

PROCESSOS K-FACTOR GARMA: PROPRIEDADES, SIMULAÇÕES E APLICAÇÃO

Data de submissão: 13/09/2023

Data de aceite: 02/10/2023

Cleber Bisognin

Departamento de Estatística – UFSM
Santa Maria – RS
<https://orcid.org/0000-0002-0714-3739>

Caroline Lopes Gonçalves

Departamento de Estatística – UFSM
Santa Maria- RS
<https://orcid.org/009-0000-6339-069X>

Keler Eliana Severo Correa

Departamento de Estatística – UFSM
Santa Maria- RS
<https://orcid.org/0000-0002-9373-6281>

RESUMO: O principal objetivo deste trabalho é estudar os processos -Factor GARMA. Esses processos exibem a característica de longa dependência e são mais gerais do que os processos ARFIMA e SARFIMA. Apresentamos resultados significativos relacionados às condições de estacionariedade, invertibilidade, representações autorregressivas e de médias móveis infinitas, sua função densidade espectral e o seu comportamento próximo às frequências de Gegenbauer, além do cálculo da previsão envolvendo esses processos. Também apresentamos um estimador para as frequências

de Gegenbauer e dois estimadores paramétricos para os parâmetros do processo, a saber, Whittle e CSS. Por meio de simulações de Monte Carlo, investigamos o desempenho dos estimadores paramétricos, nos quais constatamos que esses estimadores são não tendenciosos e consistentes. Finalmente, conduzimos uma análise da série temporal do Índice de Oscilação Sul, que é um indicador do estado da oscilação sul. Essa análise revela valores negativos do índice SOI, que estão associados ao fenômeno El Niño.

PALAVRAS-CHAVE: Série temporal, processos -Factor GARMA, longa dependência, estimação, aplicação

K-FACTOR GARMA PROCESSES: PROPERTIES, SIMULATIONS, AND APPLICATIONS

ABSTRACT: The main objective of this work is to study k-Factor GARMA processes. These processes exhibit the characteristic of long dependence and are more general than ARFIMA and SARFIMA processes. We present significant results related to stationarity, invertibility conditions, infinite autoregressive and moving average representations, their spectral density

function, and their behavior near Gegenbauer frequencies. We also provide an estimator for Gegenbauer frequencies and two parametric estimators for process parameters, namely Whittle and CSS. Through Monte Carlo simulations, we investigate the performance of the parametric estimators, confirming that these estimators are unbiased and consistent. Finally, we conduct an analysis of the time series of the Southern Oscillation Index (SOI), an indicator of the Southern Oscillation state. This analysis reveals negative values of the SOI index, which are associated with the El Niño phenomenon.

KEYWORDS: Time series, -Factor GARMA processes, long-range dependence, estimation, application

1 | INTRODUÇÃO

Atualmente, com o avanço da tecnologia e o aumento constante na disponibilidade de dados, a análise de séries temporais tem se tornado cada vez mais relevante e inovadora. De maneira geral, as séries temporais desempenham um papel crucial na análise de dados, permitindo a compreensão de tendências, integração precisa de informações temporais e embasamento para tomada de decisões. É uma área em constante evolução, à medida que novos dados emergem e novos métodos surgem para lidar com os desafios apresentados por dados temporais.

Nas últimas décadas, um estudo tem sido realizado com séries temporais que exibem propriedades de longa dependência, cujas características foram introduzidas por Hurst (1951) durante sua pesquisa relacionada à série temporal dos níveis mensais do rio Nilo. Posteriormente, Mandelbrot e Wallis (1969) definiram um processo que exibia o mesmo efeito pesquisado por Hurst, chamando-o de Ruído Gaussiano Fracionário, sendo uma versão do Movimento Browniano Fracionário a tempo discreto.

Em seguida, Granger e Joyeux (1980) e Hosking (1981 e 1984) propuseram o modelo autorregressivo fracionário integrado de média móvel, conhecido como AFIRMA(p, d, q), no qual a função densidade espectral do processo estocástico se torna ilimitada na frequência zero. Esse modelo é uma extensão dos modelos autorregressivos integrados conhecidos como ARIMA(p, d, q), propostos por G.E.P. Box e G.M. Jenkins, onde, $d \in \mathbb{N}$.

Adicionalmente, os modelos Gegenbauer, identificados como Gegenbauer(λ, u), e o modelo auto-regressivo de média móvel Gegenbauer, identificado como GARMA(p, λ, u, q), foram introduzidos por Gray, Zhang e Woodward (1989) Esses modelos apresentam uma função densidade espectral que se torna ilimitada em uma frequência arbitrária $G \in (0, \pi]$, conhecida como frequência de Gegenbauer.

Muitas séries temporais, além de exibirem características de longa dependência, também apresentam padrões sazonais. Nesse contexto, Porter-Hudak (1990) propôs o modelo SARFIMA $(0, D, 0)_s$ como uma nova abordagem, onde s é a sazonalidade. Porter-Hudak (1990) utiliza o estimador proposto por Geweke e Porter-Hudak (1983) para estimar o parâmetro de diferenciação sazonal D . Hassler (1994) emprega o estimador da

função periodograma em torno de todas as frequências sazonais para analisar modelos com abordagens flexíveis e adaptáveis usando o modelo SARFIMA $(0, D, 0)_s$. Montanari, Rosso e Taqqu (2000) utiliza o estimador de máxima verossimilhança aproximada para estimar os parâmetros do modelo SARFIMA $(p,d,q) \times (P,D,Q)$. Brietzke, Lopes e Bisognin (2005) investiga diversas propriedades do modelo SARFIMA $(0, D, 0)_s$, incluindo funções de autocovariância e autocorrelação parcial, bem como invertibilidade estacionariedade.

