

ANÁLISE DE HOMOGENEIDADE DA PRECIPITAÇÃO NA BACIA HIDROGRÁFICA DO RIO DOS SINOS - RS.

Marco Alésio Figueiredo Pereira^{1}; Katiucia Nascimento Adam²; Bruno Lippo Barbieiro³; Daniela Muller de Quevedo⁴; Daniela M. M. Osório⁵; Sabrina A. Vieira⁶*

Resumo – O objetivo deste trabalho é determinar o ajustamento da distribuição de probabilidade Gama e comparar a possível homogeneidade da precipitação, na bacia no rio dos Sinos - RS, em períodos distintos e com um número diferente de pontos monitorados. Foram utilizadas séries históricas de precipitações diárias de 14 estações, divididas em dois períodos, primeiro período de 01/1965 a 12/1978 com 11 estações e no segundo período de 01/1994 a 12/2012, com três estações. Para inferir o ajustamento da função as séries foi utilizado o teste de Shapiro-Wilk, posteriormente se realizou o agrupamento das estações pela distância Euclideana através do método de Ward, sendo propostos dois e três grupos distintos para as duas séries. Pela análise da ANOVA verificou-se que os possíveis agrupamentos das variáveis não se aplicavam, demonstrando a homogeneidade pluviométrica para a bacia em estudo. Foi possível concluir que a distribuição Gama pode ser utilizada para representar matematicamente a homogeneidade pluviométrica na bacia hidrográfica do rio dos Sinos.

Palavras-Chave – Monitoramento, Precipitação, Homogeneidade.

ANALYSIS OF RAINFALL HOMOGENEITY IN SINOS RIVER BASIN - RS.

Abstract – The objective of this work is to determine the adjustment of the Gamma probability distribution and compare the possible homogeneity of the precipitation at Sinos river basin, northeast region of the Rio Grande do Sul State, in different periods and with a different number of monitored stations. Daily historical precipitations series of fourteen stations were used, divided in two periods: the first period from 01/1965 to 12/1978 with eleven stations and the second period from 01/1994 to 12/2012, with three stations. In order of the adjustment function was used the Shapiro-Wilk's test. Posteriorly was realized the grouping of the stations by the Euclidean distance per the Ward's method. After grouping the stations through the Euclidean distance by the Ward method, two and three distinct groups for the series were established. However, the ANOVA test showed that the possible groupings of the variables were not able to be grouped, which indicates the rainfall homogeneity over the Sinos river basin. It is also concluded that the Gamma distribution can be used to define the rainfall regionalization in this basin.

Keys words: Rainfall, Gamma distribution, Homogeneidade.

¹ *Professor do Programa de Pós Graduação em Qualidade Ambiental da Universidade FEEVALE, marco@feevale.br

^{2,4 e 5} Professora do Programa de Pós Graduação da Universidade FEEVALE.

³ Bolsista de Iniciação Científica da Universidade FEEVALE

⁶ Doutoranda do Programa de Pós Graduação da Universidade FEEVALE.

1. INTRODUÇÃO

O conhecimento da oferta quantitativa e qualitativa dos recursos hídricos, ou seja, o balanço hídrico é de suma importância para o planejamento e gerenciamento dos recursos hídricos em uma bacia hidrográfica. Pois somente com o conhecimento do aporte de entradas e saídas de água na bacia hidrográfica, pode-se definir diretrizes para o uso prioritário e sustentável do meio. A WMO (2008) afirma que o objetivo final da coleta de dados em Hidrologia, seja ele medições de precipitação, registros de nível de água, medições de descarga, monitoramento de águas subterrâneas e amostragem de qualidade de água, é fornecer um conjunto de dados com qualidade suficiente para que possam ser usados na tomada de decisão em todos os aspectos da gestão dos recursos hídricos, na vasta gama de aplicações operacionais, bem como na investigação.

