



JOÃO PEDRO DE COIMBRA RIBEIRO

**REGIONALIZAÇÃO DA Q7,10 PARA A BACIA
HIDROGRÁFICA DO RIO DOCE**

**LAVRAS-MG
2022**

REGIONALIZAÇÃO DA Q7,10 PARA A BACIA HIDROGRÁFICA DO RIO DOCE

Monografia apresentada à Universidade Federal de Lavras, como parte das exigências do Curso de Engenharia Ambiental e Sanitária, para obtenção do título de Bacharel.

Prof. Dr. Carlos Rogério de Mello
Orientador
Dr. André Ferreira Rodrigues
Coorientador

JOÃO PEDRO DE COIMBRA RIBEIRO

REGIONALIZAÇÃO DA Q7,10 PARA A BACIA HIDROGRÁFICA DO RIO DOCE

REGIONALIZATION OF Q7,10 FOR THE DOCE RIVER BASIN

Monografia apresentada à Universidade Federal de Lavras, como parte das exigências do Curso de Engenharia Ambiental e Sanitária, para obtenção do título de Bacharel.

APROVADA em 8 de abril de 2022
Prof. Dr. Carlos Rogério de Mello
Prof. Dr. José Alves Junqueira Junior
Prof. Dr. Jhones da Silva Amorim

Prof. Dr. Carlos Rogério de Mello
Orientador
Dr. André Ferreira Rodrigues
Coorientador

**LAVRAS-MG
2022**

AGRADECIMENTOS

Agradeço a Deus por ter me dado forças nos momentos em que mais precisei.

À Universidade Federal de Lavras – UFLA por todos momentos vividos, auxiliando no crescimento profissional e pessoal.

Ao Departamento de Recursos Hídricos por proporcionar grandes aprendizados e grandes amizades.

À CNPq pela concessão da bolsa de iniciação científica em grande parte da minha graduação.

Ao professor orientador Carlos Rogério, por confiar e me incluir em seu laboratório de pesquisas.

Ao André pela amizade e paciência ao ensinar que ficarão guardados eternamente, gratidão.

Aos colegas de graduação que sempre estiveram juntos nos momentos de alegria e tristeza.

À todos os professores que contribuíram no meu aprendizado.

Aos amigos e colegas de laboratório e pesquisa, Renato, Vinicius de Oliveira, Tiguto, Vanessa, Vinícius Oliveira e Rafaela pelo apoio e aprendizado.

À minha mãe Ângela (in memoriam) por todos ensinamentos, que levo e vou levar para toda minha vida.

Ao meu pai José Geraldo pelo apoio e carinho em todos momentos.

Às minhas avós e avôs (in memoriam) pelo amor incondicional.

Ao meu tio Alzir por sempre estar ao meu lado e apoiar minhas decisões.

Aos meus primos e afilhados Reginaldo, Gabriela, João Gabriel e Gustavo por estarem sempre comigo e aconselhando a sempre correr atrás e nunca desistir dos meus sonhos.

À Elisângela, Liza Paula e Ueizer pelo acolhimento e apoio.

À minha sogra Roseli que me acolheu no fim da graduação quando muito precisei.

À minha companheira, Tatiane por nunca me deixar desistir de nada, me incentivando e apoiando quando mais precisei, sendo companheira e braço direito em todos momentos e me ajudando a superar obstáculos quando achei que não conseguiria. Gratidão.

Ao meu filho Murilo, que em breve estará nascendo, acredite na ciência, seja bondoso, forte e íntegro.

MUITO OBRIGADO!

RESUMO

As bacias hidrográficas são superfícies terrestres delimitadas por seus divisores de água topográficos, dentro dos quais toda a água que precipita drena para um único ponto (a foz). Entender os processos hídricos que ocorrem dentro de uma bacia hidrográfica é fundamental para a gestão dos recursos hídricos e, desta forma, para o desenvolvimento econômico de uma região. Para uma adequada gestão dos recursos hídricos, que atenda os usos múltiplos da água, a disponibilidade de informações sobre vazão e precipitação é crucial. No entanto, a baixa densidade de estações fluviométricas e pluviométricas é um problema a ser enfrentado no Brasil, sendo que muitas regiões são desprovidas destas informações e, assim, os estudos hidrológicos são prejudicados. Neste sentido, a regionalização de vazão é uma técnica amplamente aplicada para determinar os valores de vazões em bacias sem monitoramento. A outorga de água é garantida pela Política Nacional de Recursos Hídricos (nº 9.433/97) e proporciona aos interessados o acesso à água para usos múltiplos. As vazões mínimas de referência são necessárias para definir a quantidade de água a ser concedida em determinada bacia hidrográfica. Neste sentido, o objetivo deste trabalho foi realizar a regionalização das vazões médias mínimas para 7 dias consecutivos com um tempo de retorno de 10 anos ($Q_{7,10}$) para a Bacia Hidrográfica do Rio Doce e assim compreender o comportamento dos recursos hídricos na região. Utilizando a metodologia de clusterização, testando os métodos k-means e PAM com dissimilaridades Manhattan e Euclidiana, onde se dividiu as 25 regiões hidrologicamente semelhantes em clusters, após aplicar o teste de Mann-Kendall e utilizar as distribuições teóricas, os métodos k-means e PAM com dissimilaridades Manhattan e Euclides sucessivamente, para Bacia do Rio Doce, apresentou um bom valor de regionalização de vazões mínimas $Q_{7,10}$, com índices de confiança $(c) > 0,85$ para todas equações. A partir do valor da raiz quadrada do erro médio obtido para a validação cruzada é possível inferir que ambas as metodologias aplicadas foram capazes de estimar a $Q_{7,10}$ de forma satisfatória e podem ser utilizadas para a gestão dos recursos hídricos na bacia do rio Doce.

Palavras-chave: Validação cruzada, recursos hídricos, outorga.

ABSTRACT

Watersheds are land surfaces surrounded by ridges, where all the water that precipitates within these limits drains to a single point (the outlet). Understanding the water processes occurring in a watershed is fundamental for the water resources management, and therefore for the economic development of a region. The available information on streamflow and rainfall is crucial to meet the demands of multiple stakeholders and improve the management. However, gauges monitoring streamflow and rainfall are scarce in Brazil, which difficulties the water governance. In this sense, the regional frequency analysis is a widely applied technique to determine flow values in ungauged basins. The granting of water is guaranteed by the National Policy on Water Resources (n° 9433/97) and provides to stakeholders the access of water for multiple uses. Minimum reference flows are needed to define the amount of water to be granted in determined watershed. The reference flow of the Doce river watershed is the minimum value of the average flows for seven consecutive days with a return level of ten years ($Q_{7,10}$). Therefore, the aim of this study was to determine the values of $Q_{7,10}$ in ungauged regions within the Doce river watershed by means of regional frequency analysis. The k-means and PAM techniques with the Euclidian and Manhattan dissimilarity distances were considered for splitting twenty-five basins in hydrologically homogeneous regions. $Q_{7,10}$ were fit to the basin's area using linear regression models within the best two defined hydrologically homogeneous regions (k-means-Manhattan and PAM-Euclidean). The confidence index ($c > 0.85$) highlighted that the basin's area is an adequate predictor of $Q_{7,10}$ in ungauged regions. Moreover, the root-mean-square error proved the adequacy of the linear model using a k-fold cross validation. Therefore, the fitted linear models can be used for water resources management in the Doce river basin.

Keywords: Cross-validation, water resources, granting of water.

LISTA DE FIGURAS

Figura 1 – Localização da bacia hidrográfica do rio Doce no Brasil (a) e bacia hidrográfica do rio Doce (b). Fonte: Do autor (2021).....	8
Figura 2 - Estações fluviométricas localizadas na área de estudo e estações pluviométricas localizadas na área de estudo e ao entorno. Fonte: Do autor (2021).	9
Figura 3 - Bacia hidrográfica do rio Doce e sub-bacias utilizadas para o estudo. Fonte: Do autor (2021).	12
Figura 4 - Polígonos de Thiessen na Bacia hidrográfica do rio Doce e suas sub-bacias. Fonte: Do autor (2021).	15
Figura 5 - Resultado dos clusterização k-means Manhattan. Fonte: Do autor (2022).	20
Figura 6 - Resultado da clusterização PAM Euclides. Fonte: Do autor (2022).	21
Figura 7 - Regressões lineares e coeficientes de determinação (a) k-means Manhattan cluster 1, (b) k-means Manhattan cluster 2, (c) PAM Euclides cluster 1, (d) PAM Euclides cluster 2, (e) modelo global. Fonte: Do autor (2022).....	22
Figura 8 - Ajuste e incerteza na regionalização da vazão de referência (Q7,10) para as regiões homogêneas (a) k-means Manhattan cluster 1, (b) k-means Manhattan cluster 2, (c) PAM Euclides cluster 1 e (d) PAM Euclides cluster 2, (e) modelo global. Fonte: Do autor (2022).	25

LISTA DE TABELAS

Tabela 1 - Estações fluviométricas utilizadas com coordenadas e código da estação. Fonte: Do autor (2021).	9
Tabela 2 - Estações pluviométricas utilizadas com coordenadas e código da estação. Fonte: Do autor (2021).	10
Tabela 3 - Divisão das regiões semelhantes pelo método de clusters. Fonte: Do autor (2022)..	19
Tabela 4 – Equações de regionalização Q7,10 para as regiões homogêneas da BHRD e limite das áreas de drenagens para cada cluster. Fonte: Do autor (2022).....	23
Tabela 5 – Valores de RMSE (m ³ /s) para os clusters das regiões homogêneas e para o modelo global. Fonte: Do autor (2022).	24

SUMÁRIO

1	INTRODUÇÃO.....	1
1.1	Objetivo	2
2	REFERENCIAL TEÓRICO	3
2.1	Bacias Hidrográficas	3
2.2	Vazão de referência Q _{7,10} e Q _{90%}	3
2.3	Regiões hidrologicamente homogêneas.....	4
2.4	Regionalização de vazões	5
3	MATERIAL E MÉTODOS	6
3.1	Área de estudo.....	6
3.2	Dados pluviométricos e fluviométricos	8
3.3	Delimitação das sub-bacias	12
3.4	Determinação da Q _{7,10} , Q _{90%} , Q _{95%} e Q _{MLT}	13
3.5	Precipitação média na bacia	14
3.6	Regionalização de vazões	15
3.7	Validação e análise dos clusters.....	16
3.8	Validação cruzada	17
4	RESULTADOS E DISCUSSÃO.....	17
4.1	Análise de tendência	17
4.2	Distribuições Teóricas	18
4.3	Clusters	18
4.4	Modelos de regressão e regionalização de Q _{7,10}	21
4.5	Validação cruzada e incerteza nos modelos de regressão	23
5	CONCLUSÃO.....	26
	REFERÊNCIAS	27

1 INTRODUÇÃO

Estudos hidrológicos são cada vez mais necessários no Brasil, uma vez que a crescente demanda por múltiplos usos dos recursos hídricos somados a disponibilidade afetada por fenômenos climáticos diversos (e.g. secas) (CALEGARIO, 2014), os quais têm sido intensificados pelas mudanças climáticas (ALMAGRO et al., 2017), tem aumentando os conflitos pelos usos da água (PEREIRA, 2012). Neste sentido, leis federais, estaduais e municipais de gerenciamento de recursos hídricos são de suma importância na tentativa de solucionar problemas já existentes e na definição de políticas que busquem o gerenciamento integral dos recursos hídricos (DINIZ, 2008).

O gerenciamento integral dos recursos hídricos é dificultado nas regiões desprovidas de monitoramento hidrológico. Frente a esses problemas, procedimentos estatísticos têm sido adotados para determinar os valores das variáveis hidrológicas nestes locais, possibilitando, assim, a tomada de decisão.

Na gestão dos recursos hídricos uma das variáveis mais importantes é a vazão de referência, uma vez que é a partir dela que a disponibilidade hídrica para os múltiplos usos deve ser planejada. No Brasil, a depender da região onde se está realizando algum estudo de demanda hídrica, existem diferentes vazões de referências utilizadas. A maioria dos estados do Nordeste usam a vazão de referência $Q_{90\%}$, para a região Centro-Oeste $Q_{95\%}$ e para a região Sudeste a vazão de referência $Q_{7,10}$. (POMPEU et al., 2014).

