

IDENTIFICAÇÃO DE REGIÕES HOMOGÊNEAS E ANÁLISE DE REGRESSÃO MÚLTIPLA PARA REGIONALIZAÇÃO DE VAZÃO NA BACIA HIDROGRÁFICA DO RIO TAPAJÓS

Identification of Homogeneous Regions and Analysis of Multiple Regression to Flow Regionalization in Hydrographic Basin of Tapajós River

**Edkeyse Dias Gonçalves¹, Francisco Carlos Lira Pessoa²,
Raisa Rodrigues Neves¹, Rodrigo Silvano Silva Rodrigues¹
& Ana Carolina Santa Rosa de Sousa³**

¹Universidade Federal do Pará – UFPA

Programa de Pós-Graduação em Engenharia Civil

Rua Augusto Corrêa, 01 – Guamá / CEP: 66075-110 – Belém – PA – Brasil
edkeyse.jesus@hotmail.com, raisanevesufpa@gmail.com, rcsr@ufpa.br

²Universidade Federal do Pará – UFPA

Instituto de Tecnologia – Faculdade de Engenharia Sanitária e Ambiental

Rua Augusto Corrêa, 01 – Guamá / CEP: 66075-110 – Belém – PA – Brasil
fclpessoa@ufpa.br

³Universidade Federal Rural da Amazônia – UFRA

Programa de Pós-Graduação em Aquicultura e Recursos Aquáticos Tropicais

Av. Presidente Tancredo Neves, 2501 – Terra Firme / CEP: 66.077-830 – Belém - PA – Brasil
anacarolina_srsousa@outlook.com

Recebido em 12 de Setembro, 2016/ Aceito em 27 de Agosto, 2017

Received on September 12, 2016/ Accepted on August 27, 2017

RESUMO

Na gestão dos recursos hídricos é de fundamental importância conhecer o comportamento dos rios para fins de planejamento de atividades humanas, bem como para identificação de períodos de cheias e vazantes. A regionalização é importante para se obter modelos de cálculos de vazões e explorar ao máximo as séries hidrológicas por meio de métodos estatísticos. Dessa forma, o presente artigo teve como objetivo obter regiões hidrologicamente homogêneas e para cada uma das regiões propor um modelo de regionalização de vazões para a bacia hidrográfica do Rio Tapajós, uma vez que a mesma vem sendo alvo de grandes projetos hidrelétricos. Utilizou-se informações hidrológicas de 27 estações fluviométricas e pluviométrica, assim como o método de agrupamento hierárquico de Ward e a distância euclidiana quadrática como medida de similaridade resultando na formação de 3 regiões homogêneas. O método de regressão múltipla foi empregado no desenvolvimento do modelo de regionalização de vazões, no qual a vazão Q90 foi relacionada em função das características físicas (área de drenagem e comprimento do rio) e climática (precipitação mensal média). O modelo demonstrou eficiência na estimada das vazões, já que alcançou índices de determinação (R^2) e R^2 ajustado superiores a 0,75, e o coeficiente de NASH próximo de 1, provando que pode ser utilizado como ferramenta de auxílio na gestão dos recursos hídricos, bem como mensurar aspectos quantitativos de vazões à cursos de rios sem monitoramento por estações fluviométricas.

Palavras-chave: Análise de Agrupamento, Regressão Múltipla, Regionalização de Vazões.

ABSTRACT

In the management of water resources it is of fundamental importance to know the behavior of rivers for the purpose of planning human activities, as well as to identify periods of floods and droughts. Regionalization is important in order to obtain models of flow calculations and to exploit hydrological series as much as possible through statistical methods. Thus, the present article aimed to obtain hydrologically homogeneous regions and for each of the regions to propose a model of regionalization of flow to the Tapajós River basin, since it has been the subject of large hydropower projects. Hydrological information of 27 fluviometric and pluviometric stations was used, as well as Ward's hierarchical grouping method and the quadratic euclidean distance as a measure of similarity resulting in the formation of 3 homogeneous regions. The multiple regression method was used to develop the flow regionalization model, in which the Q90 flow was related to the physical characteristics (drainage area and river length) and climatic (mean monthly precipitation). The model showed efficiency in the estimation of the flow rates, since it reached indexes of determination (R²) and adjusted R² superiors to 0,75, and a mean of quadratic errors relative average percentage ($\epsilon\%$) near 3%, proving that can be used as a tool to aid in the management of water resources, as well as to measure quantitative aspects of flows to river courses without monitoring by fluviometric stations.

Keywords: Cluster Analysis, Multiple Regression, Flow Regionalization.

1. INTRODUÇÃO

Os países, bem como os estados e municípios, estão enfrentando uma árdua tarefa na gestão de recursos hídricos: lidar com problemas de escassez hídrica e conflitos pelo uso da água, além de enfrentar eventos naturais extremos, como inundações e secas.

Nas áreas metropolitanas do Brasil, a gestão dos recursos hídricos se concentra no saneamento para manter e aumentar o suprimento de água, especialmente para usos urbanos e industriais, e também para a diluição de águas residuais. Os impactos humanos muitas vezes resultam na degradação das fontes de água, comprometendo suas funções ambientais e aumentando os desafios da oferta e acesso à água, bem como a usos gerados e conflitantes da água (BROWN *et al.*, 2008).

Wu *et al.* (2015) afirmam que o aumento da demanda de água urbana surge com o crescimento da população e das atividades econômicas nas cidades, além disso, a urbanização altera a hidrologia. Desta forma, o monitoramento de dados de vazões e precipitações, obtidos em estações fluviométricas e pluviométricas, é de suma importância para o conhecimento da variabilidade hidrológica e adequada gestão dos recursos hídricos.

No entanto, a carência ou a ausência de informações hidrológicas mostra-se como uma das principais problemáticas, uma vez que o dimensionamento de projetos que façam uso da água necessita de longas séries históricas. Pessoa *et al.* (2011) comentam em seus estudos que há

uma grande dificuldade em monitorar vazões e precipitações em uma rede hidrográfica densa e extensa como a do Brasil, e falta investimento na implantação dessas estações hidrométricas.

Rodrigues *et al.* (2016) contabilizaram 1.774 estações fluviométricas monitoradas pela Agência Nacional de Água – ANA na Amazônia Legal. A Organização Meteorológica Mundial (OMM) recomenda para a região uma densidade mínima de 1.875 km² por estação. Considerando a área da Amazônia Legal, a região possui uma média de 2.830 km² por estação, ou seja, indica que a densidade de estações instaladas é inadequada, com o agravante de unidades desativadas e muitos erros e descontinuidades, dificultando estudos com longas séries históricas.

Ishihara *et al.* (2013) alertam que a Amazônia não tem uma rede de monitoramento adequada, pois a cobertura de estações pluviométricas ainda é baixa, com uma média de 3.343,91 km² por estação, enquanto que a densidade mínima recomendada pela OMM para a região é de 575 km² por estação, enfatizando que as estações não estão distribuídas homoganeamente na região, principalmente quando se leva em consideração a sua temporalidade, o que se torna um complicador para o monitoramento de uma das regiões tão importante para o Brasil como para o mundo.

Com a finalidade de suprir a insuficiência de informações hidrológicas, a regionalização de vazões vem sendo utilizada por muitos pesquisadores (MAMUN *et al.*, 2010; DAWDY *et al.*, 2012; MENDICINO & SENATORE,

2013), e tem como principal objetivo transferir informações para outros locais, buscando explorar ao máximo os dados disponíveis numa determinada área geográfica considerada hidrologicamente homogênea. O termo regiões homogêneas está associado à regiões que possuem similaridade hidrológica (TUCCI, 2009).

A análise de regressão é uma técnica bastante utilizada na construção dos modelos hidrológicos de regionalização e relaciona a variável dependente (vazão) com as variáveis independentes - fisiográficas e/ou climáticas (BARBOSA *et al.*, 2005).

