



Procesos estocásticos irregularmente espaciados: Una variación del modelo autorregresivo irregular de primer orden IAR

Wilmar Sepulveda Herrera

Universidad del valle
Facultad de Ingeniería, Escuela de Estadística
Cali, Valle del Cauca
2023

Procesos estocásticos irregularmente espaciados

by

Wilmar Sepulveda Herrera

Trabajo de investigación para optar por el título de:

Magister en

En

Estadística

Aceptamos esta tesis
Conforme a los requerimientos de la norma

Universidad del valle
Facultad de Ingeniería, Escuela de Estadística
Cali, Valle del Cauca
2023

Dedicatoria

A mi familia.

Agradecimientos

Estoy agradecido con...

Abstract

El resumen es una presentación abreviada y precisa. Se debe usar una extensión máxima de 12 renglones. Se recomienda que este resumen sea analítico, es decir, que sea completo, con información cuantitativa y cualitativa, generalmente incluyendo los siguientes aspectos: objetivos, diseño, lugar y circunstancias (u objetivo del estudio), principales resultados, y conclusiones. Al final del resumen se deben usar palabras claves tomadas del texto (mínimo 3 y máximo 7 palabras), las cuales permiten la recuperación de la información.

Palabras clave: palabras clave en español (máximo 10 palabras, preferiblemente seleccionadas de las listas internacionales que permitan el indizado cruzado)

Resumen

Es el mismo resumen pero traducido al inglés. Se debe usar una extensión máxima de 12 renglones. Al final del Abstract se deben traducir las anteriores palabras clave tomadas del texto (mínimo 3 y máximo 7 palabras), llamadas keywords. Es posible incluir el resumen en otro idioma diferente al español o al inglés, si se considera como importante dentro del tema tratado en la investigación, por ejemplo: un trabajo dedicado a problemas lingüísticos del mandarín seguramente estaría mejor con un resumen en mandarín.

Keywords: palabras clave en inglés (máximo 10 palabras, preferiblemente seleccionadas de las listas internacionales que permitan el indizado cruzado)

Contents

Abstract	vii
List of figures	xi
List of tables	xii
1 Introducción	1
1.1 Objetivos	2
1.1.1 Objetivo general	2
1.1.2 Objetivos específicos	2
2 Marco teórico	3
2.1 Procesos estocásticos irregularmente espaciados	3
2.2 Formas de describir un proceso estocástico	4
2.3 El proceso de medias móviles irregular de primer orden (IMA)	4
2.4 El proceso Autorregresivo de primer orden irregular (IAR)	6
2.4.1 Estimación	7
2.5 Modelo autoregresivo complejo irregular (CIAR)	8
3 Modelo autorregresivo de primer orden irregularmente espaciado	10
3.1 Construcción del modelo	10
3.2 Implementando propuesta de ϕ_n	12
3.3 El modelo autorregresivo irregularmente espaciado de primer orden	12
3.4 Propiedades del proceso	13
3.5 Pronóstico en el modelo IAR	13
3.6 Estimación	14
3.7 Estimación Bootstrap	15
3.8 Estudio de Simulación Monte Carlo	15

3.9 Comparación del Modelo IAR vs el CIAR	20
3.10 Aplicación en datos astronómicos	21
3.11 Aplicación en datos de espeleotemas ¹	21
4 Conclusion	24
A Appendix	25
References	26

¹Los espeleotemas son formaciones geológicas que se encuentran en cuevas y se forman a partir de la acumulación de minerales depositados por el agua que gotea del techo o fluye por las paredes de la cueva. Algunos ejemplos de espeleotemas incluyen estalactitas, estalagmitas, columnas y cortinas de piedra caliza.

List of Figures

3-1	simulación del modelo IAR con coeficientes positivos	16
3-2	simulación del modelo IAR con coeficientes negativos	17
3-3	Esquema general del estudio de Monte Carlo.	18
3-4	Esquema general del estudio de Monte Carlo.	19
3-5	Esquema general del estudio de Monte Carlo.	19
3-6	Comportamiento de la simulación, tomando como base el IAR.	20
3-7	Comportamiento de la simulación, tomando como base el CIAR.	20
3-8	Comportamiento de la simulación, tomando como base el IAR Valor de ϕ estimado.	21
3-9	Comportamiento de la simulación, tomando como base el CIAR Valor de ϕ estimado.	21
3-10	Comportamiento de la serie de tiempo.	22
3-11	Comportamiento del Δ_n de la serie.	22
3-12	Comportamiento de la serie de tiempo.	22
3-13	Comportamiento del Δ_n de la serie.	22
3-14	Comportamiento del modelo estimado con la serie original.	23
3-15	Comportamiento del modelo estimado con la serie diferenciada.	23

List of Tables

3-1 Resultados del estudio de Monte Carlo 19

Chapter 1

Introducción

Los procesos estocásticos, y en particular el análisis de series de tiempo, son metodologías estadísticas ampliamente utilizadas en la actualidad. Estas técnicas se aplican comúnmente en el pronóstico de valores futuros y en el estudio de las características de un fenómeno, incluyendo su estructura de autocorrelación. La función de autocorrelación puede ser positiva, lo que es importante en muchos contextos, como la demografía (por ejemplo, Novoseltseva et al. (2019)) o la explicación de fenómenos ecológicos (por ejemplo, Yang et al. (2019)). También pueden darse autocorrelaciones negativas, que son relevantes en el sector financiero para entender el comportamiento de los rendimientos de acciones (por ejemplo, Kuttu & Bokpin (2017)) y en la biología (por ejemplo, Rindorf et al. (2020)). En este contexto, los procesos estocásticos son importantes en diversos campos del conocimiento para caracterizar la dependencia temporal de un evento de interés y tomar mejores decisiones.

La teoría de series temporales es ampliamente utilizada en la actualidad para modelar la autocorrelación de una secuencia de interés. Uno de los métodos más populares es el modelo autorregresivo integrado de medias móviles (ARIMA). Sin embargo, este y otros métodos similares asumen que las observaciones son tomadas regularmente en el tiempo, lo cual no siempre es el caso en la práctica. Esto puede deberse a la presencia de valores faltantes o a la naturaleza de las observaciones, lo que puede hacer que la serie de tiempo sea irregularmente espaciada (Elorrieta et al. 2019). De hecho, es común encontrar series de tiempo con este tipo de irregularidades en diversas aplicaciones, como en sistemas caóticos, medicina, imágenes satelitales y calidad del aire, entre otros (Shamsan et al. 2020, Liu et al. 2019, Ghaderpour & Vujadinovic 2020, Dilmaghani et al. 2007).

En la literatura actual, existen diversos métodos para modelar series de tiempo que no son tomadas regularmente en el tiempo. Algunos autores, como Adorf (1995), abordan

este problema transformando las series irregulares en regulares a través de la interpolación, pero este enfoque puede generar sesgos. Otros, como Robinson (1977), tratan la serie de tiempo como una realización discreta de un proceso estocástico de tiempo continuo, lo cual también presenta problemas. Recientemente, Ojeda et al. (2023) propusieron el modelo de medias móviles irregulares (*IMA*). Bajo normalidad, este modelo es estrictamente estacionario, mientras que, considerando otros supuestos distribucionales, se comporta como un proceso débilmente estacionario.

En esta línea, Eyheramendy et al. (2018) proponen el Modelo Autorregresivo Irregular (IAR) de primer orden. Sin embargo, este modelo tiene una limitación importante: no acepta valores negativos en su función de autocorrelación (ACF). Para solucionar este problema, los autores presentaron el Modelo Autorregresivo Complejo Irregular (CIAR) en un estudio posterior (Elorrieta et al. 2019). En ambos casos, se han evaluado las propiedades estadísticas de los modelos y se han propuesto métodos de estimación que han obtenido buenos resultados. No obstante, tanto el modelo IAR como el IMA tienen una desventaja: su función de autocorrelación solo puede ser positiva, lo que limita su capacidad para detectar autocorrelaciones negativas, que suelen ser importantes en la práctica. Este problema no está presente en los modelos ARIMA clásicos. Por lo tanto, el objetivo de este trabajo es modificar el modelo IAR para que pueda modelar tanto autocorrelaciones positivas como negativas sin necesidad de recurrir al plano complejo.

1.1 Objetivos

1.1.1 Objetivo general

Desarrollar un modelo autorregresivo irregular de primer orden capaz de modelar estructuras de autocorrelación subyacentes con valores tanto positivos como negativos.

1.1.2 Objetivos específicos

- Encontrar estimadores por máxima verosimilitud y bootstrap para el modelo propuesto.
- Evaluar las propiedades del modelo propuesto y sus estimaciones vía simulación.
- Aplicar el modelo propuesto a un conjunto de datos reales y evaluar su capacidad para capturar la dinámica de la serie de tiempo y realizar predicciones precisas.

Chapter 2

Marco teórico

Para el tratamiento de procesos estocásticos y series de tiempo, ha sido mucha la teoría estadística implementada y se ha convertido en una herramienta importante para tratar observaciones dependientes y entender la naturaleza de dicha dependencia, la mayoría de estos métodos, asumen que las series de tiempo son regularmente espaciadas, sin embargo este escenario no siempre se da y ahí es donde cobran importancia los procesos estocásticos irregularmente espaciados.

Este capítulo tiene la siguiente estructura, en la sección 2.1 presentamos la definición matemática de un proceso estocástico irregularmente espaciado. En la sección 2.2 se presenta un proceso estocástico autorregresivo irregularmente espaciado. En la sección 2.3 se presenta un proceso estocástico irregularmente espaciado de medias móviles y en la sección 2.3 se presenta un proceso estocástico autorregresivo de medias móviles irregularmente espaciado.

