

Análise de Velocidade do Vento em Petrolina-PE Usando Cadeias de Markov de Ordem Superior

Fábio S. Santos^{1†}, Kerolly K. F. Nascimento², Jader S. Jale³, Tiago A. E. Ferreira⁴

¹Mestrando do Programa de Pós-Graduação em Biometria e Estatística Aplicada (PPGBEA) da Universidade Federal Rural de Pernambuco (UFRPE).

²Doutoranda do Programa de Pós-Graduação em Biometria e Estatística Aplicada (PPGBEA) da Universidade Federal Rural de Pernambuco (UFRPE). E-mail: kerollyfn@gmail.com

³Pós-Doutorando em Ciência da Computação na Universidade Federal Rural de Pernambuco (UFRPE). E-mail: jsjale1983@gmail.com.br

⁴Professor Associado da Universidade Federal Rural de Pernambuco (UFRPE). E-mail: tiago.espinola@ufrpe.br

Resumo: *Com a crescente demanda mundial, a velocidade do vento surge como uma das principais alternativas sustentáveis para a geração de energia. Neste trabalho, utilizamos modelos de cadeia de Markov, a tempo discreto, por meio de 13 estados de velocidade do vento, até ordem 3, com o objetivo de analisar as probabilidades de transição entre estes estados de velocidade do vento e encontrar a ordem que melhor descreve a cadeia proposta. Para isso, analisamos dados horários de velocidade do vento de uma estação meteorológica situada na cidade de Petrolina-PE. Por meio do Critério de Informação de Akaike (AIC), do Critério de Informação de Akaike corrigido (AICc), do Critério de Informação Bayesiano (BIC) e do Critério de Informação de Hannan-Quinn (HQIC), verificamos que o modelo de ordem 1 é o mais apropriado para descrever este fenômeno. Desta forma, a probabilidade de transição entre dois estados quaisquer de velocidade do vento depende apenas do estado observado atual, e independe dos estados anteriores.*

Palavras-chave: Cadeia de Markov; Energia Eólica; Velocidade do Vento; Petrolina-PE.

Abstract: *With increasing global demand, wind speed emerges as one of the main sustainable alternatives to power generation. In this work, we use discrete-time Markov chain models through 13 wind speed states, up to order 3, in order to analyze the transition probabilities between these wind speed states and find the order that best describes the proposed chain. For this, we analyzed hourly wind speed data from a meteorological station located in the city of Petrolina-PE. Through the Akaike Information Criteria (AIC), the Corrected Akaike Information Criteria (AICc), the Bayesian Information Criteria (BIC) and the Hannan-Quinn Information Criteria (HQIC), we have verified that the order model 1 is the most appropriate to describe this phenomenon. In this way, the probability of transition between two states any of wind speed depends only on the current observed state, and it does not depend on the previous states.*

Keywords: Markov Chain; Wind Energy; Wind Speed; Petrolina-PE.

[†]Autor correspondente: fabio.sandropb@gmail.com.

Introdução

Com a crescente demanda mundial por fontes de energia limpa e renovável, a velocidade do vento, dentre outras fontes, surge como uma das principais alternativas sustentáveis para a geração de energia. Para entender de forma mais abrangente o comportamento da dinâmica da velocidade do vento, é necessário utilizar vários métodos científicos para obter a maior quantidade possível de informações dessa importante variável.

Segundo o *World Energy Council* (WEC, 2019), atualmente, o Brasil se encontra na oitava posição mundial de capacidade instalada de energia eólica. Na literatura, os modelos de cadeia de Markov vêm sendo aplicados frequentemente na modelagem da velocidade do vento (D'AMICO et al., 2013; GIORGIO et al., 2015; TANG et al., 2015; TAGLIAFERRI et al., 2016; MIAO et al., 2017; MA et al., 2018).

O objetivo deste trabalho é estimar a melhor ordem que descreve o modelo de cadeia de Markov para dados de velocidade do vento horária em uma estação meteorológica da cidade de Petrolina-PE. Estabelecemos as velocidades do vento (em m/s) em 13 estados para o modelo de cadeia de Markov adotado neste trabalho: $[0,1)$, $[1,2)$, $[2,3)$, $[3,4)$, $[4,5)$, $[5,6)$, $[6,7)$, $[7,8)$, $[8,9)$, $[9,10)$, $[10,11)$, $[11,12)$ e $[12,13)$, visto que os dados obtidos fornecem velocidades entre 0 m/s e 12 m/s.