Mendicino e Senatore (2013) analisaram o desempenho de sete modelos de regionalização de curvas de permanência (dois estatísticos e cinco paramétricos), para 19 bacias calibradas, em uma região do sul da Itália conhecida como Calábria. Para a definição dos modelos regionais, utilizaram análise de regressão múltipla. Os modelos estatísticos mostraram boa confiabilidade. Já Mamun *et al.* (2010) desenvolveram um estudo de regionalização para a Península da Malásia. Os autores avaliaram vazões com durações de 1, 3 e 30 dias, com períodos de retorno que variavam entre 1 e 50 anos.

Dos métodos de análise de agrupamentos utilizados na Hidrologia para obtenção de regiões homogêneas, destacam-se o método hierárquico de Ward (FARSADNIA *et al.*, 2014; RIANNA *et al.*, 2011; YANG *et al.*, 2010). Gonçalves *et al.* (2016) identificaram seis regiões homogêneas no estado do Pará, utilizando dados de precipitação de 413 postos de monitoramento de chuvas, por intermédio do método de aglomeração hierárquica de Ward.

Tsakiris *et al.* (2011), em um estudo realizado em Massachusetts – EUA, relacionaram variáveis fisiográficas à vazões mínimas (Q7,2 e Q7,10), por meio da regressão múltipla, com a finalidade de regionalizar vazões mínimas. Os autores usaram o método de análise de agrupamento hierárquico de Ward para identificar sub-regiões homogêneas, e como medida de similaridade a distância euclidiana. Concluíram que com base nas duas sub-regiões homogêneas definidas a estimativa

das vazões mínimas foi mais precisa.

A disponibilidade hídrica na região amazônica é influenciada diretamente pela variabilidade climática e atividades humanas (RODRIGUEZ *et al.* 2010; VILLAR *et al.* 2009) como por exemplo a expansão agrícola e expansão hidroelétrica na geração de energia, visto que, a bacia do Tapajós vem sendo alvo de grandes projetos hidrelétricos conhecido por ‘Complexo Hidrelétrico do Tapajós, podendo inicialmente gerar até 10.682 MW (TAPAJÓS, 2014).

Em virtude disso, o presente trabalho teve como objetivo identificar regiões homogêneas de chuvas mensais médias na bacia hidrográfica do rio Tapajós, e para cada região, definir modelos de regionalização que permitam estimar vazões em qualquer ponto das mesmas, permitindo compreender a disponibilidade hídrica e colaborando nas tomadas de decisões na gestão integrada de recursos hídricos da bacia do Tapajós.

2. MATERIAIS E MÉTODOS

A bacia do Tapajós configura uma das principais sub-bacias da Amazônica Legal, faz fronteira com o estado do Pará, Amazonas, Mato Grosso e Rondônia (Figura 1), com uma área drenagem de 764.183 km².

Com extensão territorial de 2.000 quilômetros, o Tapajós é formado a partir do encontro dos rios Juruena e Teles Pires, na divisa dos estados de Mato Grosso, Amazonas e Pará. A partir desse ponto, avança cerca de 800 quilômetros para desaguar na margem direita do rio Amazonas (TAPAJÓS, 2014).

A bacia do Tapajós é sub dividida em 5 sub regiões hidrográficas (Jamanxim, Juruena, Médio Tapajós, Baixo Tapajós e Teles Pires), em que concentra grande potencial hidrológico e durante o ano os regimes de vazões mínimas é da ordem de 2.500 m³/s caracterizado pelo período de estiagem e vazões máximas da ordem de até 28.000 m³/s nas cheias. A estação de maior pluviosidade se dá nos meses de dezembro a maio, tendo o mês de março como período mais chuvoso, já o período de menor pluviosidade está entre junho e novembro com vazões mínimas para o mês de setembro (TAPAJÓS, 2014).

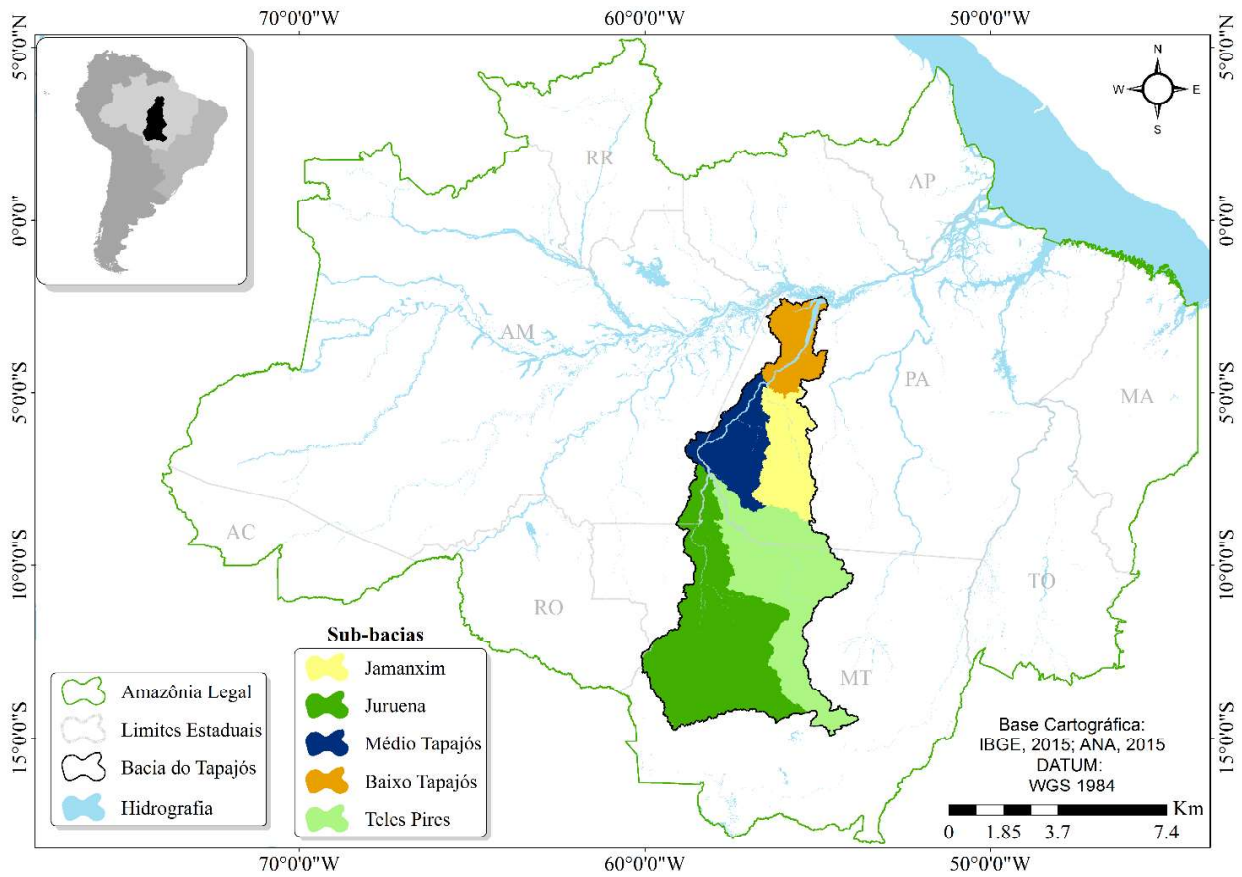


Fig. 1- Localização da bacia hidrográfica do Rio Tapajós.

