



XIII Congresso Nacional de **MEIO AMBIENTE**

de Poços de Caldas

XIII CONGRESSO NACIONAL DE MEIO AMBIENTE DE POÇOS DE CALDAS

21, 22 E 23 DE SETEMBRO DE 2016 pocos.com.br

FORMAÇÃO DE REGIÕES HOMOGÊNEAS DE PRECIPITAÇÃO POR AGRUPAMENTO FUZZY C-MEANS

**Evanice Pinheiro Gomes(1); Claudio José Cavalcante Blanco(2), Francisco Carlos Lira
Pessoa(3)**

(1) Mestranda do Programa de Pós Graduação de Engenharia Civil PPGEC. Instituto de Tecnologia. Universidade Federal do Pará UFPA. Av. Augusto Corrêa, 01 – Guamá, Belém, Pará. gomesevanice@ufpa.br;

(2) Professor Dr. Coordenador do Programa de Pós Graduação em Engenharia de Recursos Naturais da Amazônia - PRODENA/ITEC/UFPA. Av. Augusto Corrêa, 01 – Guamá, Belém, Pará. blanco@ufpa.br. (3) Professor Dr. do Instituto de Tecnologia da Universidade Federal do Pará UFPA. Av. Augusto Corrêa, 01 – Guamá, Belém, Pará. fclpessoa@ufpa.br.

Eixo temático: Gerenciamento de Recursos Hídricos e Energéticos

RESUMO – No presente estudo, utilizou-se a técnica de Análise de Agrupamento, através do método Fuzzy C-means para identificação de regiões hidrologicamente homogêneas de precipitação, na Região Hidrográfica Tocantins Araguaia - RHTA, a partir de variáveis físico climáticas, tais como a precipitação média anual, altitude e localização das estações pluviométricas (latitude e longitude). Foram definidas duas regiões homogêneas de chuva, através do agrupamento Fuzzy C-means, e as mesmas foram validadas pelo índice PBM. Essas regiões poderão auxiliar em estudos de regionalização e contribuir para o planejamento e gestão dos recursos hídricos.

Palavras-chave: Precipitações. Fuzzificação. Recursos Hídricos. Bacia Hidrográfica.

ABSTRACT – In this study, we used the cluster analysis technique using Fuzzy C-Means method for identifying hydrologically homogeneous regions of precipitation, the Hydrographic Region Tocantins Araguaia - RHTA from climate physical variables such as average rainfall annual, altitude and location of rainfall stations (latitude and longitude). two homogeneous rainfall regions were defined by grouping Fuzzy C-means, and they have been validated by the PBM index. These regions may assist in regionalization studies and contribute to the planning and management of water resources.

Key words: Precipitations. Fuzzification. Water resources. Watershed.

Introdução

A disponibilidade de precipitação em uma bacia é um fator determinante para quantificar, entre outros, a necessidade de irrigação de culturas, o abastecimento de água doméstico e industrial e o controle de inundação e erosão do solo.

Por toda a sua importância, o conhecimento da precipitação é indispensável. No entanto, um dos problemas que se apresentam em estudos de precipitação, é a



XIII Congresso Nacional de **MEIO AMBIENTE** de Poços de Caldas

XIII CONGRESSO NACIONAL DE MEIO AMBIENTE DE POÇOS DE CALDAS

21, 22 E 23 DE SETEMBRO DE 2016 pocos.com.br

falta de monitoramento pluviométrico, gerando a ausência de informação ao longo do tempo e espacialmente nas bacias hidrográficas.

Assim, para tentar preencher essa falta de dados, o desenvolvimento de modelos de estimativas de chuvas, em bacias sem monitoramento, tem se tornado muito comum. Dentre eles a regionalização hidrológica, que segundo Samuel *et. al.* (2011), consiste na utilização de um conjunto de ferramentas que buscam formas de transferência de informações de um local para outro da bacia hidrográfica, com o intuito de suprir esta carência de informação, em uma dada região considerada homogênea.

Uma das condições para a aplicação da regionalização é a identificação de regiões homogêneas, considerada uma das etapas iniciais no seu desenvolvimento. A identificação dessas regiões pode ser realizada de diversas maneiras, como por exemplo, através de funções de probabilidade, análise de agrupamento, entre outros.

