



## MÉTODOS ESTATÍSTICOS APLICADOS A TEMPERATURA AMBIENTE E UMIDADE RELATIVA DO AR NA CIDADE DE PORTO SEGURO (BA)

<sup>1</sup>Andrea de Almeida Brito (IFBA / Senai Cimatec) – [britodea@gmail.com](mailto:britodea@gmail.com); <sup>2</sup>Dênio Oliveira Cruz (UNIJORGE) – [deniomelancia@gmail.com](mailto:deniomelancia@gmail.com); <sup>3</sup>Ivan Costa da Cunha Lima (Senai Cimatec) – [ivandacunhalima@gmail.com](mailto:ivandacunhalima@gmail.com); <sup>4</sup>Gilney Figueira Zebende (UEFS) – [gfzebende@hotmail.com](mailto:gfzebende@hotmail.com)

**Resumo:** Neste trabalho, estudamos a análise das variáveis climatológicas, temperatura e a umidade relativa do ar, através dos métodos da mecânica estatística. Adotamos os métodos DFA [1], o DCCA [2] e o Coeficiente de Correlação Cruzada  $\rho_{DCCA}$  [3]. A cidade de Porto Seguro, (estação meteorológica) localizada no Sul do Estado da Bahia (Brasil) foi o nosso caso de estudo. Nossos dados foram disponibilizados hora a hora pelo Instituto Brasileiro de Meteorologia (INMET). Calculamos a temperatura do ar média diária  $\langle T \rangle$  e a umidade relativa do ar média diária  $\langle U \rangle$  por duas formas distintas apresentadas no artigo [4]. Concluímos que as variáveis climatológicas têm uma persistência ao longo do tempo para o DFA e o DCCA e uma correlação cruzada negativa para a análise do coeficiente  $\rho_{DCCA}$ .

**Palavras-Chaves:** Variáveis Climatológicas; DFA; DCCA; Coeficiente de Correlação Cruzada  $\rho_{DCCA}$

## STATISTICAL METHODS APPLIED TO ENVIRONMENTAL TEMPERATURE AND RELATIVE AIR HUMIDITY IN THE CITY PORTO SEGURO (BA)

**Abstract:** In this work, we propose an analysis of the climatological variables, ambient temperature and the relative humidity of the air through the methods of statistical mechanics. For this we adopt methods DFA [1], DCCA [2] and the cross-correlation coefficient,  $\rho_{DCCA}$  [3]. We chose the city of Porto Seguro (meteorological station) located in the south of the state Bahia (Brazil) as our case study. Our data were made available hour by hour the Brazilian Institute of Meteorology (INMET). Calculate the average daily ambient temperature  $\langle T \rangle$ , by two different forms presented in article [4] and use same procedure to calculate the relative humidity of the daily average air  $\langle U \rangle$ . We noticed that the studied climatological variables have a presitence over time for DFA and DCCA and a negative cross correlation for coefficient analysis  $\rho_{DCCA}$ .

**Keywords:** Climatological Variables; DFA; DCCA; Cross Correlation Coefficient  $\rho_{DCCA}$



## 1. INTRODUÇÃO

O município de Porto Seguro (BA) situado no sul do estado da Bahia, esta a 707 km de distância da capital, Salvador. É, oficialmente, o primeiro município que chegaram os navegadores portugueses em território brasileiro. De acordo com os dados do Instituto Brasileiro de Geografia e Estatística (IBGE), em 2017, sua população era de 149 324 habitantes, tendo área é de 2 408,327 km<sup>2</sup> e a densidade demográfica de 62 hab./km<sup>2</sup>. Porto Seguro, que está a 4 metros acima do nível do mar, limita-se com os municípios de Santa Cruz de Cabrália e Prado. Sua localização geográfica é latitude 16° 27' 00"S e longitude 39° 3' 54"W. Possui um clima tropical, sendo um clima sempre quente com temperatura compensada média elevada variando 30° C a 38° C, chegando próximo dos 42° C e ameno no inverno, com média de 26° C e mínimas 20° C. Sua umidade relativa do ar média é de 94%.

Para contemplar o objetivo deste trabalho, a próxima seção aborda um referencial teórico dos métodos abordados nesta pesquisa. A terceira seção apresenta a metodologia. A quarta seção a análise e os resultados, e por fim, na quinta, apresentamos as conclusões.

## 2. REFERENCIAL TEÓRICO

Apresentaremos um resumo sobre os métodos estatísticos para análise de séries temporais não estacionárias. O DFA, o DCCA e do coeficiente de correlação cruzada  $\rho_{DCCA}$ .