Metodologia

Descrição dos dados

Os dados de velocidade do vento (em m/s) utilizados neste trabalho foram retirados do *site* do Instituto Nacional de Meteorologia (INMET, 2019). Os registros de velocidade do vento obtidos correspondem às velocidades diárias horárias entre 0h e 23h, totalizando 123.418 observações. Estes dados são referentes à estação meteorológica localizada na cidade de Petrolina-PE (longitude: 40°28'48"W; latitude: 9°22'48"S; altitude: 370.46m).

Modelos de cadeia de Markov

Modelos de cadeia de Markov vêm sendo utilizados em diversas áreas nas mais variadas aplicações. Algumas dessas aplicações incluem processos hidrológicos (DENI; JEMAIN & IBRAHIM, 2009; YEH & LIAW, 2015; ARUMUGAM & KARTHIK, 2018; KWON, KWON & HAN, 2018; JALE et al., 2019; MITRA & SESHADRI, 2019), velocidade do vento (SONG, JIANG, ZHANG, 2014; GIORGIO, LANGELLA & TESTA, 2015; TANG, BROUSTE & TSUI, 2015; MIAO et al., 2017; LI et al., 2019) e energia solar (HOCAOGLU & SERTTAS, 2017; LI, MA & LI, 2017; MIAO et al., 2018; MUNKHAMMAR & WIDÉN, 2018a; MUNKHAMMAR & WIDÉN, 2018b; MUNKHAMMAR & WIDÉN, 2019). Os modelos de cadeia de Markov apresentam duas propriedades fundamentais: o número de estados, indicando os intervalos de valores que a variável em estudo pode assumir; e, a ordem da cadeia indicando qual é o grau de dependência entre os estados.

Sejam X_1, X_2, \dots, X_n , n variáveis aleatórias que representam a sequência de m possíveis estados para um número qualquer de dias n . Dada uma cadeia de Markov de k -ésima ordem, as probabilidades de transição estacionárias entre esses estados são dadas por

$$P_{S_{k+1}, \dots, S_1}(X_t = s_{k+1} | X_{t-1} = s_k, \dots, X_{t-k} = s_1), \quad (1)$$

em que $S = \{s_1, s_2, \dots, m\}$, com m sendo o número de estados possíveis da cadeia de Markov e S o espaço de estados. Sendo assim para a k -ésima ordem, tem-se uma matriz de probabilidade de transição de dimensão $m^k \times m$.

Para obter a ordem ótima da cadeia de Markov, pode-se aplicar critérios de informação tais como AIC, AICc, BIC, HQIC, etc. Esses critérios são obtidos por meio da função de log-verossimilhança das matrizes de probabilidade de transição das ordens consideradas. A função de máxima verossimilhança para uma cadeia de Markov de k -ésima ordem é dada por (DENI, JEMAIN & IBRAHIM, 2009)

$$L_k(X_1, X_2, \dots, X_n) = \prod_{s_k, \dots, s_1} \hat{P}_{s_k, \dots, s_1}^{n_{s_k, \dots, s_1}}, \quad (2)$$

em que $\hat{P}_{s_k, \dots, s_1}^{n_{s_k, \dots, s_1}}$ denota a probabilidade de transição estimada de cada sequência partindo do estado s_1 para o estado s_2 , do estado s_2 para s_3 e finalmente saindo do estado s_{k-1} para s_k , em que s_k é o estado observado atualmente. A notação n_{s_k, \dots, s_1} denota o número de transições ocorridas desde o estado s_1 até o estado s_k . O Estimador de Máxima Verossimilhança (EMV) na Eq. (2) é dado por

$$\hat{P}_{s_k, \dots, s_1}^{n_{s_k, \dots, s_1}} = \frac{n_{s_k, \dots, s_1}}{\sum_{s_k} n_{s_k, \dots, s_1}}. \quad (3)$$