Woodward, Cheng e Gray (1998) generalizam os processos GARMA (p,λ,u,q) propondo os processos k-Factor GARMA (p,λ,u,q) . Nos processos k-Factor GARMA a função densidade espectral torna-se ilimitada em k frequências arbitrárias $G_j \in (0,\pi]$, para $j=1,\dots,k$, chamadas frequências de Gegenbauer. As frequências de Gegenbauer podem ou não coincidir com alguma frequência de Fourier. Estes processos são uma generalização dos processos SARFIMA $(p,d,q) \times (P,D,Q)_s$ cuja função densidade espectral torna-se ilimitada nas frequências ditas de Fourier, isto é, $w_j=2\pi j/s$, para $j=1,\dots,[s/2]$, onde $[x]$ é a parte inteira de x .

O objetivo deste trabalho é estudar os processos k-Factor GARMA (p,λ,u,q) , suas propriedades, estimação de parâmetros, previsão e aplicação.

Na Seção 2 apresentamos os processos k-Factor GARMA (p,λ,u,q) e suas propriedades tais como estacionariedade, inversibilidade, função densidade espectral e previsão. Na Seção 3 são apresentados uma metodologia para estimação das frequências de Gegenbauer e dois estimadores para os parâmetros dos k-Factor GARMA (p,λ,u,q) . Algumas simulações de Monte Carlo são apresentadas na Seção 4 e uma aplicação é apresentada na Seção 5. Na Seção 6, apresentamos as conclusões.

2 | METODOLOGIA

Os processos ARFIMA (p,d,q) , onde d está no intervalo $(0,0.5;0.5)$, podem ser considerados uma extensão dos processos ARIMA (p,d,q) , onde d é um número inteiro. Tais processo são destinados a modelar dados com longa dependência, caracterizada pela função densidade espectral que é ilimitada na frequência zero. De forma semelhante, os processos GARMA (p,λ,u,q) são vistos como uma generalização dos processos ARFIMA (p,d,q) , onde sua função densidade espectral se torna ilimitada em alguma frequência G dentro do intervalo $(0,\pi]$, não necessariamente na frequência zero. No entanto, tanto os processos ARFIMA (p,d,q) quanto o mais abrangente processo GARMA (p,λ,u,q) apresentam uma limitação: suas funções densidade espectrais se tornam ilimitadas em apenas uma frequência dentro do intervalo $(0,\pi]$. Para abordar essa limitação, Giraitis e Leipus (1995), seguidos por Woodward, Cheng e Gray (1998), estenderam os modelos GARMA para criar os modelos k-Factor GARMA (p,λ,u,q) . Nesses modelos, a função densidade espectral é ilimitada em um número finito k de frequências, chamadas de frequências de Gegenbauer (ou frequências G), dentro do intervalo $(0,\pi]$. A Definição 1 apresenta os processos k-Factor

GARMA(p, λ, u, q), e informações adicionais sobre esses modelos podem ser encontradas em Giraitis e Leipus (1995) e Woodward, Cheng e Gray (1998).

Definição 1 - Seja $\{x_t\}_{t \in \mathbb{Z}}$ um processo estocástico que satisfaz a equação

$$\phi(B) \prod_{j=1}^k (1 - 2u_j B + B^2)^{\lambda_j} (X_t - \mu) = \theta(B) \varepsilon_t, \quad (1)$$

em que k é um número inteiro, $|u_j| \leq 1$, λ_j é um número fracionário, para $j=1, \dots, k$, em que $k=1, 2, \dots, \mu$, μ é a média do processo, $\{\varepsilon_t\}_{t \in \mathbb{Z}}$ é um processo ruído branco e $\phi(\cdot)$ e $\theta(\cdot)$ são os polinômios de grau p e q dados, respectivamente, por

$$\phi(z) = \sum_{l=0}^p (-\phi_l) z^l \quad e \quad \theta(z) = \sum_{m=0}^q (-\theta_m) z^m, \quad (2)$$

com $\phi_l, 1 \leq l \leq p$, e $\theta_m, 1 \leq m \leq q$, constantes reais e $\phi_0 = 1 - \theta_0$. Então, $\{x_t\}_{t \in \mathbb{Z}}$ é um processo k -Factor-GARMA de ordem (p, λ, u, q) , denotado por k -Factor GARMA(p, λ, u, q), onde $u = (u_1, \dots, u_k)'$ e $\lambda = (\lambda_1, \dots, \lambda_k)'$.

Neste trabalho estamos interessados em estudar os processos k -Factor GARMA(p, λ, u, q), apresentados na Definição 1, onde $\{\varepsilon_t\}_{t \in \mathbb{Z}}$ é um processo ruído branco gaussiano com $E(\varepsilon_t) = 0$ e $\text{Var}(\varepsilon_t) = \sigma^2$.

No Teorema 1 a seguir, apresentamos alguns resultados sobre k -Factor GARMA(p, λ, u, q) estabelecidos e provados em Giraitis e Leipus (1995) e Woodward, Cheng e Gray (1998).

Teorema 1 - Seja $\{x_t\}_{t \in \mathbb{Z}}$ um processo k -Factor GARMA (p, λ, u, q). Então as seguintes afirmações são verdadeiras.

- (i) o processo $\{x_t\}_{t \in \mathbb{Z}}$ é estacionário se todas as raízes da equação $\phi(z)=0$ estão fora do círculo unitário. Além disso, u_j são distintos e $\lambda_j < \frac{1}{2}$, quando $|u_j| < 1$, e $\lambda_j < \frac{1}{4}$, quando $|u_j|=1$, para $j=1, \dots, k$;
- (ii) o processo estacionário $\{x_t\}_{t \in \mathbb{Z}}$ possui longa dependência se satisfaz as condições do item i) desta proposição e, além disso, $\lambda_j > 0$, para $1 \leq j \leq k$;
- (iii) o processo estacionário $\{x_t\}_{t \in \mathbb{Z}}$ é causal se e somente se $\phi(z) \neq 0$, para $|z| \leq 1$;
- (iv) o processo estacionário $\{x_t\}_{t \in \mathbb{Z}}$ é inversível se e somente se $\theta(z) \neq 0$, para $|z| \leq 1$;
- (v) sob as condições dos itens (ii) e (iii) as representações média-móvel (MA(∞)) e autorregressivas (AR(∞)) infinitas, respectivamente, são dadas por

$$\psi(z) = \sum_{l \geq 0} \psi_l z^l = \frac{\theta(z)}{\phi(z)} \prod_{j=1}^k (1 - 2u_j z + z^2)^{-\lambda_j}. \quad (3)$$

e

$$\pi(z) = \sum_{l \geq 0} \pi_l z^l = \frac{\phi(z)}{\theta(z)} \prod_{j=1}^k (1 - 2u_j z + z^2)^{\lambda_j}. \quad (4)$$

satisfazendo $\sum_{l \in \mathbb{Z}_{\geq}} |\psi_l| < \infty$ e $\sum_{l \in \mathbb{Z}_{\geq}} |\pi_l| < \infty$.