O monitoramento ambiental é dispendioso tanto economicamente quanto tecnicamente, sendo, portanto, a definição de áreas hidrológicamente homogêneas uma forma de reduzir esses custos. Yadav *et al.* (2007) comentam que hidrólogos têm respondido a este desafio, desenvolvendo várias ferramentas de previsão, os quais são comumente referidos como métodos de regionalização. Alguns métodos são comumente usados para a regionalização de variáveis hidro-climáticas como a precipitação, vazão, vazões máximas e mínimas, evapotranspiração e outros componentes do ciclo hidrológico. Romero *et al.* (1999) e Ramos (2001) afirmam que técnicas multivariadas como as técnicas de agrupamento são muito comuns para classificação destas variáveis. Dentro da análise de agrupamento encontram-se na literatura vários métodos que podem ser utilizados em diversos conjuntos de dados, bem como para a regionalização hidrológica. Por exemplo, Lecce (2000), Corduas (2011), e Ramachandra e Srinivas (2006) utilizaram, respectivamente, o método de Ward, o método K-means, o modelo ARMA, e a inferência *Fuzzy* na elaboração de seus trabalhos sobre agrupamentos em regionalização hidrológica.

Em virtude da escassez de séries históricas consideráveis nos diversos campos da ciência, incluindo a Hidrologia, o uso de distribuições teóricas de probabilidades é frequentemente empregado para prever o comportamento destas variáveis. Nesse sentido Ribeiro *et al.* (2007) compararam distribuições de probabilidade para estimar precipitações na região de Barbacena-MG e concluíram que a distribuição Gama foi a que melhor se ajustou aos dados observados. Sampaio *et al.* (2007) obtiveram ajuste adequado desta distribuição para modelar a quantidade de chuva no Estado do Paraná. Oliveira *et al.* (2010) avaliaram a precipitação provável para Alegre - ES entre o período de 1940 a 2007 através da distribuição Gama, comprovando que a mesma se adequa para representar os dados de precipitação.

Nesse sentido, o objetivo deste trabalho é avaliar se a função teórica de distribuição Gama se adequa para prever as duas séries históricas observadas de precipitação diária. E verificar a homogeneidade pluviométrica da precipitação quando avaliadas com 11 estações, no período de 01/1965 a 12/1978 (série 1) e a outro no período de 01/1994 a 12/2012 (série 2), com três estações localizadas dentro e no entorno da bacia hidrográfica do rio dos Sinos, região nordeste do Estado do Rio Grande do Sul.

2. MATERIAL E MÉTODOS

2.1. Área de estudo

A bacia hidrográfica do rio dos Sinos (Figura 1) está situada no nordeste do Estado do Rio Grande do Sul, entre as coordenadas geográficas de 29°20' a 30°10' de latitude sul e 51°20' a 50°15' de longitude oeste. O rio dos Sinos, principal curso de água da bacia, nasce no município de Caraá, no litoral norte do Estado, em altitudes superiores a 600 m e possui aproximadamente 190 km de extensão. A área de drenagem total da bacia é de 3.747 km², envolvendo municípios como Campo Bom, Canoas, Gramado, Igrejinha, Novo Hamburgo, São Leopoldo, Sapucaia do Sul, Caraá, Taquara e Três Coroas. A população da bacia é estimada em 1.249.100 habitantes com 90,6% concentrada em área urbana e 9,4% na área rural (Rio Grande do Sul, 2010).