A técnica para determinar as vazões de referências em locais desprovidos de informação é a de regionalização, sendo de suma importância no gerenciamento dos recursos hídricos (SOUZA et al., 2021). De acordo com Melati e Marcuzzo (2016), a regionalização de vazões é uma ferramenta que se mostra muito útil para a obtenção de vazões em locais onde não existem dados, medições em campo ou curvas chaves realizadas através de leituras de cotas em rios. Os autores relatam que a vazão é calculada para regiões sem nenhum registro de dados através de outras regiões hidrologicamente semelhantes a ela e que possuem informações de vazão, área, região climática, dentre outras observações.

Beskow et al. (2016) em seu estudo, em uma bacia hidrográfica no estado do Rio Grande do Sul, obtiveram sucesso ao utilizarem inteligência artificial para definição de regiões homogêneas no processo de regionalização, sendo que os modelos matemáticos ajustados com base apenas na área de drenagem apresentaram uma boa estimativa na determinação da $Q_{90\%}$. Souza et al. (2021) também utilizaram a variável explicativa área de drenagem para seu estudo de regionalização utilizando o método da curva adimensional na região de Lavras-MG, pois se trata de uma variável simples e de fácil utilização em locais onde não existem calibração.

Pesquisadores em seus estudos comprovaram que a análise de agrupamento, conhecida por

clusterização, vem obtendo sucesso na divisão de regiões hidrologicamente homogêneas (PORTO et al. 2004). A clusterização busca dividir em grupos os indivíduos não classificados em grupos, onde esses grupos são chamados de clusters, onde, cada cluster possui indivíduos com alta similaridade entre eles e dissimilaridade entre indivíduos de outros clusters (NAGHETTINI E PINTO, 2007).

A Bacia Hidrográfica do Rio Doce (BHRD) está situada em sua grande maioria no estado de Minas Gerais, onde a vazão de referência utilizada é a $Q_{7,10}$. Devido à extensa área dessa bacia (83.400 km²) e sua grande importância econômica, a necessidade de definir políticas ambientais mais assertivas é fundamental, sobretudo após o desastre provocado pelo rompimento da barragem de rejeito da Samarco. Assim, a regionalização da vazão dessa bacia é uma ferramenta que assume um importante papel para a gestão de recursos hídricos dessa bacia, pois pode auxiliar os órgãos reguladores na autorização de outorga de direito de uso de águas em regiões onde não existem dados e monitoramento de vazões.

1.1 Objetivo

Objetivou-se no presente trabalho regionalizar a vazão de referência para outorga $Q_{7,10}$ na bacia hidrográfica do rio Doce pela metodologia de clusterização de regiões hidrologicamente semelhantes e validando esses resultados através da validação cruzada one-leave out e testando as incertezas através da técnica de bootstrap de modo obter resultados robustos para uma melhor gestão hídrica na bacia hidrográfica.

2 REFERENCIAL TEÓRICO

2.1 Bacias Hidrográficas

A bacia hidrográfica pode ser definida como uma área de captação natural da água de precipitação que faz convergir o escoamento para um único ponto de saída. Compõe-se de um conjunto de superfícies vertentes e de uma rede de drenagem formada por cursos de água que confluem até resultar em um leito único no seu exutório (TUCCI, 1997).

Para Barrela (2001), as bacias são um conjunto de terras drenadas por um rio e seus afluentes, formada nas regiões mais altas do relevo por divisores de água, onde as águas das chuvas, ou escoam superficialmente formando os riachos e rios, ou infiltram no solo para formação de nascentes e do lençol freático.

A Política Nacional de Recursos Hídricos (instituída pela lei 9.433/1997) adota a definição para bacias hidrográficas como uma unidade de estudo e gestão (BRASIL, 1997). Assim, é de grande importância para gestores e pesquisadores a compreensão do conceito de bacia hidrográfica e de suas subdivisões (TEODORO et al, 2007).

De acordo com Cardoso (2016) devido à complexidade de informações que uma bacia hidrográfica apresenta, é preciso ter cautela pra realizar o estudo sobre a mesma, assim, Neto (2016) evidencia que uma bacia hidrográfica emerge de um contexto sistêmico e devido a sua complexidade é necessário diferentes áreas de conhecimento devido essa necessidade de estudos integrados.

O conhecimento das características morfométricas da região de um estudo, no presente caso, uma bacia hidrográfica e suas sub-bacias é fundamental para o entendimento e conhecimento dos recursos que ali existem. Desse modo Ferreira (2012) conclui que é de suma importância conhecer a morfometria de uma bacia para a conservação dos seus recursos hídricos e compreender o comportamento dos ciclos hidrológicos.

2.2 Vazão de referência $Q_{7,10}$ e $Q_{90\%}$

Segundo o artigo 2º da Portaria IGAM N° 48 de 4 de outubro de 2019, para o estado de Minas Gerais – MG, fica estabelecido que a vazão de referência a ser utilizada para o cálculo das disponibilidades hídricas superficiais no estado de Minas Gerais é a vazão mínima de sete dias de duração e dez anos de recorrência – $Q_{7,10}$ (IGAM, 2019).

No que se refere à vazão de referência para outorgas, o IEMA emitiu a IN n° 013/2009, estabelecendo a $Q_{90\%}$ (vazão de referência com 90% de permanência, ou que é igualada ou superada em 90% da série disponível de vazões) para todos os cursos de água superficiais do

estado do Espírito Santo - ES, sendo o valor máximo outorgável em uma seção de curso de água limitado em 50% dessa vazão (PERH/ES, 2018).

A curva de permanência gerada para a Q90% e Q95% para cada posto fluviométrico pode ser construída organizando todos os registros de vazão diária em ordem decrescente. Assim, pode-se dizer que o valor máximo teve probabilidade de ser igualado ou superado igual ao inverso do número de registros. Seguindo assim, a cada registro é atribuída a probabilidade de excedência, formando a curva de permanência. As curvas de permanência relacionam a vazão com a porcentagem de tempo em que ela é igualada ou superada (MENDES, 2007).

Ribeiro, Marques e Silva (2005), regionalizaram a vazão da bacia hidrográfica do rio Doce para vazões mínimas de referência e as vazões $Q_{7,10}$ e $Q_{90\%}$ foram termos de referência utilizados pelos autores. Mendes (2007) apresenta que para o cálculo da $Q_{7,10}$ deve-se dispor de uma série histórica com pelo menos 30 anos de registros de vazões diárias, e é desejável que nos meses de estiagem não haja falha de modo a não prejudicar a representatividade da amostra.

Wolff et al. (2012) em seu estudo comparou e avaliou os valores obtidos para as vazões de referência ($Q_{7,10}$, $Q_{90\%}$, $Q_{95\%}$) pela metodologia de regionalização hidrológica do estado de São Paulo, com os calculados através das estações fluviométricas do Departamento de águas e energia elétrica – DAEE, considerando o DAEE como padrão, dada tal importância aos valores de referência para o uso da água tanto no estado de São Paulo como em outros estados do Brasil.

2.3 Regiões hidrologicamente homogêneas

O “clustering” é uma classificação não supervisionada de vários tipos de padrões como observações, dados ou vetores de recursos em grupos que são chamados de cluster. Desse modo os padrões dentro de um mesmo cluster são mais similares entre si do que com os padrões pertencentes a um cluster diferente (JAIN; MURTY; FLYNN, 1999).

Existem diversas metodologias para agrupar as bacias hidrográficas em regiões hidrologicamente homogêneas (CASSALHO et al., 2019). Estas metodologias baseiam-se em critérios de dissimilaridade, os quais “classificam” o nível de semelhança ou dessemelhança entre as bacias hidrográficas providas de monitoramento (NAGHETTINI E PINTO, 2007) utilizando parâmetros que são importantes para as vazões mínimas (e.g., precipitação média, área da bacia, latitude e longitude). Beskow et al. (2016) em seu estudo de regionalização de vazões, utilizou como medidas de dissimilaridades as distâncias: (i) Euclidiana; (ii) Manhattan; e (iii) Mahalanobis, os quais apresentaram bons resultados. A diferença entre estes métodos reside em que a distância Euclidiana é a distância entre dois pontos em um espaço qualquer desde que esse espaço seja definido por um sistema de coordenada cartesiana retangular

(HARRIS, 1955), a distância Manhattan é uma simplificação da distância Euclidiana e é um método mais eficiente para aplicações em tempo real devido a sua simplicidade (KUGLER et al., 2003). Já a distância Mahalanobis, leva em consideração as correlações dos dados de entrada entre si, e é uma boa escolha para regiões hiperelipsoidais (ZUCHINI, 2003).

Segundo Jain (2010), a clusterização k-means é muito popular devido sua simplicidade, velocidade de execução, eficiência e sucesso empírico. Outra metodologia utilizada é a k-medoides com o algoritmo Partitioning Around Medoids (PAM), que de acordo com Kaufman e Rousseeuw (1990), foi desenvolvida para minimizar os problemas com outliers no agrupamento de dados. O método k-medoides – PAM é mais robusto que o k-means na presença de outliers, pois, o método k-medoides é menos influenciado por valores discrepantes (HAN; KAMBER; PEI, 2012).

2.4 Regionalização de vazões

A regionalização é uma ferramenta de grande importância para monitorar áreas onde existe a falta de informações hidrológicas e para que ela ocorra, utiliza-se da exploração de informações de regiões onde o monitoramento é efetivo e constante (SILVA et al, 2006). Estas técnicas de transferência de informações hidrológicas denominam-se regionalização (BARBOSA et al., 2005). As vazões são as variáveis hidrológicas que mais demandam regionalização devido à escassa rede de monitoramento (quando comparada à pluviométrica) e sua importância para o planejamento integrado dos recursos hídricos (vazões mínimas) e dimensionamento de obras hidráulicas (vazões máximas), i.e., para o correto planejamento e manejo dos sistemas de recursos hídricos. Nesses casos, se faz necessária à aplicação de técnicas de transferência de informações de outros locais hidrologicamente semelhantes para a bacia hidrográfica em estudo a partir de informação de mais fácil acesso como: (i) as características fisiográficas e (ii) meteorológicas das regiões desprovidas de informações de vazão.

Existem duas formas de proceder com a regionalização de vazões: (i) regionalização de quartis (e.g., Q_{90} e Q_{95}) ou vazões de referência ($Q_{7,10}$) e (ii) dos parâmetros da distribuição de probabilidades para vazões máximas e/ou mínimas. O primeiro diz respeito ao ajuste de equações lineares ou não lineares de valores específicos de vazão (máxima ou mínima) com relação às características fisiográficas (e.g., área da bacia, comprimento do talvegue, declividade média) e meteorológicas (e.g., precipitação máxima diária anual e precipitação média anual) das bacias com monitoramento. O segundo procedimento é a regionalização dos parâmetros das distribuições de probabilidades (FDPs) ajustadas aos dados de vazões máximas ou mínimas. Após a regionalização dos parâmetros pode-se determinar o valor da vazão para o

tempo de retorno desejado (TUCCI, 1997).

A regionalização direta dos valores de vazão é o procedimento mais utilizado no Brasil devido à sua praticidade e à prévia definição das vazões de referência nas legislações estaduais (e.g., MINAS GERAIS, 2019). Vazões de máxima são normalmente regionalizadas por meio da técnica “index-flood” (CASSALHO et al., 2019; AMORIM et al., 2020; SOUZA et al., 2021), enquanto que para as vazões de mínima utilizam-se da técnica de regressão (BESKOW et al., 2016). A técnica do “index-flood” considera a premissa de que dentro de uma região hidrologicamente homogênea uma única função distribuição de probabilidade descreve o comportamento das vazões máximas em todos os pontos com monitoramento, i.e., pode-se ajustar uma FDP regional, cujos parâmetros são os mesmos para todas as bacias dentro da região homogênea (NAGHETTINI; PINTO, 2007). O que difere as bacias dentro da região é o “index-flood”, o qual é representado pela média de longo-termo das vazões monitoradas em cada bacia (NAGHETTINI; PINTO, 2007).

Para o caso das vazões mínimas, estas são diretamente relacionadas aos parâmetros fisiográficos e meteorológicos das bacias com monitoramento que se encontram dentro da mesma região homogênea (BESKOW et al., 2016). As vazões de referência são então regionalizadas a partir do levantamento dos respectivos parâmetros nas bacias sem monitoramento, os quais podem ser realizados usando apenas ferramentas de sistemas de informação geográfica. Segundo Lisboa et al., (2008), Beskow et al., (2016), Melati e Marcuzzo (2016) a área da bacia é suficiente para descrever a variabilidade espacial das vazões de referência em regiões hidrologicamente homogêneas.

O objetivo de regionalizar vazões mínimas é de suma importância para criar uma base de dados hidrológicos e oferecer subsídios para tomadores de decisão no processo de concessão de outorga do direito e uso dos recursos hídricos (DA SILVA, 2009).