(vi) seja $\{x_t\}_{t \in \mathbb{Z}}$ um processo k-Factor GARMA (p, λ, u, q) estacionário. Então a função densidade espectral do processo $\{x_t\}_{t \in \mathbb{Z}}$ é dada por

$$f_X(\omega) = \frac{\sigma_\varepsilon^2 |\theta(e^{-i\omega})|}{2\pi |\phi(e^{-i\omega})|} \prod_{j=1}^k [2(\cos(\omega) - u_j)]^{-2\lambda_j}, \quad (5)$$

onde $0 < \omega \leq \pi$ e $G_j = \cos^{-1}(u_j)$ são chamadas frequências de Gegenbauer. A expressão assintótica da função densidade espectral é dada por

$$f_X(\omega) \sim D(\ell) |w - G_\ell|^{-2\lambda_\ell}, \text{ quando } w \rightarrow G_\ell \quad (6)$$

para $\ell=1, \dots, k$, onde

$$D(\ell) = \begin{cases} \frac{\sigma_\varepsilon^2}{2\pi} |2 \sin(G_\ell)|^{-2\lambda_\ell} \prod_{\substack{j=0 \\ j \neq \ell}}^k [2|\cos(G_\ell) - \cos(G_j)|]^{-2\lambda_j}, & \text{para } 0 < G_\ell < \pi, \\ \frac{\sigma_\varepsilon^2}{2\pi} \prod_{\substack{j=0 \\ j \neq \ell}}^k [2|\cos(G_\ell) - \cos(G_j)|]^{-2\lambda_j}, & \text{para } G_\ell = 0 \text{ ou } G_\ell = \pi. \end{cases} \quad (7)$$

A Figura 1 apresenta alguns exemplos da função densidade espectral dos processos k-Factor GARMA (p, λ, u, q) , com $\lambda=(0,2;0,2)$, para $k=2$ e $\lambda=(0,2;0,2;0,3)$, para $k=3$ e diferentes valores para u e $p, q \in \{0,1\}$.

Maiores detalhes sobre os processos k-Factor GARMA (p, λ, u, q) podem ser encontrados em Giraitis e Leipus (1995), Woodward, Cheng e Gray (1998) e Ferrara e Guégan (2001).

Para os processos k-Factor GARMA (p, λ, u, q) invertíveis (ver Teorema 1), a equação de previsão é dada por

$$\hat{X}_t(h) = \sum_{j \geq 1} \pi_j \hat{X}_t(j-h), \quad (8)$$

em que $h \geq 1$ e $\{\pi_j\}_{j \geq 0}$, são os coeficientes da representação autoregressiva infinita do processo (ver Equação (4)). Pode-se perceber que $\hat{X}_t(h)$ pode ser expressa com uma soma ponderada das observações presentes e passadas $\{\hat{X}_t\}_{t \leq n}$, com n é o tamanho da série temporal. Maiores detalhes sobre previsão de processo estocásticos estacionários e invertíveis ver Bisognin e Lopes (2007).

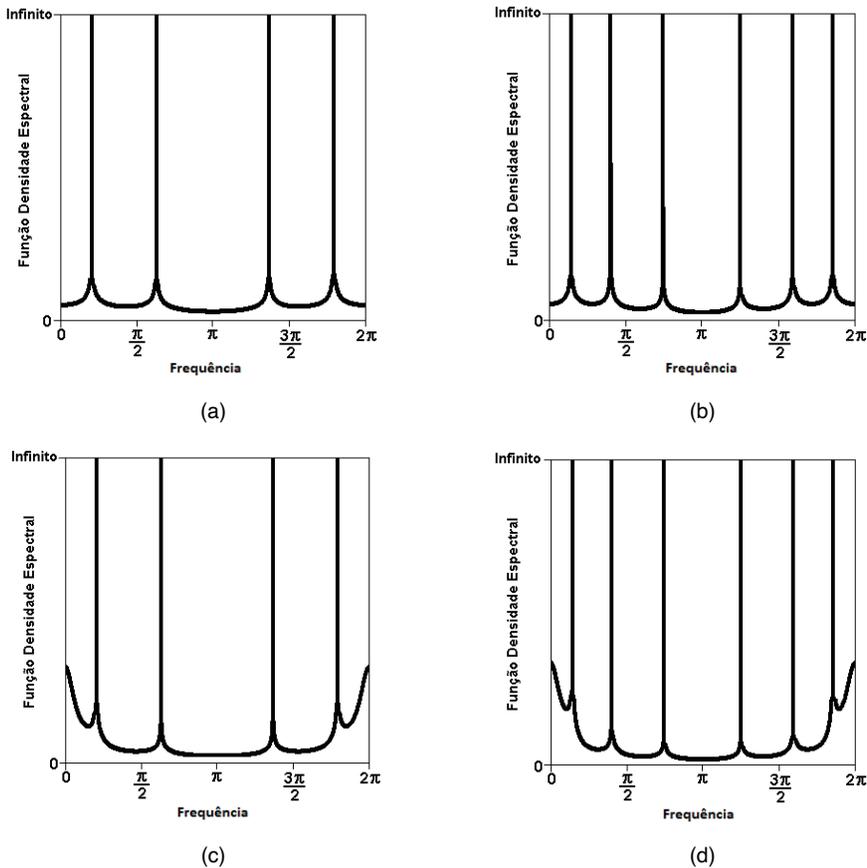


FIGURA 1 - Função densidade espectral dos processos k-Factor GARMA(p, u, λ, q), com $\lambda=(0,2;0,2)$, para $k=2$ e $\lambda=(0,2;0,2;0,3)$, para $k=3$: (a) $k=2$, $u=(-0,4; 0,8)$, $p=0=q$; (b) $k=3$, $u=(-0,7; 0,3; 0,9)$, $p=0=q$; (c) $k=2$, $u=(-0,4; 0,8)$, $p=1=q$, $\phi_1=0,8$ e $\theta_1=0,5$; (d) $k=3$, $u=(-0,7; 0,3; 0,9)$, $p=1=q$, $\phi_1=0,8$ e $\theta_1=0,5$. Fonte: Elaborada pelos Autores (2023).