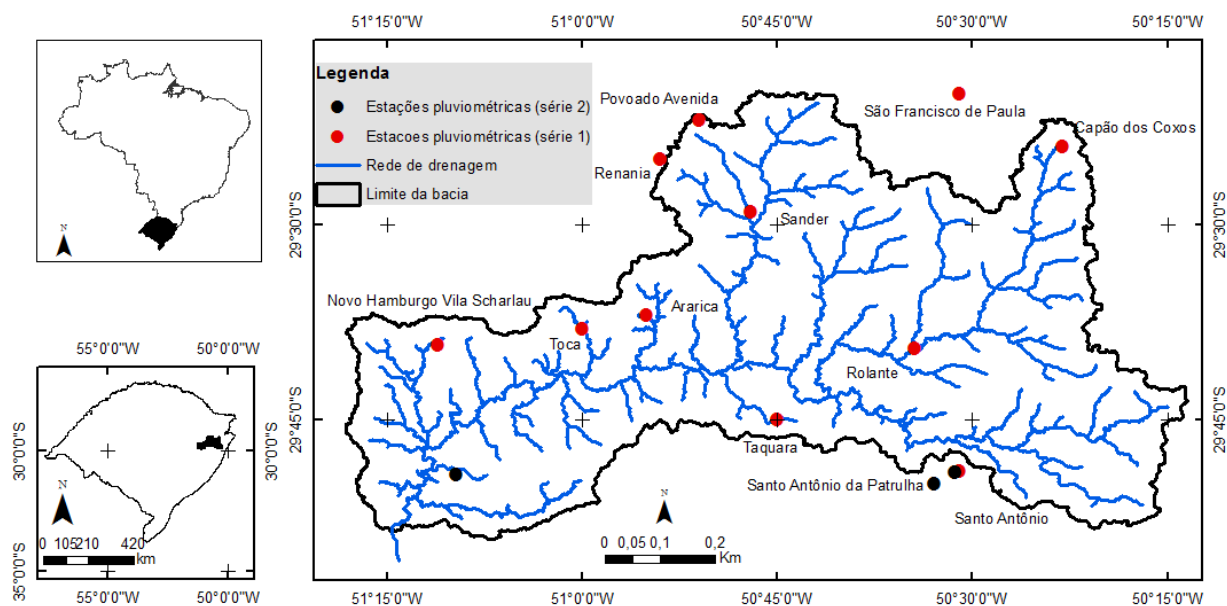


Figura 1 – Localização da bacia do rio Ijuí com a localização dos postos pluviométricos.

2.2. Dados utilizados

Os dados de precipitação diária utilizados neste trabalho foram obtidos do site da ANA (Agência Nacional de Águas) no Sistema de Informações Hidrológicas (HidroWeb). Foi utilizado o *software* Hidro 1.2 como ferramenta de suporte para visualização e seleção dos dados. A Tabela 1 apresenta o código de cada estação, o nome, município, coordenadas geográficas, altitude e sua série histórica, respectivamente. O conjunto de dados utilizados foi determinado em virtude da disponibilidade destes no site da ANA, conforme demonstra a Figura 2, sendo selecionadas 11 estações pluviométricas (série 1) com série histórica de 13 anos distribuídos nos intervalos de 01/1965 a 12/1978 e para (série 2) três estações distribuídos no intervalo de 01/1994 a 12/2012, totalizando 18 anos. Foram utilizados os respectivos períodos para as duas séries históricas em virtude do tempo de coincidência entre os dados observados de cada estação, tendo como intuito principal verificar se com um número menor de pontos monitorados a homogeneidade se aplicavam para as duas séries.

