



UNIVERSIDADE ESTADUAL DE FEIRA DE SANTANA

Autorizada pelo Decreto Federal nº 77.496 de 27/04/76
Recredenciamento pelo Decreto nº 17.228 de 25/11/2016



PRÓ-REITORIA DE PESQUISA E PÓS-GRADUAÇÃO
COORDENAÇÃO DE INICIAÇÃO CIENTÍFICA

XXVII SEMINÁRIO DE INICIAÇÃO CIENTÍFICA DA UEFS **SEMANA NACIONAL DE CIÊNCIA E TECNOLOGIA - 2023**

DCCA Correlograma na Temperatura do Ar de Feira de Santana

Bianka Lima de Oliveira¹; Natália Oliveira Santos²; Gilney Figueira Zebende³

1. Bolsista PIBIC/FAPESB, Graduando em Engenharia Civil, UEFS, e-mail: biankaoliveira37@gmail.com
2. Bolsista PIBIC/CNPq, Graduando em Engenharia Civil, UEFS, e-mail: santosnat20002@gmail.com
3. Orientador, Departamento de Física, UEFS, e-mail: gfzebende@uefs.br

PALAVRAS-CHAVE: Temperatura do Ar; Feira de Santana; Correlograma.

INTRODUÇÃO

Sabemos que a autocorrelação é a correlação cruzada de uma série temporal com uma cópia defasada dela própria em função desta defasagem. Também, a análise de autocorrelação surge como uma ferramenta matemática e estatística na intenção de se descobrir padrões, como a presença de um sinal periódico obscurecido por um ruído. Diferentemente da análise de Fourier, este tipo de análise é usada no processamento de sinais no domínio do tempo. Diferentes campos de estudo definem a autocorrelação de maneira diferente e nem todas essas definições são equivalentes. Mas, a mais famosa é aquela que utiliza o coeficiente de Karl Pearson. Processos de raiz unitária, estacionários de tendência, autorregressivos e de média móvel são formas específicas de processos com autocorrelação, porém, tais análises clássicas são bem aplicadas em séries temporais estacionária, e na sua grande maioria tais séries são não estacionárias. Para tais casos, outras alternativas (métodos) de coeficientes deverão ser aplicadas, para a obtenção de resultados mais realistas. Nesta direção, aqui neste artigo propomos realizar uma análise estatística em torno do coeficiente de autocorrelação, ρ_{DCCA}^a , para análise de autocorrelação em dados da Temperatura do Ar na cidade de Feira de Santana (BA), por meio do seu correlograma.

METODOLOGIA

➤ Auto-correlação clássica de Pearson

A análise de séries temporais tem sido uma importante metodologia para a ciência investigar o comportamento ao longo do tempo de um objeto de estudo Box (2015). Logo, a partir de uma dada variável $\{u_t\}$ (por exemplo, a temperatura do ar), sucessivos valores podem ser medidos ao longo do tempo, formando uma série temporal, que muitas das vezes pode estar correlacionada entre si. Tal propriedade ou característica, pode ser analisada no domínio do tempo ou da frequência, por modelos e métodos estatísticos. O coeficiente proposto por Pearson (1896), baseia-se se neste tipo de análise. Assim, a autocorrelação clássica, introduzida pelo coeficiente de Pearson, pode ser representado por $\rho(\tau)$, definido como:

$$\rho(\tau) \equiv \frac{Cov(u_t, u_{t+\tau})}{\sqrt{Var(u_t) * Var(u_{t+\tau})}}$$

onde, τ representa a defasagem temporal, $Cov(u_t, u_{t+\tau})$ a covariância de u_t com $u_{t+\tau}$ e $Var(u_t)$ a sua variância. Como principal característica, tem-se que $-1 \leq \rho(\tau) \leq 1$, identificativo assim o sinal da autocorrelação em negativo/positivo (antipersistência/persistência).

Porém, para séries temporais não estacionárias (caso usualmente mais encontrado no dia a dia da pesquisa), essa análise proposta por Pearson não é robusta. Para analisar estas séries não estacionárias, métodos mais robustos deveriam ser empregados, tais como: o *Detrended Fluctuation Analysis-DFA* Peng (1994), o *Detrended Cross-Correlation analysis Analysis-DCCA* Podobnik (2008) e o *Detrended Cross-Correlation Coefficient, ρ_{DCCA}* Zebende (2011). Logo a seguir, apresentamos o novo coeficiente de autocorrelação baseado em robustos métodos estatísticos, que será aqui testado por meio de um correlograma em dados da Temperatura do Ar, para a cidade de Feira de Santana.