3 MATERIAL E MÉTODOS

3.1 Área de estudo

A bacia hidrográfica do rio Doce (BHRD) encontra-se na região sudeste do Brasil (Figura 1, a), entre as latitudes 18°45' e 21°15' sul e longitudes 39°55' e 43°45' oeste, com área de drenagem aproximada de 83.400 km² (Figura 1, b). O estado de Minas Gerais – MG possui 86% do total da área da bacia, enquanto que os outros 14% pertencem ao estado do Espírito Santo – ES. O rio Doce é o principal curso d'água da bacia, estendendo-se por 879 km até desaguar no oceano Atlântico, na cidade de Linhares – ES. O rio Doce é formado pela

confluência dos rios Piranga e Carmo, cujas nascentes estão situadas na Serra da Mantiqueira e Serra do Espinhaço, respectivamente.

De acordo com o Plano Integrado de Recursos Hídricos da Bacia do rio Doce – PIRH (2010) a bacia possui 229 municípios e uma população total aproximada de 3.294.000 habitantes. Entre as atividades econômicas desenvolvidas na BHRD, merecem destaque a mineração, siderurgia, silvicultura e pecuária.

Segundo a classificação de Köppen, três tipos climáticos são identificados na BRDH: Na região das vertentes das serras do Espinhaço e Mantiqueira e na nascente do rio Doce está presente o clima tropical de altitude com chuvas de verão e verões frescos (Cwb). Nas nascentes de seus afluentes está presente o clima tropical de altitude com chuvas de verão e verões quentes (Cwa). Já nos trechos baixo e médio o clima quente com chuvas de verão (Aw) (LOUZADA, 2016).

O regime pluviométrico é caracterizado por duas estações distintas: período seco que se estende de abril a setembro com estiagem crítica de junho a agosto e o período chuvoso que se inicia em outubro e termina em março com maiores índices de precipitação no mês de dezembro (PIRH, 2010). Observando uma distribuição da precipitação heterogênea no interior da bacia com totais anuais mínimos entre 700 mm a 1000 mm em regiões deprimidas e fundo de vales e em regiões de maiores altitudes e litorâneas podendo chegar a 1500 mm. (COELHO, 2009).

De acordo com o Comitê de Bacia do Rio Doce (2021), a BHRD possui 98% de toda sua área inserida no bioma Mata Atlântica e os outros 2% no Cerrado. O solo da bacia, de acordo com dados do LAPIG MAPS, é de predominância Latossolo (60,6%), seguido de Argissolo (33,1%), Cambissolo (3,6%), Neossolo (2%) e afloramento de rochas (0,7%) (LAPIG, 2021). O uso e ocupação do solo segundo (IBGE 2014) é predominante de áreas agrícolas com remanescentes florestais (71,72%) seguido de vegetação florestal com atividades agrícolas (14,59%), pastagem com manejo (5,81%), silvicultura (4,83%), pastagem natural (0,76%), áreas artificiais (0,74%), área agrícola (0,66%), áreas agrícolas com remanescentes Campestres (0,42%), vegetação Campestre (0,26%) e corpos d'água continental (0,2%).

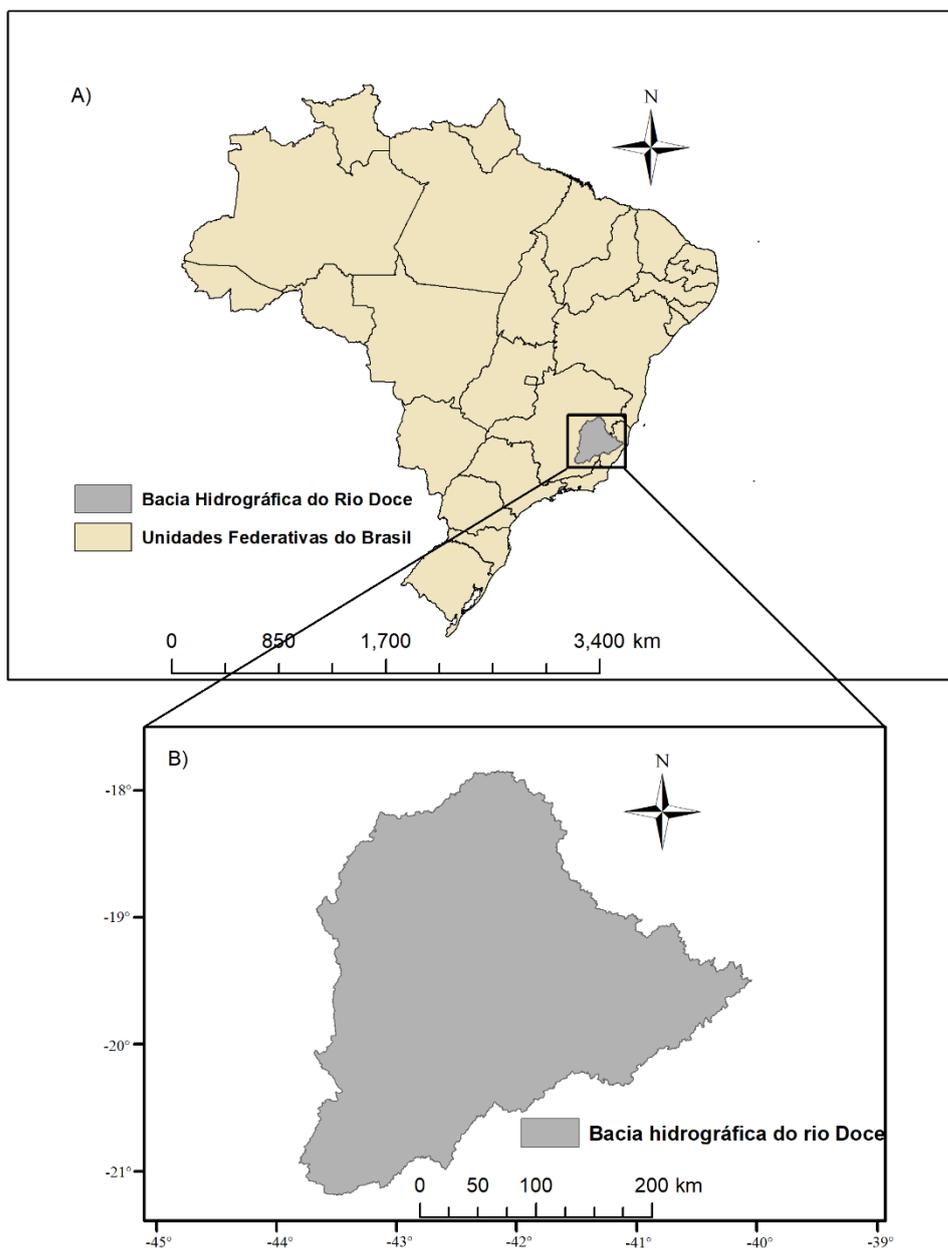


Figura 1 – Localização da bacia hidrográfica do rio Doce no Brasil (a) e bacia hidrográfica do rio Doce (b). Fonte: Do autor (2022).

3.2 Dados pluviométricos e fluviométricos

Para realização do estudo foram utilizadas 38 estações fluviométricas (Tabela 1) e 55 estações pluviométricas (Tabela 2) com dados diários abrangendo o período de 1980 até 2017. Estas foram selecionadas respeitando o critério de máximo de falhas igual a 10%. Os dados foram providenciados pela Agência Nacional de Águas – ANA, a qual é a gestora responsável por todas as estações utilizadas no presente trabalho (Figura 2).

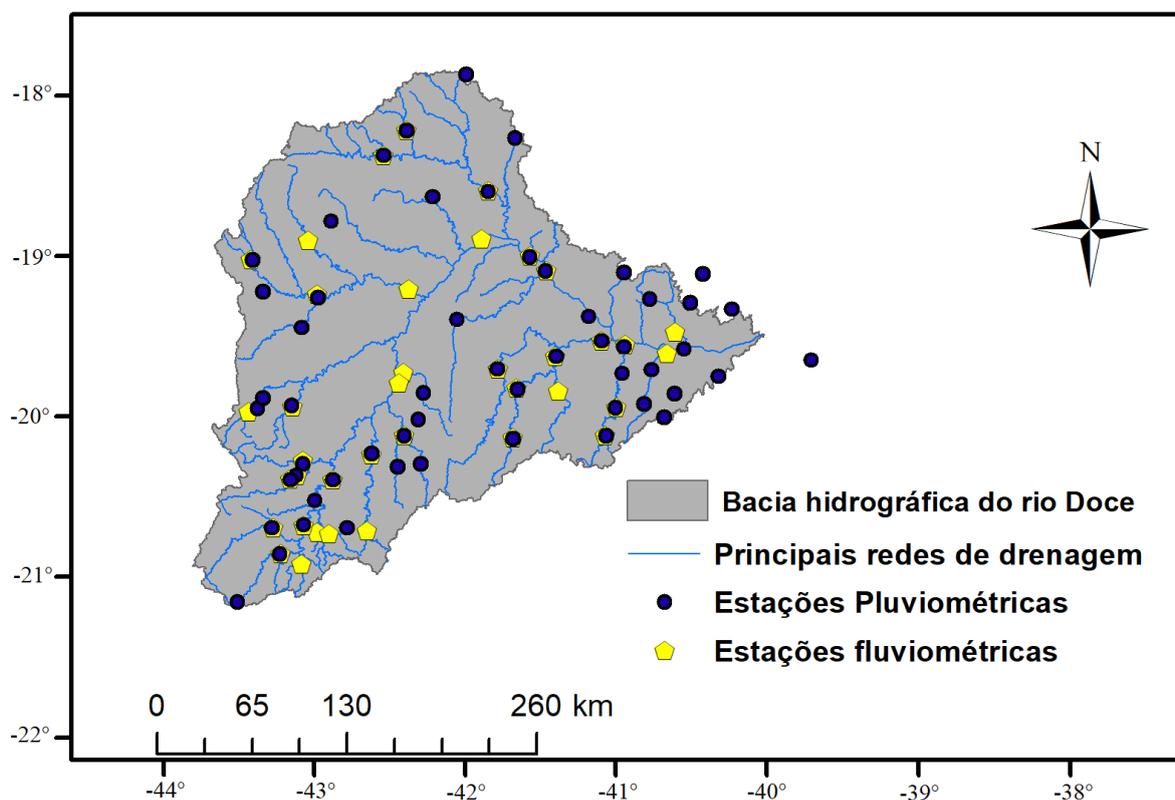


Figura 2 - Estações fluviométricas localizadas na área de estudo e estações pluviométricas localizadas na área de estudo e ao entorno. Fonte: Do autor (2022).

Tabela 1 - Estações fluviométricas utilizadas com coordenadas, código da estação e área de drenagem de cada estação. Fonte: Do autor (2022).

Estação	Código	Latitude	Longitude	Área (km ²)
Acaiaca Jusante	56335001	-20.36138889	-43.13944444	1256
Afonso Cláudio Montante	56990990	-20.07750000	-41.12416667	396.6
Assarai Montante	56989400	-19.59472222	-41.45805556	3188.7
Baixo Guandu	56992000	-19.52361111	-41.01416667	2108.6
Barra Do Cuieté Jusante	56940002	-19.06194444	-41.53333333	3202.7
Braz Pires	56055000	-20.84750000	-43.24194444	1092
Cachoeira Dos Óculos Montante	56539000	-19.77694444	-42.47638889	15532.95
Carrapato (Brumal)	56640000	-19.97166667	-43.45888889	419.8
Conceição Do Mato Dentro	56750000	-19.01444444	-43.44611111	199.13
Dores De Manhumirim	56983000	-20.10722222	-41.72833333	370.18
Fazenda Ocidente	56337000	-20.26722222	-43.10083333	529.6
Fazenda Paraíso	56240000	-20.39027778	-43.18194444	775.2
Fazenda Varginha	56090000	-20.71444444	-42.99972222	326.7
Ferros	56775000	-19.23222222	-43.02000000	3933.9
Governador Valadares	56850000	-18.88305556	-41.95083333	40103

Ipanema	56988500	-19.79888889	-41.70611111	2373
Jusante Córrego Da Piaba	56993551	-19.55861111	-40.73250000	881.7
Laranja Da Terra	56991500	-19.90111111	-41.05805556	1294.6
Mutum	56989001	-19.81055556	-41.4375	1189.2
Naque Velho	56825000	-19.18805556	-42.42277778	10035.6
Pingo D'água	56570000	-19.70861111	-42.44555556	814.7
Piranga	56028000	-20.69055556	-43.29944444	1400
Ponte Do Pancas	56995500	-19.42277778	-40.68638889	955.2
Ponte Nova Jusante	56110005	-20.38388889	-42.90277778	6198.8
Porto Firme	56075000	-20.67027778	-43.08805556	4236.9
Raul Soares Montante	56484998	-20.10361111	-42.44	1360.4
Rio Casca	56415000	-20.22611111	-42.65	2019.7
Rio Piracicaba	56610000	-19.93166667	-43.17305556	1164.5
Santa Maria Do Suaçuí	56870000	-18.20166667	-42.45472222	541.6
Santo Antônio Do Manhauçu	56978000	-19.67861111	-41.83583333	2373
São Miguel Do Anta	56385000	-20.69972222	-42.67305556	512.8
São Pedro Do Suaçuí	56860000	-18.36305556	-42.60222222	2244.7
São Sebastião Da Encruzilhada	56990000	-19.4925	-41.16166667	8712.9
Senador Firmino	56065000	-20.91166667	-43.09722222	279.7
Senhora Do Porto	56800000	-18.89472222	-43.0825	1515.3
Seriquite	56085000	-20.72333333	-42.92277778	340.89
Tumiritinga	56920000	-18.97111111	-41.64166667	52569.9
Vila Matias Montante	56891900	-18.57472222	-41.91777778	8706.1

Tabela 2 - Estações pluviométricas utilizadas com coordenadas e código da estação. Fonte: Do autor (2022).

