

08 a 11 de Outubro de 2018
Instituto Federal Fluminense
Búzios - RJ

A TIME SERIES STUDY OF INCIDENCE OF DENGUE IN MUNICIPALITIES OF BAHIA BY ONDALETAS HAAR

Vânia Gonçalves de Brito dos Santos¹ – vgbsantos@uneb.br

Stela Maria Azevedo e Ribeiro² – smaribeiro@uefs.br

¹ Universidade do Estado da Bahia, DCET-I – Salvador, BA, Brazil

² Universidade Estadual de Feira de Santana, DEXA – Feira de Santana, BA, Brazil

Abstract. Dengue is a complex public health problem common in tropical and subtropical regions. It is a viral infection transmitted through the bite of the mosquito *Aedes aegypti* (*A. aegypti*) and its incidence, in essence, has increased substantially in the last three decades. This work aims to implement an analysis via wavelets to represent and interpret data of daily incidence of dengue cases in three hundred and ninety five cities in the state of Bahia, Brazil, between the years 2000 and 2009, checking if the time series of the original data can be satisfactorily represented by the Haar wavelet approach, in order to have models defined by the coefficients of the series that parameterize them, and with these models, we can observe possible characteristics of similarities or self-affinities.

Keywords: Dengue, Wavelet Haar, Similarity.

1. INTRODUÇÃO

A dengue é considerada a mais importante das doenças virais transmitidas por artróodos, sendo também a mais comum e distribuída arbovirose do mundo e tem se destacado entre enfermidades reemergentes. O vírus da dengue tornou-se um problema de saúde pública internacional. Metade da população mundial está atualmente em risco de infecção por dengue. Aproximadamente 50 a 100 milhões de casos de infecção são registrados anualmente. Entre os infectados, 500 mil tinham infecção grave e a internação foi necessária, a maioria são crianças. Os dados de ocorrência nos municípios baianos chamam a atenção, devido ao seu comportamento recorrente e apresentando auto-similaridade na série temporal desses dados, como indica (Saba et al., 2014b).

Atualmente, com a chegada de duas novas doenças transmitidas pelo mesmo vetor, o *Aedes aegypti*, a preocupação se redobra. Essas doenças são a Chikungunya e a Zika. Apesar de possuírem o vetor em comum, os vírus são distintos. O vírus causador da chikungunya é o CHIKV vírus, e da zika, o ZIKV vírus. Além de vírus distintos, as doenças também são distintas, bem como suas consequências para população.

A transformada de Fourier é especialmente interessante para sinais estacionários, ou seja, para sinais que não variam substancialmente suas propriedades no tempo. Em trabalho recente propuseram uma análise da série temporal da dengue através do método Detrended Fluctuation Analysis – DFA, para análise de auto-afinidade em série não-estacionária, (Ribeiro, 2016). Por outro lado, a análise por ondaletas vem sendo bastante utilizada, a exemplo da parametrização de dados de velocidades sísmicas (Perin, 2010), sendo portanto, indicada para sinais de dados contendo mudanças abruptas de contraste ou outras descontinuidades. Assim, neste trabalho foi escolhida para realizar tal análise a ondaleta Haar, por ser mais simples e indicada para dados que possuem mudanças bruscas. Foi realizada uma análise por transformada discreta para verificar a possibilidade de obtenção de uma boa aproximação dos dados, e uma análise por transformada contínua de ondaleta Haar para uma observação do comportamento dos coeficientes encontrados.

