k_1(\pi_t^e - \pi_E^e) + k_2(u_t - u_E)| > \varepsilon \\ \bar{m} + k_1(\pi_t^e - \pi_E^e) + k_2(u_t - u_E) & |k_1(\pi_t^e - \pi_E^e) + k_2(u_t - u_E)| \leq \varepsilon \end{cases} \quad (2.3)$$

con  $\varepsilon > |u - u_0|$ ,  $\varepsilon \ll 1$ ,  $k_1 = -0,0718$  y  $k_2 = 23,1343$ . Nótese que la regla descrita en (2.3) es equivalente a la famosa regla de Taylor [132], aunque el autor exige  $k_i > 0$ ,  $i = 1, 2$ .

### 3. MÉTODOS DE DETECCIÓN DE NO-LINEALIDAD

#### 3.1 Test de la dimensión de correlación

En 1983, Grassberger y Procaccia [54] desarrollaron este test para la detección de caos.

Sea  $x_t$  con  $1 \leq t \leq T$  la serie temporal, a la que puede aplicarse un embebimiento  $m$ -dimensional, tal que resulten una serie de vectores  $X_t = (x_t, x_{t-1}, \dots, x_{t-m+1})'$ , con  $m \leq t \leq T$ . Asumiendo cierto que el sistema que genera  $x_t$  es  $\vartheta$ -dimensional y dado  $m \geq 2\vartheta + 1$ , entonces las  $N$   $m$ -historias pueden utilizarse para recrear la dinámica de la generación de la serie a la manera de Takens [131], con  $N = n - m + 1$ .

**Definición 20.** Sea  $x_t$  la serie temporal dada,  $t = 1, 2, \dots, T$  y su  $m$ -historia el vector  $x_t^m = (x_t, x_{t-1}, \dots, x_{t-m+1}) \in \mathbb{R}^m$ . La **integral de correlación** para una *dimensión de embebimiento*  $m$  se define como:

$$C_{T,m,\varepsilon} = \frac{2}{T_m(T_m - 1)} \sum_{m \leq s \leq t \leq T} I(x_t^m, x_s^m; \varepsilon) \quad (3.1)$$

donde  $T_m = T - m + 1$  y  $I(x_t^m, x_s^m; \varepsilon)$  es la función indicatriz que toma el valor 1 si  $|x_{t-i} - x_{s-i}| < \varepsilon$  para cada  $i = 0, 1, \dots, m - 1$  y, de otra manera, vale 0.

La función o integral de correlación es una medida de la dimensionalidad del espacio ocupado por la serie.

**Definición 21.** Sea  $C_{T,m,\varepsilon}$  la integral de correlación definida anteriormente, se define la **dimensión de correlación** como:

$$D_c^m = \lim_{\varepsilon \rightarrow 0} \frac{\log C_{T,m,\varepsilon}}{\log \varepsilon}$$

id est, la pendiente de la regresión del logaritmo de la integral de correlación frente a  $\log \varepsilon$  para valores pequeños de  $\varepsilon$ .

Así,  $D_c^m$  siempre dependerá de  $m$ . Si  $m$  incrementa su valor y la dimensión de correlación también lo hace, el sistema es estocástico. Sin embargo, si  $D_c^m$  alcanza un límite donde se satura para cierta  $m$  relativamente pequeña, el proceso es determinista. Por ello, la dimensión de correlación es una herramienta útil en la detección de caos, pero

<sup>22</sup>Para valores del sistema  $a = b = c = 0,1$ ,  $\alpha_1 = -1,14$ ,  $\alpha_2 = 5,53$ ,  $\alpha_3 = 3,68$ ; y  $m$  alrededor de  $\bar{m} = 16$ , donde el sistema presenta un atractor caótico.

presenta el problema de que sus propiedades muestrales no son conocidas y ello lleva a Brock *et al.* [17] a desarrollar un estadístico basado en este concepto del que sí conozcamos su distribución asintótica para poder realizar el contraste de hipótesis.

A partir de la noción de la dimensión de correlación, se han desarrollado numerosos tests y, además, se ha empleado con distintos fines a la hora de detectar comportamientos caóticos. Una aplicación del test propuesto por Grassberger y Procaccia sería el trabajo de Andrzejak *et al.*, donde se comparan distintas series temporales provenientes de encefalogramas para estudiar la posible dependencia que presenta la actividad eléctrica del cerebro en el estado del cerebro y la región de registro [1].

### 3.2 Test BDS

El nombre del Test BDS de 1987, proviene de sus creadores: Brock, Dechert y Scheinkman, y es publicado años más tarde (1996) por los mencionados académicos junto a LeBaron [17]. Se puede decir con certeza que ha sido el test más popular para detectar no-linealidades hasta la fecha.

En un primer momento, el test estaba diseñado para contrastar la hipótesis nula de independencia e idéntica distribución con el objetivo de detectar dinámicas caóticas. Sin embargo, se ha probado en distintos estudios por Brock, Hsieh y LeBaron [16] y Barnett, Gallant *et al.* [5] que tiene potencia contra un amplio rango de alternativas lineales y no lineales, que aquí no especificaremos.

Por último, puede ser utilizado como un test para comprobar si un modelo lineal está bien especificado, contrastando si los residuos del modelo presentan alguna dependencia no lineal, en cuyo caso dicha dependencia estaría omitida en la especificación.

Si las variables aleatorias  $x_t$  son iid<sup>23</sup>, teniendo en cuenta la definición de la integral de correlación proporcionada anteriormente, esta debe ser igual a la probabilidad conjunta de que dos puntos estén a menos de un  $\varepsilon$  de distancia:

$$C_{1,\varepsilon}^m = P(|x_t - x_s| < \varepsilon)^m$$

Así, el estadístico *BDS* se define como sigue.

**Definición 22.** Dada una serie temporal  $x_t$  con  $t = 1, 2, \dots, m$ , el **estadístico BDS** se define como:

$$V_{m,\varepsilon} = \sqrt{T} \frac{C_{m,\varepsilon} - C_{1,\varepsilon}^m}{s_{m,\varepsilon}} \quad (3.2)$$

donde  $s_{m,\varepsilon}$  es la desviación típica de  $\sqrt{T}(C_{m,\varepsilon} - C_{1,\varepsilon}^m)$ .

Bajo ciertas condiciones de regularidad<sup>24</sup> el estadístico BDS converge en distribución a una normal  $N(0, 1)$ , lo que utilizaremos para realizar un contraste bilateral de la manera usual. Brock *et al.* (1991) aconseja tomar el parámetro  $\varepsilon$  entre  $\frac{1}{3}$  y  $\frac{1}{2}$  de la desviación típica de la serie temporal  $x_t$ . Por otra parte, la dimensión del embebimiento ( $m$ ) debe tomarse en relación al número de observaciones (en su ejemplo,  $m \leq 5$  para  $T \leq 200$ ).

Numerosos artículos emplean el test BDS para detectar no linealidades. Uno de los más relevantes es el de Chen *et al.* [25] donde se construye un mercado financiero artificial estocástico. Los resultados muestran que el test BDS es muy inestable, pudiendo tanto aceptar como rechazar la hipótesis de linealidad dependiendo de la realización final del mercado (construido aleatoriamente).

<sup>23</sup>Independientes e idénticamente distribuidas.

<sup>24</sup>Véase el trabajo de Brock, Dechert, Scheinkman y LeBaron [17].

### 3.3 Test de White (o NN<sup>25</sup>)

El test de White, denominado así por Halbert White, y desarrollado en 1989 [147], es el test más importante de los basados en redes neuronales y es utilizado, por ejemplo, por Jiménez-Rodríguez para contrastar que existe una relación no-lineal entre las perturbaciones producidas en el precio del petróleo y el crecimiento del PIB<sup>26</sup> real [64].

**Definición 23.** Una **red neuronal** de una capa oculta prealimentada se define como un modelo en el que las unidades de entrada envían un vector  $\tilde{w}$  de señales  $\tilde{w}_i, i = 1, \dots, k$  a través de conexiones que atenúan o amplifican la señal original por un factor  $\gamma_{ij}$  que denominamos *peso*. La capa intermedia (oculta) de la red está formada por diferentes ( $q \in \mathbb{N}$ ) unidades. La unidad  $j$  de dicha capa recibe las señales  $w_i \gamma_{ij}, i = 1, \dots, k$  y las procesa. En general, las señales recibidas son resumidas por las unidades intermedias en un output producido a través de una función de activación  $\phi$ <sup>27</sup> y  $w = (1, \tilde{w}_1, \dots, \tilde{w}_k)$ , siendo dicho output:

$$f(w, \delta) = \beta_0 + \sum_{j=1}^q \beta_j \phi(\gamma_j' \tilde{w}_j)$$

En particular, el test NN de White emplea una de estas redes, siendo su output:

$$x_t = \pi' \tilde{w}_t + \sum_{j=1}^q \beta_j \phi(\gamma_j' \tilde{w}_t)$$

y la hipótesis nula de linealidad es equivalente a que todos los pesos óptimos sean iguales a 0. Esto es,  $H_0 : \beta_j^* = 0 \quad \forall j$ , dados  $q$  y  $\gamma_j$ .

Este contraste puede ser reemplazado en la práctica por uno basado en los multiplicadores de Lagrange, esto es:

$$\begin{cases} H_0 : E(\phi_t e_t^*) = 0 \\ H_1 : E(\phi_t e_t^*) \neq 0 \end{cases}$$

donde  $\phi_t = (\phi(\Gamma_1' \tilde{w}_t), \dots, \phi(\Gamma_q' \tilde{w}_t))$  y  $\Gamma \equiv (\Gamma_1, \dots, \Gamma_q)$  son escogidos a priori, independientemente de  $w_t$  y para un  $q$  dado. El elemento  $e_t$  se reemplaza por los residuos MCO<sup>28</sup>  $e_t = x_t - \pi' w_t$  para obtener el estadístico.

**Definición 24.** El estadístico de White (NN) para contrastar linealidad se define como:

$$M_n = \left( n^{-1/2} \sum_{t=1}^n \phi_t e_t \right)' \hat{V}_n^{-1} \left( n^{-1/2} \sum_{t=1}^n \phi_t e_t \right)$$

donde  $\hat{V}$  es un estimador consistente de  $V^* = \text{Var}(n^{-1/2} \sum_{t=1}^n \phi_t e_t^*)$ .

**Proposición 25.** Bajo  $H_0$ , el estadístico de White converge en distribución a una  $\chi^2$  cuadrado de  $q$  grados de libertad. Id est:  $M_n \xrightarrow{d} \chi_q^2$ .

Para eliminar ciertos problemas computacionales y de multicolinealidad, se adoptan dos soluciones prácticas:

1. El test se lleva a cabo para los  $q^* < q$  componentes principales de  $\phi_t, \phi_t e_t^*$ .
2. Para evitar el cálculo de  $\hat{V}_n$ , utilizamos el estadístico equivalente  $nR^2 \xrightarrow{d} \chi_q^2$ , donde  $R^2$  es el coeficiente de determinación de la regresión lineal estándar de  $\hat{e}_t$  sobre  $\phi_t^*, \tilde{w}_t$ .

<sup>25</sup>De *Neural Network*, “red neuronal” en inglés.

<sup>26</sup>Producto Interior Bruto.

<sup>27</sup>Típicamente la función logística.

<sup>28</sup>Mínimos Cuadrados Ordinarios (o lineales).

### 3.4 Test de Teräsvirta

Este test debe su nombre a uno de sus autores, T. Teräsvirta, que lo desarrolló [134] a partir del anterior de White de 1989 [147]. La hipótesis nula del test de Teräsvirta es la existencia de linealidad en media frente la alternativa en la que no. Como se ve a continuación, el estadístico se consigue a través de una expansión en serie de Taylor.

Así, considérese el siguiente modelo no lineal:

$$x_t = \phi(\gamma' w_t) + \pi' w_t + u_t \quad (3.3)$$

donde  $u_t \sim \text{nid}(0, \sigma^2)$ ,  $w_t = (1, \tilde{w}_t')'$ ,  $\tilde{w}_t = (x_{t-1}, \dots, x_{t-p})'$ ,  $\pi = (\pi_0, \pi_1, \dots, \pi_p)'$  y  $\gamma = (\gamma_0, \tilde{\gamma}')'$ , con  $\tilde{\gamma} = (\gamma_1, \dots, \gamma_p)'$ . Fijemos también

$$\phi(\gamma' w_t) = \theta_0 \psi(\gamma' w_t) \quad (3.4)$$

donde  $\psi(\gamma' w_t) = \{1 + \exp(-\gamma' w_t)\}^{-1} - 1/2$ .

La ecuación (3.3) puede ser interpretada como un modelo no lineal donde el intercepto  $\pi_0 + \theta_0 \psi(\gamma' w_t)$  varía con el tiempo y cambia suavemente desde  $\pi_0 - \theta_0/2$  hasta  $\pi_0 + \theta_0/2$  con  $\gamma' w_t$ . Así, el modelo descrito en (3.3) es un caso especial de la red neuronal con una única capa oculta:

$$x_t = \pi' w_t + \sum_{j=1}^q \theta_{0j} \{\psi(\gamma_j' w_t) - \frac{1}{2}\} + u_t \quad (3.5)$$

Considerando (3.4) y (3.5) se desea contrastar la hipótesis de que  $x_t$  (que se asume estacionaria) es lineal, *id est*,  $H_0 : x_t = \pi' w_t + u_t$  o, equivalentemente,  $H_0 : \theta_0 = 0$ . Nótese que  $\psi(0) = 0$ , por lo que es posible la hipótesis nula  $H'_0 : \gamma = 0$  frente la alternativa  $H'_1 : \gamma \neq 0$ . Para poder contrastar esto, se sustituye  $\phi$  por el desarrollo de Taylor de orden 3 alrededor de  $\gamma = 0$ , como proponen Luukkonen *et al.* [83]. Desde (3.4) se llega a:

$$\begin{aligned} t_3(\gamma' w_t) = \psi(0) + \sum_{i=0}^p \frac{\partial \psi(0)}{\partial \gamma_i} \gamma_i + \frac{1}{2} \sum_{i=0}^p \sum_{j=0}^p \frac{\partial^2 \psi(0)}{\partial \gamma_i \partial \gamma_j} \gamma_i \gamma_j \\ + \frac{1}{6} \sum_{i=0}^p \sum_{j=0}^p \sum_{k=0}^p \frac{\partial^3 \psi(0)}{\partial \gamma_i \partial \gamma_j \partial \gamma_k} \gamma_i \gamma_j \gamma_k \end{aligned} \quad (3.6)$$

Reemplazando  $\psi$  en la ecuación original (3.4) por su valor aproximado de (3.6)<sup>29</sup>:

$$x_t = \tilde{\pi}' w_t + \sum_{i=1}^p \sum_{j=1}^p \delta_{ij} x_{t-i} x_{t-j} + \sum_{i=1}^p \sum_{j=1}^p \sum_{k=1}^p \delta_{ijk} x_{t-i} x_{t-j} x_{t-k} + \tilde{u}_t \quad (3.7)$$

donde  $\delta_{ij} = d_{ij} \theta_0 \gamma_i \gamma_j \gamma_0$ , con  $d_{ij} = 1/36$  si  $i = j$  y  $d_{ij} = 1/18$  en otro caso,  $i, j \geq 1$ , y  $\delta_{ijk} = d_{ijk} \theta_0 \gamma_i \gamma_j \gamma_k$ , con  $d_{ijk} = 1/36$  si  $i = j = k$ ,  $d_{ijk} = 1/18$  si  $i = j$ ,  $i = k$  o  $k = j$ ,  $d_{ijk} = 1/6$  de otra manera,  $i, j, k \geq 1$ .

Si  $\gamma_0 = 0$ ,  $\gamma' w_t = \tilde{\gamma}' \tilde{w}_t$ , por lo que  $\delta_{ij} = 0 \quad \forall i, j$ . En ese caso, (3.7) no tiene términos de segundo orden. La hipótesis nula correspondiente a  $H'_0 : \gamma = 0$  es:

$$H'_0 : \delta_{ij} = 0, \delta_{ijk} = 0 \quad i = 1, \dots, p; j = i, \dots, p; k = j, \dots, p \quad (3.8)$$

El test se puede llevar a cabo de la manera habitual.

Nótese que el problema de desconocer el valor de  $\theta_0$  bajo la nula puede solventarse siguiendo la propuesta Davies [28] y de Luukkonen *et al.* [83], a partir de que:

$$SSR = \sum \tilde{u}_t^2 = \inf_{\theta_0} \{SSR(\theta_0)\}$$

<sup>29</sup>Las expresiones  $\frac{\partial^2 \psi}{\partial \gamma_i \partial \gamma_j}$  y  $\frac{\partial^3 \psi}{\partial \gamma_i \partial \gamma_j \partial \gamma_k}$  están desarrolladas en el anexo 3.4.

donde  $SSR(\theta_0)$  es la suma de los cuadrados de los residuos del modelo

$$x_t = \pi' w_t + \theta_0 \sum \sum \delta_{ij}^* x_{t-i} x_{t-j} + \theta_0 \sum \sum \sum \delta_{ijk}^* x_{t-i} x_{t-j} x_{t-k} + \tilde{u}_t$$

con  $\theta_0 \neq 0$ .

Así, el test de linealidad para (3.4) consiste en eliminar las series «duales» de tercer orden de la expansión de Volterra [113] de la función no lineal y contrastar la hipótesis nula de que los coeficientes de los términos cuadráticos y cúbicos son 0.

El propio Teräsvirta utiliza el test para encontrar y modelizar la no-linearidad en el ciclo económico a través de la relación entre el PNB<sup>30</sup> y el índice de indicadores adelantados provisto por el Departamento de Comercio de los Estados Unidos [53].

### 3.5 Test RESET<sup>31</sup>

Uno de los primeros tests para el contraste de linealidad es el test RESET descrito por Ramsey en 1969 [116]. La idea del test es que la no-linealidad de un proceso puede aceptarse cuando las combinaciones no-lineales de las variables explicativas (los retardos de la variable endógena en el caso de las series temporales univariantes) tienen capacidad explicativa respecto de la variable endógena. Este test va asociado a un determinado modelo inicial, como puede ser un modelo autorregresivo y realmente el test contrasta  $H_0$ : buena especificación vs  $H_1$ : mala especificación, no teniendo que provenir esta necesariamente de la existencia de no-linealidad en su dinámica.

Existen algunas complicaciones alrededor de la obtención de un vector de residuos de la regresión inicial, por lo que Ramsey lo construye a la manera de Theil<sup>32</sup> [135]. A partir de aquí, el test RESET trata de discriminar entre la hipótesis nula de que no existe una mala especificación y la alternativa de que las medias de  $\tilde{u}_i$  para  $i = 1, 2, \dots, N-K$ <sup>33</sup> pueden ser aproximadas por cierta función. Esta función a especificar debe conseguir que, bajo la hipótesis nula, los residuos de la regresión se distribuyan independientemente de los estadísticos definidos por la función. La función que propone Ramsey que cumple dicho criterio es  $q_{i1} = \mathbf{a}'_i \hat{\mathbf{y}}^{(2)}$ ,  $i = 1, 2, \dots, N-K$ . Donde  $\mathbf{a}'_i$  es la  $i$ -ésima fila de  $\mathbf{A}$ , e  $\hat{\mathbf{y}}^{(2)}$  el vector  $N \times 1$  formado por los cuadrados del vector que estima mediante MCO la variable dependiente. Esta expresión puede generalizarse definiendo  $q_{ij} = q_{1i} = \mathbf{a}'_i \hat{\mathbf{y}}^{(j+1)}$ , donde  $\hat{\mathbf{y}}^{(j)} = (\hat{y}_1^j, \hat{y}_2^j, \dots, \hat{y}_N^j)$  y, finalmente,  $\mathbf{q}_j = \mathbf{A}' \hat{\mathbf{y}}^{(j)}$ .

Así, bajo la hipótesis nula,  $\tilde{\mathbf{u}}$  se distribuye como una normal multivariante con vector media  $\mathbf{0}$  y matriz de varianzas-covarianzas  $\sigma^2 \mathbf{I}_{N-K}$ . Bajo la hipótesis alternativa,  $\tilde{\mathbf{u}}$  también se distribuye como una normal, pero con un vector de medias dado por  $\mathbf{A}' \boldsymbol{\xi}$  y matriz de varianzas-covarianzas  $\bar{\sigma} \mathbf{I}_{N-K}$ . Así, bajo la hipótesis nula, puede escribirse  $\tilde{\mathbf{u}}$  como

$$\tilde{\mathbf{u}} = \mathbf{A}' \boldsymbol{\xi} + \epsilon \quad (3.9)$$

donde  $\epsilon$  es un vector  $(N-K) \times 1$  normalmente distribuido de media cero.

El modelo especificado en (3.9) sugiere la posibilidad de realizar el contraste haciendo una regresión de  $\tilde{\mathbf{u}}$  sobre  $\boldsymbol{\zeta} = \mathbf{A}' \boldsymbol{\xi}$  y calculando el estadístico F del resultado. Sin embargo,  $\boldsymbol{\xi}$  es desconocido, por lo que debe aproximarse el vector  $\boldsymbol{\zeta}$  por una suma lineal de los vectores  $\mathbf{q}_j$  con  $j = 1, 2, \dots$ <sup>34</sup> Así, puede considerarse la regresión:

<sup>30</sup>Producto Nacional Bruto.

<sup>31</sup>Regression Equation Specification Error Test.

<sup>32</sup>Sea  $\tilde{\mathbf{u}} = \mathbf{A}\mathbf{y}$  el vector de residuos óptimo, donde  $\mathbf{A}$  es una matriz  $N \times (N-K)$  que satisface:

1.  $\mathbf{A}'\mathbf{X} = \mathbf{0}$
2.  $\mathbf{A}'\mathbf{A} = \mathbf{I}_{N-K}$
3. Sujeto a las condiciones 1 y 2,  $\mathbf{A}$  es aquella que optimiza la siguiente función objetivo:

$$\min_{\mathbf{A}} \mathbf{E}(\mathbf{u}'(\mathbf{A} - \mathbf{J})(\mathbf{A} - \mathbf{J})'\mathbf{u}) = \min_{\mathbf{A}} \sigma^2 \text{tr}(\mathbf{A} - \mathbf{J})(\mathbf{A} - \mathbf{J})' = 2\sigma^2 \min_{\mathbf{A}} (N - K - \text{tr}(\mathbf{A}'\mathbf{J}))$$

donde  $(\mathbf{A} - \mathbf{J})'\mathbf{u}$  denota el error de predicción y  $\mathbf{J}$  es la matriz  $N \times (N-K)$  que se obtiene eliminando  $K$  columnas de  $\mathbf{I}_N$ .

<sup>33</sup> $N$  y  $K$  son el número de filas y columnas respectivamente del regresor  $\mathbf{X}$ .

<sup>34</sup>No hay ninguna regla que proponga hasta que número ha de llegar  $j$ , si bien Ramsey menciona que suele ser suficiente con obtener  $\mathbf{q}_j$  para  $j = 1, 2, 3$ .

$$\tilde{u}_i = \alpha_0 + \alpha_1 q_{i1} + \alpha_2 q_{i2} + \dots + e_i \quad i = 1, 2, \dots, N - K \quad (3.10)$$

donde  $e_i$  se asume distribuido como una normal de media 0. Se calcula ahora el estadístico F para la significatividad conjunta de las  $\alpha$ 's. Bajo la hipótesis nula, los coeficientes de la regresión de (3.10) son cero. Bajo la alternativa de que la media del vector  $\tilde{\mathbf{u}}$  es  $\mathbf{A}'\xi$ , uno o más de los coeficientes tendrán media distinta de cero, es decir, el coeficiente de correlación múltiple no es cero.

Este test es tan común que lo incorporan muchos de los paquetes estadísticos más utilizados (e.g. Gretl, statsmodels en Python) y, por ello, no es difícil encontrar ejemplos donde es utilizado con una intención práctica. Entre otros, Taylor y Sarno lo aplican para contrastar la buena especificación de distintos modelos que tratan de explicar el comportamiento de los tipos de interés reales tras la conferencia de Bretton Woods [133]. Por otra parte, Dolan trata de encontrar un modelo que interpole el valor del índice del EuroQol<sup>35</sup> [31], basado en 243 estados de salud, desde valuaciones directas de un subconjunto de sus variables como “dolor”, “ansiedad” o “movilidad” y utiliza el test RESET para contrastar si el modelo recoge efectivamente las no-linealidades que puede detectar.

### 3.6 Test de Keenan

Partiendo de la misma intuición que Ramsey a la hora de buscar las no-linealidades a través de las combinaciones no-lineales de variables explicativas, Keenan [68] parte de las expansiones de Volterra [148, 15]:

$$x_t = \mu + \sum_{u=-\infty}^{\infty} a_u \varepsilon_{t-u} + \sum_{u,v=-\infty}^{\infty} a_{uv} \varepsilon_{t-u} \varepsilon_{t-v} + \sum_{u,v,w=-\infty}^{\infty} a_{uvw} \varepsilon_{t-u} \varepsilon_{t-v} \varepsilon_{t-w} + \dots \quad (3.11)$$

donde  $\{\varepsilon_t\}_{-\infty}^{\infty}$  es un proceso estacionario de media cero y las variables  $\varepsilon_t$  son independientes e idénticamente distribuidas. A través de ciertos estudios de Tukey utilizando expansiones de Fourier, trata de contrastar la linealidad de una serie temporal a través de la no existencia de términos multiplicativos en su representación como (3.11). El test se describe como sigue:

Partiendo de una expansión cuadrática<sup>36</sup> de Volterra, si el proceso es no lineal y no contiene una variable en forma no lineal, esta no linealidad se verá reflejada en la diagnosis de los residuos. El test contrasta  $H_0 : \sum_{u,v=-\infty}^{\infty} a_{uv} \varepsilon_{t-u} \varepsilon_{t-v} = 0$ , término cuadrático de la expansión de Volterra de (3.11). Para ello es necesaria una proyección en subespacios que requieren de todo el pasado de la serie, pero solo se observa una realización parcial del proceso  $(x_1, \dots, x_n)$ . Sin embargo, dichas proyecciones pueden ser aproximadas por regresiones en los  $M$  anteriores valores de  $x$  siendo  $M$  un valor arbitrario, suficientemente grande, que tomamos como dado. El proceso para su aproximación es el siguiente:

1. Se realiza la regresión MCO de  $x_s$  sobre  $\{1, x_{s-1}, \dots, x_{s-M}\}$  y se calculan los valores predichos  $\hat{x}_s$  y los residuos  $\hat{e}_s$  para  $s = M + 1, \dots, n$  y la suma de cuadrados de los residuos:  $\sum \hat{e}_s^2$ .
2. Se realiza la regresión MCO de  $x_s^2$  sobre  $\{1, x_{s-1}, \dots, x_{s-M}\}$  y se calculan los residuos  $\{\hat{\xi}_s\}$ , para  $s = M + 1, \dots, n$ .
3. Se realiza la regresión MCO de  $\hat{e} = (\hat{e}_{M+1}, \dots, \hat{e}_n)$  sobre  $\hat{\xi} = (\hat{\xi}_{M+1}, \dots, \hat{\xi}_n)$  y se obtienen  $\hat{\eta}$  y  $\hat{F}$ , siendo

$$\hat{\eta} = \hat{\eta}_0 \left( \sum_{M+1}^n \hat{\xi}_t^2 \right)^{1/2} \quad (3.12)$$

donde  $\hat{\eta}_0$  es el coeficiente de la regresión y

$$\hat{F}_{1,n-2M-2} = \frac{\hat{\eta}^2(n-2M-2)}{\langle \hat{e}, \hat{e} \rangle - \hat{\eta}^2} \quad (3.13)$$

donde los grados de libertad asociados al producto escalar  $\langle \hat{e}, \hat{e} \rangle$  son  $(n - M) - M - 1$ .

<sup>35</sup>El EuroQol es un instrumento estandarizado de medición del estado de salud general.

<sup>36</sup>Esto es, desde (3.11),  $x_t = \mu + \sum_{u=-\infty}^{\infty} a_u \varepsilon_{t-u} + \sum_{u,v=-\infty}^{\infty} a_{uv} \varepsilon_{t-u} \varepsilon_{t-v}$ .



Este test es utilizado usualmente para validar un modelo, con el fin de detectar no-linealidades no recogidas en el mismo. Un ejemplo de su aplicación sería el trabajo de Chen, donde se utilizan distintos tests además de este con el propósito mencionado [24]. El estudio parte de una caracterización lineal de las series de los saldos de las cuentas corrientes de los países de la OCDE<sup>37</sup>. Los resultados de la aplicación de los tests mencionados a los residuos del modelo lineal para cada país sugieren la presencia de no-linealidades, por lo que un enfoque no-lineal puede mejorar la caracterización de las series. Por ello, el autor propone un modelo de cambio de régimen de Markov para modelizarlas.