2.1 Procesos estocásticos irregularmente espaciados

Sea (Ω, \mathcal{B}, P) un espacio de probabilidad. Si se define un proceso estocástico como una medida que mapea $x : \Omega \mapsto \mathbb{R}^{\mathbb{T}}$, donde

$$x(\omega) = \{X_\tau(\omega), \tau \in \mathbb{T}\}$$

\mathbb{T} es llamado conjunto de índices y la variable aleatoria $X_\tau(\omega)$ es llamada coordenada del proceso o trayectoria. Consideremos ahora $\mathbb{T}' = \{t_1, t_2, t_3, \dots\}$, con $\Delta_{n+1} = t_{n+1} - t_n$, para $n \geq 1$

$$x' = \{X_\tau(\omega), \tau \in \mathbb{T}'\}$$

x' es un proceso estocástico irregularmente espaciado. Una serie de tiempo irregularmente espaciada es una realización finita de un proceso estocástico irregularmente espaciado. Note que si $\Delta_{n+1} = t_{n+1} - t_n = 1$ para $n \geq 1$, el proceso x' es un proceso regularmente espaciado, por tanto esta definición es más general.

2.2 Formas de describir un proceso estocástico

Una forma de describir un proceso estocástico x , es especificar la función de distribución conjunta de $\{X_{\tau_1}, X_{\tau_2}, \dots, X_{\tau_n}\}$ para todo n , este es llamado *punto de vista distribucional*. Por otra parte, se puede describir el proceso proporcionando una formula para el valor X_τ para cada punto τ en términos de una familia de variables aleatorias con comportamiento probabilístico conocido, esto hace que podamos ver el proceso como una función de otros procesos (o cómo familias de procesos iid); esta forma es llamada *punto de vista constructorista*.

2.3 El proceso de medias móviles irregular de primer orden (IMA)

Teniendo en cuenta el conjunto $\mathbb{T}' = \{t_1, t_2, t_3, \dots\}$ propuesto anteriormente, tal que sus diferencias Δ_{n+1} para $n \geq 1$ están acotados uniformemente lejos de cero. Ahora, sea $m : \mathbb{T}' \mapsto \mathbb{R}$ una función tal que $m(t_n) = 0$, para cualquier $t_n \in \mathbb{T}'$. A continuación, sea $\Gamma : \mathbb{T}' \times \mathbb{T}' \mapsto \mathbb{R}$ una función tal que para cualquier pareja $t_n, t_s \in \mathbb{T}'$,

$$\Gamma(t_n, t_s) = \begin{cases} \gamma_0 & |n - s| = 0 \\ \gamma_1, \Delta_{\max(n,s)} & |n - s| = 1 \\ 0 & |n - s| \geq 1 \end{cases}$$

Note que Γ puede ser representada cómo una matriz diagonal así

$$\Gamma = \begin{bmatrix} \gamma_0 & \gamma_1, \Delta_2 & 0 & 0 & \dots \\ \gamma_1, \Delta_2 & \gamma_0 & \gamma_1, \Delta_3 & 0 & \\ 0 & \gamma_1, \Delta_3 & \gamma_0 & \gamma_1, \Delta_4 & \\ \cdot & & & \cdot & \\ \cdot & & & & \cdot \\ \cdot & & & & & \cdot \end{bmatrix} \quad (2-1)$$

2.3. EL PROCESO DE MEDIAS MÓVILES IRREGULAR DE PRIMER ORDEN (IMA) 5

Ser Γ_n la truncación $n \times n$ de Γ y asumiendo $\gamma_1, \Delta_j \neq 0$, para $j = 2, \dots, n$, Γ_n es definida positiva si γ_0 y $(\frac{\gamma_1, \Delta_{n+1}}{\gamma_0})^2 \leq 1/4$, para $j = 2, \dots, n$

Esto implica que existe un proceso Gaussiano estacionario $\{X_{t_n}, t_n, \tau \in \mathbb{T}\}$, único hasta la equivalencia, con media 0 y covarianza Γ . Este proceso es llamado, proceso de medias móviles de primer orden irregularmente espaciado de forma general. A continuación se darán las expresiones particulares de este proceso desde los dos puntos de vista antes definidos.

El punto de vista distribucional

En (2-1), γ_0 y γ_1, Δ_{n+1} , para $n \geq 1$, representa la varianzas y las covarianzas de primer orden respectivamente. Definimos la varianzas como $\gamma_0 = \sigma^2(1 + \theta^2)$ y las covarianzas de primer orden como $\gamma_1, \Delta_{n+1} = \sigma^2 \theta^{\Delta_{n+1}}$, $\sigma^2 > 0$ y $0 < \theta < 1$. Por tanto, obtenemos el proceso estocástico irregularmente espaciado de primer orden con matriz de covarianzas

$$\Gamma = \begin{bmatrix} 1 + \theta^2 & \theta^{\Delta_2} & 0 & 0 & \dots \\ \theta^{\Delta_2} & 1 + \theta^2 & \theta^{\Delta_3} & 0 & \\ 0 & \theta^{\Delta_3} & 1 + \theta^2 & \theta^{\Delta_4} & \\ \cdot & & & & \cdot \\ \cdot & & & & \cdot \\ \cdot & & & & \cdot \end{bmatrix} \quad (2-2)$$

el cual contiene el modelo de medias móviles convencional como caso especial. Este es llamado proceso Gaussiano irregular de medias móviles de primer orden.

El punto de vista constructor

Ahora, como es usual, especificaremos el proceso IMA como función de otros procesos estocásticos. Sea $\{\epsilon_{t_n}\}_{n \geq 1}$ variables aleatorias independientes que siguen una distribución normal $N(0, \sigma^2 c_n(\theta))$ con $\sigma^2 > 0, 0 < \theta < 1, c_1(\theta) = 1 + \theta^2$ y

$$c_n(\theta) = 1 + \theta^2 - \frac{\theta^{2\Delta_n}}{c_{n-1}(\theta)} \text{ para } n \geq 2$$

donde $\Delta_n = t_n - t_{n-1}$. El proceso $\{X_{t_n}, t_n \in \mathbb{T}'\}$, es decirse tiene un proceso IMA si $X_{t_1} = \epsilon_{t_1}$ y para $n \geq 2$

$$X_{t_n} = \epsilon_{t_n} + \frac{\theta^{\Delta_n}}{c_{n-1}(\theta)} \epsilon_{t_{n-1}} \quad (2-3)$$

Decimos que $\{X_{t_n}, t_n \in \mathbb{T}'\}$ es un proceso IMA con media μ si $\{X_{t_n} - \mu, t_n \in \mathbb{T}'\}$ es un proceso IMA

2.4 El proceso Autorregresivo de primer orden irregular (IAR)

Se denota X_{t_n} a una observación medida en el tiempo t_n y consideramos que $\{t_n\}$ para $n = 1, \dots, n$, se define el proceso IAR de primer orden así:

$$X_{t_n} = \phi^{t_n - t_{n-1}} X_{t_{n-1}} + \sigma \sqrt{1 - \phi^{2(t_n - t_{n-1})}} \varepsilon_{t_n} \quad (2-4)$$

Donde ε_{t_n} son variables aleatorias independientes con media cero y varianza unitaria, observe que $E(X_{t_n}) = 0$ y $Var(X_{t_n}) = \sigma^2$ para todo X_{t_n} .

Así, la covarianza entre X_{t_k} y X_{t_j} es $E(X_{t_k}, X_{t_j}) = \sigma^2 \phi^{t_k - t_j}$ para $k \geq j$. La matriz de varianzas y autocovarianzas se muestra en 2-5.

$$\begin{bmatrix} 1 & \phi^{t_2 - t_1} & \phi^{t_3 - t_1} & \cdot & \cdot & \cdot & \cdot & \phi^{t_n - t_1} \\ & 1 & \phi^{t_3 - t_2} & \phi^{t_4 - t_2} & \cdot & \cdot & \cdot & \phi^{t_n - t_2} \\ & & 1 & \phi^{t_4 - t_3} & \phi^{t_5 - t_3} & \cdot & \cdot & \phi^{t_n - t_3} \\ & & & 1 & \cdot & \cdot & \cdot & \cdot \\ & & & & \cdot & \cdot & \cdot & \cdot \\ & & & & & \cdot & \cdot & \cdot \\ & & & & & & \phi^{t_n - t_{n-1}} & \cdot \\ & & & & & & & 1 \end{bmatrix} \quad (2-5)$$

Así, para cualquier par de tiempos observacionales $s < t$, podemos definir la función autocovarianza cómo:

$$\gamma(t - s) = E(X_t, X_s) = \sigma^2 \phi^{t-s}$$

También, la función de autocorrelación (ACF) está dada por $\rho(t - s) = \frac{\gamma(t-s)}{\gamma(0)} = \phi^{t-s}$

La secuencia X_{t_n} corresponde a un proceso debilmente estacionario de segundo orden, también se puede demostrar que bajo algunas condiciones el proceso es estacionario y

ergódico. Ver (Eyheramendy et al. 2018).

