

UNIVERSIDADE FEDERAL DE MINAS GERAIS
Escola de Engenharia
Programa de Pós-Graduação em Engenharia Elétrica
Especialização em Fontes Renováveis: Geração, Operação e Integração

Lucas Bacha Pereira Horta

**DESENVOLVIMENTO DE MODELO COMPUTACIONAL PARA
TRANSPOSIÇÃO DE DADOS FLUVIOMÉTRICOS PARA APLICAÇÃO NO
CÁLCULO DA GARANTIA FÍSICA DE CENTRAIS GERADORAS
HIDRELÉTRICAS**

Belo Horizonte

2019

Lucas Bacha Pereira Horta

**DESENVOLVIMENTO DE MODELO COMPUTACIONAL PARA
TRANSPOSIÇÃO DE DADOS FLUVIOMÉTRICOS PARA APLICAÇÃO NO
CÁLCULO DA GARANTIA FÍSICA DE CENTRAIS GERADORAS
HIDRELÉTRICAS**

Versão Final

Monografia de especialização apresentada à Banca Examinadora designada pela Comissão Coordenadora do Curso de Especialização em Fontes Renováveis do Programa de Pós-Graduação em Engenharia Elétrica da Universidade Federal de Minas Gerais, como requisito parcial à obtenção do Certificado de Especialista em Fontes Renováveis.

Orientador: Prof. Dr. Carlos Barreira Martinez

Belo Horizonte

2019

H821d	<p>Horta, Lucas Bacha Pereira. Desenvolvimento de modelo computacional para transposição de dados fluviométricos para aplicação no cálculo da garantia física de centrais geradoras hidrelétricas [recurso eletrônico] / Lucas Bacha Pereira Horta. – 2019. 1 recurso online (93 f. : il., color.) : pdf.</p> <p>Orientador: Carlos Barreira Martinez.</p> <p>“Monografia apresentada ao Curso de Especialização em Fontes renováveis: geração, operação e integração da Escola de Engenharia da Universidade Federal de Minas Gerais”</p> <p>Anexos: f. 67-93. Bibliografia: f. 62-66.</p> <p>Exigências do sistema: Adobe Acrobat Reader.</p> <p>1. Engenharia elétrica. 2. Energia renovável. 3. Hidrologia. 4. Hidrometria. 5 Incerteza. I. Martinez, Carlos Barreira. II. Universidade Federal de Minas Gerais. Escola de Engenharia. III. Título.</p> <p style="text-align: right;">CDU: 621.3</p>
-------	---



ATA DA DEFESA DA MONOGRAFIA DO ALUNO LUCAS BACHA PEREIRA HORTA

Realizou-se, no dia 02 de dezembro de 2019, às 10:00 horas, Auditório do CPH, da Universidade Federal de Minas Gerais, a defesa de monografia, intitulada *Desenvolvimento de modelo computacional para transposição de dados fluviométricos para aplicação no cálculo da garantia física de Centrais Geradoras Hidrelétricas*, apresentada por LUCAS BACHA PEREIRA HORTA, número de registro 2018704880, graduado no curso de ENGENHARIA DE ENERGIA, como requisito parcial para a obtenção do certificado de Especialista em FONTES RENOVÁVEIS - GERAÇÃO, OPERAÇÃO E INTEGRAÇÃO, à seguinte Comissão Examinadora: Prof(a). Carlos Barreira Martinez - Orientador (Universidade Federal de Itajubá), Prof(a). Aloysio Portugal Maia Saliba (UFMG).

A Comissão considerou a monografia:

Aprovada

Reprovada

Finalizados os trabalhos, lavrei a presente ata que, lida e aprovada, vai assinada por mim e pelos membros da Comissão.

Belo Horizonte, 02 de dezembro de 2019.

Prof(a). Carlos Barreira Martinez (Doutor)

Prof(a). Aloysio Portugal Maia Saliba (Doutor)

RESUMO

O presente trabalho retrata o estado da arte de metodologias para regionalização de dados fluviométricos em bacias sem monitoramento hidrológico. Ademais, discorre sobre as incertezas associadas aos métodos de coleta e transporte de dados propondo a implementação e teste de um modelo simplificado para a regionalização de dados fluviométricos, com o objetivo de avaliar as incertezas alcançadas. O modelo apresentado foi desenvolvido na plataforma MATLAB® e tem como proposta a automação completa do processo de regionalização de dados fluviométricos. A fim de verificar o nível de incerteza dos resultados obtidos com o software, foi conduzida a análise de 271 postos fluviométricos distribuídos pelo Brasil, concluindo que o modelo é aplicável a 40% dos locais analisados, apresentando erros abaixo de 30%. No entanto, para os 60% dos locais restantes o modelo não foi capaz de explicar corretamente as descargas líquidas locais, apresentando erros superiores a 30% e, portanto, recomenda-se a introdução de um número maior de informações, como dados de entrada, como objetivo de minimizar os erros.

Palavras-chave: Regionalização de dados fluviométricos. Hidrologia. Hidrometria. Incertezas. Bacias sem medições.

ABSTRACT

The present study displays the state-of-the-art methodologies for estimation of fluviometric data in ungauged catchments besides discussing the uncertainties associated with data collection and streamflow estimation methods. In order to evaluate the uncertainties achieved with this type of methods, it was proposed the implementation and test of a simplified model for streamflow data estimation in ungauged catchments. The proposed model was developed on MATLAB® platform and seeks for complete automation of the process. In order to verify the uncertainty level of the results obtained with the software, an analysis of 271 gauges distributed around the Brazilian territory was conducted, concluding that the model is applicable to 40% of the sites, presenting errors below 30%. However, for the remaining 60% of the sites the model was unable to explain the discharges with great accuracy, showing errors greater than 30%. Therefore, it is recommended an increase in the number of variables as input data, seeking to minimize errors.

Keywords: Streamflow Regionalization. Hydrology. Hydrometry. Uncertainty. Ungauged catchments.

LISTA DE ILUSTRAÇÕES

Figura 1 – Potência instalada por fonte da matriz elétrica brasileira. (ANEEL, 2019a)	12
Figura 2 – Estações fluviométricas inventariadas pela ANA, comparativo: nov/1949 a atualidade. (ANA, 2019).....	13
Figura 3 – Exemplo de período crítico hidrológico. (Fonte: o autor).....	15
Figura 4 - Método de medição e integração da distribuição de velocidade. (Fonte: o autor)	18
Figura 5 - método de medição de vazão acústico. (Fonte: o autor).....	20
Figura 6 - Exemplo teórico de curva-chave. (Fonte: o autor)	23
Figura 7 - Fluxograma simplificado da consistência de dados fluviométricos. (fonte: o autor)	25
Figura 8 - Fluxograma simplificado da regionalização de dados hidrológicos. (Fonte: o autor)	27
Figura 9 – Representação esquemática de um dendrograma. (Fonte: o autor)	30
Figura 10 – Exemplo genérico de uma das etapas do algoritmo de “ <i>k-means</i> ” para agrupamento de dados. (Fonte: o autor).....	31
Figura 11 - Diferentes casos abordados pela metodologia de interpolação em GIS. Pontos em preto: postos fluviométricos; Pontos em branco: pontos de interesse. (Fonte: o autor).	36
Figura 12 - Fluxograma do algoritmo proposto. (Fonte: o autor)	40
Figura 13 - Definido o local de estudo (x), o software identifica a sub-bacia a qual o ponto pertence (no exemplo da figura a sub-bacia 39 – “Paraíba e outros”). (Fonte: o autor).....	41
Figura 14 - Filtro de dados através do ajuste da curva-chave. (Fonte: o autor)	43
Figura 15 – Exemplos de correlações entre dados de estações fluviométricas. (Fonte: o autor)	44
Figura 16 – Curva estimada de rendimento específico para a bacia na qual o local de estudo está inserido. Círculos: Dados das estações; Quadrado: previsão para um local genérico na bacia. (Fonte: o autor)	46
Figura 17 – Distribuição acumulada dos erros: visão geral (Fonte: o autor)	49