### 3.7 Test de Tsay

En 1985, Tsay [139] generaliza el resultado obtenido por Keenan [68] permitiendo recoger la no-linealidad producida por la interacción de los distintos retardos a través de sus productos cruzados (i.e.  $x_{t-i}x_{t-j}$ ,  $i, j = 1, \dots, p$ ). Así, los pasos en este nuevo test de Tsay son los siguientes<sup>38</sup>:

1. Se realiza la regresión MCO de  $x_t$  sobre  $\{1, x_{t-1}, \dots, x_{t-M}\}$  y se calculan los valores predichos  $\hat{x}_t$  y los residuos  $\hat{e}_t$  para  $t = M + 1, \dots, n$  y la suma de cuadrados de los residuos:  $\sum \hat{e}_t^2$ . El modelo resultante se denota por

$$x_t = W_t \Phi + e_t \quad (3.14)$$

donde  $W_t = (1, x_{t-1}, \dots, x_{t-M})$  y  $\Phi = (\Phi_0, \Phi_1, \dots, \Phi_M)'$  con  $M \in \mathbb{N}$  y  $n$  el tamaño de la muestra.

2. Se realiza la regresión MCO de  $Z_t$  sobre  $\{1, x_{t-1}, \dots, x_{t-M}\}$  y se obtiene el vector residual  $\hat{Y}_t$  para  $t = M + 1, \dots, n$ . Aquí, el modelo multivariante es:

$$Z_t = W_t H + Y_t \quad (3.15)$$

donde  $Z_t$ <sup>39</sup> es un vector de dimensión  $m = \frac{1}{2}M(M+1)$  definido por  $Z_t' = \text{vech}(U_t' U_t)$  con  $U_t = (x_{t-1}, \dots, x_{t-M})$  y  $\text{vech}$  denotando la mitad del vector apilado.

Se realiza la regresión MCO de  $\hat{e}_t$  sobre  $\hat{Y}_t$ . Sea  $\hat{F}$  el  $F$ -ratio de la regresión:

$$\hat{e}_t = \hat{Y}_t \beta + \varepsilon_t \quad (3.16)$$

definido como

$$\hat{F} = \frac{(\sum_{t=M+1}^n \hat{Y}_t \hat{e}_t)(\sum_{t=M+1}^n \hat{Y}_t' \hat{Y}_t)^{-1}(\sum_{t=M+1}^n \frac{\hat{Y}_t' \hat{e}_t}{m})}{\sum_{t=M+1}^n \frac{\hat{e}_t^2}{n - M - m - 1}}$$

Este contraste es utilizado, por ejemplo, en el trabajo de Maheu y McCurdy, donde el objetivo del estudio es identificar distintos tipos de mercados financieros (alcistas o *bull* y bajistas o *bear*) [84]. También es utilizado por Panagiotidis y Rutledge en un artículo acerca de los mercados del petróleo y del gas, en el que se descubre que existe un equilibrio a largo plazo entre los precios de ambos mercados a través del estudio de la cointegración [105]. El test de Tsay se utiliza aquí como herramienta de validación, en el sentido de que es usado para contrastar si existe un componente no-lineal no especificado en su modelo propuesto.

<sup>37</sup>Organización para la Cooperación y el Desarrollo Económicos.

<sup>38</sup>Nótese que el primer paso es idéntico al propuesto en Keenan (1984).

<sup>39</sup>Si agregamos  $Z_t$  por los pesos determinados por la estimación MCO en (3.14), obteniendo una variable escalar, todo el procedimiento es idéntico al de Keenan [68].



#### 4. MÉTODOS DE DETECCIÓN DE CAOS

##### 4.1 Métodos con muestras subrogadas

Esta técnica es utilizada por varios autores a la hora de diseñar los contrastes para detección de caos y puede verse una explicación detallada en el manual de Kantz y Schreiber [66]. Aquí se expondrán los aspectos claves de la herramienta. La idea es que normalmente se desconocen los valores de la distribución de cierto estadístico  $\lambda$  para un proceso estocástico lineal dado.

Primero de todo debe especificarse la hipótesis nula que se quiere contrastar (e.g., que los datos proceden de un proceso estocástico lineal) y un nivel de significación (que usualmente toma el valor  $\alpha = 0,05$ ). Dados entonces  $H_0$  y  $\alpha$ , deben crearse las  $K$  muestras subrogadas. Para cada una de ellas se calculará el estadístico  $\lambda_i$ , con  $i = 1, \dots, K$ , siendo  $\lambda_0$  el asociado a los datos originales. A partir de aquí si, por ejemplo, se diera el caso en que tenemos cierta seguridad de que  $\lambda$  sigue una distribución gaussiana, podríamos computar tanto  $\hat{\lambda}$  y  $\hat{\sigma}$  de la manera habitual y rechazar la hipótesis nula cuando  $|\lambda_0| > 2\hat{\sigma}_\lambda + \hat{\lambda}$ .

Generalmente no se dispone de información acerca de la distribución de  $\lambda$ , pero puede utilizarse como sustituto el ranking de la ordenación de los distintos  $\lambda$ 's. Así, se rechaza la hipótesis nula en el contraste bilateral cuando  $\lambda_0$  es uno de los  $(K+1)\frac{\alpha}{2}$  mayores (o menores)  $\lambda_i$ , obtenidos a partir de muestras subrogadas cuando  $i \neq 0$ .

Una vez explicado el sistema, se muestran algunos ejemplos de cómo obtener los datos subrogados. Si consideramos que los datos son independientes y procedentes de una distribución dada pero desconocida, un muestreo aleatorio simple sin repetición de los datos forma la misma distribución empírica y es, por construcción, independiente. Sin embargo, si consideráramos que los datos provienen de un proceso lineal estocástico con inputs gaussianos, un muestreo adecuado debe mantener el mismo espectro de potencia, y podría utilizarse diferentes transformadas de Fourier de los datos originales tal como sigue:

Sea  $\tilde{s}_k$  la transformada rápida de Fourier de los datos originales  $x_t$  (FFT):

$$\tilde{s}_k = \frac{1}{\sqrt{T}} \sum_{t=1}^T x_t e^{i2\pi tk/T} \quad (4.1)$$

Se multiplican los componentes complejos de  $\tilde{s}_k$  por fases aleatorias:  $\tilde{s}_k' = \tilde{s}_k e^{i\phi_k}$ , donde  $\phi_k$  sigue una distribución  $\mathcal{U}(0, 2\pi)$ . Se computa la inversa de la transformada:

$$s_t' = \frac{1}{\sqrt{T}} \sum_{k=1}^T \tilde{s}_k' e^{-i2\pi tk/T} \quad (4.2)$$

La nueva serie  $s_t'$  contiene ahora números aleatorios con el espectro prescrito y variando  $\phi_k$  se obtienen más muestras subrogadas.

Un último ejemplo que se expone es cuando los datos no siguen una distribución gaussiana. La hipótesis más general que permite esa desviación, pero que debemos contrastar de cualquier manera, es que el output de una serie que sí es gaussiana ha sido distorsionado por una función monótona y dependiente del tiempo  $s$  tal que

$$s_t = s(x_t) \quad (4.3)$$

donde  $x_t$  sigue un proceso estocástico lineal gaussiano. El enfoque habitual en estos casos es invertir  $s$  empíricamente reescalando los valores para que sigan una distribución gaussiana a través de la transformada de Fourier de amplitud ajustada (AAFT), esto es:  $\hat{x}_t = \hat{s}^{-1}(s_t)$ . Ahora la aleatorización se produce en  $\hat{x}_t$ , obteniendo distintas subrogadas  $\hat{x}_t'$  que son reescaladas de nuevo a la distribución original  $\hat{s}_t' = \hat{s}(\hat{x}_t')$ . En este proceso, aunque la distribución original se mantiene, el espectro de potencia cambia ligeramente, dado que  $s$  ha de estimarse y que  $\hat{x}_t$  difiere de la  $x_t$  real. Estas estimaciones, sin embargo, suelen ser bastante buenas. Si se quisiera mejorar el proceso podría implementarse

un filtro como el de Schreiber y Schmitz [121], similar a un filtro de Wiener, que resulta en el espectro correcto de la series subrogadas.

Miller *et al.* aplican los algoritmos basados en subrogación [97] propuestos por Theiler [137, 136] y Small [126] al paso (manera de andar) de los individuos para detectar si es un proceso caótico. Las diferencias entre los resultados de los distintos algoritmos son notables: el algoritmo de Theiler cambia drásticamente la dinámica intracíclica del paso, obteniendo mayores valores en las métricas de caos utilizadas (exponente de Lyapunov y entropía) en los datos subrogados que en la serie original.

En la sección 5 se aplica un método basado en muestras subrogadas<sup>40</sup> que utiliza el embebimiento de Takens para representar las series temporales como imágenes, crea otras imágenes a partir de muestras de la serie original y compara si existe diferencias entre las primeras y las segundas (i.e. si existe un patrón observable en la serie original) a través de un algoritmo de clasificación creado a partir de una red neuronal convolucional<sup>41</sup>.

## 4.2 Test de Shintani-Linton

Este test, publicado por Shintani y Linton en 2003 [80], se basa en la definición de proceso caótico a partir de la positividad del exponente de Lyapunov en un sistema no-lineal acotado y disipativo<sup>42</sup> [33]. El planteamiento es utilizar los exponentes para contrastar la hipótesis nula de presencia de caos, junto a un estimador consistente de la varianza. La ventaja de este método es que no se queda en el análisis descriptivo del exponente, sino que establece un test estadístico.

Whang y Linton [146] derivan la normalidad asintótica del estimador del exponente de Lyapunov basado en el jacobiano de Eckmann *et al.* [34]. Shintani y Linton aquí, apoyándose en este resultado y en el de Chen y White [23] sobre redes neuronales, consiguen la distribución asintótica del estimador del exponente de Lyapunov de McCaffrey *et al.* [91].

Antes de definir el estimador del exponente de Lyapunov, debemos asumir que la serie temporal  $\{x_t\}_{t=1}^T$  viene generada por el modelo no-lineal autorregresivo:

$$x_t = \theta_0(x_{t-1}, \dots, x_{t-d}) + u_t \quad (4.4)$$

donde  $\theta_0 : \mathbb{R}^d \rightarrow \mathbb{R}$  es un mapeo dinámico no-lineal y  $\{u_t\}$  una secuencia de variables aleatorias. Así, el modelo expresado en (4.4) puede ser escrito en términos de un mapeo con vector de error  $U_t = (u_t, 0, \dots, 0)'$  y una función  $F : \mathbb{R}^d \rightarrow \mathbb{R}$  tal que:

$$Z_t = F(Z_{t-1}) + U_t \quad (4.5)$$

donde  $Z_t = (x_t, \dots, x_{t-d+1})' \in \mathbb{R}^d$ .

**Definición 26.** La **matriz jacobiana** del mapeo  $F$  de (4.5) evaluado en  $Z_t$  se define como:

$$J_t = \begin{pmatrix} \Delta\theta_{0,1,t} & \Delta\theta_{0,2,t} & \dots & \Delta\theta_{0,d-1,t} & \Delta\theta_{0,d,t} \\ 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & 0 \end{pmatrix} \quad (4.6)$$

para  $t = 0, 1, \dots, T-1$ , donde  $\Delta\theta_{0,j,t} = D^{e_j}\theta_0(Z_t)$  para  $j = 1, \dots, d$ , con  $e_j$  el  $j$ -ésimo vector canónico de  $\mathbb{R}^d$ .

<sup>40</sup>Denominado *Método de las imágenes*.

<sup>41</sup>Puede verse el procedimiento con más detalle en la nota al repositorio en [GitHub](https://github.com/i-rb/img4chaos) <https://github.com/i-rb/img4chaos> así como distintos ejemplos del uso de la biblioteca creada para su implementación y su aplicación al caso de la serie de IBM.

<sup>42</sup>Una **estructura disipativa** es aquella que se sostiene a través de tomar energía del ambiente y luego disipándola como entropía inútil [18].

Sea  $\hat{\theta}$  el estimador basado en redes neuronales de la función objetivo  $\theta_0$  en (4.4), considerando únicamente redes neuronales prealimentadas (*feed-forward*) de una única capa oculta y un único output.

**Definición 27.** El **estimador basado en redes neuronales** del  $i$ -ésimo mayor estimador de Lyapunov viene determinado por:

$$\hat{\lambda}_{iM} = \frac{1}{2M} \ln v_i(\hat{\mathbf{T}}'_M \hat{\mathbf{T}}_M) \quad (4.7)$$

para  $1 \leq i \leq d$ , donde  $v_i(A)$  es el  $i$ -ésimo mayor autovalor de la matriz  $A$ ,  $\mathbf{T}_M = \prod_{t=1}^M \hat{\mathbf{J}}_{M-t} = \hat{\mathbf{J}}_{M-1} \cdot \hat{\mathbf{J}}_{M-2} \cdots \hat{\mathbf{J}}_0$ ,

$$\hat{\mathbf{J}}_t = \begin{pmatrix} \Delta\hat{\theta}_{0,1,t} & \Delta\hat{\theta}_{0,2,t} & \dots & \Delta\hat{\theta}_{0,d-1,t} & \Delta\hat{\theta}_{0,d,t} \\ 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & 0 \end{pmatrix} \quad (4.8)$$

y  $\Delta\hat{\theta}_{j,t} = D^{e_j} \hat{\theta}(Z_t)$ , para  $t = 0, 1, \dots, M-1$  y  $M$  el numero de observaciones tomadas en  $t = \{T, T-1, \dots, T-M\}$ .

Bajo ciertas asunciones sobre la dinámica y la red neuronal<sup>43</sup> se obtienen algunos resultados teóricos. Se enuncia el teorema fundamental acerca de la distribución de el estimador.

**Definición 28.** Sea  $x_t$  la serie temporal dada y  $\lambda$  el mayor exponente del sistema (4.6), definido cada uno como  $\lambda_i = \lim_{M \rightarrow \infty} \frac{1}{2M} \ln v_i(\mathbf{T}'_M \mathbf{T}_M) < \infty$  y  $\mathbf{T}_M = \prod_{t=1}^M \mathbf{J}_{M-t}$ :

$$\Phi = \lim_{M \rightarrow \infty} V \left[ \frac{1}{\sqrt{M}} \sum_{t=1}^M \eta_t \right] \quad (4.9)$$

donde  $\eta_t = \ln |D\theta_0(x_{t-1})| - \lambda$ .

**Teorema 29.** Bajo ciertos supuestos,  $M \rightarrow \infty$  y  $M = \mathcal{O}([T/\ln T]^{1/(2+4\phi)})$ ,  $\phi \in \mathbb{R}_+$ , para  $1 \leq i \leq d$  se tiene que:

$$\sqrt{M}(\hat{\lambda}_{iM} - \lambda_i) \longrightarrow N(0, \Phi_i) \quad (4.10)$$

donde  $\lambda_M = \frac{1}{2M} \sum_{t=1}^M \ln[D\theta(x_{t-1})^2]$  y  $\hat{\lambda}_M = \frac{1}{2M} \sum_{t=1}^M \ln[D\hat{\theta}(x_{t-1})^2]$ .

Dado cierto  $M < \infty$ , al término  $\hat{\lambda}_M$  se le denomina el *estimador local del exponente de Lyapunov* y mide la divergencia de la serie en el corto plazo. A partir del teorema 29 se puede construir como sigue un test unilateral. La hipótesis nula a contrastar es ser un proceso caótico frente a no serlo, *id est*, basándonos en la definición que parte de los exponentes de Lyapunov:  $H_0 : \lambda \geq 0$  vs  $H_1 : \lambda < 0$ . Para ello, se define el estadístico  $\hat{t}$ .

**Definición 30.** Sea  $\hat{\lambda}_M$  el estimador local del exponente de Lyapunov descrito anteriormente y  $\hat{\Phi}$  un estimador consistente de  $\Phi$ , definimos el **estadístico**  $\hat{t}$  como:

$$\hat{t} = \frac{\hat{\lambda}_M}{\sqrt{\hat{\Phi}/M}} \quad (4.11)$$

Se rechaza la hipótesis nula si  $\hat{t} \leq -z_\alpha$  donde  $z_\alpha$  es el valor crítico que satisface que  $P(Z \geq z_\alpha) = \alpha$ , con  $Z$  siguiendo una normal estándar. Finalmente, se define el estimador consistente de  $\Phi$ .

<sup>43</sup>Véase el trabajo de Linton y Shintani [80], p. 6-11.

**Definición 31.** Sea  $\hat{\eta}_t = \ln |D\hat{\theta}_0(x_{t-1})| - \hat{\lambda}_M$ ,  $w(x)$  una función kernel<sup>44</sup> y  $S_M$  un parámetro de truncamiento de los retardos. El **estimador consistente de  $\Phi$**  se define como:

$$\hat{\Phi} = \sum_{j=-M+1}^{M-1} w(j/S_M) \hat{\gamma}(j) \quad (4.12)$$

$$\text{con } \hat{\gamma}(j) = \frac{1}{M} \sum_{t=|j|+1}^M \hat{\eta}_t \hat{\eta}_{t-|j|}.$$

Los autores recomiendan escoger el número de unidades  $r$  en la capa oculta de la red neuronal y el número de retardos  $d$  que minimicen el criterio BIC<sup>45</sup>.

Finalmente, se da un ejemplo de su aplicación. El test recién descrito ha sido utilizado por Kyrtsov *et al.* para contrastar la posible presencia de caos en los precios de la energía (gasolina, crudo, propano y gas natural) y en ninguno de ellos se encuentra evidencia de caos [76].

### 4.3 Test 0-1

El Test 0-1 de Gottwald y Melbourne [48, 49, 50, 51, 52] está diseñado para discriminar entre dinámicas regulares, i.e., periódicas o cuasiperiódicas y dinámicas caóticas. El método ha sido utilizado en diferentes contextos como ingeniería [81], electrónica [130], epidemiología [93], o también económico: Kriz analiza la variable macroeconómica fundamental, el PIB (de la economía checa) y encuentra que existe comportamiento caótico en su serie temporal [73]; Xin y Li presentan un sistema financiero que resulta caótico al introducir incentivos a la inversión y utilizan el test 0-1 para corroborarlo [152].

Para su construcción, partimos de una serie temporal  $x_t$  con  $t = 1, 2, \dots, T$ . A partir de ella se define el sistema bidimensional:

$$\begin{cases} p(t+1) = p(t) + x_t \cos ct \\ q(t+1) = q(t) + x_t \sin ct \end{cases} \quad (4.13)$$

dado  $c \in (0, 2\pi)$ . Además, tomamos el desplazamiento de la media cuadrática.

**Definición 32.** El **desplazamiento de la media cuadrática** (ponderada por el tiempo) se define como:

$$M(t) = \lim_{T \rightarrow \infty} \frac{1}{T} \sum_{j=1}^T [p(j+t) - p(j)]^2 + [q(j+t) - q(j)]^2, \quad t = 1, 2, 3, \dots$$

y su **tasa de crecimiento**:

$$K = \lim_{t \rightarrow \infty} \frac{\log M(t)}{\log t} \quad (4.14)$$

Bajo ciertas condiciones generales,  $M(t)$  y  $K$  existen.  $K$  puede tomar dos valores, si  $K = 0$ , el proceso tiene una dinámica regular. En cambio, para  $K = 1$  el proceso es caótico.

A continuación se introduce mínimamente el marco teórico para una mayor comprensión del test. Previamente existían algunos resultados acerca de las tasas de crecimiento en sistemas con simetría euclídea. En concreto, una clase de sistemas en tiempo discreto con simetría euclídea plana viene dada por ecuaciones de la forma:

<sup>44</sup>Se emplean las funciones en la clase  $\mathcal{W}$ ,  $w : \mathbb{R} \rightarrow [-1, 1]$ , similares a las usadas en Andrews(1991), e.g. el kernel de Barlett  $w(u) = 1 - |u|$ .

$\mathcal{W} = \{w : w(0) = 1, w(-x) = w(x) \forall x \in \mathbb{R}, \int_{\mathbb{R}} |w(x)| dx < \infty, w(x) \text{ continua en } 0 \text{ y en el resto salvo un número finito de puntos}\}$

<sup>45</sup>Bayesian Information Criterion, de Schwartz [122].

$$\begin{cases} m(t+1) = f(m(t)) \\ \vartheta(t+1) = \vartheta(t) + h(m(t)) \\ p(t+1) = p(t) + \phi(m(t))\cos(\vartheta(t)) - \psi(m(t))\sin(\vartheta(t)) \\ q(t+1) = q(t) + \phi(m(t))\sin(\vartheta(t)) - \psi(m(t))\cos(\vartheta(t)) \end{cases} \quad (4.15)$$

Aquí,  $f : X \rightarrow X$  define la dinámica base (perpendicular a las variables de simetría) mientras  $\vartheta(t)$  representa las rotaciones bidimensionales y  $p(t)$  y  $q(t)$  las traslaciones en el plano. Asumimos que  $h, \phi, \psi : X \rightarrow \mathbb{R}$  son diferenciables. Para esta clase de sistemas descrita, se tienen ciertos resultados que justifican el test. En [101] se muestra que, para dinámicas en  $X$  periódicas o cuasiperiódicas,  $p(t)$  y  $q(t)$  se mantienen acotadas. Sin embargo, si las dinámicas son suficientemente caóticas, las variables de traslación presentan un comportamiento difusivo [95, 51, 39]. Desde aquí, tomando la definición de (4.14) y computando su valor, se obtiene  $K = 0$  y  $K = 1$  en los respectivos casos<sup>46</sup>

#### 4.4 Test de Matilla-García y Ruiz Marín

El siguiente test, descrito por Matilla-García y Ruiz Marín [88], contrasta si una serie temporal dada presenta una dinámica determinista (incluido procesos de caos a bajo nivel) o estocástica y se basa en la dinámica simbólica<sup>47</sup>.

Para la comprensión del test, algunos conceptos previos de dinámica simbólica son necesarios. La idea es dividir el espacio de estados donde se presenta la dinámica en un número finito de particiones y, posteriormente, comparar los valores vecinos de una observación en dicha partición.

**Definición 33.** El **patrón ordinal** de dimensión de embebimiento  $m$ , en un tiempo  $t$  dado, se define como la permutación única:  $\pi_m(t) \equiv (r_0 r_1 \dots r_{m-1})$  del conjunto  $\{0, 1, \dots, m-1\}$  que satisface:

$$x_{t+r_0} \leq x_{t+r_1} \leq \dots \leq x_{t+r_{m-1}} \quad (4.16)$$

$$r_{s-1} \leq r_s \text{ si } x_{t+r_{s-1}} = x_{t+r_s} \quad (4.17)$$

Como se ve, (4.17) garantiza la unicidad de la permutación definida. Nótese que, en particular,  $\pi_m(t) = \pi_m(s)$  con  $t \neq s$  si y sólo si  $\forall k, l \in \{1, \dots, m-1\}$  con  $k \neq l$  se da que  $x_{t+l} < x_{t+k} \Leftrightarrow x_{s+l} < x_{s+k}$ . *Id est*, el patrón de ordenación, desde un momento  $t$  es idéntico a aquel producido desde el momento  $s$ .

**Definición 34.** Dada una serie temporal  $x_t$  y una dimensión de embebimiento  $m$ , la **frecuencia relativa** de un símbolo  $\pi$  se define de la manera usual como:

$$p(\pi) = \frac{\text{card}\{t | 0 \leq t \leq T - (m-1) : x_{m(t)} \text{ es de tipo } \pi\}}{T - (m-1)} \quad (4.18)$$

Una vez que se tienen estas definiciones, se presenta una medida de complejidad para realizar un estudio basado en la teoría de la información de Shannon [125]. Con esta medida se trata de distinguir si la serie temporal tiene un comportamiento regular, caótico o aleatorio. Hay diversas medidas, muchas de ellas basadas en la entropía<sup>48</sup> y aquí se utiliza la *entropía de permutación*.

<sup>46</sup>Nótese que en la construcción del test se parte de la serie  $x_t$ , generada por un sistema dinámico desconocido  $f$ , al que se le adjuntan las variables  $\vartheta$ ,  $p$  y  $q$  para crear un sistema como el descrito al que aplicar los resultados teóricos. El hecho de escoger  $h = ct$ ,  $\psi = 0$  y  $\phi(m(t)) = x_t$  reduce el sistema de (4.15) al descrito en (4.13).

<sup>47</sup>La dinámica simbólica es la modelización de un sistema dinámico a través de un espacio discreto consistente en infinitas secuencias de símbolos abstractos, cada uno de los cuales corresponde a un estado del sistema, con la dinámica (evolución) determinada por el operador de traslación que toma  $x \rightarrow f(x)$  y devuelve  $x \rightarrow f(x+a)$ .

<sup>48</sup>La entropía de la información o entropía de Shannon es la medida de la cantidad de información o incertidumbre inherente a las posibles realizaciones de la variable aleatoria a la que va asociada.

**Definición 35.** Sea  $\pi = (r_0 r_1 \dots r_{m-1}) \in S_m$  un vector aleatorio, donde  $S_m$  denota el grupo simétrico<sup>49</sup> de orden  $m!$ . A través de (4.18) se asigna la frecuencia relativa para cada símbolo de  $S_m$ , así  $\sum_{i=1}^{m!} p(\pi_i) = 1$ . Definimos entonces la **entropía de permutación** como:

$$h(m) \equiv - \sum_{i=1}^{m!} p(\pi_i) \log p(\pi_i) \quad (4.19)$$

Nótese que, desde esta definición,  $-\log p(\pi_i) = 0$  si  $\pi = \pi_i$  con completa certeza, que se denomina *falta de información*. Por otro lado, si  $p(\pi_i) \rightarrow 0$ ,  $-\log p(\pi_i) \rightarrow \infty$ <sup>50</sup>

A partir de estas definiciones, se construye un test para discriminar si la serie temporal dada es un proceso determinista o estocástico. Sea  $x_t$  con  $t = 1, \dots, T < \infty$  dicha serie temporal y  $m$  la dimensión del embebimiento que suponemos dada, estudiamos el comportamiento de  $h(m)$  cuando se incrementan el número de símbolos observados.

Sean dados  $w, k \in \mathbb{N}$  tal que  $w = \frac{m!}{k}$  y  $w \ll m!$ . Definimos  $\mathcal{W}_1 \subseteq \mathcal{W}_2 \subseteq \dots \subseteq \mathcal{W}_k$  como los subconjuntos de  $S_m$  contruidos recursivamente como sigue:

1.  $\mathcal{W}_1$  es una muestra de  $w$  símbolos escogidos aleatoriamente en  $S_m$ .
2.  $\mathcal{W}_j = \mathcal{W}_{j-1} \cup \{w \text{ símbolos aleatoriamente en } S_m \setminus \mathcal{W}_{j-1}\}$  para  $j = 2, \dots, k$ .

A partir de estos subconjuntos y de (4.19) obtenemos la entropía de permutación *modificada* como:

$$h^{\mathcal{W}_j}(m) = - \sum_{\pi_i \in \mathcal{W}_j} p(\pi_i) \log p(\pi_i)$$

Desde la secuencia  $\{h^{\mathcal{W}_j}(m)\}_{j=1}^k$  obtenemos el siguiente resultado:

**Teorema 36.** Sea  $x_t$  una serie temporal estacionada y  $m$  una dimensión de embebimiento dada. Entonces, una de las dos condiciones debe darse:

1.  $x_t$  es estocástica y  $h^{\mathcal{W}_j}(m) < h^{\mathcal{W}_{j+1}}(m) \quad \forall j = 1, \dots, k$
2.  $x_t$  es determinista y  $\exists j_0 : h^{\mathcal{W}_{j_0}}(m) = h^{\mathcal{W}_j}(m) \quad \forall j_0 \leq j \leq k$

La prueba al teorema está disponible en el anexo 4.4, tal cual se explicita en el artículo original [88].

Debe definirse la pendiente numérica de  $h^{\mathcal{W}_j}(m)$ <sup>51</sup> que servirá de base para la formalización del test.

**Definición 37.** Dada  $x_t$  serie temporal y  $m$  la dimensión del embebimiento, se define la **pendiente numérica** de  $h^{\mathcal{W}_j}(m)$  como:

$$dh^{\mathcal{W}_j}(m) = \frac{h^{\mathcal{W}_{j+1}}(m) - h^{\mathcal{W}_j}(m)}{\log((j+1)/j)} \quad (4.20)$$

La propiedad de que la pendiente numérica de  $h^{\mathcal{W}_j}(m)$  crece  $\log(jw)$  en un proceso estocástico frente a un proceso determinista, donde puede variar de manera no creciente, puede contrastarse utilizando una regresión con tendencia lineal:

$$dh^{\mathcal{W}_j}(m) = \alpha_0 + \alpha_1 j + \varepsilon_j, \quad j = 1, 2, \dots, k-1 \quad (4.21)$$

donde  $\varepsilon_j$  es un ruido blanco con  $E(\varepsilon_j^2) = \sigma^2$  y  $E(\varepsilon_j^4) < \infty$ .

Así, llegamos a la definición de las hipótesis, que son las siguientes:

- $H_0 : \alpha_1 = 0$ , *id est*, se trata de un proceso determinista.
- $H_1 : \alpha_1 > 0$ , *id est*, se trata de un proceso estocástico.

Nótese que dicho contraste puede realizarse mediante el habitual test T unilateral (debido a que  $\alpha_1$  no puede ser negativo por construcción).

