

Artigo

Aplicação da Distribuição Nakagami na Análise de Dados de Precipitação

Josmar Mazucheli¹, Isabele Picada Emanuelli^{2,3}¹Departamento de Estatística, Universidade Estadual de Maringá, Maringá, PR, Brasil.²Programa de Pós-Graduação em Tecnologias Limpas, Centro Universitário de Maringá, Maringá, PR, Brasil.³Instituto Cesumar de Ciência, Tecnologia e Inovação, Maringá, PR, Brasil.

Recebido em: 8 de Março de 2016 - Aceito em: 7 de Abril de 2018

Resumo

Este trabalho teve como objetivo avaliar o desempenho da distribuição Nakagami na análise de séries de precipitação mensal, ao longo de vários anos, visando à seleção de uma distribuição útil para o planejamento e gestão de atividades dependentes dos índices de precipitação na Região Sul do Brasil. Para tanto, compara-se a mesma com cinco distribuições alternativas: Weibull, Gama, Log-Normal, Log-Logística e Inversa-Gaussiana. Foram utilizadas séries históricas de 33 estações meteorológicas observadas entre janeiro de 1970 a dezembro de 2014, totalizando 396 séries (33 estações × 12 meses). Para a escolha da distribuição, que forneceu o melhor ajuste, foram utilizados os valores dos critérios de informação de Akaike, de Kolmogorov-Smirnov, de Anderson-Darling e de Cramér-von Mises. Segundo estes critérios se encontrou que as distribuições Nakagami e Weibull foram selecionadas o maior número de vezes (Nakagami: 146 vezes e Weibull: 100 vezes). Embora a distribuição Nakagami não seja muito utilizada, na descrição de dados de precipitação, recomenda-se sua utilização na descrição do comportamento da pluviosidade mensal como alternativa para distribuições tradicionalmente utilizadas.

Palavras-chave: ajuste de distribuições, variações climáticas, dados de precipitação, distribuição Nakagami, verossimilhança.

The Nakagami Distribution Applied in Precipitation Data Analysis

Abstract

This research had as its objective evaluating if Nakagami distribution achieves in precipitation data analysis, by considering monthly data series over several years, in an attempt to select a useful distribution that helps in the management and administration of those activities that depends on Southern Brazil precipitation index. For this purpose Nakagami distribution was compared with five alternatives: Weibull, Gamma, Log-Normal, Log-Logistic e Inverse-Gaussian. In the analysis, data series belonging to 33 weather stations between January 1970 and December 2014 were used, amounting 396 data series (33 weather × stations 12 months). In order to choose the best distribution were used the Akaike Information, the Kolmogorov-Smirnov, then Anderson-Darling and the Cramér-von Mises criterions. According to these criterions, it can be found that Nakagami and Weibull distributions were selected the greatest number of times (Nakagami: 146 times and Weibull: 100 times). Even though Nakagami distribution is not too used in rain precipitation data description, we recommend its application in the description of monthly precipitation behavior as an alternative of distributions that have been traditionally applied.

Keywords: climate variation, distributions adjustment, maximum likelihood, Nakagami distribution, precipitation data.

1. Introdução

A Região Sul do Brasil tem grande parte de sua economia baseada na produção agropecuária, o que gera preocupações em torno das variações climáticas ao longo

do ano. A intensidade e a frequência com que os fenômenos climáticos ocorrem influenciam diretamente a sustentabilidade ambiental e econômica dos setores do agronegócio (Beyruth, 2008).

Para realização de estudos com variáveis climáticas, a World Meteorological Organization (WMO) recomenda séries históricas com o mínimo de trinta anos. De maneira geral, dentre os elementos climáticos, é possível observar que a precipitação é o elemento que possui menor precisão das estimativas tomando como base trinta anos de observação (Cargnelutti Filho *et al.*, 2009). Um estudo recente observou que trinta anos são suficientes para prever a precipitação na maioria dos meses do ano, com exceção de junho, julho e agosto, uma vez que para esses meses podem ser necessários até sessenta e três anos (Torres *et al.*, 2016).

A caracterização da distribuição da precipitação constitui um importante instrumento de apoio para estudos e desenvolvimento de cultivares. Para descrever dados de precipitação e compreender esse fenômeno climatológico, ao longo do tempo, um grande número de distribuições de probabilidades vem sendo utilizado. O estudo da distribuição de variáveis climáticas determina padrões de ocorrência, permitindo uma previsão razoável do comportamento climático de uma região (Dourado Neto *et al.*, 2005).

As distribuições de probabilidade como a Weibull, a Gama, a Log-Normal, a Log-Logística e a Inversa-Gaussiana, geralmente especificadas por dois parâmetros, são tradicionalmente usadas na modelagem de dados de natureza climática (Coe e Stern, 1982). Ramos e Moala (2014) comparam as distribuições Gama, Weibull e Log-Normal com a distribuição Exponencial Geométrica Estendida mostrando que essa distribuição, menos usual para dados climatológicos, pode ser aplicada com sucesso em dados de precipitação. Uma família de distribuições também muito utilizada, só que em geral na modelagem de extremos, é a distribuição Valor Extremo Generalizada (Kotz e Nadarajah, 2000).