Estação	Código	Latitude	Longitude
Abre Campo	2042010	-20.29888889	-42.47805556
Acaiaca - Jusante	2043009	-20.3625	-43.14388889
Afonso Cláudio - Montante	2041023	-20.07861111	-41.12138889
Alto Rio Novo	1941012	-19.05916667	-41.0275
Assarai - Montante	1941006	-19.59472222	-41.45805556
Baixo Guandu	1941003	-19.52361111	-41.01416667
Barra De São Gabriel	1940016	-19.05777778	-40.51638889
Barra Do Cuieté - Jusante	1941005	-19.06166667	-41.53277778
Bom Jesus Do Galho	1942002	-19.83361111	-42.31777778
Braz Pires	2043026	-20.8475	-43.24194444
Caldeirão	1940020	-19.955	-40.74166667
Campanário	1841003	-18.23861111	-41.74861111
Cavalinho	1940005	-19.69222222	-40.39805556

Colatina - Corpo De Bombeiros	1940006	-19.53083333	-40.62305556
Conceição Do Mato Dentro	1943002	-19.01666667	-43.44416667
Coroaci	1842005	-18.61194444	-42.27861111
Desterro Do Melo	2143003	-21.14916667	-43.52
Dom Cavati	1942008	-19.37361111	-42.105
Dores Do Manhumirim	2041008	-20.10805556	-41.72833333
Fazenda Ocidente	2043027	-20.28555556	-43.09888889
Fazenda Paraíso	2043011	-20.39	-43.18027778
Ferros	1943003	-19.25027778	-43.01444444
Guanhães	1842007	-18.77222222	-42.93111111
Ibituba	1941009	-19.69111111	-41.02
Ipanema	1941000	-19.79888889	-41.70611111
Itaimbé	1940012	-19.66361111	-40.83527778
Itarana	1940000	-19.87444444	-40.87444444
Laranja Da Terra	1941008	-19.90111111	-41.05805556
Malacacheta	1742017	-17.84555556	-42.07555556
Matipó	2042017	-20.27722222	-42.32555556
Morro Do Pilar	1943025	-19.2175	-43.37416667
Novo Brasil	1940013	-19.2375	-40.59138889
Pancas	1940009	-19.22027778	-40.85333333
Piranga	2043010	-20.69055556	-43.29944444
Ponte Nova Jusante	2042018	-20.38472222	-42.90277778
Porto Firme	2043014	-20.67027778	-43.08805556
Povoação	1939002	-19.5775	-39.79444444
Raul Soares - Montante	2042008	-20.10361111	-42.44
Resplendor - Jusante	1941004	-19.34305556	-41.24611111
Rio Bananal	1940023	-19.27416667	-40.32083333
Rio Casca	2042011	-20.21472222	-42.65222222
Rio Piracicaba	1943001	-19.92277778	-43.17777778
Santa Bárbara	1943007	-19.94527778	-43.40111111
Santa Maria Do Itabira	1943008	-19.44027778	-43.11861111
Santa Maria Do Suaçuí	1842008	-18.20111111	-42.45527778
Santo Antônio Do Manhuaçu	1941011	-19.67833333	-41.83611111
São João De Petrópolis	1940001	-19.80527778	-40.67888889
São Miguel Do Anta	2042016	-20.6825	-42.80666667
São Pedro Do Suaçuí	1842004	-18.36305556	-42.60222222
São Sebastião Da Encruzilhada	1941010	-19.4925	-41.16166667
Tumiritinga	1841011	-18.97638889	-41.64027778
Usina Da Brecha	2043025	-20.51666667	-43.01666667
Usina Peti	1943027	-19.88083333	-43.3675
Vermelho Velho	1942006	-19.99888889	-42.3475
Vila Matias - Montante	1841001	-18.57472222	-41.91777778

3.3 Delimitação das sub-bacias

O modelo digital de elevação (MDE) utilizado foi o produto da missão Aster (<https://search.earthdata.nasa.gov/search>), cuja resolução espacial é de 30m x 30m. O elipsoide de referência utilizado foi o SAD 1969. O procedimento para delimitação da BHRD por meio do MDE, a (Figura 1-b) seguiu as seguintes etapas: (i) preenchimento de depressões espúrias (fill), (ii) direção de fluxo (flow direction), (iii) fluxo acumulado (flow accumulation), (iv) criação das redes de drenagem através da calculadora raster (raster calculator) com condição de fluxo acumulado ≥ 3000 (foi definido o fluxo acumulado ≥ 3000 pois valores menores geraram cursos d'água efêmeros e esse valor foi o que mais se aproximou dos cursos d'água mapeados pela ANA), (v) criação do ponto de fluidez ("snap pour point") e (vi) delimitação da bacia (Watershed) nos pacotes "Hydrology" e "Map, Algebra". As sub-bacias (referentes a cada estação fluviométrica) foram delimitadas utilizando o MDE da bacia do rio Doce criada no processo acima, utilizando as ferramentas criação do ponto de fluidez (snap pour point) para coordenada de cada estação fluviométrica (Tabela 1) e delimitação da bacia (Watershed) do pacote "Hydrology". Todos estes procedimentos foram realizados no programa ArcGis 10.5. (Figura 3). Para cada uma das 38 sub-bacias geradas foram extraídas as informações de área e perímetro usando a função "calculate geometry" na tabela de atributos.

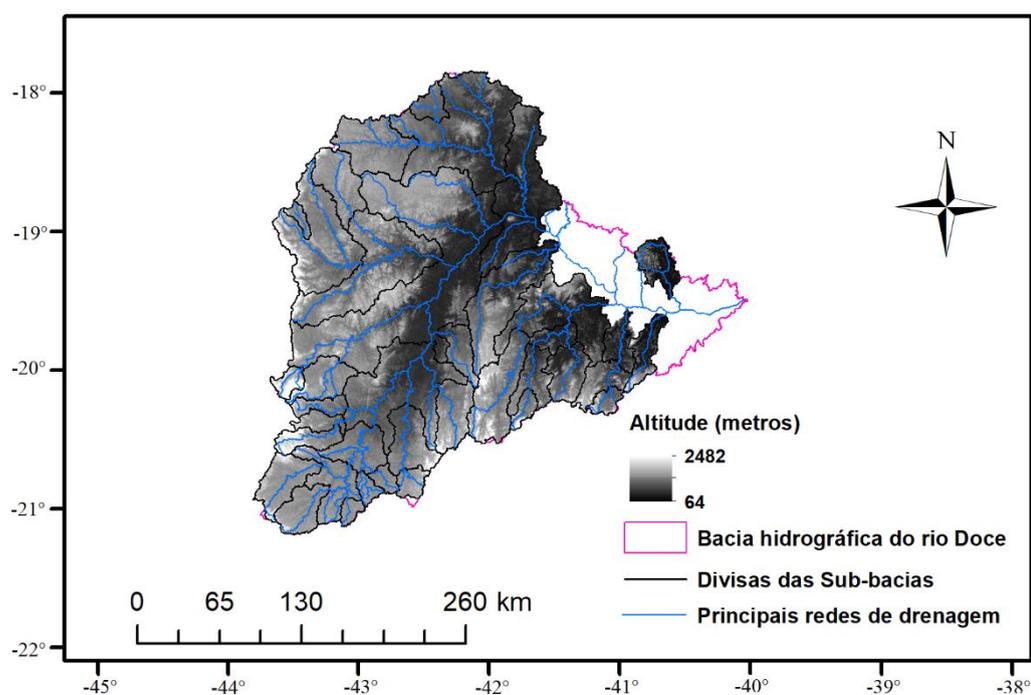


Figura 3 - Bacia hidrográfica do rio Doce e sub-bacias utilizadas para o estudo. Fonte: Do autor (2022).

3.4 Determinação da $Q_{7,10}$, $Q_{90\%}$, $Q_{95\%}$ e Q_{MLT}

Antes dos valores de $Q_{7,10}$, $Q_{90\%}$, $Q_{95\%}$ e Q_{MLT} (vazão média de longo termo) serem determinados, os dados diários de vazão foram primeiramente filtrados através do aplicativo Hidro1.4 disponibilizado pelo site da Hidroweb – ANA (<https://www.snirh.gov.br/hidroweb/download>) e organizados em planilhas do Microsoft Office Excel - 2016 (para a identificação e eliminação de “outliers”) que segundo Mello e Silva (2013) é recomendável pois esses dados podem comprometer os testes estatísticos de aderência e posterior aplicação das distribuições de probabilidade.

Os valores de Q_7 foram calculados através da média móvel das vazões de 7 dias consecutivos para cada ano da série histórica de cada estação fluviométrica (Tabela 1). Após a média móvel ser calculada, obteve-se então o menor valor de vazão para cada ano, construindo assim a série histórica de Q_7 .

O teste de Mann-Kendall a 5% de significância foi aplicado para averiguar a premissa de estacionariedade nas séries históricas de Q_7 . Nove funções densidade de probabilidade (FDPs) foram testadas para determinar àquela que melhor descreve a distribuição dos valores de Q_7 para cada estação fluviométrica, objetivando assim a maior assertividade na determinação dos valores de $Q_{7,10}$ (NAGHETTINI E PINTO, 2007). A adequabilidade das FDPs Gumbel, Weibull, Gamma, GEV, Logística, Pareto, Wakeby, Pearson e LN 3P foi avaliada por meio do teste de Anderson-Darling a 5% de significância para medir o quão bem os dados seguem as distribuições em particular.

Os valores de $Q_{90\%}$ e $Q_{95\%}$ foram determinados a partir da construção da curva de permanência. Para isso, os dados diários de vazão foram organizados em ordem decrescente e as respectivas frequências de excedência determinadas (Equação 1). A vazão que é igualada ou superada em 90% (95%) do tempo (frequência de excedência) equivale a $Q_{90\%}$ ($Q_{95\%}$).

$$\text{FreqObs} = \frac{\text{Ordem}}{N+1} \quad \text{Eq. (1)}$$

Onde “Ordem” é a posição em que o número se encontra e “N” é o número total de dados existentes na coluna.

A vazão média de longo termo (Q_{MLT}) (Equação 2) foi obtida através da média de todos os valores de vazão para todos os anos em estudo de determinada estação fluviométrica.

$$Q_{MLT} = \frac{Q_1+Q_2+Q_3+Q_4+\dots+Q_N}{N} \quad \text{Eq. (2)}$$

Onde os valores de Q são os dados diários de vazão para a série histórica inteira e N é o número total de dados que compõe a série histórica.

$$R_{90} = \frac{Q_{90}}{Q_{MLT}} \quad \text{Eq. (3)}$$

$$R_{95} = \frac{Q_{95}}{Q_{MLT}} \quad \text{Eq. (4)}$$

O R_{90} e R_{95} (Equações 3 e 4) representam a capacidade natural de regularização das vazões de uma bacia hidrográfica, onde, quanto menor for os valores dessas relações, menor é a capacidade dessa bacia de regularizar sua vazão no período seco.