A identificação de regiões hidrologicamente homogêneas possui duas finalidades: uma descritiva, cujo objetivo é a caracterização hidrológica, e outra como subsídio essencial na regionalização. Assim, diversos trabalhos de identificação de regiões homogêneas têm sido realizados, no mundo inteiro, objetivando a aplicação da regionalização, como por exemplo, os trabalhos desenvolvidos por: Dikbas *et al.*, (2012) na Turquia; Goyal e Gupta, (2014) no nordeste indiano; Santos (2014) na Amazônia brasileira e Parracho e Rocha, (2015) na Península Ibérica.

Assim o objetivo deste trabalho é identificar regiões homogêneas de chuva na Região Hidrográfica Tocantins-Araguaia – RHTA, através do método de análise de agrupamento, Fuzzy C-means, utilizando algumas características físico climáticas da bacia, como a localização geográfica das estações pluviométricas (latitude, longitude), altitude e precipitação média anual, permitindo, assim, auxiliar em estudos de regionalização, contribuir para o entendimento do comportamento da precipitação na bacia hidrográfica e apoio na gestão dos recursos hídricos.

Material e Métodos

Área de Estudo e base de Dados

A RHTA localiza-se entre o paralelo sul 0° 30' e 18° 05' e os meridianos de longitude oeste 45° 45' e 56° 20'. Possui uma área de 918.822 km², vazão média de 13.799 m³/s, possui clima tropical, com temperatura média anual de 26° C. A precipitação média na região é de 1.744 mm (ANA, 2009).

Adotou-se as séries históricas de chuvas de 84 estações pluviométricas do banco de dados da Agência Nacional de Águas - ANA, localizadas na RHTA (Figura 1). Utilizou-se as variáveis de precipitação anual média, a altitude e a localização (latitude e longitude) de cada posto pluviométrico, para aplicar a análise de agrupamento e assim identificar as regiões homogêneas de chuva.



XIII Congresso Nacional de **MEIO AMBIENTE**

de Poços de Caldas

XIII CONGRESSO NACIONAL DE MEIO AMBIENTE DE POÇOS DE CALDAS

21, 22 E 23 DE SETEMBRO DE 2016 pocos.com.br

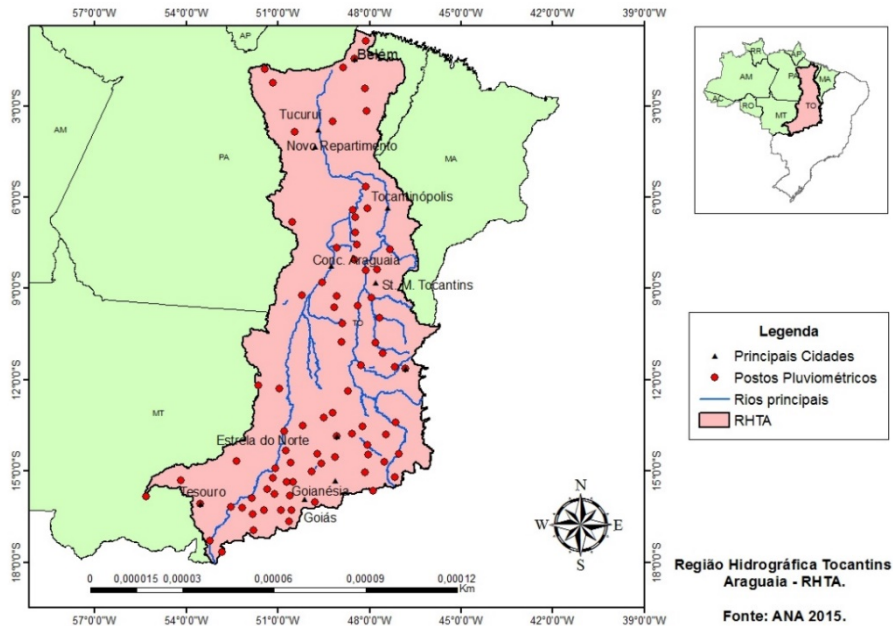


Figura 1 – Região Hidrográfica Tocantins Araguaia.

Análise de Agrupamento Fuzzy C-means

Para a determinação das regiões hidrológicamente homogêneas de chuva, foi utilizado o método de análise de agrupamento, Fuzzy c-means, do tipo não hierárquico, proposto por Dunn (1973) e depois generalizado por Bezdek (1981).