Figura 18 - Resultados de oito estações distintas: (i) melhor resultado da amostra; (ii) pior resultado da amostra; (iii) erro (RMSE) de 10%; (iv) erro de 20%; (v) erro de 37%; (vi) erro de 50%; (vii) erro de 81%; e (viii) erro de 569%. (Fonte: o autor)	52
Figura 19 - Distribuição acumulada dos R^2 : visão geral (Fonte: o autor).....	53
Figura 20 - RMSE vs n° de estações c/ $R^2 > 80\%$. (Fonte: o autor).....	54
Figura 21 - RMSE vs n° de anos c/ dados disponíveis na estação de referência. (Fonte: o autor)	55
Figura 22 - RMSE versus diferença entre as áreas de drenagem. (Fonte: o autor)	57
Figura 23 – RMSE versus distância da estação de referência. (Fonte: o autor).....	58
Figura 24 - Comportamento do RMSE para diferentes faixas de vazão. (Fonte: o autor)...	59

LISTA DE TABELAS

Tabela 1 - Incertezas relacionadas ao método de medição e integração da distribuição de velocidade (SAUER; MEYER, 1992).	18
Tabela 2 – Sistema de pontuação para classificação das estações e escolha da estação de referência. (Fonte: o autor).....	45

LISTA DE ABREVIATURAS OU SIGLAS

AD	Área de drenagem
ADCP	Acoustic Doppler Current Profiling
AHE	Aproveitamento Hidrelétrico
ANA	Agência Nacional de Águas
ANEEL	Agência Nacional de Energia Elétrica
CCEE	Câmara de Comercialização de Energia Elétrica
CGH	Central Geradora Hidrelétrica de capacidade reduzida
EIH	Estudos de Inventário Hidrelétricos
EPE	Empresa de Pesquisa Energética
EUA	Estados Unidos da América
FC	Fator de Capacidade
GCPS	Grupo Coordenador do Planejamento dos Sistemas Elétricos
GF	Garantia Física
GFR	Geometria Fractal
GH	Geometria Hidráulica
HUIG	Hidrograma Unitário Instantâneo Geomorfológico
kW	Quilowatt
MME	Ministério de Minas e Energia
MRE	Mecanismo de Realocação de Energia
MSUI	Modelo de Simulação a Usinas Individualizadas
MW	Megawatt
ONS	Operador Nacional do Sistema
PCH	Pequena Central Hidrelétrica
PNE	Plano Nacional de Energia
Q50	Vazão com permanência em 50% do tempo
Q7,10	Vazão mínima da média móvel de sete dias de duração, anual, com período de recorrência de dez anos
Q80	Vazão com permanência em 80% do tempo
Q90	Vazão com permanência em 90% do tempo
Q95	Vazão com permanência em 95% do tempo
R ²	Coefficiente de determinação
RMSE	Raiz do erro médio quadrado

SIG Sistemas de Informações Geográficas
SIPOT Sistema de Informações do Potencial Hidrelétrico Brasileiro
USGS United States Geological Survey

SUMÁRIO

1. Introdução.....	12
2. Energia firme e garantia física.....	14
2.1 Período crítico hidrológico.....	14
2.2 Garantia física.....	16
3. Métodos para medição de vazão e incertezas associadas	17
3.1 Medição e integração da distribuição de velocidade	17
3.2 Método de medição de vazão acústico.....	19
3.3 Método de medição de vazão volumétrico.....	21
3.4 Método químico de medição de vazão.....	21
3.5 Medição de vazão com uso de dispositivos com geometria regular.....	22
3.6 Medição de níveis em uma seção de medição de descarga líquida	22
4. Consistência de dados fluviométricos.....	24
5. Metodologias para Regionalização de dados fluviométricos.....	26
5.1 Regressão Paramétrica	28
5.2 Análise de agrupamento de dados (“Cluster analysis”).....	29
5.3 Modelo de índices.....	32
5.4 Geometria hidráulica e fractal	33
5.5 Programação de expressão de genes	33
5.6 Interpolação em sistemas de informações geográficas (GIS)	35
6. Proposta de algoritmo para regionalização de dados fluviométricos em escala local	
36	
7. Verificação dos resultados e eficácia do algoritmo de regionalização	47
8. Conclusões	59
Referências Bibliográficas	62
Anexos.....	67

1. INTRODUÇÃO

A utilização da água, para produção de trabalho através de rodas d'água, remonta ao século III a.C. (MIRANDA, 2007), a primeira hidrelétrica, no entanto, remonta ao século XIX, construída em Cragside, na Inglaterra (EPE, 2018). Hoje a hidroeletricidade representa cerca de 64% da matriz energética brasileira, com aproximadamente 105 GW de potência instalada. (ANEEL, 2019a)

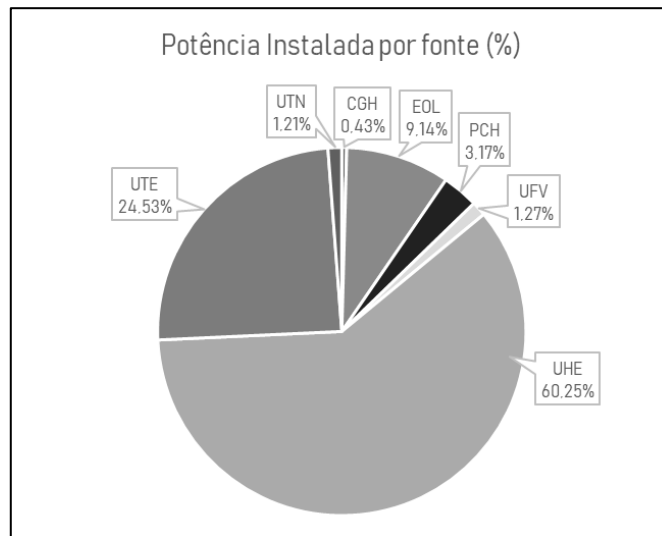


Figura 1 – Potência instalada por fonte da matriz elétrica brasileira. (ANEEL, 2019a)

De acordo com Plano Nacional de Energia 2050 (PNE – 2050) o potencial hidroenergético brasileiro é da ordem de 176 GW (EPE, 2018) e, portanto, é esperado que ainda sejam explorados aproximadamente 71 GW em hidroeletricidade. Este montante, é calculado com base nos Estudos de Inventários Hidrelétricos (EIH) apresentados e aprovados pela Agência Nacional de Energia Elétrica (ANEEL).