Por outro lado, o mesmo não ocorre com a distribuição Nakagami, cujas aplicações se concentram, principalmente, na área de engenharia de comunicações. Essa distribuição foi proposta pela primeira vez por Nakagami, em 1960, para descrever a propagação de ondas e, atualmente, é muito utilizada para modelar ambientes micro celulares e celulares, nos quais há desvanecimento rápido (Antônio, 2003). Se comparada com outras distribuições, a distribuição Nakagami é bastante simples e flexível em termos do comportamento de sua função densidade de probabilidade. No entanto, para variáveis climatológicas, esta distribuição parece ser pouco usual, principalmente, para dados de precipitação que não constam na literatura em nenhuma aplicação.

Baseado nessas informações, este trabalho teve como objetivo avaliar o desempenho da distribuição Nakagami na análise de séries de precipitação mensal ao longo dos anos. Para tanto, o presente estudo comparou o uso da distribuição Nakagami como alternativa para cinco distribuições convencionais: Weibull, Gama, Log-Normal, Log-Logística e Inversa-Gaussiana.

2. Material e Métodos

2.1. Os dados

Foram utilizadas 396 séries históricas da precipitação mensal de 33 estações meteorológicas localizadas na Região Sul do Brasil: Paraná (8 estações, 96 séries), Santa Catarina (7 estações, 84 séries) e Rio Grande do Sul (18 estações, 216 séries). As séries foram obtidas no Portal do Instituto Nacional de Meteorologia (INMET), no qual são disponibilizadas séries históricas de 265 estações meteorológicas convencionais do Brasil. Na Tabela 1 apresenta-se o código WMO, a latitude e a longitude das estações utilizadas.

De cada estação foram geradas doze séries para cada ano, uma para cada mês, com período de observação compreendido ente janeiro de 1970 a dezembro de 2014, totalizando 44 anos. Possivelmente, por algum problema operacional nas estações, as séries utilizadas não possuem o mesmo número de observações e variam de 22 a 44, com mediana de 39 observações.

Antes de conduzir o ajuste e a discriminação das distribuições foram avaliadas a validade das suposições de homogeneidade (as observações são identicamente distribuídas) e aleatoriedade (as observações são independentes), respectivamente, pela aplicação dos testes de Von-Neumann (Kangand e Yusof, 2012; Buishand, 1982) e Wald-Wolfowitz (Gibbons e Chakraborti, 2001).

O método da máxima verossimilhança foi o método adotado para a obtenção das estimativas dos parâmetros, enquanto que a qualidade do ajuste da distribuição selecionada, segundo os critérios descritos abaixo, foi avaliada a partir da aplicação dos testes: Kolmogorov-Smirnov, Anderson-Darling e Cramér-von Mises. Na avaliação da qualidade do ajuste, por esses critérios, pelo menos um deles não rejeitou o ajuste da distribuição selecionada em nível de significância de 5%.

Utiliza-se a função *mledist* da biblioteca *fitdistrplus* (Delignette-Muller e Dutang, 2015), do ambiente *R*, para encontrar as estimativas de máxima verossimilhança e os valores das estatísticas usadas na discriminação das seis distribuições. Os valores-p associados aos testes Kolmogorov-Smirnov, Anderson-Darling e Cramér-von Mises foram calculados pelas funções *ks.test*, *ad.test* e *cvm.test*, disponíveis na biblioteca *gofest* (Faraway *et al.*, 2014). A biblioteca *randtests* (Caeiro e Mateus, 2014) foi utilizada na aplicação dos testes de homogeneidade e de aleatoriedade.

2.2 As distribuições

Para análise da pluviosidade total mensal são consideradas, além da distribuição Nakagami, as distribuições Weibull, Gama, Log-Normal, Log-Logística e Inversa-Gaussiana. A seguir são descritas todas elas, enfatizando a distribuição Nakagami, objeto de estudo do presente artigo.

Tabela 1 - Número das estações meteorológicas e suas respectivas coordenadas geográficas situadas na região sul do Brasil.

Rio Grande do sul			Paraná			Santa Catarina		
Estação	Latitude	Longitude	Estação	Latitude	Longitude	Estação	Latitude	Longitude
83919	-28.66	-50.43	83836	-25.46	-50.63	83891	-27.81	-50.33
83916	-28.21	-51.50	83813	-24.78	-50.00	83872	-26.90	-49.21
83914	-28.21	-52.40	83783	-24.05	-52.36	83897	-27.58	-48.56
83985	-31.78	-52.41	83811	-25.00	-50.86	83887	-27.38	-51.20
83927	-29.75	-57.08	83766	-23.31	-51.13	83923	-28.51	-49.31
83967	-30.05	-51.16	83842	-25.43	-49.26	83883	-27.11	-52.61
83907	-28.40	-55.01	83767	-23.40	-51.91	83920	-28.30	-49.93
83953	-30.83	-55.60	83844	-25.53	-48.51			
83964	-30.53	-52.51						
83912	-28.63	-53.60						
83980	-31.33	-54.10						
83948	-29.35	-49.73						
83881	-27.18	-53.23						
83942	-29.16	-51.20						
83941	-29.15	-51.20						
83936	-29.70	-53.70						
83995	-32.03	-52.11						
83997	-33.51	-53.35						