Teorema 2.4.1. *Considere el proceso definido por la ecuación 2-4 y supongamos que el ruido de entrada es una secuencia iid de variables aleatorias con media cero y varianza unitaria. Además, suponga que $t_j - t_{j-n} \geq C \log n$ como $n \rightarrow \infty$, $0 \leq \phi \leq 1$, donde C es una constante positiva que satisface $C \log \phi^2 \geq -1$. Luego, existe una solución al proceso definido por la ecuación 2-4 y la secuencia $\{X_{t_j}\}$ es estacionaria y ergódica.*

Note también que si $t_n - t_{n-1} = 1$ para todo n , la ecuación 2-4 se convierte en:

$$X_{t_n} = \phi X_{t_{n-1}} + \sigma \sqrt{1 - \phi^2} \varepsilon_{t_n} \quad (2-6)$$

La ecuación 2-6 corresponde al modelo autorregresivo de primer orden para datos regularmente espaciados AR(1). Por lo tanto, el modelo IAR es una extensión del modelo autorregresivo regular

2.4.1 Estimación

La verosimilitud de una muestra aleatoria $X_{t_1} \dots X_{t_n}$ puede ser expresada como

$$f(X_{t_1} \dots X_{t_n}; \theta) = f(X_{t_1}; \theta) * f(X_{t_2} | X_{t_1}; \theta) * \dots * f(X_{t_n} | X_{t_{n-1}}; \theta)$$

donde $\theta = (\sigma^2, \phi)$ es el vector de parámetros del modelo, para poder describir el proceso de estimación, se asume que las distribuciones condicionales son Gaussianas, es decir, se asume que

$$f(X_{t_1}) \sim N(0, \sigma^2) \text{ y } f(X_{t_n} | X_{t_{n-1}}; \theta) \sim N(\phi^{t_j - t_{j-1}} X_{t_{j-1}}, \sigma^2(1 - \phi^{2(t_j - t_{j-1})})) \text{ para } j = 1, 2, 3, \dots, n$$

basado en la ecuación 2-4, la log-verosimilitud del proceso puede ser escrita como

$$l(\theta) = \frac{n}{2} \log(2\pi) + \frac{1}{2} \sum_{j=1}^n \log v_{t_j} + \frac{1}{2} \sum_{j=1}^n \frac{e_{t_j}}{v_{t_j}} \quad (2-7)$$

se define $e_{t_1} = X_{t_1}$, $e_{t_j} = X_{t_j} - \phi^{t_j - t_{j-1}} X_{t_{j-1}}$ para $j > 1$ y sus varianzas $v_{t_1} = \text{Var}(e_{t_1})$.

Observe que el predictor pasado de tiempo finito del proceso en el tiempo t_j es dado por $\hat{X}_{t_1} = 0$, y $\hat{X}_{t_j} = \phi^{t_j - t_{j-1}} X_{t_{j-1}}$, para $j = 2, \dots, n$

Por tanto, $e_{t_j} = X_{t_j} - \hat{X}_{t_j}$ es el error de predicción con varianza $v_{t_1} = \text{Var}(e_{t_1} = \sigma^2)$ y $v_{t_j} = \text{var}(e_{t_j}) = \sigma^2[1 - \phi^{2(t_j - t_{j-1})}]$, para $j = 2, \dots, n$

Por maximización directa de la función de log-verosimilitud (2-7), se puede obtener el estimador de máxima verosimilitud de σ^2

$$\hat{\sigma}^2 = \frac{1}{n} \sum_{j=1}^n \frac{(X_{t_j} - \hat{X}_{t_j})}{\tau_{t_j}}, \text{ donde } \tau_{t_j} = v_{t_j}/\sigma^2. \quad (2-8)$$

Pero no es posible encontrar $\hat{\phi}$, el estimador por máxima verosimilitud de ϕ , por maximización directa de la verosimilitud, es necesario recurrir a metodos iterativos.

2.5 Modelo autoregresivo complejo irregular (CIAR)

Para derivar una extensión compleja del modelo 2-4, se sigue el mismo proceso que para series de tiempo regulares. Suponiendo que x_{t_j} es una secuencia de valores complejos, tal que $x_{t_j} = y_{t_j} + iz_{t_j} \forall j = 1, \dots, n$, y de manera similar, $\phi = \phi^R + i\phi^I$ es el coeficiente complejo del modelo y $\varepsilon_{t_j} = \varepsilon_{t_j}^R + i\varepsilon_{t_j}^I$ es un ruido blanco complejo, definimos el proceso CIAR como

$$y_{t_j} + iz_{t_j} = (\phi^R + i\phi^I)^{t_j - t_{j-1}} (y_{t_{j-1}} + iz_{t_{j-1}}) + \sigma_{t_j} (\varepsilon_{t_j}^R + i\varepsilon_{t_j}^I) \quad (2-9)$$

donde $\sigma_{t_j} = \sigma \sqrt{1 - |\phi^{t_j - t_{j-1}}|^2}$ y $|\cdot|$ es el módulo de un número complejo. Suponemos que solo se observa la parte real y_{t_j} y que la parte imaginaria z_{t_j} es un proceso latente. Además, se asume que $\varepsilon_{t_j}^R$ y $\varepsilon_{t_j}^I$, la parte real e imaginaria de ε_{t_j} , respectivamente, son independientes con media cero y varianza positiva $\mathbb{V}(\varepsilon_t^R) = 1$ y $\mathbb{V}(\varepsilon_t^I) = c$, respectivamente, donde c es un parámetro fijo que toma valores en \mathbb{R}^+ . Generalmente, se supone que $c = 1$, y los valores iniciales son $y_{t_1} = \sigma \varepsilon_{t_1}^R$ y $z_{t_1} = \sigma \varepsilon_{t_1}^I$. En el siguiente lema, se presentan algunas de las propiedades de este proceso.

Lemma 2.5.1. *Considere el proceso CIAR x_{t_j} descrito por la ecuación 2-9. Defina $\gamma_0 = \mathbb{E}(\bar{x}_{t_j} x_{t_j})$, $\gamma_k = \mathbb{E}(\bar{x}_{t_{j+k}} x_{t_j})$ y ρ_k como la varianza, autocovarianza y autocorrelación, respectivamente, del proceso x_{t_j} . A continuación, el valor esperado, la varianza, la autocovarianza y la autocorrelación del proceso*

- $\mathbb{E}(x_{t_j}) = 0$
- $\mathbb{V}(x_{t_j}) = \gamma_0 = \mathbb{E}(\bar{x}_{t_j} x_{t_j}) = \sigma^2(1 + c)$,

- $\gamma_k = \mathbb{E}(\bar{x}_{t_{j+k}} x_{t_j}) = \overline{\phi^{\Delta_k}} \sigma^2 (1 + c),$
- $\rho_k = \overline{\phi^{\Delta_k}}$

donde $\Delta_k = t_{j+k} - t_j$ denota las diferencias de tiempo entre los tiempos de observación t_{j+k} y t_j . Además, \bar{x}_{t_j} es el conjugado complejo de x_{t_j} .

Si $|\phi| = |\phi^R + i\phi^I| < 1$, los resultados en el Lema 2.5.1 muestran que la secuencia compleja x_{t_j} es un proceso débilmente estacionario. Observamos que la ACF ρ_k del proceso CIAR decae a una tasa $\phi^{t_{j+k}-t_j}$ (la llamada decaimiento exponencial). Esta estructura de autocorrelación es diferente a la de un proceso CARFIMA de memoria antipersistente o de memoria intermedia, ya que la ACF de estos últimos decae más lentamente que un decaimiento exponencial. Por lo tanto, aunque ambos modelos pueden ajustar series de tiempo irregulares con valores negativos de la ACF, el uso apropiado de estos modelos dependerá de la estructura de correlación de los datos. Para tomar una decisión sobre el modelo más adecuado para los datos, se pueden utilizar técnicas basadas en la verosimilitud de los datos, como AIC, BIC, y así sucesivamente.

Chapter 3

Modelo autorregresivo de primer orden irregularmente espaciado

En este capítulo, proponemos un modelo basado en la propuesta de Eyheramendy et al. (2018), la cual permite modelar unicamente estructuras de autocorrelacion positivas, el llamado $IAR(1)$. Nosotros generalizamos esta propuesta para que se permitan estructuras tanto positivas como negativas, teniendo en cuenta que nuestra novedosa propuesta, debe tener como caso particular el modelo $IAR(1)$ y el clasico modelo $AR(1)$

3.1 Construcción del modelo

En esta sección, construiremos un proceso estocástico estacionario que tiene una estructura autorregresiva y que además considera espacios de tiempo irregulares. Suponemos que el comportamiento irregularmente espaciado es independiente de las propiedades estocásticas del proceso.

Sea $\{\varepsilon_\tau, \tau \in \mathbb{T}\}$ una secuencia iid de variables aleatorias cada una con distribución $N(0, 1)$ y defino:

$$X_{\tau_1} = W_1^{1/2} \varepsilon_{\tau_1}$$

$$X_{\tau_{n+1}} = \phi_{n+1} X_{\tau_n} + W_{n+1}^{1/2} \varepsilon_{\tau_{n+1}} \quad (3-1)$$

Donde, $0 \leq \phi \leq 1$, $E(X_{\tau_n}) = 0$ y $\{W_n\}_{n \geq 1}$ es una secuencia que caracteriza los momentos del proceso, ϕ_n es una funcion de Δ_n y del parametro ϕ del modelo autorregresivo que

definiremos mas adelante. así, para $n \geq 1$, tenemos $E(X_{\tau_n}) = 0$,

$$V(X_{\tau_1}) = W_1$$

$$V(X_{\tau_{n+1}}) = \phi_{n+1}^2 V(X_{\tau_n}) + W_{n+1}$$

Ahora, bajo este modelo, calculemos $Cov(X_{\tau_n}, X_{\tau_{n+k}})$, esto puede hacerse mirando casos particulares

- si $k=1$

$$Cov(X_{\tau_n}, X_{\tau_{n+1}}) = Cov(X_{\tau}, \phi_{n+1}X_{t_n} + W_{n+1}^{1/2}) = \phi_{n+1}Cov(X_{t_n}, X_{t_n}) = \phi_{n+1}Var(X_{t_n})$$