Os EIHs definem a potência instalada dos aproveitamentos hidrelétricos (AHEs) conforme metodologia descrita no Manual de Inventário Hidroelétrico de Bacias Hidrográficas, desenvolvido pelo Ministério de Minas e Energia (MME). Para definir o valor de potência instalada dos AHEs, deve-se determinar um Fator de Capacidade (FC) de referência para os empreendimentos. O FC é definido como a razão entre a energia firme e a potência instalada de um empreendimento, normalmente adotado como 55% na fase de pré-dimensionamento de pequenas hidrelétricas (esse valor é recomendado pelo MME no Manual de Inventário, caso o empreendedor desconheça os valores de fator de carga do mercado e perdas na transmissão). A partir desse número, e do cálculo da energia firme do aproveitamento, é definida a potência a ser instalada (MME, 2007).

A energia firme, por sua vez, é calculada utilizando-se a queda líquida média do aproveitamento e a descarga líquida média do período crítico do sistema de referência no local do aproveitamento, esta última, é definida como a vazão afluyente média do local, entre os meses de junho de 1949 a novembro de 1956, período histórico em que ocorre o deplecionamento máximo dos reservatórios do sistema elétrico brasileiro (MME, 2007). No entanto, são poucos os cursos d'água que possuem dados de vazão neste período, de fato, uma breve pesquisa nos dados de inventário da Agência Nacional de Águas (ANA) revela que apenas 3,8% dos rios inventariados pela agência possuem dados fluviométricos neste período (ANA, 2019).



Figura 2 – Estações fluviométricas inventariadas pela ANA, comparativo: nov/1949 a atualidade. (ANA, 2019).

Mesmo atualmente, apenas 28,8% dos cursos d'água inventariados pela ANA, possuem ou possuíram algum monitoramento fluviométrico até o presente. Portanto, fica clara a importância de técnicas de regionalização e transferência de dados hidrológicos entre bacias, para a estimativa de vazões em regiões onde não existem dados medidos, a fim de possibilitar o cálculo hidroenergético de futuros empreendimentos.

O presente trabalho tem como objetivo apresentar algumas destas técnicas, e propor um algoritmo computacional a fim de automatizar grande parte do processo de regionalização e transferência de dados fluviométricos. Por fim, pretende-se validar o algoritmo proposto identificando as principais incertezas associadas.

2. ENERGIA FIRME E GARANTIA FÍSICA

O setor elétrico brasileiro é hoje regulamentado pela Agência Nacional de Energia Elétrica – ANEEL, uma autarquia em regime especial vinculada ao Ministério de Minas e Energia, criada por meio da Lei nº 9.427/1996 e do Decreto nº 2.335/1997. A ANEEL tem como principais atribuições: (i) regular a geração (produção), transmissão, distribuição e comercialização de energia elétrica; (ii) fiscalizar, as concessões, as permissões e os serviços de energia elétrica; (iii) implementar as políticas e diretrizes do governo federal relativas à exploração da energia elétrica e ao aproveitamento dos potenciais hidráulicos; (iv) estabelecer tarifas; (v) dirimir as divergências, na esfera administrativa, entre os agentes e entre esses agentes e os consumidores, e (vi) promover as atividades de outorgas de concessão, permissão e autorização de empreendimentos e serviços de energia elétrica, por delegação do Governo Federal (ANEEL, 2019b).

A ANEEL é incumbida de receber, analisar e aprovar os inventários e projetos básicos de pequenas centrais hidrelétricas (PCHs) e registrar a construção de centrais geradoras hidrelétricas de capacidade reduzida (CGHs). A Energia Firme de um empreendimento é definida, no caso de PCHs quando é feito o inventário hidrelétrico, utilizando o período crítico do histórico de vazões para o cálculo hidroenergético da usina e consequente definição da potência instalada do aproveitamento (MME, 2007). Na essência a Energia Firme representa o pior cenário histórico de geração média teórica do empreendimento, ou seja, espera-se que, no mínimo, a usina seja capaz de gerar esta quantia. Para empreendimentos hidrelétricos despachados centralizadamente, a Energia Firme também é utilizada para determinar a parcela de participação de atendimento da carga crítica do sistema pelo empreendimento e consequente definição da Garantia Física do mesmo, conforme metodologia definida pela Portaria MME nº 101, de 22 de março de 2016 (MME, 2016).

2.1 PERÍODO CRÍTICO HIDROLÓGICO

O período crítico de um dado empreendimento, corresponde ao maior período em que ocorre o deplecionamento máximo do reservatório deste empreendimento, sem que, neste período o

reservatório passe pelo volume máximo mais de uma vez. O período se inicia com o reservatório em capacidade máxima (100% do armazenamento) e termina quando está totalmente deplecionado (0% do armazenamento - EPE, 2007). Normalmente, este conceito é aplicado a um grupo de reservatórios que atuam em conjunto em um mesmo sistema elétrico, dado que muitas vezes haverá influências entre os reservatórios, que se simulados de forma separada, não seriam consideradas.

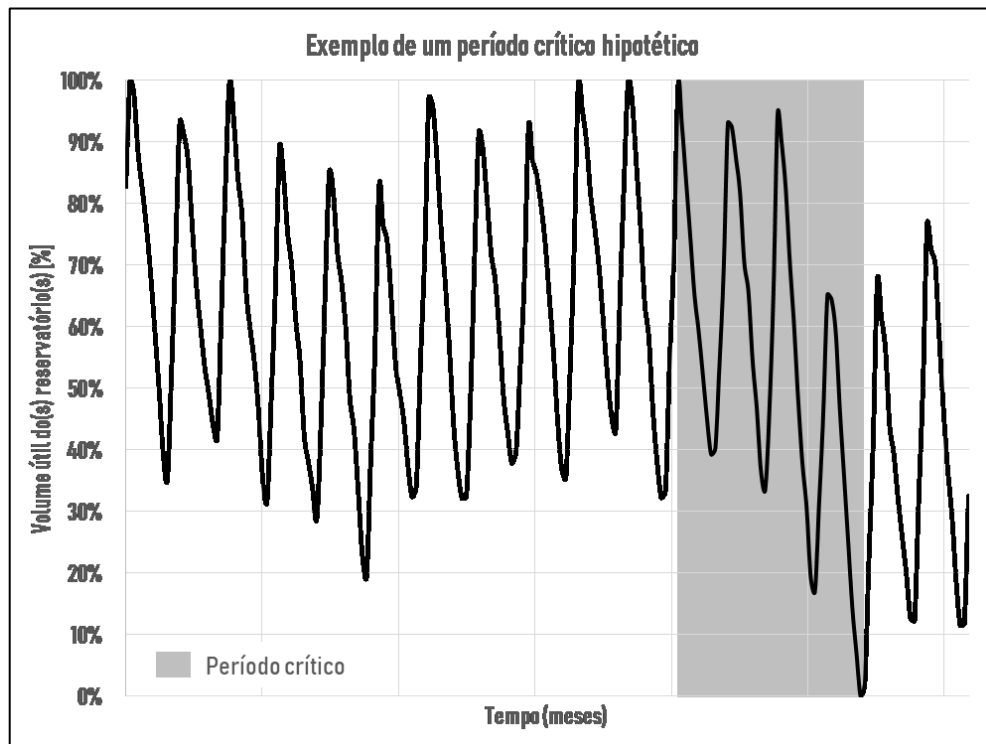


Figura 3 – Exemplo de período crítico hidrológico. (Fonte: o autor)

No Brasil o período crítico do sistema interligado pode ser calculado através do Modelo de Simulação a Usinas Individualizadas – MSUI, desenvolvido pela Eletrobrás. O software simula a operação do sistema mês a mês, para uma dada condição de carga (carga crítica), considerando os dados de (i) curva de cota dos reservatórios em função do volume, (ii) curva de área dos reservatórios em função da cota, (iii) curva do nível de jusante em função da vazão, (iv) perdas hidráulicas, por evaporação, por rendimento e fator de potência, para todas as usinas disponíveis no Sistema de Informações do Potencial Hidrelétrico Brasileiro – SIPOT (OLIVEIRA, et al., 2005).

