

AUTO-AFINIDADE NA SÉRIE TEMPORAL DA DENGUE

Stela Azevedo¹ Hugo Saba² e Marcelo Moret³

¹Universidade Estadual de Feira de Santana

²Universidade do Estado da Bahia

³Faculdade Senai Cimatec

E-mails: stelamariaa@gmail.com, huqosaba@gmail.com, mamoret@gmail.com

RESUMO

A dengue é uma infecção viral transmitida por mosquito *Aedes aegypti* (*A. aegypti*) predominantemente em regiões tropicais e subtropicais, com uma incidência e alcance que têm aumentado substancialmente nas últimas três décadas. Este artigo propõe a aplicação do método de análise de flutuações sem tendências (Detrended Fluctuations Analysis - DFA), para analisar o comportamento de escala nas séries temporais obtidas por meio do número de ocorrência de casos de dengue nos municípios do Estado da Bahia, Brasil. Nossos resultados mostram que há dois comportamentos distintos para flutuação, ou seja, dentro de um mês a série temporal é persistente e de um mês a um ano, esta se comporta como um processo sub-difusivo.

1. INTRODUÇÃO

A dengue é considerada a mais importante das doenças virais transmitidas por artrópodos (Braga *et al.*, 2007), (Rigau-Pérez *et al.*, 1998) sendo também a mais comum e distribuída arbovirose no mundo (Rathor, 2002) e tem se destacado entre enfermidades reemergentes.

A palavra dengue é de origem espanhola 'melindre', 'manha'. Refere-se ao estado de moleza e prostração em que fica a pessoa contaminada pelo arbovírus. A transmissão ocorre pela picada do mosquito *A. aegypti*, uma espécie hematófaga originária da África que chegou ao continente americano na época da colonização (SVS, 2013).

O primeiro relato da dengue foi no final do século XVIII, no Sudoeste Asiático, em Java, e nos Estados Unidos, na Filadélfia. Entretanto, a Organização Mundial de Saúde (OMS) somente a reconheceu como doença no século XX (SVS, 2013).

Na segunda metade do século XX a dengue emergiu como o vetor urbano da doença mais prevalente transmitida de humanos, devido à elevada propagação entre as populações urbanas de seres humanos e mosquitos *Aedes aegypti*. A intervenção contra *A. aegypti* em habitats domésticos, conhecida como redução na fonte, é fundamental para as atividades de prevenção da maioria dos departamentos de saúde pública das cidades endêmicas (Barrera *et al.*, 2006, Farrar *et al.*, 2007, Horstick *et al.*, 2010, Padmanabha *et al.*, 2012).

No Brasil, há referências de epidemias desde 1846, sendo que a primeira epidemia, documentada clínica e laboratorialmente, ocorreu entre os anos de 1981 e 1982, em Boa Vista, Roraima – essas epidemias foram causadas pelos sorotipos 1 e 4, considerados os mais perigosos. A partir de 1986, ocorreram epidemias, atingindo o Rio de Janeiro e algumas capitais da região Nordeste (SVS, 2013).

O Boletim Epidemiológico da Secretaria de Vigilância em Saúde (SVS, 2012) demonstrou que, no período de janeiro a abril de 2012, foram registrados 286.011 casos de dengue no Brasil. Esses dados demonstram uma redução de 44% dos casos de dengue no país. Contudo, apesar de todas as regiões apresentarem redução na transmissão da doença, algumas unidades federadas apresentaram aumento na incidência em 2012, entre elas destaca-se o Estado da Bahia com 200,9 casos por 100 mil habitantes.

No ano de 2011, foram registrados 22.424 casos de dengue na Bahia, em 2012 esse número subiu para 28.154. Portanto, em 2011 a incidência de dengue no Estado da Bahia era de 160,0 casos por 100 mil habitantes, sendo elevadas em 2012 para 200,9 casos por 100 mil habitantes (SVS, 2012).

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 50-100 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. Aproximadamente 2,5% dos infectados morreram (Pongpan *et al.*, 2013).

Não existem tratamentos específicos exceto para a sintomática, sendo eficazes na detecção precoce. Em pacientes com uma infecção grave, choque e hemorragia geralmente complementam o quadro clínico. Se não for tratada, a morte pode ser uma consequência. A detecção precoce ou prognóstico correto podem evitar tais complicações graves (Pongpan *et al.*, 2013).