Este tipo de agrupamento é realizado através de um algoritmo conhecido como FCM, que tem como objetivo encontrar grupos fuzzy, baseados no grau de pertinência, para um conjunto de dados. Para alcançar este objetivo, o algoritmo precisa minimizar uma função que diz respeito à minimização das distâncias entre os dados e os centros dos grupos, aos quais tais dados pertencem com algum grau de pertinência (XU & WUNSCH, 2005).

O agrupamento Fuzzy C-means procura a partição que minimiza a função objetivo, representada pela Equação 1.

(1)

Em que n é o número de dados; p é o número de grupos; μ_j é o grau de pertinência da amostra j -ésimo cluster; m é o parâmetro de fuzificação; d_{ij} é a distância euclidiana entre x_i e μ_j ; x_i é um vetor de dados, onde $i=1, 2, \dots, n$, e representa um atributo de dados; C_j é o centro de um agrupamento Fuzzy.

Segundo Bezdek (1992), o desempenho do algoritmo é influenciado pela escolha do número de classes c , dos centros de clusters (grupos) iniciais, da ordem na qual os vetores de dados são processados, da medida de distância, do critério de parada e do parâmetro de fuzificação.



XIII Congresso Nacional de **MEIO AMBIENTE**

de Poços de Caldas

XIII CONGRESSO NACIONAL DE MEIO AMBIENTE DE POÇOS DE CALDAS

21, 22 E 23 DE SETEMBRO DE 2016 pocos.com.br

O parâmetro de fuzificação (m) é conhecido, também, como o expoente de peso Fuzzy. Para Hall e Mins (1999), o índice de fuzificação (m) é um parâmetro que controla o nível de difusividade no processo de classificação. Desta forma, para $m = 1$, os agrupamentos apresentam limites rígidos equivalentes aos do K-means e, à medida que o valor cresce, os limites tornam-se mais difusos. Segundo Cox (2005), usualmente, m está no intervalo de 1,25 a 2,0.

Para executar o algoritmo FCM são utilizados, as variáveis físico climáticas (precipitação anual média, altitude e localização), de cada estação pluviométrica (X_i), o parâmetro de fuzificação (m) e a indicação do número de grupos (p). Assim para um dado X_i , o seu maior grau de pertinência u_{ij} , determinará a qual agrupamento este X_i pertencerá.

Índice PBM

Para avaliar se a partição gerada pelo agrupamento Fuzzy C-means é adequada e de boa qualidade para os dados, adotou-se o Índice PBM, proposto por PAKHIRA *et al.* (2004). Esse índice serve para validar o agrupamento, indicando qual o melhor número de grupos.

O Índice PBM é definido como o produto de três fatores, conforme a Equação 2, dos quais a maximização assegura que a partição tem um pequeno número de grupos compactos com grande separação entre pelo menos dois deles PAKHIRA *et al.* (2004).

(2)

Em que K é o número de grupos; E_1 é a soma das distâncias de cada amostra ao centro geométrico de todas as amostras; E_k é a soma das distâncias entre grupos de K agrupamentos e representa a máxima separação de cada par de agrupamentos.

Quanto maior o índice PBM, melhor a partição. Assim, o índice PBM é um índice de otimização, de forma que para se obter a melhor partição, deve-se processar o algoritmo para diversos valores de K e escolher aquele que resultar no maior valor do índice PBM (MODENESI, 2008).

Resultados e Discussão

Com as variáveis de chuva média anual, altitude e localização (latitude e longitude), aplicou-se a Análise de agrupamento pelo método Fuzzy c-means, em ambiente computacional, onde foram testados 7 agrupamentos, variando o número de grupos, de 2 a 8, e 9 parâmetros de fuzificação (m), de 1,2 a 2,0. O resultado do agrupamento são os graus de pertinência de cada estação.

A estação E1, por exemplo, (Figura 2), apresenta para o Grupo 1, grau de pertinência de 0,82 e para o grupo 2 de 0,19. Isso significa que esta estação possui aproximadamente 80% de probabilidade de pertencer ao grupo 1. As estações foram alocadas em grupos conforme seus graus de pertinência, onde a maior pertinência, classifica-a como pertencente ao grupo.