- si $k=2$

$$Cov(X_{\tau_n}, X_{\tau_{n+2}}) = Cov(X_{\tau}, \phi_{n+2}X_{t_{n+1}} + W_{n+2}^{1/2}) = \phi_{n+2}Cov(X_{t_n}, X_{t_{n+1}}) = \phi_{n+2}\phi_{n+1}Var(X_{t_n})$$

- si $k=3$

$$Cov(X_{\tau_n}, X_{\tau_{n+3}}) = Cov(X_{\tau}, \phi_{n+3}X_{t_{n+2}} + W_{n+3}^{1/2}) = \phi_{n+3}Cov(X_{t_n}, X_{t_{n+2}}) = \phi_{n+3}\phi_{n+2}\phi_{n+1}Var(X_{t_n})$$

- Reemplazando recursivamente hasta $k = n$ tenemos:

$$Cov(X_{\tau_n}, X_{\tau_{n+k}}) = \prod_{i=1}^k \phi_{n+i} Var(X_{\tau_n}) \quad (3-2)$$

Para que este proceso sea estacionario, debemos demostrar que la varianza es constante en cualquier momento τ , por tanto, debemos mostrar que, para $n \geq 1$; $Var(X_{\tau_{n+1}}) = Var(X_{\tau_1}) = \gamma_0$, es decir:

$$\phi_{n+1}^2 \gamma_0 + W_{n+1} = W_1 = \gamma_0 \quad (3-3)$$

despejando W_{n+1} de 3-3 tenemos:

$$W_{n+1} = \gamma_0(1 - \phi_{n+1}^2) \quad (3-4)$$

Finalmente, para que nuestro modelo coincida con el modelo $AR(1)$ en el momento en que $\Delta_{n+1} = 1$ para todo n , hacemos $\gamma_0 = \frac{\sigma^2}{1-\phi^2}$

3.2 Implementando propuesta de ϕ_n

El foco de la investigación está en generar una propuesta estable y teóricamente coherente para la secuencia ϕ_{n+1} , después de mucho estudiarlo, llegamos a esta expresión.

$$\phi_{n+1} = \text{sign}(\phi)|\phi|^{\Delta_{n+1}} \quad (3-5)$$

Si reemplazamos 3-5 en 3-4 obtenemos:

$$W_{n+1} = \frac{\sigma^2}{1-\phi^2} \left(1 - \text{sign}(\phi)^2 |\phi|^{2\Delta_{n+1}} \right) \quad (3-6)$$

El modelo final, reemplazando 3-5 y 3-6 en 3-1 tiene la expresión:

$$X_{\tau_{n+1}} = \text{sign}(\phi)|\phi|^{\Delta_{n+1}} X_{\tau_n} + \left[\frac{\sigma^2}{1-\phi^2} \left(1 - \text{sign}(\phi)^2 |\phi|^{2\Delta_{n+1}} \right) \right]^{1/2} \varepsilon_{\tau_{n+1}} \quad (3-7)$$

por definición la varianza de la expresión 3-7 es:

$$V(X_{\tau_{n+1}}) = \frac{\sigma^2}{1-\phi^2} \text{ y } Cov(X_{\tau_n}, X_{\tau_{n+k}}) = \left(\frac{\sigma^2}{1-\phi^2} \right) \text{sign}(\phi)^k |\phi|^{\sum_{i=1}^k \Delta_{n+i}}$$

3.3 El modelo autorregresivo irregularmente espaciado de primer orden

Basado en los calculos y definiciones anteriores, esta nueva propuesta de modelo autorregresivo irregularmente espaciado IAR es realizada desde un punto de vista constructorista. Más adelante en este documento presentamos las propiedades del proceso.

Definición 3.3.1 (Proceso IAR - Autocorrelaciones positivas y negativas). Sea $\{\varepsilon_{\tau_n}\}_{n \geq 1}$ variables aleatorias, cada una con distribución $N(0,1)$, siendo $\sigma^2 > 0$ y $-1 < \phi < 1$, se dice es un proceso IAR Si $X_1 = \left(\frac{\sigma^2}{1-\phi^2} \right)^{1/2} \varepsilon_{\tau_1}$ y para $n \geq 2$.

$$X_{\tau_n} = \text{sign}(\phi)|\phi|^{\Delta_n} X_{\tau_{n-1}} + \left[\frac{\sigma^2}{1-\phi^2} \left(1 - \text{sign}(\phi)^2 |\phi|^{2\Delta_n} \right) \right]^{1/2} \varepsilon_{\tau_n} \quad (3-8)$$

Donde $\Delta_n = \tau_n - \tau_{n-1}$

3.4 Propiedades del proceso

If $X_n = [X_{\tau_1}, \dots, X_{\tau_n}]'$ es un vector aleatorio de un proceso IAR, luego X_n es un vector aleatorio Gaussiano con $m_n = 0$ y matriz de varianzas y covarianza dada por

$$\Gamma_n = \begin{bmatrix} \gamma_0 & & & & & \\ \phi_2 \gamma_0 & \gamma_0 & & & & \\ \phi_2 \phi_3 \gamma_0 & \phi_3 \gamma_0 & \ddots & & & \\ \vdots & \vdots & & \gamma_0 & & \\ \prod_{i=2}^{n-1} \phi_i \gamma_0 & \prod_{i=3}^{n-1} \phi_i \gamma_0 & \cdots & \phi_{n-1} \gamma_0 & \gamma_0 & \\ \prod_{i=2}^n \phi_i \gamma_0 & \prod_{i=3}^n \phi_i \gamma_0 & \cdots & \phi_{n-1} \phi_n \gamma_0 & \phi_n \gamma_0 & \gamma_0 \end{bmatrix},$$

donde $\phi_n = \text{sign}(\phi)|\phi|^{\Delta_n}$ y $\gamma_0 = \frac{\sigma^2}{1-\phi^2}$, para $n \geq 2$. Así, el proceso IAR es, por definicion un proceso Gaussiano debilmente estacionario y por tanto estrictamente estacionario. Además, note que cuando $\Delta_n = 1$, para $n \geq 2$, obtenemos el proceso AR convencional.

3.5 Pronóstico en el modelo IAR

El objetivo es predecir los valores futuros de la serie $X_{\tau_{n+m}}$, donde $m = 1, 2, \dots$, basados en la información recolectada $X_{\tau_{1:n}} = X_{\tau_1}, \dots, X_{\tau_n}$. El predictor de error cuadrático medio mínimo para $X_{\tau_{n+m}}$ es:

$$\hat{X}_{\tau_{n+m}} = E(X_{\tau_{n+m}} | X_{\tau_{1:n}}) \quad (3-9)$$

Esto se debe a que la esperanza condicional minimiza el error cuadrático medio, como se demuestra en el apéndice.

$$E[X_{\tau_{n+m}} - g(X_{\tau_{1:n}})]. \quad (3-10)$$

Si consideramos la predicción a un paso, $\hat{X}_{\tau_{n+1}}$, basada en el modelo 3-7, tenemos que:

$$\hat{X}_{n+1} = E(X_{\tau_{n+1}} | X_{\tau_{1:n}}) = \text{sign}(\phi)|\phi|^{\Delta_{n+1}} X_{\tau_n}. \quad (3-11)$$

Además, $X_{\tau_1} = W_1^{1/2} \varepsilon_{\tau_1}$, por lo tanto $\hat{X}_{\tau_1} = 0$. Con estos resultados, el error de predicción es:

$$(X_{\tau_k} - \hat{X}_{\tau_k}) = W_k^{1/2} \varepsilon_{\tau_k}, \quad k = 1, \dots, n \quad (3-12)$$

Podemos ver que, debido a la forma en que se define el modelo, el error de predicción sigue una distribución normal con media 0 y varianza W_k .

3.6 Estimación

Sea X_τ observado en los puntos τ_1, \dots, τ_n , la función de verosimilitud para estimar los parámetros σ y ϕ es la siguiente:

$$L(\phi, \sigma) = \prod_{i=1}^n \frac{1}{\sqrt{2\pi W_i}} e^{-\frac{(X_{\tau_i} - \hat{X}_{\tau_i})^2}{2W_i}} \quad (3-13)$$

Ahora, aplicamos logaritmo para calcular la log-verosimilitud $l(\phi, \sigma)$

$$\begin{aligned} l(\phi, \sigma) &= \sum_{i=1}^n \log\left(\frac{1}{\sqrt{2\pi W_i}}\right) - \frac{1}{2} \sum_{i=1}^n \frac{(X_{\tau_i} - \hat{X}_{\tau_i})^2}{W_i} \\ &= -\frac{1}{2} \sum_{i=1}^n \log(2\pi W_i) - \frac{1}{2} \sum_{i=1}^n \frac{(X_{\tau_i} - \hat{X}_{\tau_i})^2}{W_i} \\ &= -\frac{n}{2} \log(2\pi) - \frac{1}{2} \sum_{i=1}^n \log(W_i) - \frac{1}{2} \sum_{i=1}^n \frac{(X_{\tau_i} - \hat{X}_{\tau_i})^2}{W_i} \end{aligned} \quad (3-14)$$

Donde W_i es definido por la ecuación 3-4. Por maximización directa de la función de log verosimilitud (3-14), podemos obtener el estimador de σ^2

$$\frac{\partial(l)}{\partial W_i} = \frac{1}{2} \sum_{i=1}^n \frac{1}{W_i} - \frac{1}{2} \sum_{i=1}^n \frac{(X_{\tau_i} - \hat{X}_{\tau_i})^2}{W_i^2} = 0 \quad (3-15)$$

$$\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n \frac{(X_{\tau_i} - \hat{X}_{\tau_i})^2}{W_i/\sigma} \quad (3-16)$$

Pero, la estimación de ϕ no es posible calcularla por maximización directa, por tanto, es necesario acudir a métodos iterativos.