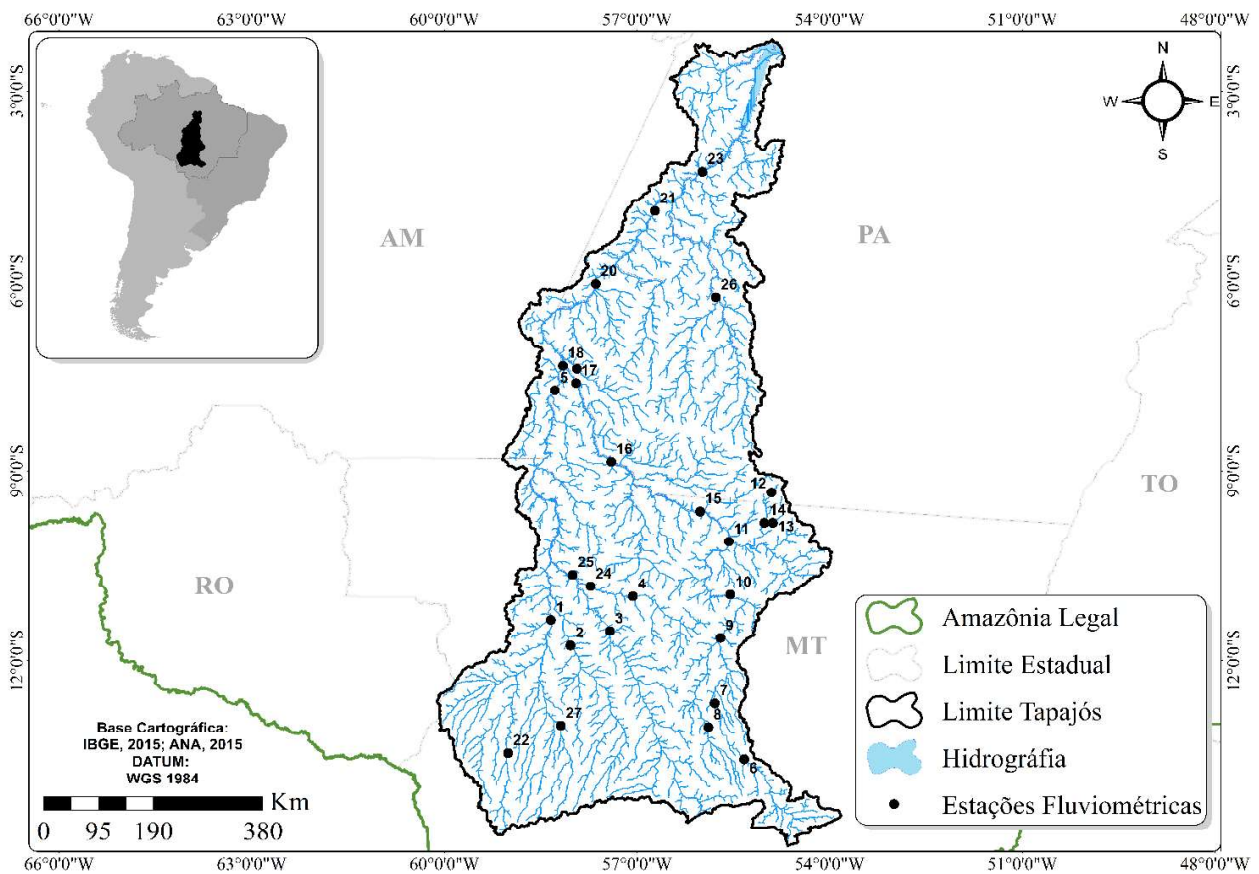


Fig. 2 - Localização da bacia do Tapajós.

2.1 Dados Utilizados

O presente estudo envolveu 27 estações fluviométricas e pluviométricas sendo todas localizadas nos limites da bacia do Tapajós. Os dados foram adquiridos na rede hidrometeorológica do Sistema de Informações Hidrológicas - Hidroweb (2015), da Agência Nacional de Águas – ANA. Na Figura 2 pode

ser verificada a distribuição espacial dos postos fluviométricos utilizados no estudo.

A Tabela 1 apresenta os dados de extensão da série em anos, vazão de referência (Q90) - m³/s, área de drenagem (km²), precipitação mensal média (mm) e o comprimento do curso d'água principal (km), obtidos para cada estação fluviométrica da bacia.

Tabela 1: Dados observados para as estações de chuva e vazão

Nº	Código	Extensão da série de dados (anos)	Nome	UF	Q ₉₀ (m ³ /s)	Área de drenagem (km ²)	Precipitação média mensal (mm)	Comprimento do rio (km)
1	17093000	27	Fontanilhas	MT	1.181,83	55.900	156,02	75,26
2	17095000	16	F a z e n d a Tombador	MT	382,08	24.700	156,9	55,73
3	17120000	27	P o r t o dos Gaúchos	MT	449,22	37.100	164,98	235,75
4	17121000	9	Caiabis (Balsa – Itapaiuna)	MT	104,89	9.750	171,31	184,35
5	17130000	6	F o z dos Gaúchos	MT	2.605,72	182.000	178,47	417,29
6	17200000	29	P o r t o Roncador	MT	117,34	10.800	145,58	258,08
7	17210000	21	Teles Pires	MT	414,31	13.900	140,08	181,91
8	17230000	20	PCH Canoa Quabrada Lucas do Rio Verde	MT	79,24	5.435	154,72	40,14
9	17280000	33	Cachoeirão	MT	456,01	34.600	141,22	16,23
10	17300000	24	F a z e n d a Tratex	MT	532,96	40.700	167,45	126,24
11	17340000	14	INDECO	MT	588,42	52.200	170,54	195,06
12	17345000	40	B a s e do Cachimbo	MT	18,1	465	156,67	79,56
13	17350000	27	Cachimbo	MT	12,25	1.010	191,65	21,21
14	17355000	19	PCH Braço Norte Jusante	MT	52,5	3.100	191,65	104,3
15	17380000	28	Jusante Foz Peixoto de Azevedo	MT	844,44	81.600	181,93	90,98
16	17410000	17	Santa Rosa	MT	1.640,49	131.000	177,6	54,72
17	17420000	20	Três Marias	MT	1.554,05	138.000	207,26	188,64
18	17430000	23	Barra do São Manuel	AM	18.331,45	333.000	199,41	55,52
19	17450000	7	Missão Cururu	PA	57,8	7.345	215,11	45,55

Nº	Código	Extensão da série de dados (anos)	Nome	UF	Q ₉₀ (m ³ /s)	Área de drenagem (km ²)	Precipitação média mensal (mm)	Comprimento do rio (km)
20	17500000	17	Fortaleza	PA	4.514,58	363.000	181,34	226,61
21	17650002	7	Acará do Tapajós	PA	4.038,66	390.000	177,35	189,99
22	17091000	13	Fazenda Tucunaré	MT	133,33	4.360	180,09	246,6
23	17730000	18	Itaituba	PA	3.626,13	458.000	176,73	142,66
24	17122000	9	Rio dos peixes	MT	144	14.000	160,55	113,77
25	17123000	8	Rio Arinos	MT	653,62	57.100	167,62	167,94
26	17675000	9	Jardim do ouro	PA	230,24	37.400	169,74	118,66
27	17092800	8	Aldeia Sacre II	MT	153,98	6.770	139,3	85,46

Fonte: HIDROWEB (2015).

A base cartográfica como os arquivos shapefile foram obtidos através do Sistema Nacional de Informações sobre Recursos Hídricos (SNIRH), para confecção de mapas.

Foram utilizados dados diários de vazões e posteriormente, foram construídas curvas de permanência para cada uma das estações objetivando determinar a vazão de referência Q90. Por intermédio das séries históricas de chuvas diárias, calculou-se as médias mensais de precipitação. A obtenção do comprimento do rio principal foi realizado a vetorização espacial dos rios resultando em um arquivo shapefile de hidrografia em uma escala de 1:10000, com uso de imagens de alta resolução espacial SPOT-5, disponibilizada por meio da ferramenta ArcGis online. O software de processamento espacial utilizado foi o ArcGis 10.1. Para a elaboração da análise estatística de agrupamento Ward fez-se uso do software Statistical Package for Social Sciences - SPSS.