3.5 Precipitação média na bacia

Para obtenção da precipitação média de cada sub-bacia foram utilizadas as estações pluviométricas (Tabela 2) distribuídas em toda extensão e ao redor da BHRD. Com dados de precipitação anual para cada estação, foi realizada a precipitação média de cada pluviômetro entre os anos 1980 - 2017. Com os pontos de coordenadas de cada estação pluviométrica definidos, foi realizado através do método “Create Thiessen Polygons”, na aba “Analysis Tools”, “Proximity” a criação dos Polígonos de Thiessen (Figura 4). Esse método é indicado quando não existe distribuições uniformes dos pluviômetros dentro da bacia. A metodologia de Thiessen atribui um fator peso aos totais precipitados em cada pluviômetro e esse peso é proporcional a área de influência de cada ponto pluviométrico. Com os polígonos de Thiessen definidos, foi estabelecida a área de influência dos pluviômetros em cada sub-bacias. Assim, com as influências determinadas, calculou-se a precipitação média anual de cada sub-bacia.

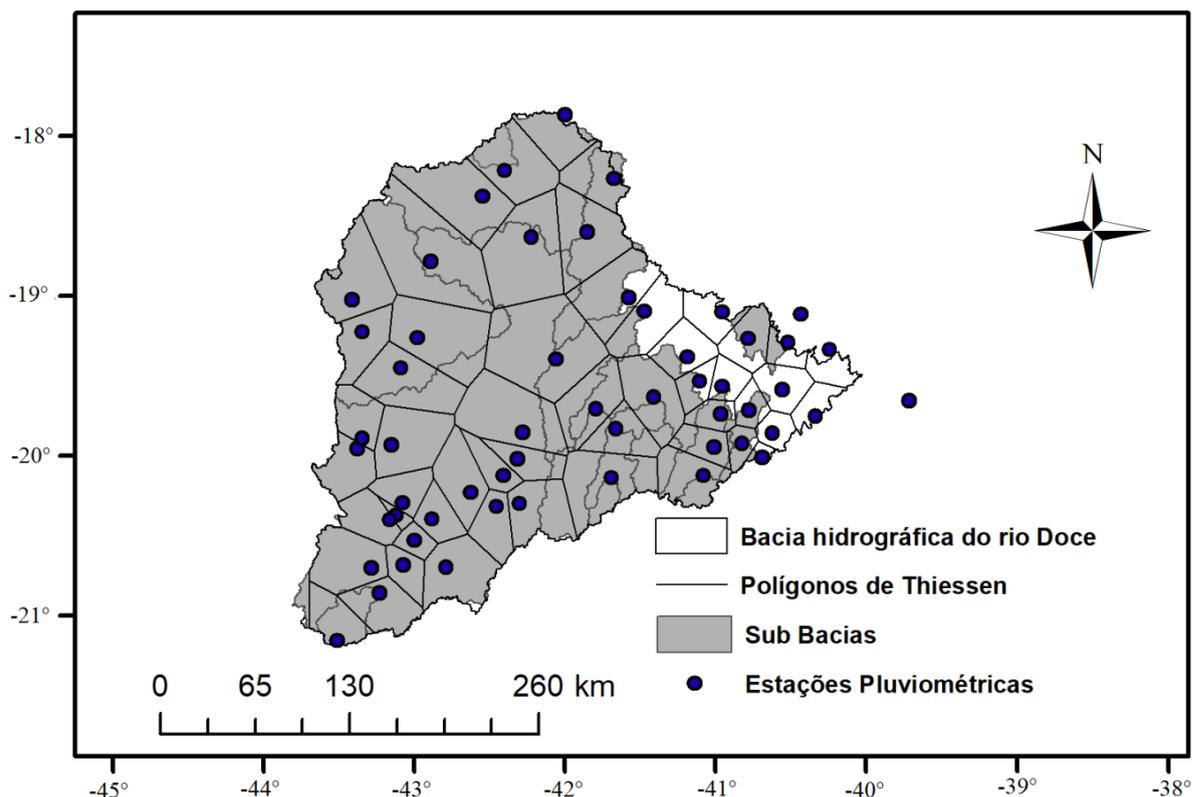


Figura 4 - Polígonos de Thiessen na Bacia hidrográfica do rio Doce e suas sub-bacias.

Fonte: Do autor (2022).

3.6 Regionalização de vazões

A definição de regiões hidrológicamente homogêneas é premissa para o desenvolvimento da regionalização de vazão. Os métodos de clusterização k-means e Partitioning Around Medoids (PAM) foram aplicados para definir as regiões homogêneas utilizando as variáveis: área, perímetro, precipitação média da bacia, Q90%/Qmlt, Q95%/Qmlt, latitude e longitude. Duas medidas de dissimilaridade foram consideradas para formação dos clusters e definição das regiões hidrológicamente homogêneas: distância Euclidiana (Equação 5) e distância Manhattan (Equação 6).

$$d(x,y)_{\text{Euclidiana}} = \sqrt{\sum_{j=1}^n (x_j - y_j)^2} \quad \text{Eq. (5)}$$

$$d(x,y)_{\text{Manhattan}} = \sum_{j=1}^n |x_j - y_j| \quad \text{Eq. (6)}$$

Onde: d é o número de atributos de cada objeto e x_j e y_j são as variáveis que representam dois objetos do mesmo conjunto de dados com d atributos cada.

Dentro de cada região homogênea a regionalização da $Q_{7,10}$ foi realizada a partir do ajuste da função linear relacionando a $Q_{7,10}$ com a área da bacia. Um modelo global também foi ajustado de acordo com a metodologia proposta afim de comparação. Neste caso, todas as bacias foram usadas, sem considerar as regiões hidrologicamente homogêneas.

3.7 Validação e análise dos clusters

Para avaliar o desempenho dos modelos lineares de regionalização da $Q_{7,10}$ aplicou-se o índice de confiança (c) (CAMARGO; SENTELHAS, 1997) (Equação 8). Este índice é o resultado do produto entre o coeficiente de correlação linear (r_{correl}) e o coeficiente de precisão (d). O coeficiente de precisão (Equação 7) (WILLMOTT; CKLESON; DAVIS, 1985):

$$d = 1 - \left[\frac{\sum(Q_{est} - Q_{obs})^2}{\sum(|Q_{est} - Q_{med}| + |Q_{obs} - Q_{med}|)^2} \right] \quad \text{Eq. (7)}$$

Onde Q_{est} e Q_{obs} são os valores estimados e observados de $Q_{7,10}$, respectivamente; e Q_{med} é o valor médio de $Q_{7,10}$ observado.

$$c = r_{correl} * d \quad \text{Eq. (8)}$$

Camargo e Sentelhas (1997) propuseram uma classificação para os modelos lineares de acordo com os valores do índice de confiança (c). A classificação é definida como ótima, muito bom, bom, média, tolerável, ruim e terrível com valores de (c) sendo, respectivamente, ($c > 0,85$), ($0,76 \leq c \leq 0,85$), ($0,66 \leq c \leq 0,75$), ($0,61 \leq c \leq 0,65$), ($0,51 \leq c \leq 0,60$), ($0,41 \leq c \leq 0,50$) e ($c \leq 0,40$).

A qualidade de um modelo também reside nas incertezas associadas às predições (SOUZA et al., 2021). Neste sentido, a técnica de bootstrap foi aplicada no intuito de avaliar as incertezas associadas aos modelos ajustados para as regiões hidrologicamente homogêneas e considerando todas as bacias (global). Essa técnica reside em amostrar o conjunto de dados (com substituição) para criar séries históricas sintéticas do mesmo tamanho das originais (SOUZA et al., 2021). Dessa forma foram geradas 500 séries históricas sintéticas para cada conjunto de dados para avaliar a incerteza dos modelos ajustados.

3.8 Validação cruzada

Além de avaliar a qualidade dos ajustados, é imprescindível a validação dos modelos de regionalização. Segundo Vezza et al. (2010), os métodos para estimar em locais não medidos não pode ser totalmente avaliado por estatísticas de adequação, como por exemplo o coeficiente de adequação R^2 . Dessa forma, a técnica de validação cruzada Leave-One-Out foi utilizada. Essa técnica separa os dados em n “folds” (i.e. pacotes) de igual tamanho (OLIVEIRA et al., 2021) e procede da seguinte forma: (i) seleciona $n-1$ folds para ajustar o modelo de regionalização; (ii) a qualidade de predição deste modelo é testada no fold remanescente, o qual não foi utilizado na calibração. O processo se repete n vezes até que todas observações sejam utilizadas para validação. A raiz quadrada do erro médio (RMSE) foi então aplicada para avaliar a capacidade preditiva dos modelos de regionalização:

$$RMSE = \frac{1}{N} \sum_{i=1}^N \sqrt{(Q7_{obs_i} - Q7_{est_i})^2} \quad \text{Eq. (9)}$$

Onde N é o número de bacias hidrográficas; e i refere-se a cada bacia utilizada no ajuste do modelo.

4 RESULTADOS E DISCUSSÃO

4.1 Análise de tendência

Primeiramente, o teste foi aplicado para todas as 38 séries completas (1980 – 2017). Como resultado apenas 2 séries se mostraram estacionárias ($p\text{-value} < 0,05$). Devido ao grande número de séries que não foram aprovadas no teste de Mann-Kendall, foi analisado graficamente o comportamento desses dados. Concluiu-se que os últimos cinco anos (2013 – 2017) estavam causando esse grande número de séries não estacionárias. Segundo Marengo et al. (2015) a crise hídrica no sudeste do Brasil nos anos de 2013 e 2014 foi ocasionada por baixos índices pluviométricos durante os verões, alta demanda de água e falta de gerenciamento dos recursos hídricos. Assim, a crise hídrica na região sudeste do Brasil pode ter afetado a estacionariedade das séries. Dessa forma, procedeu-se com a remoção destes anos antes de prosseguir com a regionalização das vazões.

Após a remoção dos últimos cinco anos de dados das séries históricas, o teste foi refeito e 26 séries fluviométricas se mostraram estacionárias, provando que os últimos anos eram os responsáveis pela não estacionariedade. Cada série completa continha 37 anos de dados diários (13.500 dados pluviométricos e 13.500 dados fluviométricos) e após a remoção dos últimos

anos (2013 – 2017) ela passou a obter 11.680 dados pluviométricos e 11.680 dados fluviométricos (32 anos de dados diários).

4.2 Distribuições Teóricas

A partir do resultado do teste de Mann Kendall, as nove FDPs, foram testadas para cada estação, e através do teste de Anderson-Darling a 5%, seis FDPs melhor se ajustaram às 26 séries históricas: Logística - 15 subbacias, LN 3P - 3 subbacias, Pareto - 2 subbacias, Weibull - 2 subbacias, Pearson - 2 subbacias e GEV – 2 subbacias.

Segundo Amorim et al (2020), FDPs como Logística, GEV, LN 3P apresentam bons resultados tanto para dados de séries de vazões máximas quanto para vazões mínimas. Ainda de acordo com os autores a FDP Weibull apresenta bons resultados para séries de Q_7 . Assim, as FDPs que se ajustaram as séries históricas podem ser recomendadas para estimar a vazão da bacia do rio Doce.

4.3 Clusters

Com os valores de $Q_{7,10}$ definidos, as 26 sub-bacias foram divididas em clusters para uma melhor compreensão de regiões semelhantes e, conseqüentemente, realizar a regionalização. Um modelo global também foi ajustado afim de comparação com os resultados gerados pelos clusters.

Primeiramente, as regiões hidrologicamente homogêneas foram geradas utilizando as metodologias k-means e PAM, considerando as distâncias de dissimilaridade Manhattan e Euclidiana. Após o resultado gerado, a sub-bacia Tumiritinga (56920000) foi retirada devido a sua extensa área, destoando das demais, o que impossibilitou a formação dos clusters. Com a retirada de Tumiritinga, as metodologias foram novamente aplicadas considerando as 25 subbacias restantes. Dentre as quatro possibilidades de formação de clusters, os métodos k-means e PAM, com as dissimilaridades Manhattan e Euclidiana, respectivamente, apresentaram a melhor separação das sub-bacias nas regiões homogêneas (gerando 2 regiões cada um) (Tabela 3).