### 2.1. Método DFA - Método da Autocorrelação

O método Detrended Fluctuation Analysis (DFA), desenvolvido por [1], análise das flutuações sem tendências, tem como objetivo identificar e mensurar autocorrelação de longo alcance em séries temporais não-estacionárias.

A modelagem consiste nos seguintes passos: De posse de uma série temporal  $u_i$ , com  $i$  variando de 1 a  $N$  ( $N$  é o número total de pontos da série), primeiro calculamos o desvio padrão de cada registro em relação à incidência média da série completa, integramos o sinal  $u_i$  e obtemos em relação à incidência média na série completa, desta forma, obtemos a série integrada  $y(k)$  pela expressão

$$y(k) = \sum_{i=1}^k [u(i) - \langle u \rangle], \quad k = 1, 2, \dots, N, \text{ em que } \langle u \rangle \text{ denota o valor médio de } u_i.$$

Em seguida, dividimos a série integrada  $y(k)$  em intervalos de igual amplitude  $n$  sobrepostos. Para cada intervalo de tamanho  $n$ , ajustamos a um polinômio de grau  $\geq 1$  que denotamos por  $y_n(k)$ . Esta estatística representa a tendência local na amplitude de tamanho  $n$ . Próximo passo, subtraímos a série integrada  $y(k)$  da série ajustada  $y_n(k)$ , em cada amplitude de tamanho  $n$ , e, após a subtração, calculamos

a função  $F_{DFA}(n)$  expressa pela equação 
$$F_{DFA}(n) = \sqrt{\frac{1}{N} \sum_{k=1}^N [y(k) - y_n(k)]^2}.$$



Este cálculo deve ser repetido para diferentes amplitudes de tamanho  $n$ , com  $4 < n < \frac{N}{4}$  (o valor mínimo e máximo de  $n$  foi definido empiricamente).

Por fim, verificamos se a função  $F_{DFA}(n)$  tem comportamento do tipo lei de potência, ou seja,  $F_{DFA}(n) \sim n^\alpha$ . Neste caso  $\alpha$  será o expoente de auto correlação de longo alcance.

Desta forma, se o expoente  $\alpha = 0,5$ , a série será aleatória, ou seja, decorrelacionada; se o expoente  $\alpha > 0,5$ , a série apresenta persistência de longo alcance e se o expoente  $\alpha < 0,5$ , a série apresenta um comportamento antipersistente de longo alcance.

## 2.2. Método DCCA - Método da Correlação Cruzada

Proposto por [2], o método Detrended Cross-Correlation Analysis (DCCA), que é a generalização do DFA, tem como objetivo analisar correlações cruzadas entre duas séries temporais distintas não estacionárias de mesmo tamanho  $N$ . A finalidade do DCCA é estimar a existência ou não de correlação cruzada de longo alcance entre as duas séries temporais analisadas, por meio da função de covariância sem tendência  $F^2_{DCCA}(n)$ . A análise duas séries temporais  $y_i$  e  $y'_i$  com o mesmo tamanho  $N$  pelo método DCCA é feita da seguinte maneira: Primeiramente, integramos as duas séries,  $y_i$  e  $y'_i$ , de acordo com as seguintes expressões  $R_k = S_1 + S_2 + S_3 + \dots + S_k$  e  $R'_k = S'_1 + S'_2 + S'_3 + \dots + S'_k$ , em que  $k = 1, 2, \dots, N$ . Em seguida, dividimos os sinais integrados  $R_k$  e  $R'_k$  em  $(N-n)$  boxes (com superposição) de tamanho  $n$  cada um, contendo  $n+1$  valores. Para as duas séries temporais, em cada box (que inicia em  $i$  e termina em  $n+i$ ) calculamos as tendências

em cada box, isto é,  $\tilde{R}_{k,i}$  e  $\tilde{R}'_{k,i}$ , ( $i < k < i+n$ ) como sendo a ordenada do ajuste linear (processo dos mínimos quadrados) da série somada. Próximo passo, a partir do cálculo anterior calculamos a covariância dos resíduos em cada intervalo, dada pela expressão

$f^2_{DCCA}(n,i) \equiv \frac{1}{(n+1)} \sum_{k=i}^{i+n} \left[ R_k - \tilde{R}_{k,i} \right] \left[ R'_k - \tilde{R}'_{k,i} \right]$ . E por último, calculamos

a função de correlação cruzada dada pela expressão

$f^2_{DCCA}(n,i) \equiv \frac{1}{(N-n)} \sum_{i=1}^{N-n} f^2_{DCCA}(n,i)$ . Dadas duas séries temporais de mesmo tamanho