<sup>49</sup>El grupo formado por todas las permutaciones de longitud  $m$ .

<sup>50</sup>Para una explicación con detalle, véase el trabajo de Ullah [140].

<sup>51</sup>Esto es así porque, de acuerdo con el Teorema 36, si el proceso es estocástico, dicha secuencia aumentará  $\log(jw)$  y no ocurrirá si el proceso es determinista.

Este test de Matilla y Ruiz ha sido aplicado, por ejemplo, en el trabajo de Inglada-Pérez, donde se utiliza para contrastar si las series temporales de ciertos índices bursátiles (Dow Jones, Ibex35, Nasdaq y Nikkei) presentan un componente caótico [60]. Para ello utiliza distintos métodos (e.g. Lyapunov, Hurst, 0/1), y en lo que concierne al test de Matilla-García y Ruiz Marín, se rechaza la hipótesis nula de determinismo en todos los casos.

#### 4.5 Algoritmo MESAH<sup>52</sup>

El siguiente método, desarrollado por McDonough, Noonan y Hall [92] fue creado con la intención de realizar un algoritmo simple y menos intensivo computacionalmente que los métodos anteriores, sin perder por ello robustez. La idea principal es parecida a la propuesta de Matilla-García y Ruiz Marín [88]: buscar secuencias cortas repetidas en el tiempo. Como el nombre indica, el algoritmo utiliza los errores medios al cuadrado para crear un histograma tal como sigue:

1. Dada una serie temporal  $x_t$  para  $t = 1, 2, \dots, T$ ,  
definimos  $x_{max} = \max\{x_1, \dots, x_T\}$  y, análogamente,  $x_{min} = \min\{x_1, \dots, x_T\}$ .
2. Para  $t = 1, \dots, T - 1$ :  
$$value(t) = \frac{(x_t - x_{max})^2 + (x_{t+1} - x_{min})^2}{2}$$
3. Se obtiene el histograma de la variable  $value$ .

Para discriminar comportamientos caóticos de otras señales debe haber presencia de picos aislados en el histograma creado. Esto revela que ciertas secuencias cortas se repiten con recurrencia, aun cuando comienzan desde distintos valores previos y acaban en distintos valores posteriores, lo que se suele entender como muestra de un proceso caótico.

A pesar de tratarse de un algoritmo de relativa sencillez en la implementación y de suficiente antigüedad (1995) es difícil encontrar usos del mismo. Kotwal presenta un estudio acerca de la aplicabilidad de distintas modificaciones de la ecuación logística a la criptografía [72]. Utiliza el algoritmo MESAH, entre otros, para cuantificar la detectabilidad del componente caótico de la encriptación de la que se deriva su predictibilidad. *Id est*, cuanto más fácil sea detectar el componente caótico de la señal, más sencillo será decodificarla. Además, también es utilizado para la propia decodificación de la encriptación basada en la ecuación logística.

#### 4.6 El test de los tres estados

El test de los tres estados, desarrollado por Eyebe, Effa, Kom y Ali en [37] se basa en el análisis del patrón de la serie (como [88, 92]). La mejora que presenta este test frente a otros es que es capaz de discernir si la serie tiene un comportamiento caótico, periódico o cuasiperiódico<sup>53</sup>. En líneas generales, el método estudia la distribución de los estados del sistema y clasifica el patrón (forma) de la serie dada según se asemeje al que generaría una serie de cierto tipo (caótico, cuasiperiódico, periódico).

Una señal periódica está caracterizada por un patrón básico se repite con recurrencia separando las apariciones cierto periodo. Una señal cuasiperiódica presenta un patrón que tiende a repetirse periódicamente después de cierto tiempo de observación (grande). En el caso de dinámicas no-regulares, el periodo no puede ser determinado y el patrón básico que se repite no es tan obvio.

Sea  $x_j(k) = \phi(x(k))$ ,  $k \in \mathbb{N}$  un evento observable, con  $x(k) = (x_1, x_2, \dots, x_j, \dots, x_M)(k)$  el vector de estado y  $M$  su dimensionalidad.  $x_j$  es la serie temporal a caracterizar, con  $1 \leq j \leq M$ . Con el fin de determinar el patrón, definimos la

<sup>52</sup>Acrónimo de *MEan Squared error Histogram*.

<sup>53</sup>Estos son los tres estados.



función  $g$  tal que  $u_j = g(x_j)$  es la serie temporal en orden ascendente y la función  $q$  tal que  $v_j = q(u_j)$ , la distribución de los índices ordenados de la función  $g$ <sup>54</sup>.

**Proposición 38.** Si la distribución de los estados de  $x_j$  es periódica, entonces la correspondiente distribución de los índices  $v_j = q(g(x_j))$  es una función definida a trozos lineal donde todas las funciones internas que la componen tienen la misma pendiente. Esta pendiente es igual al ciclo de la función observable.

*Demostración.* Sea  $\phi(x)$  una función de periodo  $L$  observable y considérense  $P$  periodos. La longitud de la serie temporal  $x_j = \phi(x)$  es igual a  $N = LxP$  y el número de distintos estados es igual a  $L$ . Así,  $u_j = g(x_j)$  es una función escalón que presenta  $L$  escalones de longitud  $P$ , que podemos expresar como sigue<sup>55</sup>.

$$u_j(k) = x_j(L \cdot i + b) \quad \text{donde} \quad \begin{cases} k = 0, 1, \dots, LxP - 1 \\ i = k \text{ mód}(P) \\ b = \text{floor}(k/P) \end{cases} \quad (4.22)$$

Aplicando la función  $q$  a  $u_j$  tenemos que:

$$v_j(k) = L \cdot i + b \quad \text{donde} \quad \begin{cases} k = 0, 1, \dots, LxP - 1 \\ i = k \text{ mód}(P) \\ b = \text{floor}(k/P) \end{cases} \quad (4.23)$$

La ecuación (4.23) muestra que  $v_j$  es una función definida a trozos lineales que presenta  $L$  pendientes iguales todas ellas: el patrón de una señal periódica está caracterizado por una pendiente constante. Así,  $L$  es equivalente al número de estados distintos de  $\phi(x)$ .  $\square$

En el caso de ecuaciones cuasi-periódicas,  $v_j(k)$  sigue siendo una función definida a trozos lineales. Sin embargo, el patrón básico no se repite periódicamente y las pendientes en la distribución de  $v_j(k)$  depende de cada estado en  $\phi(x)$ , mientras que el número de estados depende del tiempo observacional  $N$ . Por tanto:

$$v_j(k, N) = L(b) \cdot i + b(N)$$

Así, aunque la distribución en  $v_j(k)$  no corresponda a un patrón de señales periódicas, presenta un límite tal que la señal cuasiperiódica puede ser caracterizada por una de sus pendientes. Con el fin de tener en cuenta la dependencia temporal de  $v_j(k, N)$ , se considera como caracterización del patrón la pendiente maximal ( $LS$ )<sup>56</sup>, definida como sigue:

$$LS(N) = \max_{1 \leq k \leq N-1} (v_j(k+1, N) - v_j(k, N))$$

La caracterización del patrón de una dinámica cuasiperiódica se caracteriza por  $L = \lim_{n \rightarrow \infty} LS(N)$ .

En ocasiones, como en el caso de datos brownianos,  $LS$  no tiene por qué estar acotado. Por ello, puede utilizarse su error cuadrático medio para detectar caos en la serie temporal.

**Definición 39.** Definimos el **error cuadrático medio de LS**,  $\sigma_{LS}(N)$  como:

$$\sigma_{LS}(N, n) = \sqrt{\frac{1}{N} \sum_{j=0}^p (LS(jN_0 + n) - \bar{LS})^2}$$

donde  $\bar{LS} = \frac{1}{N} \sum_{j=0}^p LS(jN_0 + n)$ , siendo  $N = pN_0 + n$ ,  $N_0$  el paso de integración y  $p \in \mathbb{N}$ .  $\sigma_{LS}(N, n)$  mide la capacidad del sistema dinámico de generar nuevos patrones en el tiempo. El valor  $n$  es la menor duración de la observación para la cual la pendiente maximal está bien evaluada y debe verificar  $n \ll N$ .

<sup>54</sup>E.g., sea  $x_j = x_j(1), x_j(2), x_j(3), x_j(4)$  con  $u_j = g(x_j) = (x_j(4), x_j(2), x_j(3), x_j(1))$  su ordenación ascendente, entonces  $v_j = q(u_j) = (4, 2, 3, 1)$ .

<sup>55</sup>La función  $\text{floor} : \mathbb{R} \rightarrow \mathbb{Z}$ ,  $\text{floor}(x) = \max\{m \in \mathbb{Z} : m \leq x\}$

<sup>56</sup>Del inglés *Largest Slope*.



Asumiendo que  $\sigma_{LS}(N, n)$  varía como una potencia de  $N$ , si  $\sigma_{LS}(N, n)$  está acotado, la dinámica es no-caótica, de acuerdo al comportamiento de  $LS$ . Con tal fin, se define:

$$\mu(N, n) = \frac{\log(1 + \sigma_{LS}(N, n))}{\log N}$$

$$K(n) = \lim_{N \rightarrow \infty} \mu(N, n)$$

$K$  representa la tasa de crecimiento asintótica de  $\sigma_{LS}(N, n)$ . Así,  $K$  debe ser igual a 0 en el caso de señales periódicas y cuasiperiódicas y mayor que cero con comportamientos caóticos.

Para distinguir entre comportamientos periódicos y cuasiperiódicos, se tiene en cuenta la variación global de  $\mu(N, n)$ . Por un lado,  $\sigma_{LS}(N, n) = 0$  para señales periódicas, consecuentemente  $\mu(N, n) = 0$ . Para señales cuasiperiódicas,  $\sigma_{LS}(N, n) < \infty$ , por lo que  $\mu(N, n) \rightarrow 0$ . Para el caso en que la dinámica es caótica,  $\mu(N, n)$  es creciente. Por tanto, el signo de la derivada de  $\mu(N, n)$  puede discernir entre los tres estados.

**Definición 40.** Definimos la **derivada de  $\sigma_{LS}(N, n)$**  como  $\lambda(n)$ :

$$\lambda(n) = \lim_{N_1 \rightarrow \infty} \sum_{k=1}^{m-1} (\mu(N_{k+1}, n) - \mu(N_k, n)), \quad N_1 \ll N_m$$

siendo  $N_1$  la menor duración de integración y  $N_m$  la mayor.<sup>57</sup>

Finalmente, el test de los tres estados o *3ST* asocia un tipo de dinámica a la serie temporal dependiendo del signo de  $\lambda$  como sigue.<sup>58</sup>

$$\begin{cases} \text{caótica} & \lambda(n) > 0 \\ \text{cuasiperiódica} & \lambda(n) < 0 \\ \text{periódica} & \lambda(n) = 0 \end{cases}$$

Un ejemplo de su aplicación puede encontrarse en el trabajo de Wontchui *et al.* donde es utilizado para contrastar el comportamiento caótico del mapa de Peter-De-Jong [150]. Dicho mapa se describe por medio de un sistema de dos ecuaciones:

$$\begin{cases} x_{n+1} = \sin(ax_n) - \cos(bx_n) \\ y_{n+1} = \sin(cx_n) - \cos(dy_n) \end{cases}$$

donde  $a, b, c, d \in \mathbb{R}$  son distintos parámetros del sistema. Una representación del mapa de Peter-De-Jong puede verse en el anexo 4.5 donde se muestran todas las coordenadas posibles que la serie puede generar, es decir, el atractor del sistema caótico. Este sistema, en general, no depende de las condiciones iniciales, esto es, cualquier elección de  $x_0, y_0$  resulta en la misma imagen, pero cuyos puntos se dibujan en distinto orden.

El resultado al que llegan Wontchui *et al.* es que el test de los tres estados tiene un buen desempeño en general y, para casi cualquier parametrización del mapa de Peter-De-Jong, el  $\lambda$  resultante es mayor que 0. Existen, sin embargo, casos donde  $\lambda \approx 0$ , pero los autores afirman que las parametrizaciones asociadas a estos casos pueden producir sistemas cuasiperiódicos o con un componente caótico muy débil.

#### 4.7 La técnica de la valoración del ruido

Este método ha sido desarrollado por Poon y Barahona a partir del contraste de no-linealidad de los propios autores [111] y aplicado, por ejemplo, por Caraianni [20], por Hu y Raman [59] o por Wysocki *et al.* [151]. En concreto, en el último artículo referido, los autores tratan de contrastar si la respiración de los humanos en reposo sigue un proceso caótico. El componente caótico resulta en todos los casos estudiados (nueve individuos) a través del test de

<sup>57</sup>En la práctica, se obtienen buenos resultados eligiendo un tiempo de retardo tal que  $n \leq \frac{N_1}{2}$ .

<sup>58</sup>Nótese que, además de distinguir entre los distintos estados, el test 3ST obtiene la periodicidad  $L$ , en su caso.

la valoración del ruido, mientras que los valores obtenidos mediante esta técnica tienen una correlación baja, pero positiva, con la dimensión de la correlación y el exponente maximal de Lyapunov.

Este test está diseñado con el fin de estudiar series cortas y con presencia de ruido, justamente dos características que comparten muchas series financieras y económicas. Por un lado, los test clásicos que generalmente obtienen buenos resultados son muy sensibles al ruido y, por otro, los tests que son robustos frente al ruido satisfacen únicamente condiciones de caos necesarias, por lo que son tests parciales. Además, al tratar el ruido por separado, mediante algún filtro previo, no se tiene la certeza de que este tipo de mecanismos preserve bien la dinámica caótica subyacente, como se desprende de los trabajos de Kostelich y Schreiber [71] y Walker [142]. La idea de este método, sin embargo, es la contraria: añadir nuevo ruido.

Como el contraste de no-linealidad es necesario en el test de detección de caos<sup>59</sup>, se presenta un breve resumen:

Sea la serie temporal  $x_t, t = 1, \dots, T$ , se construye la expansión de Volterra de grado  $d \in \mathbb{N}_0$  y con memoria  $k \in \mathbb{N}_0$ :

$$x_t^p = \alpha_0 + \alpha_1 x_{t-1} + \alpha_2 x_{t-2} + \dots + \alpha_k x_{t-k} + \alpha_{k+1} x_{t-1}^2 + \dots + \alpha_M x_{t-k}^d$$

donde  $M = \frac{(k+d)!}{k!d!}$  es la dimensión total. En el siguiente paso, se construye una medida de bondad de ajuste, la varianza normalizada de los errores:

$$\varepsilon(k, d)^2 = \frac{\sum_{t=1}^T (x_t^p(k, d) - x_t)^2}{\sum_{t=1}^T (x_t - \mu)^2}$$

donde  $\mu = \frac{1}{T} \sum_{t=1}^T x_t$ .

El mejor modelo se escoge siguiendo el criterio de información de Akaike:  $C(r) = \log \varepsilon(r) + \frac{r}{N}$ , siendo  $r$  el número de términos de la expansión truncada de Volterra. Dicho mejor modelo se compara con el mejor modelo lineal (*id est*,  $d = 1$ ) usando el estadístico F. Basado en este estadístico, si el p-valor es menor que 0,10, se puede afirmar que existe una clara componente no-lineal.

El corazón de la técnica de detección de caos es un algoritmo simple que se describe a continuación: Se añade ruido blanco (o linealmente correlado) de desviación típica  $\sigma$  creciente a los datos hasta que la no-linealidad deja de ser detectada<sup>60</sup>. El primer  $\sigma$  que lo consigue se denomina  $NL$ <sup>61</sup>. Si  $NL > 0$ , el proceso es caótico y el valor de  $NL$  da una medida de su intensidad. Por el contrario, si  $NL = 0$ , o bien la dinámica no es caótica o el ruido presente en los datos originales ya había neutralizado dicha componente. Por ello,  $NL > 0$  es una condición suficiente pero no necesaria para la presencia de caos, aunque los autores remarcan que la potencia es muy alta de cualquier manera.

#### 4.8 El árbol de decisión para procesos caóticos.

Toker, Sommer y D'Esposito desarrollan un algoritmo para detectar caos en series temporales [138]. En este caso, el contraste está diseñado de tal manera que evita los principales problemas que presentan los tests clásicos: alta sensibilidad al ruido y malos resultados en casos extremos, como también mencionan Poon y Barahona [111]. La idea de este método es combinar diferentes herramientas que solucionan los problemas habituales en un procesamiento automatizado mediante un árbol de decisión.

Dicho árbol de decisión puede verse gráficamente en el anexo 4.8 y el algoritmo, que denominamos *AADC*<sup>62</sup> se construye como sigue:

1. Primero se contrasta si los datos proceden de un proceso estocástico. Esto se realiza mediante un test basado en la idea de obtener muestras subrogadas de los datos originales. Esto es, se compara la entropía de permutación de los datos originales y de combinaciones aleatorias de muestras subrogadas siguiendo el esquema

<sup>59</sup>No tendría por qué utilizarse el test de no-linealidad de Poon y Barahona aunque, lógicamente, ellos lo recomiendan.

<sup>60</sup>Bajo un nivel estadístico de confianza previamente fijado.

<sup>61</sup>Noise Limit.

<sup>62</sup>De *Algoritmo del Árbol de Decisión para detección de Caos*.

planteado por Zuzino y Kulp [155]. *Cualquier* algoritmo puede servir tanto para crear muestras subrogadas como para determinar la posible estocasticidad. Para lo primero, el artículo original utiliza una combinación de la Transformada de Fourier con Amplitud Ajustada [137] y la Permutación de Fase Cíclica [63], mientras que Zuzino y Kulp se sirven únicamente del primero de estos algoritmos. Una breve descripción de ambos se da a continuación:

a) *Transformada de Fourier con Amplitud Ajustada.*

- 1) Sea  $x_t$  la serie temporal original, con  $t = 0, \dots, T-1$ .
- 2) Se computa la transformada de Fourier discreta:

$$z_n = \sum_{t=0}^{T-1} e^{2\pi i n t / T} x_t$$

- 3) Se aleatorizan las fases tal que:  $z'_n = z_n e^{i\phi_n}$  donde  $\phi_n \sim U(0, 2\pi)$
- 4) Se simetrizan dichas fases:

$$\begin{cases} \text{Re}(z''_n) = \frac{1}{2} \text{Re}(z'_n + z'_{T-n}) \\ \text{Im}(z''_n) = \frac{1}{2} \text{Im}(z'_n + z'_{T-n}) \end{cases}$$

- 5) Se invierte la transformada:

$$x'_t = \frac{1}{T} \sum_{n=0}^{T-1} e^{-2\pi i n t / T} z''_n$$

Debido a la simetría en las fases,  $x'_t$  es real y es nuestra muestra subrogada.

b) *Permutación de Fase Cíclica.*

- 1) Se extrae la fase instantánea de la serie original utilizando la transformada de Hilbert<sup>63</sup>
- 2) Si la fase extraída no está acotada, se estandariza entre 0 y  $2\pi$ .
- 3) Se buscan los ciclos individuales en la señal y se divide en ciclos seleccionando los puntos de discontinuidad, *id est*, los puntos de la señal estandarizada en los que la fase decae en exceso.
- 4) El principio y el final de la serie presentarán ciclos incompletos. Estos deben mantenerse en la muestra subrogada.
- 5) Se reconstruye una nueva serie con estas fases iniciales y finales, pero se permutan el resto de ciclos interiores aleatoriamente.
- 6) Se repite para cada muestra.

Tras ello, si la serie original presenta una entropía de permutación *similar* a la de las muestras subrogadas, dicha serie temporal se clasifica como estocástica. En caso contrario, el algoritmo pasa al punto 2.

2. En este punto se trata de eliminar el ruido de la serie original para realzar la posible dinámica caótica. Existen varias rutinas que pueden utilizarse en este punto, pero los autores recomiendan el algoritmo de reducción de ruido de Schreiber [120]:

a) *Algoritmo de reducción de ruido de Schreiber.*

Sea  $s_t$  la señal recibida, formada como  $s_t = x_t + \eta_t$ , siendo  $\eta_t$  el componente del ruido (no correlado con  $x_t$  y con autocorrelaciones que, al menos, decaen rápidamente) y  $x_t$  la serie subyacente tal que  $x_t = f(x_{t-m}, \dots, x_{t-1})$ , siendo  $f$  el mapeo determinista desconocido. Definimos  $\sigma = \sqrt{\langle \eta^2 \rangle}$  como la amplitud del ruido o el valor absoluto del nivel de ruido.

Desde aquí se trata de resolver la relación implícita de  $x_t - f(x_{t-m}, \dots, x_{t-1}) = 0$  para una de las coordenadas centrales (e.g.,  $x_{t-m/2}$ , con  $m$  par). Con el fin de obtener una estimación de  $\hat{x}_{t_0-m/2}$  para el valor  $x_{t_0-m/2}$  (ya que desconocemos  $f$ ), formamos vectores retardados de la señal:  $\mathbf{s}_t = (s_{t-m+1}, \dots, s_t)$  y determinamos

<sup>63</sup>La transformada de Hilbert convierte una señal real en una compleja (su *representación analítica*) a través de la expresión:

$$H(t) = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{f(u)}{u-t} du$$

aquellos que más cerca estén de  $\mathbf{s}_{t_0}$ . El valor medio de  $s_{t-m/2}$  es utilizado como estimador  $\hat{x}_{t_0-m/2}$ :

$$\hat{x}_{t_0-m/2} = \frac{1}{|\mathcal{U}_c(\mathbf{s}_{t_0})|} \sum_{\mathbf{s}_t \in \mathcal{U}_c(\mathbf{s}_{t_0})} s_{t-m/2}$$

donde la norma es la norma máxima. Para  $m$  grande se recomienda la norma euclídea.

3. Una vez eliminado el ruido, se comprueba si existe sobremuestreo<sup>64</sup> a través del test de Matthews, pues el posterior test de detección de caos puede ser inexacto en presencia de sobremuestreo, como indican Gottwald y Melbourne [50]. El test de Matthews devuelve una medida  $\eta$  y se considera que da un resultado positivo en sobremuestreo siempre que  $\eta > 10$ . En ese caso, se corrige iterativamente a través del *downsampling* de la serie por un factor 2 hasta que el test de sobremuestreo resulte negativo.

■ *Test de Matthews.*

Este test, que Matthews denomina como *tosco*<sup>65</sup> [88], simplemente chequea cuántas veces mayor ( $\eta$ ) es la diferencia entre el máximo y el mínimo global que la media cuadrática del desplazamiento de la misma serie según la definición que dan Gottwald y Melbourne [50]. Primero se define el siguiente sistema:

$$\begin{cases} p_c(t) = \sum_{j=1}^t x_j \cos jc \\ q_c(t) = \sum_{j=1}^t x_j \sin jc \end{cases} \quad (4.24)$$

y la media cuadrática del desplazamiento para una muestra temporal  $\tau_s$  como:

$$M_c = \frac{1}{T} \sum_{j=1}^T ([p_{c\tau_s}(j+t) - p_{c\tau_s}(j)]^2 + [q_{c\tau_s}(j+t) - q_{c\tau_s}(j)]^2) \tau_s^2 \quad (4.25)$$

con  $c$  escogido aleatoriamente tal que  $c \in (0, c^*)$ , donde  $c^* = 2\pi \frac{v_s^*}{v_s}$ ,  $v_s = \frac{1}{\tau_s}$  y  $v_s^* = 2v_{max}$ .

Así, la medida  $\eta$  queda definida como:  $\eta = \frac{\max_{t \in [1, T]} \{x_t\} - \min_{t \in [1, T]} \{x_t\}}{M_c}$

4. Finalmente, el AADC realiza un test 0-1 modificado con el fin de ser menos sensible al ruido, suprimir correlaciones causadas por la cuasiperiodicidad y normalizar la desviación típica de la señal del test.

a) *Test 0-1 modificado.* Sea  $x_t$  la serie temporal, se crea entonces el sistema bidimensional:

$$\begin{cases} p(t+1) = p(t) + x_t \cos ct \\ q(t+1) = q(t) + x_t \sin ct \end{cases} \quad (4.26)$$

donde  $c \in \mathbb{R}$  es un valor aleatorio entre 0 y  $2\pi$ .

Gottwald y Melbourne muestran que, si  $x_t$  es regular,  $p$  y  $q$  están acotadas; mientras que si muestran un comportamiento browniano,  $x_t$  sigue un proceso caótico. Dawes y Freeland [29] proponen una media cuadrática del desplazamiento ponderado por el tiempo al que añaden un término del ruido:

$$M_c = \frac{1}{T} \sum_{j=1}^T ([p_c(j+t) - p_c(j)]^2 + [q_c(j+t) - q_c(j)]^2) + \sigma \eta_t \quad (4.27)$$

donde  $p_c(t)$  y  $q_c(t)$  se definen como en (4.13) y  $\eta_t$  es una variable aleatoria con distribución  $\eta_t \sim U(\frac{1}{2}, \frac{1}{2})$ , mientras que  $\sigma$  es el *nivel de ruido*<sup>66</sup>. Finalmente, el estadístico  $K$  resultante del test 0-1 se define como

$$K_c = \text{corr}(n, M_c) \quad (4.28)$$

y se computa para una muestra  $C$  de 100 valores distintos de  $c$ , aleatoriamente seleccionados entre 0 y  $2\pi$ , siendo  $K = \text{mediana}(\{K_c : c \in C\})$ .

<sup>64</sup>El sobremuestreo, en procesamiento de señales, es aquel proceso de muestreo (digital) a una frecuencia más alta que el ratio de Nyquist.

<sup>65</sup>“Crude” en inglés.

<sup>66</sup>Se recomienda utilizar  $\sigma = 0,5$

b) *Elección de  $K$ .*

El resultado del test 0-1 es un valor  $K$  que tiende a 0 en sistemas periódicos y a 1 para procesos caóticos. Puede prefijarse un valor  $K^*$  para el cual se acepta o se rechaza la dinámica caótica (*id est*, si  $K > K^*$ , se clasifica como un proceso caótico. En otro caso, como periódico), pero también puede escogerse en base a la longitud de la serie temporal.

Se trata de un artículo reciente (2020) y, por ello, aún nadie ha utilizado esta metodología en ninguna aplicación.

#### 4.9 Test NEGM<sup>67</sup>

Este método se basa en la medida fundamental de la teoría de caos: el exponente de Lyapunov. Desarrollado en 1992 por Nychka *et al.* [102], comparte algunos aspectos con el Test de Shintani y Linton [80], como el uso de una red neuronal. El propio Shintani lo utiliza junto a Serletis [123] con el fin de contrastar si la bolsa de valores norteamericana (en concreto, la serie diaria del Dow Jones industrial) presenta un componente caótico. Los resultados son negativos, pero habiendo aplicado previamente el test BDS consiguen rechazar la nula de que la serie se comporta como un ruido blanco, por lo que hay dependencia aunque no se trate de un proceso caótico.

Sea  $x_t$  la serie temporal que viene generada por un modelo autorregresivo no-lineal de la forma:

$$x_t = f(x_{t-L}, x_{t-2L}, \dots, x_{t-mL}) + e_t \quad (4.29)$$

para  $1 \leq t \leq T$ , donde  $L$  es el parámetro de retardo y  $m$  la longitud de la regresión. Aquí  $f$  es una función desconocida derivable y  $e_t$  sigue una distribución de media cero y varianza constante (y finita). Desde (4.29) se produce una representación en forma del espacio de los estados:

$$X_t = F(X_{t-L}) + E_t, \quad F: \mathbb{R}^m \rightarrow \mathbb{R}^m$$

donde  $X_t = (x_t, x_{t-L}, \dots, x_{t-mL+L})'$ ,  $F(X_{t-L}) = (f(x_{t-L}, \dots, x_{t-mL}), x_{t-L}, \dots, x_{t-mL+L})'$  y  $E_t = (e_t, 0, \dots, 0)'$ .

Para conseguir una estimación de  $\lambda$ , exponente de Lyapunov, necesitamos primero estimar las matrices jacobianas  $J_t = \frac{\partial F(X_t)}{\partial X'}$ . La recomendación es utilizar redes neuronales, siguiendo a McCaffrey *et al.* [91], tal que:

$$f(X_{t-L}, \theta) = \beta_0 + \sum_{j=1}^q \beta_j \psi \left( \gamma_{0j} + \sum_{i=1}^m \gamma_{ij} x_{t-iL} \right)$$

donde  $\psi$  es la función de activación (conocida<sup>68</sup>)  $\theta$  un vector de parámetros que sirve para ajustar los datos al modelo via mínimos cuadrados no-lineales que minimiza  $S(\theta) = \sum_{t=1}^T [x_t - f(X_{t-1}, \theta)]^2$  y utiliza  $\hat{F}(X_t) = (f(x_{t-L}, \dots, x_{t-mL}, \hat{\theta}))$  para aproximar  $F(X_t)$ .

Como los valores apropiados de  $L$ ,  $m$  y  $q$  son desconocidos, se recomienda utilizar aquellos que minimicen el criterio BIC. Una vez se tiene  $\hat{F}(X_t)$  puede utilizarse directamente  $\hat{J} = \frac{\partial \hat{F}(X_t)}{\partial X'}$  como un estimador no paramétrico de  $J_t$ , como proponen White y Gallant [43].