2.2 Análise de agrupamento

A Análise Multivariada é um conjunto de técnicas usadas para resolver problemas relacionados ao agrupamento de itens (JOHNSON & WICHERN, 2007). Uma das formas de obtenção de resultados quando se deseja dividir uma área de estudo em regiões homogêneas é por meio da análise de agrupamento (cluster analysis), método estatístico multivariado, cuja finalidade

primária é agregar objetos com base em uma medida de similaridade e nas características que eles possuem (HAIR *et al.*, 2009).

Segundo Pereira (1999), a cluster analysis pode ser resumida nos seguintes procedimentos: 1.Cálculo das distâncias euclidianas entre os objetos estudados no espaço multiplano de todas as variáveis consideradas. As principais medidas de distâncias são: euclidiana; euclidiana ao quadrado; city block; chebychev; 2.Sequência de agrupamento por proximidade geométrica. As regras para agrupamento mais utilizadas são o single linkage, que é o cálculo a partir do objeto mais próximo, e o complete linkage, que é computado pela maior distância entre dois objetos; 3.Reconhecimento dos passos de agrupamento para identificação coerente de grupos dentro do universo de objetos estudados.

A técnica é dividida em medidas de similaridade e de dissimilaridade. O coeficiente de correlação é uma medida de similaridade, enquanto que a distância euclidiana é uma medida de dissimilaridade.

Para a identificação das regiões homogêneas de chuvas mensais médias, foram utilizadas as características físico-climáticas (precipitação mensal média, área de drenagem e comprimento do rio principal).

Utilizou-se a distância euclidiana quadrática como medida de similaridade do método

hierárquico de agrupamento de Ward. A distância euclidiana é uma medida constantemente utilizada na combinação de objetos, visto que é uma distância geométrica tomada em um espaço de p dimensões. Sendo x_{ij} a observação da i -ésima estação fluviométrica ($i = 1, 2, \dots, n$), com referência a j -ésima variável em cada classe ($j = 1, 2, \dots, p$) estudada, define-se a distância euclidiana padronizada entre dois postos i e j , dada pela Equação 1 (GOMES *et al.*, 2013):

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

em que x_{ij} é a j -ésima característica do i -ésimo indivíduo; e x_{ij} é a j -ésima característica do i -ésimo indivíduo.

O método de Ward (WARD, 1963) consiste em um procedimento de agrupamento hierárquico no qual a medida de similaridade, usada para juntar agrupamentos, é calculado com a soma de quadrados, denominada de SQ, entre os dois agrupamentos e aplicada para todas as variáveis (HAIR *et al.*, 2009). A reunião definida dos objetos irá contemplar os menores valores de SQ, dado pela Equação 2:

$$E_{(G1G2)} = \sum_{v=j=1}^p \sum_{i \in G1}^n (x_{iv} - \bar{x}_v)^2 \quad (2)$$

no qual o valor de E para dois grupos, $G1$ e $G2$; \bar{x}_v é a média do grupo para cada variável v .

O método de Ward apresenta algumas características com tendência, que combinam grupos com poucos elementos, é um método que tende a resultar em agrupamentos de tamanhos aproximadamente iguais devido a sua minimização de variação interna (SEIDEL *et al.*, 2008).

Por ser tratar de uma das técnicas estatísticas mais robustas, diversos autores utilizam a análise de agrupamento (cluster analysis) como base ou associação metodológica para estudo de gestão ambiental (ZHANG *et al.*, 2016; YAN *et al.*, 2016; KUMAR *et al.*, 2016; CARMO *et al.*, 2016; PARENTE *et al.*, 2016; DONG *et al.*, 2016), bem como para estudos hidrológicos (LAVARINI *et al.*, 2014; KULT *et al.*, 2014; MASCARO *et al.*, 2015; NGUYEN *et al.*, 2015; HAWKINS *et al.*, 2015; HUANG *et al.*, 2016; PEÑAS *et al.*, 2016;

AGARWAL *et al.*, 2016; BELL *et al.*, 2016).

2.3 Análise de Regressão Linear Múltipla

A análise de regressão múltipla é uma técnica estatística que pode ser usada para verificar a relação entre uma única variável dependente e várias independentes. O objetivo da análise da regressão múltipla é usar as variáveis independentes cujos valores são conhecidos para prever os valores da variável dependente selecionada pelo pesquisador (HAIR *et al.*, 2009; LANDIM, 2003).

Para se obter o valor de uma variável Y em função de outras variáveis X independentes entre si, usa-se um modelo de Regressão Linear Múltipla, dado pela Equação 3.

$$y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \beta_n X_{in} + \varepsilon_i \quad (3)$$

em que y_i é variável dependente (Q90) a resposta no i -ésimo ensaio, X_{i1}, X_{i2}, X_{in} são os valores das variáveis explicativa ou independentes (área de drenagem, precipitação mensal média e comprimento do rio principal) no i -ésimo ensaio. Os coeficientes de regressão são $\beta_0, \beta_1, \beta_2, \beta_n$, e ε_i é o termo do erro, isto é, a variabilidade não explicada pela relação linear.

De maneira direta a regressão múltipla se vale da utilização destas variáveis independentes adicionais para melhorar a capacidade de predição em confronto com a regressão linear simples, sendo um modelo matemático utilizado em estudos hidrológicos atuais por diversos autores, como Griffis e Stedinger (2007); Matteau *et al.* (2009); Regonda *et al.* (2013); Molowny-Horas *et al.* (2015); Liu e Cheng (2016); Sridharan e Sandhya (2016).

2.4 Critérios de Desempenho do modelo

O Coeficiente de correlação múltipla (R^2) mede a fração da variabilidade nas observações de Y explicada pela regressão dos valores de X (MONTGOMERY *et al.*, 2012). Contudo, mesmo que o R^2 utilize uma variável que não esteja correlacionada com a saída e não melhore o modelo, seu valor tende a aumentar com a adição de variáveis.

Desde modo, o coeficiente mais utilizado para medir quanto o valor estimado se assemelha ao medido, sem a interferência do número de

variáveis, é o R^2 ajustado (ESQUERRE *et al.*, 2004).

O desempenho do modelo foi avaliado por intermédio do coeficiente de determinação R^2 ajustado (Equação 4), por este ter melhor capacidade explicativa do modelo do que o coeficiente de determinação R^2 . Esta medida de ajuste mostra quanto da variável dependente é explicado pela variação conjunta das variáveis independentes consideradas no modelo (FÁVERO *et al.*, 2009).

$$R^2_{ajustado} = 1 - \frac{n-1}{n-p-1} \cdot (1 - R^2) \quad (4)$$

no qual n é o número de valores observados, p é o número de variáveis independentes e R^2 é o coeficiente de determinação.

Outro coeficiente analisado foi Nash-Sutcliffe (Nash) para validar o desempenho do modelo de regionalização. Nash e Sutcliffe (1970) propuseram uma equação que compara a redução do desvio quadrático do erro do modelo com desvio quadrático do modelo alternativo de prever sempre a média dos valores, conforme apresenta-se na Equação 5.

$$Nash = 1 - \frac{\sum (Q_{obs} - Q_{sim})^2}{\sum (Q_{obs} - Q_{med})^2} \quad (5)$$

em que Q_{obs} é a vazão observada (m^3/s), Q_{sim} é a vazão simulada pelo modelo de regionalização (m^3/s), e Q_{med} é a média das vazões observadas.

O coeficiente de Nash pode variar de $-\infty$ a 1, sendo o valor de 1 indicativo de um perfeito ajuste, enquanto valores inferiores a 0,36 fazem com que o modelo seja julgado como inaceitável (COLLISCHONN, 2001; SILVA, P. et al., 2008).