Tabela 3 – Divisão das regiões semelhantes pelo método de clusters. Fonte: Do autor (2022).

k-means Manhattan cluster 1	k-means Manhattan cluster 2	PAM Euclidiana cluster 1	PAM Euclidiana cluster 2
Acaia	Afonso Cláudio	Acaia	Afonso Cláudio
Braz Pires	Assaraí	Braz Pires	Assaraí
Cachoeira dos óculos	Baixo Guandu	Cachoeira dos Óculos	Baixo Guandu
Carrapato (Brumal)	Barra do Cuieté	Carrapato (Brumal)	Barra do Cuieté
Fazenda Ocidente	Dores de Manhumirim	Fazenda Ocidente	Dores de Manhumirim
Fazenda Paraíso	Ferros	Fazenda Paraíso	Ferros
Fazenda Varginha	Ipanema	Fazenda Varginha	Ipanema
Piranga	Mutum	Piranga	Mutum
Ponte Nova Jusante	Pingo Água	Ponte Nova Jusante	Pingo Água
Porto Firme	Raul Soares Montante	Porto Firme	Raul Soares Montante
Rio Casca	Santo Antônio do Manhuaçu	São Miguel do Anta	Rio Casca
São Miguel do Anta	São Sebastião da Encruzilhada	Seriquite	Santo Antônio do Manhuaçu
Seriquite			São Sebastião da Encruzilhada

De acordo com Hosking e Wallis (1997), um cluster não pode ser muito pequeno em número de sub-bacias a ponto de ter baixa precisão no ajuste dos modelos de regressão, e nem grande demais, correndo o risco de incorrer em heterogeneidade. Para estes autores, um cluster é considerado muito grande quando existem mais de 20 regiões e é considerado pequeno quando contém menos de sete regiões. Desse modo corroborando com os autores, cada metodologia de cluster (k-means Manhattan e PAM Euclidiana) desse estudo geraram 2 clusters (Figuras 5 e 6) com o número de regiões compatíveis ao proposto por Hosking e Wallis (1997).

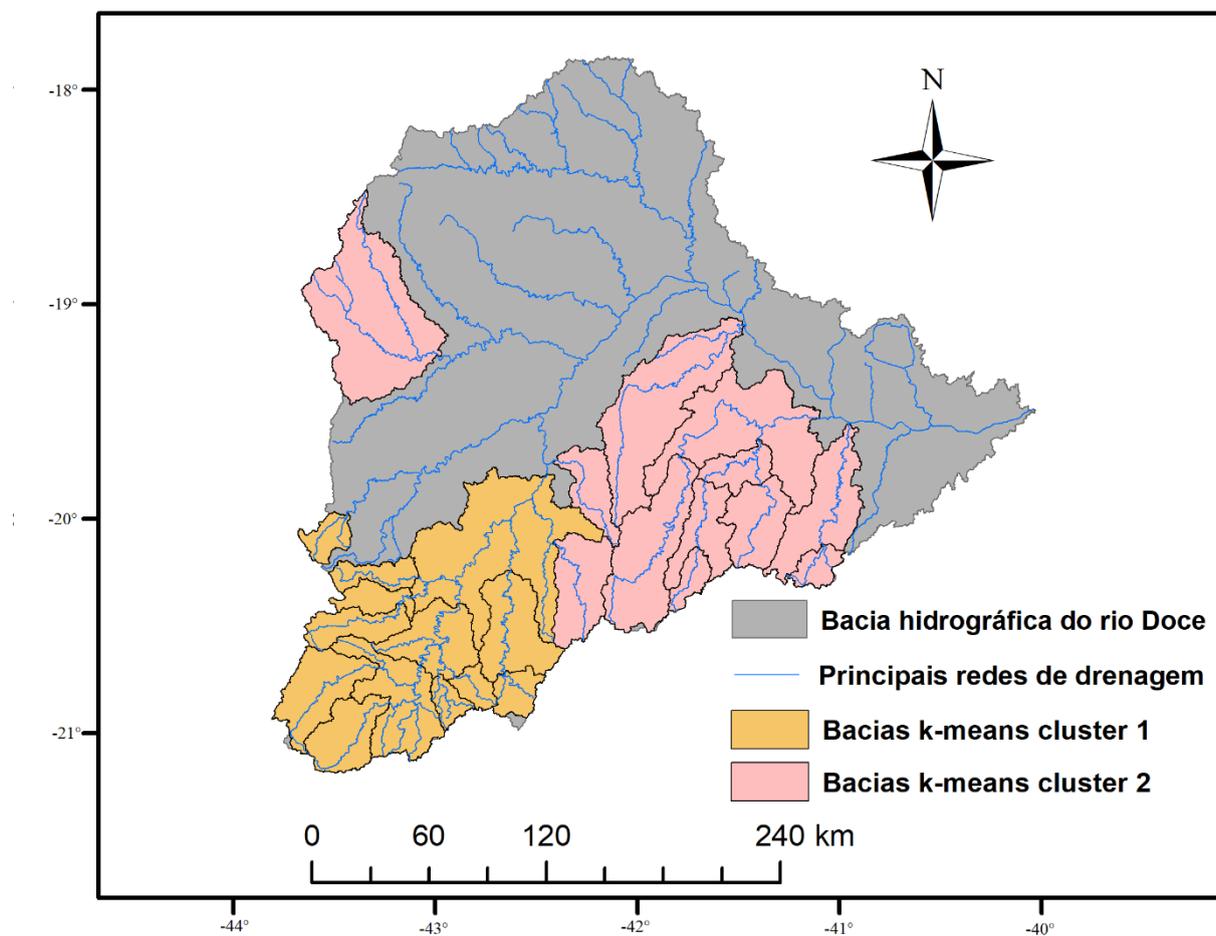


Figura 5 - Regiões hidrologicamente homogêneas formadas a partir da metodologia k-means com a medida de dissimilaridade Manhattan. Fonte: Do autor (2022).

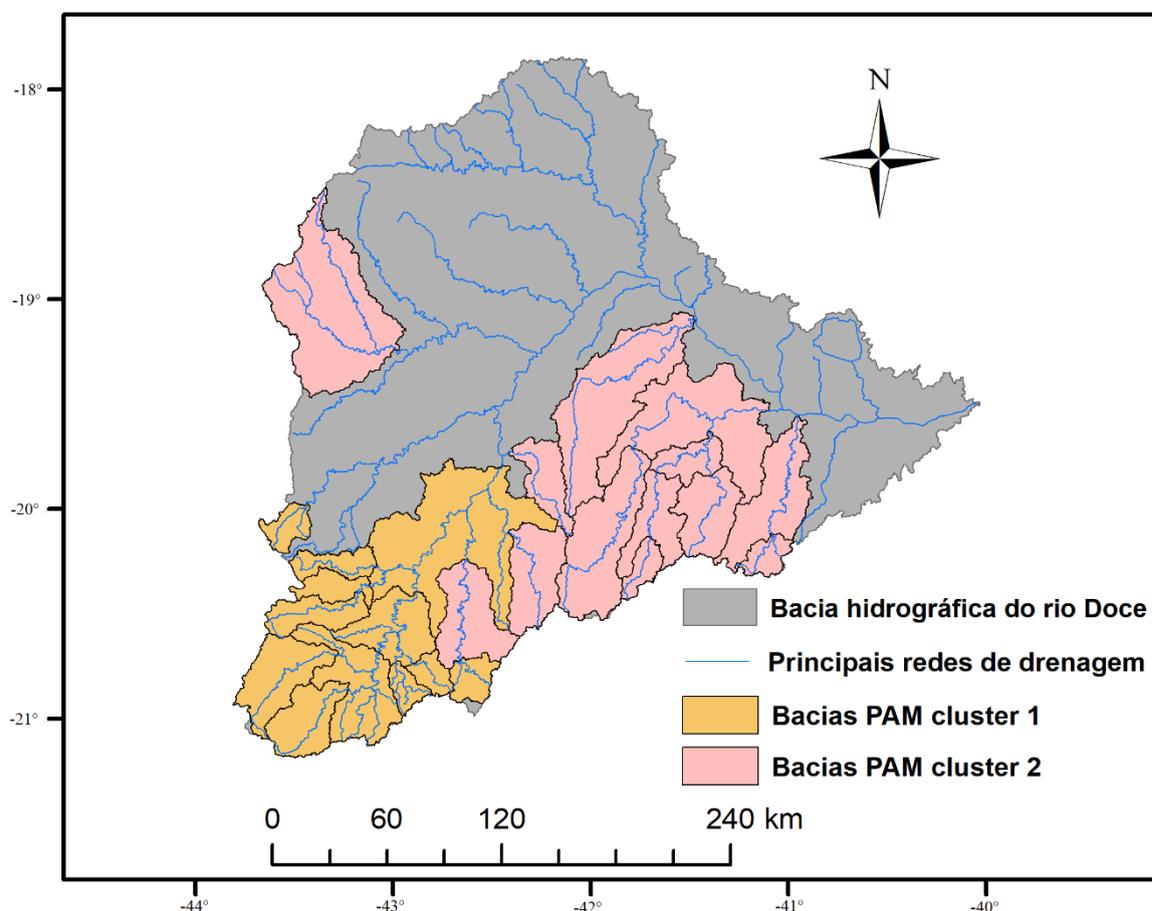


Figura 6 - Regiões hidrologicamente homogêneas formadas a partir da metodologia PAM com a medida de dissimilaridade Euclidiana. Fonte: Do autor (2022).

Para ambos os cluster (Figuras 5 e 6) é nitidamente observado que ao norte da BHRD não existem sub-bacias hidrologicamente semelhantes. Das 12 bacias que não passaram no teste de estacionariedade, cinco delas estão localizadas mais ao norte da BHRD. Cassalho et al. (2019) também observou dificuldades na formação de regiões hidrologicamente homogêneas para o extremo sul do estado do Rio Grande do Sul quando a metodologia PAM foi considerada. Por outro lado, estes autores obtiveram resultados satisfatórios com a metodologia de k-means, uma vez que a abrangência espacial abordou praticamente todo o estado. Para incluir a região norte da BHRD na regionalização das vazões, as séries históricas não estacionárias devem ser reavaliadas no intuito de eliminar a parcela que esteja causando a não estacionariedade, i.e., tornar as séries mais curtas para garantir a estacionariedade.

4.4 Modelos de regressão e regionalização de $Q_{7,10}$

Segundo Naghettini e Pinto (2007) as regressões mais utilizadas são dos modelos potencial, exponencial e logarítmico para quaisquer vazões. Beskow et al. (2016) também exaltou a aplicação de modelos potenciais para relacionar as vazões mínimas às características das sub-bacias. No entanto, devido à baixa diferença no coeficiente de determinação entre os

modelos linear e potencial, e a modo de usar um modelo mais simplificado, o modelo linear foi utilizado no presente trabalho. (Figura 7).