2. FUNDAMENTOS TEÓRICOS

A grande utilidade das wavelets ou ondaletas, é que as ondaletas são funções mais simples que as funções já conhecidas e utilizadas para representar séries, além disso, a ondaleta consegue dar ênfase às similaridades que ocorrem nos dados, tornando mais visíveis as tendências ou mais fácil localizar grandes similaridades. Por exemplo, a ondaleta Haar é dada por duas funções, a função escala ϕ (ondaleta pai) e a função ondaleta mãe ψ :

$$\phi(t) = \begin{cases} 1, & 0 \leq t < 1 \\ 0, & \text{outros casos} \end{cases} \quad (1)$$

$$\psi(t) = \begin{cases} +1, & 0 \leq t < 1/2 \\ -1, & 1/2 \leq t < 1 \\ 0, & \text{outros casos} \end{cases} \quad (2)$$

através da função ψ , é possível realizar dilatações (compressões) e translações desta, levando a obter a ondaleta mãe $\psi(t)$:

$$\phi_{j,k}(t) = \begin{cases} 2^{j/2}, & 2^{-j}k \leq t < 2^{-j}(k+1) \\ 0, & \text{outros casos,} \end{cases} \quad (3)$$

e

$$\psi_{j,k}(t) = \begin{cases} 2^{j/2}, & 2^{-j}k \leq t < 2^{-j}(k+1/2) \\ -2^{j/2}, & 2^{-j}(k+1/2) \leq t < 2^{-j}(k+1) \\ 0, & \text{outros casos.} \end{cases} \quad (4)$$

A Transformada discreta de ondaleta convolve a entrada do sinal por meio dos deslocamentos (translações no tempo) e escalonamento (dilatações e contrações) da ondaleta.

O sistema $\psi_{j,k}(t)$ forma uma base de $L_2(R)$, espaço das funções reais de variável real quadrado integráveis. Existem coeficientes $c_{j,k}$ que tornam possível representar uma função $f(t)$ da seguinte forma:

$$f(t) = \sum_{j=-\infty}^{+\infty} \sum_{k=-\infty}^{+\infty} c_{j,k} \psi_{j,k}(t). \quad (5)$$

A Equação (5) é denominada série de ondaletas de $f(t)$.

A Transformada contínua de ondaleta é representada por:

$$W_f(a,b) = |a|^{-1/2} \int f(t) \psi^* \left(\frac{t-b}{a} \right) dt, \quad (6)$$

em que $f(t)$ é a função a ser analisada, ψ é a ondaleta mãe e $\psi((t-b)/a)$ é a versão deslocada e escalonada da ondaleta no tempo b e na escala a .

3. METODOLOGIA

Foram selecionados 395 municípios baianos para realização desse estudo, e sobre esses municípios foram coletados dados de ocorrência da dengue entre os anos de 2000 e 2009, os dados foram coletados no Sistema de Notificação de Agravos (SINAN). A incidência foi calculada tomando como população média a população de 2006, no apêndice pode ser encontrada uma planilha com dados de alguns municípios. A título de exemplo, segue cálculo do índice de dengue no município de Feira de Santana:

População em 2006 = 571.997

Número de Ocorrências de dengue registrado entre 2000-2009 = 12.734

$$\text{Índice de dengue} = \frac{12734}{571997} = 0.022769252$$

Para citar alguns dos 395 municípios estão as cidades: Abaíra, Acajutiba, Adustina, Água Fria, Amargosa, Amélia Rodrigues, Anagé, Andaraí, Barreiras, Barro Alto, Barro Preto, Belmonte, Belo Campo, Biritinga, Bom Jesus da Lapa, Bom Jesus da Serra, Bonito, Cachoeira, Caculé, Caém, Caetanos, Caetité, Camaçari, Candeias, Catu, Caturama, Central, Chorrochó, Conde, Condeúba, Dias D'Ávila, Dom Basílio, Entre Rios, Feira de Santana, Guanambi, entre outras.

Todos os cálculos, gráficos e resultados foram obtidos por programas implementados através do Matlab, versão 7.11 (R2010b).

A partir dos dados da série de incidência da dengue foi realizada uma análise discreta via ondaleta Haar. Para entender como funciona a decomposição de um sinal em ondaletas, basta lembrar que a Transformada de Fourier é utilizada para encontrar um conjunto de senóides cuja soma corresponde ao sinal original, da mesma forma, a Transformada discreta de ondaleta (DWT) gera “pequenas ondas” a partir de um sinal, produzindo um conjunto de subsinais que se somam para formar uma cópia da entrada do sinal.