3. RESULTADOS E DISCUSSÃO

Apresenta-se, a seguir, os resultados e discussão da aplicação da técnica de agrupamento hierárquico de Ward e do método de regressão múltipla a partir das características físico-climáticas da bacia hidrográfica do Rio Tapajós. Esses métodos foram escolhidos por serem os mais usados em estudos de identificação de regiões homogêneas, da mesma maneira que, em estudos de regionalização de vazões na Hidrologia.

3.1 Análise de agrupamento (Cluster)

Na Figura 3 está representado o dendrograma, também conhecido como árvore de decisão, com as 27 estações com suas respectivas variáveis explicativas agrupadas: vazão (Q) em $m^3.s^{-1}$, precipitação média mensal (P) em mm, comprimento do rio principal (C) em km e área de drenagem (A) em km^2 todas submetidas à medida da distância euclidiana quadrática pelo método de agrupamento hierárquico de Ward.

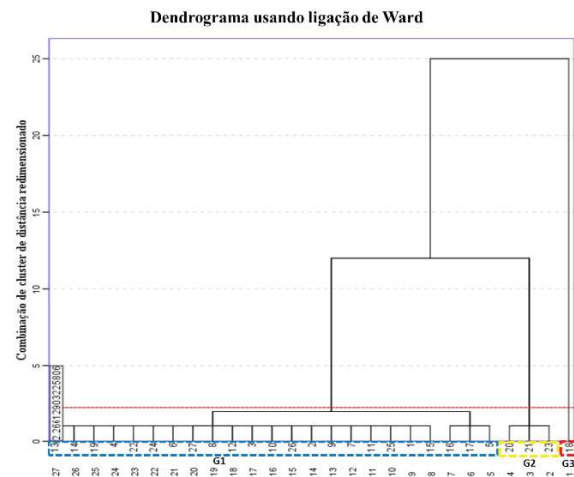


Fig. 3 - Dendrograma aplicado ao Método de Ward.

A análise do dendrograma resultou na formação de 3 grupos homogêneas (G1, G2 e G3), considerou-se um corte na distância euclidiana de aproximadamente 2,262.

Na Figura 4 é apresentado o mapa das 3 regiões homogêneas de precipitações mensais médias definidas no estudo. As três regiões foram bem definidas apresentando comportamentos bem dissimilares quando comparadas.

Na região homogênea 1, observou-se que os dados de 23 estações fluviométricas foram agrupados, representando 85% de todas as informações disponíveis na bacia do Tapajós. Apresenta menores alturas de precipitações mensais médias, com média em torno de 169 mm e vazões (Q90) entre 12,25 m^3/s e 2.605,72 m^3/s . Considerando as sub bacias, a região 1 compreendeu 81,6% da área da sub bacia do Jamanxim, 20,2%, 99,7% e 98,8% respectivamente, das sub bacias Médio Tapajós, Teles Pires e Juruena. As duas últimas sub bacias citadas apresentam as mesmas características climáticas, considerando a classificação por

Köppen-Geiger em Aw, além disto, a vegetação predominante é do tipo savana. Segundo Domingues e Bermann (2012), o sul da bacia do rio Tapajós representa a zona de maior stress hídrico da bacia e coincidentemente é a área definida como Arco do desmatamento.

A intensa expansão agrícola, a exploração madeireira e a expansão de pastagens são ações que provocam desequilíbrio na fauna, flora e nos recursos hídricos evidenciados pelas menores vazões e precipitações, fato este confirmado por (SANTOS *et al.*, 2014).

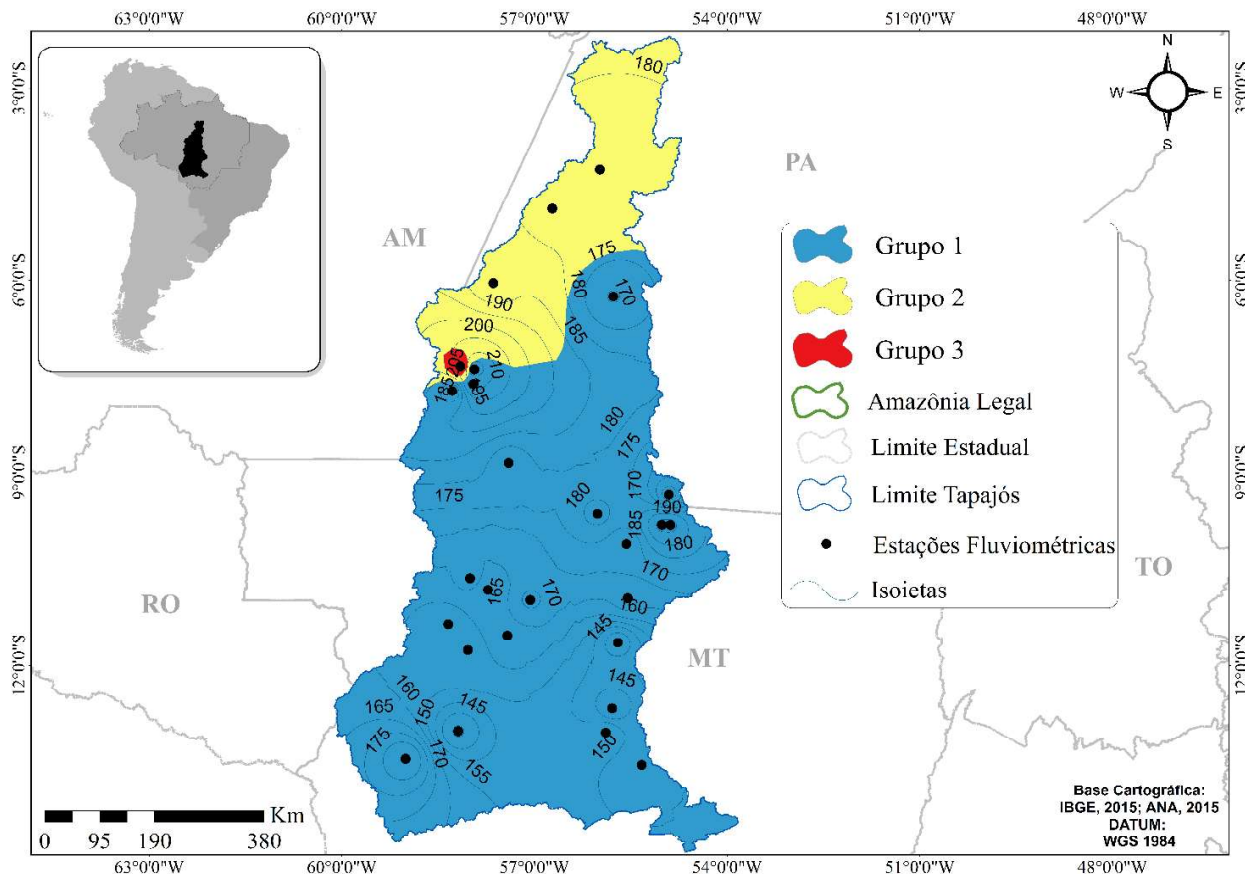


Fig. 4 – Regiões Homogêneas de precipitações mensais médias.

As características climáticas das regiões 2 e 3 podem ser classificadas como Am e clima tropical de monção, com uma breve estação seca e chuvas intensas durante o resto do ano, o mês menos quente tem temperatura média superior a 18°C (SANTOS *et al.*, 2014).

A região 2 representa 20,5% da bacia do Tapajós, as maiores alturas de precipitações mensais médias foram representadas nesta região, com média de 178,5mm, assim como as maiores vazões (Q_{90}) variando entre 3.626 m^3/s a 4.514,58 m^3/s . As grandes vazões encontradas são consequência da localização das estações fluviométricas, uma vez que estão instaladas próximas da foz do rio Tapajós. A região abrange uma porção de 18,3% da

sub bacia Jamanxim, além de 0,68%, 78,4%, 99,5% e 0,10% das sub bacias do Juruena, médio Tapajós e baixo Tapajós e Teles Pires respectivamente. A região 3 comporta 0,5% da bacia do Tapajós.