El estimador del exponente maximal de Lyapunov es:

$$\hat{\lambda} = \frac{1}{2T} \log |\hat{v}_1(T)|$$

donde  $\hat{v}_1(T)$  es el mayor autovalor de la matriz  $\hat{\Lambda}_T' \hat{\Lambda}_T$  donde  $\hat{\Lambda}_T = \hat{J}_T \hat{J}_{T-1} \cdots \hat{J}_1$ .

<sup>67</sup>Nychka, Ellner, Gallant, MacCaffrey.

<sup>68</sup>Usualmente, como ya se ha mencionado, la función logística.

#### 4.10 Gráficos de recurrencia

Un atractor caótico contiene infinitas órbitas periódicas. Las órbitas pertenecientes al atractor son la clausura del conjunto de órbitas periódicas e inestables. A pesar de que dicho conjunto sea dinámicamente invariante, es difícil de hallar debido a la inestabilidad intrínseca y los algoritmos conocidos para ello son computacionalmente complejos.

Por ello, Bradley y Mantilla proponen una aproximación a través de los gráficos de recurrencia [12].

Según definen Eckmann y Ruelle [33], un gráfico de recurrencia de una serie temporal  $x_1, \dots, x_N$  es una visualización de la matriz de recurrencia de dicha secuencia: los píxeles de las coordenadas  $(i, j)$  y  $(j, i)$  se pintan en negro si la distancia entre los elementos  $i$ -ésimo y  $j$ -ésimo pertenece a cierto umbral tal que

$$\delta_l < \|x_i - x_j\| < \delta_h \quad (4.30)$$

con una norma apropiada. Los píxeles se dejan en blanco en otro caso.

Esta técnica permite observar qué frecuencias presenta la serie y dónde exactamente se registran las señales correspondientes. Además, los autores indican que es bastante robusto en presencia de ruido, preserva la invariabilidad de la dinámica y su estructura es independiente del embebimiento electo.

Si bien los gráficos de recurrencia de procesos caóticos tienen una estructura muy característica, Zbilut y Webber definen distintas métricas con el fin de cuantificar esta estructura en lo que se denomina el RQA<sup>69</sup> [154]. Sin embargo, este análisis cuantifica propiedades de la dinámica de manera similar a la que lo hace el exponente de Lyapunov o la dimensión de la correlación, pero no consigue dilucidar los detalles espacio temporales de la dinámica. Además, como muestran Iwanski y Bradley [62], dos gráficos de recurrencia estructuralmente distintos pueden obtener en la práctica unas métricas muy similares.

Las órbitas periódicas inestables constituyen la base del atractor caótico. Con el fin de analizarlo, existen algunos algoritmos para encontrar estas órbitas periódicas inestables (UPOs<sup>70</sup>), pero su complejidad computacional es demasiado grande, a pesar de poder utilizarse estimaciones locales de la dinámica.

Por ello, los autores recomiendan que, con el fin de encontrar una UPO en la práctica, simplemente se traten de detectar estructuras repetidas en el gráfico de recurrencia construido a partir de una trayectoria dada y utilizar esa información con el fin de encontrar los valores de la variable de estado asociados a esa estructura.

Algunas características de la serie pueden ser deducidas directamente del gráfico de recurrencia: un gráfico de recurrencia homogéneo es característico de los procesos estocásticos; si se desvanece siguiendo la diagonal, la serie no es estacionaria y presenta tendencia; si aparecen bandas blancas, algunos estados están demasiado lejos de lo esperable, podría ser un indicio de no estacionariedad; si aparecen líneas paralelas a la diagonal puede tratarse de un proceso periódico o caótico, siendo este segundo más probable cuando las líneas emergen de un punto aislado; si aparecen líneas curvadas, la evolución de los estados es similar pero con diferente velocidad, siendo un proceso determinista.

Los gráficos de recurrencia (de la señal original o modificada) son usuales en el proceso de detección y análisis de caos en series temporales. Por ejemplo, en [56], Harrison *et al.* tratan de modelizar la ritmicidad presente en los ataques epilépticos y sugerir una métrica para su detección y cuantificación.

<sup>69</sup>Recurrence Quantification Analysis.

<sup>70</sup>Unstable Periodic Orbits.

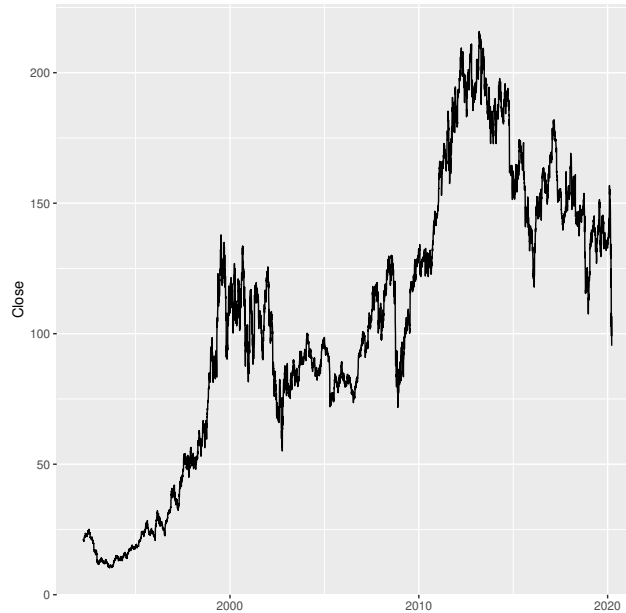
## 5. APLICACIÓN DE LOS MÉTODOS

### 5.1 Datos y metodología

#### 5.1.1 Datos

La serie temporal a la que se aplicarán los métodos anteriores es el valor de la cotización diaria de IBM desde el día 23 de marzo de 1992 hasta el día 20 de marzo de 2020<sup>71</sup> en el momento del cierre de la sesión. Su evolución temporal puede observarse en la figura 5.1.

FIGURA 5.1. Gráfico de la evolución de la cotización de IBM



Una posible explicación de esta evolución puede consultarse en el anexo 5.1.1.

#### 5.1.2 Metodología

La metodología será la propuesta por Inglada-Pérez [61]. En primer lugar, al carecer esta serie de estacionariedad, se toma una transformación de la serie original  $y_t$ : la diferencia de logaritmos  $z_t$ . Es decir:

$$z_t = \nabla \ln(y_t) = \ln y_t - \ln y_{t-1}$$

Además, para valores pequeños,  $z_t$  es una buena aproximación de la tasa de variación, por lo que sigue siendo fácil de interpretar. Los valores  $z_t$  se denominan *retornos*. La serie presenta el exceso de curtosis característico de algunas series financieras y algo de asimetría (tabla 1).

A continuación, se especifica la modelización ARMA con el fin de eliminar la dependencia lineal existente en la serie y se realizan sobre sus residuos los contrastes de raíz unitaria y estacionariedad: Dickey-Fuller aumentado [30], Phillips-Perron [109] y Kwiatkowski-Phillips-Schmidt-Shin [75]. En teoría, una modelización mediante modelos ARCH (GARCH y EGARCH) sería necesaria debido a que las series financieras suelen caracterizarse por una heterocedasticidad producida por la existencia de periodos de alta y de baja volatilidad.

<sup>71</sup>Estos datos están disponibles en el repositorio de GitHub junto al script utilizado en el análisis. Puede consultarse en <https://github.com/i-rb/tfgmath>. En un primer momento se obtuvieron de Yahoo <https://finance.yahoo.com/quote/IBM/>.

TABLA 1. Estadísticos y test básicos de las series

	Serie Original	Retornos	ARMA(1,1)	GARCH(1,1)	EGARCH(2,1)
Media	104,60	0,00022	1,40e-07	0,01032	0,01161
Mediana	105,18	0,00015	3,28e-05	0,01298	0,01308
D.T.	54,24	0,01765	0,01764	0,99958	0,99937
Mínimo	10,25	-0,16892	-0,16871	-7,85142	-6,93648
Máximo	215,80	0,12366	0,12376	9,59241	10,9442
Asimetría	-0,0805	-0,10812	-0,11809	-0,16738	-0,04383
Curtosis	-0,8867	7,80614	7,82982	6,26980	6,21917
Jarque-Bera	238,35***	17930***	18042***	11590***	11375***
N	7052	7051	7051	7050	7051

**Nota:** DT corresponde a la desviación típica. “\*”, “\*\*” y “\*\*\*” indican que el resultado es significativo al 90 %, 95 % y 99 % de nivel de confianza.

TABLA 2. Resultados de los contrastes de estacionariedad.

	Test	Serie Original	Retornos
	Dickey-Fuller Aumentado	-1,7808 (0,67)	-17,525 (0,01)***
	Phillips-Perron	-8,7222 (0,6234)	-7256 (0,01)***
	Kwiatkowski-Phillips-Schmidt-Shin	45,971 (0,01)***	0,28635 (0,1)*

**Nota:** DT corresponde a la desviación típica. “\*”, “\*\*” y “\*\*\*” indican que el resultado es significativo al 90 %, 95 % y 99 % de nivel de confianza. El p-valor se muestra entre paréntesis.

Finalmente, sobre los retornos, los residuos del modelo ARMA y los residuos estandarizados por sus desviaciones estándar condicionales de los distintos modelos no-lineales se aplican los métodos de detección de no-linealidad y caos.

## 5.2 Resultados empíricos

### 5.2.1 Modelización lineal y ARCH

Los modelos elegidos son aquellos que minimizan el criterio BIC de Schwarz<sup>72</sup> [122]. Son, respectivamente, el ARMA(1,1)<sup>73</sup>, el GARCH(1,1) y el EGARCH(2,1)<sup>74,75</sup>. En particular, basando la decisión exclusivamente en este criterio, se preferiría la especificación EGARCH(2,1) –con BIC -55082– frente al GARCH(1,1) –cuyo BIC es de -38647–, al ARMA(1,1) –con BIC 36902– y al modelo ARMA(0,0) sin constante –cuyo BIC es de -36910–. No es extraño que la mejor

<sup>72</sup>El criterio de información bayesiano (BIC) o de Schwarz penaliza la sobreparametrización por encima de los criterios de información usuales (AIC, Hannan-Quinn), de manera que suele utilizarse para evaluar la capacidad predictiva de un modelo.

<sup>73</sup>Realmente, el ARMA(0,0) sin constante mejora mínimamente este criterio y el autocorrelograma y la función de autocorrelación parcial son similares, siendo esto una evidencia más de que se está ante un ruido blanco. Sin embargo, el primer retardo del autocorrelograma y de la función de autocorrelación es significativo, por lo que se decide probar el ARMA(1,1) cuyos parámetros no son nada significativos. Se estudia el ARMA(1,1) dado que los residuos del ARMA(0,0) sin constante no dejan de ser la serie de retornos que se analiza de cualquier manera.

<sup>74</sup>Sin modelo de la media, i.e. un ARMA(0,0) sin constante tal que  $y_t = u_t$ .

<sup>75</sup>ARMA: *AutoRegressive Moving Average*.

GARCH: *Generalized AutoRegressive Conditional Heteroskedasticity*.

EGARCH: *Exponential Generalized AutoRegressive Conditional Heteroskedasticity*.



modelización de las propuestas sea la de tipo EGARCH, debido en gran parte a que estos modelos consiguen recoger el hecho apreciable en distintas series bursátiles de que *la volatilidad se comporta diferente frente a innovaciones positivas que frente a negativas*<sup>76</sup> *No reacciona de la misma manera frente a un alza en el precio de un activo que con respecto a una caída en el precio de este*, como indica Casas Monsegny <sup>[21]</sup>, quien entiende que es una característica producida como efecto del apalancamiento<sup>77</sup>

Las representaciones de la distribución de los residuos se encuentran en el anexo 5.2.1. Ninguna de las series de los residuos de los distintos modelos parece provenir de una distribución normal, debido a un exceso de curtosis en todas ellas, además de cierta asimetría. Los modelos tipo ARCH se han realizado después de comprobar que la serie de retornos muestra indicios de heterocedasticidad a través del test de Breusch-Godfrey <sup>[14]</sup> para órdenes del 1 al 8<sup>78</sup>. Por otro lado, el test no rechaza la hipótesis nula de homocedasticidad para los residuos de los dos modelos ARCH ni para el ARMA(1,1), debido a que la fuente de heterocedasticidad parece estar en la volatilidad heredada del periodo inmediatamente anterior. Las ecuaciones estimadas para cada modelo siguen a continuación y pueden comprobarse los detalles de la estimación de los parámetros en el anexo 5.2.1. Sea  $z_t$  es la serie en diferencia de logaritmos<sup>79</sup>

$$\text{Modelo ARMA(1,1):} \quad z_t = 0,00021 - 0,0121z_{t-1} - 0,01294\hat{u}_{t-1} + \hat{u}_t$$

(0,00021)      (0,38394)      (0,38711)

$$\text{Modelo GARCH(1,1):} \quad \begin{cases} z_t = \hat{u}_t \\ \epsilon_t = \frac{\hat{u}_t}{\hat{\sigma}_t}, \quad \epsilon_t \sim N(0,1) \\ \hat{\sigma}_t^2 = 6,3402E-06 + 0,088723 \epsilon_{t-1}^2 + 0,89448 \sigma_{t-1}^2 \end{cases}$$

(3,2399E-07)      (3,8473E-03)      (4,2009E-03)

$$\text{Modelo EGARCH(2,1):} \quad \begin{cases} z_t = \hat{u}_t \\ \epsilon_t = \frac{\hat{u}_t}{\hat{\sigma}_t}, \quad \epsilon_t \sim N(0,1) \\ \ln(\hat{\sigma}_t^2) = -0,13382 + \sum_{i=1}^2 \{ \hat{\alpha}_i (|\epsilon_{t-i}| - E[|\epsilon_{t-i}|]) + \hat{\gamma}_i \epsilon_{t-i} \} + 0,98273 \ln(\sigma_{t-1}^2) \\ \hat{\alpha}_1 = -0,03514, \hat{\alpha}_2 = -0,02088, \hat{\gamma}_1 = 0,21949, \hat{\gamma}_2 = -0,09226 \end{cases}$$

(0,002813)      (0,000309)      (0,016017)      (0,015203)      (0,007696)      (0,016559)

### 5.2.2 Análisis de la no linealidad

Tras la modelización de la serie de retornos, se comprueba la existencia de diferentes tipos de no-linealidades a través de algunos de los métodos descritos en la sección 3. La implementación práctica puede consultarse en los cuadernos del repositorio. Los resultados se muestran en la tabla 6 y de manera pormenorizada en el anexo Ra.

### 5.2.3 Análisis del comportamiento caótico

La existencia de no-linealidad es una condición necesaria para que el comportamiento del sistema sea caótico, pero no suficiente. En este punto, los residuos de los modelos ARCH parecen no provenir de un sistema no-lineal. O,

<sup>76</sup>Considerando, s.p.g., un EGARCH(1,1), la parte de la función que afecta a las innovaciones o residuos puede definirse como una función  $g(\epsilon_t) = \alpha|\epsilon_{t-1}| - E[|\epsilon_{t-1}|] + \gamma\epsilon_{t-1}$ . Así:

$$g(\epsilon_t) = \begin{cases} (\gamma + \alpha)\epsilon_{t-1} + E[|\epsilon_{t-1}|] & \epsilon_{t-1} > 0 \\ (\gamma - \alpha)\epsilon_{t-1} + E[|\epsilon_{t-1}|] & \epsilon_{t-1} < 0 \end{cases}$$

dependiendo del signo de la innovación su efecto en la modelización, mientras que en el GARCH(1,1)  $\epsilon_{t-1}$  se recoge al cuadrado, provocando un efecto simétrico.

<sup>77</sup>En el marco del mercado de valores, «el apalancamiento podría definirse como un tipo de facilidad operativa que ofrece el 'broker' (o intermediario financiero) al inversor, permitiendo hacer 'trading' con posiciones mayores que la cuantía de fondos requeridos.» <sup>[87]</sup>

<sup>78</sup>Los resultados pueden consultarse en el anexo.

<sup>79</sup>Entre paréntesis se muestra la desviación típica del parámetro estimado.

TABLA 3. Resultados de los contrastes de linealidad.

Test	H <sub>0</sub>	Retornos	ARMA(1,1)	GARCH(1,1)	EGARCH(2,1)
BDS	Linealidad	Rechaza***	Rechaza***	Mixto <sup>a</sup>	Mixto <sup>a</sup>
Keenan	Linealidad	Rechaza**	Rechaza**	No rechaza	No rechaza
Tsay	Linealidad	Rechaza**	Rechaza**	No rechaza	No rechaza
White	Linealidad	Rechaza***	Rechaza***	No rechaza	No rechaza
Teräsvirta	Linealidad	Rechaza**	Rechaza**	No rechaza	No rechaza
RESET	Linealidad	No rechaza	No rechaza	No rechaza	Rechaza**

**Nota:** DT corresponde a la desviación típica. “\*”, “\*\*” y “\*\*\*” indican que el resultado es significativo al 90 %, 95 % y 99 %. <sup>a</sup> El resultado depende de la elección paramétrica.

dicho de otra forma, sus no-linealidades ya están recogidas a través de la modelización de la varianza. No es así, sin embargo, en el caso del modelo lineal ni de la serie de retornos original.

En la tabla 7 se recogen las conclusiones a las que se accede tras la ejecución de distintos métodos para la detección de caos descritos en la sección 4. Todos los resultados se detallan en el anexo Rb.

TABLA 4. Resultados de los métodos de detección de caos.

Método	Contraste	Conclusión
G. de recurrencia	Determinismo	No se aprecia ningún patrón, parecen series estocásticas.
Test 0-1	Caos	El valor de $K > 0.99$ para todas las series: o se trata de ruido o de un componente caótico.
Coef. Hurst	Memoria a largo plazo	Cercanos a 0.5, por lo que puede considerarse que no existe memoria a largo plazo.
Exp. Lyapunov	Caos	Todas las series presentan exponentes $\lambda \lesssim -4$ . No es un proceso caótico.
Dim. Correlación	Determinismo	La dim. de correlación aumenta proporcionalmente con la de embebimiento, característica de los procesos aleatorios.
MESAH	Determinismo	No existen picos aislados: proceso aleatorio.
Mapa de Poincaré	-	No parece reconstruirse ningún atractor.
Valoración ruido	Caos	$NL \approx 0,8$ , para los retornos y el ARMA(1,1): cierto componente caótico.
Test de los tres estados	Caos	$K \approx 0,6$ . El proceso es caótico y no periódico.
Matilla y Ruiz	Determinismo	Se rechaza la nula de que el proceso es determinista para todas las series.
Árbol de decisión	Determinismo + Caos	Todas las series se clasifican como estocásticas.
Mét. de las Imágenes	Determinismo	No se detecta ningún patrón observable.

Nota: Los contrastes de *caos* pretenden detectar si un proceso es caótico asumiendo que este es determinista, siendo la alternativa que la dinámica es (cuasi)periódica. Los de *determinismo* disciernen entre si un proceso es estocástico o determinista.

## 6. CONCLUSIÓN

En el trabajo, tras presentar los conceptos y definiciones básicas de la teoría de caos, se han expuesto distintos mecanismos que permiten contrastar si una serie temporal determinada proviene de un proceso determinista o estocástico, lineal o no lineal, caótico o regular, según el método en cuestión.

Una vez realizados distintos contrastes a la serie elegida, el histórico de la cotización de IBM en la bolsa neoyorquina, se tienen claros indicios de que, por un lado, la serie es estocástica, como así resulta de todos los métodos que tratan de discernir entre determinismo y estocasticidad: el test de Matilla y Ruiz, el árbol de decisión, el algoritmo MESAH, la dimensión de correlación y los gráficos de recurrencia. Por otro lado, la serie presenta un componente no-lineal, que es modelizado mediante los modelos ARCH, como muestran los resultados de los distintos test de linealidad (e.g. Keenan, Tsay, White...).

También se ha comprobado si la serie presenta memoria prolongada a largo plazo a través del coeficiente de Hurst, pero no se puede rechazar estadísticamente que el valor real del exponente sea 0,5, i.e. efectivamente no parece que la serie sea persistente o antipersistente.

Los mecanismos expuestos parecen no contradecirse, en general, por lo que la utilización de esta batería de métodos puede ser de gran utilidad para todo aquel que desee modelizar cualquier serie temporal y, en concreto, una serie bursátil. El objetivo sería comprobar qué metodología es la más conveniente: para las series estocásticas, por ejemplo, en este trabajo se propone un modelo lineal ARMA asumiendo homocedasticidad y autorregresivos con heterocedasticidad condicional -ARCH- en el caso heterocedástico. Para series deterministas-caóticas, predecibles sólo en el corto plazo, se debería seguir el enfoque planteado en el algoritmo de la sección 2.4.

Por ello, en el caso concreto de la serie de IBM, dado que puede concluirse que se trata de una serie temporal estocástica que presenta heterocedasticidad condicional, es recomendable la especificación de un modelo de tipo ARCH para la modelización de la varianza. En concreto, y como se ha visto, EGARCH(2,1).

**Futuras líneas de investigación.** Las posibles futuras investigaciones son diversas. Una primera línea de investigación podría comparar los distintos métodos expuestos, tomando un gran conjunto de series temporales -tanto estocásticas como deterministas- y tratando de analizar con qué técnicas de detección de caos se tienden a obtener unos u otros resultados y la correlación entre los resultados de los métodos, probablemente diferenciando según el tipo de serie o su longitud.

Una segunda línea, similar a la que plantea Inglada-Pérez para los índices bursátiles [60], podría enfocarse en el estudio de la cotización de IBM y realizar el análisis de la relación de esta serie con otras series bursátiles y de la comparativa de los resultados de las predicciones de los distintos modelos lineales y no lineales propuestos.

Una última línea podría ser el desarrollo de un nuevo método de detección de caos, por ejemplo partiendo de los últimos resultados en clasificación de series temporales (TSC)<sup>80</sup>. Uno de los problemas centrales de este enfoque es el tratamiento de la serie para utilizarla como datos de entrada del algoritmo, con el fin de no perder la información de la ordenación temporal. Wang y Oates plantean el uso de distintas imágenes [143], mientras que otros, como Cui *et al.* [27] o Fawaz *et al.* [38], proponen clasificar directamente a través del aprendizaje profundo.

<sup>80</sup>Véase, por ejemplo, el trabajo de Bagnall *et al.* [3].

## REFERENCIAS

- [1] Andrzejak, R. G. *et al.* (2001). Indications of nonlinear deterministic and finite-dimensional structures in time series of brain electrical activity: dependence on recording region and brain state. *Phys. Rev. E* **64**, 061907.
- [2] Aznar, A. and Trivez, F. J. 1993. *Métodos de predicción en economía II*. Madrid: Editorial Ariel.
- [3] Bagnall, A., Bostrom, A., Large, J., and Lines, J. (2016) The great time series classification bake off: An experimental evaluation of recently proposed algorithms. *arXiv preprint*, arXiv:1602.01711.
- [4] Barnett, W. A. and Choi, S. S. (1989). A comparison between the conventional econometric approach to structural inference and the nonparametric chaotic attractor approach. *Economic Complexity: Chaos, Sunspots, Bubbles, and Nonlinearity*, 141- 212.
- [5] Barnett, W. A., Gallant, A. R., Hinich, M. J., Jungeilges, J. A., Kaplan, D. T., & Jensen, M. J. (1997). A single-blind controlled competition among tests for nonlinearity and chaos. *Journal of econometrics* **82**(1), 157-192.
- [6] Battin, T. J., Kaplan, L. A. *et al.* (2009). Supplementary information 1. *Nature Geoscience* **1**, 955–1000.
- [7] Baumol, W. J. and Benhabib J. (1989). Chaos: significance, mechanism, and economic applications. *The Journal of Economic Perspectives*, 77-79.
- [8] Baumol, W. J. and Quandt, R. E., 1996. *Chaos and their implications for forecasting en Trippi, R.R.* (ed.): *Chaos & nonlinear dynamics in the financial markets*. Irwin Professional Publishing.
- [9] Berryman, S. (2004). Leucippus in <https://plato.stanford.edu/entries/leucippus/>.
- [10] Blank, S. C. (1991). “Chaos” in futures markets? A nonlinear dynamical analysis. *Journal of Futures Markets* **11**(6), 711-728.
- [11] Boccaletti, S., Grebogi, C., Lai, Y.-C., Mancini, H., and Maza, D. (2000). The control of chaos: theory and applications. *Physics Reports* **329**.
- [12] Bradley, E. and Mantilla, R. (2002). Recurrence plots and unstable periodic orbits. *Chaos* **12**(3), 596-600.
- [13] Brainman, Y., and Goldhirsch, I. (1991). Taming chaotic dynamics with weak periodic perturbations. *Physics Review Letters* **68**.
- [14] Breusch, T. S. (1978). Testing for Autocorrelation in Dynamic Linear Models. *Australian Economic Papers* **17**, 334–355.
- [15] Brillinger, D. R. (1970). The identification of polynomial systems by means of higher order spectra. *J. Sound Vib* **12**(3), 301–313.
- [16] Brock, W. A., Hsieh, D. A. and LeBaron, B. D. (1991). *Nonlinear dynamics, chaos, and instability: statistical theory and economic evidence*. Cambridge: MIT press.
- [17] Brock, W. A., Scheinkman, J. A., Dechert W. D. and LeBaron, B. (1996). A test for independence based on the correlation dimension. *Econometric Reviews* **15**(3), 197-235, DOI: 10.1080/07474939608800353.
- [18] Buvet, R., Allen, M., & Massue, J. P. 1977. *Living Systems as Energy Converters*. Amsterdam: North-Holland.
- [19] Bueno, G. (1995). Principios de una teoría filosófica política materialista. *Anuario hispano cubano de filosofía*, **1995-96**.
- [20] Caraianni, P. (2013). Testing for nonlinearity and chaos in economic time series with noise titration. *Economic Letters* **120**, 192-194.
- [21] Casas Monsegny, M. and Cepeda Cuervo, E. (2008). Modelos ARCH, GARCH y EGARCH: aplicaciones a series financieras. *Cuadernos de Economía XXVII* **48**, 287-319.
- [22] Chaparro Guevara, G. (2016). *El control de sistemas dinámicos caóticos en economía: una aplicación a las reglas de política monetaria*. (Tesis doctoral). Universidad Complutense de Madrid, Madrid.
- [23] Chen, X., and White, H. (1999). Improved rates and asymptotic normality for nonparametric neural network estimators. *IEEE Transactions on Information Theory*, **45**(2), 682-691.
- [24] Chen, S. H. (2011). Current account deficits and sustainability: evidence from the OECD countries. *Economic Modelling* **28**, 1455-1464.
- [25] Chen, S. H., Lux, T. and Marchesi, M. (2001). Testing for non-linear structure in an artificial market. *Journal of Economic Behavior & Organization* **46**(3), 327-342.
- [26] Çinar, A. (1995). Nonlinear time series models for multivariable dynamic processes. *Chemometrics and Intelligent Laboratory Systems* **30**(1), 147–158, DOI: 10.1016/0169-7439(95)00060-7.
- [27] Cui, Z., Chen, W., and Chen, Y. (2016). Multi-scale convolutional neural networks for time series classification. *arXiv preprint*, arXiv:1603.06995.
- [28] Davies, R. B. (1987). Hypothesis testing when a nuisance parameter is present only under the alternative. *Biometrika* **74**(1), 33-43.
- [29] Dawes, J. H. P. and Freeland, M. C. (2020). The ‘0–1 test for chaos’ and strange nonchaotic attractors. *Preprint*.
- [30] Dickey, D. A. and Fuller, W. A. (1979). Distribution of the Estimators for Autoregressive Time Series with a Unit Root. *Journal of the American Statistical Association* **74**(366), 427–431.
- [31] Dolan, P. (1997). Modeling valuations for EuroQol health states. *Medical Care* **35**(11), 1095-1108.
- [32] Dorf, R., 1989. *Sistemas modernos de control*. Nueva York: Addison-Wesley.
- [33] Eckmann, J. P. and Ruelle, D. (1985). Ergodic theory of chaos and strange attractors. *Rev. Mod. Phys* **57**, 973-977.
- [34] Eckmann, J. P., Kamphorst, S. O., Ruelle, D., and Ciliberto, S. (1986). Lyapunov exponents from time series. *Physical Review A*, **34**(6), 4971.
- [35] Einstein, A. (1926). Letter to Max Born. *The Born-Einstein Letters*. Publicado en 1971. Nueva York: Walker and Company.
- [36] Escrivá de Balaguer, J. 1973. *Es Cristo que pasa*. Madrid: Rialp.
- [37] Eyebe Fouda, J. S. A., Effa, J. Y., *et al.* (2013). The three-state test for chaos detection in discrete maps. *Applied Soft Computing* **13**, 4731-4737.