Também foi utilizada a Transformada contínua de ondaleta Haar (CWT), com o objetivo de visualizar o comportamento do sinal através de seus coeficientes gerados, e comparar com o resultado obtido na análise discreta, observando ainda se há ocorrência de similaridades ou não.

4. DISCUSSÃO DOS RESULTADOS

Foi gerado um gráfico com os dados de incidência da dengue nos municípios da Bahia, e obteve-se o resultado apresentado na Fig. 1, assim, essa série temporal representa os índices de ocorrência da dengue no período de 2000 à 2009, aqui entende-se sinal original como o sinal de dados dos índices de ocorrência calculados.

Em seguida, os dados foram submetidos a uma análise discreta por compressão via ondaleta Haar, com 5 oitavas ou aproximação nível 5 (Fig. 2). Percebe-se que após a compressão discreta do sinal via ondaleta Haar e eliminando os coeficientes próximos de zero (Fig. 3), a superposição de sinais indica que há um certo erro em alguns pontos do sinal aproximado. Já quando se utiliza todos os coeficientes calculados, inclusive os de valor muito pequeno (próximos de zero) a aproximação se torna quase perfeita (Fig. 4).

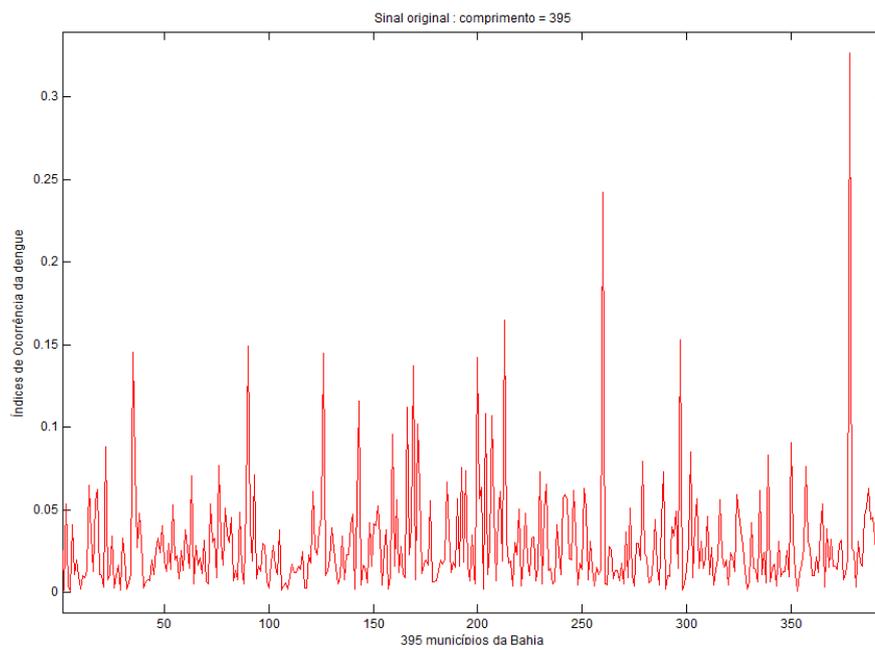


Figura 1 – Gráfico de incidência de Dengue em 395 municípios bahianos em 2000-2009.