No Brasil, a dengue apresenta um padrão sazonal, com maior incidência de casos nos primeiros cinco meses do ano, período mais quente e úmido, típico dos climas tropicais, (Braga *et al.*, 2007, Rigau-Pérez *et al.*, 1998).

Este estudo trata da aplicação do método DFA em séries temporais obtidas através do número de ocorrência de casos de dengue nos municípios do Estado da Bahia, Brasil. Está estruturado da seguinte forma; na primeira seção é feita uma introdução do problema a ser estudado; na segunda seção é apresentado os "Conceitos Fundamentais necessários para entender o DFA." que é o utilizado no trabalho; a terceira seção apresenta os resultados e discussão e na quarta seção tem-se a conclusão do estudo feito por este artigo.

2. METODOLOGIA

Neste artigo foram selecionadas séries temporais de incidência diária dos casos de dengue em vinte e cinco cidades do estado da Bahia, entre os anos de 2000 a 2010, para analisarmos a auto-afinidade dessa epidemia. Para tanto usaremos o método Detrended Fluctuation Analysis (DFA).

O Método DFA

O método Detrended Fluctuation Analysis (DFA), ou seja, análise de flutuações retirando tendências foi proposto por Peng *et al.* (1994) e tem como ideia central subtrair possíveis tendências determinísticas da série temporal original analisando seus dados destendenciados. Originalmente o método DFA, foi baseado na teoria de caminhos aleatórios (Mandelbrot, 1983, Mandelbrot, 1997) e um aperfeiçoamento do método Fluctuation Analysis

FA (Peng *et al.*, 1992). Inicialmente proposto para aplicações em análise sequencial de DNA (Peng *et al.*, 1994), o método DFA tem aplicações na análise de séries temporais nas mais diversas áreas (Bashan *et al.*, 2008, Kantelhardt *et al.*, 2001) tais como: estrutura de nuvens (Ivanova *et al.*, 2000), geologia (Malamud *et al.*, 1999), etnologia (Alados *et al.*, 2000), variabilidade cardíaca (Ivanova *et al.*, 1996); flutuações em canais iônicos (Siwy *et al.*, 2002); climatologia (Ivanova *et al.*, 1999) intervalo entre passos sucessivos, como forma de avaliar uma doença através do modo que o paciente caminha (Hausdorff *et al.*, 1996) e séries temporais financeiras (Liu *et al.*, 1997, Costa *et al.*, 2003). Uma das razões de se empregar o método DFA é evitar a detecção falsa de correlações que são artefatos de não-estacionaridades nas séries temporais.

Para aplicação do método DFA (Peng *et al.*, 1992), é necessário seguir os seguintes passos:

- i) Dada uma série temporal original obtém-se a série acumulada.
- ii) Dividir a série acumulada em janelas ou subconjuntos de tamanho s .
- iii) Remover a tendência local em cada janela por meio de um ajuste polinomial.
- iv) Calcular a função de flutuação $F'(s)$ em termos do tamanho das janelas.

Considere uma série temporal $\{r\}$, cujos valores são r_i , com i variando de 1 à N (comprimento total da série), que podem ser por exemplo o número de casos de dengue em um município baiano coletada diariamente.

- i) Para obtemos uma série com menor número de oscilações e apropriada para aplicar o método, calcula-se a somatória do desvio padrão de cada registro em relação à incidência média na série completa e obtendo assim a série integrada $Y(k)$ dado por:

$$Y(k) = \sum_{i=1}^k r_i - \langle r \rangle,$$

onde $\langle r \rangle$, é o valor médio de r_i .

- ii) Em seguida divide-se a série integrada $Y(k)$ em intervalos (janelas) de tempos iguais de amplitude n não sobrepostos.