3.7 Estimación Bootstrap

Ahora, aplicamos el método bootstrap descrito en ?? para estimar ϕ_0 en el proceso IMA con $\sigma_0^2 = 1$. Sea X_τ observado en los puntos τ_1, \dots, τ_N y considere $\hat{\phi}_N$ como la estimación MLE respectiva. Las innovaciones estandarizadas estimadas son

$$e_{\tau_n}^s = \frac{X_{\tau_n} - \hat{X}_{\phi_n}(\hat{\phi}_N)}{\sqrt{W_n(\hat{\phi}_N)}},$$

para $n = 2, \dots, N$. El llamado remuestreo basado en el modelo podría proceder por muestreo equiprobable con reemplazo de los residuos centrados $e_{t_2}^s - \bar{e}, \dots, e_{t_N}^s - \bar{e}$, donde $\bar{e} = \sum_{n=2}^N e_{t_n}^s / (N-1)$, para obtener innovaciones simuladas $\zeta_{t_1}^*, \dots, \zeta_{t_N}^*$, y luego establecer:

$$X_{\tau_1}^* = \sqrt{W(\hat{\phi}_N)} \zeta_{\tau_1}^*, \quad (3-17)$$

$$X_{\tau_n}^* = \hat{\phi}_n X_{\tau_{n-1}}^* + \sqrt{W(\hat{\phi}_N)} \zeta_{\tau_n}^* \quad (3-18)$$

A continuación, estimamos los parámetros a través de MLE suponiendo que los datos son $X_{t_n}^*$. Por lo tanto, podemos repetir este proceso un gran número, B , de veces generando una colección de estimaciones de parámetros bootstrap. A continuación, podemos aproximar la distribución de muestra finita del estimador, $\hat{\phi}_N$, a partir de los valores de los parámetros bootstrap.

3.8 Estudio de Simulación Monte Carlo

Para probar el modelo 3-8 se genera un escenario de simulación, donde se muestran sus propiedades y las de los estimadores por máxima verosimilitud y bootstrap, la simulación se realiza considerando $\sigma_0 = 1$, $\phi_0 \in \{\pm 0.1, \pm 0.5, \pm 0.9\}$ y $n \in \{20, 50, 100\}$ para que pueda observarse el comportamiento irregularmente, los tiempos τ_1, \dots, τ_n pueden ser considerados tanto regulares como irregulares con este modelo, en este caso, consideramos $\Delta_n = \tau_n - \tau_{n-1} \sim 1 + \text{poisson}(\lambda = 2)$ (para asegurar que no haya 0). La figura 3-1 muestra el comportamiento de una trayectoria simulada para cada una de las combinaciones de los parámetros establecidos en el caso de $\phi > 0$, las marcas rojas en la parte de arriba del gráfico muestran mas claramente el comportamiento irregular. La figura 3-2 muestra el

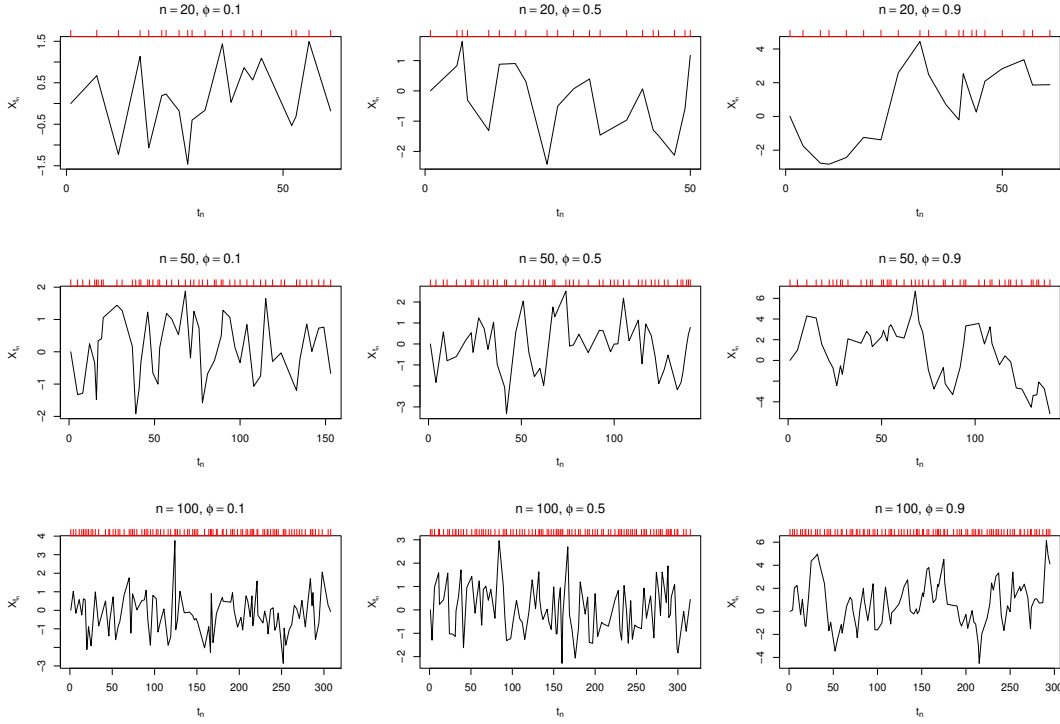


Figure 3-1: simulación del modelo IAR con coeficientes positivos

mismo comportamiento en el caso de $\phi < 0$. Para efectos de la simulación montecarlo simulamos $M = 1000$ trayectorias, $\{S_m\}_{m=1}^M$ y estimamos ϕ_0 para cada configuración. Sea $\hat{\phi}_m^{\text{MLE}}$ la estimación ML y $\widehat{\text{se}}(\hat{\phi}_m^{\text{MLE}})$ el error estandar estimado para la m -ésima trayectoria. El error estandar se estima por la curvatura de la superficie de verosimilitud en $\hat{\phi}_m^{\text{MLE}}$. Resumimos las estimaciones de máxima verosimilitud M mediante

$$\hat{\phi}^{\text{MLE}} = \frac{1}{M} \sum_{m=1}^M \hat{\phi}_m^{\text{MLE}} \quad \text{y} \quad \widehat{\text{se}}(\hat{\phi}^{\text{MLE}}) = \frac{1}{M} \sum_{m=1}^M \widehat{\text{se}}(\hat{\phi}_m^{\text{MLE}}).$$

Por otra parte, para cada trayectoria, hemos simulado $B = 500$ trayectorias bootstrap representadas por $\{\{S_{m,b}\}_{b=1}^B\}_{m=1}^M$. Entonces, obtuvimos $\{\{\hat{\phi}_{m,b}^b\}_{b=1}^B\}_{m=1}^M$ (las estimaciones ML). La estimación bootstrap y su error estándar estimado se definen como

$$\hat{\phi}_m^b = \frac{1}{B} \sum_{b=1}^B \hat{\phi}_{m,b}^b \quad \text{and} \quad \widehat{\text{se}}^2(\hat{\phi}_m^b) = \frac{1}{B-1} \sum_{b=1}^B (\hat{\phi}_{m,b}^b - \hat{\phi}_m^b)^2,$$

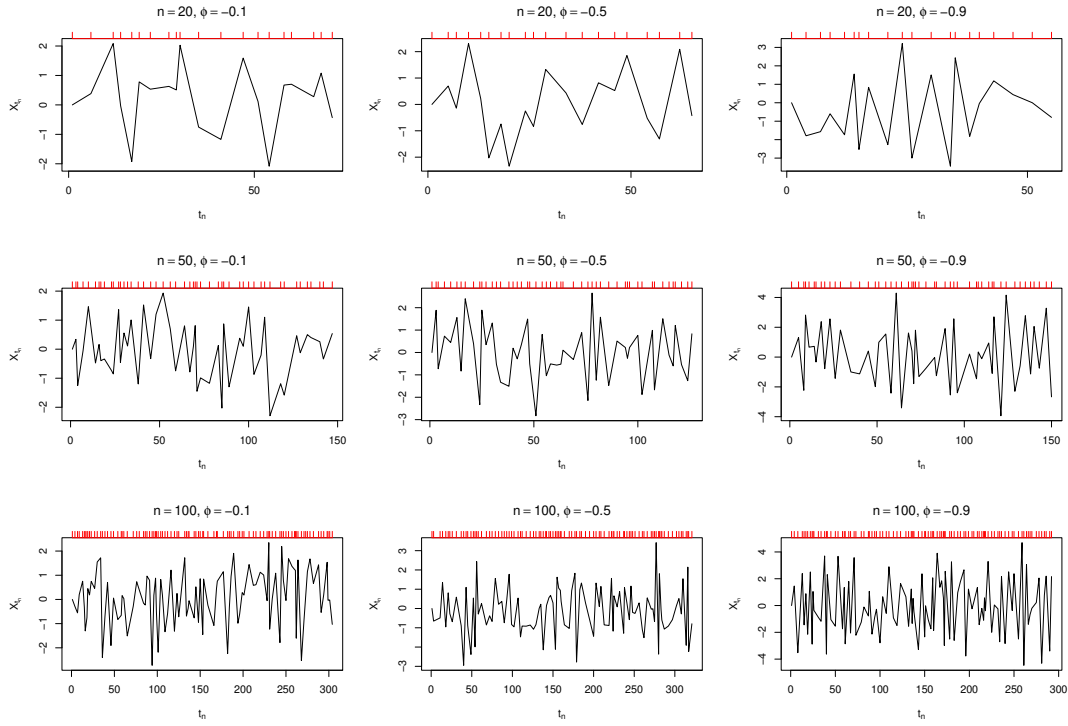


Figure 3-2: simulación del modelo IAR con coeficientes negativos

para $m = 1, \dots, M$. Por último, resumimos las M estimaciones bootstrap mediante

$$\hat{\phi}^b = \frac{1}{M} \sum_{m=1}^M \hat{\phi}_m^b \quad \text{and} \quad \widehat{\text{se}}(\hat{\phi}^b) = \frac{1}{M} \sum_{m=1}^M \widehat{\text{se}}(\hat{\phi}_m^b).$$