2.2.1 A Distribuição de Nakagami

Uma variável aleatória não negativa X com distribuição Nakagami tem, respectivamente, função de densidade e de distribuição escritas nas formas:

$$f(x|\alpha, \theta) = \frac{2}{\Gamma(\alpha)} \left(\frac{\alpha}{\theta}\right)^\alpha x^{2\alpha-1} \exp\left[-\frac{\alpha}{\theta}x^2\right] \quad (1)$$

e

$$F(x|\alpha, \theta) = \Gamma\left(\frac{\alpha}{\theta}x^2, \alpha, 1\right) \quad (2)$$

em que $\theta > 0$ é o parâmetro de escala, $\alpha \geq 0.5$ é o parâmetro de forma, $\Gamma(\cdot)$ é a função gama e $\Gamma(\cdot, a, b)$ é a função de distribuição acumulada de uma variável aleatória com distribuição Gama, com parâmetro de forma a e escala b . Para $\alpha = 1$ tem-se a distribuição seminormal e para $\alpha = 0.5$ a distribuição Rayleigh. A partir do método de transformação de variáveis aleatórias é possível mostrar que se, $X \sim \text{Nakagami}(\theta, \alpha)$ então $Y = X^2$ tem distribuição Gama com parâmetro de escala $\frac{\theta}{\alpha}$ e parâmetro de forma α .

O momento de ordem k da distribuição Nakagami é dado por:

$$E(X^k) = \frac{\Gamma(\alpha + 0.5k)}{\Gamma(\alpha)} \sqrt[k]{\left(\frac{\theta}{\alpha}\right)}. \quad (3)$$

Em particular, para $k = 1, 2$ tem-se:

$$E(X) = \frac{\Gamma(\alpha + 0.5)}{\Gamma(\alpha)} \sqrt{\frac{\theta}{\alpha}} \quad (4)$$

e

$$E(X^2) = \theta. \quad (5)$$

Consequentemente:

$$V(X) = \theta \left[1 - \frac{1}{\alpha} \left(\frac{\Gamma(\alpha + 0.5)}{\Gamma(\alpha)} \right)^2 \right]. \quad (6)$$

Dado um conjunto de observações independentes, $x = (x_1, \dots, x_n)$ e supondo que estas são realizações de uma variável aleatória X com distribuição Nakagami (θ, α) , tem-se que as estimativas de θ e α podem ser encontradas a partir do método da máxima verossimilhança.

De (1) tem-se a função de verossimilhança:

$$L(\alpha, \theta|x) = \left[\frac{2}{\Gamma[\alpha]} \left(\frac{\alpha}{\theta}\right)^\alpha \right]^n \prod_{i=1}^n x_i^{2\alpha-1} \exp\left[-\frac{\alpha}{\theta} \sum_{i=1}^n x_i^2\right] \quad (7)$$

cujo logaritmo pode ser escrito na forma:

$$l(\alpha, \theta|x) = n \log \left[\frac{2}{\Gamma[\alpha]} \left(\frac{\alpha}{\theta}\right)^\alpha \right] + (2\alpha - 1) \sum_{i=1}^n \log(x_i) - \frac{\alpha}{\theta} \sum_{i=1}^n x_i^2. \quad (8)$$

As estimativas de máxima verossimilhança de θ e α são obtidas pela maximização da função de verossimilhança ou log-verossimilhança. Explicitamente, (8) é maximizada

resolvendo, simultaneamente em α e θ o sistema formado pelas equações $\frac{\partial}{\partial \alpha} l(\alpha, \theta | \mathbf{x}) = 0$ e $\frac{\partial}{\partial \theta} l(\alpha, \theta | \mathbf{x}) = 0$ em que:

$$\frac{\partial}{\partial \alpha} l(\alpha, \theta | \mathbf{x}) = n \left[1 + \log \left(\frac{\alpha}{\theta} \right) - \psi(\alpha) \right] + 2 \sum_{i=1}^n \log(x_i) + \frac{1}{\theta} \sum_{i=1}^n x_i^2, \quad (9)$$

$$\frac{\partial}{\partial \theta} l(\alpha, \theta | \mathbf{x}) = \frac{n\alpha}{\theta} + \frac{\alpha}{\theta^2} \sum_{i=1}^n x_i^2 \quad (10)$$

e que $\psi(\alpha) = \frac{d}{d\alpha} \log \Gamma(\alpha)$ é a função digama.