O atual período crítico do sistema elétrico brasileiro, considerado para fins de cálculo da Energia Firme dos empreendimentos, foi determinado na década de 80 através de estudos realizados pelo Grupo Coordenador do Planejamento dos Sistemas Elétricos (GCPS), e é

definido como o período entre os meses de junho de 1949 a novembro de 1956 (OLIVEIRA, et al., 2005). Oliveira, et al. (2005) utilizaram séries de vazões atualizadas de janeiro de 1931, até dezembro de 2003 com o intuito de verificar a validade do período crítico determinado pelo GCPS, o estudo considerou a entrada em operação de novos empreendimentos hidrelétricos inventariados e planejados à época. Os resultados obtidos indicaram pouca ou nenhuma alteração no período crítico calculado anteriormente.

2.2 GARANTIA FÍSICA

A garantia física (GF), no caso de PCHs e CGHs não despachadas centralizadamente, objetiva representar a capacidade média de produção hidrelétrica de um empreendimento e tem como principais finalidades, (i) delimitar o montante de energia que poderá ser negociado em contratos de venda de energia elétrica por uma dada usina (lastro) e (ii) determinar a parcela de participação da usina no Mecanismo de Realocação de Energia – MRE. A garantia física é normalmente expressa em unidade de potência média (MW médios - CCEE, 2019).

O Ministério de Minas e Energia, através das instruções estabelecidas pela Portaria nº 463 de 3 de dezembro de 2009, recebe dos empreendedores as informações necessárias para o cálculo da garantia física de seus empreendimentos hidrelétricos não despachados centralizadamente pelo Operador Nacional do Sistema (ONS). Para usinas em operação a mais de quarenta e oito meses, com dados medidos de geração, a GF é calculada aplicando-se a média de geração do período após o primeiro ano até o último mês completo de medição.

Para empreendimentos que ainda irão entrar em operação, a garantia física é calculada a partir de uma simulação hidroenergética da usina, considerando: (i) a potência instalada total; (ii) a potência instalada por gerador e fator de potência dos mesmos; (iii) a potência instalada por turbina e o engolimento mínimo das mesmas; (iv) o rendimento das turbinas e geradores; (v) as taxas de indisponibilidade forçada e programada previstas para a usina; (vi) as perdas hidráulicas previstas; (vii) a queda bruta nominal; (viii) as perdas elétricas até o ponto de conexão com o sistema interligado nacional; (ix) o consumo de energia elétrica interno da usina; (x) a vazão remanescente no trecho de vazão reduzida (se aplicável – corresponde à vazão determinada pelo órgão ambiental ou Agência Nacional das Águas a ser mantida no trecho entre o barramento e o canal de fuga da usina); (xi) a vazão de usos consuntivos a montante do empreendimento; e (xii) o histórico de vazões médias mensais, não inferior a trinta anos.

Com exceção do histórico de vazões, todas as demais variáveis utilizadas no cálculo da GF são determinadas pelo projetista da usina, portanto, a variabilidade da garantia física está intimamente associada à disponibilidade hídrica. Tal observação, evidencia a importância da determinação de uma série histórica representativa e da realização de um estudo hidrológico que apresente de forma clara as incertezas associadas à série de vazões e conseqüentemente à garantia física da usina, informando ao empreendedor o risco hidrológico ao qual o empreendimento está exposto.

Este tipo de análise é comum para a determinação da garantia física de empreendimentos despachados centralizadamente, nos quais o risco de não atendimento à carga é fixado em 5% (OLIVEIRA, et al., 2009) e séries sintéticas de vazões, geradas a partir de modelos estocásticos, como, por exemplo, os apresentados por Kelman (1983), são utilizadas para calcular a garantia física dos empreendimentos (CEPEL, 2019). No entanto, para empreendimentos enquadrados na Portaria MME nº 463/2009, tal análise não é exigida.

3. MÉTODOS PARA MEDIÇÃO DE VAZÃO E INCERTEZAS ASSOCIADAS

Existem diversos métodos para medição de vazão em cursos d'água superficiais, variando principalmente com relação às ferramentas utilizadas para realizar as medições, a finalidade da medição e as dimensões do curso d'água. Segundo Santos et al. (2001) os principais métodos de medição são: (i) medição e integração da distribuição de velocidade; (ii) acústico; (iii) volumétrico; (iv) químico; (v) uso de dispositivos com geometria regular (calhas de Parshal e vertedores) e (vi) medição com flutuadores. Sendo este último extremamente incerto e aplicado apenas com o objetivo de obter estimativas preliminares da vazão ou em situações específicas nas quais a velocidade da corrente for superior a 5 m/s, colocando em risco o operador e os equipamentos tradicionais de medição.

3.1 MEDIÇÃO E INTEGRAÇÃO DA DISTRIBUIÇÃO DE VELOCIDADE

O método mais utilizado é o de medição e integração da distribuição de velocidade. Consiste em identificar através de equipamentos topográficos a profundidade em diversos pontos espaçados ao longo de uma seção selecionada de um curso d'água, criando-se subseções e em seguida medir a velocidade do fluxo em diferentes profundidades em cada subseção. O produto “área” e “velocidade” do fluxo permite calcular a vazão. Para medir a velocidade do fluxo, normalmente é utilizado um molinete (SANTOS et al., 2001). A Figura 4 apresenta a metodologia de forma esquemática.