Además, como medida del rendimiento del estimador, utilizamos el error cuadrático medio (RMSE), y el coeficiente de variación (CV) estimado definido por

$$\widehat{\text{RMSE}}_{\hat{\phi}^{\text{MLE}}} = \left(\widehat{\text{se}}(\hat{\phi}^{\text{MLE}})^2 + \widehat{\text{bias}}_{\hat{\phi}^{\text{MLE}}}^2 \right)^{1/2}, \text{ and}$$

$$\widehat{\text{CV}}_{\hat{\phi}^{\text{MLE}}} = \frac{\widehat{\text{se}}(\hat{\phi}^{\text{MLE}})}{|\hat{\phi}^{\text{MLE}}|},$$

donde $\widehat{\text{bias}}_{\hat{\phi}^{\text{MLE}}} = \hat{\phi}^{\text{MLE}} - \phi$. Por último, estimamos la varianza del estimador mediante

$$\widehat{\text{se}}^2(\hat{\phi}^{\text{MLE}}) = \frac{1}{M-1} \sum_{m=1}^M (\hat{\phi}_m^{\text{MLE}} - \hat{\phi}^{\text{MLE}})^2$$

De la misma manera, definimos $\widehat{\text{RMSE}}_{\hat{\phi}^b}$, $\widehat{\text{CV}}_{\hat{\phi}^b}$, $\widehat{\text{bias}}_{\hat{\phi}^b}$ y $\widehat{\text{se}}^2(\hat{\phi}^b)$ para el caso Bootstrap. La figura 3-3 muestra flujo de trabajo de este estudio de simulación

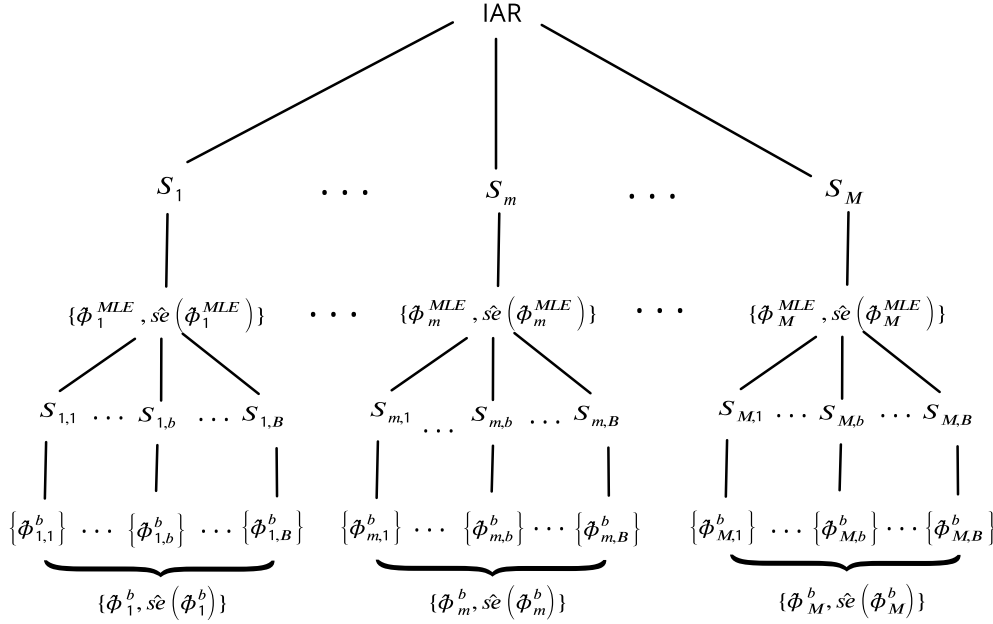


Figure 3-3: Esquema general del estudio de Monte Carlo.

En este estudio, se evaluaron los resultados del modelo en el caso de tiempo irregularmente espaciado, aunque también se comprobó que el modelo es válido para tiempos regulares con $\Delta_n = 1$ para todos los valores de n .

Para analizar el comportamiento de los estimadores de máxima verosimilitud y bootstrap, se realizaron simulaciones de distribuciones de muestras finitas. Los resultados de estas simulaciones se muestran en las figuras 3-4 y 3-5. Se puede observar que ambos estimadores presentan un comportamiento satisfactorio, parecen ser insesgados y consistentes en sus resultados.

Note que para valores pequeños de n y para valores cercanos a cero de ϕ , los estimadores tienden a tener una mayor variabilidad. Este hallazgo es importante para el análisis de los resultados, ya que sugiere que la precisión de los estimadores puede verse afectada por factores como la cantidad de datos disponibles y la magnitud del parámetro de interés.

En segundo lugar, para ambos estimadores, presentamos las medidas de rendimiento para ambos métodos de estimación las cuales se ven reflejadas en la tabla 3-1

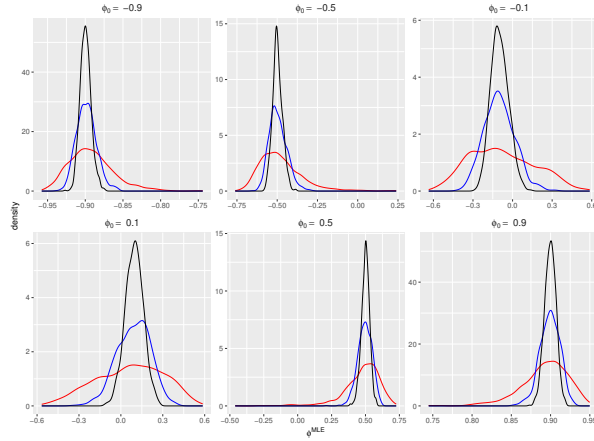


Figure 3-4: Esquema general del estudio de Monte Carlo.

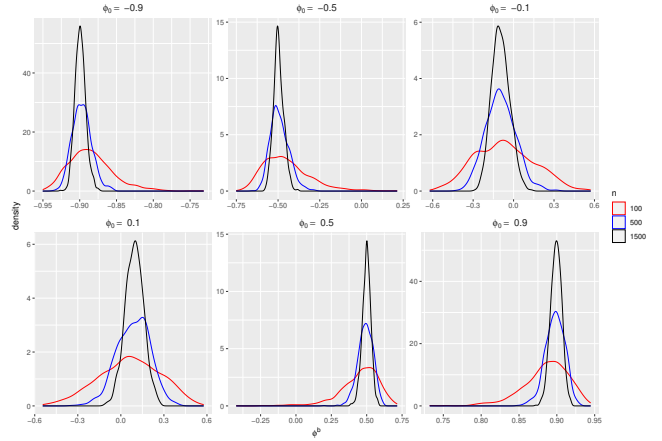


Figure 3-5: Esquema general del estudio de Monte Carlo.

Table 3-1: Resultados del estudio de Monte Carlo

n	Máxima Verosimilitud						Bootstrap				
	ϕ	$\hat{\phi}^{MLE}$	$\hat{se}(\phi^{MLE})$	$\widehat{bias}_{\hat{\phi}^{MLE}}$	$\widehat{RMSE}_{\hat{\phi}^{MLE}}$	$\widehat{CV}_{\hat{\phi}^{MLE}}$	$\hat{\phi}^b$	$\hat{se}(\phi^b)$	$\widehat{bias}_{\hat{\phi}^b}$	$\widehat{RMSE}_{\hat{\phi}^b}$	$\widehat{CV}_{\hat{\phi}^b}$
100	-0.9	-0.892	0.029	-0.008	0.031	0.033	-0.885	0.032	-0.015	0.035	0.036
	-0.5	-0.479	0.126	-0.021	0.128	0.263	-0.450	0.153	-0.050	0.161	0.340
	-0.1	-0.081	0.210	-0.019	0.211	2.586	-0.068	0.222	-0.032	0.224	3.267
	0.1	0.062	0.211	0.038	0.214	3.388	0.060	0.222	0.040	0.226	3.726
	0.5	0.473	0.126	0.027	0.129	0.267	0.450	0.145	0.050	0.154	0.323
	0.9	0.892	0.029	0.008	0.030	0.033	0.886	0.032	0.014	0.035	0.036
500	-0.9	-0.898	0.013	-0.002	0.013	0.014	-0.897	0.013	-0.003	0.013	0.014
	-0.5	-0.495	0.052	-0.005	0.053	0.106	-0.491	0.054	-0.009	0.055	0.110
	-0.1	-0.101	0.110	0.001	0.110	1.095	-0.097	0.111	-0.003	0.111	1.153
	0.1	0.091	0.111	0.009	0.112	1.227	0.087	0.111	0.013	0.112	1.274
	0.5	0.495	0.052	0.005	0.052	0.105	0.491	0.054	0.009	0.054	0.109
	0.9	0.899	0.013	0.001	0.013	0.014	0.898	0.013	0.002	0.013	0.014
1500	-0.9	-0.900	0.007	0.000	0.007	0.008	-0.899	0.007	-0.001	0.007	0.008
	-0.5	-0.498	0.030	-0.002	0.030	0.060	-0.497	0.030	-0.003	0.030	0.060
	-0.1	-0.098	0.066	-0.002	0.066	0.673	-0.096	0.066	-0.004	0.066	0.688
	0.1	0.096	0.067	0.004	0.067	0.695	0.094	0.066	0.006	0.067	0.705
	0.5	0.500	0.030	0.000	0.030	0.059	0.498	0.030	0.002	0.030	0.060
	0.9	0.900	0.007	0.000	0.007	0.008	0.899	0.007	0.001	0.007	0.008