Para tomar a decisão de qual é a ordem ótima entre dois modelos de cadeia de Markov distintos, digamos, k -ésima e r -ésima ordem, por exemplo, com $k < r$, em que $k = 0, 1, \dots, r - 1$, pode-se utilizar a estatística da razão de máxima verossimilhança dada por

$$\eta_{k,r} = -2 \log \lambda_{k,r}, \quad (4)$$

em que

$$\lambda_{k,r} = \frac{L_k(X_1, \dots, X_n)}{L_r(X_1, \dots, X_n)}. \quad (5)$$

A tomada de decisão quanto a ordem ótima da cadeia de Markov pode ser obtida a partir de critérios de informação, tais como AIC, AICc, BIC, HQIC, etc. Atribuindo a k -ésima ordem à cadeia de Markov, pode-se definir a função perda AIC(k) da seguinte forma (TONG, 1975)

$$AIC(k) = \eta_{k,r} - 2\nu, \quad (6)$$

em que $\nu = (m^r - m^k)(m - 1)$ denota os graus de liberdade e m o número de estados possíveis da cadeia de Markov. Outro critério de informação adotado neste trabalho, trata-se do AICc (AIC corrigido) (HURVICH & TSAI, 1989), dado por

$$AICc(k) = \eta_{k,r} - 2\nu \frac{n}{(n - \nu - 1)}, \quad (7)$$

em que n é o tamanho amostral. A função perda baseada no BIC (SCHWARZ, 1978) é dada por

$$BIC(k) = \eta_{k,r} - \nu \log n. \quad (8)$$

Sigmae, Alfenas, v.8, n,2, p. 385-392, 2019.

64ª Reunião da Região Brasileira da Sociedade Internacional de Biometria (RBRAS).

18º Simpósio de Estatística Aplicada à Experimentação Agronômica (SEAGRO).

Outro critério adotado neste trabalho, trata-se da função perda baseada no HQIC (HANNAN & QUINN, 1979), dado por

$$HQIC(k) = \eta_{k,r} - 2\nu \log(\log n). \quad (9)$$

Portanto, o valor que minimize AIC(k), AICc(k), BIC(k) e HQIC(k) será escolhido como o valor ótimo para a ordem do modelo da cadeia de Markov.

Resultados

As Figuras 1 e 2 exibem o histograma de frequências e a série temporal da velocidade do vento para a estação meteorológica de Petrolina-PE.

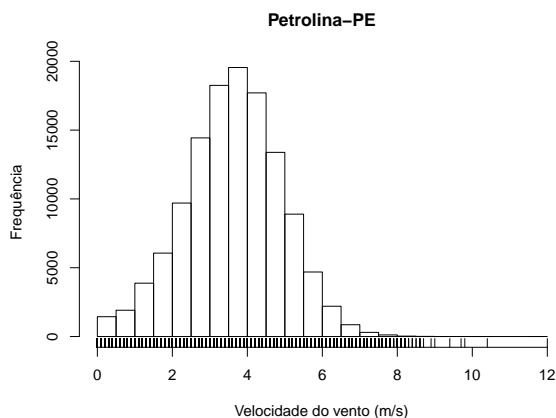


Figura 1: Histograma.

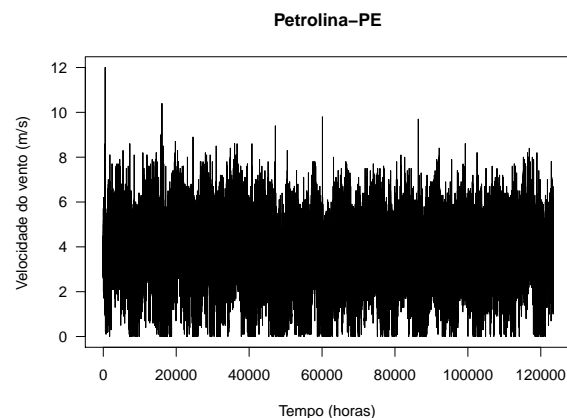


Figura 2: Série temporal.

Fonte: próprios autores.

A Tabela 1 exhibe as estatísticas descritivas da velocidade do vento (em m/s) para a estação meteorológica de Petrolina-PE.

Tabela 1: Estatísticas descritivas para velocidade do vento (em m/s) em Petrolina-PE.

| Mínimo | Q_1 | Mediana | Média | Q_3 | Máximo | DP | CV | Assimetria | Curtose |
|--------|-------|---------|-------|-------|--------|------|------|------------|---------|
| 0.00 | 2.80 | 3.70 | 3.67 | 4.50 | 12.00 | 1.30 | 0.35 | -0.12 | 3.06 |

Fonte: próprios autores.

