

Rebeca R. H. Pinafo, Patrícia R. Cirilo, Karen de Lolo G. Paulino, Elbert E. N. Macau

Programa de Pós-Graduação em Computação Aplicada - INPE  
Programa de Pós-Graduação Matemática Pura e Aplicada - UNIFESP

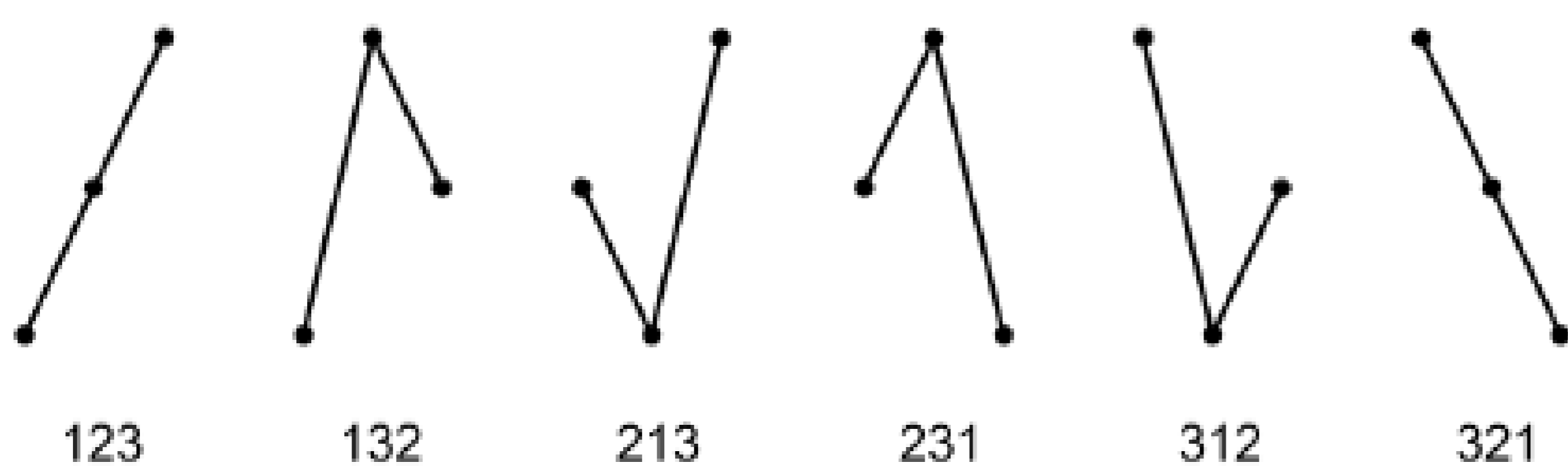
rebeca.pinafo@inpe.br, pcirilo@unifesp.com, klgpaulino@unifesp.br,  
elbert.macau@unifesp.br

## Motivação

Em 2020, o mundo se deparou com um cenário pandêmico com o surgimento da COVID-19, doença causada pelo vírus Sars-CoV-19. A pandemia apresentou em diversos lugares do mundo, uma dinâmica caracterizada por diferentes ondas de contágio, incluindo a região metropolitana do Rio de Janeiro. Este trabalho propõe entender a dinâmica dos óbitos por COVID-19 em diferentes ondas de contágio nesta região, pela perspectiva de medidas de complexidade com entropia de permutação [1].

## Desenvolvimento

Analisamos uma série temporal de óbitos acumulados de 890 dias e a subdividimos em 6 períodos diferentes, cada um correspondendo à uma onda de contágio de acordo com [3]. Os dados foram tratados com média móvel de sete dias para manejar subnotificações, e a frequência dos padrões de ordem de 3 símbolos como a distribuição de probabilidade necessária aos cálculos das medidas de complexidade e de suas funções componentes.



Padrões de ordem com três símbolos.