3.9 Comparación del Modelo IAR vs el CIAR

En este apartado, nos disponemos a comparar nuestra solución, con la solución propuesta en Elorrieta et al. (2019) descrita en la sección 2.5 (El modelo CIAR), para ello, se plantea un escenario de simulación similar al de la sección anterior, donde, $\sigma_0 = 1$, $\phi_0 \in \{-0.1, -0.5, -0.9\}$, $n \in \{100, 5000, 1500\}$ y $\Delta_n = \tau_n - \tau_{n-1} \sim 1 + \text{poisson}(\lambda = 2)$. En este caso solo simulamos muestras aleatorias con parámetro negativo ya que es el foco de este estudio. En el caso del IAR, se asumen que $\phi^I = 0$ (la parte imaginaria es cero) y que $c = 0$. Para comparar nuestro modelo con el CIAR, utilizaremos el *Error cuadrático medio* (ECM), que definiremos así

$$ECM = \sqrt{\frac{\sum_{i=1}^n (e_i - \bar{e})^2}{n}}$$

donde $e_i = X_{t_i} - \hat{X}_{\tau_i}$ para $i = 1, \dots, n$.

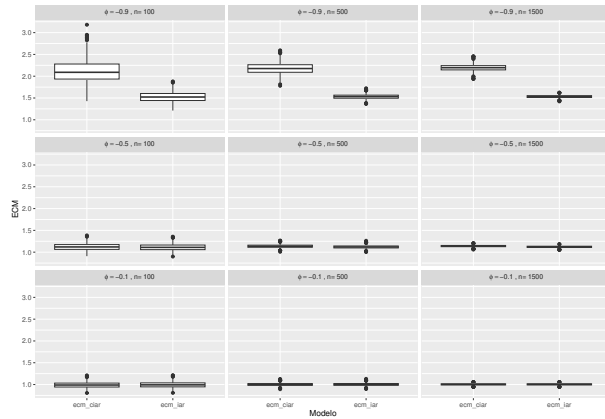


Figure 3-6: Comportamiento de la simulación, tomando como base el IAR.

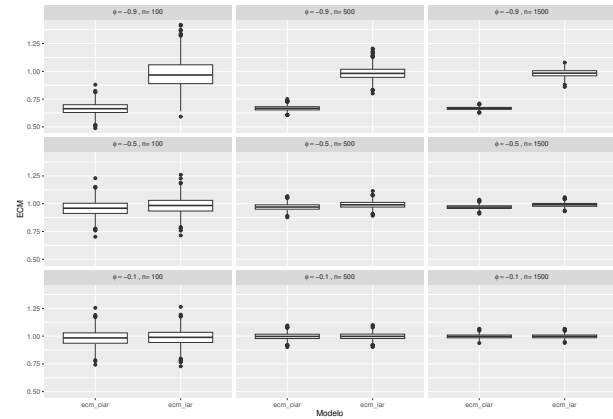


Figure 3-7: Comportamiento de la simulación, tomando como base el CIAR.

la figura 3-6 muestra el resultado de la simulación de un IAR con las condiciones dadas, y estimando tanto el IAR como el CIAR con los mismos datos simulados y midiendo el ECM de la predicción. Por otra parte la figura 3-7 muestra los mismos resultados pero esta vez, los datos fueron simulados a partir del CIAR.

Esta simulación permite llevarnos a la conclusión de que, cuando el parámetro $\phi_n \rightarrow -1$ el modelo IAR tiene un rendimiento mucho mejor en cuanto al ECM. Por otra parte cuando el parámetro $\phi \rightarrow 0$ el CIAR presenta un mejor comportamiento, sobre todo en muestras grandes. Y cuando ϕ está al rededor de 0.5, el CIAR muestra un mejor rendimiento en

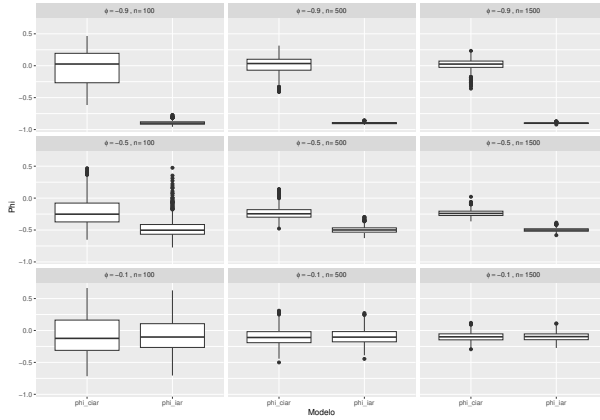


Figure 3-8: Comportamiento de la simulación, tomando como base el IAR Valor de ϕ estimado.

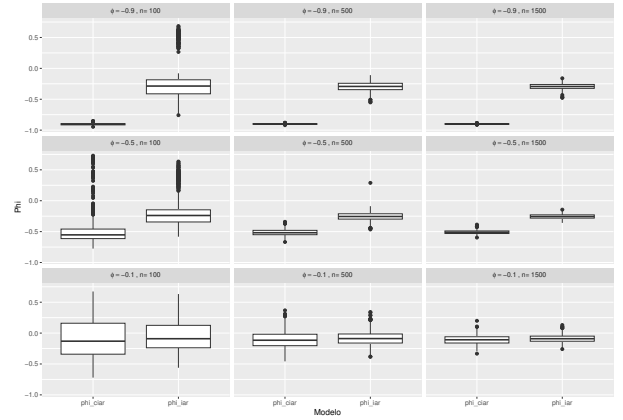


Figure 3-9: Comportamiento de la simulación, tomando como base el CIAR Valor de ϕ estimado.

muestras pequeñas, pero el IAR es mejor para muestras grandes, este comportamiento es idéntico tanto simulando IAR como CIAR.

3.10 Aplicación en datos astronómicos

Estos datos describen una serie temporal de la cantidad de energía en ergios que llega por cm^2 en un ancho de banda de un Amstrong ($10^{(-15)} ergs/s/cm^2/A$) desde la galaxia activa (AGN) MCG-6-30-15 medida en la banda K entre agosto de 2006 y julio de 2011. La medición se realizó utilizando la cámara ANDICAM, la cual estaba montada en el telescopio de 1.3 metros ubicado en el Observatorio Interamericano Cerro Tololo (CTIO).

3.11 Aplicación en datos de espeleotemas¹

Estos datos han sido descritos y trabajados por Mudelse (2014) en su libro y los hemos encontrado disponibles en la pagina web del autor. Este conjunto de datos cuenta con 1345 datos donde se miden el pocentaje de isotopos de oxigeno 18 y 16 estables; $\delta^{18}O$ a travez de los años, la escala de tiempo (para años antes de 1950) es basada en se basada en 18 edades espectrométricas de masa U/Th . La figura 3-12 muestra el comportamiento

¹Los espeleotemas son formaciones geológicas que se encuentran en cuevas y se forman a partir de la acumulación de minerales depositados por el agua que gotea del techo o fluye por las paredes de la cueva. Algunos ejemplos de espeleotemas incluyen estalactitas, estalagmitas, columnas y cortinas de piedra caliza.

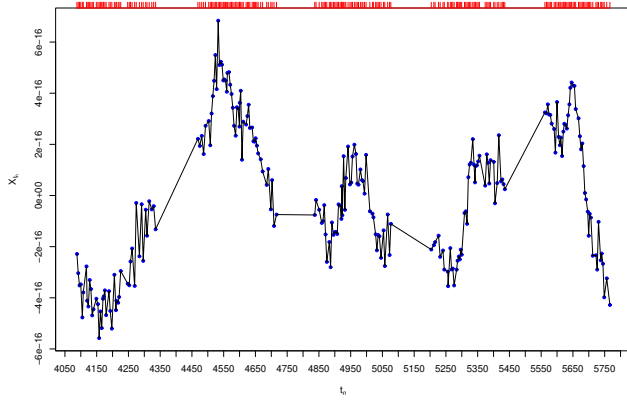


Figure 3-10: Comportamiento de la serie de tiempo.