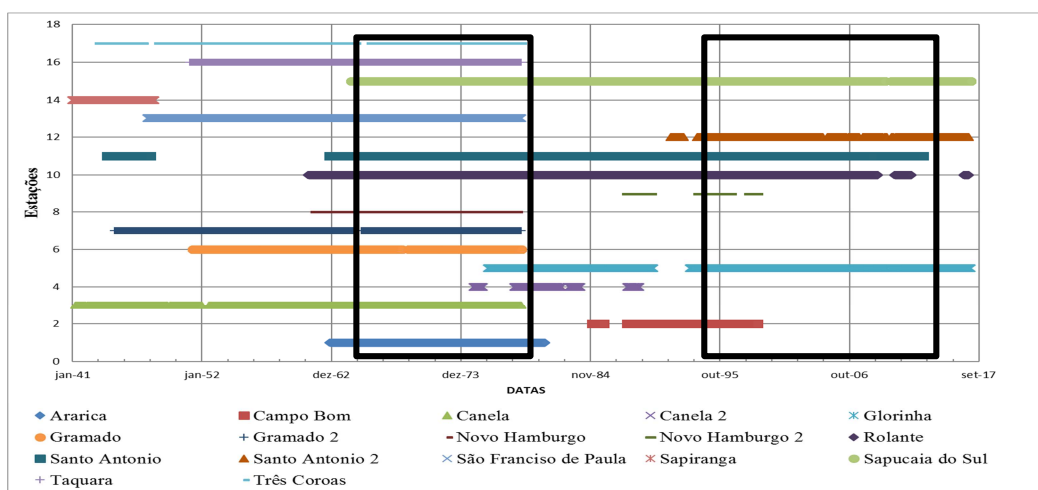


Figura 2 – Série histórica de todas as estações com dados disponíveis para download no site hidroweb. Acesso em 01/05/2017.

Tabela 1 – Dados básicos das estações pluviométricas.

Código	Nome	Município	Latitude (S)	Longitude (W)	Altitude (m)	Série histórica
2950001	Araricá	Sapiranga	29°37'01"	50°55'01"	50	13 anos
2950010	Capão dos Coxos	São Francisco de Paula	29°24'00"	50°22'59"	920	13 anos
2950025	Povoado Avenida	Gramado	29°22'01"	50°51'00"	830	13 anos
2950026	Renania	Gramado	29°25'01"	50°54'00"	660	13 anos
2950028	Rolante	Rolante	29°39'33"	50°34'22"	50	13 anos
2950029	Sander	Três Coroas	29°28'59"	50°46'59"	150	13 anos
2950031	Santo Antônio	Santo Antônio da Patrulha	29°49'01"	50°31'01"	80	13 anos
2950032	Canela	Canela	29°22'00"	50°48'00"	830	13 anos
2950036	Taquara	Taquara	29°45'00"	50°45'00"	35	13 anos
2951036	Vila Scharlau	Novo Hamburgo	29°39'14"	51°11'05"	50	13 anos
2951040	Sapucaia do Sul	Sapucaia do Sul	29°49'12"	51°09'40"	20	13 anos
2950031	Santo Antônio	Santo Antônio da Patrulha	29°49'01"	50°31'17"	80	18 anos
2950060	Santo Antônio da Patrulha	Santo Antônio da Patrulha	29°49'54"	50°32'55"	80	18 anos
2951028	Sapucaia do Sul	Sapucaia do Sul	29°49'12"	51°09'40"	20	18 anos

2.3. Análise estatística

A função de distribuição de probabilidade Gama é expressa como:

$$f_x(x) = \frac{(x/\theta)^{\eta-1} \exp(-x/\theta)}{\theta \Gamma(\eta)} \text{ para } X, \theta, \eta > 0 \quad (1)$$

$$E[X] = \eta\theta ; \quad (2)$$

$$\text{Var}[X] = \eta\theta^2 \quad (3)$$

onde θ e η são os parâmetros de escala e forma, respectivamente; $\mathcal{A}(\eta)$ denota o fator de normalização que obriga a área total da densidade ser igual a 1; $E[X] = \eta\theta$ a média; e $Var[X] = \eta\theta^2$ a variância.

Para averiguar o grau de ajuste dos dados observados com a função teórica de probabilidade, foi empregado o teste proposto por Shapiro e Wilk (1965), (equação 4).

$$SW = \left(\sum_{i=1}^n a_i \cdot y_i \right)^2 / \sum_{i=1}^n (y_i - \bar{y})^2 \quad (4)$$

onde SW é a evidência de quão se adéqua a distribuição teórica a distribuição observada; a_i é um coeficiente linear; y_i é o valor da variável; e \bar{y} é a média dos valores das variáveis. Quanto mais próximo a zero forem os valores de SW mais adequada será a distribuição teórica à distribuição observada.