A partir da Eq. (10) se tem que, $\hat{\theta} = \frac{1}{n} \sum_{i=1}^n x_i^2$ enquanto que $\hat{\alpha}$ é obtido numericamente a partir da equação:

$$n [\log(\alpha) - n \log(\hat{\theta}) - \psi(\alpha)] + 2 \sum_{i=1}^n \log(x_i) = 0. \quad (11)$$

As estimativas intervalares para θ e α podem ser obtidas a partir da normalidade assintótica dos estimadores de máxima verossimilhança e pela propriedade de invariância, funções de θ e α , diga-se, $g(\alpha, \theta)$ são estimadas por $g(\hat{\alpha}, \hat{\theta})$. Por exemplo, a média, $E(X)$ é estimada por $\hat{E}(X) = \frac{\Gamma(\hat{\alpha} + 0.5)}{\Gamma(\hat{\alpha})} \sqrt{\frac{\hat{\theta}}{\hat{\alpha}}}$ enquanto que a estimativa de $\text{Var}(\hat{E}(X))$ é calculada a partir da aplicação do método delta (Casella e Berger, 1990).

Em se tratando de dados de precipitação total (diária, decendial, mensal, etc.), a precipitação total mais provável ou máxima no período, representada aqui por, X_{max} obtida da equação, $\frac{d}{dX_{max}} f(X_{max} | \hat{\alpha}, \hat{\theta}) = 0$ e estimada por $\hat{X}_{max} = \sqrt{\frac{\hat{\theta}(2\hat{\alpha}-1)}{2\hat{\alpha}}}$.

2.2.2. As demais distribuições

A título de avaliar a eficiência da distribuição Nakagami, na descrição de dados de precipitação, a mesma foi comparada com cinco distribuições comuns na análise de dados de precipitação, sendo essas: Weibull, Gama, Log-Normal, Log-Logística e Inversa-Gaussiana, resumidamente caracterizadas a seguir pela função de densidade, média e moda, representada por X_{max} .

Distribuição Weibull:

$$f(x | \alpha, \theta) = \frac{\alpha}{\theta^\alpha} x^{\alpha-1} \exp \left[- \left(\frac{x}{\theta} \right)^\alpha \right], \quad (12)$$

$$E(X) = \theta \Gamma \left(\frac{\alpha+1}{\alpha} \right) e, \text{ para } \alpha > 1, \quad X_{max} = \theta \left(\frac{\alpha-1}{\alpha} \right).$$

Distribuição Gama:

$$f(x | \alpha, \theta) = \frac{1}{\theta^\alpha \Gamma(\alpha)} x^{\alpha-1} \exp \left(- \frac{x}{\theta} \right), \quad (13)$$

$$E(X) = \alpha \theta \text{ e para } \alpha > 1, \quad X_{max} = \theta(\alpha - 1)$$

Distribuição Inversa-Gaussiana:

$$f(x | \alpha, \mu) = \sqrt{\frac{\alpha}{2\pi x^3}} \exp \left[- \frac{\alpha(x - \mu)^2}{2\mu^2 x} \right], \quad (14)$$

$$E(X) = \mu \text{ e } X_{max} = \mu \sqrt{1 + \frac{9\mu^2}{4\alpha^2}} - \frac{3\mu^2}{2\alpha}.$$

Distribuição Log-Logística:

$$f(x | \alpha, \theta) = \frac{\alpha}{\theta^\alpha} x^{\alpha-1} \left[1 + \left(\frac{x}{\theta} \right)^\alpha \right]^{-2}, \quad (15)$$

$$E(X) = \frac{\theta\pi}{\text{asen}(\frac{\pi}{2})}, \text{ e para } \alpha > 1, \quad X_{max} = \theta \left(\frac{\alpha-1}{\alpha+1} \right)^{\frac{1}{\alpha}}$$

Distribuição Log-Normal:

$$f(x | \mu, \theta) = \frac{1}{\sqrt{2\pi\theta x}} \exp \left[- \frac{1}{2} \left(\frac{\log(x) - \mu}{\theta} \right)^2 \right], \quad (16)$$

$$E(X) = \exp(\mu + 0.5\theta^2) \text{ e } X_{max} = \exp(\mu - \theta^2).$$

Vale lembrar que os parâmetros, $\theta > 0$, $\alpha > 0$ e $\mu > 0$ descrevem, respectivamente, a escala, a forma e a locação destas distribuições. Apenas na distribuição Log-Normal o parâmetro μ tem suporte nos números reais.

Uma caracterização completa destas distribuições apresentadas pode ser encontrada, por exemplo, em Johnsson *et al.* (1994 e 1995).

2.3. Os critérios de discriminação

Depois de estimados os parâmetros das distribuições candidatas pelo método da máxima verossimilhança foram utilizados quatro critérios para a escolha da distribuição com melhor ajuste. A distribuição escolhida foi a que mais vezes se mostrou vencedora, segundo os valores das quatro estatísticas descritas no decorrer desta seção.