As regiões 2 e 3 são mais preservadas quanto a cobertura vegetal, uma vez que podem ser caracterizadas como bioma do tipo amazônico. Nota-se, especificamente, que as mesmas foram representadas pelas sub bacias do baixo e médio Tapajós. Em consequência, maior equilíbrio na fauna, flora e nos recursos hídricos evidenciados pelas maiores incidências de chuvas e vazões. Oliveira *et al.*, (2016) mostram que a cobertura vegetal se concentra na parte norte da bacia do Tapajós (identificadas pelas sub bacias

do baixo e médio tapajós) devido a existência de terras indígenas e unidades de conservação, assim neutralizando e amortecendo os impactos gerados pela degradação e fragmentação causadas pelo desmatamento e pela mudança de uso do solo.

3.2 Modelo de Regionalização de Vazões (Q₉₀)

No desenvolvimento do modelo de regionalização de vazões, por meio de regressão, a vazão (Q₉₀) foi relacionada em função das variáveis área de drenagem e precipitação mensal média. Para cada uma das três regiões homogêneas formadas tentou-se elaborar um modelo, no entanto, só foi possível propor um modelo de regionalização de vazões para região 1, uma vez que concentrou um maior número de estações fluviométricas estudadas, na qual comporta 79% de toda a área da bacia do Tapajós. Já para as regiões 2 e 3, não foi possível desenvolver modelos regionais, devido à baixa quantidade de estações concentradas nestas regiões.

Devido a variável “comprimento do rio” não ter apresentado boa correlação com a variável dependente Q₉₀, não foi preciso inseri-la no modelo. Assim a regressão linear múltipla foi aplicada às variáveis área de drenagem e a precipitação mensal média, no qual foram capazes de prever a vazão de referência Q₉₀ resultando em um modelo estatisticamente significativo. A Equação 6 representa o modelo de regionalização de vazões proposto para a região homogênea 1.

$$Q_{90} = 519,456 + 0,013.(A) - 3,111.(P) \quad (6)$$

em que Q₉₀ é a vazão de referência, A – área de drenagem (km²) e P – precipitação mensal média.

Uma vez que o valor do coeficiente de determinação (R²) entre as variáveis vazão (Q₉₀) e área de drenagem indicam uma forte correlação, como mostra a Figura 5. Sendo assim, pode-se dizer que a área de drenagem explica 94,3% da vazão observada. Portanto no teste F o grau de liberdade dado F (2, 20) =

195,895; p<0,001 hipótese alternativa é aceita, logo há correlação linear significativa. Portanto o modelo proposto atende as exigências de distribuição Normal.

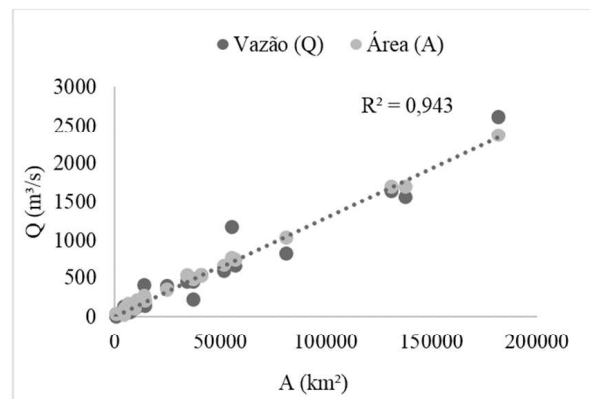


Fig. 5 - Correlação entre as variáveis vazão e área.

Sendo assim, o modelo de regionalização de vazões aqui proposto mostrou-se satisfatório na estimativa de vazões Q₉₀ na região homogênea 1, uma vez que na regressão múltipla alcançou índice de determinação (R²) Ajustado igual 0,946 e coeficiente de Nash igual 0,97 como mostra a Tabela 2.

A análise gráfica dos resíduos, dá indicação de que os resíduos parecem distribuir-se com dispersão constante, sugerindo que não há violações sérias dos pressupostos de homocedasticidade e apresentou média zero. Os resíduos padronizados encontram-se na faixa de -1 a 3 e sem a presença de outliers, como mostra a Figura 6.

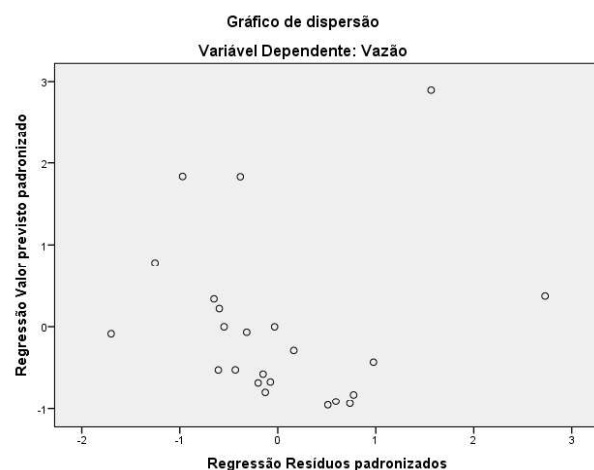


Fig. 6 – Resíduos Padronizados.

Tabela 2: Cálculo da vazão simulada e do coeficiente de Nash

Grupo	Código	Estações	Qobs	Qsim	(Qobs - Qsim) ²	(Qobs - Qmed) ²
	17093000	1	1181.83	760.81	177263	166227
	17095000	2	382.08	352.49	876	145987
	17120000	3	449.22	488.54	1547	201795
	17121000	4	104.89	113.30	71	11002
	17200000	6	117.34	206.99	8038	13768
	17210000	7	414.31	264.42	22466	171653
	17230000	8	79.24	108.82	875	6279
	17280000	9	456.01	529.97	5470	207942
	17300000	10	532.96	527.67	28	284044
	17340000	11	588.42	667.53	6260	346235
	17345000	12	18.10	38.15	402	328
G1	17350000	13	12.25	-63.58	5750	150
	17355000	14	52.50	-36.41	7905	2756
	17380000	15	844.44	1014.30	28853	713076
	17450000	19	57.80	-54.21	12545	3341
	17091000	22	133.33	15.92	13786	17778
	17122000	24	144.00	207.23	3998	20736
	17123000	25	653.62	740.33	7518	427222
	17675000	26	230.24	477.63	61203	53010
	17092800	27	153.98	174.15	407	23711
	17130000	5	2605.72	2330.27	75871	6789776
	17410000	16	1640.49	1669.98	870	2691210
	17420000	17	1554.05	1668.72	13150	2415061
Somatório (Σ)					455152	14713086
Coeficiente de Nash					0,97	

Fonte: Autores, 2016.

4. CONCLUSÕES

O presente estudo mostrou a eficiência da análise de agrupamento hierárquico de Ward com a formação de três regiões homogêneas bem definidas para a bacia do Tapajós. As regiões 2 e 3 apresentaram semelhança em suas características físico-climáticas, uma vez que apresentaram as maiores alturas pluviométricas e maiores vazões Q90, além de apresentarem uma vegetação mais intensa, visto que representam, quase no total, o exutório do rio Tapajós (bioma amazônico). Já a região 1 apresentou as menores alturas de precipitações mensais médias, assim como as menores vazões, justificado pela forte degradação da supressão da vegetação em função dos usos alternativos do solo, fator predominante

nas áreas do estado do Mato Grosso.