Na seção a seguir, apresentamos os estimadores para os parâmetros dos processos k-Factor GARMA (p, λ, u, q).

3 I ESTIMAÇÃO DE PARÂMETROS

Na literatura de processos estocásticos existem diversos métodos de estimação propostos para estimar seus parâmetros. Estes métodos podem ser semiparamétricos, paramétricos e não-paramétricos. Neste trabalho consideramos estimadores na classe dos métodos semiparamétricos e paramétricos. Nesta seção apresentamos um estimador semiparamétrico para estimar o vetor de parâmetros $u=(u_1, \dots, u_k)'$, um estimador para estimar o vetor de parâmetros $\lambda=(\lambda_1, \dots, \lambda_k)'$ considerando as estimativas do vetor $u=(u_1, \dots, u_k)'$ e dois métodos paramétricos os quais estimam todos os parâmetros dos processos k-Factor GARMA (p, λ, u, q).

A seguir, apresentamos um estimador para as frequências de Gegenbauer de um processo k-Factor GARMA (p, λ, u, q) . Inicialmente definimos a função periodograma que é um estimador não viesado, mas não consistente, para a função densidade espectral de um processo estacionário.

Definição 2 – Seja $\{x_t\}_{t=1}^n$ uma série temporal obtida a partir do processo estocástico $\{x_t\}_{t \in \mathbb{Z}}$. A função periodograma, denotado por $I(\cdot)$, da série temporal $\{x_t\}_{t=1}^n$, é definida por

$$I(\lambda) = \frac{1}{2\pi n} \left| \sum_{t=1}^n X_t e^{-i\lambda t} \right|^2, \text{ para } -\pi \leq \lambda \leq \pi. \quad (9)$$

A função periodograma $I(\cdot)$ é um estimador não viesado, mas não consistente, para a função densidade espectral de um processo.

3.1 Estimador Semiparamétrico

Utilizando a função periodograma, Yajima (1996) propõe um estimador gráfico para as frequências de Gegenbauer. As frequências de Gegenbauer são aquelas em que o periodograma possui um máximo local.

Seja $\{x_t\}_{t \in \mathbb{Z}}$ um processo k-Factor GARMA (p, λ, u, q) dado pela Equação (1) cuja função densidade espectral possui k singularidades, nas frequências $G_j = \cos^{-1}(u_j)$, para $j=1, \dots, k$ as quais são chamadas de frequências de Gegenbauer. A função densidade espectral possui a seguinte expressão

$$f_X(w) = f_U(w) \prod_{j=1}^k [2(\cos(w) - u_j)]^{-2\lambda_j} \quad (10)$$

em que

$$f_U(w) = \frac{\sigma_\varepsilon^2 |\theta(e^{-iw})|^2}{2\pi |\phi(e^{-iw})|^2}, \text{ para todo } w \in (0, \pi]. \quad (11)$$

Os parâmetros u_j e λ_j , para todo $j=1, \dots, k$, satisfazem as condições da Teorema 1.

Seja $\{V_\nu\}_{\nu=1}^k$ uma sequência de intervalos definidos em que: para todo $\varepsilon > 0$ fixado e para todo $\nu=1, \dots, k$,

$$V_\nu = (G_\nu - a_\nu, G_\nu - \varepsilon) \cup (G_\nu + \varepsilon, G_\nu + b_\nu), \quad (12)$$

onde $\{a_\nu\}_{\nu=1}^k$ e $\{b_\nu\}_{\nu=1}^k$ são sequências reais, tais que $\{V_\nu\}_{\nu=1}^k$ é uma cobertura do intervalo $(0, \pi]$ quando $\varepsilon \rightarrow 0$, isto é,

$$\bigcup_{\nu=1}^k V_\nu = (0, \pi] \setminus \{G_1, \dots, G_k\}. \quad (13)$$

O estimador para as frequências de Gegenbauer G_j , denotado por \hat{G}_j , para $j=1, \dots, k$ é definido por

$$\hat{G}_j = \frac{2\pi}{n} \operatorname{argmax}_{w \in V_j} I(w) \quad (14)$$

Suposição 1 - As funções $f_x(\cdot)$ e $f_u(\cdot)$ satisfazem as seguintes afirmações:

- I. $f_x(\cdot)$ é uma função par;
- II. $f_u(\cdot)$ é uma função positiva no intervalo $[0, \pi]$;
- III. para todo $v=1, \dots, k$ e $w \in V_v$, $f_u(\cdot)$ satisfaz

$$\left| \frac{f'_u(w)}{f_u(w)} \right| = \mathcal{O}(|w - G_v|^{-1}), \quad (15)$$

para $w \in (0, \pi]$.

O seguinte Teorema 2 a seguir foi estabelecido por Yajima (1996).

Teorema 2 – Seja $\{x_t\}_{t \in \mathbb{Z}}$ um processo k-Factor GARMA (p, λ, u, q) dado pela Equação (1) cuja função densidade espectral satisfaz a Equação (5). Sob as condições estabelecidas na Suposição 1, para todo $\alpha \in (0, 1)$, quando $n \rightarrow \infty$, temos que

$$n^\alpha (\hat{G}_j - G_j) \xrightarrow{P} 0, \text{ para cada } j = 1, \dots, k. \quad (16)$$

A seguir, apresentamos os estimadores para o vetor parâmetro $\lambda = (\lambda_1, \dots, \lambda_k)$ para um processo k-Factor GARMA (p, λ, u, q) .