XIII Congresso Nacional de MEIO AMBIENTE

de Poços de Caldas

XIII CONGRESSO NACIONAL DE MEIO AMBIENTE DE POÇOS DE CALDAS

21, 22 E 23 DE SETEMBRO DE 2016 www.pocos.com.br

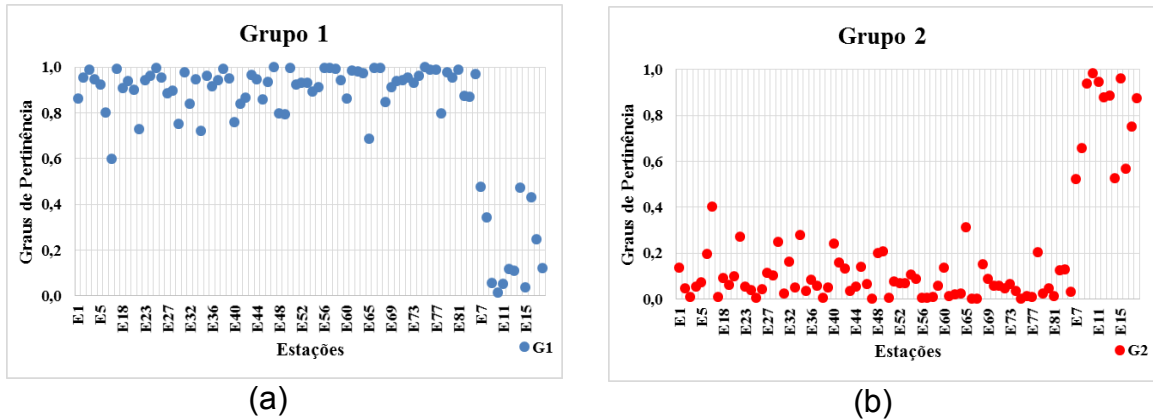


Figura 2 – Graus de pertinência das Estações.

O índice PBM indicou como melhor agrupamento, entre os 63 realizados, o formado por dois grupos, tendo como $K = 0,0072$ e parâmetro de fuzzificação (m) de 1,2. A Tabela 1 mostra o comportamento dos índices PBM de cada grupo formado. Nota-se que para o agrupamento desses postos pluviométricos, o índice PBM foi maior para todos os agrupamentos com dois grupos ($p = 2$) e menor para oito grupos, indicando que o agrupamento mais adequado seria o formado por dois grupos com máximos índice PBM.

Os grupos formados pela análise de agrupamento Fuzzy C-means, são considerados como as regiões homogêneas de chuva (Figura 3). A Região 1 é formada por 72 estações, que apresentam chuvas com média de 1.610 mm, a mínima de 1.144 mm e a máxima de 1.948 mm. Estas estações estão concentradas na porção central e sul da RHTA, mais especificamente nas sub bacias do Alto Tocantins e Araguaia, onde predomina o bioma do cerrado, de clima menos úmido, com baixo índice pluviométrico.

A Região 2 é formada por 12 estações, que apresentam chuvas de maior intensidade, com média de 2.370 mm, a mínima de 1.931 mm e a máxima de 2.843 mm. As estações deste grupo estão concentradas na porção norte da RHTA, sub bacia Alto Tocantins, onde predomina o bioma Amazônico, com clima quente e úmido e alto índice pluviométrico.

Tabela 1 - Índice PBM para cada agrupamento e parâmetro de fuzzificação.

p	$m=1,2$	$m=1,3$	$m=1,4$	$m=1,5$	$m=1,6$	$m=1,7$	$m=1,8$	$m=1,9$	$m=2,0$
2	0,007	0,006	0,006	0,006	0,006	0,006	0,006	0,006	0,005
3	2	9	9	6	4	3	1	1	6
4	0,003	0,003	0,003	0,003	0,003	0,003	0,003	0,003	0,003
5	3	2	2	2	1	2	2	2	2
6	0,002	0,002	0,002	0,002	0,002	0,002	0,002	0,002	0,002
7	5	6	7	7	8	9	9	9	9
8	0,002	0,001	0,001	0,001	0,001	0,002	0,002	0,002	0,002
9	0	5	8	8	9	0	0	0	0
10	0,000	0,001	0,001	0,001	0,000	0,001	0,001	0,001	0,001



XIII Congresso Nacional de MEIO AMBIENTE de Poços de Caldas

XIII CONGRESSO NACIONAL DE MEIO AMBIENTE DE POÇOS DE CALDAS

21, 22 E 23 DE SETEMBRO DE 2016 www.pocos.com.br

	9	5	0	5	8	4	4	4	5
7	0,000	0,001	0,000	0,001	0,000	0,000	0,000	0,001	0,001
	8	1	8	1	8	9	9	2	1
8	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
	6	6	6	1	6	7	6	9	9

Fonte: Autor, 2016.