Na Tabela 2 apresentam-se os resultados referentes a ordem ótima do modelo de cadeia de Markov, por meio dos critérios de informação AIC, AICc, BIC e HQIC, para $r = 3$. Observe que a ordem ótima do modelo de cadeia de Markov é dado por $k = 1$, pois este é o valor que minimiza AICc(k), BIC(k) e HQIC(k), embora o AIC(k) tenha reportado ordem 2. Já que a maioria dos critérios indicam ordem 1 para ordem ótima da cadeia de Markov, adota-se a ordem $k = 1$ como a ordem ótima da cadeia de Markov (destacam-se os menores valores por *). Este resultado indica que existe dependência de primeira ordem entre os estados da cadeia de Markov. Portanto, a ocorrência do estado seguinte

depende apenas do estado atualmente observado e não depende da ocorrência de estados anteriores.

Tabela 2: Critérios de informação para estimação da ordem ótima para $r = 3$.

| k | $\eta_{k,r=3}$ | AIC | AICc | BIC | HQIC |
|-----|----------------|------------|------------|-------------|-------------|
| 0 | 103582.78 | 50878.78 | 36569.72 | -205350.47 | -26152.39 |
| 1 | 8478.41 | -43937.60 | -58069.73* | -298766.68* | -120547.83* |
| 2 | 3101.51 | -45570.50* | -57525.66 | -282197.50 | -116708.56 |

Fonte: próprios autores.

A Tabela 3 apresenta as probabilidades de transição entre os estados da cadeia de Markov analisada neste trabalho.

Tabela 3: Matriz de probabilidade de transição do modelo de cadeia de Markov de ordem 1.

| | [0, 1) | [1, 2) | [2, 3) | [3, 4) | [4, 5) | [5, 6) | [6, 7) | [7, 8) | [8, 9) | [9, 10) | [10, 11) | [11, 12) | [12, 13) |
|----------|---------------|---------------|---------------|---------------|---------------|---------------|---------------|---------------|--------|---------|----------|----------|----------|
| [0, 1) | 0,3946 | 0,3720 | 0,1590 | 0,0483 | 0,0197 | 0,0049 | 0,0014 | 0,0000 | 0,0000 | 0,0000 | 0,0000 | 0,0000 | 0,0000 |
| [1, 2) | 0,1412 | 0,3992 | 0,3226 | 0,1025 | 0,0242 | 0,0065 | 0,0028 | 0,0008 | 0,0003 | 0,0000 | 0,0000 | 0,0000 | 0,0000 |
| [2, 3) | 0,0157 | 0,1531 | 0,4278 | 0,3140 | 0,0723 | 0,0136 | 0,0028 | 0,0004 | 0,0002 | 0,0000 | 0,0000 | 0,0000 | 0,0000 |
| [3, 4) | 0,0017 | 0,0226 | 0,2086 | 0,4747 | 0,2415 | 0,0448 | 0,0052 | 0,0007 | 0,0001 | 0,0000 | 0,0000 | 0,0000 | 0,0000 |
| [4, 5) | 0,0003 | 0,0031 | 0,0442 | 0,3043 | 0,4644 | 0,1617 | 0,0199 | 0,0019 | 0,0002 | 0,0000 | 0,0000 | 0,0000 | 0,0000 |
| [5, 6) | 0,0007 | 0,0020 | 0,0081 | 0,0914 | 0,3955 | 0,3996 | 0,0931 | 0,0087 | 0,0007 | 0,0001 | 0,0000 | 0,0000 | 0,0000 |
| [6, 7) | 0,0006 | 0,0028 | 0,0064 | 0,0324 | 0,1783 | 0,4401 | 0,2837 | 0,0519 | 0,0039 | 0,0000 | 0,0000 | 0,0000 | 0,0000 |
| [7, 8) | 0,0019 | 0,0019 | 0,0132 | 0,0113 | 0,0907 | 0,2836 | 0,4216 | 0,1607 | 0,0151 | 0,0000 | 0,0000 | 0,0000 | 0,0000 |
| [8, 9) | 0,0000 | 0,0192 | 0,0385 | 0,0385 | 0,0577 | 0,1731 | 0,2500 | 0,3654 | 0,0577 | 0,0000 | 0,0000 | 0,0000 | 0,0000 |
| [9, 10) | 0,0000 | 0,0000 | 0,5000 | 0,0000 | 0,5000 | 0,0000 | 0,0000 | 0,0000 | 0,0000 | 0,0000 | 0,0000 | 0,0000 | 0,0000 |
| [10, 11) | 1,0000 | 0,0000 | 0,0000 | 0,0000 | 0,0000 | 0,0000 | 0,0000 | 0,0000 | 0,0000 | 0,0000 | 0,0000 | 0,0000 | 0,0000 |
| [11, 12) | 0,0000 | 0,0000 | 0,0000 | 0,0000 | 0,0000 | 0,0000 | 0,0000 | 0,0000 | 0,0000 | 0,0000 | 0,0000 | 0,0000 | 0,0000 |
| [12, 13) | 0,0000 | 0,0000 | 0,0000 | 0,0000 | 0,0000 | 0,0000 | 0,0000 | 1,0000 | 0,0000 | 0,0000 | 0,0000 | 0,0000 | 0,0000 |