$N$ , é possível obter o mesmo resultado do método DFA aplicado para uma única série temporal. Desta forma, quando repetido o procedimento descrito anteriormente para tamanhos diferentes de boxes  $N$ , podemos verificar a existência ou não da lei de potência, através da expressão  $f^2_{DCCA}(n) \sim n^{2\lambda}$  em que  $\lambda$  descreve o expoente de correlação cruzada de longo alcance. Assim, o expoente  $\lambda_{DCCA}$  quantifica a correlação cruzada entre duas séries temporais distintas e mesmo tamanho, tem-se como análise as seguintes características: para um expoente  $\lambda_{DCCA} = 0,5$ , as séries



são não apresentam correlação cruzada, ou seja, são descorrelacionadas; se o expoente  $\lambda_{DCCA} > 0,5$  as séries apresentam persistência de longo alcance e se o expoente  $\lambda_{DCCA} < 0,5$ , as séries apresentam comportamento antipersistente de longo alcance.

### 2.3. Coeficiente de Correlação Cruzada $\rho_{DCCA}$

O coeficiente de correlação cruzada,  $\rho_{DCCA}$ , proposto por [3], tem como objetivo quantificar e mensurar o nível de correlação cruzada entre duas séries temporais distintas e não estacionárias, tendo como base o DFA e o DCCA. O coeficiente de correlação cruzada,  $\rho_{DCCA}$ , é a razão entre a função de covariância sem tendência,  $F^2_{DCCA}$  e a função de variância sem tendência,  $F_{DFA}$ , definido por:

$$\rho_{DCCA}(n) \equiv \frac{F^2_{DCCA}(n)}{F_{DFA_{y_i}}(n)F_{DFA_{y_i}}(n)} \quad (1)$$

Este coeficiente define uma nova escala de correlação cruzadas em séries temporais não estacionárias. O coeficiente de correlação cruzada,  $\rho_{DCCA}$ , é um coeficiente adimensional com intervalo que varia entre  $-1 \leq \rho_{DCCA}(n) \leq 1$ . Se  $\rho_{DCCA} = 0$ , não existe correlação cruzada entre as séries, dividindo o nível de correlação cruzada entre positivo e negativo. Assim, caso  $\rho_{DCCA} > 0$  existe correlação cruzada entre as séries e será perfeita se  $\rho_{DCCA} = 1$ . Se  $\rho_{DCCA} < 0$  existe uma anticorrelação cruzada, sendo ela anticorrelação perfeita caso  $\rho_{DCCA} = -1$ .

Dentre os diversos trabalhos utilizando o coeficiente  $\rho_{DCCA}$ , nas diversas áreas científicas, podemos citar: [5], [6], [7], [8], [9], [10], [11], [12].

## 3. METODOLOGIA

Para analisar o comportamento das variáveis climatológicas (temperatura ambiente e umidade relativa do ar) através dos métodos da mecânica estatística, o DFA, o DCCA e o  $\rho_{DCCA}$ , da cidade de Porto Seguro, utilizamos 803 dados medidos hora a hora da estação meteorológica automática, disponibilizados pelo INMET referentes ao período de 12/07/2008 a 30/09/2010. De posse dos dados hora a hora, calculamos a temperatura ambiente média diária,  $\langle T \rangle$ , por duas formas distintas, tomando como base para o cálculo da temperatura ambiente média o exposto no artigo [4] e utilizamos a mesma forma para calcular a umidade relativa do ar média diária,  $\langle U \rangle$ .



### 3.1. Temperatura Ambiente e Umidade Relativa do Ar Média Diária Hora a Hora

$$\langle T_m \rangle = \frac{\sum_{i=1}^{24} T_i}{24} \text{ e } \langle U_m \rangle = \frac{\sum_{i=1}^{24} U_i}{24} \quad (2)$$

em que  $i$  é a hora,  $T_i$ , a temperatura na hora  $i$  e  $U_i$ , a Umidade relativa do ar na hora  $i$ .