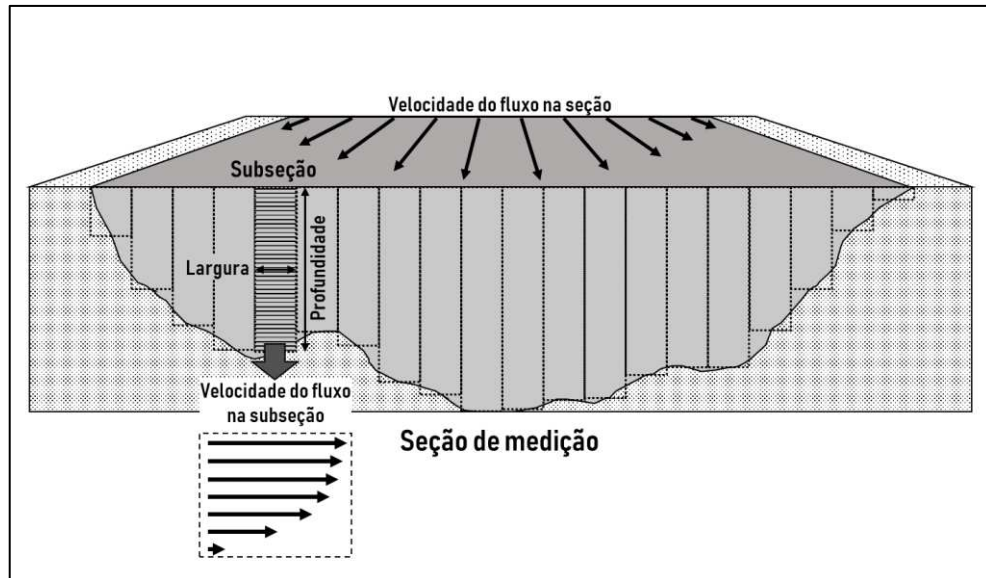


Figura 4 - Método de medição e integração da distribuição de velocidade. (Fonte: o autor)

As incertezas associadas ao método de medição e integração da distribuição de velocidade estão relacionadas a quatro diferentes fontes: (i) erros relacionados com a área da seção; (ii) erros relacionados com a velocidade do fluxo; (iii) erros relacionados com o método de cálculo e (iv) erros provocados por alterações no nível do curso d'água durante as medições (SAUER; MEYER, 1992). A Tabela 1 apresenta alguns resultados encontrados pelo Serviço Geológico dos Estados Unidos (USGS) com relação às incertezas da referida metodologia.

Tabela 1 - Incertezas relacionadas ao método de medição e integração da distribuição de velocidade (SAUER; MEYER, 1992).

Fontes de incerteza	Incertezas observadas
(i) Erros relacionados com a área da seção: Erros na medida da largura.	< 1%
(i) Erros relacionados com a área da seção: Erros na medida da profundidade.	2 a 15%
(ii) Erros relacionados com a velocidade do fluxo: Erros dos equipamentos de medição (molinetes).	0,32 a 5,14%
(ii) Erros relacionados com a velocidade do fluxo: Erros devido ao tempo de medição (exposição).	2,1 a 10,3 %
(ii) Erros relacionados com a velocidade do fluxo: Erros devido a distribuição vertical de velocidade e método de medição empregado (1 ou 2 pontos). Métodos mais precisos não foram avaliados no relatório da USGS.	4,3 a 11,2%
(ii) Erros relacionados com a velocidade do fluxo: Erros devido ao mal posicionamento do equipamento de medição.	< 1%
(iii) Erros relacionados com o método de cálculo: Premissa de que a velocidade em uma subseção é constante ou	1,02 a 7,7%

Fontes de incerteza	Incertezas observadas
varia de acordo com uma distribuição pré-estabelecida e não pontualmente. (erro diminui com o aumento no número de subseções).	
(iv) Erros provocados por alterações no nível do curso d'água durante as medições	Não mensuráveis
(v) Erros sistemáticos: provocados por equipamentos mal calibrados ou mal-uso dos equipamentos	0,5% (para cada medição: largura, profundidade e velocidade)

De acordo com Sauer e Meyer (1992), na maioria dos casos, as incertezas se mantêm na faixa de 3 a 6%, e podem ser consideradas boas medições, no entanto, em alguns casos onde várias condições ruins de medição se somam, as incertezas podem chegar a praticamente 20%. McMillan et al. (2012) reuniram diversos estudos sobre o tema e reportaram incertezas entre 2 e 17% sendo que para locais em que a velocidade do fluxo d'água é muito baixa (entre 0,5 e 0,05 m/s) foram calculadas incertezas de até 40%.

3.2 MÉTODO DE MEDIÇÃO DE VAZÃO ACÚSTICO

O método de medição de vazão acústico, mais conhecido como ADCP, da sigla em inglês: *Acoustic Doppler Current Profiling* (perfilador acústico de correntes por efeito Doppler), consiste na utilização do efeito doppler provocado pelo eco de pulsos de ultrassom refletidos em partículas sólidas suspensas e em movimento de uma dada seção do curso d'água, para calcular a velocidade do fluxo em pequenas áreas (chamadas de “células”) que subdividem a seção de medição. O produto “área” e “velocidade” do fluxo permite calcular a vazão. Diferentemente do método anterior, apenas um equipamento é utilizado para medir a velocidade e área. A Figura 5 apresenta a metodologia de forma esquemática.

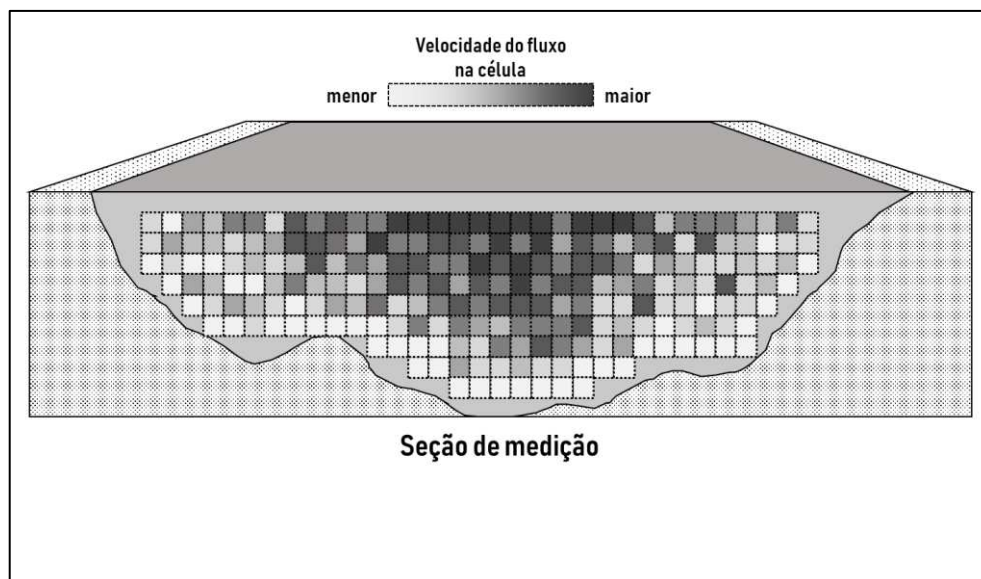


Figura 5 - método de medição de vazão acústico. (Fonte: o autor)

Huang (2012) apresenta como principais fontes de incertezas do método de medição de vazão acústico: (i) erros relacionados à calibração dos equipamentos; (ii) erros aleatórios relacionados à limitação do número de verticais; (iii) erros aleatórios relacionados com a medição da largura da seção; (iv) erros relacionados a variações de descarga durante as medições entre subseções; (v) erros de posicionamento do equipamento; (vi) leituras incorretas de temperatura e salinidade; (vii) má escolha do modelo de extrapolação do perfil de velocidade vertical e (viii) erros do operador. Simpson (2001) apresenta de forma mais detalhada as diversas fontes de incertezas relacionadas ao método acústico, e ressalta a importância da escolha correta da velocidade da embarcação ao realizar as leituras, dado que a escolha incorreta pode levar a erros superiores a 20%.