O modelo de regionalização de vazões, proposto para a região homogênea 1, mostrou-se satisfatório na estimativa das vazões Q90, visto que resultou bom desempenho do coeficiente de determinação ajustado (R^2 ajustado) e do coeficiente de Nash, provando que pode ser utilizado como ferramenta de auxílio na gestão dos recursos hídricos, bem como mensurar aspectos quantitativos de vazões à cursos de rios sem monitoramento por estações fluviométricas.

REFERÊNCIAS BIBLIOGRÁFICAS

AGARWAL, A.; MAHESWARAN, R.; V. SEHGAL, V.; KHOSA, R.; SIVAKUMAR, B.; BERNHOFER, C. Hydrologic regionalization

- using wavelet-based multiscale entropy method. **Journal of Hydrology**, v. 538, p. 22–32, 2016.
- BARBOSA, S. E. S.; BARBOSA JUNIOR, A. R.; SILVA, G. Q.; CAMPOS, E. N. B.; RODRIGUES, V. C. Geração de Modelos de Regionalização de Vazões Máximas, Médias de Longo Período e Mínimas de sete dias para a bacia do Rio do Carmo, Minas Gerais. **Revista de Engenharia Sanitária e Ambiental**, v.10, n.1, p. 64-71, 2005.
- BELL, C. D.; MCMILLAN, S. K.; CLINTON, S. M.; JEFFERSON, A. J. Hydrologic response to stormwater control measures in urban watersheds. **Journal of Hydrology**. v.541, p. 1488–1500, 2016.
- BROWN, R.; KEATH, N.; WONG, T. Transitioning to water sensitive cities: historical, current and future transition states. In: **11th international conference on urban drainage**, 2008. 10p.
- CARMO, C. M. R.; CHRISTENSEN, T. H. Cluster analysis of residential heat load profiles and the role of technical and household characteristics. **Energy and Buildings**, v. 125, p. 171–180, 2016.
- COLLISCHONN, W. **Simulação hidrológica em grandes bacias**. Tese (Doutorado em Engenharia de Recursos Hídricos e Saneamento Ambiental) – Universidade Federal do Rio Grande do Sul – UFRGS, Porto Alegre, 2001. 194p.
- DAWDY, D. R.; GRIFFIS, V. W.; GUPTA, V. K. Regional flood-frequency analysis: How we got here and where we are going. **Journal of Hydrologic Engineering**, v. 17, n. 9, p. 953-959, 2012.
- DOMINGUES, M.; BERMANN, C. O arco de desflorestamento na Amazônia: da pecuária à soja. **Ambiente & Sociedade**, v. 15, n. 2, p. 1-22, 2012.
- DONG, L.; WANG, L.; KHAHRO, S. F.; GAO, S. LIAO, X. Wind power day-ahead prediction with cluster analysis of NWP. **Renewable and Sustainable Energy Reviews**, v. 60, p. 1206–1212, 2016.
- ESQUERRE, O. P. K.; SEBORG, E. D.; BRUNS, E. ROY.; MORI, M. Application of steady-state and dynamic modeling for the prediction of the BOD of an aerated lagoon at a pulp and paper mill Part I. Linear approaches. **Chemical Engineering Journal**. v. 104, n. 1, p. 73-81, 2004.
- FARSADNIA, F. KAMROOD, M. R.; NIA, A. M.; MODARRES, R.; BRAY, M. T.; HAN, D.; SADATINEJAD, J. Identification of homogeneous regions for regionalization of watersheds by two-level self-organizing feature maps. **Journal of Hydrology**, v. 509, p. 387-397, 2014.
- FÁVERO, L. P.; BELFIORE, P.; SILVA, F. L.; CHAN, B. L. **Análise de dados: modelagem multivariada para tomada de decisões**. Rio de Janeiro: Elsevier, 2009. 646p.
- GOMES, E. P., BLANCO, C. J. C., PESSOA, F. C. L. Análise de Agrupamento para Identificação de Regiões Hidrológicamente Homogêneas na Amazônia. In: **XX Simpósio Brasileiro de Recursos Hídricos**. Bento Gonçalves-RS, 2013.
- GONÇALVES, M. F. BLANCO, C. J. C.; SANTOS, V. C.; OLIVEIRA, L. L. S.; PESSOA, F. C. L. Identification of Rainfall Homogenous Regions taking into account El Niño and La Niña and Rainfall Decrease in the state of Pará, Brazilian Amazon. **Acta Scientiarum. Technology**, v. 38, n. 2, p. 209, 2016.
- GRIFFIS, V. W.; STEDINGER, J. R. The use of GLS regression in regional hydrologic analyses. **Journal of Hydrology**, v. 344, p. 82–95, 2007.
- HAIR, J. F.; BLACK, W. C.; BABIN, B. J.; ANDERSON, R. E.; TATHAM, R. L. **Análise multivariada de dados**. Trad. Adonai S. Sant’Anna e Anselmo C. Neto. 6 ed. Porto Alegre: Bookman, 2009. 688p.
- HAWKINS, G. A.; VIVONI, E. R.; ROBLES-MORUA, A.; MASCARO, G.; RIVERA, E.; DOMINGUEZ, F. A climate change projection for summer hydrologic conditions in a semiarid watershed of central Arizona. **Journal of Arid Environments**. v. 118, p. 9–20, 2015.
- HIDROWEB – Sistemas de Informações Hidrológicas – **Agência Nacional de Águas - ANA**. Disponível em: <http://hidroweb.ana.gov.br>. Acesso em: 2015.
- HUANG, P.; LI, Z.; CHEN, J.; LI, Q.; YAO, C. Event-based hydrological modeling for detecting dominant hydrological process and suitable model strategy for semi-arid catchments. **Journal of Hydrology**. v.542, p. 292–303, 2016.
- ISHIHARA, J. H.; FERNANDES, L. L.; DUARTE, A. A. A. M.; LOUREIRO, G. E.