### 3.2. Temperatura Ambiente e Umidade Relativa do Ar Média Diária Máx/Min

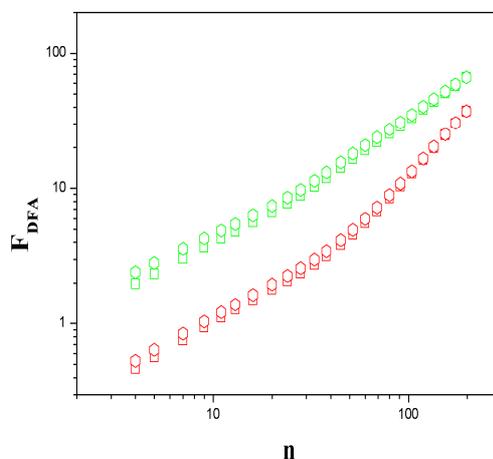
$$\langle T_{\text{máx / min}} \rangle = \frac{T_{\text{min}} + T_{\text{máx}}}{2} \text{ e } \langle U_{\text{máx / min}} \rangle = \frac{U_{\text{min}} + U_{\text{máx}}}{2} \quad (3)$$

em que  $T_{\text{min}}$ , é a temperatura ambiente mínima diária e  $T_{\text{máx}}$ , é a temperatura ambiente máxima diária do local e  $U_{\text{min}}$ , é a umidade relativa do ar mínima diária e  $U_{\text{máx}}$ , é a umidade relativa do ar máxima diária do local. A temperatura ambiente está medida em graus Celsius ( $^{\circ}C$ ) e umidade relativa do ar em valor percentual (%).

## 4. ANÁLISE E RESULTADOS

Para o método DFA, a Figura 1, mostra que as médias diárias calculadas pelas formas (2) e (3) propostas por [4], para a variável temperatura ambiente representada pela cor vermelha e a umidade relativa do ar, pela cor verde, tem um comportamento persistente e não estacionário nas séries temporais estudadas, ou seja,  $\alpha > 0,5$ .

Figura 1. Gráfico do DFA para a temperatura ambiente e umidade relativa do ar

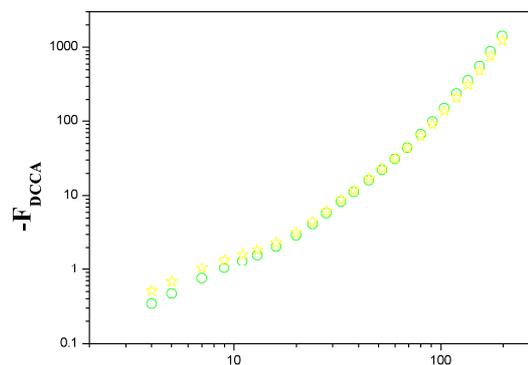


Na modelagem do expoente de correlação cruzada,  $\lambda_{DCCA}$ , observamos através da Figura 2, uma correlação cruzada negativa para todas as escalas temporais



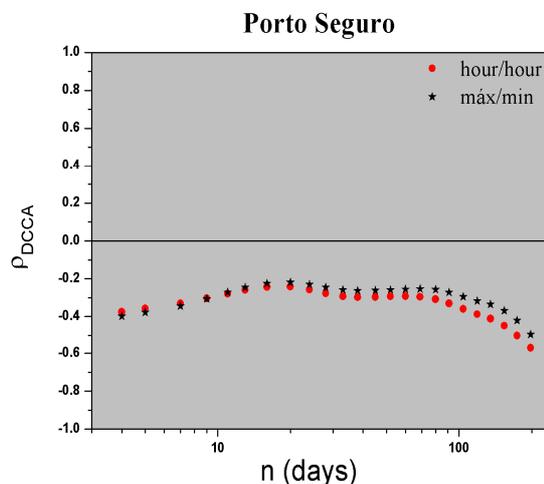
independente da forma que fora calculada a média diária das variáveis. Este resultado nos permite concluir que existe uma relação de antiperistência entre as duas variáveis climatológicas, pois a função  $F^2_{DCCA}$  é negativa ao longo do tempo.

Figura 2. Gráfico DCCA para a temperatura ambiente e umidade relativa do ar



Na Figura 3, apresentamos o gráfico com os coeficientes  $\rho_{DCCA}$ , calculados pelas formas (2) e (3) propostas por [4], e através desta percebemos que a correlação cruzada do ponto de vista do coeficiente  $\rho_{DCCA}$  é negativa para toda a escala de tempo n (dias) independente da forma em que a média diária das variáveis fora calculada.

Figura 3.  $\rho_{DCCA}$  aplicados aos valores médios diários da temperatura ambiente e umidade relativa do ar





#### 4. CONCLUSÕES

Após minerarmos os dados, calculamos as médias diárias da temperatura ambiente por duas maneiras distintas, equação (2) e (3), segundo [4]. Utilizamos a mesma técnica para o cálculo das médias diárias para a umidade relativa do ar. Com as médias calculadas procedemos à análise pelos métodos supracitados.