A literatura ainda carece de estudos que apresentem valores médios de incertezas diretamente associadas ao método de medição acústico, no entanto, existem diversas publicações (SIMPSON, 2001; HUANG, 2016; MUELLER, 2003; OBERG E MUELLER, 2007) que adentram o tema e algumas delas apresentam incertezas relativas da metodologia, comparada com o método de medição e integração da distribuição de velocidade. De acordo com McMillan et al. (2012) a bibliografia apresenta valores de incertezas relativas para a metodologia de medição acústica variando entre 3 e 7% com relação ao método anterior. Em estudo realizado em 20 estações brasileiras, por Gomes e Santos (1999), com o objetivo de comparar as metodologias, foram identificadas diferenças superiores a 10% em 5 das 26 medições realizadas (o estudo, no entanto, conclui que não existe uma preferência entre um método e outro).

3.3 MÉTODO DE MEDIÇÃO DE VAZÃO VOLUMÉTRICO

O método de medição volumétrico é (potencialmente) o mais preciso entre os métodos de medição de vazão, consistindo na medição direta do volume d'água despejado dentro de um reservatório em um dado período. A divisão do volume pelo período da coleta tem como resultado a vazão. Esta metodologia não é aplicável a qualquer curso d'água, sendo mais utilizada em duas situações: (i) se o curso d'água for muito pequeno, a ponto de se conseguir um recipiente grande o suficiente para coletar a água por alguns segundos, ou (ii) utilizando um reservatório de hidrelétrica com curva cota x volume conhecida de forma precisa, nesse caso, utiliza-se desta relação matemática para o cálculo da vazão, observando o aumento do nível do reservatório ao longo de um período de tempo medido (SANTOS et al., 2001).

As principais incertezas associadas ao método de medição volumétrico são: (i) erros de medição do volume do reservatório (técnica utilizada e precisão dos equipamentos); (ii) erros de medição do intervalo de tempo para enchimento (completo ou parcial) do reservatório (técnica utilizada e precisão dos equipamentos) e (iii) no caso de medições feitas a partir de grandes reservatórios: o cálculo de evaporação, precipitação e infiltração no solo de uma parcela da vazão afluyente. Tais incertezas variam significativamente a depender do tamanho do reservatório e das técnicas utilizadas para coleta de dados. A premissa de que a curva cota-área-volume do reservatório está atualizada e de que o reservatório não sofreu assoreamento desde a última atualização da curva é fundamental para a precisão da medição. Não foram identificados estudos específicos ponderando os erros médios obtidos com esta metodologia.

3.4 MÉTODO QUÍMICO DE MEDIÇÃO DE VAZÃO

O método químico de medição de vazão é mais utilizado em rios de montanha com escoamentos turbulentos, onde é mais difícil utilizar o método tradicional com molinetes. O método consiste em despejar no curso d'água uma solução homogênea de um determinado produto químico solúvel em água, com concentração, local e horário conhecidos e a depender da metodologia utilizada, também é preciso conhecer a vazão ou o volume da solução despejada. A jusante deste local, a uma distância suficiente para que a solução despejada tenha se tornado homogênea no corpo hídrico, são coletadas amostras de água ao longo de um período de tempo pré-determinado a fim de se determinar a concentração resultante do produto despejado. Com as informações coletadas, é possível calcular o volume final da solução (produto químico solvido de forma homogênea no corpo hídrico) estabelecendo uma relação entre a concentração inicial e final do produto químico e a vazão ou volume inicial da solução (antes de entrar em contato

com o corpo hídrico). Com o volume final em mãos e o tempo registrado entre o despejo e a coleta, é possível determinar a vazão no trecho (SANTOS et al., 2001).

As principais incertezas associadas ao método químico de medição de vazão são: (i) risco do produto químico não se misturar de forma homogênea no corpo d'água principal; (ii) erros de medição da vazão de injeção da solução e (iii) erros de medição da concentração do produto químico (inicial e final). Santos et al. (2001) apontam incertezas relativas identificadas na literatura entre 1% e 7%, comparando o método químico com a utilização do método tradicional, com molinetes. McMillan et al. (2012) apontam valores de incertezas relativas entre 5% a 7,1% para essa mesma condição.

3.5 MEDIÇÃO DE VAZÃO COM USO DE DISPOSITIVOS COM GEOMETRIA REGULAR

A medição de vazão com o uso de dispositivos com geometria regular, utiliza de relações matemáticas teóricas a respeito do comportamento do fluxo de água através de uma dada estrutura para a partir da observação de apenas uma variável (normalmente o nível d'água em uma seção específica da estrutura) determinar a vazão teórica correspondente. Exemplos deste tipo de dispositivos são Calhas de Parshall, vertedouros triangulares, trapezoidais e retangulares com contração (SANTOS et al., 2001).

As incertezas associadas ao método de medição com dispositivos de geometria regular são: (i) erros de execução da construção do dispositivo; (ii) erros de leitura do nível e (iii) incerteza associadas as equações (coeficientes de descarga). De acordo com Herschy (2009) as incertezas consideradas para este tipo de medição devem ser da ordem de 5%. No entanto, a instalação incorreta e a ocorrência de cheias que provoquem a submergência excessiva da seção de jusante podem provocar erros superiores a 35% (WAHL, 2002).

3.6 MEDIÇÃO DE NÍVEIS EM UMA SEÇÃO DE MEDIÇÃO DE DESCARGA LÍQUIDA

A obtenção de séries de vazões ao longo de períodos extensos, com o objetivo de estudar o comportamento hidrológico de um certo local, envolve a coleta recorrente de dados de descarga líquida em uma dada seção do curso d'água. Na maioria das vezes é inviável manter uma equipe dedicada a medir vazões por longos períodos utilizando os métodos anteriormente citados, nesses casos opta-se por realizar leituras periódicas do nível d'água no local, com frequência contínua, horária ou diária e, posteriormente, correlacionar os dados de nível com medições de

vazões pontuais realizadas no mesmo local, ajustando uma curva (conhecida como “curva-chave” ou “curva de descarga”) que passa a definir valores de vazão para cada medição de nível realizada. A Figura 6 exemplifica uma curva de descarga.