Fonte: próprios autores.

Note que as probabilidades de transição entre os estados $[10, 11) \rightarrow [0, 1)$ e $[12, 13) \rightarrow [7, 8)$ é igual 1, ou seja, se no instante t , a velocidade do vento estiver no intervalo $[10, 11)$, certamente no instante $t + 1$ ela estará no intervalo $[0, 1)$. O mesmo ocorre na transição $[12, 13) \rightarrow [7, 8)$. Um comportamento interessante ocorre na probabilidade de transição entre os estados $[9, 10) \rightarrow [2, 3)$ e $[9, 10) \rightarrow [4, 5)$, cujas transições ocorrem com probabilidade igual a 0,5. Além disso, destacamos em negrito as maiores probabilidades encontradas em cada linha da matriz, indicando que grande parte dos estados tem maior probabilidade de recorrência (acessam a si mesmo), como por exemplo: $[0, 1)$, $[1, 2)$, $[2, 3)$, $[3, 4)$, $[4, 5)$ e $[5, 6)$.

Conclusões

Neste trabalho investigamos a série histórica de dados de velocidade do vento medidas de hora em hora para uma estação meteorológica situada na cidade de Petrolina-PE. Analisamos as matrizes de probabilidade de transição de modelos de cadeia de Markov até ordem 3, utilizando 13 estados de velocidade do vento. Nossos resultados indicam que a ordem ótima do modelo de cadeia de Markov analisado é $k = 1$, utilizando os critérios de informação AIC, AICc, BIC e HQIC. Portanto, a ocorrência do estado seguinte depende apenas do estado atual observado e não de estados anteriores. Em trabalhos futuros, essas informações podem ser úteis para simulação e/ou previsão de séries de velocidade

do vento, de acordo com a matriz de probabilidade de transição do modelo de cadeia de Markov de ordem 1 encontrado.

Agradecimentos

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001, com o apoio da Fundação de Amparo à Ciência e Tecnologia do Estado de Pernambuco (FACEPE) e com o apoio do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq). Os autores agradecem ao Instituto Nacional de Meteorologia (INMET) pela concessão dos dados.

Referências bibliográficas

- ARUMUGAM, P.; KARTHIK, S. M. Stochastic Modelling in Yearly Rainfall at Tirunelveli District, Tamil Nadu, India. *Materials Today: Proceedings*, v.5, n.1, p.1852-1858, 2018.
- D'AMICO, G.; PETRONI, F.; PRATTICO, F. First and second order semi-Markov chains for wind speed modeling. *Physica A: Statistical Mechanics and its Applications*, v.392, n.5, p.1194-1201, 2013.
- DENI, S. M.; JEMAIN, A. A.; IBRAHIM, K. Fitting optimum order of Markov chain models for daily rainfall occurrences in Peninsular Malaysia. *Theoretical and Applied Climatology*, v.97, n.1-2, p.109-121, 2009.
- GIORGIO, M.; LANGELLA, R; TESTA, A. Markov chain modeling for very-short-term wind power forecasting. *Electric Power Systems Research*, v.122, p.152-158, 2015.
- HANNAN, E. J.; QUINN, B. G. The determination of the order of an autoregression. *Journal of the Royal Statistical Society: Series B (Methodological)*, v.41, n.2, p.190-195, 1979.
- HOCAOGLU, F. O.; SERTTAS, F. A novel hybrid (Mycielski-Markov) model for hourly solar radiation forecasting. *Renewable Energy*, v.108, p.635-643, 2017.
- HURVICH, C. M.; TSAI, C.-L. Regression and time series model selection in small samples. *Biometrika*, v.76, n.2, p.297-307, 1989.
- INSTITUTO NACIONAL DE METEOROLOGIA. 2019. URL <http://www.inmet.gov.br/portal/>.
- JALE, J. S.; JÚNIOR, S. F. A. X.; XAVIER, É. F. M.; STOŠIĆ, T.; STOŠIĆ, B; FERREIRA, T. A. E. Application of Markov chain on daily rainfall data in Paraíba-Brazil from 1995-2015. *Acta Scientiarum. Technology*, v.41, p.e37186, 2019.