Seja $\{x_t\}_{t \in \mathbb{Z}}$ um processo k-Factor GARMA (p, λ, u, q) com média $\mu=0$, cuja função densidade espectral é dada pela Equação (5) e sua expressão assintótica pela Equação (6). Os parâmetros $u_j = \cos(G_j)$, para $j=1, \dots, k$, são previamente estimados por $\hat{u}_j = \cos(\hat{G}_j)$, onde \hat{G}_j , $j=1, \dots, k$, são dados através da Equação (14). Pela Equação (6), uma vez estimados os parâmetros \hat{G}_j a expressão assintótica da função densidade espectral para este processo é dada por

$$f_x(\omega) \sim D(\ell) |w - \hat{G}_\ell|^{-2\lambda_\ell}, \text{ quando } w \rightarrow G_\ell \quad (17)$$

com $D(\ell)$ dado pela Equação (7), substituindo G_ℓ por \hat{G}_ℓ .

Utilizando o método de estimação de Geweke e Porter-Hudak (1983) e Artech e Robinson (1998), temos que o estimador semiparamétrico para $\lambda = (\lambda_1, \dots, \lambda_k)$ é dado por

$$\log(I(w_v)) \sim c - \sum_{j=1}^k \lambda_j \ln[2(w_v - \hat{u}_j)]^2 + \varepsilon_v, \quad (18)$$

em que $B = \{0, 1, \dots, g(n) | v \neq v_j = \operatorname{argmax}_{w \in V_j} I(w), j = 1, \dots, k\}$ onde $g(n)$ é tal que $g(n) \rightarrow \infty$ e $\frac{g(n)}{n} \rightarrow 0$, quando $n \rightarrow \infty$ e $w_v = \frac{2\pi v}{n}$, para $v \in B$, são as frequências de Fourier. Denominamos este estimador por log-periodograma, denotado por LPR.

A segue definimos os estimadores de Whittle e CSS, respectivamente.

Estimador Whittle

Este estimador é baseado no estimador de Whittle (1951) utilizando o periodograma dado na Definição 2. Considere os processos k-Factor GARMA (p, λ, u, q) (ver Definição 1). Seja $\sigma_n^2(\boldsymbol{\eta})$ função de $\boldsymbol{\eta} = \{\boldsymbol{\phi}, \mathbf{u}, \boldsymbol{\lambda}, \boldsymbol{\theta}\}$, com $\boldsymbol{\phi} = \{\phi_1, \dots, \phi_p\}$, $\mathbf{u} = \{u_1, \dots, u_k\}$, $\boldsymbol{\lambda} = \{\lambda_1, \dots, \lambda_k\}$ e $\boldsymbol{\theta} = \{\theta_1, \dots, \theta_q\}$, dada por

$$\sigma_n^2(\boldsymbol{\eta}) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{I(\omega)}{f_X(\omega, \boldsymbol{\eta})} d\omega, \quad (19)$$

onde $I(\omega)$ é o periodograma normalizado (ver Definição 2) e $f_X(\omega, \boldsymbol{\eta})$ é a função densidade espectral (ver Equação 5). O valor de $\boldsymbol{\eta}$, denotado por $\hat{\boldsymbol{\eta}}$, que minimiza $\sigma_n^2(\boldsymbol{\eta})$ é o estimador para o verdadeiro valor de $\boldsymbol{\eta}$. Na prática, utilizarmos a soma nas frequências de Fourier, ao invés da integral, assim minimizamos a seguinte função com relação a $\boldsymbol{\eta}$

$$\hat{\sigma}_n^2 = \frac{1}{n} \sum_t \frac{I(\omega_t)}{f_X(\omega_t, \boldsymbol{\eta})} \quad (20)$$

em que $\omega_t = \frac{2\pi t}{n}$ são as frequências de Fourier, para $t \in Z, -\lfloor \frac{n}{2} \rfloor < t \leq \lfloor \frac{n}{2} \rfloor$ e $\lfloor x \rfloor$ é a parte inteira de x .

Estimador CSS

Este estimador denominado “soma de quadrados condicionais” foi proposto por Beran (1995) para os processos ARFIMA (p, d, q) . Posteriormente, Chung (1996) estende este método de estimação para os processos GARMA (p, λ, u, q) . Estendemos este método para os processos k-Factor GARMA (p, λ, u, q) . Chung (1996) sugere estimar o vetor de parâmetros desconhecidos $\boldsymbol{\eta}$ minimizando a soma dos resíduos ao quadrado, isto é, minimizar a expressão

$$S(\boldsymbol{\eta}) = -\frac{n}{2} \log(2\pi) - \frac{n}{2} \log(\sigma^2) - \frac{1}{2\sigma^2} \sum_{t=2}^n \hat{e}_t^2(\boldsymbol{\eta}), \quad (21)$$

em relação a $\boldsymbol{\eta}$, onde os resíduos são calculados por

$$\hat{e}_t^2(\boldsymbol{\eta}) = \sum_{j=0}^{t-1} \pi_j X_{t-j},$$

com $\{X_t\}_{t=1}^n$ a série temporal de interesse. Os coeficientes $\{\pi_j\}_{j \in Z_+}$ da representação auto regressiva infinita são dados pela Equação (4).

Na seção a seguir, apresentamos algumas simulações de Monte Carlo onde são avaliadas as propriedades dos estimadores Whittle e CSS para os parâmetros dos processos k-Factor GARMA (p, λ, u, q) .

4 | SIMULAÇÕES DE MONTE CARLO

Nesta seção apresentamos um estudo de simulações de Monte Carlo para avaliar

os estimadores de Whittle e CSS para os parâmetros dos processos k-Factor GARMA (ρ, λ, u, q).

Para tanto, consideramos um cenário com 3.000 replicações, diferentes tamanhos amostrais $n \in \{500, 1000\}$ Para o cenário fixamos os parâmetros em $k=2, \lambda=(0,2;0,4), u=(0,4;0,8), \rho=0=q$. Para avaliação dos estimadores utilizamos as seguintes medidas: média, viés, desvio padrão, viés relativo e erro quadrático médio. As simulações foram realizadas no software R(2023). Para as simulações de Monte Carlo, foram desenvolvidas rotinas próprias para a geração de amostras, a estimação e as medidas usadas na avaliação dos estimadores CSS e Whittle.