Após o teste de Shapiro e Wilk realizou-se o agrupamento das variáveis através do método proposto por Scott & Symons (1971) (*Cluster analyses*) que consiste no agrupamento de um conjunto de observações multivariadas, divididas em um número pequeno de grupos relativamente homogêneos. Para determinação dos grupos, utilizou-se a metodologia da distância Euclideana (equação 5) pelo método de Ward.

$$DE = \sum_{j=1}^p (x_{ij} - x_{ij'})^2 \quad (5)$$

onde DE é a distância euclideana; X_{ij} é a j -ésima característica do i -ésimo indivíduo; $X_{ij'}$ é a j -ésima característica do i -ésimo indivíduo. Quanto mais próximo de zero for DE , mais similares são os objetos comparados.

O método de Ward Júnior (1963), equação 6, emprega a análise de variância para determinar as distâncias entre grupos, através da soma dos quadrados entre os dois agrupamentos feita sobre todas as variáveis (Naghettini e Pinto, 2007). Nesse método a formação dos grupos se dá pela maximização da homogeneidade dentro dos grupos.

$$W = \sum_{i=1}^n X_i^2 - \frac{1}{n} (\sum X_i)^2 \quad (6)$$

onde W é a soma de quadrado dos desvios; n é o número de valores analisados; e X_i é o i -ésimo elemento do agrupamento.

Para se determinar a sustentabilidade dos grupos propostos, determinou-se a ANOVA (análise de variância). Este procedimento tem por objetivo determinar se todos os grupos têm valores médios idênticos, comparando duas estimativas de variância total. Se a hipótese nula (H_0) é verdadeira a média dos grupos será muito similar à variância dentro de um grupo em torno da média desse grupo. Se a média dos grupos é diferente, alguns deles serão suficientemente diferentes da média geral, sendo a variância dentro dos grupos não mais igual à variância total. A variância dentro dos grupos é calculada pela estimativa do erro quadrado médio Hirsch *et al.* (1992).

$$MSE = \sum_{j=1}^k \sum_{i=1}^{n_j} \left(y_{ij} - \bar{y}_j \right)^2 / (N - K) \quad (7)$$

onde MSE é o erro quadrado médio; y_{ij} é o valor observado no grupo; \bar{y}_j é a média do j -ésimo grupo; N é o número de observações em cada grupo; e K é o número de grupos.

As análises estatísticas multivariadas deste trabalho foram realizadas pelo aplicativo Statistica® 7.0, desenvolvido pela “StatSoft”. O programa auxilia o usuário na identificação e determinação de parâmetros estatísticos.

3. RESULTADOS E DISCUSSÕES

3.1. Teste de aderência à 5% de significância

Para determinar a distribuição que melhor se adéqua aos dados da série histórica em análise foi realizado o teste de Shapiro-Wilk. A Tabela 2 apresenta os resultados (p -value) do referente teste de aderência para os dados teóricos e observados. Salienta-se que neste teste quanto mais próximo a zero for à inferência, maior será a hipótese que a distribuição teórica se adéqua a distribuição observada. A distribuição Gama demonstra bom desempenho na inferência de ajuste entre a série teórica e a observada, sendo a hipótese nula aceita para a referida distribuição.

Tabela 2 – Teste de aderência pelo método de Shapiro – Wilk.