- [38] Fawaz, H. I., Forestier, G., Weber, J., Idoumghar, L., and Muller, P. A. (2019). Deep learning for time series classification: a review. *Data Mining and Knowledge Discovery* **33**(4), 917-963.
- [39] Field, M., Melbourne, I., and Török, A. (2003). Decay of correlations, central limit theorems and approximation by Brownian motion for compact Lie group extensions. *Ergodic Theory Dyn. Syst* **23**, 87-110.
- [40] Fradkov, A. and Evans, R. (2005). Control of chaos: methods and applications in engineering. *Annual Reviews in Control* **29**.
- [41] Fradkov, A., Evans, R. and Andrievsky, B. (2006). Control of chaos: methods and applications in mechanics. *Philosophical Transactions: Mathematical, Physical and Engineering Sciences* **364**(1846).
- [42] Frazer A. M. and Swinney H. L. (1986). Independent coordinates for strange attractors from mutual information. *Phys Rev A* **33**(2), 1134-40.
- [43] Gallant, A. R., & White, H. (1992). On learning the derivatives of an unknown mapping with multilayer feedforward networks. *Neural Networks*, **5**(1), 129-138.
- [44] Gencer, M. and Unal, G. (2016). Testing non-linear dynamics, long memory and chaotic behaviour of energy Commodities. *Munich Personal RePEc Archive*.
- [45] Gerolimetto, M., and Bisaglia, L. (2014). Testing for (non)linearity in economic time series: a Montecarlo comparison. *Quaderni Di Statistica*, **16**(16), 5-32.
- [46] Gimeno, R., Mateos, R., Escot, L., Olmedo, E., and Grau, P. (2004). Detección de comportamientos caóticos en modelos TAR aplicados a series temporales de coyuntura económica española. *Revista Venezolana de Análisis de Coyuntura* **10**(1), 123-142.
- [47] Gómez-Corral, A. "Procesos estocásticos". Universidad Complutense de Madrid. Madrid. Septiembre de 2018.
- [48] Gottwald, G. A. and Melbourne, I. (2004). A new test for chaos in deterministic systems. *Proc. Roy. Soc. A* **460**, 603-611.
- [49] Gottwald, G. A. and Melbourne, I. (2005). Testing for chaos in deterministic systems with noise. *Physica D* **212**(1-2), 100-110.
- [50] Gottwald, G. A. and Melbourne, I. (2009). On the implementation of the 0-1 test for chaos. *SIAM J. Appl. Dyn.* **8**, 129-145.
- [51] Gottwald, G. A. and Melbourne, I. (2013). Central limit theorems and suppression of anomalous diffusion for systems with symmetry.
- [52] Gottwald, G. A. and Melbourne, I. (2016). The 0-1 test for chaos: A review. *Lecture Notes in Physics*, 915, 221-247. DOI: 10.1007/978-3-662-48410-4\_7.
- [53] Granger, C. W., Terasvirta, T. and Anderson, H. M. (1993). "Modeling nonlinearity over the business cycle." sh.8 in J. H. Stock and M. W. Watson (eds.), *Business cycles, indicators and forecasting*. Chicago: University of Chicago Press for the NBER, 311-326.
- [54] Grassberger, P. and Procaccia, I. (1983). Characterization of strange attractors. *Physical Review Letters* **50**, 346-349.
- [55] Greulich, P. E. (30 de septiembre del 2019). Comparing Market Valuation. *MBIConcepts*. <https://www.mbicconcepts.com/comparing-market-valuation.html>
- [56] Harrison, M. A., Frei, M. G. and Osorio, I. (2008). Detection of seizure rhythmicity by recurrences. *Chaos* **18**, 033124
- [57] Hübler, A., Georgii, R., Kuckler, M., Stelzl, W., and Lüscher, E. (1988). Resonant stimulation of nonlinear damped oscillators by Poincaré maps. *Helvetica. Physica Acta* **61**.
- [58] Hübler, A., and Plapp, B. (1990). Nonlinear resonances and suppression of chaos in the rf-Biased Josephson Junction. *Physical Review Letter* **65**(18).
- [59] Hu, S. and Raman, A. (2005). Chaos in atomic force microscopy. *Physical Review Letters* **96**(036107).
- [60] Inglada-Pérez, L. S. (2014). *Cinco ensayos sobre la complejidad en economía*. (Tesis doctoral). Universidad de Cantabria, Santander.
- [61] Inglada-Pérez, L. S. (2016). Uncovering nonlinear dynamics in air transport demand. *International Journal of Transport Economics* **43**(1-2).
- [62] Iwanski, J. and Bradley, E. (1998). Recurrence plots of experimental data: To embed or not to embed?. *Chaos* **8**, 861-871.
- [63] Jamšek, J., Paluš, M. and Stefanovska, A. (2010). Detecting couplings between interacting oscillators with time-varying basic frequencies: *Instantaneous wavelet bispectrum and information theoretic approach*. *Phys. Rev. E* **81**, 036207.
- [64] Jiménez-Rodríguez, R. (2009). Oil price shocks and real GDP growth: testing for non-linearity. *The Energy Journal* **30**(1).
- [65] Kantz, H. (1994). A robust method to estimate the maximal Lyapunov exponent of a time series. *Physiscs letters A* **185**(1), 77-87.
- [66] Kantz, H. and Schreiber, T., 2003. *Nonlinear Time Series Analysis*. Cambridge: Cambridge University Press.
- [67] Karakasidis, T. E. and Charakopoulos, A. C. (2008). Detection of low-dimensional chaos in wind time series. *Chaos, Solitons and Fractals* **41**, 1723-1732.
- [68] Keenan, D. (1985). A tukey non additivity-type test for time series. *Biometrika* **72**, 39-44.
- [69] Kelsey, D. (1988). The economics of chaos or the chaos of economics. *Oxford Economic Papers* **40**, 1-31
- [70] Kennel, M., Brown, R. and Abarbanel, H. D. I. (1992). Determining embedding dimension for phase-space reconstruction using a geometrical construction. *Phys Rev A* **45**, 3403-11.
- [71] Kostelich, E. J., and Schreiber, T. (1993). Noise reduction in chaotic time-series data: A survey of common methods. *Physical Review E* **48**(3), 1752.
- [72] Kotwal, J. (2008). *The Modified Logistic Map And Its Application To Cryptography*. Tufts University.
- [73] Kriz, R. (2013). Chaotic analysis of the GDP time series. In I. Zelinka, G. Chen, O. E. Röessler, V. Snašel, and A. Abraham, editors, *Nostradamus 2013: Prediction, Modeling and Analysis of Complex Systems*, volume 210 of *Advances in Intelligent Systems and Computing*, pages 353-362.

- [74] Kurths, J., and Herzel, H. (1987). An attractor in a solar time series. *Physica D: Nonlinear Phenomena*, **25**(1-3), 165-172.
- [75] Kwiatkowski, D., Phillips, P. C. B., Schmidt, P. and Shin, Y. (1992). Testing the null hypothesis of stationarity against the alternative of a unit root. *Journal of Econometrics* **54**(1-3), 159-178.
- [76] Kyrtsov, C., Malliaris, A. G. and Serletis, A. (2009). Energy sector pricing: on the role of the neglected nonlinearity. *Energy Economics* **31**, 492-502.
- [77] Lancaster, G., Iatsenko, D., Pidde, A., Ticcinelli, V. and Stefanovsk. A. (2018). Surrogate data for hypothesis testing of physical systems. *Phys. Rep.* **748**, 1-60.
- [78] Lahmiri, S. and Bekiros, S. (2019). Cryptocurrency forecasting with deep learning chaotic neural networks. *Chaos, Solitons and Fractals* **118**, 35-40.
- [79] Lindner, J. F., Kohar, V. *et al.* (2015). Strange nonchaotic stars. *Physical Review Letters* **114**(5), 054101.
- [80] Linton, O., and Shintani, M. (2003). Nonparametric Neural Estimation of Lyapunov. Discussion Paper EM/03/455.
- [81] Litak, G. and Syta, A. & Wiercigroch, M. (2009). Identification of chaos in a cutting process by the 0-1 test. *Chaos, Solitons & Fractals* **40**, 2095-2101.
- [82] Lorenz, H. W., 1993. *Nonlinear dynamical economics and chaotic motion*. Berlin: Springer-Verlag
- [83] Luukkonen, R., Saikkonen, P. and Teräsvirta, T. (1988). Testing linearity against smooth transition autoregressive models. *Biometrika* **75**(3), 491-499.
- [84] Maheu, J. M. and McCurdy, T. H. (2000). Identifying bull and bear markets in stock returns. *Journal of Business & Economic Statistics* **18**(1).
- [85] Mandelbrot, B. B., 1983. *The fractal geometry of nature*. Nueva York: Freeman.
- [86] Markoff, J. (29 de mayo de 1991). IBM Chief Gives Staff Tough Talk. *The New York Times*. Disponible en <https://www.nytimes.com/1991/05/29/business/ibm-chief-gives-staff-tough-talk.html>
- [87] Martínez Fariña, M. Apalancamiento: un concepto fundamental en el 'trading'. *BBVA*. Disponible en <https://www.bbva.com/es/apalancamiento-concepto-fundamental-trading/>
- [88] Matilla-García, M. and Marín, M. R. (2010). A new test for chaos and determinism based on symbolic dynamics. *Journal of Economic Behavior and Organization* **76**(3), 600-614. DOI: /10.1016/j.jebo.2010.09.017.
- [89] Matthews, P. (2020). 0 - 1 test for chaos. *MATLAB Central File Exchange*. Retrieved April 13, 2020.
- [90] Maxwell, J.C. (1860). Illustrations of the dynamical theory of gases. Part I. On the motions and collisions of perfectly elastic spheres. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* **19**(4), 19-32.
- [91] McCaffrey, D. F., Ellner, S., Gallant, A. R., & Nychka, D. W. (1992). Estimating the Lyapunov exponent of a chaotic system with nonparametric regression. *Journal of the American Statistical Association* **87**(419), 682-695.
- [92] McDonough, P., Noonan, P., and Hall, G. (1995). A new chaos detector. *Computers Elect. Engng* **21**, 6, 417-431.
- [93] McLennan-Smith, T. A. and Mercer, G. N. (2014). Complex behaviour in a dengue model with a seasonally varying vector population. *Mathematical Biosciences* **248**(0), 22 - 30.
- [94] Medio, A. 1992. *Chaotic dynamics. Theory and applications to Economics*. Cambridge: Cambridge University Press
- [95] Melbourne, I. and Török, A. (2004). Statistical limit theorems for suspension flows. *Israel Journal of Math.* **144**, 191-209.
- [96] Meseguer-Ruiz, O., Osborn, T. J. *et al.* (2019). Definition of a temporal distribution index for high temporal resolution precipitation data over Peninsular Spain and the Balearic Islands: the fractal dimension; and its synoptic implications. *Climate Dynamics* **52**, 430-456.
- [97] Miller, D. J., Stergiou, N. and Kurz, J. (2006). An improved surrogate method for detecting the presence of chaos in gait. *Journal of Biomechanics* **39**, 2873-2876.
- [98] Milnor, J. (1985). On the concept of attractor. *Comm. Math. Phys* **99**, 177-195.
- [99] Myrdal, G. (1972). The place of values in social policy. *Journal of Social Policy* **1**(1), 1-14. DOI: 10.1017/S0047279400002154.
- [100] Naruse, M., Matsubara, T. *et al.* (2019). Generative adversarial network based on chaotic time series. *Scientific Reports* **9**, 12963.
- [101] Nicol, M., Melbourne, I., and Ashwin, P. (2001). Euclidean extensions for dynamical systems. *Nonlinearity* **14**, 275-300.
- [102] Nychka, D., Ellner, S., Gallant, R. and McCaffrey, D., (1992). Finding chaos in noisy systems. *Journal of the Royal Statistical Society B* **54**, 399-426.
- [103] Ogata, K., 1998. *Ingeniería de control moderna*. Ciudad de México: Prentice-Hall.
- [104] Ott, E., Grebogi, C., & Yorke, J. (1990). Controlling chaos. *Physical Review Letters* **64**(11).
- [105] Panagiotidis, T. and Rutledge, E. (2007). Oil and gas markets in the UK: evidence from a cointegrating approach. *Energy Economics* **29**, 329-347.
- [106] Petkov, B. H. (2015). Detecting chaos in a complex system. *Applied Mathematics and Computation* **260**, 35-36, DOI: 10.1016/j.amc.2015.03.034
- [107] Pettini, M. (1989). Controlling chaos through excitations. En: Lima, R., Streit, L., and Mendes, R. (Eds.), *Dynamics and Stochastic Processes*. Nueva York: Springer-Verlag.
- [108] Phillips, A. (1954). Stabilization policy in a closed economy. *Economic Journal* **64**.
- [109] Phillips, P. C. B. and Perron, P. (1988). Testing for a Unit Root in Time Series Regression. *Biometrika* **75**(2), 335-346.



- [110] Poincaré, H. (1912). Chance. *The Monist* **22**(1), 31-52.
- [111] Poon, C. S., Barahona, M., (2001). Titration of chaos with added noise. *Proceedings of the National Academy of Sciences of the United States of America* **98**, 7107–7112.
- [112] Popper, K. R. (1992). *The open universe: An argument for indeterminism*, Vol. 2. East Sussex, RU: Psychology Press.
- [113] Priestley, M. B. (1980). State-dependent models: A general approach to non-linear time series analysis. *Journal of Time Series Analysis* **1**(1), 47-71.
- [114] Prigogine, I. 1997a. *Las leyes del caos*. Barcelona: Crítica.
- [115] Prigogine, I. 1997b. *End of Certainty*, pp. 162–85 (1 ed). Nueva York: Free Press.
- [116] Ramsey, J. (1969). Tests for specification errors in classical linear least squares regression analysis. *Journal of the Royal Statistical Society. Series B* **31**, 350–371.
- [117] Rosensetein, M. T., Collins, J. J. and De Luca, C. J. (1993). A practical method for calculating largest Lyapunov exponents from small data sets. *Physica D* **65**(1-2), 117-134.
- [118] Rothbard, M. N. (1988). Chaos theory: destroying mathematical economics from within? *The Free Market* **6**(3), 1.
- [119] Sarton, G. (1941). Second Preface to Volume XXXIII: Laplace's Religion. *Isis* **33**(3).
- [120] Schreiber, T. (1993). Extremely simple nonlinear noise-reduction method. *Phys. Rev. E* **47**, 2401.
- [121] Schreiber, T., and Schmitz, A. (1996). Improved surrogate data for nonlinearity tests. *Physical review letters* **77**(4), 635.
- [122] Schwarz, G. E. (1978). Estimating the dimension of a model. *Annals of Statistics* **6**(2), 461–464.
- [123] Serletis, A. and Shintani, M. (2003). No evidence of chaos but some evidence of dependence in the US stock market. *Chaos, Solitons and Fractals* **17**, 449-454.
- [124] Shalizi, Cosma R., 2006. *Methods and Techniques of Complex Systems Science: An Overview*. In Deisboeck, T. S, Kresh, J., Yasha (eds.). *Complex Systems Science in Biomedicine*. Nueva York: Springer, pp. 33–114.
- [125] Shannon, C. E. (1948). A mathematical theory of communication. *The Bell system technical journal* **27**(3), 379-423.
- [126] Small, M., Yu, D. and Harrison, R. G. (2001). Surrogate test for pseudoperiodic time series data. *Physical Review Letter* **87**(18) 188101-188104.
- [127] Soliman, A. S. (1996a). Transitions from stable equilibrium points to periodic cycles to chaos in a Phillips curve system. *Journal of Macroeconomics* **18**(1).
- [128] Stavarache, L. L., Narbutis, D. et al. (2019). Exploring Multi-Banking Customer-to-Customer Relations in AML Context with Poincaré Embeddings. *arXiv preprint*, 1912.07701.
- [129] Stewart, I. and King, G. 1991. *Symmetric chaos*. In *Nonlinear Equations in the Applied Sciences*, Eds: Ames, WF and Rogers, CF. 257–315. New York: Academic Press.
- [130] Swathy, P. S. and Thamilaran, K. (2014). Dynamics of SC-CNN based variant of MLC circuit: An experimental study. *International Journal of Bifurcation and Chaos*, 24(02), 1430008.
- [131] Takens, F. (1981). Detecting strange attractors in turbulence. *Dynamical systems and turbulence* **898**, 366–81.
- [132] Taylor, J. (1979). Estimation and control of a macroeconomic model with rational expectations. *Econometrica* **47**.
- [133] Taylor, M. P. and Sarno, L. (1998). The behavior of real exchange rates during the post-Bretton Woods period. *Journal of International Economics* **46**(2), 281-312.
- [134] Teräsvirta, T., Lin, C. F. and Granger, C. W. J. (1993). Power of the Neural Network Linearity Test. *Journal of Time Series Analysis* **14**(2), 209–220. DOI: 10.1111/j.1467-9892.1993.tb00139.x
- [135] Theil, H. (1965). The analysis of disturbances in regression analysis. *Journal of the American Statistical Association* **60**(312), 1067-1079.
- [136] Theiler, J. and Rapp, P. E. (1996). Re-examination of the evidence for low-dimensional, nonlinear structure in the human electroencephalogram. *Electroencephalography and Clinical Neurophysiology* **98**, 213-222.
- [137] Theiler, J., Eubank, S., Longtin, A., Galdrikian, B. and Farmer, J. D. (1992). Testing for nonlinearity in time series: the method of surrogate data. *Phys. D: Nonlinear Phenom.* **58**, 77–94.
- [138] Toker, D., Sommer, F.T. and D'Esposito, M. A. (2020). Simple method for detecting chaos in nature. *Commun Biol* **3**, 11. DOI: /10.1038/s42003-019-0715-9.
- [139] Tsay, R. (1986). Non-linearity tests for time series. *Biometrika* **73**, 461–466.
- [140] Ullah, A. (1996). Entropy, divergence and distance measures with econometric applications. *Journal of Statistical Planning and Inference* **49**(1), 137-162.
- [141] Van Thuan, V. (2015). From Micro Gravitational Waves to the Quantum Indeterminism. *Communications in Physics*, **25**(3), 247.
- [142] Walker, D. M. (1998). Local filtering of noisy nonlinear time series. *Phys. Lett. A* **249**, 209 –217.
- [143] Wang, Z., and Oates, T. (2015). Imaging time-series to improve classification and imputation. *arXiv preprint*, arXiv:1506.00327.
- [144] Weber, M., Chen, J., et al. (2018). Scalable graph learning for anti-money laundering: A first look. *arXiv preprint*, 1812.00076.
- [145] Wen, C. and Yang, J. (2019). Complexity evolution of chaotic financial systems based on fractional calculus. *Chaos, Solitons and Fractals* **128**, 242-251.

- [146] Whang, Y. J., and Linton, O. (1999). The asymptotic distribution of nonparametric estimates of the Lyapunov exponent for stochastic time series. *Journal of Econometrics*, **91**(1), 1-42.
- [147] White, H. (1989). An additional hidden unit test for neglect non-linearity in multilayer feedforward networks. *Proceedings of the International Joint Conference on Neural Networks*, 451–455.
- [148] Wiener, N. (1958). *Nonlinear problems in random theory*. New York: John Wiley
- [149] Wolf, A. (1986). Quantifying chaos with Lyapunov exponents. *Chaos* **16**, 285-317.
- [150] Wontchui, T. T., Effa, J. Y. *et al.* (2017). Dynamical behavior of Peter-De-Jong map using the modified 0-1 and 3ST tests for chaos. *Annual review of chaos theory, bifurcations and dynamical systems* **7**, 1-21.
- [151] Wysocki, M., Fiamma, M. N. *et al.* (2005). Chaotic dynamics of resting ventilatory flow in humans assessed through noise titration. *Respiratory Physiology & Neurobiology* **153**, 54-65.
- [152] Xin, B. and Li, Y. (2013). 0-1 test for chaos in a fractional order financial system with investment incentive. *Abstract and Applied Analysis*, 876298.
- [153] Yang, S. R. and Brorsen, B. W. (1993). Nonlinear dynamics of daily futures prices: conditional heteroskedasticity or chaos?. *Journal of Futures Markets* **13**(2), 175- 191.
- [154] Zbilut, J. and Webber, C. (1992). Embeddings and delays as derived from recurrence quantification analysis. *Phys. Lett. A* **171**, 199-203.
- [155] Zunino, L. and Kulp, C. W. (2017). Detecting nonlinearity in short and noisy time series using the permutation entropy. *Physics Letters A* **381**, 3627-3635.



## ANEXO

**Anexo 0. Información de la computación de los métodos y del repositorio.**

Todas las implementaciones y los datos están disponibles en el siguiente repositorio de **GitHub**: <https://github.com/i-rb/tfgmath>.

En este repositorio existen dos carpetas: *data* y *scripts* y un archivo readme.

1. En la carpeta ***data*** se encuentran los siguientes archivos:
  - a) **IBM.csv**, cotización de IBM desde el 23/3/92 hasta el 20/3/20 (utilizado en el análisis).
  - b) **IBM\_com.csv**, cotización de IBM desde el 2/1/62 hasta el 20/3/20.
  - c) **retorn.csv**, los retornos de la serie, la diferencia de logaritmo de la variable Close presenten en IBM.csv.
  - d) **arma11.csv**, los residuos del modelo ARMA(1,1) estimado.
  - e) **garc11.csv**, los residuos del modelo GARCH(1,1) estimado.
  - f) **egar21.csv**, los residuos del modelo EGARCH(2,1) estimado.
2. En la carpeta ***scripts*** se encuentran los guiones de las implementaciones de los distintos métodos, siendo el principal **TFG\_matematicas.R**. Todos, excepto uno, tienen asociado un cuaderno ejecutable<sup>81</sup> en Google Colaboratory al que se puede acceder libremente a través del enlace que se indica. La excepción es la copia de un programa de Matlab desarrollada por Toker *et al.* para el árbol de decisión [138].
  - a) **TFG\_matematicas.R**, el guion principal, está escrito en R y en él se implementan todos los métodos excepto aquellos que se especifican a continuación, que se realizan en otros cuadernos por claridad. Su enlace al cuaderno de Google Colaboratory es el siguiente: [https://colab.research.google.com/drive/1X3CjvW2J2UDByI1\\_\\_xd48oDYboMA2ouQ?usp=sharing](https://colab.research.google.com/drive/1X3CjvW2J2UDByI1__xd48oDYboMA2ouQ?usp=sharing).
  - b) **Valoracion\_ruido.R**, en este guion se implementa en R la técnica de valoración del ruido. Su enlace al cuaderno de Google Colaboratory es el siguiente: <https://colab.research.google.com/drive/10KAYB3aUQKAcImzdAAFIIEIC6TWoLzvPN?usp=sharing>.
  - c) **3ST\_Test.R**, en este fichero se encuentra la implementación del test de los tres estados en R. Su enlace al cuaderno de Google Colaboratory es el siguiente: <https://colab.research.google.com/drive/1RDmh-0NE4ZIsJsSmOhuSvTqLlyGh0qom?usp=sharing>.
  - d) **Matilla\_Ruiz.py**, en este archivo se encuentra la implementación en Python del test de Matilla y Ruiz basado en la entropía. Su enlace al cuaderno de Google Colaboratory es el siguiente: <https://colab.research.google.com/drive/1DRzGkx6cG7PYtbQAVat0RZhdLXVxOrVr?usp=sharing>.
  - e) **Graf\_recurrencia.py**, con este guion se obtiene uno de las cuatro formas de gráficos de recurrencia que se encuentran en el anexo. Su enlace al cuaderno de Google Colaboratory es el siguiente: [https://colab.research.google.com/drive/1nTRZQ70jctFSQ\\_8kGVTTXFVAYb9MdOLn?usp=sharing](https://colab.research.google.com/drive/1nTRZQ70jctFSQ_8kGVTTXFVAYb9MdOLn?usp=sharing).
  - f) **Chaos.m**, en este programa se implementa el árbol de decisión para la detección de caos en Matlab. Está programado directamente por D. Toker, autor del artículo, y puede descargarse libremente en <https://figshare.com/s/80891dfb34c6ee9c8b34>.

<sup>81</sup>Deben subirse los archivos correspondientes al entorno de ejecución.

**Anexo 3.4**

$$\frac{\partial^2 \psi}{\partial \gamma_i \partial \gamma_j} = - \frac{[\exp(-\gamma' w_t) \{1 + \exp(-\gamma' w_t)\}^2 - 2\{1 + \exp(-\gamma' w_t)\} \exp(-2\gamma' w_t)] x_{t-i} x_{t-j}}{\{1 + \exp(-\gamma' w_t)\}^4} - \frac{\{\exp(-\gamma' w_t) - \exp(-2\gamma' w_t)\} x_{t-i} x_{t-j}}{\{1 + \exp(-\gamma' w_t)\}^3}$$

para  $i, j \geq 1$ .

$$\frac{\partial^3 \psi}{\partial \gamma_i \partial \gamma_j \partial \gamma_k} = - \frac{\{\exp(-\gamma' w_t) - 2\exp(-\gamma' w_t)\} \{1 - 2\exp(-\gamma' w_t)\} x_{t-i} x_{t-j} x_{t-k}}{\{1 + \exp(-\gamma' w_t)\}^4}$$

para  $i, j, k \geq 1$ .

$$\frac{\partial^2 \psi(0)}{\partial \gamma_i \partial \gamma_j} = 0; \quad \frac{\partial^3 \psi(0)}{\partial \gamma_i \partial \gamma_j \partial \gamma_k} = \frac{1}{16} x_{t-i} x_{t-j} x_{t-k}$$

para  $i, j, k \geq 1$ .