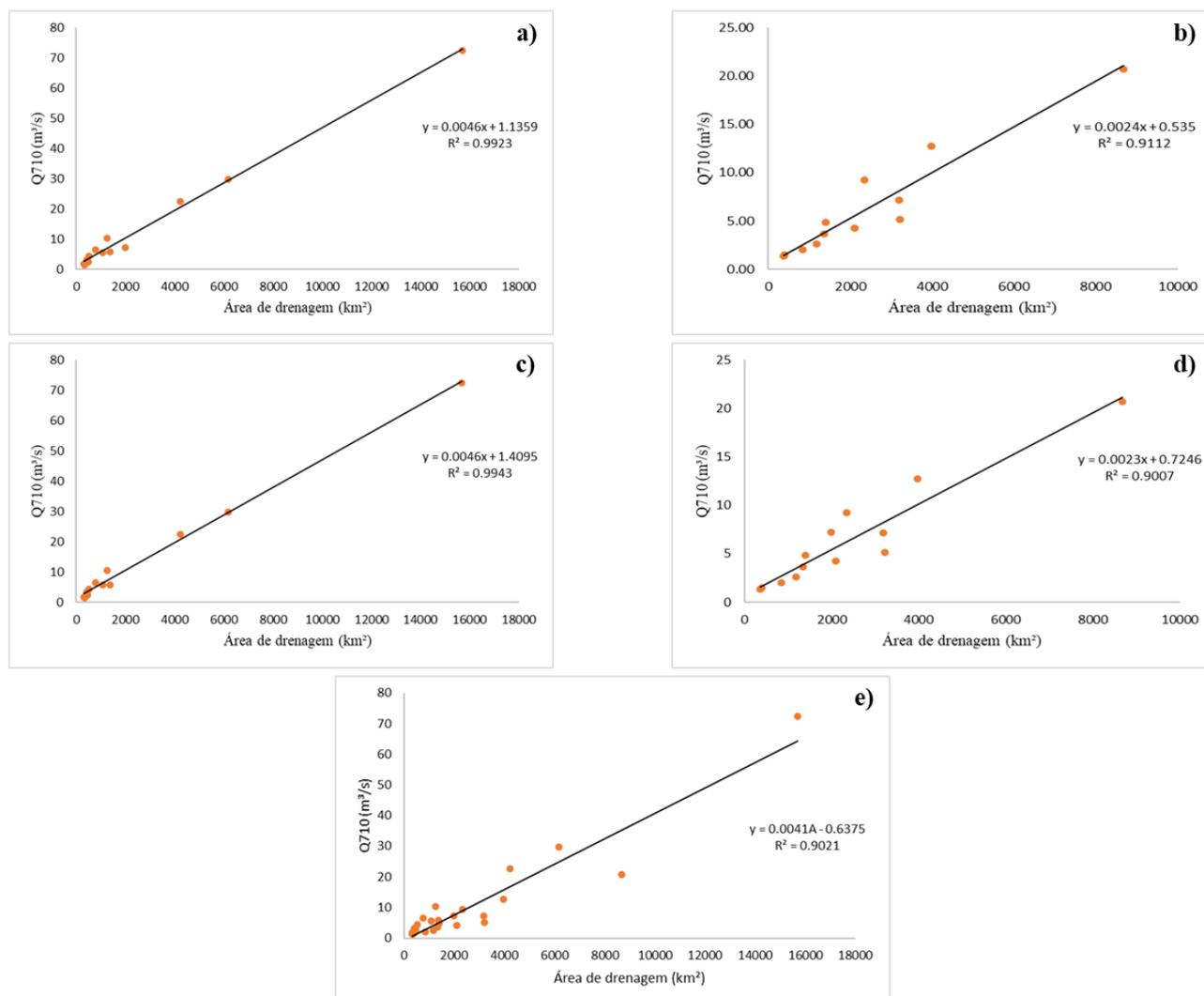


Figura 7 - Regressões lineares e coeficientes de determinação (a) k-means Manhattan cluster 1, (b) k-means Manhattan cluster 2, (c) PAM Euclides cluster 1, (d) PAM Euclides cluster 2, (e) modelo global. Fonte: Do autor (2022).

A área de drenagem tem sido considerada o fator principal na regionalização de vazões mínimas (LISBOA et al., 2008; BESKOW et al., 2016; MELATI E MARCUZZO 2016) devido à sua relação intrínseca com a recarga e a produção de vazão. Considerando a área de drenagem como variável explicativa para a distribuição de $Q_{7,10}$, as regressões apresentaram $R^2 > 0.9$ para todas as regiões hidrologicamente homogêneas, i.e., a área da bacia é responsável por mais de 90% da variabilidade da $Q_{7,10}$ dentro das regiões homogêneas. Estes resultados corroboram com outros estudos que indicam a área de drenagem como fator primordial na regionalização de vazões mínimas (LISBOA et al., 2008; BESKOW et al., 2016;

MELATI E MARCUZZO 2016).

Os resultados do índice de confiança (c) para os modelos ajustados foram: $c=0,994$ (a), $c=0,925$ (b), $c=0,995$ (c) e $c=0,920$ (d), de modo que todos os ajustes foram classificados como ótimos ($c > 0,85$). Estes resultados corroboram com os anteriores sobre a importância da área de drenagem como variável explicativa, como também observado por Beskow et al. (2016). Desse modo os modelos matemáticos para as regiões homogêneas formadas pelas duas metodologias (k-means Manhattan e PAM Euclidiana) resultaram nas equações de regionalização (Tabela 4). Apesar dos ajustes terem apresentado resultado satisfatório, a aplicação dos modelos de regionalização fica restringida entre os limites de áreas de drenagem utilizados em seu ajuste (Tabela 4).

Tabela 4 – Equações de regionalização $Q_{7,10}$ para as regiões homogêneas da BHRD e limite das áreas de drenagens para cada cluster (km²). Fonte: Do autor (2022).

	k-means Manhattan cluster 1	k-means Manhattan cluster 2	PAM Euclides cluster 1	PAM Euclides cluster 2
Equação	$Q_{7,10}=0,0046A+1,1359$	$Q_{7,10}=0,0024A+0,535$	$Q_{7,10}=0,0046A+1,4095$	$Q_{7,10}=0,0023A+0,7246$
Área (km ²)	327,23 - 15686	372,61 - 8681	327,23 - 15686	372,61 - 8681

Afim de comparação com o método de clusters, a regionalização através do modelo global, onde as 25 regiões não são separadas em regiões homogêneas gerou como resultado equação de regressão linear $y=0,0041A-0,6375$ com coeficiente de determinação $R^2=0,9021$.

O índice de confiança (c) proposto por Camargo e Sentelhas (1997) para o modelo global foi de $c=0,859$, classificado como ótimo ($c>0,85$), mas, obtendo um índice de confiança menor que os obtidos através da regionalização por clusters.

4.5 Validação cruzada e incerteza nos modelos de regressão

Como resposta da validação cruzada, a raiz quadrada do erro médio - RMSE foi calculada de modo a testar a qualidade preditiva do modelo de regionalização (Tabela 5). Os resultados indicam resultados satisfatórios para a regionalização da $Q_{7,10}$ para ambas as metodologias de formação de regiões homogêneas utilizadas. Cassalho et al. (2019) também utilizou a técnica de validação cruzada Leave-One-Out em seu estudo de modo a avaliar o

poder preditivo das equações de regressão. Beskow et al. (2016) conclui que a validação cruzada é um forte instrumento para avaliar a regionalização para as regiões homogêneas de modo a dar mais consistência no ponto de vista regional e verificar a capacidade preditiva do modelo de regionalização.

Tabela 5 – Valores de RMSE (m^3/s) para os clusters das regiões homogêneas e para o modelo global. Fonte: Do autor (2022).

	k-means Manhaattan cluster 1	k-means Manhaattan cluster 2	PAM Euclides cluster 1	PAM Euclides cluster 2	Modelo global
RMSE (m^3/s)	1,672	1,396	1,616	1,528	4

A técnica de bootstrap foi aplicada para avaliar a importância da definição das regiões hidrologicamente homogêneas nas incertezas dos modelos de regionalização. A incerteza dos modelos ajustados para as regiões hidrologicamente homogêneas foi consideravelmente menor ao modelo global como pode ser observada pela amplitude do intervalo de confiança na (Figura 8). As incertezas são aumentadas pela presença de poucas sub-bacias com maiores áreas de drenagem, i.e., os maiores erros dos valores de $Q_{7,10}$ regionalizados tendem a ocorrer nas áreas de drenagem maiores que 5000 km^2 . Dessa forma, recomenda-se a determinação da $Q_{7,10}$ em regiões desprovidas de monitoramento que façam parte da mesma região homogênea e, ainda, que os valores para bacias $> 5000 \text{ km}^2$ sejam adotados com cautela.

Apesar do modelo global ter obtido um bom ajuste com um bom índice de confiança, após a validação cruzada o modelo apresentou uma incerteza muito alta, comparada com a metodologia dos clusters, apresentando como validação $RMSE = 4 \text{ m}^3/s$. Essa incerteza para sub bacias com baixas vazões não apresenta uma boa confiabilidade no modelo (Figura 8).

Laaha e Blöschl (2006), em seu estudo, ao comparar 4 métodos de regionalização por agrupamentos com um método global sem agrupamento, concluíram que o estudo tem um melhor resultado nas metodologias onde ocorreu a formação de clusters com regiões homogêneas.

Veza et al. (2010) e Beskow et al. (2016) corroboram com os resultados, mostrando que, dividir a bacia em sub regiões homogêneas, melhora o resultado da regionalização por meio de regressões regionais.

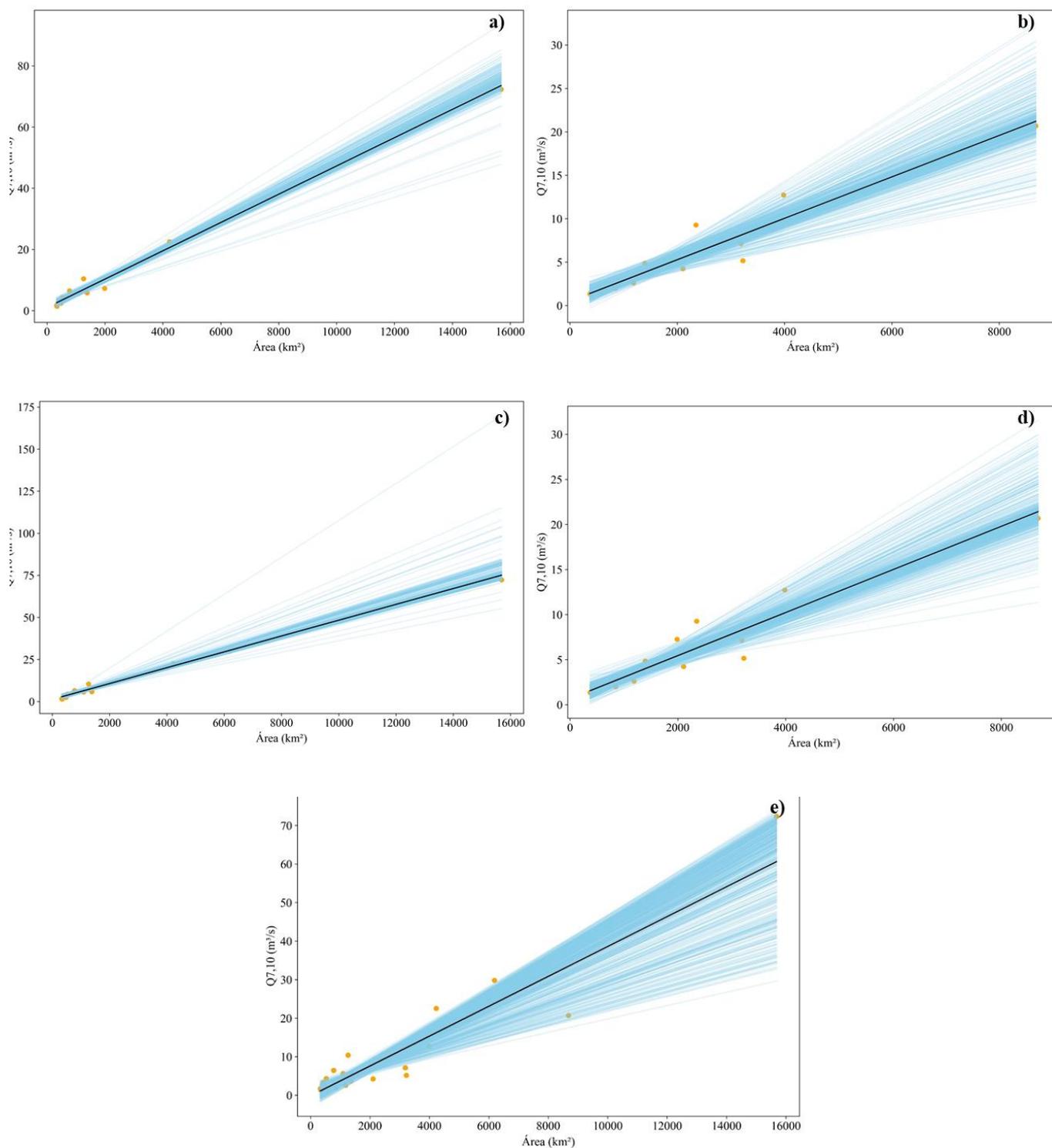


Figura 8 - Ajuste e incerteza na regionalização da vazão de referência (Q7,10) para as regiões homogêneas (a) k-means Manhattan cluster 1, (b) k-means Manhattan cluster 2, (c) PAM Euclides cluster 1 e (d) PAM Euclides cluster 2, (e) modelo global. As linhas azuis indicam a incerteza dos modelos de regionalização. Fonte: Do autor (2022).

5 CONCLUSÃO

A regionalização pela metodologia de clusterização, por utilizar séries históricas conhecidas e dividir a bacia em regiões homogêneas, garante uma boa eficiência em prever vazões uma vez que utiliza dados já monitorados para estimar a vazão em uma região semelhante à aquela onde as informações foram extraídas.

A regionalização de vazões utilizando o modelo global, onde não divide a bacia em sub regiões, em comparação com a metodologia de clusters, gerou uma incerteza muito alta, na média 4 m³/s, um valor que para vazões baixas modifica significativamente a precisão da regionalização.

A utilização de um modelo global para prever vazões mínimas Q_{7,10}, provou ser um método que produz uma maior incerteza devido a não trabalhar com regiões hidrologicamente homogêneas mas pra áreas pequenas ela consegue ser mais assertiva que em grandes áreas.