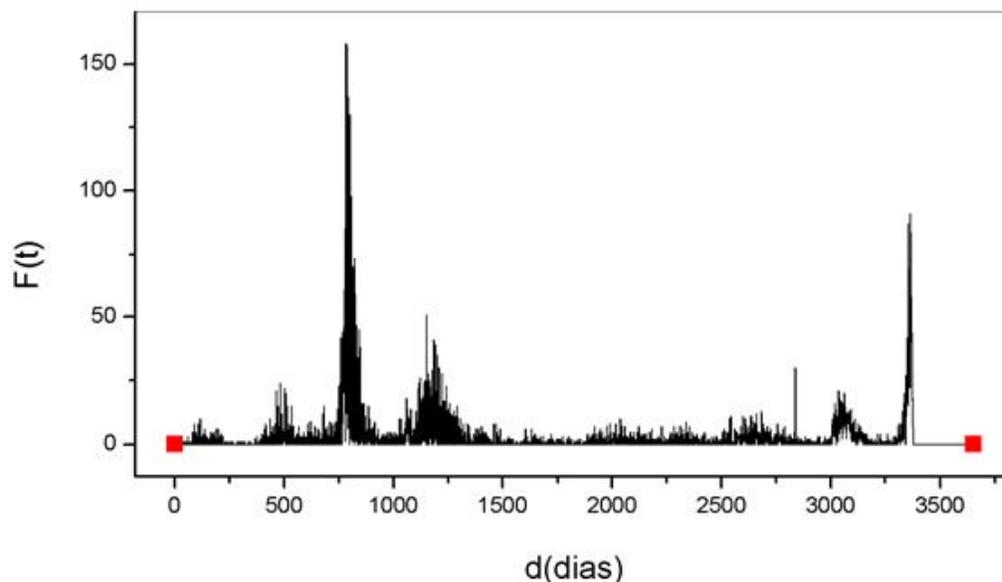


Figura 1 – Série original de casos de incidência diária da dengue na cidade de Feira de Santana.



Figura 2 – Série integrada após ter sido retirada a média de casos de incidência diária da dengue na cidade de Feira de Santana.

- iii) Ajusta-se os dados por um polinômio de tendência local de grau ≥ 1 a $Y(k)$, denominado $Y_n(k)$ para cada janela de tamanho n .
- iv) A série integrada $Y(k)$ é subtraída da série ajustada $Y_n(k)$ em cada janela com tamanho n . Assim, é calculada a raiz quadrática média $F(n)$ para cada janela de tamanho n , isto é,

$$F(n) = \sqrt{\frac{1}{N} \sum_{k=1}^N [Y(k) - Y_n(k)]^2}$$

Deve-se repetir o cálculo apresentado anteriormente para diferentes janelas de tamanho n . E, além disso, o ajuste $Y_n(k)$ pode ser modelado para diferentes polinômios de grau m (Perazzo *et al.*, 2004).

Finalmente, verifica-se se a função $F(n)$ comporta-se como uma lei de potência, ou seja, se é do tipo $F(n) \sim n^\alpha$. Em caso afirmativo, indica que a série possui fractalidade ou auto-similaridade e α será o coeficiente de correlação de longo alcance. O coeficiente α poderá ser obtido por meio de um ajuste linear ($y = \alpha x + b$) de um gráfico $\log [F(n)]$ por $\log(n)$.

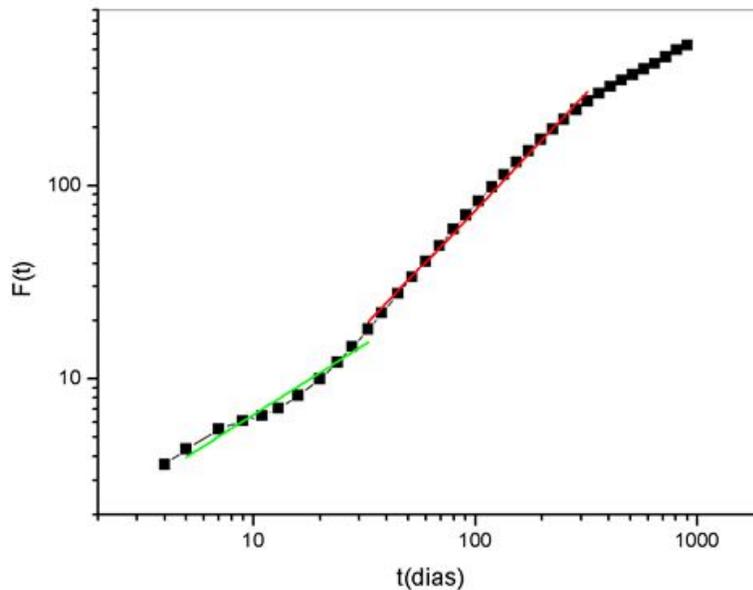


Figura 3 – Obtenção do coeficiente α a partir do DFA da incidência da dengue na cidade de Feira de Santana, de 2000 a 2010. A curva verde referente há trinta dias $\alpha = 0,72 \pm 0,05$. A curva vermelha referente há um ano $\alpha = 1,20 \pm 0,02$.