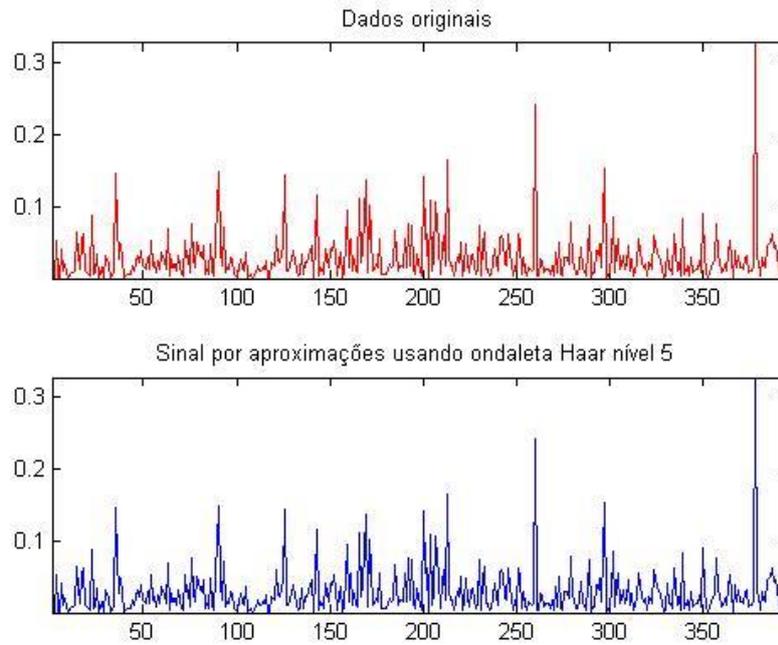


Figura 2 - Sinal de dados originais e sinal por aproximações usando análise discreta por ondaleta Haar nível 5 (dados originais em vermelho, dados aproximados por ondaletas Haar em azul)

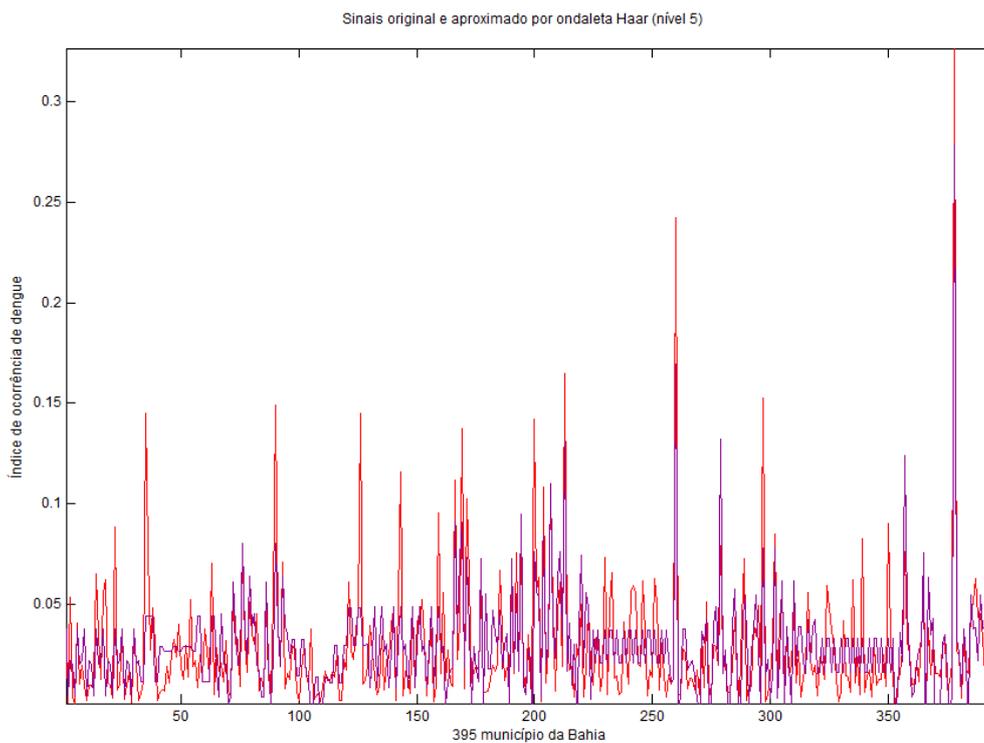


Figura 3 - Superposição de sinais originais e aproximados por ondaleta Haar nível 5 (1 – D), removendo os coeficientes próximos de zero.