Para o método DFA, Figura 1, observou-se que para a cidade analisada, Porto Seguro (BA), as séries temporais (temperatura ambiente e umidade relativa do ar) o expoente de autocorrelação,  $\alpha_{DFA}$ , foi maior do que 0,5. Dessa forma, classificamos as séries como persistentes ao longo do tempo.

Para o expoente de correlação cruzada,  $\lambda_{DCCA}$ , Figura 2, identificou-se a existência de uma lei de potência para toda escala de tempo  $n$  (dias),  $-F^2_{DCCA} \sim n^{2\lambda}$ . A função de correlação cruzada,  $F^2_{DCCA}$ , apresentou-se negativa no período analisado.

E, com relação aos coeficientes  $\rho_{DCCA}$ , Figura 3, que tem por objetivo mensurar e quantificar o nível de correlação cruzada entre as séries temporais não estacionárias (temperatura ambiente e umidade relativa do ar) observou-se que os mesmos foram negativos para todas as escalas de tempo  $n$  (dias), implicando numa anticorrelação cruzada entre estas variáveis climatológicas, isto é, quanto maior (menor) a temperatura do ar, menor (maior) será a sua umidade relativa do ar.

#### Agradecimentos

A. A. BRITO e I. C. CUNHA LIMA, agradecem a FAPESB (Fundação de Amparo à Pesquisa do Estado da Bahia) (Grant BOL 0262/2017 e BOL 097/2016) e ao INMET (Instituto Brasileiro de Meteorologia) e G. F. ZEBENDE ao CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico) (Grant 304362-2017-4).

#### 5. REFERÊNCIAS

<sup>1</sup> PENG, C. K.; BULDYREV, V. SIMONS, M.; STANLEY, H. E. GOLDBERGER, L. **Mosaic organization of dna nucleotides.**, Physical Review E, v. 49, p. 1685-1689, 1994.

<sup>2</sup> PODOBNIK, B.; STANLEY, H. E. **Detrended cross-correlation analysis: A new method for analyzing two nonstationary times series**, PHYSICAL REVIEW LETTERS, v. 100, p. 084102, 2008.

<sup>3</sup> ZEBENDE, G. F. **Dcca cross-correlation coefficient: Quantifying level of cross-correlation**, Physica A, v. 390, p.614-618, 2011.



- <sup>4</sup> WEISS, A.; HAYS, C. J. **Calculating daily mean air temperature by different methods: implications from a non-linear algorithm**, Agricultural and Forest Meteorology, v.128, p. 57-65, 2005.
- <sup>5</sup> ZEBENDE, G. F.; SILVA, A. P.; DA SILVA FILHO, A. M. **Study of corss correlation in a self a\_ne time series of taxi accidents**. Physica A, v. 390, p.1677-1683, 2011.
- <sup>6</sup> MACHADO, A. SILVA, M.; ZEBENDE, G. F. **Autocorrelation and cross-correlation in time series of homicide and attempted homicide.**, Physica A, v. 400, p12-19, 2014.
- <sup>7</sup> G. F. ZEBENDE, A. A. BRITO; A. M. SILVA MACHADO; A. P. N. CASTRO. .; **Dcca applied betwewn air temperature and relative humidity: An hour/hour.**, Physica A, v.494, p. 17-26, 2018.
- <sup>8</sup> G. F. ZEBENDE, G. F.; VASSOLER, R. T **Dcca cross-correlation coefficient apply in time series of air temperature and air relative humidity.**, Physica A, v. 391, p. 2438-2443, 2012.
- <sup>9</sup> BRITO A. A.; MACHADO, A.; ZEBENDE, G. F. **Análise temporal de dados climatológicos de Feira de Santana (Ba), Conjuntura & Planejamento**, v. 189, p. 50-59, 2015.
- <sup>10</sup> ZEBENDE, G. F.; A. MACHADO, A. **Cross-correlation between time series of vehicles and passengers.**, Physica A v. 88 p. 4863-4866, 2009.
- <sup>11</sup> PODOBNIK, B.; JIANG, Z. -Q.; ZHOU,W. -X.; STANLEY, H. E. **Statistical tests for power-law cross-correlated processes.**, Physical Review E, v. 84 p. 066118, 2011.
- <sup>12</sup> ZEBENDE, G. F.; MACHADO, A.; SILVA, M. F. **Autocorrelation and cross-correlation in time series of homicide and attempted homicide.**, Physica A, v. 400 p. 12-19, 2014.