➤ Coeficiente de Autocorrelação sem Tendência - ρ_{DCCA}^a

Considere uma dada série temporal não estacionária, $\{u_t\}$, com $t = 1, 2, 3, \dots, N$ (tamanho total da série). Uma nova série temporal é criada (clonada) com os mesmos elementos da série original, porém com seus valores defasados, por um tempo τ . Logo, o coeficiente de autocorrelação sem tendência é definido segundo Zebende (2022) por:

$$\rho_{DCCA}^a(n, \tau) \equiv \frac{F_{u_t, u_{t+\tau}}^2(n)}{F_{u_t}(n) * F_{u_{t+\tau}}(n)}$$

As funções de flutuação $F_{u_t, u_{t+\tau}}^2(n)$, $F_{u_t}(n)$ e $F_{u_{t+\tau}}(n)$, são baseadas no cálculo do coeficiente de correlação cruzada, ρ_{DCCA} , proposto em Zebende (2011). Aplicação de alguns passos do algoritmo, foram exemplificados neste breve vídeo do YouTube:

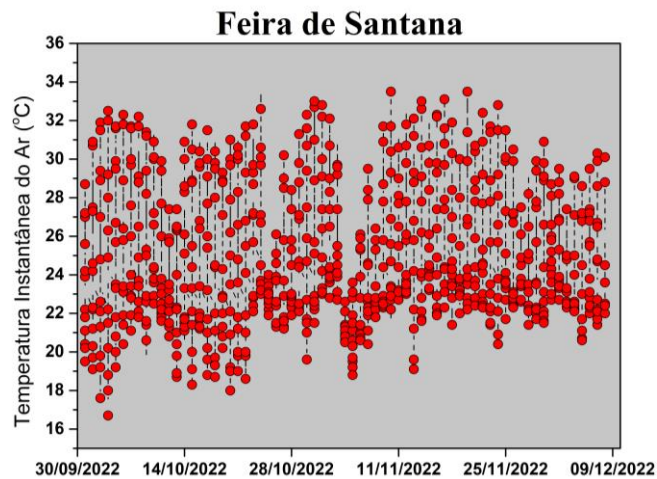
<https://www.youtube.com/watch?v=RQL7Db74yG0&t=216s>

O coeficiente, $\rho_{DCCA}^a(n, \tau)$ depende da escala temporal n (tamanho da caixa) e da defasagem temporal τ e seus valores variam entre $-1 \leq \rho_{DCCA}^a \leq 1$. O valor de $\rho_{DCCA}^a = 0$ significa que não há autocorrelação na série, caso contrário ρ_{DCCA}^a se divide em valores positivos e negativos.

Fechando assim a parte teórica, logo abaixo destacamos a obtenção de nossos dados a partir do INMET.

RESULTADOS E DISCUSSÃO

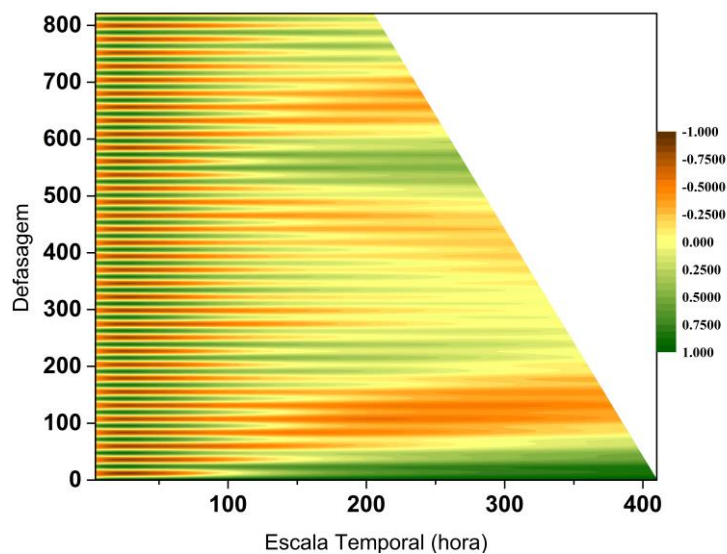
Aqui aplicaremos o correlograma em dados produzidos pelo Instituto Nacional de Meteorologia, INMET, que é o órgão federal do Brasil responsável por fornecer informações meteorológicas. Um dos itens disposto no Sistema de Coleta e Distribuição de Dados Meteorológico são as estações meteorológicas, onde as estações automáticas do INMET registram informações a cada hora, enquanto as estações convencionais coletam dados três vezes ao longo do dia. Neste artigo, selecionamos dados da temperatura do Ar (hora a hora) da estação automática, localizada na cidade de Feira de Santana (BA), entre os dias 01/10/2022 e 08/12/2022, ver fig.1 abaixo.