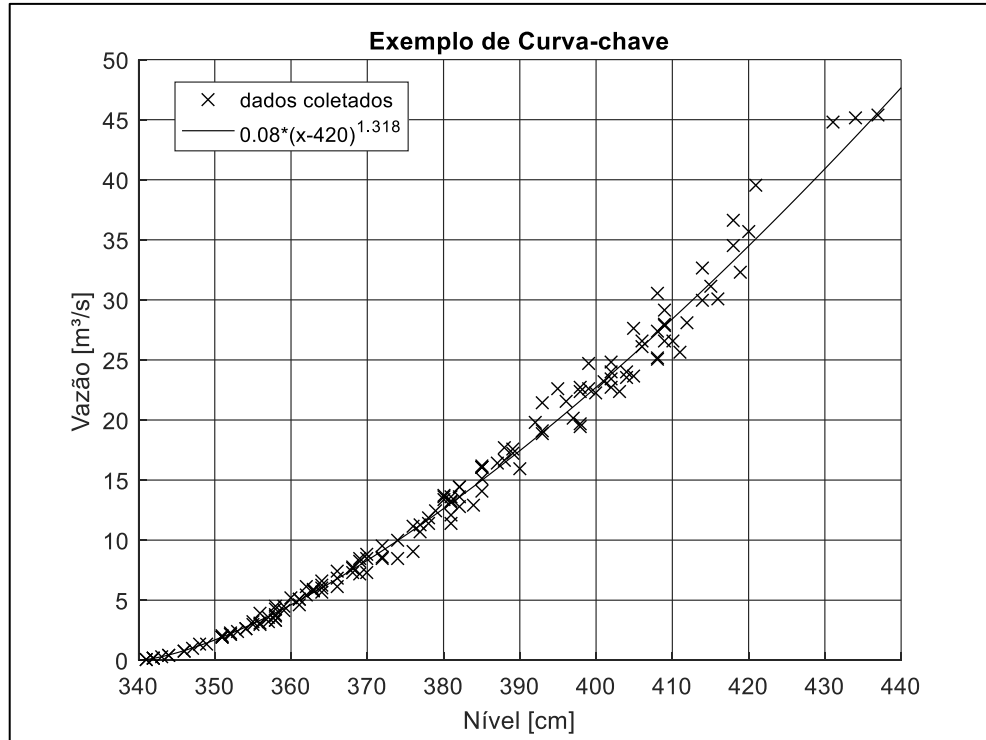


Figura 6 - Exemplo teórico de curva-chave. (Fonte: o autor)

As incertezas relacionadas aos dados de vazão gerados a partir de curvas-chave, variam de acordo com o método de medição empregado, podendo ser: linígrafos analógicos com boia, transdutores de pressão submersos, linígrafos de bolhas ou leitura de régua por operador. As principais incertezas citadas na literatura são: (i) erros de instalação das réguas; (ii) variações no leito do rio (assoreamento, alterando a forma da seção); (iii) erros de medição (problemas mecânicos no linígrafo, anotações erradas, invenção de registros, etc.); (iv) extrapolações erradas da curva-chave para níveis de vazões não medidos; (v) no caso de medições por leitura de régua, pode ocorrer uma cheia entre uma medição e outra do mesmo dia, e que acabaria não sendo documentada). Santos et al. (2001) afirma que para a aplicação desta metodologia, devem ser consideradas incertezas da ordem de 5% a 10% para vazões obtidas a partir de níveis de água observados e 10% a 20% para vazões no trecho extrapolado (não observado) da curva. McMillan et al. (2012) reuniram diferentes estudos apresentando incertezas entre 46% e 100% para vazões mais baixas (início da curva) e entre 10% e 20% para vazões médias a altas.

4. CONSISTÊNCIA DE DADOS FLUVIOMÉTRICOS

Após a coleta de dados de nível que serão convertidos em vazões através de uma curva-chave, é preciso analisar os resultados obtidos a fim de identificar possíveis erros que comprometam a credibilidade do estudo. Para tal, a Agência Nacional de Águas requer que os dados coletados passem por um processo de consistência, que envolve: (i) detecção de erros; (ii) correção de erros; e (iii) preenchimento de falhas.

A detecção de erros deve contemplar a possibilidade de ocorrência de erros grosseiros, sistemáticos e fortuitos (SANTOS et al., 2001) alguns exemplos de erros grosseiros são: leitura efetuada de maneira errada, seja interpretando um lance de régua como sendo outro ou na visualização incorreta das demarcações da régua, entupimento de condutos no caso de linígrafo, leitura efetuada em horários distintos dos definidos, invenção do registro, boia furada, entre outros. Alguns exemplos de erros sistemáticos também são apresentados por Santos et al. (2001), como: alterações no leito do rio que provoquem a alteração do “zero da régua”, mudança de local ou de margem, influência de pontes ou outras obras que provoquem remanso etc. Os erros fortuitos, por sua vez, são normalmente muito menores e estão dentro da margem de precisão dos equipamentos de medição ou da capacidade humana de observação (SANTOS et al., 2001) alguns exemplos de erros fortuitos são: pequenas variações de nível provocadas por ondas, escorregamento do cabo de aço na roldana (no caso de linígrafos que utilizam esse tipo de equipamento), variações de nível mais rápidas que a inércia do linígrafo ou da frequência de medição ou registro do transdutor ou *datalogger* utilizados, etc.

A análise de consistência deve, sempre que possível, considerar os dados coletados nas demais estações situadas na mesma bacia, possibilitando uma visão sistêmica dos eventos e facilitando a identificação de dados incorretos (SANTOS et al., 2001). As informações coletadas devem ser separadas em dois grupos distintos: (i) estações principais: locais onde os registros serão utilizados para o estudo e (ii) estações secundárias: que possuem como finalidade verificar a existência de erros nas estações principais.

As estações principais são identificadas como aquelas que possuem localização adequada na bacia para o propósito do estudo, possuem grande extensão do período de observação (em geral superior a 25 anos), possuem curvas de descarga estáveis (característica hidráulica não varia demais ao longo do tempo), possuem poucos períodos de interrupção na coleta de dados, possuem medições de vazões extremas máximas e mínimas, e são bem operadas, sem que tenha sofrido muitas alterações ao longo dos anos.

Uma visita técnica ao local da estação, assim como um relatório fotográfico, também é importante para identificação de possíveis erros (como, por exemplo: verificar se as réguas e sensores estão posicionadas corretamente). Utilizar estudos hidrológicos anteriores, realizados para a mesma região, também são fonte de informações interessantes que podem auxiliar no desenvolvimento do trabalho. (SANTOS et al., 2001). Na Figura 7 encontra-se um fluxograma simplificado das etapas do estudo de consistência de dados fluviométricos descrito por Santos et al. (2001).