É neste contexto que o expoente α revela propriedades de correlação na série em estudo. Ele nos permite avaliar em que medida a tendência observada na série temporal passada implica na manutenção do comportamento futuro, indicando um efeito de memória de longa duração da série.

O expoente de flutuação pode ser classificado de acordo com uma faixa dinâmica de valores (Bashan *et al.*, 2008, Kantelhardt *et al.*, 2001, Zeng *et al.*, 2008, Galhardo *et al.*, 2009).

- Se $0 < \alpha < 0,50$, a série apresenta um comportamento anti-correlacionado, sinal anti-persistente, significa que os valores grandes (pequenos) têm maior probabilidade de serem seguidos por valores pequenos (grandes);
- $\alpha \cong 0,50$, a série será aleatória, descorrelacionada, ruído branco, não apresenta memória e a série original apresenta valor zero para a função de auto-correlação, isto é, a função de correlação terá decaimento exponencial;
- $0,50 < \alpha < 1$, será um parâmetro de auto-similaridade e a série apresenta correlação de longo alcance;
- $\alpha \cong 1$ a série apresenta um ruído do tipo $1/f$;
- $1 < \alpha < 1,5$, não estacionários, como um processo subdifusivo (Galhardo *et al.*, 2009);
- $\alpha \cong 1,5$, indica um ruído marrom, que é a integração do ruído branco, ruído tipo browniano (Galhardo *et al.*, 2009);
- $\alpha > 1,5$ corresponde a um processo superdifusivo (Galhardo *et al.*, 2009).

O parâmetro α pode ser interpretado como indicador da suavidade da série original. Quanto maior seu valor, maior a suavidade dá série.

3. RESULTADOS E DISCUSSÃO

Com base nos dados do Sistema de Notificação de Agravos (SINAN), obtivemos os dados da incidência diárias das cidades do estado da Bahia. Após aplicarmos o método DFA,

nas séries de vinte e cinco cidades selecionadas pelo maior índice populacional, obtivemos os dados da Tabela 1. Analisando os resultados podemos observar que existem dois comportamentos distintos para todas as cidades, a saber, até o trigésimo dia o coeficiente de correlação de longo alcance varia entre $0,50 < \alpha < 1,00$ e para o período compreendidos entre um mês e um ano o α varia entre 1,00 e 1,50.

Tabela 1 – Expoentes de correlação de longo alcance, α e o desvio padrão, σ , das vinte e cinco cidades selecionadas e o erro referente ao mês e ao ano. Dados de 2000 a 2010.

Município	α (mês)	σ (desvio padrão)	α (ano)	σ (desvio padrão)
Salvador	0,98	0,04	1,35	0,03
Feira de Santana	0,72	0,05	1,20	0,02
Vitória da Conquista	0,55	0,02	1,23	0,01
Camacari	0,50	0,02	1,34	0,02
Itabuna	0,76	0,06	1,31	0,03
Juazeiro	1,04	0,06	1,09	0,02
Ilhéus	0,68	0,04	1,28	0,02
Lauro de Freitas	0,88	0,02	1,18	0,01
Jequié	0,88	0,06	1,49	0,03
Teixeira de Freitas	0,54	0,01	1,18	0,02
Alagoinhas	0,59	0,05	1,26	0,03
Barreiras	0,74	0,03	1,09	0,02
Porto Seguro	0,79	0,04	1,02	0,03
Simões Filho	0,63	0,02	1,45	0,02
Paulo Afonso	0,65	0,03	1,11	0,01
Eunápolis	0,58	0,01	1,16	0,01
Santo Antônio de Jesus	0,58	0,03	1,22	0,02
Valença	0,68	0,02	1,05	0,01
Candeias	0,65	0,02	1,08	0,01
Guanambi	0,76	0,01	1,14	0,01
Jacobina	0,57	0,04	1,27	0,02
Serrinha	0,67	0,02	1,09	0,01
Senhor do Bonfim	0,71	0,05	1,11	0,02
Dias d'Ávila	0,54	0,01	1,13	0,01
Itapetinga	0,62	0,02	1,32	0,02

A Tabela 1 ainda nos fornece o valor médio do expoente $\alpha = 0,69 \pm 0,14$, para períodos inferiores a um mês e $\alpha = 1,21 \pm 0,12$ para períodos compreendidos entre um mês e um ano. Desta análise podemos afirmar que há dois comportamentos bem determinados para os dados, ou seja, o comportamento é persistente para períodos de até um mês e subdifusivos para períodos de até um ano.