Estação	Ararica	Canela	Povoado Avenida	Renania	Vila Scharlau	Rolante	Santo Antônio	Capão dos Coxos	Toca	Taquara	Sander	Santo Antônio	Santo Antonio da Patrulha	Sapucaia do Sul
Observada	0.200	0.198	0.166	0.218	0.207	0.191	0.194	0.165	0.177	0.262	0.262	0.147	0.192	0.241
Teórica	0.052	0.058	0.044	0.051	0.06	0.06	0.048	0.051	0.059	0.057	0.073	0.042	0.061	0.061

3.2. Agrupamento

Para representar o agrupamento proposto pelo método *cluster analysis* das variáveis, agora representada pela distribuição Gama, utilizou-se o dendograma (Figura 3a,b), com dois e três grupos. Na Figura 3a observa-se que quando o dendograma é dividido em dois grupos a estação de Santo Antônio compõe um grupo único. Já quando se faz a divisão em três grupos, forma-se um novo grupo da estação Toca. Na figura 3b quando separadas em dois grupos as estações de Sapucaia e Santo Antônio da Patrulha formam um único grupo.

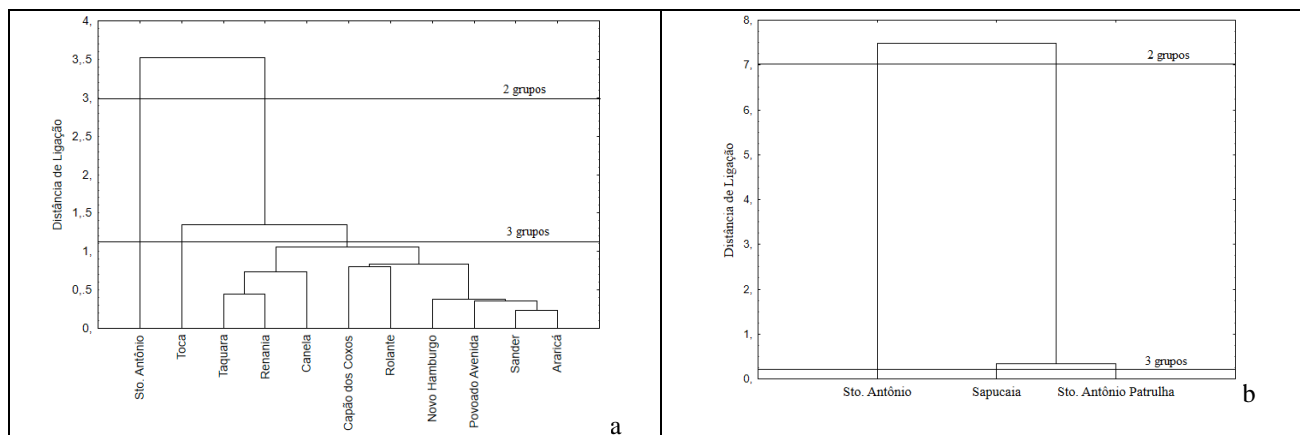


Figura 3 -. (a) Dendograma da série histórica da série 1; (b) Dendograma da série histórica da série 2.

Para verificar a consistência do agrupamento proposto pelo método *cluster analysis* foi realizada o teste da ANOVA. Esta análise tem como objetivo comprovar as diferenças entre os grupos através da homogeneidade das variâncias das médias dentro de cada grupo (Figura 4). Analisando a referida figura observa-se que em todas as situações não houve diferenças significativas entre as médias das séries, conseqüentemente, evidenciando a homogeneidade entre as séries históricas, ou seja, a aplicação da ANOVA nas séries em períodos distintos comprovou que a precipitação que ocorre na bacia não é discrepante, podendo, portanto, um número reduzido de estações representar a precipitação que ocorre em toda a bacia do rio dos Sinos.