Dado um conjunto de observações, $x = (x_1, \dots, x_n)$ representando a precipitação total mensal, e as seis distribuições concorrentes, decidiu-se em favor de uma delas, a partir dos valores das estatísticas:

- Critério de informação de Akaike: $AIC = -2l + 2p$;
- Kolmogorov-Smirnov: $KS = \max_{1 \leq i \leq n} \left(\hat{z}_i - \frac{i}{n}, \frac{i}{n} - \hat{z}_i \right)$ para $i = 1, \dots, n$;
- Anderson-Darling: $AD = -n - \frac{1}{n} \sum_{i=1}^n (2i-1) [\log(\hat{z}_i) + \log(1 - \hat{z}_{n+1-i})]$;
- Cramér-von Mises: $CvM = -\frac{1}{12n} + \sum_{i=1}^n \left(\hat{z}_i - \frac{2i-1}{2n} \right)^2$.

Nestas expressões se tem que l é a verossimilhança localmente nas estimativas dos parâmetros, n é o tamanho da amostra, p o número de parâmetros, $x_{(i)}$ a i -ésima estatística de ordem da amostra e $\hat{z}_i = F(x_{(i)})$ é a estimativa da função de distribuição. Para todos os critérios decide-se

em favor da distribuição que apresenta o menor valor da estatística.

Além do critério de informação de Akaike (1974) encontra-se na literatura, dentre outros, o critério de informação de Akaike corrigido, $AICc = -2n + \frac{2np}{n-p-1}$, o critério de informação de Hannane-Quinn (1979), $HQIC = -2l + 2p \log[\log(n)]$ o critério de informação de Akaike consistente, $CAIC = -2l + p[\log(n) + 1]$ e o critério de informação de Schwarz (1978), $BIC = -2l + p \log(n)$.

Neste artigo, como as distribuições rivais são especificadas pelo mesmo número de parâmetros se pode, simplesmente, usar como critério de seleção o valor de $-2l$. Sob esta circunstância, as estatísticas AIC, AICc, HQIC, CAIC, BIC e $-2l$ sempre irão selecionar a mesma distribuição.

Embora as estatísticas de Kolmogorov-Smirnov, Anderson-Darling e Cramér-von Mises sejam, geralmente, usadas para avaliar a plausibilidade de uma dada distribuição, elas também podem ser usadas na escolha de uma dentre duas ou mais distribuições concorrentes. Ao leitor interessado em detalhes destas estatísticas se recomenda, entre outros, D'Agostino e Stephens (1986) e Stephens (1974).

3. Resultados e Discussões

A primeira avaliação realizada nos dados ocorreu quanto à validação das suposições de homogeneidade e de aleatoriedade. Todas as 396 séries foram submetidas aos testes de Von-Neumann e Wald-Wolfowitz. Apresenta-se, na Tabela 2, a distribuição do nível mínimo de significância associado a ambos os testes. Em nível de significância de 5%, 23(26) séries não satisfizeram a suposição de homogeneidade(aleatoriedade). Esses números representam 5.8% e 6.6% das 396 séries. No geral, 37 séries, 9.3%, não satisfizeram a suposição de homogeneidade, de aleatoriedade ou ambas.

Em adição aos testes Von-Neumann e Wald-Wolfowitz foram aplicados os testes de Bartels (Bartels, 1982) e Mann-Kendall (Mann, 1945; Kendall e Gibbons, 1990). As conclusões oriundas dos mesmos corroboram as conclusões obtidas com os testes de Von-Neumann e Wald-Wolfowitz; por esse motivo, seus resultados não são reportados. Uma detalhada revisão desses e vários outros testes é apresentada em Machiwal e Jha (2012, Capítulo 4, página 51).

Nas Tabelas 3 e 4 são apresentados o número e a porcentagem de vezes que cada uma das seis distribuições foi selecionada, segundo os quatro critérios de discriminação. Na Tabela 3 foram consideradas todas as 396 séries, enquanto que na Tabela 4 constam somente as 359 séries que passaram, simultaneamente, nos testes de Von-Neumann e Wald-Wolfowitz (homogeneidade e aleatoriedade, respectivamente).

Tabela 2 - Distribuição do nível mínimo de significância associado aos testes Von-Neumann e Wald-Wolfowitz.

Intervalo para o valor-p	Von-Neumann		Wald-Wolfowitz	
	Frequência Absoluta	Frequência Relativa	Frequência Absoluta	Frequência Relativa
$0.00 \leq p < 0.01$	5	0.01	6	0.02
$0.01 \leq p < 0.05$	18	0.05	20	0.05
$0.05 \leq p < 0.20$	62	0.16	67	0.17
$0.20 \leq p < 0.50$	113	0.29	108	0.27
$0.50 \leq p < 0.80$	151	0.38	118	0.30
≥ 0.80	47	0.11	77	0.19
Total	396	1.00	396	1.00

Com relação à qualidade do ajuste, avaliada pelos testes de Kolmogorov-Smirnov, Anderson-Darling e Cramér-von Mises se constata que o número de vezes que a distribuição obteve maior valor-p coincidiu com os valores observados nas Tabelas 3 e 4.