Como exemplo, mostramos os comportamentos para o expoente α da cidade de Feira de Santana na Figura 3. Observamos pelos ajustes mostrados na Figura 3 que Feira de Santana segue o mesmo padrão descrito na Tabela 1, o que também é verificado quando analisamos a série em todo o estado da Bahia.

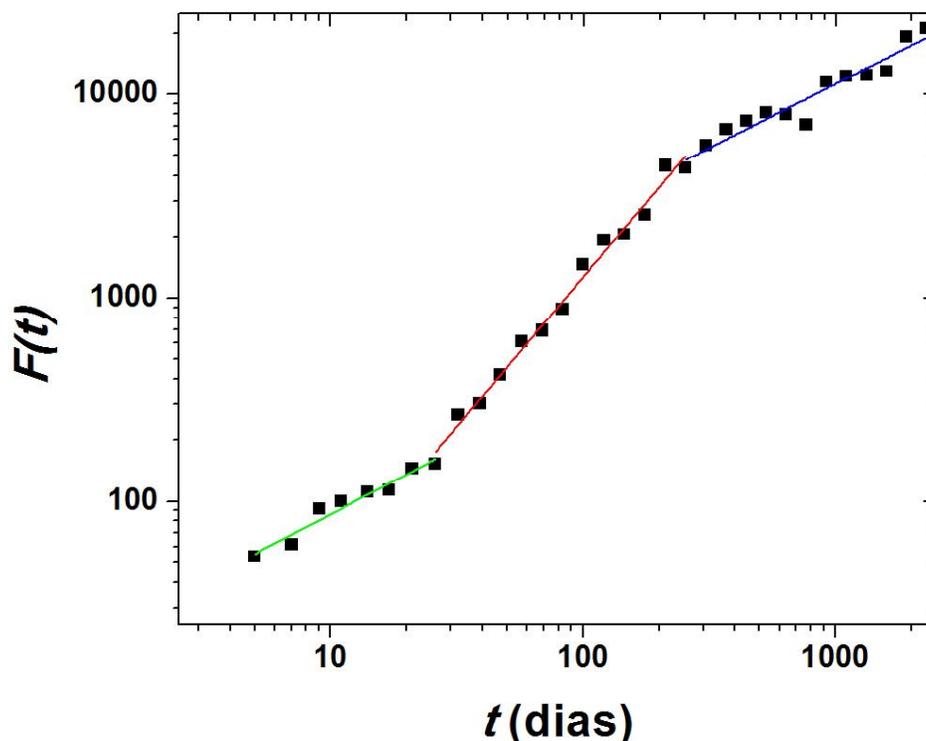


Figura 4 - Série da incidência diária de casos da dengue no estado da Bahia, de 2000 a 2010. A curva verde referente há trinta dias $\alpha = 0,63 \pm 0,05$. A curva vermelha referente há um ano $\alpha = 1,45 \pm 0,04$.

A Figura 4 mostra o comportamento auto-afim da Dengue em relação aos dados gravados no SINAM do estado da Bahia. O perfil do expoente de correlação se mantém persistente em períodos inferiores a um mês e subdifusivo para períodos compreendidos entre um mês e um ano.

4. CONCLUSÃO

Ao aplicar o algoritmo do método *Detrended Fluctuation Analysis* – DFA nos dados obtidos das vinte e cinco cidades, listadas na Tabela 1, foi possível identificar as correlações em séries temporais, e obtivemos as seguintes conclusões:

No caso específico da cidade de Feira de Santana, percebe-se que existem dois comportamentos distintos da série, conforme mostra a Figura 3. Até o trigésimo dia o valor do expoente α obtido pelo método DFA varia entre 0,50 e 1,00, o que indica um parâmetro de

auto-similaridade e que a série original apresenta auto-correlações de longo alcance persistentes, significando que os valores grandes (pequenos) têm probabilidade de serem seguidos por valores grandes (pequenos).