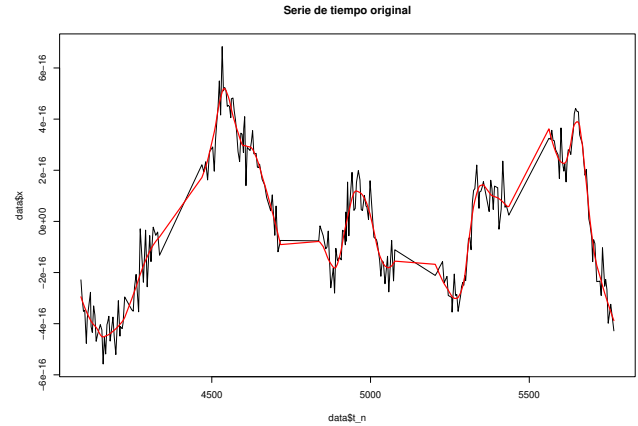


Figure 3-11: Comportamiento del Δ_n de la serie.

de la serie de tiempo descrita y la figura 3-13 muestra el comportamiento irregularmente espaciado entre los tiempos, note que si esta fuera una serie regular, este grafico sería una línea constante horizontal en $y = 1$

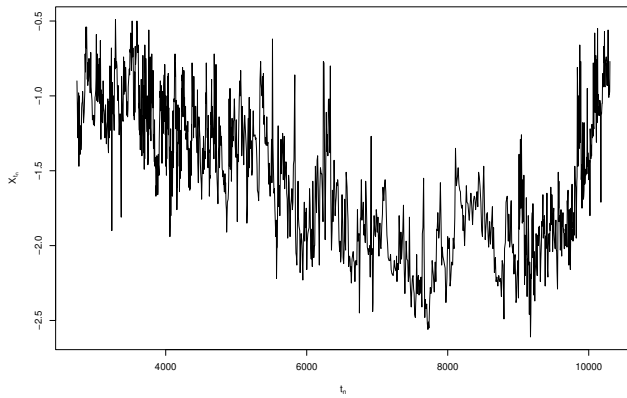


Figure 3-12: Comportamiento de la serie de tiempo.

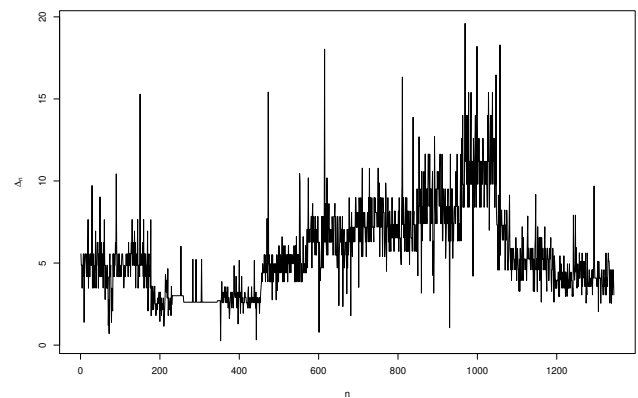


Figure 3-13: Comportamiento del Δ_n de la serie.

Ahora procedemos a realizar nuestras estimaciones tanto por Bootstrap como por Máxima Verosimilitud, en este caso, ejemplificamos el uso tanto en la serie original como en la serie diferenciada, esta diferenciación se realiza con el fin de eliminar tendencias y comportamientos cíclicos en nuestra serie de tiempo, este caso, $\nabla X_{\tau_n} = X_{\tau_n} - X_{\tau_{n-1}}$ es la serie diferenciada. La figura 3-15 muestra el resultado de la serie diferenciada y la figura 3-14

muestra los resultados para la serie original.

Los parametros estimados para la serie original son los siguientes:

$$\begin{aligned}\hat{\phi}^{MLE} &= 0.997 & \hat{se}(\phi^{MLE}) &= 0.0007 & \hat{\sigma}_{MLE}^2 &= 0.0112 & \hat{se}(\hat{\sigma}_{MLE}^2) &= 0.0020 \\ \hat{\phi}^b &= 0.997 & \hat{se}(\phi^b) &= 0.00086 & \hat{\sigma}_b^2 &= 0.0069 & \hat{se}(\hat{\sigma}_b^2) &= 0.0020\end{aligned}$$

Seguido a esto, calculamos los parametros para la serie diferenciada obteniendo:

$$\begin{aligned}\hat{\phi}^{MLE} &= -0.790 & \hat{se}(\phi^{MLE}) &= 0.0136 & \hat{\sigma}_{MLE}^2 &= 0.0190 & \hat{se}(\hat{\sigma}_{MLE}^2) &= 0.0037 \\ \hat{\phi}^b &= -0.789 & \hat{se}(\phi^b) &= 0.0149 & \hat{\sigma}_b^2 &= 0.0167 & \hat{se}(\hat{\sigma}_b^2) &= 0.04785\end{aligned}$$

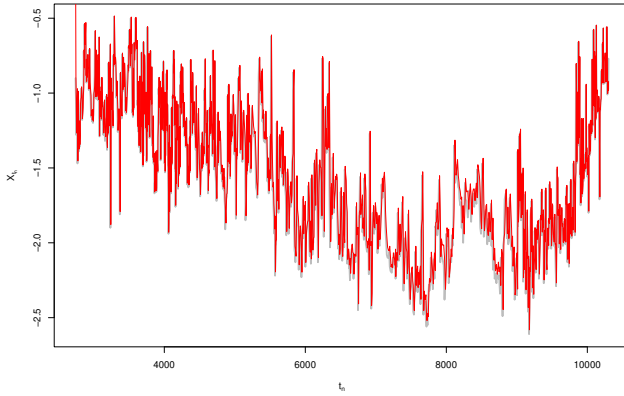


Figure 3-14: Comportamiento del modelo estimado con la serie original.

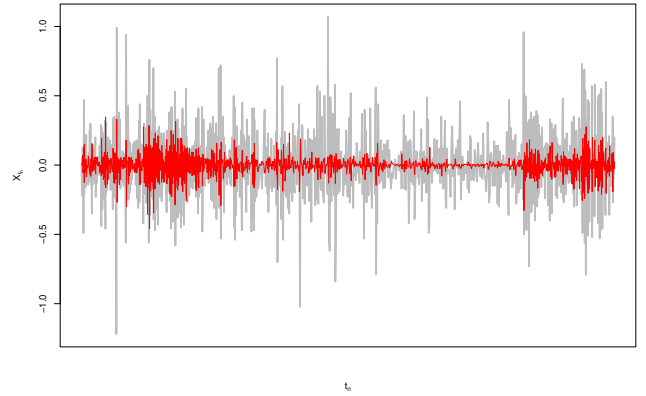


Figure 3-15: Comportamiento del modelo estimado con la serie diferenciada.

Chapter 4

Conclusion

Las conclusiones constituyen un capítulo independiente y presentan, en forma lógica, los resultados del trabajo de grado. Las conclusiones deben ser la respuesta a los objetivos o propósitos planteados. Se deben titular con la palabra conclusiones en el mismo formato de los títulos de los capítulos anteriores (Títulos primer nivel), precedida por el numeral correspondiente (según la presente plantilla).

Appendix A

Appendix

Los Anexos son documentos o elementos que complementan el cuerpo de la tesis o trabajo de investigación y que se relacionan, directa o indirectamente, con el trabajo de grado, tales como acetatos, cd, normas, etc.

References

- Adorf, H.-M. (1995), Interpolation of irregularly sampled data series—a survey, *in* ‘Astronomical Data Analysis Software and Systems IV’, Vol. 77, p. 460.
- Dilmaghani, S., Henry, I. C., Soonthornnonda, P., Christensen, E. R. & Henry, R. C. (2007), ‘Harmonic analysis of environmental time series with missing data or irregular sample spacing’, *Environmental science & technology* **41**(20), 7030–7038.
- Elorrieta, F., Eyheramendy, S. & Palma, W. (2019), ‘Discrete-time autoregressive model for unequally spaced time-series observations’, *Astronomy & Astrophysics* **627**, A120.
- Eyheramendy, S., Elorrieta, F. & Palma, W. (2018), ‘An irregular discrete time series model to identify residuals with autocorrelation in astronomical light curves’, *Monthly Notices of the Royal Astronomical Society* **481**(4), 4311–4322.
- Ghaderpour, E. & Vujadinovic, T. (2020), ‘Change detection within remotely sensed satellite image time series via spectral analysis’, *Remote Sensing* **12**(23), 4001.
- Kuttu, S. & Bokpin, G. A. (2017), ‘Feedback trading and autocorrelation patterns in sub-saharan african equity markets’, *Emerging Markets Finance and Trade* **53**(1), 213–225.
- Liu, M., Stella, F., Hommersom, A., Lucas, P. J., Boer, L. & Bischoff, E. (2019), ‘A comparison between discrete and continuous time bayesian networks in learning from clinical time series data with irregularity’, *Artificial intelligence in medicine* **95**, 104–117.
- Mudelse, M. (2014), *Climate Time Series Analysis: Classical Statistical and Bootstrap Methods*, Springer, Berlin, Heidelberg.
- Novoseltseva, M., Gutova, S. & Glinchikov, K. (2019), Spatial and econometric modeling of the demographic situation in the kemerovo region, *in* ‘IOP Conference Series: Earth and Environmental Science’, Vol. 272, IOP Publishing, p. 032155.

- Ojeda, C., Palma, W., Eyheramendy, S. & Elorrieta, F. (2023), A novel first-order autoregressive moving average model to analyze discrete-time series irregularly observed, *in* O. Valenzuela, F. Rojas, L. J. Herrera, H. Pomares & I. Rojas, eds, 'Theory and Applications of Time Series Analysis and Forecasting', Springer International Publishing, Cham, pp. 91–103.
- Rindorf, A., Cadigan, N., Howell, D., Eero, M. & Gislason, H. (2020), 'Periodic fluctuations in recruitment success of atlantic cod', *Canadian Journal of Fisheries and Aquatic Sciences* **77**(2), 236–246.
- Robinson, P. (1977), 'Estimation of a time series model from unequally spaced data', *Stochastic Processes and their Applications* **6**(1), 9–24.
- Shamsan, A., Wu, X., Liu, P. & Cheng, C. (2020), 'Intrinsic recurrence quantification analysis of nonlinear and nonstationary short-term time series', *Chaos: An Interdisciplinary Journal of Nonlinear Science* **30**(9), 093104.
- Yang, Q., Fowler, M. S., Jackson, A. L. & Donohue, I. (2019), 'The predictability of ecological stability in a noisy world', *Nature ecology & evolution* **3**(2), 251–259.