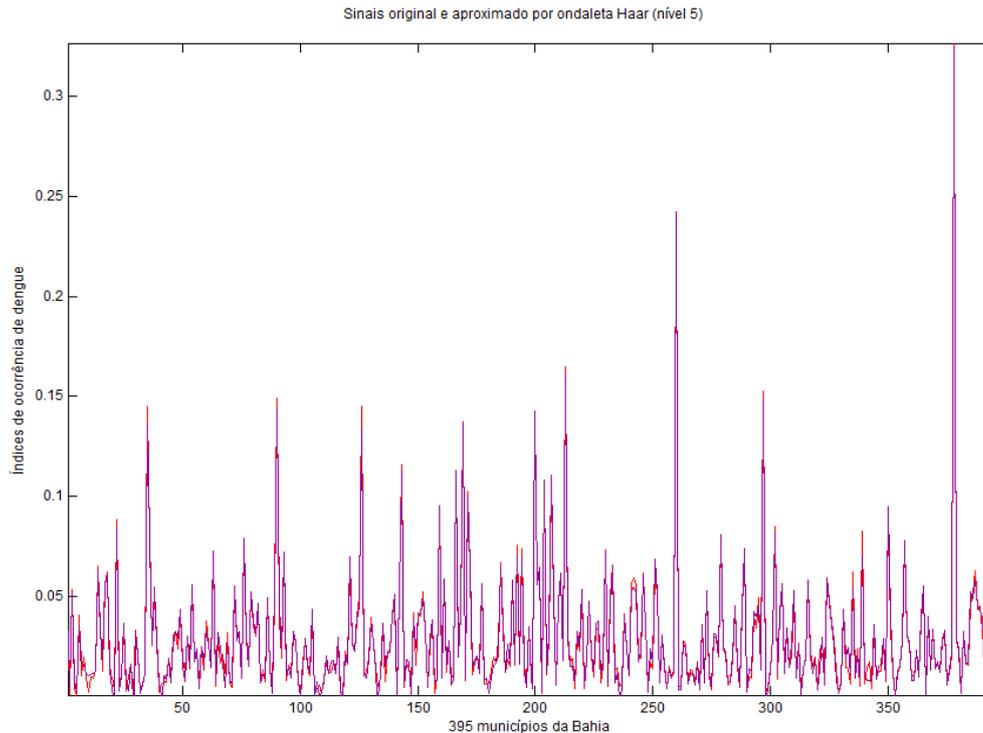


Figura 4 - Superposição dos sinais originais e aproximados por ondaleta Haar nível 5 (1 – D), sem remoção de coeficientes.

Foi realizado um comparativo do resíduo obtido entre o sinal de dados originais e o sinal aproximado pela ondaleta Haar (nível 5) sem remoção dos coeficientes bem pequenos, e em seguida a autocorrelação, que caracteriza uma série temporal em relação a sua interdependência temporal, ou seja, a autocorrelação tem o propósito de mensurar a interdependência existente entre observações de uma série temporal, obteve-se os seguintes valores: média do valor residual = 5.06×10^{-6} , máximo residual = 0.0127 e mínimo residual = -0.01068, indicando que existe auto correlação na serie.

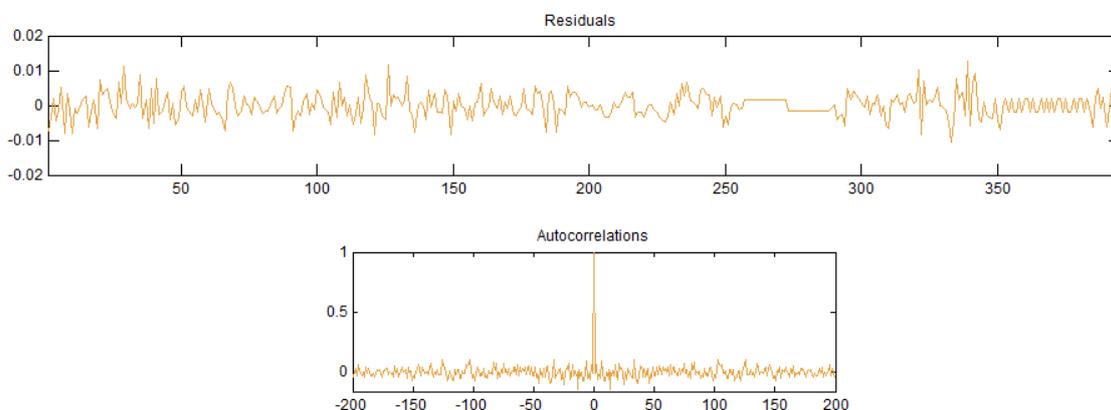


Figura 5 – Resíduo entre sinal original e sinal aproximado por ondaleta Haar e autocorrelação.