Para o período anual α varia entre 1,00 e 1,50, que caracteriza uma série não estacionária, similar ao comportamento de caminhos aleatórios não estacionários em um processo subdifusivo, ou seja, seu comportamento apresenta uma certa sazonalidade, sem, contudo, apresentar epidemias semelhantes ano a ano.

Observa-se ainda, na Tabela 1, que o comportamento dos expoentes das vinte e quatro cidades tem o mesmo comportamento da cidade Feira de Santana. Nos primeiros trinta dias de todas as cidades os coeficientes do DFA estão entre 0,50 e 1,00, indicando que todas as séries são persistentes, o que significa que valores altos tem maior probabilidade de serem seguidos por valores altos, e valores baixos tem maior probabilidade de serem seguidos por valores baixos, e no período anual os coeficientes variam entre 1,00 e 1,50, mantendo a característica de série não estacionária, com o comportamento similar ao de caminhos aleatórios.

Portanto, a análise de dados estatísticos de incidência diária da dengue pode ser tratada por séries temporais, e a aplicação do método *Detrended Fluctuation Analysis* – DFA nos permite obter o coeficiente α , o qual nos informa o comportamento da função de correlação (auto-afinidade), ou seja, indica se a série tem um comportamento linear: valores altos (baixos) tem maior probabilidade de serem seguidos por valores altos (baixos), para prever o comportamento da incidência da dengue em períodos futuros.

Concluindo, a análise dos dados do número de infectados declarados a partir das vinte e cinco cidades estudadas usando método *Detrended Fluctuation Analysis* – DFA, pode ser útil como instrumento de previsão da incidência da dengue em períodos futuros, através de uma extrapolação do comportamento do expoente de correlação de longo alcance (α) possibilitando com antecedência a ação de medidas efetivas de manutenção adequada à prevenção da doença. Estas ações podem promover a minimização de casos da dengue e a previsão de demanda hospitalar nessas comunidades. O comportamento de longo alcance da Dengue apresenta uma maior previsibilidade em períodos relativamente curtos (aproximadamente um mês), pois a ocorrência da infecção (incidência diária) tem um comportamento de longo alcance persistente, segundo o coeficiente de correlação de longo alcance. Para períodos maiores do que um mês o método somente é capaz de prever que há uma tendência a sazonalidade, contudo não é possível se prever uma epidemia futura somente com a informação atual de infectados.

REFERÊNCIAS

- Addison, P. S. *Fractals and Chaos: An Illustrated Course*. CRC Press, 256 (1997).
- Barrera, R., Amador, M. & Clark, G. G. Use of the pupal survey technique for measuring *Aedes aegypti* (Diptera: Culicidae) productivity in Puerto Rico. *Am. J. Trop. Med. Hyg.*, **74**(2), 290-302 (2006).
- Bashan, A., Bartsch, R., Kantelhardt, J. W. & Havlin, S. Comparison of detrending methods for fluctuation analysis, *Physica A*. **387**(21), 5080-5090 (2008).
- Malamud, B. D. & Turcotte, D. L., Self-affine time series: measures of weak and strong persistence, *J. Stat. Plan. Infer.* **80** (1-2), 173 (1999).
- Braga, I. A. & Valle, D. *Aedes aegypti*: histórico do controle no Brasil. *Epidemiologia e Serviços de Saúde* **16**(2), 113-118 (2007).

Alados, C. L. & Huffman, M. A. Fractal long-range correlations in behavioural sequences of wild chimpanzees: A non-invasive analytical tool for the evaluation of health, *Ethnology* **106**(2), 105-116 (2000).

Costa, R. L. & Vasconcelos, G. L. Long-range correlations and nonstationarity in Brazilian stock market. *Physica A*. **329**(1-2), 231-248 (2003).

Ivanov, P. Ch., Rosenblum, M. G., Peng, C.-K., Mietus, J., Havlin, S., Stanley, H. E. & Goldberger, A. L. Scaling behaviors of Heartbeat Intervals obtained by Wavelet-Based Time-Series Analysis. *Nature*. **383**(6598), 323-327 (1996).

Farrar, J., Focks, D., Gubler, D., Barrera, R., Guzman, M. G., Simmons, C., Kalayanarooj, S., Lum, L., McCall, P. J., Lloyd, L., Horstick, O., Dayal-Drager, R., Nathan, M. B., Kroeger, A. WHO/TDR Dengue Scientific Working Group. Towards a global dengue research agenda. *Trop. Med. Int. Health*. **12**(6), 695-699 (2007).