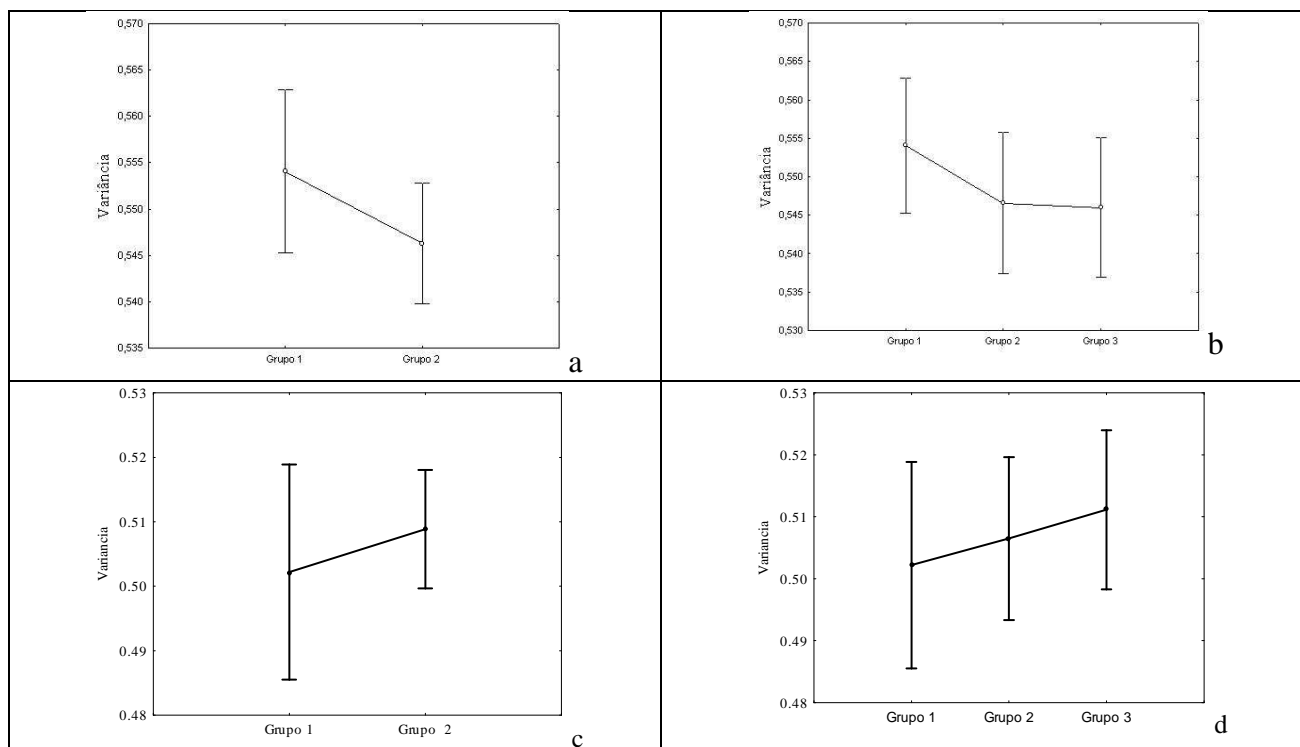


Figura 4 - Teste de homogeneidade pelos gráficos da variância da média para distribuição Gama.

4. CONCLUSÕES

As análises estatísticas utilizadas no presente trabalho foram úteis para compreensão da variabilidade espacial da precipitação diária de 14 estações pluviométricas, localizadas dentro e no entorno da bacia do rio dos Sinos, no primeiro período de 01/1965 a 12/1978 e no segundo período de 01/1994 a 12/2012. Os resultados demonstram que ocorre homogeneidade espacial na distribuição pluviométrica das duas séries históricas analisadas. Em relação aos testes propostos para verificar o ajustamento da distribuição Gama, constata-se que a mesma se ajusta adequadamente para representar dados diários de precipitação.

Quando se realizou o agrupamento da variância pela distância Euclidiana pelo método (W) foi definido que as variáveis se enquadravam em dois e três grupos. Porém, Com a aplicação da ANOVA, verificou-se que os possíveis agrupamentos das variáveis não se aplicavam, demonstrando, portanto, que as precipitações que ocorrem na bacia apresentam uma homogeneidade pluviométrica. Avaliando o desempenho da função Gama se conclui que esta distribuição pode ser utilizada para definir a regionalização hidrológica da bacia hidrográfica dos Sinos.