Estimadores	Whittle				CSS			
Parâmetros	u_1	u_2	λ_1	λ_2	u_1	u_2	λ_1	λ_2
Medidas	n=500							
Média	0,4034	0,9361	0,1987	0,3173	0,3993	0,8000	0,1994	0,4119
Viés	0,0034	0,1361	-0,0013	-0,0827	-0,0007	0,0000	-0,0006	0,0119
VR(%)	8,2227	4,0403	-0,7353	-6,6187	-3,7635	0,4243	-1,8015	47,3055
DP	0,0414	3,3680	0,1811	1,2502	0,0176	0,0048	0,0346	0,0252
EQM	0,0017	11,3619	0,0328	1,5699	0,0003	0,0000	0,0012	0,0008
Medidas	n=1000							
Média	0,4015	0,8010	0,2001	0,4006	0,3993	0,8000	0,1990	0,4128
Viés	0,0015	0,0010	0,0001	0,0006	-0,0007	0,0000	-0,0010	0,0128
VR(%)	19,2632	21,8566	1,5124	7,7055	-3,7670	-0,1418	-2,8374	51,7156
DP	0,0080	0,0046	0,0072	0,0075	0,0174	0,0048	0,0349	0,0248
EQM	0,0001	0,0000	0,0001	0,0001	0,0003	0,0000	0,0012	0,0008

TABELA 1 – Resultados das simulações de Monte Carlo para um processo -Factor GARMA, pelos métodos de Whittle e CSS, com $n \in \{500, 1000\}, k=2, \lambda=(0,2;0,4), u=(0,4;0,8), \rho=0=q$.

Fonte: Elaborada pelos Autores (2023)

Os resultados das simulações de Monte Carlo são apresentados na Tabela 1. Ambos os estimadores possuem um viés pequeno, o qual tende a diminuir à medida que o tamanho da amostra aumenta, indicando que os estimadores são não tendenciosos. O mesmo ocorre com o erro quadrático médio, o que indica que os estimadores são consistentes e precisos.

Na seção a seguir apresentamos a aplicação dos processos k-Factor GARMA (ρ, λ, u, q) a série temporal do Índice de Oscilação Sul (SOI).

5 | APLICAÇÃO

O Índice de Oscilação Sul (SOI) é um indicador do estado da Oscilação Sul, um fenômeno climático que afeta o clima global. O SOI é calculado com base nas diferenças de pressão do ar entre o oeste e o leste do Pacífico tropical. Valores negativos do SOI

estão associados a El Niño, um período de aquecimento anormal das águas do Oceano Pacífico tropical. Valores positivos do SOI estão associados a La Niña, um período de resfriamento anormal das águas do Oceano Pacífico tropical. Dada a importância do SOI, estamos interessados em prever seus os futuros valores de agosto de 2023 a julho de 2024, utilizando o período de janeiro de 1951 a julho de 2023.

A etapa da obtenção de dados consistiu em resgatar dados históricos do site do NOAA - *National Oceanic and Atmospheric Administration* (<https://www.nci.noaa.gov/access/monitoring/enso/soi>). A Figura 3 apresenta o gráfico e a função de autocorrelação amostral da série temporal que será alvo da parte do estudo aplicado neste trabalho.

Analisando o gráfico da série temporal e da função de autocorrelação amostral (ver Figura 2) nos leva a concluir que a série temporal é estacionária.

Com o propósito de verificar a estacionariedade da série temporal, foram empregados dois testes de raiz unitária: o Teste de Dickey-Fuller (ADF) e o Teste de Phillips-Perron (PP). Estes testes formulam as seguintes hipóteses: H_0 : A série possui uma raiz unitária, logo não é estacionária, em oposição a H_1 : A série não possui raiz unitária, portanto é estacionária. Ambos os testes resultaram em um valor-p $<0,01$, indicando que a série temporal é estacionária.

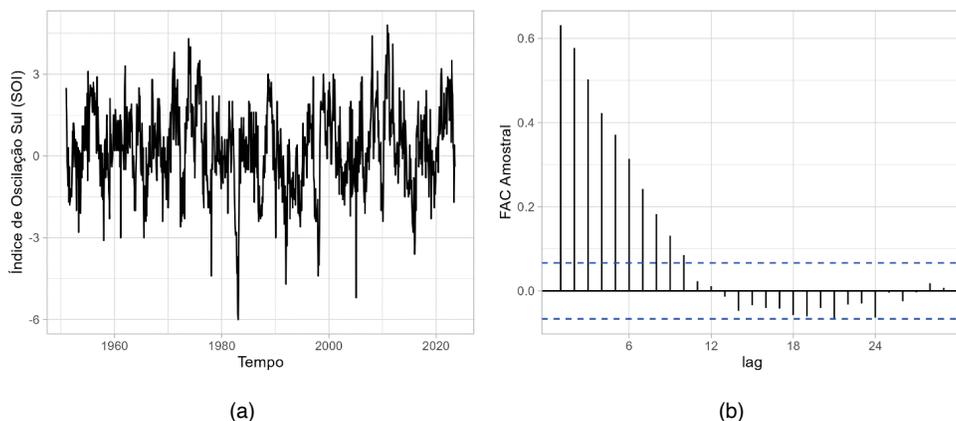


FIGURA 2 - Índice de Oscilação Sul (SOI), no período de janeiro de 1951 a julho de 2023: (a) gráfico da série temporal; (b) função de autocorrerlação amostral. Fonte: Elaborada pelos Autores (2023).

Foi conduzida uma análise de tendência determinística nos dados utilizando o teste de Cox-Stuart, com as seguintes hipóteses: H_0 : Não há presença de tendência determinística e H_1 : Existe tendência determinística nos dados. O resultado do teste teve um valor-p = 0,99, não rejeitando, dessa forma a hipótese nula e indicando não haver tendência determinística.