Legenda: 1. Dados da temperatura instantânea do ar utilizados para a análise

A partir de uma estatística descritiva breve vemos que a temperatura do ar tem um valor médio de $24,9^{\circ}C$, desvio padrão de $3,5^{\circ}C$, Assimetria de 0,49 e Curtose de -0,68. Tais valores demonstram ser esta uma distribuição onde a não estacionaridade se faz presente. Portanto, uma estatística mais robusta se mostra interessante para ser aplicada, como veremos a seguir com o DCCA correlograma.

O resultado para o DCCA correlograma pode ser visto logo na figura abaixo,



Legenda :2. Visualização da autocorrelação da temperatura do ar, por meio do DCCA correlograma

Aqui percebe-se que para pequenas escalas temporais há uma periodicidade nos valores de ρ_{DCCA}^a , independentemente da defasagem, que se perde para grandes escalas temporais.

CONSIDERAÇÕES FINAIS

Neste trabalho, propomos aplicar ρ_{DCCA}^a , com seu correlograma, em dados da temperatura do Ar para a cidade de Feira de Santana, com o intuito de se aprofundar mais no estudo dessa importante variável meteorológica. A utilização do DCCA correlograma tornou-se essencial, pois permite a obtenção de estatísticas mais robustas na análise da série temporal correlacionada entre si, possibilitando um progresso nas investigações

relacionadas à climatologia e no estudo previsões ou inferências sobre o comportamento de variáveis.

Os resultados nesta análise foram apresentados em um mapa de cores inédito para melhor visualização e entendimento.

Assim, por meio da modelagem apresentada neste artigo, é possível obter um maior entendimento sobre o comportamento de séries temporais, incluindo a identificação da presença ou ausência de um padrão de periodicidade.

REFERÊNCIAS

- Boris Podobnik, Zhi-Qiang Jiang, Wei-Xing Zhou, and H Eugene Stanley. *Statistical tests for power-law cross-correlated processes*. Physical Review E, 84(6):066118, 2011.
- C.-K. Peng, S. V. Buldyrev, S. Havlin, M. Simons, H. E. Stanley, and A. L. Goldberger. *Mosaic organization of DNA nucleotides*. Phys. Rev. E, 49:1685–1689, 1994.
- Chris Chatfield and Haipeng Xing. *The analysis of time series: an introduction with R*. CRC press, 2019.
- C-K Peng, SV Buldyrev, AL Goldberger, S Havlin, RN Mantegna, M Simons, and HE Stanley. *Statistical properties of DNA sequences*. Physica A: Statistical Mechanics and its Applications, 221(1-3):180–192, 1995.
- G. F. Zebende. *DCCA cross-correlation coefficient: Quantifying level of cross-correlation*. Physica A, 390(4):614–618, 2011.
- G.F. Zebende and A.M. da Silva-Filho. *Detrended multiple cross-correlation coefficient*. Physica A, 510:91 – 97, 2018. ISSN 0378-4371.
- George EP Box, Gwilym M Jenkins, Gregory C Reinsel, and Greta M Ljung. *Time Series Analysis: forecasting and control*. John Wiley & Sons, 2015.
- GF Zebende and EF Guedes. *Detrended correlogram method for non-stationary time-series analysis*. Fluctuation and Noise Letters, 21(02):2250012, 2022.
- GF Zebende, AA Brito, and AP Castro. *DCCA cross-correlation analysis in time-series with removed parts*. Physica A, 123472, 2019.
- Joseph Lee Rodgers and W. Alan Nicewander. *Thirteen ways to look at the correlation coefficient*. The American Statistician, 42(1):59–66, 1988. doi: 0.1080/00031305.1988.10475524.
- Karl Pearson. *Notes on the history of correlation*. Biometrika, 13(1):25–45, 1920.
- Karl Pearson. *Vii. mathematical contributions to the theory of evolution. iii. regression, heredity, and panmixia*. Philosophical Transactions of the Royal Society of London. Series A, containing papers of a mathematical or physical character, (187):253–318, 1896.
- Karl Pearson. *Vii. note on regression and inheritance in the case of two parents*. Proceedings of the Royal Society of London, 58(347-352):240–242, 1895.
- Ladislav Kristoufek. *Measuring correlations between non-stationary series with DCCA coefficient*. Physica A, 402:291–298, 2014.
- William T Vetterling, Saul A Teukolsky, William H Press, and Brian P Flannery. *Numerical recipes in C: the art of scientific computing, volume 1*. Cambridge university press, 2 edition, 1999. ISBN 0-521-43108-5.