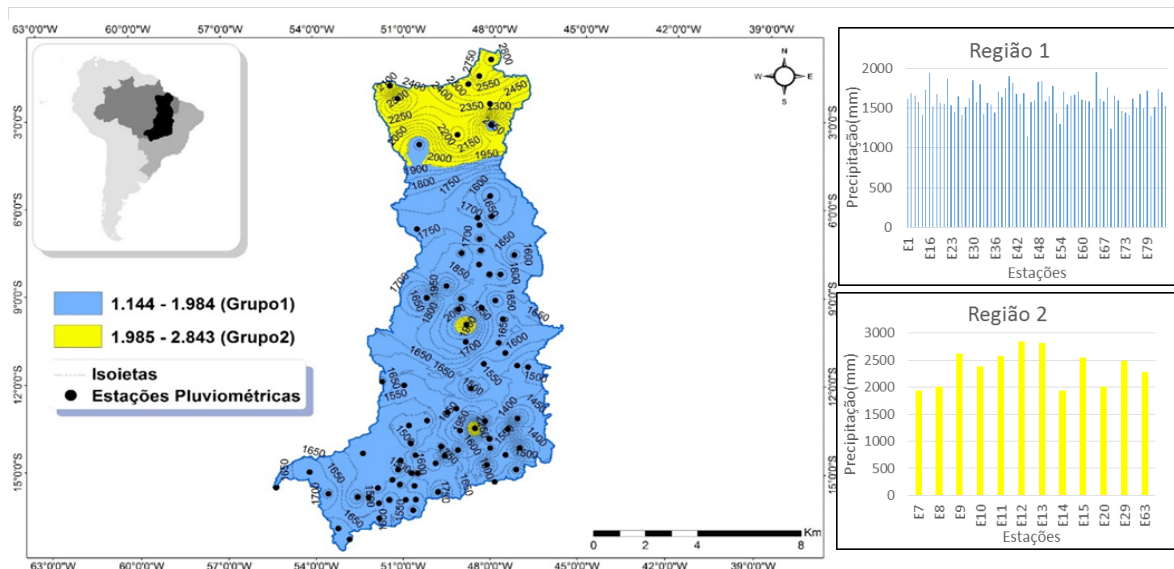


Figura 3 – Regiões Homogêneas de Chuva na RHTA.

Fonte: Autor, 2016.

Conclusões

O método de Análise de Agrupamento Fuzzy C-means, mostrou-se satisfatório, pois a partir de sua aplicação foi possível identificar duas regiões homogêneas de precipitação bem definidas e com características bem similares entre os grupos formados. Além disso o índice de validação PBM, conseguiu validar o agrupamento formado, auxiliando na escolha dos grupos, tornando a decisão mais rápida, eficiente e sem subjetividade.

As Regiões homogêneas de precipitação formaram-se de maneira bem compactas e representativas do ponto de vista quantitativo de precipitação, onde a Região 2 pode ser caracterizada como uma região com maiores volumes de precipitações, e a Região 1 com menores volumes de precipitações. Esse agrupamento representou a realidade do comportamento da precipitação nesta região, pois conseguiu ratificar estudos já existentes de precipitação, realizados pela Agência Nacional de Águas - ANA (2009), nesta região.

A identificação dessas regiões homogêneas de precipitação, contribuem para o entendimento do comportamento pluviométrico na RHTA e podem auxiliar na gestão dos recursos hídricos.

Agradecimento(s)

Agradece-se ao CNPq pela bolsa através do processo nº 133296/2015-6 – Modalidade Mestrado-GM.