Para examinar a presença de sazonalidade na série temporal, foram empregados os testes de Kruskal-Wallis e Friedman, com o intuito de avaliar as seguintes hipóteses: H_0 :

Não há sazonalidade e H_1 : Existe sazonalidade na série temporal. Os testes apresentaram valor-p=0,2899 e 0,5793, respectivamente, ou seja, maior que o índice de significância (0). Com isso não se rejeita a hipótese de ausência de sazonalidade. Dado que a série temporal em questão é estacionária, podemos utilizar os modelos k-Factor GARMA (p, λ, u, q).

Na Figura 3 apresentamos o periodograma da série temporal em estudo. Podemos perceber que dois valores se destacam (maiores) e locais diferentes e não no tempo zero. Utilizando o estimador semiparamétrico, dado pela Equação (13), com $\alpha=0,8$, temos o indicativo de um processo k-Factor GARMA (p, λ, u, q), com $k=2$, em que $\hat{\lambda}_1=0,2144$, $\hat{u}_1=0,9896$ [$\hat{G}_1=0,1443$] e $\hat{\lambda}_2=0,1750$, $\hat{u}_2=0,9990$ [$\hat{G}_2=0,0433$] que satisfaz as condições do Teorema 1.

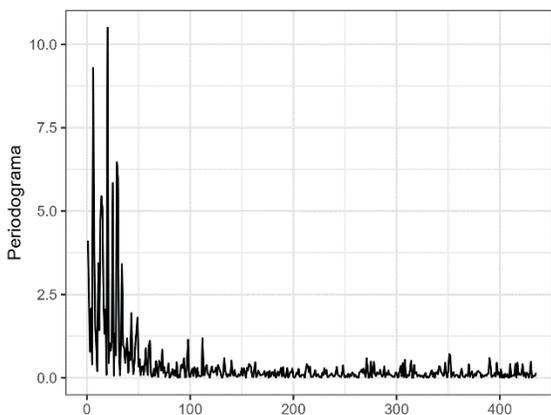


FIGURA 3 – Gráfico da Função Periodograma (ver Definição 2) da Série Temporal do Índice de Oscilação Sul (SOI), no período de janeiro de 1951 a julho de 2023. Fonte: Elaborada pelos Autores (2023).

Desta forma, conclui-se que um processo k-Factor GARMA é um processo adequado para a série temporal do Índice de Oscilação Sul (SOI). Assim, utilizando a função *garma*, do pacote de mesmo nome e utilizando o Critério de Informação de Akaike (AIC), o modelo mais adequado é o modelo k-Factor GARMA (p, λ, u, q), com $k=2$, $p=1$ e $q=0$. A Tabela 2 a seguir, apresenta os resultados do ajuste do processo.

A Tabela 3 e a Figura 4 apresentam as previsões para a série temporal do Índice de Oscilação Sul (SOI), para o período de agosto de 2023 a julho de 2024, calculada utilizando o modelo k-Factor GARMA (p, λ, u, q), com $k=2$, $p=1$ e $q=0$. Podemos perceber que todas as previsões geradas pelo modelo ajustado indicam um período do Índice de Oscilação Sul negativo, indicando uma atuação do fenômeno El Niño, ou seja, um período de aquecimento anormal das águas do Oceano Pacífico tropical.

Estimación Via CSS				
Parâmetro	Estimativa	Erro Padrão	Estatística	Pr(> z)
u_1	-0,8196	0,0026	-321,2100	<0,0001
λ_1	0,1735	0,0030	58,5390	<0,0001
u_2	-0,4003	0,0037	-107,0400	<0,0001
λ_2	0,0994	0,0087	11,4160	<0,0001
ϕ_1	0,8221	0,0209	39,3100	<0,0001
Estimación Via Whittle				
u_1	-0,8199	0,0049	-168,8158	<0,0001
λ_1	0,1666	0,0154	10,8477	<0,0001
u_2	-0,4013	0,0117	-34,3496	<0,0001
λ_2	0,1111	0,0207	5,3713	<0,0001
ϕ_1	0,8177	0,0162	50,3845	<0,0001

Log-Verossimilhança = -1357,0404

AIC=2726,081 e BIC= 8603,669

Teste de Normalidade de Shapiro-Wilk: p-valor= 0,1125

Teste de Ljung-Box (df=20): p-valor=0,1265

TABELA 2 – Modelo k-Factor GARMA (ρ, λ, u, q) ajustado a série temporal do Índice de Oscilação Sul (SOI), com $k=2$, $p=1$ e $q=0$.

Fonte: Elaborada pelos Autores (2023)

Ago/23	Set/23	Out/23	Nov/23	Dez/23	Jan/24
-0,4596	-0,2108	-0,0958	0,0305	-0,1670	-0,1179
Fev/24	Mar/24	Abr/24	Mai/24	Jun/24	Jul/24
0,0060	-0,0446	-0,1099	-0,0145	-0,2435	-0,2029

TABELA 3 – Previsão da Série Temporal do Índice de Oscilação Sul (SOI), para o período de agosto de 2023 a julho de 2024.

Fonte: Elaborada pelos Autores (2023)

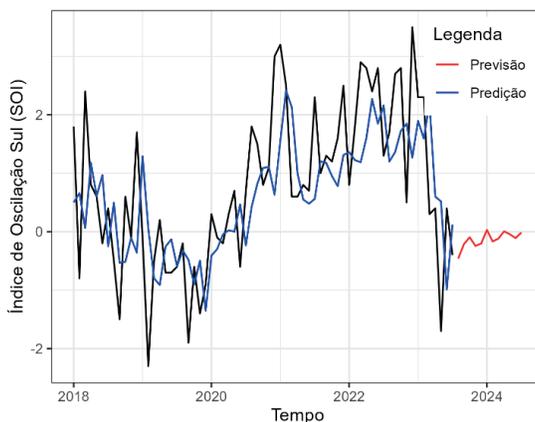


FIGURA 4 – Gráfico da Predição e da Previsão da Série Temporal do Índice de Oscilação Sul (SOI), para o período de agosto de 2023 a julho de 2024. Fonte: Elaborada pelos Autores (2023).