Si  $k = 0$ :

$$\frac{\partial^3 \psi(0)}{\partial \gamma_i \partial \gamma_j \partial \gamma_k} = \frac{1}{16} x_{t-i} x_{t-j}$$

**Anexo 4.4**

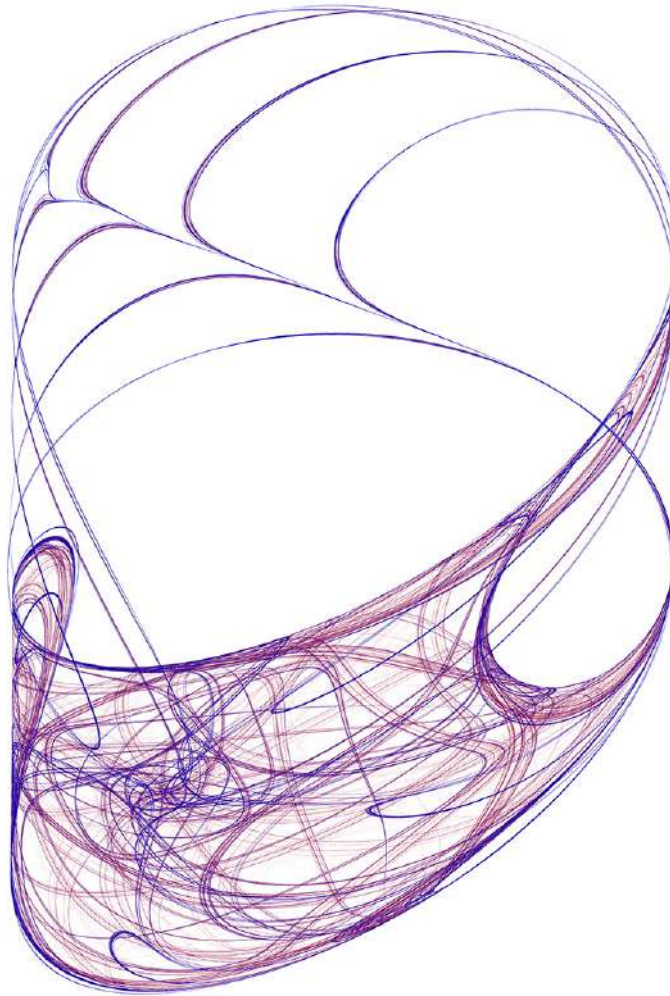
*Demostración* (del Teorema 16). Si el proceso subyacente es no-determinista, la máxima ganancia de información  $h^{M_j}(m)$  solo se alcanza cuando  $h^{\mathcal{W}_j}(m)$  alcanza el punto límite  $h(m) = \lim_{j \rightarrow k} h^{\mathcal{W}_j}(m)$ . En concreto, para cualquier subconjunto  $K \subseteq S_m$  se tiene que  $p(K) < p(K \cup \{\pi\})$  para cualquier  $\pi \in S_m \setminus K$ . Por ello:

$$- \sum_{\sigma \in K} p(\sigma) \log(p(\sigma)) < - \sum_{\sigma \in K \cup \{\pi\}} p(\sigma) \log(p(\sigma))$$

para cada  $\pi \in S_m$ . Por lo tanto,  $h^{\mathcal{W}_{j+1}}(m) - h^{\mathcal{W}_j}(m) > 0$  para  $j = 1, 2, \dots, k-1$ . Sin embargo, si el proceso es caótico,  $h^{M_j}(m)$  se alcanza *mucho antes* del límite. De hecho, existe cierto subconjunto  $H \subseteq S_m$  tal que  $p(H) = p(H \cup \{\pi\})$  para cada  $\pi \in S_m$ . Por ello:

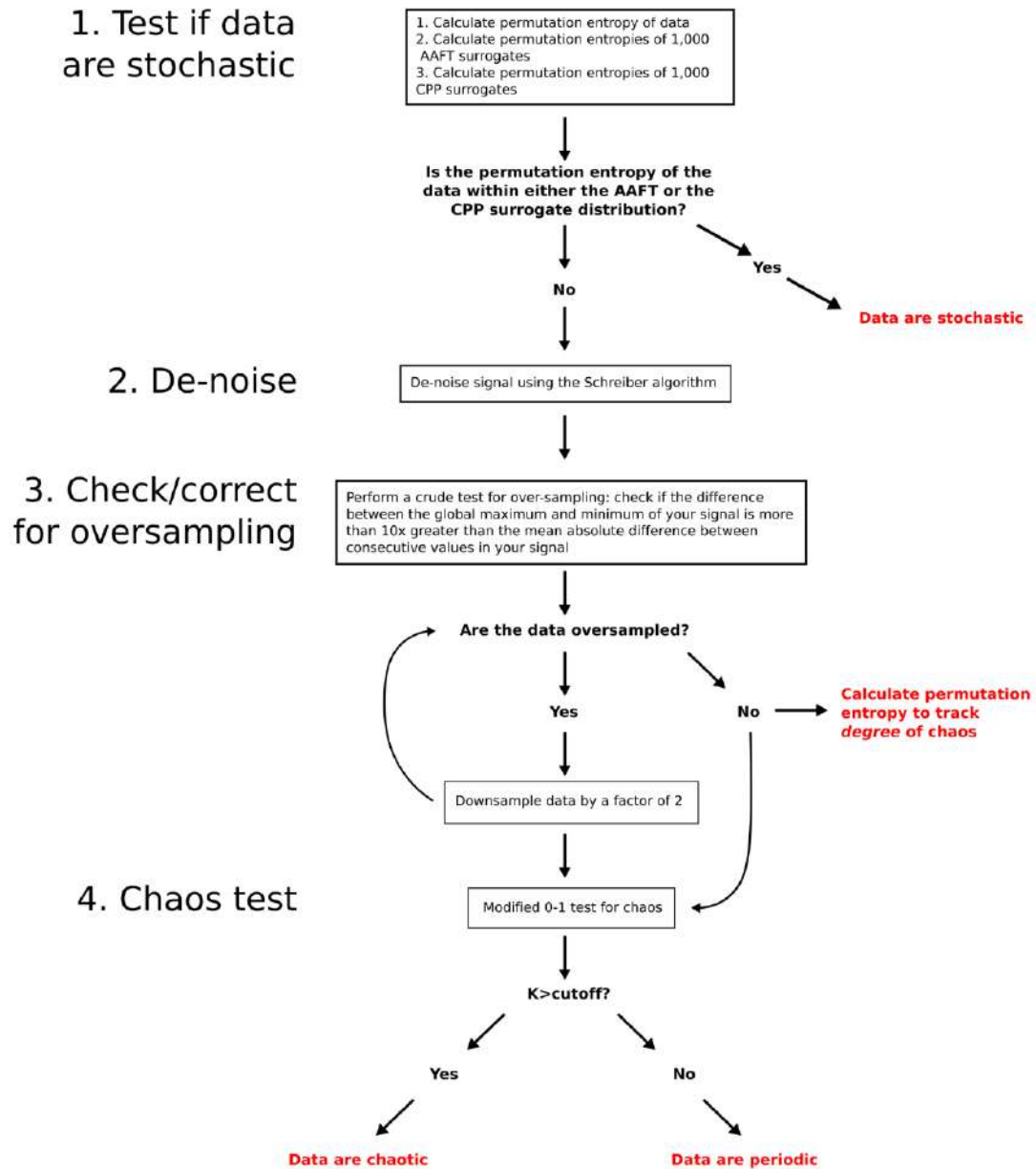
$$- \sum_{\sigma \in H} p(\sigma) \log(p(\sigma)) = - \sum_{\sigma \in H \cup \{\pi\}} p(\sigma) \log(p(\sigma))$$

para cada  $\pi \in S_m$ . Así, existe  $1 \leq j_0 \leq k$  tal que  $h^{\mathcal{W}_{j_0+s}}(m) = h^{\mathcal{W}_{j_0}}(m)$  para todo  $0 \leq s \leq k - j_0$ , haciendo que no se mantenga el crecimiento de  $h^{\mathcal{W}_j}(m)$ . De hecho, la serie  $h^{\mathcal{W}_j}(m)$  alcanza su límite varias veces, no únicamente una como ocurre en el caso estocástico.  $\square$

**Anexo 4.5**

*Fuente:* <http://paulbourke.net/fractals/peterdejong/>;  $a = -2.7$ ,  $b = -0.09$ ,  $c = -0.86$ ,  $d = -2.2$ .

## Anexo 4.8



Fuente: Toker et al. (2020)

### Anexo 5.1.1. Explicación de la evolución de la cotización de IBM.

Greulich ofrece una explicación a la evolución de la cotización de IBM basada en las actuaciones de los distintos CEO [55], variable fundamental en este tipo de análisis. Para contextualizar, a mediados de los ochenta, IBM entra en una pequeña crisis bajo la dirección de John F. Akers, quien dudó de sus empleados<sup>82</sup>, acerca de reducir los gastos y trató de vender las acciones de IBM como *babyblues*<sup>83</sup>. Por ello, el valor de mercado, que había alcanzado los 96B<sup>84</sup> de \$ en 1985, cae a los 29B de \$ en 1992, momento de su marcha.

Un nuevo CEO, Louis V. Gerstner, consigue recuperar la confianza de los empleados y su DeepBlue bate a Kasparov en 1996. IBM diversifica sus productos entre software, hardware y servicios y, en el momento de dejar la compañía, esta vale 208 B de \$ en el mercado. Sin embargo, los expertos creen que los resultados no son tan buenos en el contexto *la gran expansión* de la economía norteamericana y la relación de IBM con las *puntocom*, cuya burbuja estallaría en 2001.

Bajo la dirección de Samuel J. Palmisano se vende, en 2004, la división de PCs y el valor de la compañía alcanza en 2008 -en el contexto de los peores momentos de la crisis financiera global- un mínimo en el valor de mercado de 113B de \$.

Tras la marcha de Palmisano (IBM valía entonces 214B de \$ en el mercado) toma el relevo Virginia M. Rometty, quien parece repetir los errores de J. F. Akers, llevando el valor de mercado de la compañía hasta los 101B de \$. Recientemente (el 3 de febrero del 2020), Arvind Krishna ha sido nombrado nuevo director mundial de IBM. Un perfil mucho más técnico que, se supone, obtendrá mejores resultados para la compañía en un mundo cada vez más técnico.

### Anexo 5.2.1

#### 0. Resultados de la estimación

TABLA 5. Resultados de la estimación ARMA.

Modelo ARMA(1,1)				
	Coefficiente	DT	Estadístico t	P-valor
$\mu$	0,00021	0,00021	1,0326	0,3018
$\phi_1$	-0,01210	0,38394	-0,0315	0,9749
$\theta_1$	-0,01294	0,38711	-0,0334	0,9733

La notación refiere al siguiente modelo:

$$y_t = \phi_1 y_{t-1} + \theta_1 u_{t-1} + \mu + u_t$$

<sup>82</sup>Los llegó a acusar a sus empleados de estar *parados frente a las máquinas de agua*. [86]

<sup>83</sup>Término con el que se conoce a la depresión postparto en Estados Unidos.

<sup>84</sup>Billones norteamericanos. Esto es, 96 mil millones de \$.

TABLA 6. Resultado de la estimación GARCH.

<b>Modelo GARCH(1,1)</b>				
	Coefficiente	DT	Estadístico z	P-valor
$\omega$	6,3402e-06	3,2399e-07	19,569	< 2,2e-16
$\alpha_1$	8,8723e-02	3,8473e-03	23,061	< 2,2e-16
$\beta_1$	8,9448e-01	4,2009e-03	212,927	< 2,2e-16

La notación refiere al siguiente modelo:

$$\begin{cases} y_t = u_t \\ \epsilon_t = \frac{u_t}{\sigma_t}, \quad \epsilon_t \sim N(0,1) \\ \sigma_t^2 = \omega + \alpha_1 \epsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2 \end{cases}$$

TABLA 7. Resultado de la estimación EGARCH.

<b>Modelo EGARCH(2,1)</b>				
	Coefficiente	DT	Estadístico t	P-valor
$\omega$	-0,13382	0,002813	-47,5674	0,000000
$\alpha_1$	-0,03514	0,016017	-2,1940	0,028239
$\alpha_2$	-0,02088	0,015203	-1,3733	0,169667
$\beta_1$	0,98273	0,000309	3181,76	0,000000
$\gamma_1$	0,21949	0,007696	28,5203	0,000000
$\gamma_2$	-0,09226	0,016559	-5,5718	0,000000

La notación refiere al siguiente modelo:

$$\begin{cases} y_t = u_t \\ \epsilon_t = \frac{u_t}{\sigma_t}, \quad \epsilon_t \sim N(0, \omega) \\ \ln(\sigma_t^2) = \omega + \sum_{i=1}^2 \{ \alpha_i (|\epsilon_{t-i}| - \mathbb{E}[|\epsilon_{t-i}|]) + \gamma_i \epsilon_{t-i} \} + \beta_1 \ln(\sigma_{t-1}^2) \end{cases}$$

## A. Autocorrelogramas y función de autocorrelación parcial

FIGURA 6.1. *Autocorrelograma de la serie de retornos.*

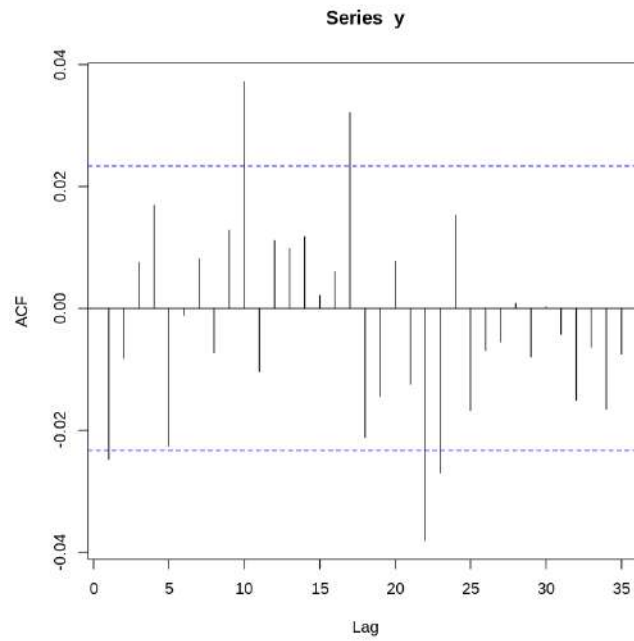


FIGURA 6.2. *Función de autocorrelación parcial de la serie de retornos.*

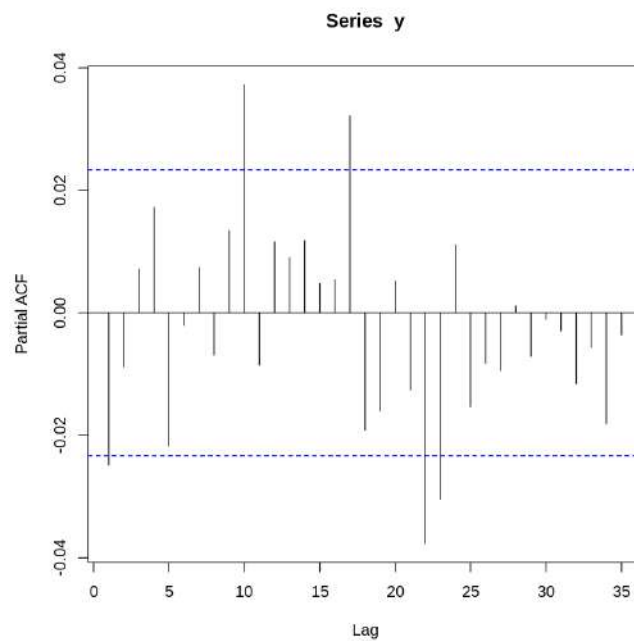
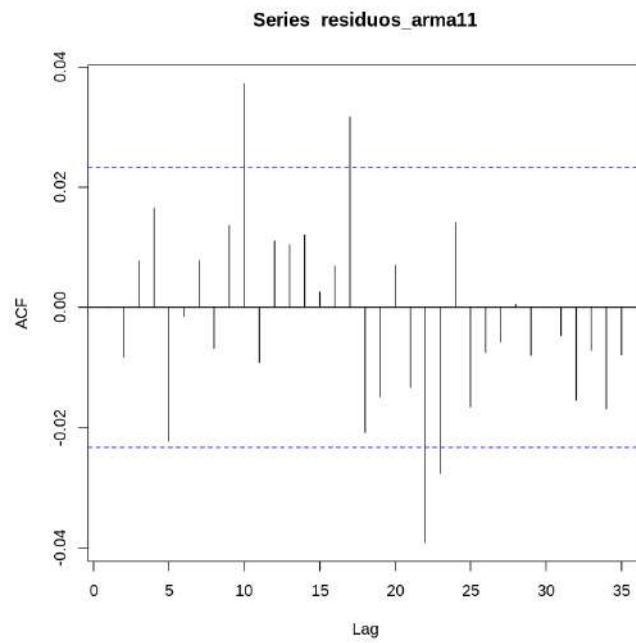
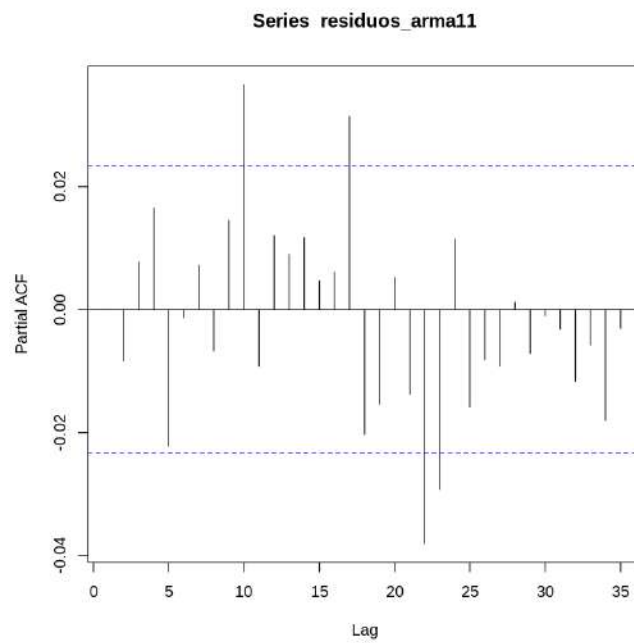


FIGURA 6.3. Autocorrelograma de los residuos del modelo  $ARMA(1,1)$ .FIGURA 6.4. Función de autocorrelación parcial de los residuos del modelo  $ARMA(1,1)$ .



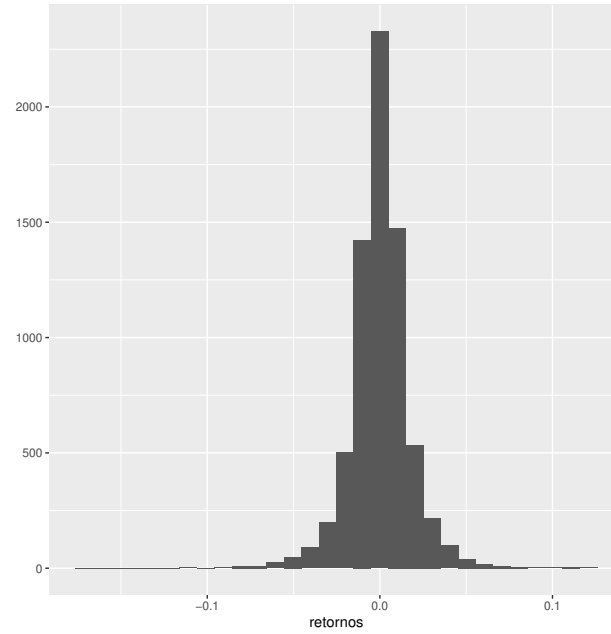
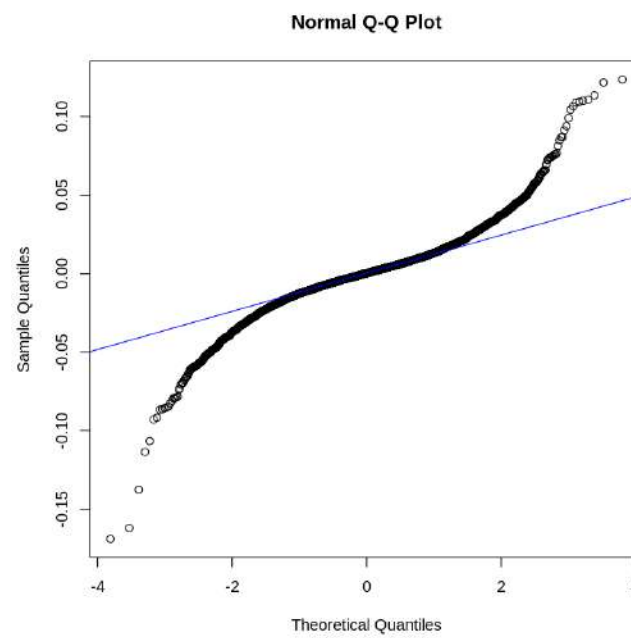
**B. Histogramas y gráficos Qq**FIGURA 6.5. *Histograma de los retornos.*FIGURA 6.6. *Gráfico Qq de los retornos con los cuantiles teóricos de una normal.*

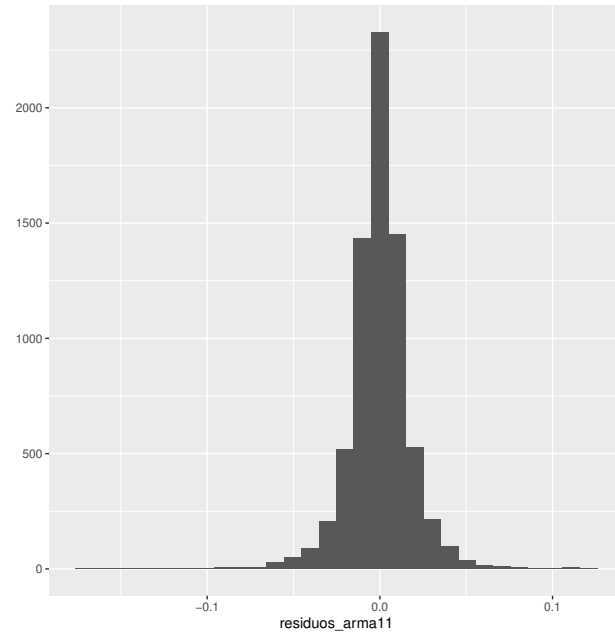
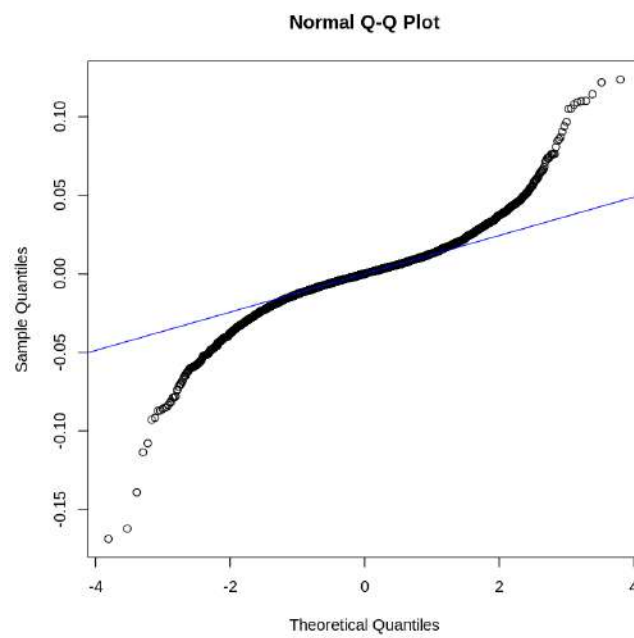
FIGURA 6.7. *Histograma de los residuos del ARMA(1,1).*FIGURA 6.8. *Gráfico Qq de los residuos del ARMA(1,1) con los cuantiles teóricos de una normal.*

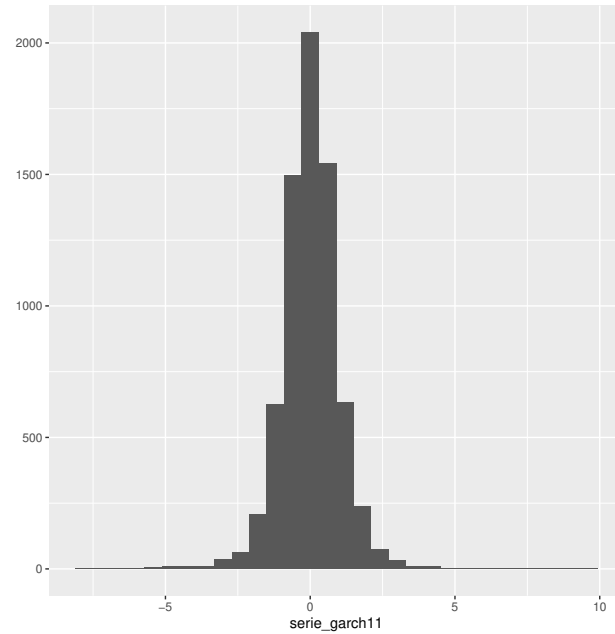
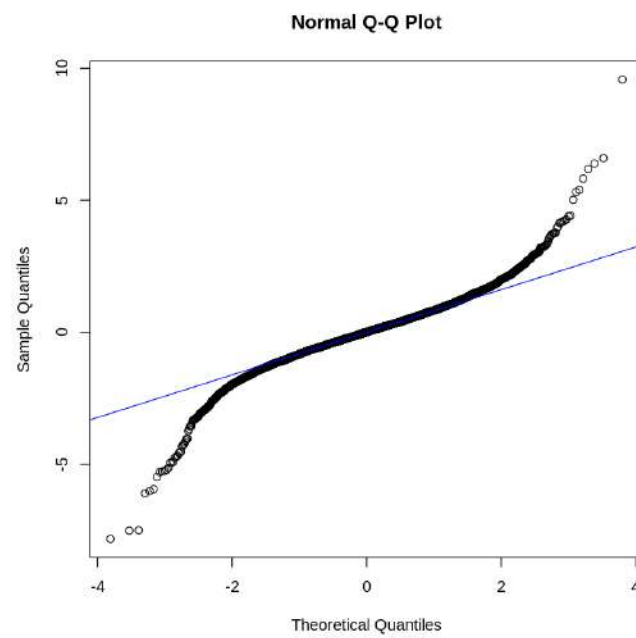
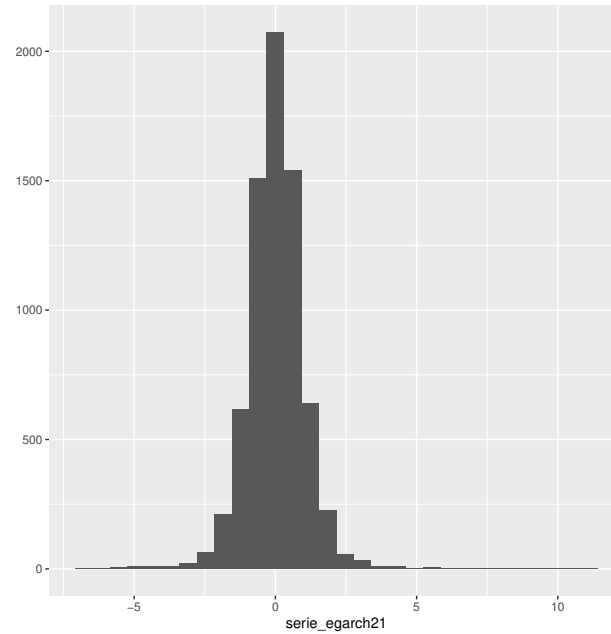
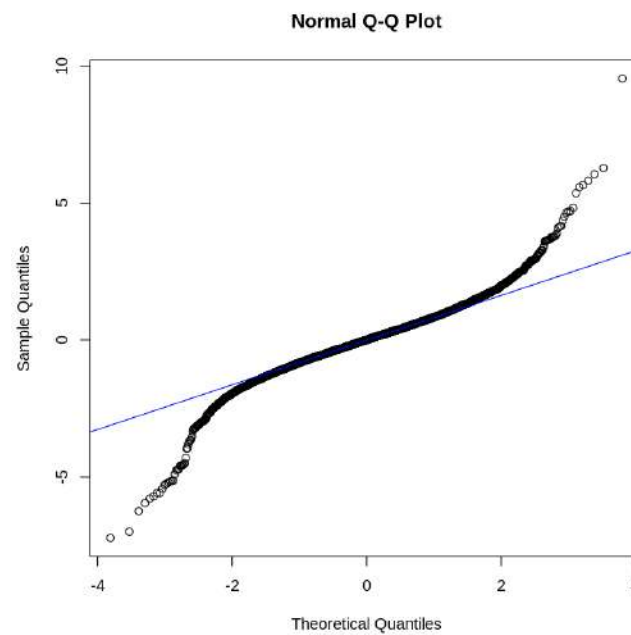
FIGURA 6.9. *Histograma de los residuos del GARCH(1,1).*FIGURA 6.10. *Gráfico Qq de los residuos del GARCH(1,1) con los cuantiles teóricos de una normal.*

FIGURA 6.11. *Histograma de los residuos del EGARCH(2,1).*FIGURA 6.12. *Gráfico Qq de los residuos del EGARCH(2,1) con los cuantiles teóricos de una normal.*

### C. Resultados del test de heterocedasticidad Breusch-Godfrey

TABLA 8. Resultados del contraste Breusch-Godfrey.

Orden	Retornos	ARMA(1,1)	GARCH(1,1)	EGARCH(2,1)
1	(4,3403; 0,03722)**	(9,44E-07; 0,9992)	(0,011067; 0,9162)	(0,001433; 0,9698)
2	(4,8948; 0,08652)**	(0,49339; 0,7814)	(0,04402; 0,9782)	(0,039393; 0,9805)
3	(5,2492; 0,1544)	(0,91472; 0,8219)	(0,36051; 0,9483)	(0,80353; 0,8486)
4	(7,3549; 0,1183)	(2,846; 0,5839)	(0,81091; 0,937)	(1,294; 0,8624)
5	(10,714; 0,05736)*	(6,3506; 0,2736)	(3,1165; 0,6820)	(3,4842; 0,6258)
6	(10,74; 0,09676)*	(6,3607; 0,3840)	(3,256; 0,7761)	(3,7453; 0,7111)
7	(11,102; 0,1342)	(6,7095; 0,4597)	(3,6024; 0,8242)	(4,4427; 0,7275)
8	(11,442; 0,1779)	(7,0325; 0,5331)	(4,3187; 0,8273)	(4,7402; 0,7848)

**Nota:** El primer término indica el valor del estadístico, el segundo el p-valor asociado.  
(\*) y (\*\*) indican, respectivamente, significatividad al 10 % y 5 % de nivel de significación.

#### Ra.1. Resultados Test BDS

TABLA 9. Test BDS para los retornos

Test BDS para los retornos				
m/ε	0,5σ	σ	1,5σ	2σ
2	(18,67; 0)***	(18,91; 0)***	(18,24; 0)***	(17,11; 0)***
3	(26,19; 0)***	(25,21; 0)***	(22,99; 0)***	(20,42; 0)***
4	(32,75; 0)***	(29,82; 0)***	(25,95; 0)***	(22,55; 0)***
5	(41,34; 0)***	(34,95; 0)***	(28,91; 0)***	(24,39; 0)***
6	(51,58; 0)***	(40,34; 0)***	(31,46; 0)***	(25,63; 0)***
7	(65,00; 0)***	(46,44; 0)***	(34,00; 0)***	(26,83; 0)***
8	(81,47; 0)***	(53,09; 0)***	(36,51; 0)***	(27,88; 0)***
9	(103,93; 0)***	(61,16; 0)***	(39,17; 0)***	(28,87; 0)***
10	(135,83; 0)***	(70,86; 0)***	(42,03; 0)***	(29,91; 0)***

**Nota:** El primer término indica el valor del estadístico, el segundo el p-valor asociado(\*), (\*\*) y (\*\*\*) indican, respectivamente, significatividad al 10%, 5% y 1% de nivel de significación.

TABLA 10. Test BDS para el modelo ARMA(1,1)

Test BDS para el modelo ARMA(1,1)				
m/ε	0,5σ	σ	1,5σ	2σ
2	(18,37; 0)***	(18,80; 0)***	(18,15; 0)***	(16,94; 0)***
3	(25,84; 0)***	(25,14; 0)***	(22,83; 0)***	(20,21; 0)***
4	(32,36; 0)***	(29,66; 0)***	(25,79; 0)***	(22,34; 0)***
5	(40,83; 0)***	(34,76; 0)***	(28,76; 0)***	(24,18; 0)***
6	(50,86; 0)***	(40,11; 0)***	(31,29; 0)***	(25,43; 0)***
7	(64,02; 0)***	(46,16; 0)***	(33,83; 0)***	(26,61; 0)***
8	(80,18; 0)***	(52,77; 0)***	(36,33; 0)***	(27,65; 0)***
9	(102,44; 0)***	(60,79; 0)***	(38,97; 0)***	(28,64; 0)***
10	(133,25; 0)***	(70,45; 0)***	(41,84; 0)***	(29,67; 0)***

**Nota:** El primer término indica el valor del estadístico, el segundo el p-valor asociado. (\*), (\*\*) y (\*\*\*) indican, respectivamente, significatividad al 10%, 5% y 1% de nivel de significación.