Sigmae, Alfenas, v.8, n.2, p. 385-392, 2019.

64ª Reunião da Região Brasileira da Sociedade Internacional de Biometria (RBRAS).
18º Simpósio de Estatística Aplicada à Experimentação Agronômica (SEAGRO).

- KWON, M.; KWON, H.-H.; HAN, D. A spatial downscaling of soil moisture from rainfall, temperature, and AMSR2 using a Gaussian-mixture nonstationary hidden Markov model. *Journal of hydrology*, v.564, p.1194-1207, 2018.
- LI, S.; MA, H.; LI, W. Typical solar radiation year construction using k-means clustering and discrete-time Markov chain. *Applied Energy*, v.205, p.720-731, 2017.
- LI, J.; ZHANG, X.; ZHOU, X.; LU, L. Reliability assessment of wind turbine bearing based on the degradation-Hidden-Markov model. *Renewable Energy*, v.132, p.1076-1087, 2019.
- MA, J.; FOULADIRAD, M.; GRALL, A. Flexible wind speed generation model: Markov chain with an embedded diffusion process. *Energy*, v.164, p.316-328, 2018.
- MIAO, S.; XIE, K.; YANG, H.; TAI, H.-M.; HU, B. A. Markovian wind farm generation model and its application to adequacy assessment. *Renewable Energy*, v.113, p.1447-1461, 2017.
- MIAO, S.; NING, G.; GU, Y.; YAN, J.; MA, B. Markov Chain model for solar farm generation and its application to generation performance evaluation. *Journal of Cleaner Production*, v.186, p.905-917, 2018.
- MITRA, A.; SESHADRI, A. K. Detection of spatiotemporally coherent rainfall anomalies using Markov Random Fields. *Computers & Geosciences*, v.122, p.45-53, 2019.
- MUNKHAMMAR, J.; WIDÉN, J. A Markov-chain probability distribution mixture approach to the clear-sky index. *Solar Energy*, v.170, p.174-183, 2018a.
- MUNKHAMMAR, J.; WIDÉN, J. An N-state Markov-chain mixture distribution model of the clear-sky index. *Solar Energy*, v.173, p.487-495, 2018b.
- MUNKHAMMAR, J.; WIDÉN, J. A spatiotemporal Markov-chain mixture distribution model of the clear-sky index. *Solar Energy*, v.179, p.398-409, 2019.
- SCHWARZ, G. Estimating the dimension of a model. *The Annals of Statistics*, v.6, n.2, p.461-464, 1978.
- SONG, Z.; JIANG, Y.; ZHANG, Z. Short-term wind speed forecasting with Markov-switching model. *Applied Energy*, v.130, p.103-112, 2014.
- TANG, J.; BROUSTE, A.; TSUI, K. L. Some improvements of wind speed Markov chain modeling. *Renewable Energy*, v.81, p.52-56, 2015.
- TAGLIAFERRI, F.; HAYES, B. P.; VIOLA, I. M.; DJOKIĆ, S. Z.. Wind modelling with nested Markov chains. *Journal of Wind Engineering and Industrial Aerodynamics*, v.157, p.118-124, 2016.

Sigmae, Alfenas, v.8, n,2, p. 385-392, 2019.

64^a Reunião da Região Brasileira da Sociedade Internacional de Biometria (RBRAS).
18^o Simpósio de Estatística Aplicada à Experimentação Agronômica (SEAGRO).

TONG, H. Determination of the order of a Markov chain by Akaike's information criterion. *Journal of Applied Probability*, v.12, n.3, p.488-497, 1975.

WORLD ENERGY COUNCIL. 2019. URL <https://www.worldenergy.org/>.

YEH, C.-K.; LIAW, S.-C. Application of landscape metrics and a Markov chain model to assess land cover changes within a forested watershed, Taiwan. *Hydrological Processes*, v.29, n.24, p.5031-5043, 2015.