Na Tabela 5 apresentam-se os números de séries que não rejeitaram, para um nível de significância de 5%, a distribuição ajustada como sendo apropriada. É importante observar o que ocorreu com as distribuições Weibull e Nakagami. Estas distribuições não foram rejeitadas nenhuma vez, ou seja, pelo ajuste, foram vencedoras em todas as séries utilizadas, 396 vezes (33 estações \times 12 meses). Segundo os critérios utilizados na discriminação se encontra que as distribuições Nakagami e Weibull foram as distribuições selecionadas o maior número de vezes. Das 396 séries, a distribuição Nakagami foi selecionada 146 vezes, seguida da Weibull selecionada 100 vezes. Observa-se também que a escolha está relacionada ao critério de discriminação utilizado. O mesmo fato foi observado por Ashkar e Zoglat (1997). Em Ashkar e Aucoin (2012) é apontado que a razão das verossimilhanças maximizadas e o critério baseado na estatística Anderson-Darling são os critérios mais poderosos. Segundo esses dois critérios se tem que a distribuição Nakagami leva vantagens sobre as demais distribuições (ver, Tabelas 3 e 4).

É importante ressaltar, neste ponto, que a distribuição Nakagami é amplamente utilizada em várias áreas do conhecimento, especialmente, na análise do desvanecimento de sinais de rádio e de ultrassom (Karagiannidis *et al.*, 2005; Antônio, 2003; Guedes, 1994). Existe uma vasta literatura referente à aplicação na área de engenharia de comunicações. Por outro lado, quando se buscam aplicações envolvendo variáveis climatológicas, um número bastante reduzido de publicações é encontrado (Schwartz *et al.*, 2013; Singh e Sarkar, 2013; Sarkar e Rai, 2011; Sarkar e Mathur, 2010; Sarkar e Mathur, 2009). No entanto, nenhum deles utilizou dados de precipitação.

É importante enfatizar que, independentemente da frequência com que cada uma das distribuições foi

Tabela 3 - Número de séries (porcentagem) em que as distribuições de probabilidade (Weibull, Gamma, LogN, Log-Log, IGauss, Nakagami) foram selecionadas (vencedoras) em cada critério de seleção: *Akaike (AIC)*, *Kolmogorov-Smirnov (KS)*, *Anderson-Darling (AD)* e *Cramér-von Mises (CvM)*, dentre as 396 séries utilizadas. Números marcados em negrito indicam a distribuição vencedora em cada critério de seleção.

Critérios de seleção	Distribuição de probabilidade candidata					
	Weibull	Gamma	LogN	Log-Log	IGauss	Nakagami
<i>AIC</i>	100 (25,25%)	96 (24,24%)	7 (1,77%)	21 (5,30%)	26 (6,57%)	146 (36,87%)
<i>KS</i>	113 (28,54%)	63 (15,91%)	14 (3,54%)	96 (24,24%)	16 (4,04%)	94 (23,73%)
<i>AD</i>	123 (31,06%)	76 (19,19%)	11 (2,78%)	49 (12,37%)	13 (3,28%)	124 (31,32%)
<i>CvM</i>	140 (35,35%)	57 (14,39%)	13 (3,29%)	74 (18,69%)	13 (3,28%)	99 (25,00%)

Tabela 4 - Número de séries (porcentagem) em que as distribuições de probabilidade (Weibull, Gamma, LogN, Log-Log, IGauss, Nakagami) foram selecionadas (vencedoras) em cada critério de seleção: *Akaike (AIC)*, *Kolmogorov-Smirnov (KS)*, *Anderson-Darling (AD)* e *Cramér-von Mises (CvM)*, somente das séries que passaram, simultaneamente, nos testes de von-Neumann (homogeneidade) e Wald-Wolfowitz (aleatoriedade). Números marcados em negrito indicam a distribuição vencedora em cada critério de seleção.

Critérios de seleção	Distribuição de probabilidade candidata					
	Weibull	Gamma	LogN	Log-Log	IGauss	Nakagami
<i>AIC</i>	96 (26,74%)	87 (24,23%)	1,95 (7,00%)	19 (5,29%)	24 (6,69%)	126 (35,10%)
<i>KS</i>	103 (28,69%)	54 (15,04%)	14 (3,90%)	92 (25,63%)	14 (3,90%)	82 (22,84%)
<i>AD</i>	111 (30,92%)	68 (18,94%)	11 (3,06%)	47 (13,09%)	11 (3,06%)	111 (30,92%)
<i>CvM</i>	131 (36,49%)	50 (13,93%)	11 (3,06%)	70 (19,50%)	11 (3,06%)	86 (23,96%)

Tabela 5 - Número de vezes em que as distribuições sob hipótese nula não foi rejeitada em três testes de critérios de seleção [*Akaike (AIC)*, *Kolmogorov-Smirnov (KS)*, *Anderson-Darling (AD)* e *Cramér-von Mises (CvM)*], dentre as 396 séries utilizadas. Números marcados em negrito indicam as distribuições que não tiveram a hipótese nula rejeitada nenhuma vez.