XIII Congresso Nacional de **MEIO AMBIENTE** de Poços de Caldas

XIII CONGRESSO NACIONAL DE MEIO AMBIENTE DE POÇOS DE CALDAS

21, 22 E 23 DE SETEMBRO DE 2016 www.meioambiente.pocos.com.br

Referências

ANA, Agência Nacional de Águas (2009). Plano Nacional de Recursos Hídricos da Região Hidrográfica Tocantins Araguaia. Disponível em: <http://www.ana.gov.br/om>. Acessado em 09 de abril, 2016.

BEZDEK, J. Pattern Recognition with Fuzzy Objective Function Algorithms, Plenum Press, New York, 1981.

BEZDEK, J. C. PAL, S. K. Fuzzy models for pattern recognition: methods that search for structures in data. Methods that search for structures in data. IEEE Press, New York, 1992.

COX, E. Fuzzy modeling and genetic algorithms for data mining and exploration [S.1]: Elsevier/Morgan Kaufmann, 2005. Hardcover. (Morgan Kaufmann series in data management systems).

DIKBAS, Fatin at al. Classification of precipitation series using fuzzy cluster method. Journal of Climatology, v. 32, p. 156-1603, abril 2011. Disponível em: <http://onlinelibrary.wiley.com/doi/10.1002/joc.2350/abstract>. Acessado em 09 de abril 2016.

DUNN, J. C. A. A fuzzy relative of the ISODATA process and its use in detecting compact well-separated clusters. Cybernetics and Systems, 1973.

GOYAL, M. K., GUPTA V., Identification of Homogeneous Rainfall Regimes in Northeast Region of India using Fuzzy Cluster Analysis. In: Water Resour Manage. DOI 10.1007/s11269-014-0699-7. Agosto de 2014.

HALL, M. J.; MINNS, W. A. The classification of hydrologically homogeneous regions. Hydrological Sciences Journal, v. 44, n. 5, p. 693-704, outubro 1999. Disponível em: <http://www.tandfonline.com/doi/pdf/10.1080/02626669909492268>. Acessado em 12 de abril 2016.

HOSKING, J., WALLIS J. Regional Frequency Analysis: Na Approach Based on L-moments. Cambridge University, 1997.

MONDENESI M. V. Análise de Agrupamentos FCM utilizando processamento Paralelo. COOPE/UFRJ. Rio de Janeiro, 2008.

PAKHIRA M. K.; BANDYOPADHYAY S.; MAULIK K. Validity index for crisp and fuzzy clusters, Pattern Recognition, v. 37, n. 3, p. 481-501, março de 2004. Disponível em: <http://www.sciencedirect.com/science/article/pii/S0031320303002838>. Acessado em 21 de março 2016.

PARRACHO, A. C.; ROCHA, MELO GONÇALVES, P., ROCHA, A. Regionalization of precipitation for the Iberian Peninsula and climate change. In: Physics and Chemistry of the Earth. Julho, 2015. Disponível em <http://www.sciencedirect.com/science/article/pii/S1474706515000777>.



XIII Congresso Nacional de **MEIO AMBIENTE** de Poços de Caldas

XIII CONGRESSO NACIONAL DE MEIO AMBIENTE DE POÇOS DE CALDAS

21, 22 E 23 DE SETEMBRO DE 2016 www.meioambiente.pocos.com.br

SANTOS, et al. Precipitation regionalization of the Brazilian Amazon. In: Atmospheric Science Letters. Let.16, p. 185-192, setembro 2014. Disponível em: <http://onlinelibrary.wiley.com/doi/10.1002/joc.2350/abstract>. Acessado em 20 de março de 2016.

SAMUEL, J.; COULIBALY, P.; METCALFE, R. A. Estimation of continuous streamflow in Ontario ungauged basins: comparison of regionalization methods. Journal of Hydrologic Engineering, v. 16, n. 5, p. 447-459, Abril de 2011. Disponível em [http://ascelibrary.org/doi/abs/10.1061/\(ASCE\)HE.1943-5584.0000338](http://ascelibrary.org/doi/abs/10.1061/(ASCE)HE.1943-5584.0000338). Acessado em 23 de março 2016.

XU, R. E WUNSCH, D. II survey of clustering algorithms. IEEE Transactions on Neural Networks, v. 16, n. 3, p. 645-678, maio 2005. Disponível em: <http://ieeexplore.ieee.org/xpl/login.jsp>. Acessado em 24 março de 2016.