REFERÊNCIAS

- CORDUAS, C. (2011). Clustering streamflow time series for regional classification. *Journal of Hydrology* 407, pp.73–80.
- HIRSCH, T. A.; HESEL, D. R.; COHN, T. A.; GILROY, E. J. (1992). Statistical treatment of data of hydrologic data.(cap. 17) In. *Handbook of hydrology*. U.S. Geological Survey. Ed. Donnelley&Sons Company. New York, pp. 645-696.
- LECCE, S. A. (2000). Spatial variations in the timing of annual floods in the southeastern United States. *Journal of Hydrology*, 235, pp. 151–169.
- NAGHETTINI, M. PINTO, É. J. de A. (2007) *Hidrologia Estatística. Belo Horizonte*, CPRM, 552p.
- OLIVEIRA, J. P. B. de; CECÍLIO, R. A.; XAVIER, A. C.; JASPER, A. P. dos S.; OLIVEIRA, L. B. (2010). Precipitação provável para Alegre – ES através da distribuição de probabilidade Gama. *Engenharia Ambiental - Espírito Santo do Pinhal*, v. 7, n. 2, pp.204-211.
- RAMACHANDRA RAO, A.; SRINIVAS, V.V. (2006). Regionalization of watersheds by fuzzy cluster analysis. *Journal of Hydrology* v. 318, pp. 57–79.
- RAMOS, M.C., 2001. Divisive and hierarchical clustering techniques to analyze variability of rainfall distribution patterns in a Mediterranean region. *Journal of Hydrology* v. 57, 123–138.
- RIBEIRO, B. T.; AVANZI, J. C.; MELLO, C. R. de ; LIMA, J. M. de ; SILVA, M. L. N.; (2007). Comparação de distribuições de probabilidade e estimativa da precipitação provável para região de Barbacena, MG. *Ciênc. agrotec.*, Lavras, v.31, n.5, pp.1297-1302.
- RIO GRANDE DO SUL. Secretaria Estadual do Meio Ambiente. (2010) Portal do meio ambiente RS. Bacia Hidrográfica do rio dos Sinos. Disponível em: Acesso em: maio 2017.
- ROMERO, R., SUMMER, G., RAMIS, C., GENOVES, A., (1999). A classification of the atmospheric circulation patterns producing significant daily rainfall in the Spanish Mediterranean area. *Int. J. Climatol.* 19, pp. 765–785.
- SAMPAIO, S. C.; QUEIROZ, M. M. F. de; FRIGO, E. P.; LONGO, A. J.; SUSZEK, M. (2007). Estimativa e distribuição de precipitações decendiais para o Estado do Paraná. *Irriga*, Botucatu, v. 12, n. 1, pp. 38-53.
- SCOTT, A. J.; SYMONS, M. J. (1971). On the Edwards and Cavalli-Sforza Method of Cluster Analysis. *Biometrics*. Vol27. Nº1, University of North Carolina, Chapel Hill, U.S.A., pp.217-219.
- SHAPIRO, S.S.; WILK, M. B. (1965). An Analysis of Variance Test for Normality (Complete Samples). *Biometrika*, Vol. 52, pp. 591-611.
- TUKEY, J. W. (1949). Comparing Individual Means in the Analysis of Variance, *Biometrics*, 5, 99 p.
- WARD JÚNIOR, J. H. (1963). Hierarchical grouping to optimize an objective function. *Journal of the American Statistical Association*, v.58, n.301, pp.236-244.
- WMO – WORLD METEOROLOGICAL ORGANIZATION. (2008). *Guide to hydrological Practices, v.I, Hydrology – From Measurement to Hydrological Information*, 296p.
- YADAV, M.; WAGNER, T.; GUPTA, H.; (2007). Regionalization of constraints on expected watershed response behavior for improved predictions in ungauged basins. *Advances in Water Resources*, v. 30, p. 1756-1774.