Teste	Distribuição sob Hipótese Nula					
	Weibull	Gamma	LogN	Log-Log	IGauss	Nakagami
<i>KS</i>	396	395	378	396	275	396
<i>AD</i>	396	395	376	395	282	396
<i>AvM</i>	396	395	380	396	277	396

selecionada, todas elas são importantes e, sempre que possível, o maior número delas deve ser ajustada e escolhida, segundo os critérios discutidos aqui e na literatura.

4. Conclusões

Embora a distribuição Nakagami não seja muito utilizada na descrição de dados de precipitação ou de outras variáveis climatológicas, os resultados deste trabalho evidenciam seu potencial na análise de séries de precipitação.

A distribuição Nakagami foi escolhida o maior número de vezes como vencedora, obtendo o melhor ajuste para as estações climatológicas da Região Sul do Brasil para série histórica de até 44 anos. Esta distribuição permite previsões de precipitação climática, que podem servir de suporte para tomada de decisões, planejamento e gestão no agronegócio de uma região. Recomenda-se sua utilização na descrição do comportamento da pluviosidade

mensal como concorrente a algumas distribuições tradicionalmente utilizadas.

Agradecimentos

Agradecemos ao Instituto Nacional de Meteorologia (INMET) por disponibilizar no portal INMET os dados utilizados neste trabalho. O segundo autor agradece ao Instituto Cesumar de Ciência, Tecnologia e Inovação (ICETI) pela Bolsa de Produtividade em Pesquisa concedida.

Referências

- AKAIKE, H. A new look at the statistical model identification. **IEEE Trans. Automat. Contr.**, v. 19, n. 6, p. 716-723, 1974.
- ANTÔNIO, J.R. **Área de cobertura em ambiente de propagação modelado com a distribuição k-u**. 2003. 88 p. Dissertação (Mestrado em Engenharia Elétrica) - Instituto Nacional de Telecomunicações (INATEL), Santa Rita do Sapucaí. 2003. Disponível em: <http://www.wisstek.org/ie708/Jamil.pdf>. Acesso em: 10 jan. de 2016
- ASHKAR, F.; AUCOIN, F. Choice between competitive pairs of frequency models for use in hydrology: a review and some new results. **Hydrol. Sci. J.** v. 57, n. 6, p. 1092-1106, 2012.
- ASHKAR, F.; ZOGLAT A.M. On the discrimination between statistical distributions for hydrological frequency analysis. In: CSCE ANNUAL CONFERENCE. **Major Public Works: Key Technologies for the 21st Century**, Quebec, 1997. p. 169-178