Ivanova K.; Ausloos, M. Application of the detrended fluctuation analysis (DFA) method for describing cloud breaking. *Physica A*. **274**(1), 349-354 (1999).

Hausdorff, J. M., Purdon, P. L., Peng, C.-K., Ladin, Z., Wei, J. Y., & Goldberger, A. L. Fractal dynamics of human gait: stability of long-range correlations in stride interval fluctuations. *Journal of Applied Physiology*. **80**(5), 1448-1457 (1996).

Horstick, O., Runge-Ranzinger, S., Nathan, M. B. & Kroeger, A. Dengue vector-control services: how do they work? A systematic literature review and country case studies. *Trans. R. Soc. Trop. Med. Hyg.* **104**(6), 379-386 (2010).

Kantelhardt, J. W., Koscielny-Bunde, E., Rego, H. H. A., Havlin, S. & Bunde, A. Detecting long-range correlations with detrended fluctuation analysis. *Physica A*. **295**(3-4), 441-454, (2001).

Ivanova, K., Ausloos, M., Clothiaux, E. E., Ackerman, T. P. Break-up of stratus cloud structure predicted from non-Brownian motion liquid water and brightness temperature fluctuations. *Europhys. Lett.* **52**(1), 40-46 (2000).

Liu, Y., Cizeau, P., Meyer, M., Peng, C.-K. & Stanley, H. E. Correlations in economic time series. *Physica A*. **245**(3-4), 437-440 (1997).

Mandelbrot, B. *The Fractal Geometry of Nature*. New York: W.H. Freeman and Company. 1983.420p.

Padmanabha, H., Durham, D., Correa, F., Diuk-Wasser, M. & Galvani, A. The interactive roles of *Aedes aegypti* super-production and human density in dengue transmission. *PLoS Negl Trop Dis.* **6**(8), e1799 (2012).

Peng, C.-K., Buldyrev, S. V., Havlin, S., Simons, M., Stanley, H. E., Goldberger, A. L. Mosaic organization of DNA sequences. *Phys. Rev. E*. **49**(2), 1685-1689 (1994).

Peng, C.-K.; Buldyrev, S.V.; Goldberger, A.L.; Havlin, S.; Sciortino, F.; Simons, M.; Stanley, H.E. Long-range correlations in nucleotide sequences. *Nature*, **356**(6365), 168-170 (1992).

Perazzo, C. A., Fernandez, E. A., Willshaw, Y. P. Determinación de Correlaciones a Largo Plazo Utilizando DFA. In: *Procesamiento de señales e imágenes: Teoría y Aplicaciones*, Ed.: Facultad Regional Buenos Aires, Universidad Tecnológica Nacional, 2004.

Pongpan, S., Wisitwong, A., Tawichasri, C., Patumanond, J. & Namwongprom, S. Development of Dengue Infection Severity Score. *ISRN Pediatr.* **2013**, 845876 (2013).

Rathor, H. R. The role of vectors in emerging and re-emerging diseases in the Eastern Mediterranean region. *East. Mediterr. Health J.* **2**(1), 61-62 (1998).

Rigau-Pérez, J. G., Clark, G. G., Gubler, D. J., Reiter, P., Sanders, E. J. & Vorndam, A. V. Dengue and dengue hemorrhagic fever. *Lancet*, **352**(9132), 971-977 (1998).

Secretaria de Vigilância em Saúde. Boletim epidemiológico-dengue **43**, 11-15 (2012) Disponível em <http://portalsaude.saude.gov.br/portalsaude/arquivos/bolepi_vol_43_n1.pdf>, acesso em dezembro, 2013.

Secretaria Vigilância Sanitária. Boletim SVS. Disponível em <Ministério da saúde - <http://portal.saude.gov.br/portal/saude>>, acesso em dezembro, 2013. Siwy, Z., Ausloos, M. & Ivanova, K. Correlation studies of open and closed state fluctuations in an ion channel: Analysis of ion current through a large conductance locust potassium channel. Phys. Rev. E, **65**(1), 031907 (2002).

Zheng, H., Song, W. & Wang, J. Detrended fluctuation analysis of forest fires and related weather parameters, Physica A. **387**(9), 2091-2099 (2008).