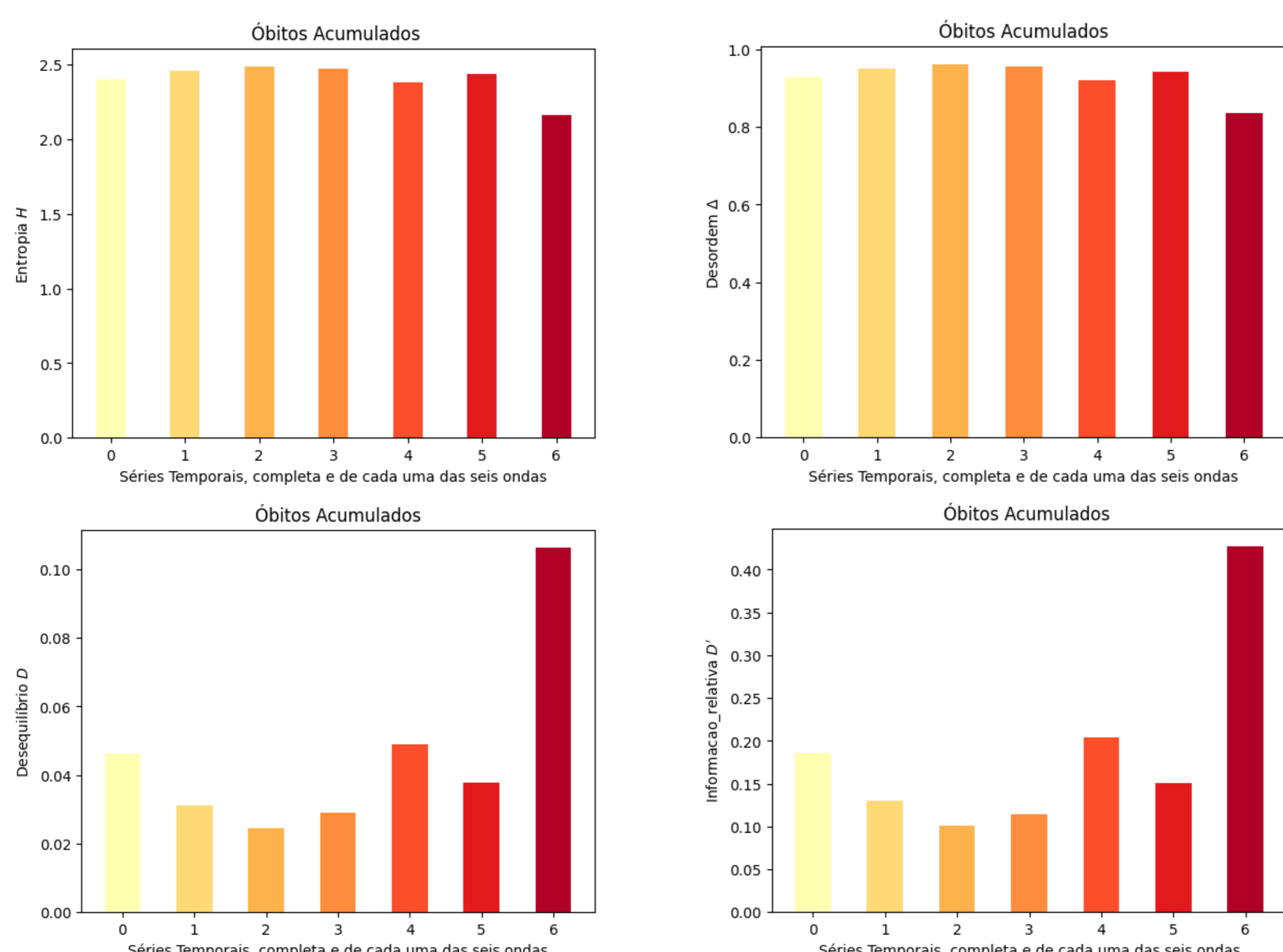
padrão $\pi$	123	132	213	231	312	321	n
Inteira	271	86	88	100	99	244	888
1ª Onda CA	24	15	16	10	9	30	104
2ª Onda CA	29	13	11	14	16	27	110
3ª Onda CA	24	14	16	13	12	34	113
4ª Onda CA	38	15	17	12	9	36	127
5ª Onda CA	38	16	16	19	18	47	154
6ª Onda CA	144	13	11	31	34	67	270

Frequência dos padrões de ordem na série temporal.

A entropia de permutação  $H(X) = \sum_{i=1}^N p_i \log_2(p_i)$ , desordem  $\Delta = \frac{H(X)}{H_{max}(X)}$ ,

o desequilíbrio  $D = \sum_{i=1}^N (p_i - \frac{1}{N})^2$  e a informação relativa  $D' = \sum_{i=1}^N p_i \log_2(\frac{p_i}{1/N})$

são as funções componentes das medidas de complexidade. As duas primeiras quantificam a informação armazenada pelo sistema dinâmico, enquanto as duas últimas estimam como a informação esta armazenada, tendo a distribuição de probabilidade uniforme como referência.



Histograma dos valores de entropia de permutação, desequilíbrio, informação relativa e desordem.

As medidas de complexidade são dadas por:

- Medida de Complexidade  $C_{LMC}$  [2]:

$$C_{LMC}(X) = H(X)D(X) \quad (1)$$

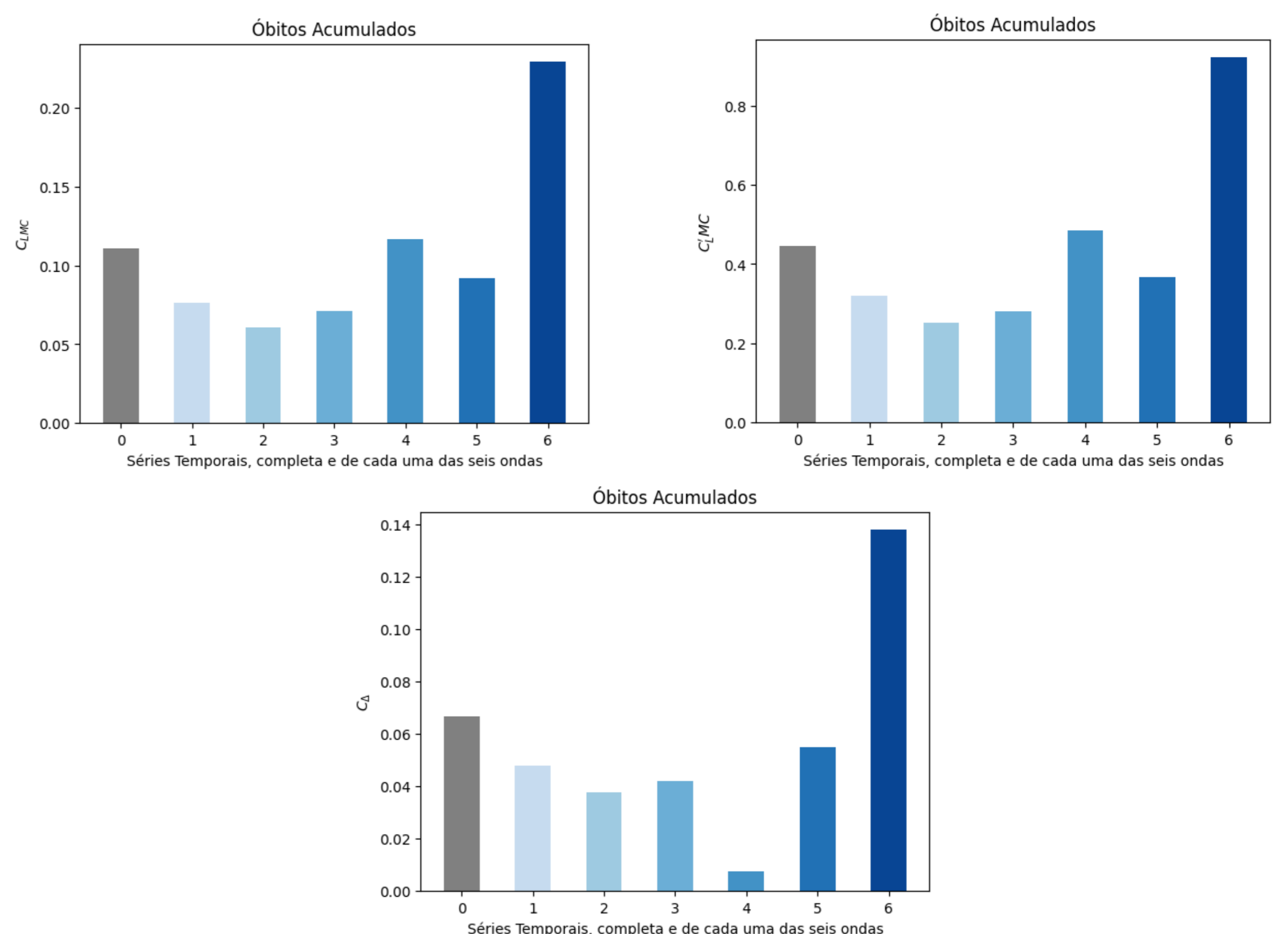
- Medida de Complexidade  $C'_{LMC}$  modificada [2]:

$$C'_{LMC}(X) = H(X)D'(X) \quad (2)$$

- Medida de Complexidade  $C_{\Delta}$  [4]:

$$C_{\Delta} = (1 - \Delta)^{k_1} \Delta^{k_2} \quad (3)$$

em que  $k_2, k_1$  são pesos escolhidos adequadamente de acordo com o contexto.



Histograma dos valores das medidas de complexidade  $C_{LMC}$ ,  $C'_{LMC}$  e  $C_{\Delta}$ .

## Resultados

No contexto destas séries temporais, a medida de complexidade indica a previsibilidade do seu comportamento. Assim, identificamos que durante a quarta onda, a série temporal se tornou mais previsível devido à queda nos números de mortes como resultado do início da campanha de vacinação.

## Agradecimentos

À Capes e CAP/INPE pelo apoio financeiro.

## References

- [1] Christoph Bandt. Small order patterns in big time series: A practical guide. *Entropy*, 21(6), 2019.
- [2] P Feldman David and P Crutchfield James. Measures of statistical complexity: Why? *Physics Letters A*, 238(4):244–252, 1998.
- [3] Paulo Roberto de Lima Gianfelice, Ricardo Sovek Oyarzabal, Americo Cunha, Jose Mario Vicensi Grzybowski, Fernando da Conceição Batista, and Elbert EN Macau. The starting dates of covid-19 multiple waves. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 32(3), 2022.
- [4] José Roberto Castilho Piqueira. A mathematical view of biological complexity. *Communications in Nonlinear Science and Numerical Simulation*, 14(6):2581–2586, 2009.