TABLA 11. Test BDS para el modelo GARCH(1,1)

Test BDS para el modelo GARCH(1,1)				
m/ε	0,5σ	σ	1,5σ	2σ
2	(2,72; 0,0067)***	(2,75; 0,0123)**	(2,15; 0,0319)**	(1,43; 0,1538)
3	(3,62; 0,0003)***	(3,54; 0,0004)***	(3,07; 0,0021)***	(2,12; 0,0342)**
4	(4,04; 0,0001)***	(3,78; 0,0002)***	(3,02; 0,0025)***	(1,82; 0,0690)*
5	(4,91; 0)***	(4,58; 0)***	(3,58; 0,0003)***	(2,07; 0,0384)**
6	(5,39; 0)***	(5,93; 0)***	(3,77; 0,0002)***	(2,05; 0,0402)**
7	(6,02; 0)***	(5,34; 0)***	(3,89; 0,0001)***	(1,94; 0,0517)*
8	(5,83; 0)***	(5,43; 0)***	(3,89; 0,0001)***	(1,84; 0,0664)*
9	(5,62; 0)***	(5,59; 0)***	(3,92; 0,0001)***	(1,73; 0,0832)*
10	(6,60; 0)***	(5,86; 0)***	(3,98; 0,0001)***	(1,64; 0,1002)

**Nota:** El primer término indica el valor del estadístico, el segundo el p-valor asociado. (\*), (\*\*) y (\*\*\*) indican, respectivamente, significatividad al 10%, 5% y 1% de nivel de significación.

TABLA 12. Test BDS para el modelo EGARCH(2,1)

Test BDS para el modelo EGARCH(2,1)				
m/ε	0,5σ	σ	1,5σ	2σ
2	(-0,33; 0,7454)	(-0,13; 0,8934)	(0,08; 0,9427)	(0,14; 0,8922)
3	(0,79; 0,4309)	(1,17; 0,2411)	(1,47; 0,1405)	(1,51; 0,1299)
4	(1,24; 0,2142)	(1,62; 0,1052)	(1,82; 0,0692)*	(1,76; 0,0791)*
5	(2,10; 0,0355)**	(2,56; 0,0105)**	(2,64; 0,0084)***	(2,38; 0,0175)**
6	(2,53; 0,0113)**	(3,05; 0,0023)***	(3,07; 0,0021)***	(2,66; 0,0079)***
7	(3,20; 0,0014)***	(3,60; 0,0003)***	(3,39; 0,0007)***	(2,72; 0,0064)***
8	(3,08; 0,0020)***	(3,82; 0,0001)***	(3,56; 0,0004)***	(2,76; 0,0058)***
9	(2,82; 0,0047)***	(4,09; 0)***	(3,76; 0,0002)***	(2,80; 0,0051)***
10	(3,60; 0,0003)***	(4,43; 0)***	(3,98; 0,0001)***	(2,86; 0,0043)***

**Nota:** El primer término indica el valor del estadístico, el segundo el p-valor asociado. (\*), (\*\*) y (\*\*\*) indican, respectivamente, significatividad al 10%, 5% y 1% de nivel de significación.

## Ra.2. Resultados de los tests de linealidad (excepto BDS)

TABLA 13. Contrastes de linealidad excepto BDS

Contrastes de linealidad excepto BDS				
Test	Retornos	ARMA(1,1)	GARCH(1,1)	EGARCH(1,1)
Keenan <sup>a</sup>	(6,37; 0,0116)**	(5,21; 0,0224)**	(0,43; 0,5103)	(0,09; 0,7646)
Tsay	(5,37; 0,0204)**	(5,73; 0,0166)**	(1,00; 0,3175)	(0,43; 0,5101)
White	(9,24; 0,0098)***	(12,30; 0,0021)***	(0,59; 0,744)	(0,78; 0,6785)
Teräsvirta	(7,41; 0,0246)**	(7,78; 0,0204)**	(2,01; 0,366)	(1,69; 0,4296)
RESET	(0,04; 0,9613)	(2,17; 0,1144)	(0,22; 0,8059)	(3,36; 0,0347)**

**Nota:** El primer término indica el valor del estadístico, el segundo el p-valor asociado. (\*), (\*\*) y (\*\*\*) indican, respectivamente, significatividad al 10%, 5% y 1% de nivel de significación. <sup>a</sup>El orden elegido es 1.

### Rb.1. Gráficos de recurrencia

FIGURA 6.13. *Gráfico de recurrencia de los retornos ( $f_{Nonlinear}$ ,  $m=1$ ,  $d=1$ ,  $eps=0,01$ )*

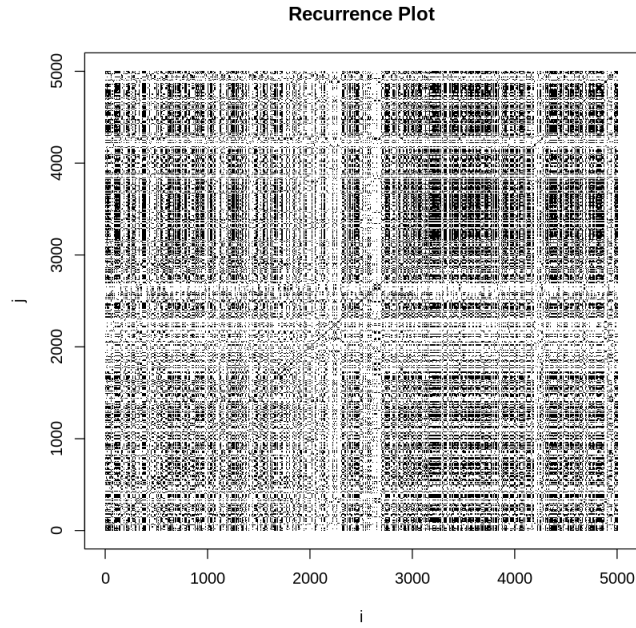


FIGURA 6.14. *Gráfico de recurrencia del modelo ARMA(1,1) ( $f_{Nonlinear}$ ,  $m=1$ ,  $d=1$ ,  $eps=0,01$ )*

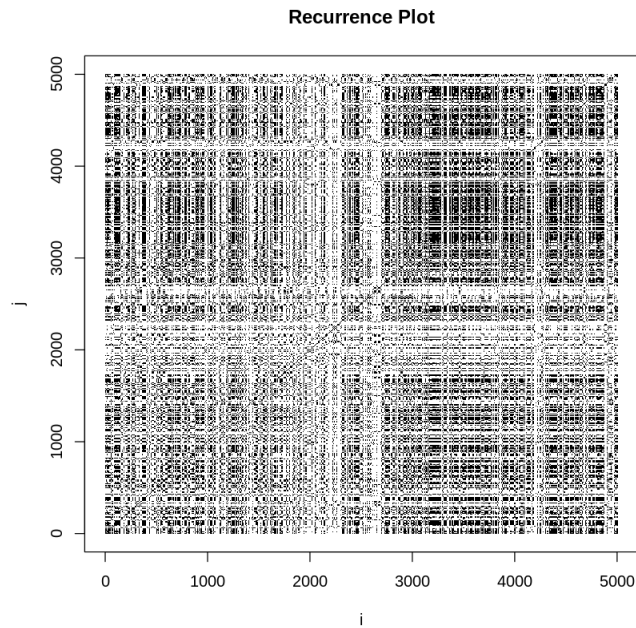




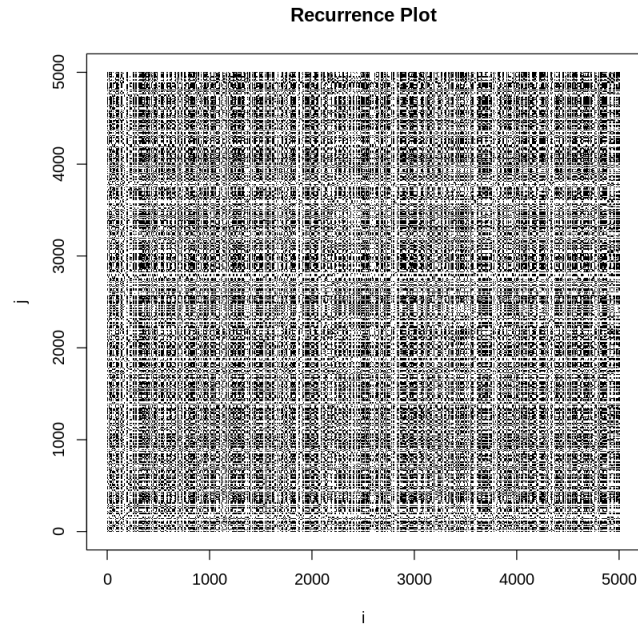
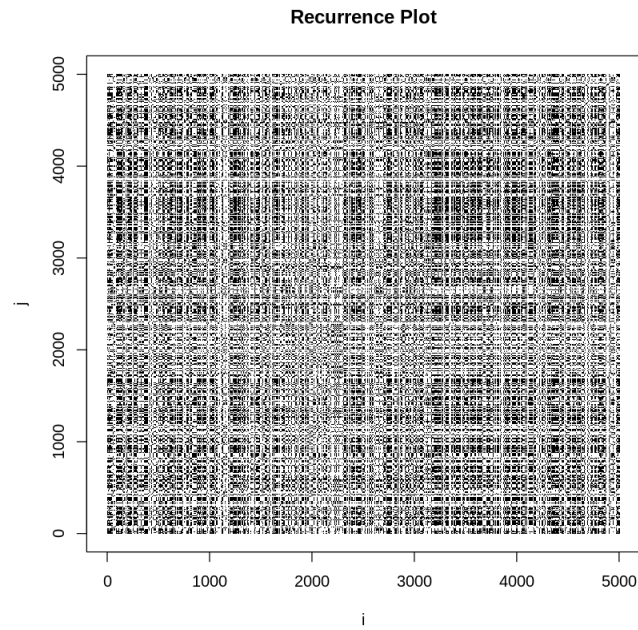
FIGURA 6.15. *Gráfico de recurrencia del modelo GARCH(1,1) (fNonlinear, m=1, d=1, eps=0,7)*FIGURA 6.16. *Gráfico de recurrencia del modelo EGARCH(2,1) (fNonlinear, m=1, d=1, eps=0,7)*

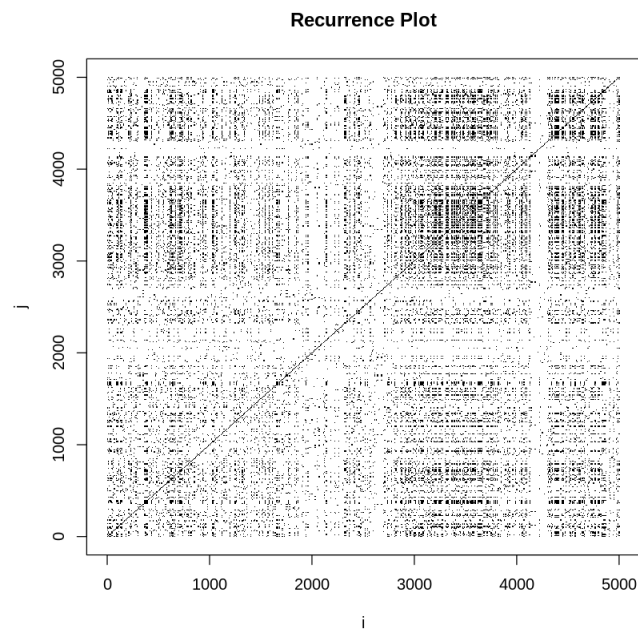
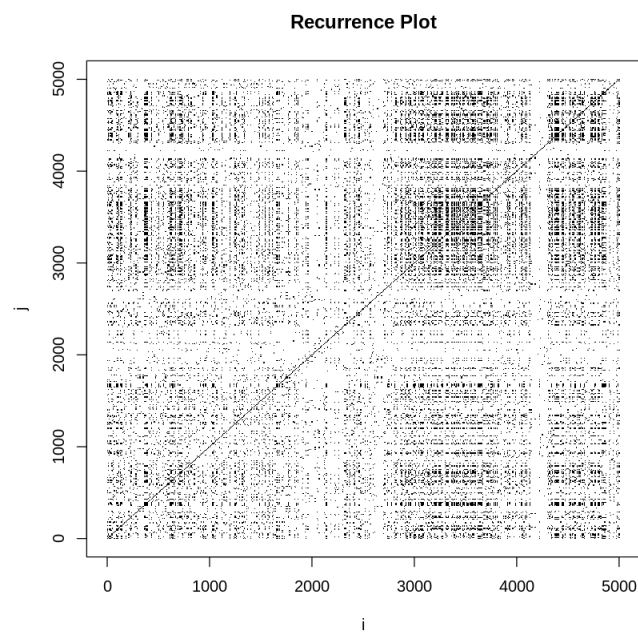
FIGURA 6.17. *Gráfico de recurrencia de los retornos ( $f_{\text{Nonlinear}}$ ,  $m=2$ ,  $d=2$ ,  $\text{eps}=0,01$ )*FIGURA 6.18. *Gráfico de recurrencia del modelo ARMA(1,1) ( $f_{\text{Nonlinear}}$ ,  $m=2$ ,  $d=2$ ,  $\text{eps}=0,01$ )*

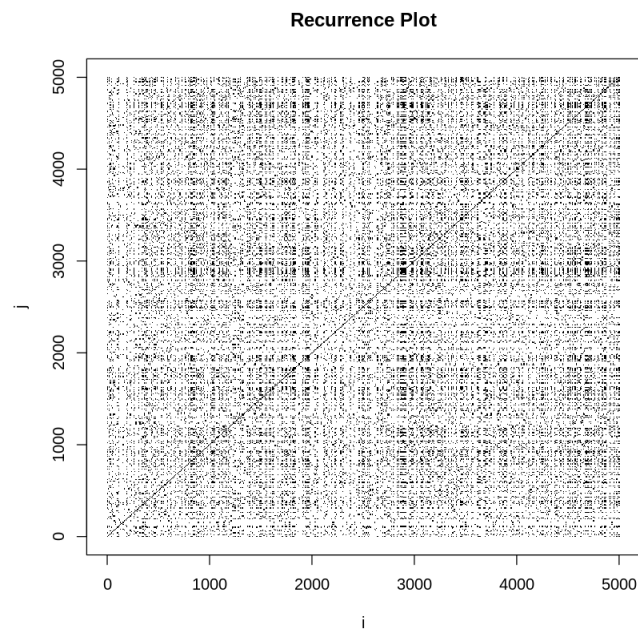
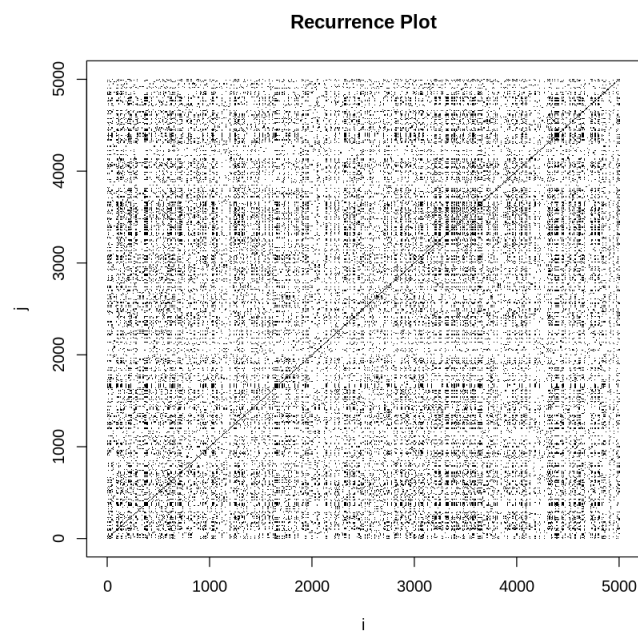
FIGURA 6.19. *Gráfico de recurrencia del modelo GARCH(1,1) (fNonlinear, m=2, d=2, eps=0,7)*FIGURA 6.20. *Gráfico de recurrencia del modelo EGARCH(2,1) (fNonlinear, m=2, d=2, eps=0,7)*

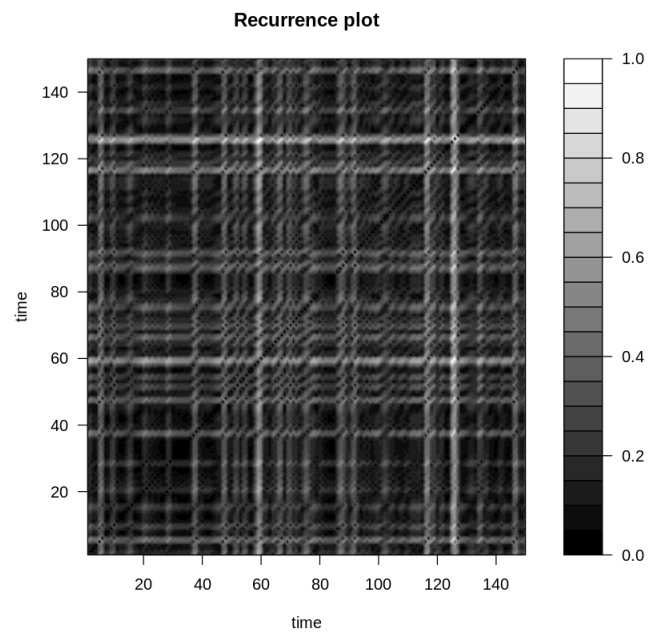
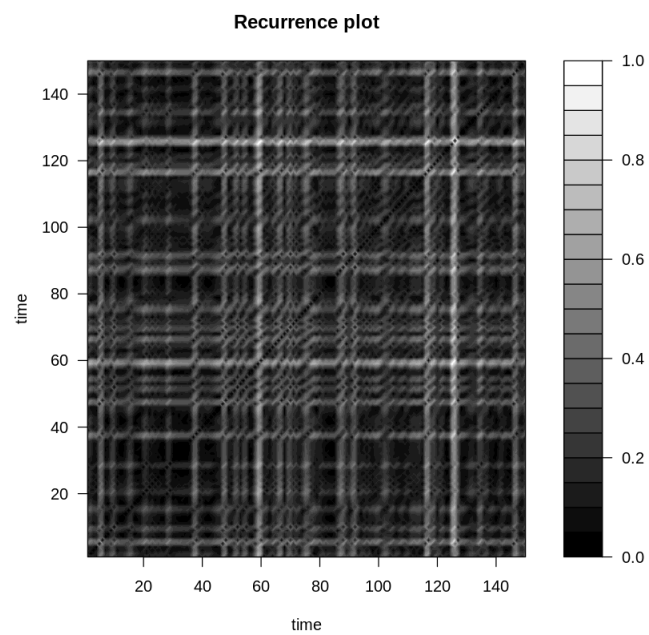
FIGURA 6.21. *Gráfico de recurrencia de los retornos ( $tseriesChaos$ ,  $m=2$ ,  $d=1$ )*FIGURA 6.22. *Gráfico de recurrencia del modelo  $ARMA(1,1)$  ( $tseriesChaos$ ,  $m=2$ ,  $d=1$ )*

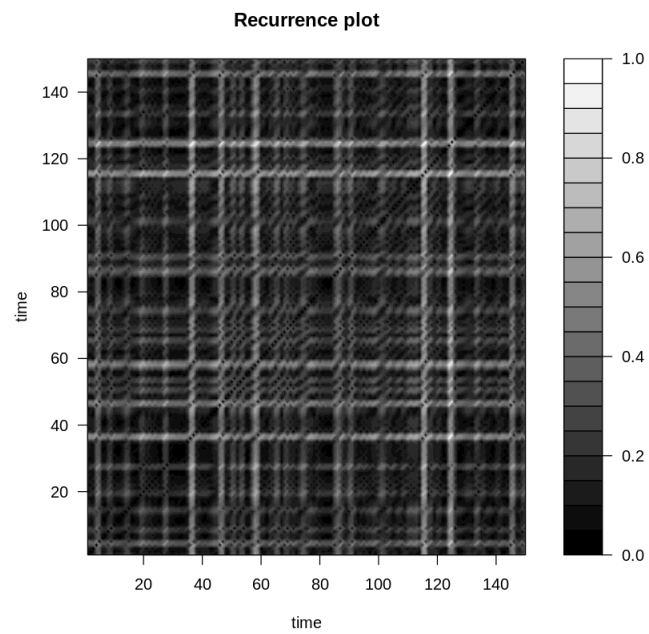
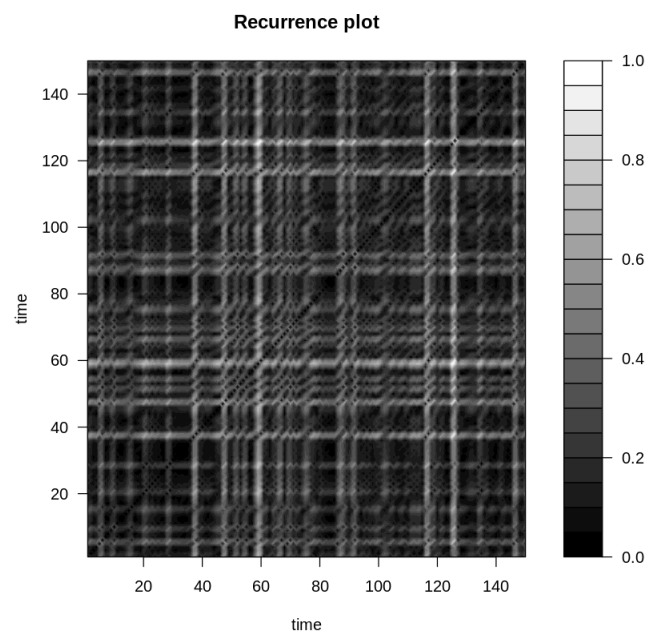
FIGURA 6.23. *Gráfico de recurrencia del modelo GARCH(1,1) (tseriesChaos, m=2, d=1)*FIGURA 6.24. *Gráfico de recurrencia del modelo EGARCH(2,1) (tseriesChaos, m=2, d=1)*

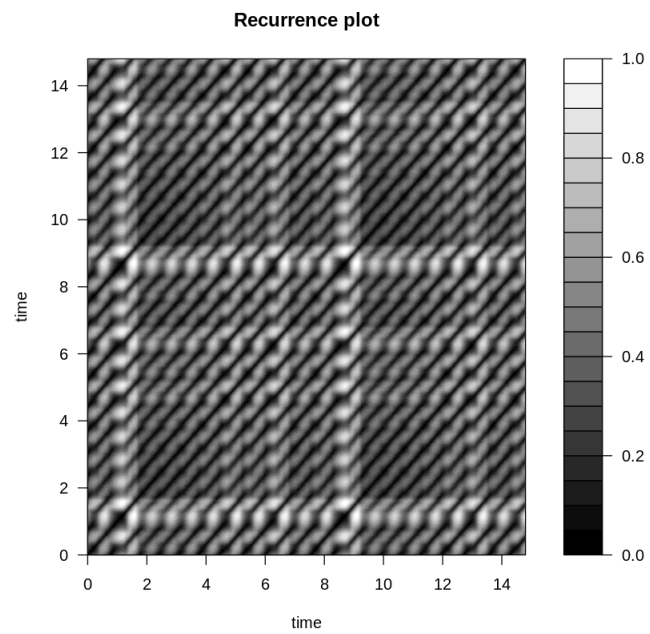
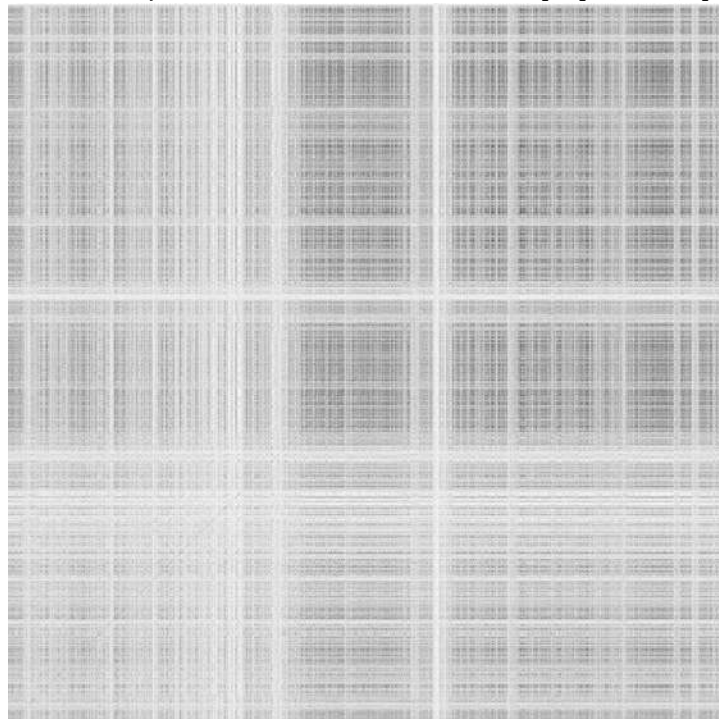
FIGURA 6.25. *Gráfico de recurrencia caótico. Sistema de ecuaciones de Lorenz. (tseriesChaos, m=2, d=1)*FIGURA 6.26. *Gráfico de recurrencia de los retornos (propio, m=1, eps=0,05)*

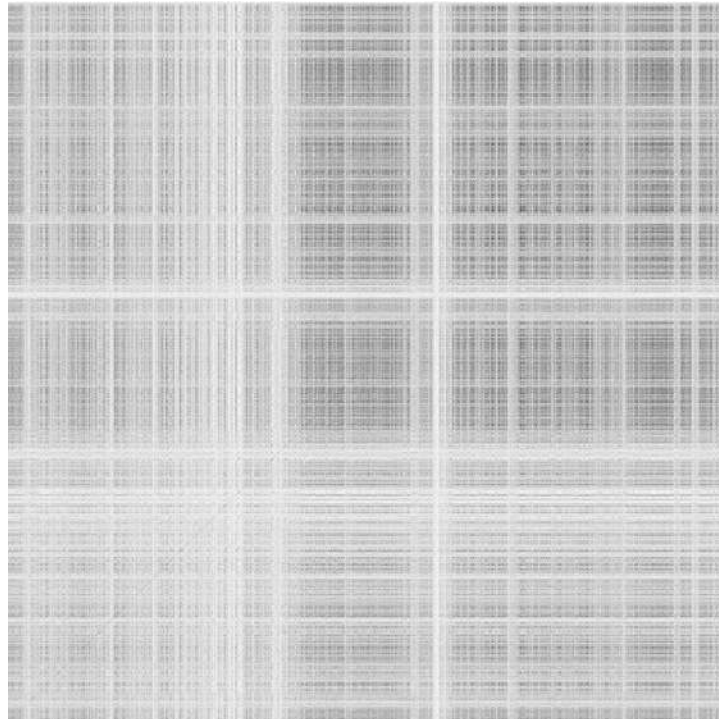
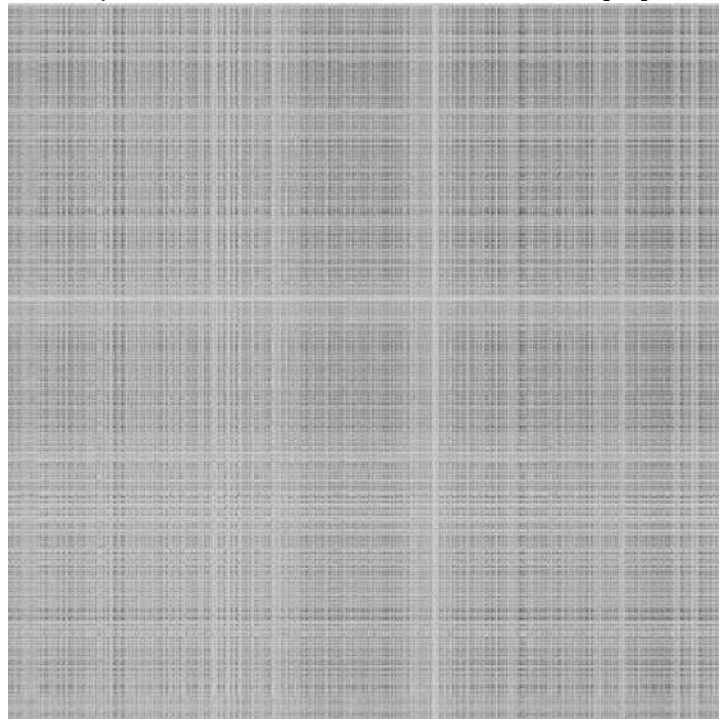
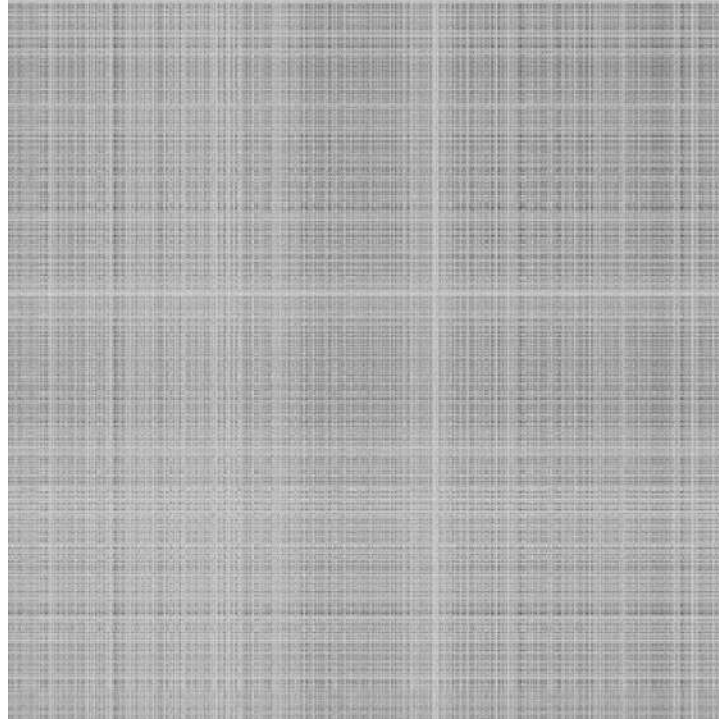
FIGURA 6.27. *Gráfico de recurrencia del modelo ARMA(1,1) (propio,  $m=1$ ,  $\epsilon=0,05$ )*FIGURA 6.28. *Gráfico de recurrencia del modelo GARCH(1,1) (propio,  $m=1$ ,  $\epsilon=0,5$ )*