- BARTELS, R. The rank version of von Neumann's ratio test for randomness. **J. Am. Stat. Assoc.**, v. 77, n. 377, p. 40-46, 1982.
- BEYRUTH, Z. Água, agricultura e as alterações climáticas globais. **Rev. Tec. & Agrop.**, 318 n. 6, p. 74-89, 2008.
- BUISSHAND, T.A. Some methods for testing the homogeneity of rainfall records. **J. Hydrol.**, v. 58, n. 1-2, p. 11-27, 1982.
- CAEIRO, F.; MATEUS, A. Randtests: Testing randomness in R. R package version 1.0. 2014. Disponível em: URL <http://CRAN.R-project.org/package=randtests>. Acesso em: 30 de abr. 2015.
- CARGNELUTTI-FILHO, A.; MATZENAUER, R.; MALUF, J. R.T.; RADIN, B. Variabilidade temporal e espacial da precisão das estimativas de elementos meteorológicos no Rio Grande do Sul. **Cienc. Rural.**, v. 39, n. 4, p. 962-970, 2009.
- CASELLA, G.; BERGER, R.L. **Statistical inference**. Série: Statistics/Probability Serie, Ed. Wadsworth & Brooks/Cole Advanced Books & Software, Pacific Grove, CA. 1990
- COE, R.; STERN, R.D. Fitting models to daily rainfall data. **J. Appl. Meteorol.**, v. 21, n. 7, p. 337 1024-1031, 1982.
- D'AGOSTINO, R.B.; STEPHENS, M.A. Goodness-of-fit techniques. **Statistics: Textbooks and Monographs**. v. 68. New York, Marcel Dekker, 1986.
- DELIGNETTE-MULLER, M.L.; DUTANG, C. Fitdistrplus: An R package for fitting distributions. **J. Stat. Software**, v. 64, n. 4, p. 1-34. 2015
- DOURADO NETO D.; ASSIS, J.P. de; TIMM, L.C.; MANFRON, P.A.; SPAROVEK, G.; FARAWAY, J.; MARSAGLIA, G.; MARSAGLIA, J.; BADDELEY, A. Gofest: Classical Goodness-of-Fit Tests for Univariate Distributions. **R package version 1.0-2**. 2014. Disponível em: URL <http://CRAN.R-project.org/package=gofest>. Acesso em: 30 de abr. 2015.
- GIBBONS, J.D.; CHAKRABORTI, S. **Nonparametric statistical inference**, 5th Edition. **Statistics: Textbooks and Monographs**. Boca Raton: CRC Press. 2011.
- GUEDES, L.G.R. **Área de cobertura e sobreposição de células no ambiente Nakagami**. 1994. 123p. Dissertação (Mestrado em Engenharia Elétrica) - Faculdade de Engenharia Elétrica e de Computação UNICAMP, Campinas. 1994. Disponível em: http://repositorio.unicamp.br/jspui/bitstream/REPOSIP/259670/1/Guedes_LeonardoGuerradeRezende_M.pdf. Acesso em: 30 de abr. 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.
- JOHNSON, N.L.; KOTZ, S.; BALAKRISHNAN, N. **Continuous univariate distributions**. v. 1, 2ed. New York: Wiley-Interscience Publication. 1994.
- JOHNSON, N. L.; KOTZ, S.; BALAKRISHNAN, N. **Continuous univariate distributions**. v. 2, 2ed. New York: Wiley-Interscience Publication. 1995.
- KANG, H.M.; YUSOF, F. Homogeneity tests on daily rainfall series in peninsular malaysia. **Int. J. Contemp. Math. Sci.**, v. 7, n. 1, p. 9-22, 2012.
- KARAGIANNIDIS, G.K.; SAGIAS, N.C.; MATHIOPOULOS P.T. The $N \times$ Nakagami Fading Channel Model. In: **Wireless Communication Systems of 2nd International Symposium**, 2005. **Anais 2nd International Symposium**; p. 185-189.
- KENDALL, M.; GIBBONS, J.D. **Rank correlation methods**, 5th Edition. A Charles Griffin Title. Edward Arnold, London. 1990.
- KOLAR, R.; JIRIK, R.; JAN, J. Estimator comparison of the Nakagami-m parameter and its application in echocardiography. **Radioengineering**, v. 13, n. 1, p. 8-12. 2004.
- KOTZ, S.; NADARAJAH, S. **Extreme value distributions**. London: Imperial College Press. 2000
- MACHIWAL, D.; JHA, M.K. **Hydrologic time series analysis: theory and practice**. Springer, Dordrecht; Capital Publishing Company, New Delhi, 2012.
- MANN, H.B. Nonparametric tests against trend. **Econometrica**, v. 13, n. 3, p. 245-259, 1945.
- NAKAGAMI, M., **The m-distribution - A general formula of intensity distribution of rapid fading**. In: HOFFMAN, W. C. (Ed.), **Statistical Methods in Radio Wave Propagation**. Pergamon, 1960. p. 3-36.
- R CORE TEAM, 2013. **R: A Language and Environment for Statistical Computing**. R Foundation for Statistical Computing, Vienna, Austria. 2013.
- RAMOS, P.L.; MOALA, F.A. A aplicação da distribuição exponencial geométrica estendida para modelagem de dados pluviométricos. **Rev. bras. meteorol.**, v. 29, n. 4, p. 613-620, 2014.
- SARKAR, S.; GOEL, N.K.; MATHUR, B. Adequacy of Nakagami-m distribution function to derive GIUH. **J. Hydraul. Eng.**, v. 14, n. 10, p. 1070-1079, 2009.
- SARKAR, S.; GOEL, N.K.; MATHUR, B. Performance investigation of Nakagami-m distribution to derive flood hydrograph by genetic algorithm optimization approach. **J. Hydraul. Eng.**, v. 15, n. 8, p. 658-666, 2010.
- SARKAR, S.; RAI, R.K. Flood inundation modeling using Nakagami-m distribution based GIUH for a partially gauged catchment. **Water Resour. Manag.**, v. 25, n. 14, p. 3805-3835, 2011.
- SCHWARTZ, J.; GODWIN, R.T.; GILES, D.E. Improved maximum-likelihood estimation of the shape parameter in the Nakagami distribution. **J. Stat. Comput. Simul.**, v. 83, n. 3, p. 434-445, 2013.
- SCHWARZ, G. Estimating the dimension of a model. **The Annals of Statistics**, v. 6, n. 2, p. 461-464, 1978.
- SINGH, K.J.; SARKAR, S. Development of GIUH for the catchment contributing to Loktak lake, North East India. **J. Indian Society Remote Sens.**, v. 41, n.2, p. 447-459, 2013.
- STEPHENS, M.A. EDF statistics for goodness of fit and some comparisons. **J. Am. Stat. Assoc.**, v. 69, n. 347, p. 730-737, 1974.
- TORRES, F.E.; CARGNELUTTI FILHO, A.; TEODORO, P.E.; CORRÊA, C.C.G.; RIBEIRO, L.P.; CUNHA, E.R. Dimensionamento amostral para a estimação da média de precipitação pluvial mensal em locais do Estado do Mato Grosso do Sul. **Cienc. Rural**, v. 46, n. 1, p. 60-69, 2016.

Endereços de Internet

Portal do Instituto Nacional de Meteorologia (INMET), Disponível em: <http://www.inmet.gov.br>.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.