Para efeito de comparação da análise e visualização, os mesmos dados foram submetidos a uma análise por Transformada contínua por ondaleta Haar, obtendo-se então uma

representação dos seus coeficientes (Fig. 6), nessa análise foram calculados 25.280 coeficientes. Observa-se uma imagem com um grande número de ‘feixes’ periódicos sobre toda a série.

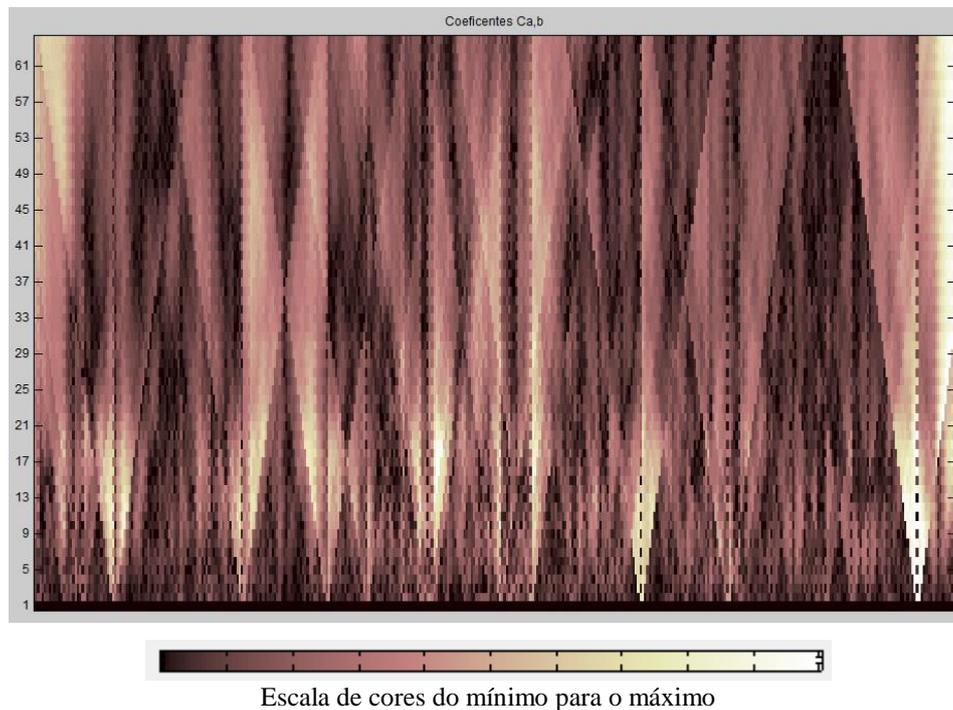


Figura 6 - Coeficientes obtidos por Transformada contínua por ondaleta Haar 1- D.

A escala de cor indica que a região mais clara (cor branca) representa valores maiores e a cor escura representa valores pequenos, muito próximos de zero, sendo $C_{a,b} = 0.2118$ o maior valor encontrado, e $C_{a,b} = - 0.2166$, o menor valor.

4. CONCLUSÕES

As Figuras 3 e 4 mostram a superposição dos sinais, indicando uma forte coincidência entre o sinal original e o reconstruído pelas somas dos subsinais ou aproximações das ondaletas, isto é, houve um erro muito pequeno na aproximação utilizando todos os coeficientes calculados, levando a uma recuperação quase perfeita do sinal. Isso induz a possibilidade de usar o tratamento e estudo dos dados via ondaleta Haar. A representação desses dados por ondaleta é importante por tornar possível outros tipos de análises, pois o sinal de dados agora passa a ter uma função que o representa, seguindo o modelo dado pela Eq. (5).