FIGURA 6.29. *Gráfico de recurrencia del EGARCH(2,1) (propio,  $m=1$ ,  $\epsilon=0,5$ )*

## Rb.2. Resultados Test 0-1

TABLA 14. Resultados Test 0-1

Test 0-1	$K$
Retornos	0,9974
ARMA(1,1)	0,9978
GARCH(1,1)	0,9979
EGARCH(2,1)	0,9983

## 7B.3. Coeficientes de Hurst

TABLA 15. Exponentes de Hurst (pracma)

	R/S	R/S Corregido	Empírico	Empírico Corregido	Teórico
Retornos	0,55	0,56	0,53	0,51	0,53
ARMA(1,1)	0,55	0,56	0,53	0,51	0,53
GARCH(1,1)	0,54	0,56	0,55	0,52	0,53
EGARCH(2,1)	0,54	0,56	0,54	0,51	0,53



TABLA 16. Exponentes de Hurst e intervalos de confianza (liftLRD)

	Exponente	Extremo menor	Extremo mayor	Desv. Típica
Retornos	0,4598	0,3593	0,5706	0,05957
ARMA(1,1)	0,4862	0,3739	0,5920	0,06128
GARCH(1,1)	0,4963	0,4188	0,5908	0,04876
EGARCH(2,1)	0,4924	0,3975	0,5874	0,06264

#### Rb.4. Dimensión de la correlación

TABLA 17. Dimensión de la correlación (nonlinearTseries)

m	1	2	3	4	5
Retornos	1,009	1,992	3,046	3,046	4,8
ARMA(1,1)	1,000	1,989	2,968	4,037	4,782
GARCH(1,1)	1,003	1,998	3,018	3,018	4,851
EGARCH(2,1)	1,006	1,987	3,025	3,999	4,775

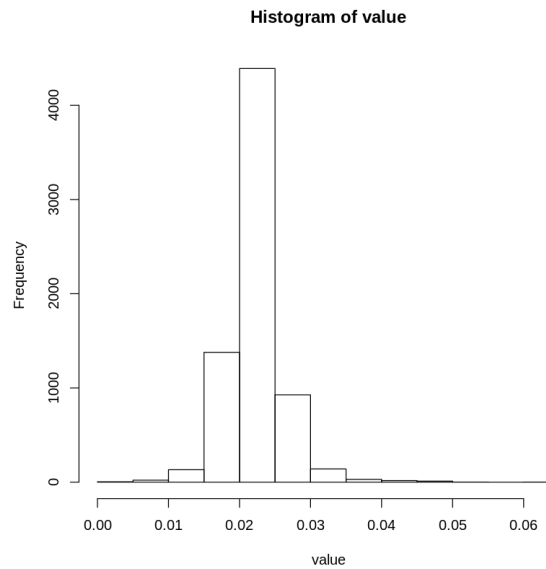
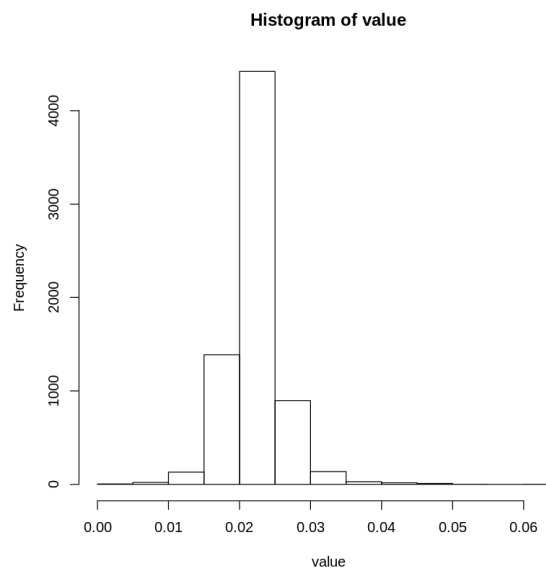
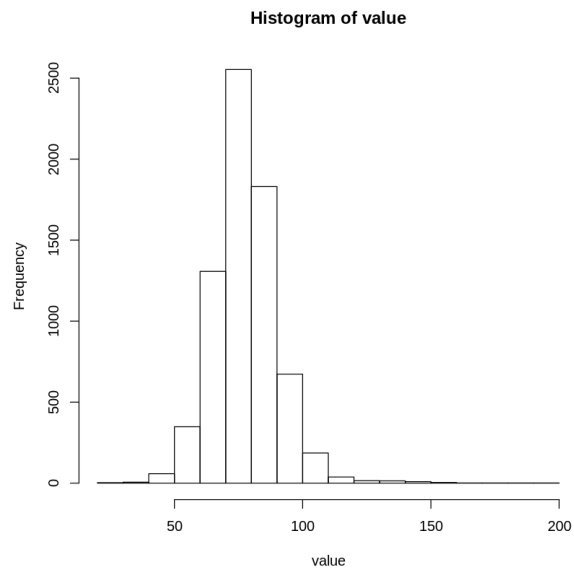
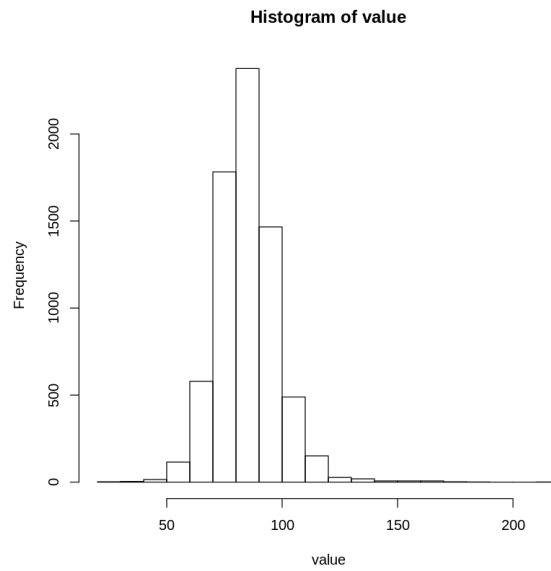
**Rb.5. Histogramas del algoritmo MESAH**FIGURA 6.30. *Histograma MESAH de los retornos*FIGURA 6.31. *Histograma MESAH del modelo ARMA(1,1)*

FIGURA 6.32. *Histograma MESA*H del modelo  $GARCH(1,1)$ FIGURA 6.33. *Histograma MESA*H del modelo  $EGARCH(2,1)$ 

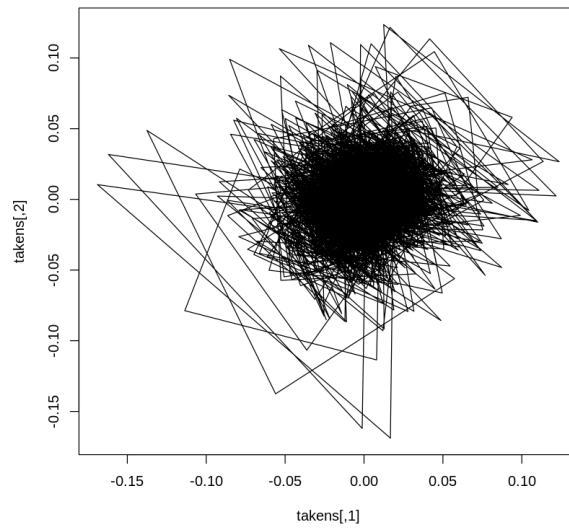
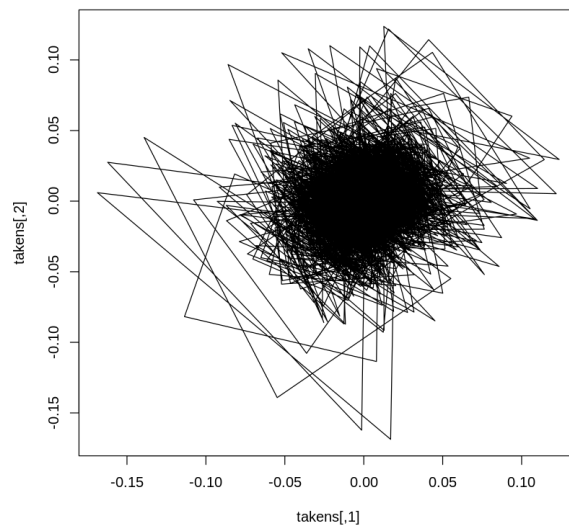
**Rb.6. Mapas de Poincaré y del espacio de fases**FIGURA 6.34. *Mapa del espacio de fases de los retornos ( $d=1$ ,  $m=2$ )*FIGURA 6.35. *Mapa del espacio de fases del modelo ARMA(1,1) ( $d=1$ ,  $m=2$ )*

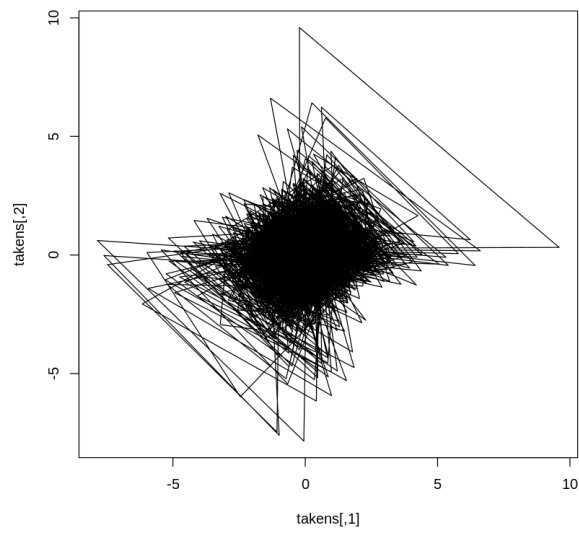
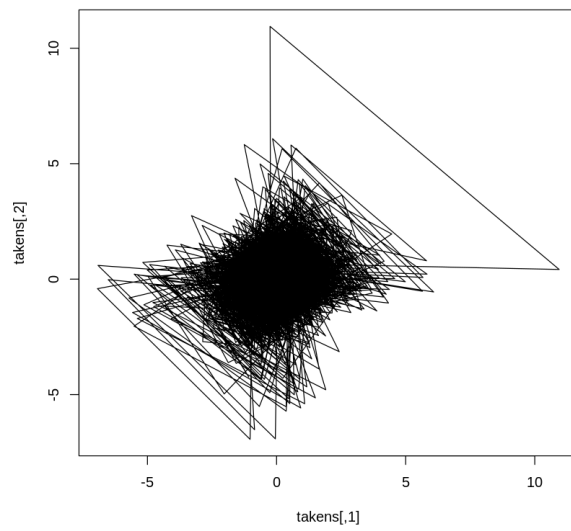
FIGURA 6.36. *Mapa del espacio de fases del modelo GARCH(1,1) ( $d=1$ ,  $m=2$ )*FIGURA 6.37. *Mapa del espacio de fases del modelo EGARCH(2,1) ( $d=1$ ,  $m=2$ )*

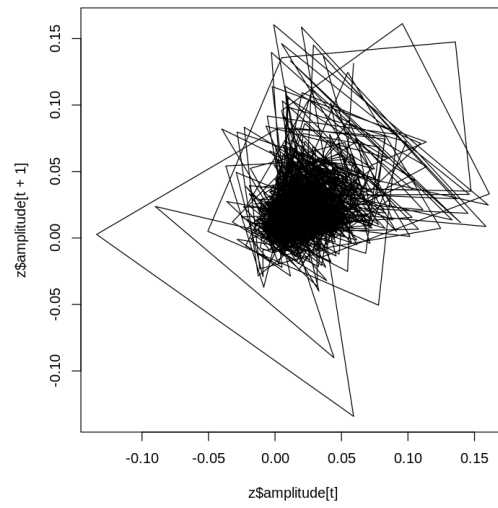
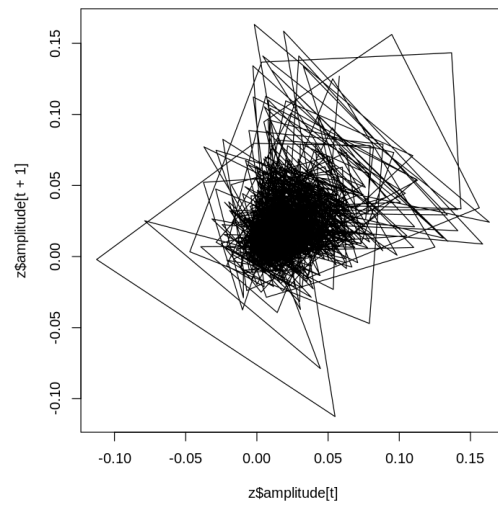
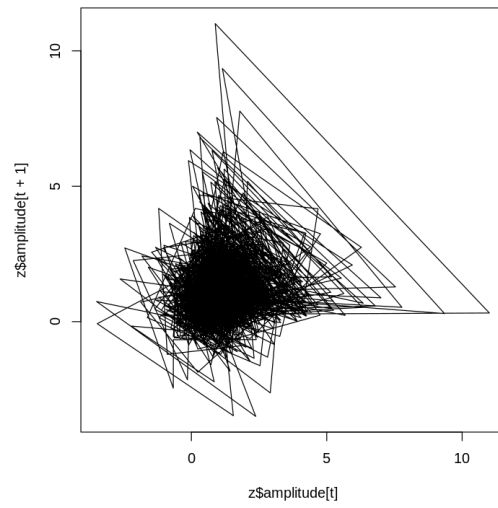
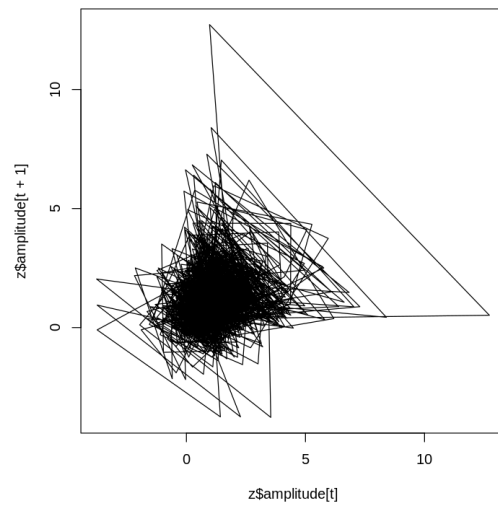
FIGURA 6.38. *Mapa de Poincaré de los retornos (nonlinearTseries)*FIGURA 6.39. *Mapa de Poincaré del modelo ARMA(1,1) (nonlinearTseries)*

FIGURA 6.40. *Mapa de Poincaré del modelo GARCH(1,1) (nonlinearTseries)*FIGURA 6.41. *Mapa de Poincaré del modelo EGARCH(2,1) (nonlinearTseries)*

**Rb.7. Resultados de la técnica de valoración del ruido**

TABLA 18. Resultados de la técnica de valoración del ruido

factor	0,0	0,1	0,4	0,8	1,1	1,5	1,8
Retornos	0,00	0,02	0,04	0,16	0,92	0,95	0,90
ARMA(1,1)	0,02	0,02	0,02	0,09	0,26	0,62	0,87
GARCH(1,1) <sup>a</sup>	0,84	No se detecta no linealidad					
EGARCH(2,1)	0,51	No se detecta no linealidad					

*Resultados del test de Tsay al añadir ruido gaussiano de desviación típica factor \*  $\sigma$  a las series.  $\sigma$  denota la d.t. de cada serie. Los elementos de la tabla son los p-valores.*

<sup>a</sup>No se rechaza  $H_0$  ya para la serie original.

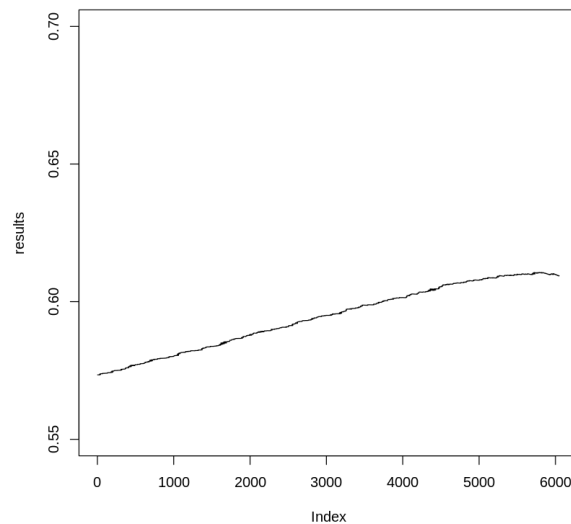
**Rb.8. Resultados del test de los tres estados**FIGURA 6.42. Convergencia a  $K$  en el test de los tres estados para los retornos



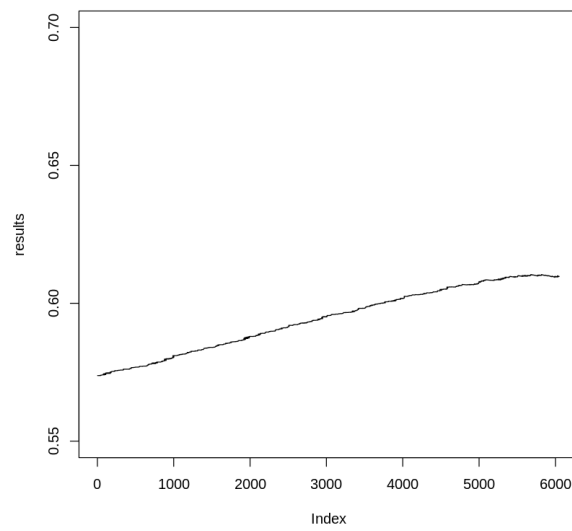
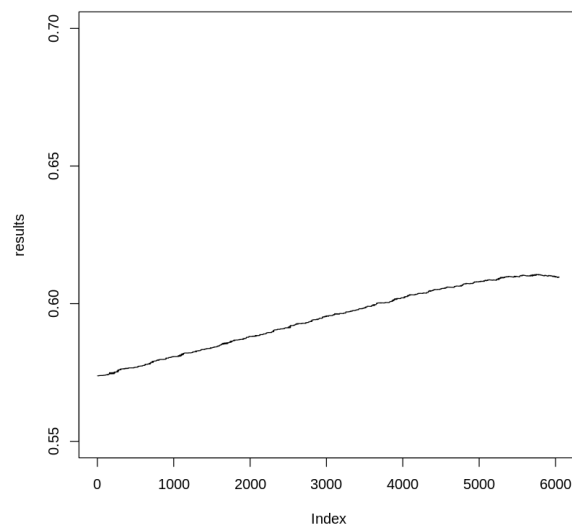
FIGURA 6.43. *Convergencia a  $K$  en el test de los tres estados para el modelo ARMA(1,1)*FIGURA 6.44. *Convergencia a  $K$  en el test de los tres estados para el modelo GARCH(1,1)*

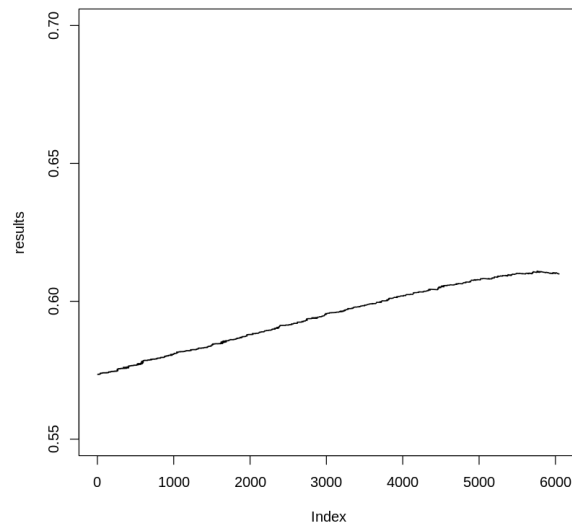
FIGURA 6.45. *Convergencia a  $K$  en el test de los tres estados para el modelo EGARCH(2,1)***Rb.9. Resultados del test de Matilla y Ruiz**

TABLA 19. Resultados del test de Matilla y Ruiz

	Constante ( $\alpha_0$ )	Pendiente ( $\alpha_1$ )	Estadístico T	P-valor
Retornos	0,1588 (0,0505)	0,4023 (0,0126)	31,8696	$2,98E-13$
ARMA(1,1)	0,2046 (0,0518)	0,3926 (0,0146)	26,8812	$2,15E-12$
GARCH(1,1) <sup>a</sup>	0,2397 (0,0493)	0,3848 (0,0098)	39,2229	$2,43E-14$
EGARCH(2,1)	0,1566 (0,0238)	0,4016 (0,0039)	104,3275	0,0

*Resultados del test de Matilla y Ruiz basado en la entropía a las series.  
 $w=10$ ,  $m=5$ . Entre paréntesis las desviaciones típicas de cada parámetro.*

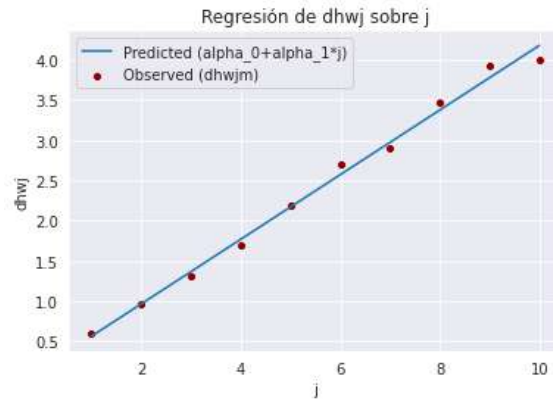
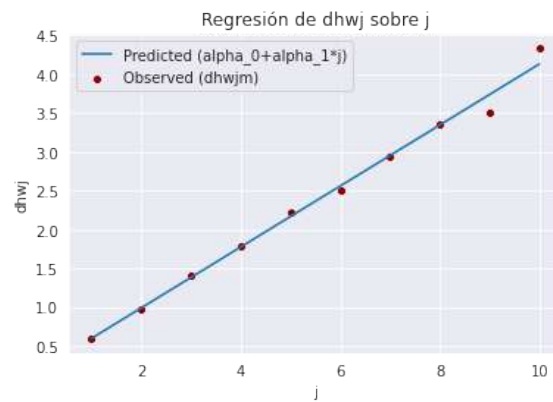
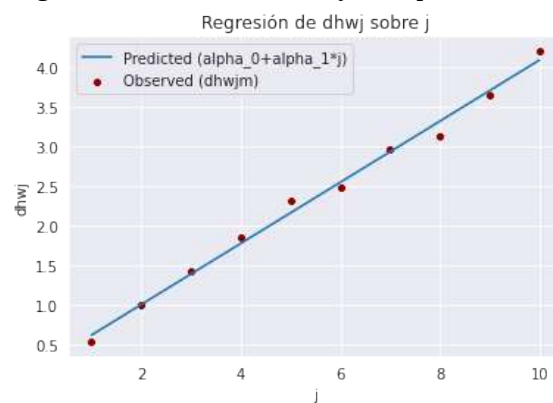
FIGURA 6.46. *Regresión del test de Matilla y Ruiz para los retornos*FIGURA 6.47. *Regresión del test de Matilla y Ruiz para el modelo ARMA(1,1)*FIGURA 6.48. *Regresión del test de Matilla y Ruiz para el modelo GARCH(1,1)*

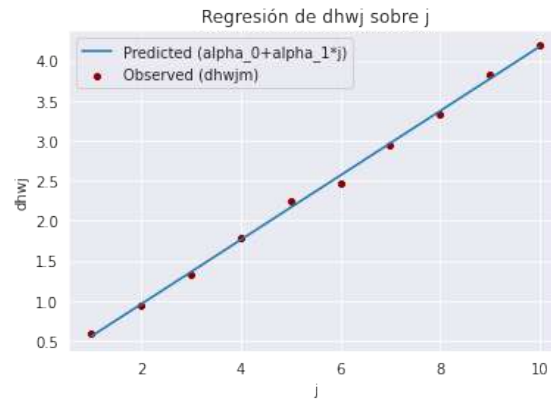
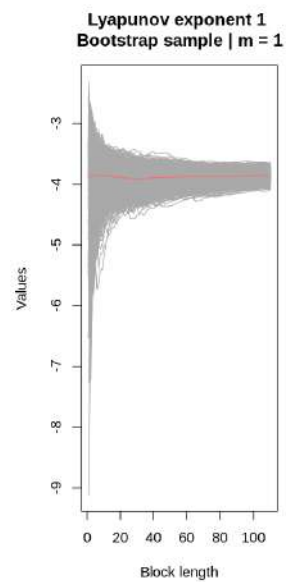
FIGURA 6.49. *Regresión del test de Matilla y Ruiz para el modelo EGARCH(2,1)***Rb.10. Exponentes de Lyapunov estimados mediante redes neuronales.**FIGURA 6.50. *Exponente de Lyapunov estimados para los retornos (DChaos)*

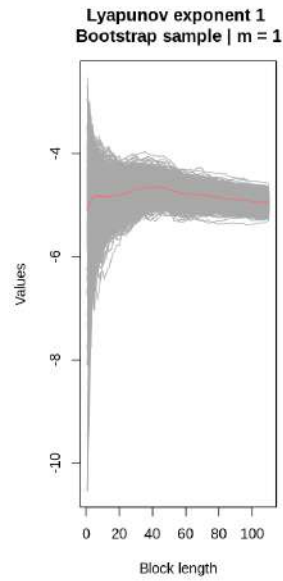
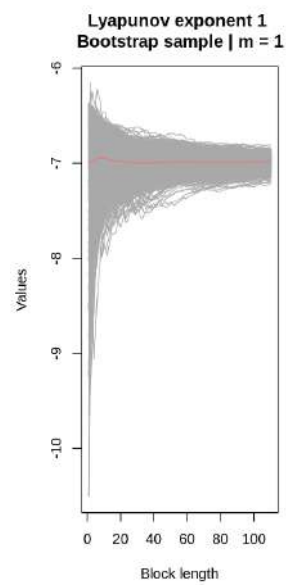
FIGURA 6.51. *Exponente de Lyapunov estimados para el modelo ARMA(1,1) (DChaos)*FIGURA 6.52. *Exponente de Lyapunov estimados para el modelo GARCH(1,1) (DChaos)*

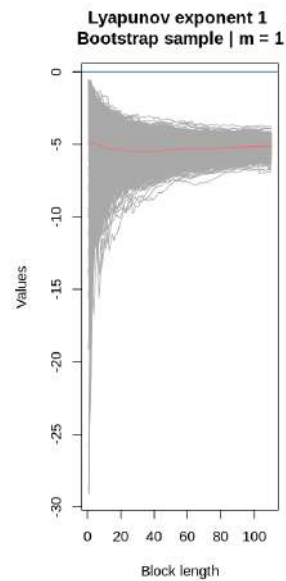
FIGURA 6.53. *Exponente de Lyapunov estimados para el modelo EGARCH(2,1) (DChaos)***Rb.11. Resultados del árbol de decisión de procesos caóticos.**

TABLA 20. Resultados del árbol de decisión de procesos caóticos

Resultado	
Retornos	Estocástico
ARMA(1,1)	Estocástico
GARCH(1,1)	Estocástico
EGARCH(2,1)	Estocástico

*Resultados del árbol de decisión a través de la rutina implementada en Matlab por los propios autores.*

La rutina está disponible para su [descarga](https://figshare.com/articles/The_Chaos_Decision_Tree_Algorithm/7476362/7) en [https://figshare.com/articles/The\\_Chaos\\_Decision\\_Tree\\_Algorithm/7476362/7](https://figshare.com/articles/The_Chaos_Decision_Tree_Algorithm/7476362/7)