A utilização de modelos digitais de elevação – MDE e de programas como ArcGis no processo de obtenção de dados para a regionalização é de grande ajuda, uma vez que algumas variáveis utilizadas nos modelos nem sempre estão disponíveis para acesso.

As equações de regionalização de vazões obtidas pelo modelo de clusters por possuir um alto índice de confiança (c), tem como única variável explicativa a sua área de drenagem, obtendo uma equação linear simples, mas gerando bons resultados.

REFERÊNCIAS

- ALMAGRO, A. et al. **Projected climate change impacts in rainfall erosivity over Brazil**. Scientific Reports, v.7, n.1, p.1-12, 2017. doi: 10.1038/s41598-017-08298-y.
- AMORIM, J. S. et al. **Streamflow regionalization for the Mortes River Basin upstream from the Funil Hydropower Plant, MG**. Revista Ambiente e Agua, v. 15, p. 1-11, 2020.
- BARBOSA, S. E. da S. et al. **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**. Engenharia Sanitária e Ambiental, v. 10, n. 1, p. 64-71, 2005.
- BARRELLA, W. et al. As relações entre as matas ciliares, os rios e os peixes. In: RODRIGUES, R. R; FILHO, H. F. L. (Org.). **Matas ciliares: conservação e recuperação**, v. 2, p. 187-207, 2001.
- BESKOW, Samuel et al. **Artificial intelligence techniques coupled with seasonality measures for hydrological regionalization of Q90 under Brazilian conditions**. Journal of Hydrology, v. 541, p. 1406-1419, 2016.
- BRASIL. **Lei nº 9.433**, de 8 de janeiro de 1997. Institui a Política Nacional de Recursos Hídricos. Diário Oficial da União, Brasília, DF, 9 jan. 1997. Disponível em: http://www.planalto.gov.br/ccivil_03/LEIS/L9433.htm. Acesso em: 17 set. 2021.
- CALEGARIO, A. T. **Identificação de regiões hidrologicamente homogêneas e comportamento das vazões mínima e média regionalizadas**. 2014. 80 p. Dissertação (Mestrado em Recursos Hídricos) – Universidade Federal de Viçosa, Viçosa, 2014.
- CAMARGO, A. P.; SENTELHAS, P. C. **Avaliação do desempenho de diferentes métodos de estimativas da evapotranspiração potencial no Estado de São Paulo, Brasil**. Revista Brasileira Agrometeorologia, v. 1, n.1, p. 89–97, 1997.
- CARDOSO, F. A. C. et al. **Modelo digital de elevação hidrologicamente consistente para a bacia do rio Doce: elaboração e análise**. Revista Brasileira de Geografia Física, v. 9, n. 6, p. 1978-1989, 2016.
- CASSALHO, F. et al. **Artificial intelligence for identifying hydrologically homogeneous regions: A state-of-the-art regional flood frequency analysis**. Hydrological Processes, v.33, n.7, p. 1101-1116, 2019. doi: 10.1002/hyp.13388.
- COELHO, A. L. N. **Bacia hidrográfica do Rio Doce (MG/ES): uma análise socioambiental integrada**. Geografares, 2009. doi: 10.7147/GEO7.156.
- COMITÊ DA BACIA HIDROGRÁFICA DO RIO DOCE (CBH-DOCE) Disponível em: <https://www.cbhdoce.org.br/institucional/a-bacia>, 2021.
- COMITÊ DA BACIA HIDROGRÁFICA DO RIO DOCE (CBH-DOCE) Disponível em: http://www.cbhdoce.org.br/wp-content/uploads/2016/12/PIRH_Doce_Volume_1.pdf. Acesso em: 8 set. 2021.
- DA SILVA, D. D.; DE AZEVEDO MARQUES, F.; LEMOS, A. F. **Avaliação de metodologias de regionalização de vazões mínimas de referência para bacia do Rio São Francisco**. Revista Engenharia na Agricultura-Reveng, v. 17, n. 5, p. 392-403, 2009.

DINIZ, L. S. **Regionalização de parâmetros de modelos chuva-vazão usando redes neurais**. 2008. 230 p. Tese de doutorado. Instituto de Pesquisas Hidráulicas da UFRGS, Porto Alegre – RS, 2008.

FERREIRA, R.; MOURA, M. C.; CASTRO, F. **Caracterização morfométrica da sub-bacia do Ribeirão Panquinhas, ES**. Enciclopédia Biosfera, v. 8, n. 15, 2012.

HAN, J.; PEI, J.; KAMBER, M. **Data mining: concepts and techniques**. 3. ed. Elsevier, 2012.

HARRIS, C. W. **Characteristics of two measures of profile similarity**. Psychometrika, v. 20, n.4, p.289-297, 1955. doi: 10.1007/BF02289035.

HOSKING, J., & WALLIS, J. Identification of homogeneous regions. In **Regional Frequency Analysis: An Approach Based on L-Moments**. Cambridge: Cambridge University Press, p. 54-72, 1997. doi:10.1017/CBO9780511529443.006.

IGAM. INSTITUTO MINEIRO DE GESTÃO DAS ÁGUAS. **Portaria nº 48**, de 4 de outubro de 2019. Estabelece normas suplementares para a regularização dos recursos hídricos de domínio do Estado de Minas Gerais e dá outras providências. 2019. Disponível em: <http://www.siam.mg.gov.br/sla/download.pdf?idNorma=49719>. Acesso em: 20 set. 2021.

Instituto Brasileiro de Geografia e Estatística – IBGE **Disponível em:** <https://www.ibge.gov.br/geociencias/downloads-geociencias.html>, 2021.

JAIN, A. K. **Data clustering: 50 years beyond K-means**, Pattern Recognition Letters, v.31, n.8, p. 651-666, 2010. doi: 10.1016/j.patrec.2009.09.011.

JAIN, A. K.; MURTY, M. N.; FLYNN, P. J. **Data clustering: a review**, ACM Computing Surveys, v. 31, p. 264-323, 1999. Doi: 10.1145/331499.331504.

KAUFMAN, L.; ROUSSEEUW, P. J. **Partitioning around medoids (program pam). Finding groups in data: an introduction to cluster analysis**, Wiley-Interscience, v.344, p.68-125, 1990.

KENDALL, M. G. **Rank Correlation Methods**. 4. ed. Londres: Charles Griffin, 1975.

KUGLER, M.; JÚNIOR, J. T.; LOPES, H. S. **Desenvolvimento de uma Rede Neural LVQ em Linguagem VHDL para Aplicações em Tempo-Real**. In: Proceedings of the VI Brazilian Conference on Neural Networks-VI Congresso Brasileiro de Redes Neurais, p.103–108, 2003.

LAAHA, G.; BLÖSCHL, G. **A comparison of low flow regionalisation methods - catchment grouping**. Journal of Hydrology, v. 323, n.1-4, p.193-214, 2006.

Laboratório de Processamento de Imagens e GeoProcessamento – LAPIG. **Disponível em:** <https://maps.lapig.iesa.ufg.br>, 2021.

LISBOA, L. et al. **Estimativa e regionalização das vazões mínimas e média na bacia do Rio Paracatu**. Engenharia na Agricultura, v.16, n.4, p.471-479, 2008.

LOUZADA, F. L. R. de O. **Balanço hídrico climatológico com dados estimados pelo**

Satélite Tropical Rainfall Measuring Mission para bacia hidrográfica do Rio Doce. 2016. 107 p. Tese (Doutorado em Produção Vegetal) – Universidade Federal do Espírito Santo, Centro de Ciências Agrárias. Espírito Santo, 2016.

MANN, H. B. **Nonparametric tests against trend.** *Econometrica Journal of the Econometric Society*, p. 245-259, 1945.

MARENGO, J. A., et al. **A seca e a crise hídrica de 2014-2015 em São Paulo.** *Revista USP* 106 (2015): 31-44.

MELATI, M. D.; MARCUZZO, F. F. N. **Regressões simples e robusta na regionalização da vazão Q95 na Bacia Hidrográfica do Taquari-Antas,** 2016. Disponível em: <http://rigeo.cprm.gov.br/jspui/handle/doc/16858>.

MELLO, C. R. de; SILVA, A. M. da. **Hidrologia: princípios e aplicações em sistemas agrícolas.** 1. ed. Lavras: Editora UFLA, v. 1, p. 455, 2013.

MENDES, L. A. **Análise dos critérios de outorga de direito de usos consuntivos dos recursos hídricos baseados em vazões mínimas e em vazões de permanência.** 2007. 187 p. Dissertação (Mestrado em Recursos Hídricos (Gerenciamento; Instrumentos)) - Universidade de São Paulo, São Paulo, 2007.

NAGHETTINI, M.; PINTO, E. J. de A. **Hidrologia estatística.** CPRM, p. 552, 2007.
NATIONAL AERONAUTICS AND SPACE ADMINISTRATION - NASA. (2016). Mission. Disponível em: <https://www2.jpl.nasa.gov/srtm/mission.htm>. Acesso em: 22 mar. 2017.

NETO, M. C. P. **Análise areal como subsidio aos estudos integrados da Bacia Hidrográfica do rio Seridó (RN/PB).** *Revista Equador*, v. 5, n. 4, p. 250-261, 2016.

OLIVEIRA, V. A. et al. **Spatiotemporal modelling of soil moisture in an Atlantic forest through machine learning algorithms.** *European Journal of Soil Science*, v.72, n.5, p.1969-1987, 2021. doi: 10.1111/ejss.13123.

PEREIRA, E. M. **Análise de conflitos pelo uso da água relacionados à oferta e à demanda: Bacia do Rio Piracicaba – MG.** 2012. 56 p. Dissertação (Mestrado em Análise e Modelagem de Sistemas Ambientais) – Universidade Federal de Minas Gerais, Instituto de Geociências, 2012.

POMPEU, R. D. M. et al. A outorga de direito de uso de recursos hídricos. In: SIMPÓSIO DE RECURSOS HÍDRICOS DO NORDESTE, 7., 2014, Natal. **Anais...** Natal: ABRH, 2014. p. 1 – 9.

PORTO, M. M. et al. **Identificação de bacias hidrográficas com características físicas similares no Estado do Ceará, Brasil.** *Revista Ciência Agronômica*, v. 35, n.1, p. 17-25, 2004.

RIBEIRO, C. B. de M.; MARQUES, F. de A.; SILVA, D. D. da. **Estimativa e regionalização de vazões mínimas de referência para a bacia do rio Doce.** *Engenharia na Agricultura*, v. 13, n. 2, p. 103-117, 2005.

SILVA, A. M. da et al. **Vazões mínimas e de referência para outorga na região do Alto Rio Grande, Minas Gerais.** *Revista Brasileira de Engenharia Agrícola e Ambiental*, v. 10, p. 374-380, 2006.

SOUZA, G. R. et al. **Regional flood frequency analysis and uncertainties: Maximum streamflow estimates in ungauged basins in the region of Lavras, MG, Brazil.** Catena, v.197, p. 104970, 2021. doi: 10.1016/j.catena.2020.104970.

TEODORO, V. L. I. et al. **O conceito de bacia hidrográfica e a importância da caracterização morfométrica para o entendimento da dinâmica ambiental local.** Revista Brasileira Multidisciplinar, v. 11, n. 1, p. 137-156, 2007.

TUCCI, C. E. M. **Hidrologia: ciência e aplicação.** 2.ed. Porto Alegre: ABRH/Editora da UFRGS, 1997.

VEZZA, P. et al. **Low flows regionalization in north-western Italy.** Water Resources Management, v. 24, n.14, p. 4049-4074, 2010.

WILLMOTT, C. J.; CKLESON, S. G.; DAVIS, R. E. **Statistics for evaluation and comparisons of models.** Journal of Geophysical Research, Ottawa, v. 90, n. C5. p. 8995-9005, 1985.

WOLFF, W. et al. Avaliação do modelo de regionalização hidrológica do estado de são paulo para o cálculo da Q7,10, Q90, Q95, QM. In: X Congresso Latinoamericano y del Caribe de Ingeniaria Agrícola e XLI Congresso Brasileiro de Engenharia Agrícola, 2012, Londrina-PR. **A Engenharia Agrícola na Evolução dos Sistemas de Produção.** Jaboticabal-SP: SBEA, 2012. v. 1.

ZUCHINI, M. H. **Aplicações de mapas auto-organizáveis em mineração de dados e recuperação de informações.** 2003. 215 p. Dissertação (Mestrado) - Universidade Estadual de Campinas, Faculdade de Engenharia Elétrica e de Computação. Campinas, 2003.