O resíduo entre sinal original e sinal aproximado por ondaleta Haar (nível 5), mostra um erro médio bem pequeno, 5.06×10^{-6} , confirmando que a recuperação do sinal foi excelente, e para confirmar tal acurácia foi gerada a autocorrelação, conforme mostra a Fig. 5.

A Figura 6 confirma a existência de similaridades, a informação da escala de cor: cor branca representa valores maiores, e a cor mais escura representa valores pequenos, quase nulos, percebe-se que nos pontos onde há picos de incidência com alta magnitude, os coeficientes não são valores pequenos, já que essa característica importante não pode ser suavizada nem descartada. O comportamento repetido, com simetria de ‘feixes’ por toda a

série, revelam que os dados apresentam uma forte similaridade, mais ainda, há uma repetição periódica no comportamento dos coeficientes, algo já esperado por se tratar de dados de incidência de dengue, que segue um comportamento periódico, de acordo com o clima apropriado para reprodução do mosquito.

Assim, é possível representar uma série temporal de incidência da dengue por série de ondaleta Haar por Transformada discreta, sem perdas efetivas, pois o sinal é recuperado com erro muito pequeno. O estudo feito por transformada contínua mostrou ser um bom parâmetro para verificação do comportamento da série de dados através de seus coeficientes, de forma a detectar similaridades ou autoafinidades que o caracterizam, podendo ser utilizada para futuras pesquisas sobre a propagação da doença, para estabelecer iniciativas e metas de combate e controle da mesma.

REFERENCES

- Ivanov, P. Ch.; Rosenblum, M. G.; Peng, C.K.; Mietus, J.; Havlin, S.; Stanley, H. E.; Goldberger, A. L. (1996), “Scaling behaviors of Heartbeat Intervals obtained by Wavelet-Based Time-Series Analysis”. *Nature*. 383 (6598), 323-327.
- Morettin, P. A. (1999), “*Ondas e Ondaletas: Da Análise de Fourier à Análise de Ondaletas*”, EDUSP, São Paulo.
- Oliveira, H. M. (2007), “*Análise de sinais para engenheiros: uma abordagem via wavelets*”, Brasport, Rio de Janeiro.
- Perin, H. M. e Figueiró, W. M. (2010), “Parametrização de campo unidimensional de velocidades sísmicas usando ondaletas”, Resumos Expandidos, *IV SimBGf - Simpósio Brasileiro de Geofísica*, CDROM, ISBN 2179-0965, Brasília.
- Ribeiro, S.M.A., (2016), “*Auto-afinidade na Série Temporal da Dengue*”, Tese de Doutorado, SENAI CIMATEC, Salvador.
- Saba, H.; Vale, V. C., Moret, M. A.; Miranda, J. G. (2014a), “Spatio-temporal correlation networks of dengue in the state of Bahia”, *BMC Public Health* (Online), v. 14, p. 1085, 2014b.
- Saba, H.; Miranda, J. G. V.; Freitas, J. E. M.; Moret, M. A. (2014b), Correlation between Transport and Occurrence of Dengue Cases in Bahia. *Journal of Communication and Computer*, v. 11, p. 217-221.
- Teixeira, M. G.; Costa, M. C. N.; Barreto, M. L.; Mota, E. L. (2005) “Dengue and Dengue Hemorrhagic Fever Epidemics in Brazil: what research needs are indicated by its trend, surveillance and control experiences?”, *Cadernos de Saúde Pública*, 21 (5), 1307-1315, Rio de Janeiro.
- Teixeira, M. G.; Barreto, M. L.; Guerra, Z. (1999) “Epidemiologia e medidas de prevenção do dengue”, *Inf. Epidemiol. Sus.*, v. 8, n. 4, p. 5-33.
- Weeks, Michael. (2012), “*Processamento Digital de Sinais – Utilizando Matlab e Wavelets*”, 2ªed., LTC, Rio de Janeiro.