Foram calculadas as medidas de acurácia RMSE (raiz quadrada do erro quadrático médio) e MAE (erro médio absoluto). Os valores calculados foram 1,1461 e 0,9063, respectivamente, indicando um bom ajuste do processo processos k-Factor GARMA (p, λ, u, q) a série temporal SOI.

6 | CONCLUSÕES

Este trabalho teve como objetivo estudar os processos k-Factor GARMA (p, λ, u, q) . Inicialmente, foi apresentada a definição dos processos k-Factor GARMA (p, λ, u, q) , seguida das condições de estacionariedade, longa dependência, causalidade e inversibilidade. Também foram apresentadas suas representações média-móvel e autorregressiva infinitas, a função densidade espectral e seu comportamento próximo às frequências de Gegenbauer, além da equação de previsão.

Para a estimação dos parâmetros dos processos k-Factor GARMA (p, λ, u, q) , foram apresentados o periodograma, o estimador semiparamétrico para as frequências de Gegenbauer e os estimadores paramétricos de Whittle e CSS. Foram realizadas simulações de Monte Carlo para avaliar as propriedades dos estimadores paramétricos, nas quais foram encontrados indícios indicando que os estimadores são não tendenciosos, consistentes e apresentam distribuição assintoticamente normal.

Foi realizada uma aplicação dos processos k-Factor GARMA (p, λ, u, q) à série temporal do Índice de Oscilação Sul (SOI) para o período de agosto de 2023 a julho de 2024, utilizando uma série temporal mensal no período entre janeiro de 1951 a julho de 2023. Foram realizados testes de raiz unitária e tendência determinística para verificar a estacionariedade da mesma, além de testes para demonstrar que a série temporal não apresenta sazonalidade. Através do periodograma e do estimador semiparamétrico, verificou-se que um processo k-Factor GARMA (p, λ, u, q) era o mais adequado para descrever a série em estudo. Foram utilizados os estimadores Whittle e CSS para estimação dos parâmetros do modelo, bem como os critérios AIC e BIC para selecionar o melhor modelo. Também foram calculadas as previsões, as quais possuem magnitude negativa, indicando que o período previsto estará associado ao El Niño. Isso é corroborado pelas notícias divulgadas atualmente por diversos sites de meteorologia nacionais e internacionais.

REFERÊNCIAS

Arteche, J. ; Robinson, P. M. **Semiparametric inference in seasonal and cyclical long memory processes**. *Journal of Time Series Analysis*, v. 21, n. 1, p. 1-25, 1998.

Beran, J. **Statistics for long-memory processes**. *New York: Chapman & Hall*, 1994.

Bisognin, C. ; Lopes, S. R. C. **Estimating and forecasting the long memory parameter in the presence of periodicity**. *Journal of Forecasting*, v. 26, n. 6, p. 405-427, 2007.

Brietzke, E. H. M. ; Lopes, S. R. C. ; Bisognin, C. **A closed formula for the durbin-levinson's algorithm in seasonal fractionally integrated processes.** *Mathematical and Computer Modelling*, v. 42, n.12, p. 1191-1206, 2005.

Chung, C. F. .. **Estimating a generalized long memory process.** *Journal of Econometrics*, v. 73, p. 237-259, 1996.

Ferrara, L. ; Guégan, D. **Forecasting with k-factor gegenbauer processes: theory and applications.** *Journal of Forecasting*, v. 20, n. 8, p. 581-601, 2001.

Geweke, J. ; Porter-Hudak, S. **The estimation and application of long memory time series models.** *Journal of Time Series Analysis*, v. 4, n. 4, p. 221-238, 1983.

Giraitis, L. ; Leipus, R. **A generalized fractionally differencing approach in long-memory modeling.** *Lithuanian Mathematical Journal*, v. 35, n. 1, p. 53-65, 1995.

Granger, C. W. ; Joyeux, R. **An introduction to long-memory time series models and fractional differencing.** *Journal of Time Series Analysis*, v. 1, n. 1, p. 15-29, 1980.

Gray, H. L. ; Zhang, N. ; Woodward, W. A. **On generalized fractional processes.** *Journal of Time Series Analysis*, v. 10, n. 3, p. 233-257, 1989.

Hassler, U. **(Mis)specification of long memory in seasonal time series.** *Journal of Time Series Analysis*, v. 15, n. 1, p. 19-30, 1994.

Hosking, J. R. **Fractional differencing.** *Biometrika*, v. 68, n. 1, p. 165-176, 1981.

Hosking, J. R. **Modeling persistence in hydrological time series using fractional differencing.** *Water resources research*, v. 20, n. 12, p. 1898-1908, 1984.

Hurst, H. E. **Long-term storage capacity of reservoirs.** *Trans. Amer. Soc. Civil Eng*, v. 116, p. 770-808, 1951.

Mandelbrot, B. B. ; Wallis, J. R. **Computer experiments with fractional gaussian noises. part one, averages and variances.** *Water Resources Research*, v. 5, n. 1, p. 228-241, 1969.

Montanari, A.; Rosso, R. ; Taqqu, M. S.. **A seasonal fractional arima model applied to the Nile river monthly flows at aswan.** *Water Resources Research*, v. 36, n. 5, p. 1249-1259, 2000.

Woodward, W. A. ; Cheng, Q. C. ; Gray, H. L. **A k-factor gamma long-memory model.** *Journal of Time Series Analysis*, v. 19, n. 4, p. 485-504, 1998.

Porter-Hudak, S. **An application of the seasonal fractionally differenced model to the monetary aggregates.** *Journal of the American Statistical Association*, v. 85, n. 410, p. 338-344, 1990.

R Core Team (2023). **R: A language and environment for statistical computing.** R Foundation for Statistical Computing, Vienna, Austria. <<https://www.R-project.org/>>.

Whittle, P. **Hypothesis testing in time series analysis.** New York: Hafner, 1951.

Yajima, Y. **Estimation of the Frequency of Unbounded Spectral Densities.** *Proceedings of the Business and Economic Statistics Section 4-7, American Statistical Association, 1996.*