- Avaliação do monitoramento pluviométrico da Amazônia Legal. **Engenharia Ambiental: Pesquisa e Tecnologia**, v. 10, n. 3, p. 132-144, 2013.
- JOHNSON, R.A., WICHERN, D. W. **Applied multivariate statistical analysis**. New Jersey: Prentice-Hall, 6th.ed., 2007.
- KULT, J. M.; FRY, L. M.; GRONWOLD, A. D.; CHOI, W. Regionalization of hydrologic response in the Great Lakes basin: Considerations of temporal scales of analysis. **Journal of Hydrology**, v. 519 B, p. 2224–2237, 2014.
- KUMAR, D. B.; VERMA, S. Potential emission flux to aerosol pollutants over Bengal Gangetic plain through combined trajectory clustering and aerosol source fields analysis. **Atmospheric Research**, v. 178–179, p. 415–425, 2016.
- LANDIM, P. M. B. **Análise estatística de dados geológicos**. 2ª ed. revista e ampliada. São Paulo: Editora UNESP, 2003. 266p.
- LAVARINI, C.; MAGALHÃES JÚNIOR, A. P.; OLIVEIRA, F. S. Análise de agrupamentos de bacias hidrográficas como procedimento-guia para o estudo e a amostragem de coberturas superficiais. **Revista Brasileira de Cartografia**, Rio de Janeiro, n. 66/5, p. 1039-1050, 2014.
- LIU, L.; CHENG, Y. Efficient system reliability analysis of soil slopes using multivariate adaptive regression splines-based Monte Carlo simulation. **Computers and Geotechnics**. v.79, p. 41–54, 2016.
- MAMUN, A. A.; HASHIM, A.; DAOUD, J. I. Regionalisation of low flow frequency curves for the Peninsular Malaysia. **Journal of Hydrology**, v. 381, n. 1, p. 174-180, 2010.
- MASCARO, G.; VIVONI, E. R.; MÉNDEZ-BARROSO, L. A. Hyperresolution hydrologic modeling in a regional watershed and its interpretation using empirical orthogonal functions. **Advances in Water Resources**, v. 83, p. 190–206, 2015.
- MATTEAU, M.; ASSANI, A. A.; MESFIOUI, M. Application of multivariate statistical analysis methods to the dam hydrologic impact studies. **Journal of Hydrology**, v. 371, p. 120–128, 2009.
- MENDICINO, G.; SENATORE, A. Evaluation of parametric and statistical approaches for the regionalization of flow duration curves in intermittent regimes. **Journal of hydrology**, v. 480, p. 19-32, 2013.
- MOLOWNY-HORAS, R.; BASNOU, C.; PINO, J. A multivariate fractional regression approach to modeling land use and cover dynamics in a Mediterranean landscape. **Computers, Environment and Urban Systems**, v.54, p. 47–55, 2015.
- MONTGOMERY, D. C.; PECK, E. A.; VINING, G. G. **Introduction to linear regression analysis**, 5 th ed. John Wiley & Sons, 2012.
- NASH, J. E.; SUTCLIFFE, J. V. River flow forecasting through conceptual models part I—A discussion of principles. **Journal of hydrology**, v. 10, n. 3, p. 282-290, 1970.
- NGUYEN, P.; THORSTENSEN, A.; SOROOSHIAN, S.; HSU, K.; AGHAKOUCHAK, A.; SANDERS, B.; KOREN, V.; CUI, Z.; SMITH, M. A high resolution coupled hydrologic–hydraulic model (HiResFlood-UCI) for flash flood modeling. **Journal of Hydrology**. v.541, p. 401-420, 2016.
- OLIVEIRA, R. R. S.; VENTURIERI, A. V.; SAMPAIO, S. M. N.; LIMA, A. M. M.; ROCHA, E. J. P. Dinâmica de uso e cobertura da terra das regiões de integração do Araguaia e Tapajós/PA, para os anos de 2008 e 2010. **Revista Brasileira de Cartografia**, V. 68, n. 7, p. 1411-1424, 2016.
- PARENTE, J.; PEREIRA, M. G.; TONINI, M. Space-time clustering analysis of wildfires: The influence of dataset characteristics, fire prevention policy decisions, weather and climate. **Science of The Total Environment**, v. 559, p. 151–165, 2016.
- PEÑAS, F. J.; BARQUÍN, J.; ÁLVAREZ, C. Assessing hydrologic alteration: Evaluation of different alternatives according to data availability. **Ecological Indicators**, v. 60, p. 470–482, 2016.
- PEREIRA, J. C. R. **Análise de dados qualitativos**. São Paulo: Editora da Universidade de São Paulo/FAPESP, 1999.
- PESSOA, F. C. L. BLANCO, C. J. C. MARTINS, J. R. Regionalização de Curvas de Permanência de Vazões da Região da Calha Norte no Estado do Pará. **Revista Brasileira de Recursos Hídricos - RBRH**. Volume 16 n.2, p. 65-74, 2011.
- REGONDA, S. K.; SEO, D.; LAWRENCE, B.; BROWN, J. D.; DEMARGNE, J. Short-term ensemble streamflow forecasting using operationally-produced single-valued streamflow forecasts – A Hydrologic Model Output Statistics

- (HMOS) approach. **Journal of Hydrology**, v.497, p. 80–96, 2013.
- RIANNA, M.; RUSSO, F.; NAPOLITANO, F. Stochastic index model for intermittent regimes: from preliminary analysis to regionalisation. **Natural Hazards and Earth System Sciences**, v. 11, n. 4, p. 1189-1203, 2011.
- RODRIGUES, R. S. S.; FERNANDES, L. L.; CRISPIM, D. L., VIEIRA, A. S. A, PESSOA, F. C. L. Caracterização morfométrica da bacia hidrográfica do Igarapé da Prata, Capitão Poço – Pará – Brasil. **Revista Verde - Pombal – PB**. v. 11, n.3, p.143-150, 2016.
- RODRIGUEZ, D. A.; TOMASSELLA, J.; LINHARES, C. Is the forest conversion to pasture affecting the hydrological response of Amazonian catchments? Signals in the Ji-Paraná Basin. **Hydrological Processes**, v. 24, n. 10, p. 1254-1269, 2010.
- SANTOS, C. A.; SERRÃO, E. A. O.; GONÇALVES, L.J. M.; WANZELER, R. T. S.; LIMA, A. M. M. Zoneamento da distribuição da precipitação pluviométrica na bacia hidrográfica do rio Tapajós. **Enciclopédia Biosfera, Centro Científico Conhecer - Goiânia**, v. 10, n. 18, p. 3092, 2014.
- SEIDEL, E. J.; MOREIRA JÚNIOR, F. J.; ANSUJ, A. P.; NOAL, M. R. C. Comparação entre o método Ward e o método K-médias no agrupamento de produtores de leite. **Ciência e Natura – UFSM**. v.1, p.7-15, 2008.
- SILVA, P. M. O.; MELLO, C. R.; SILVA, A. M.; COELHO, G. Modelagem da hidrógrafa de cheia em uma bacia hidrográfica da região Alto do Rio Grande. **Revista Brasileira Engenharia Agrícola e Ambiental**. v.12, p. 258-265, 2008.
- SRIDHARAN, S.; SANDHYA, M. Long-term (2004–2015) tendencies and variabilities of tropical UTLS water vapor mixing ratio and temperature observed by AURA/MLS using multivariate regression analysis. **Journal of Atmospheric and Solar-Terrestrial Physics**, v. 147, p. 156–165, 2016.
- TAPAJÓS: **Hidrelétricas, infraestrutura e caos. Elementos para a governança da sustentabilidade em uma região singular**. Organizado por Wilson Cabral de Sousa Júnior. 1.ed. São José dos Campos: ITA/CTA, 2014. p. 92.
- TSAKIRIS, G.; NALBANTIS, I.; CAVADIAS, G. Regionalization of low flows based on canonical correlation analysis. **Advances in Water Resources**, v. 34, n. 7, p. 865-872, 2011.
- TUCCI, C. E. M. **Hidrologia: Ciência e Aplicação**. Porto Alegre, Ed. ABRH / UFRGS, 2009. 944p.
- VILLAR, J. C. E.; GUYOT, J. L.; RONCHAIL, J.; COCHONNEAU, G.; FILIZOLA, N.; FRAIZY, P.; LABAT, D.; OLIVEIRA, E.; ORDOÑEZ, J. J.; VAUCHEL, P. Contrasting regional discharge evolutions in the Amazon basin (1974–2004). **Journal of Hydrology**, v. 375, n. 3, p. 297-311, 2009.
- WARD, J, H. Hierarchical grouping to optimize an objective function. **Journal American Association**. v.58, p.236 – 244, 1963.
- WU, F.; ZHAN, J.; GÜNERALP, I. Present and future of urban water balance in the rapidly urbanizing Heihe River Basin, Northwest China. **Ecological Modelling**, v. 318, p. 254–264, 2015.
- YAN, R.; MA, Z.; KOKOGIANNAKIS, G.; ZHAO, Y. A sensor fault detection strategy for air handling units using cluster analysis. **Automation in Construction**, v. 70, p. 77–88, 2016.
- YANG, T.; SHAO, Q.; HAO, Z.; CHEN, X.; ZHANG, Z.; XU, C.; SUN, L. Regional frequency analysis and spatio-temporal pattern characterization of rainfall extremes in the Pearl River Basin, China. **Journal of Hydrology**, v. 380, n. 3, p. 386-405, 2010.
- ZHANG, J.; ZHANG, L.; DU, M.; ZHANG, W.; HUANG, X.; ZHANG, Y.; YANG, Y.; ZHANG, J-M.; DENG, S.; SHEN, F.; LI, F.; XIAO, H. Identifying the major air pollutants base on factor and cluster analysis, a case study in 74 Chinese cities. **Atmospheric Environment**, v. 144, p. 37–46, 2016.