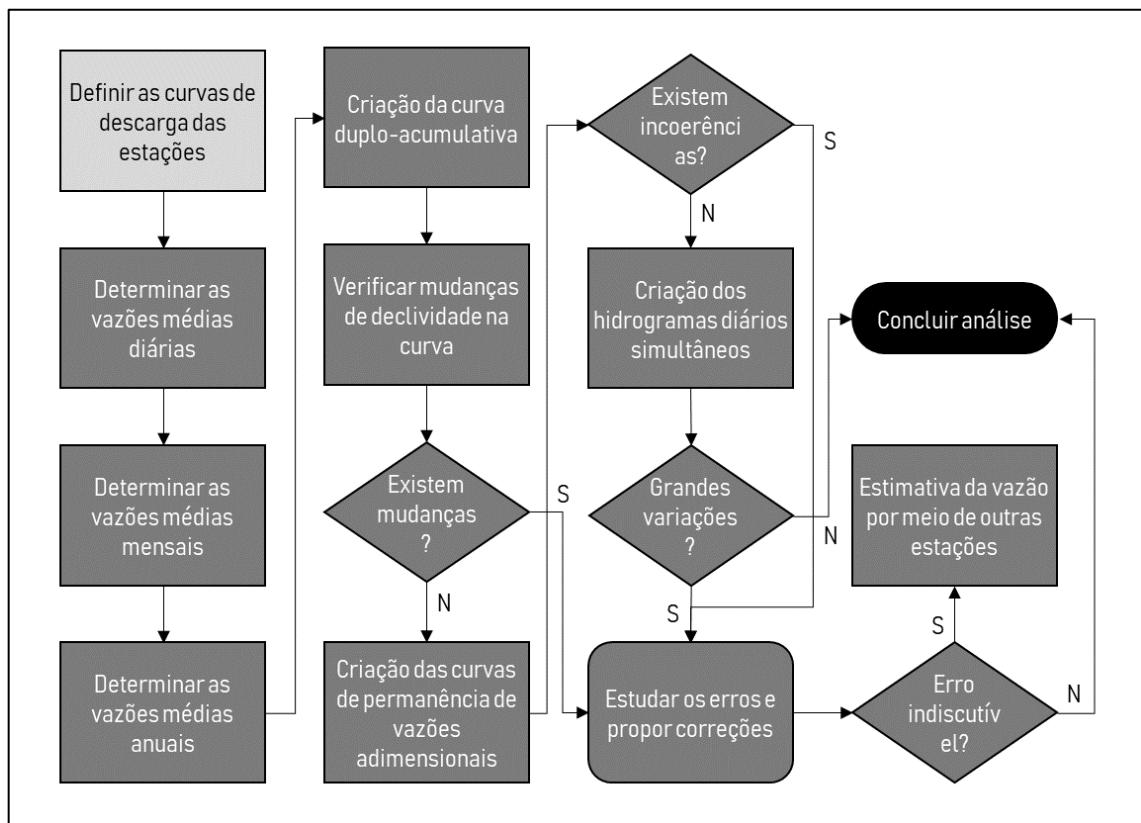


Figura 7 - Fluxograma simplificado da consistência de dados fluviométricos. (fonte: o autor)

A curva duplo-cumulativa, referenciada na Figura 7 é traçada a partir dos pontos formados pela média acumulada anual da estação principal e soma das médias acumuladas anuais das demais estações da bacia. As curvas de permanência de vazões, por sua vez, são obtidas ordenando-se as vazões de cada estação para o mesmo período de dados, da maior vazão (que é igualada apenas uma vez em todo o histórico e, portanto, possui permanência próxima a 0%) até a menor vazão (que é superada ou igualada 100% do tempo e, portanto, possui permanência de 100%).

Toda a metodologia utilizada para a análise de consistência das cotas lidas, medições realizadas, traçado de curvas-chave, extrapolação de curvas-chave, vazões, preenchimento de falhas etc.

deve ser descrito em um relatório e encaminhado para a ANA juntamente com os dados consistidos (ANA, 2012).

5. METODOLOGIAS PARA REGIONALIZAÇÃO DE DADOS FLUVIOMÉTRICOS

Devido à natureza de dimensões geográficas das bacias hidrográficas e das limitações tecnológicas e econômicas para se realizar medições hidrológicas, muitas bacias não possuem equipamentos para medição de dados fluviométricos ou pluviométricos, em muitos casos, inclusive, as redes de coleta de dados hidrológicos vêm diminuindo (SIVAPALAN et al., 2003). Métodos para regionalização de dados característicos de bacias hidrográficas são frequentemente estudados na hidrologia e servem ao propósito de estimar informações em áreas de drenagem que não possuem nenhuma medição ou medições insuficientes para o estudo desejado.

De forma geral, a regionalização de dados hidrológicos entre bacias pode ser resumida nas seguintes etapas: (i) Seleção da Bacia que se pretende estudar; (ii) Levantamento dos dados disponíveis desta bacia; (iii) Seleção dos dados que pretende-se conhecer; (iv) Seleção de uma ou mais bacias que possuem dados suficientes para o objetivo pretendido; (v) Estabelecer uma relação entre os dados existentes e os dados que se deseja conhecer; (vi) Criar um modelo para gerenciar essas relações e aplica-las aos dados conhecidos da bacia estudada; (vii) A partir do modelo, gerar os dados estimados para a bacia em estudo. A Figura 8 apresenta um fluxograma destas etapas.

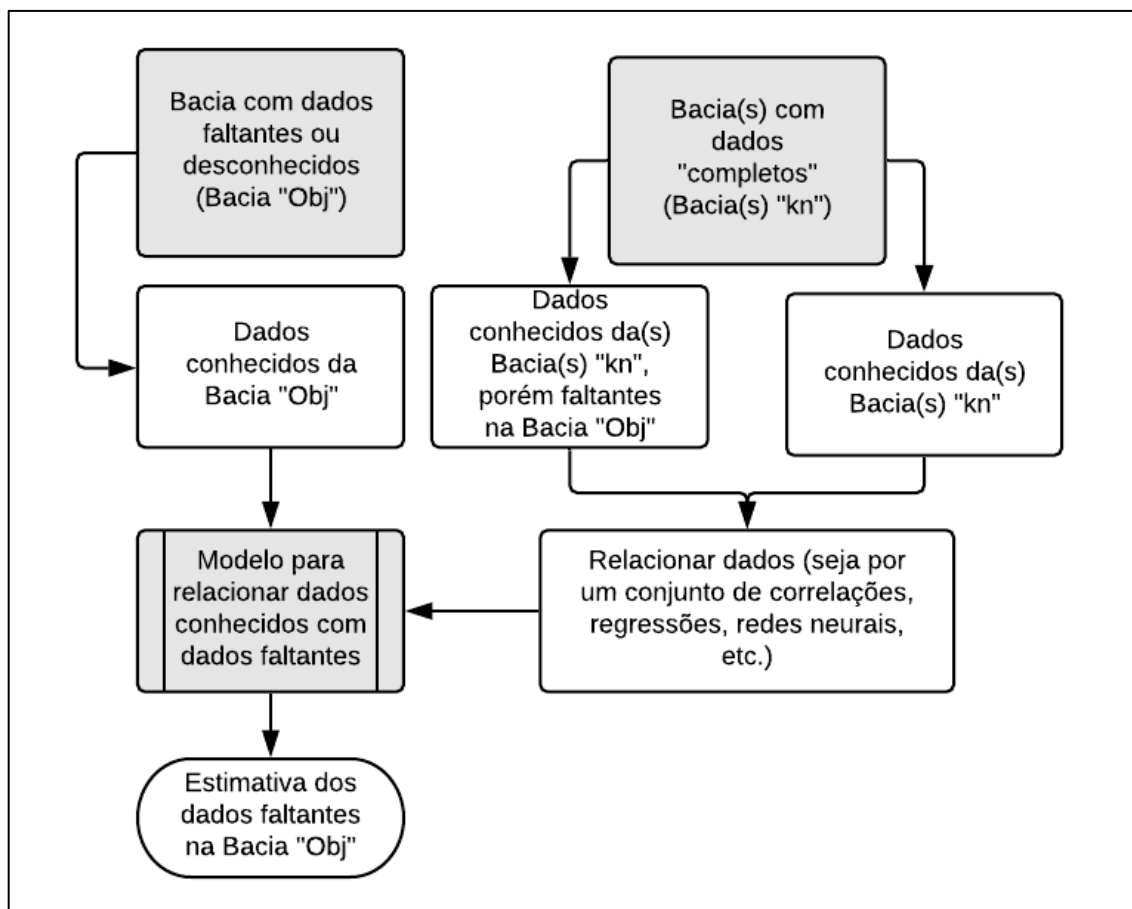


Figura 8 - Fluxograma simplificado da regionalização de dados hidrológicos. (Fonte: o autor)

Na literatura são encontradas diferentes abordagens metodológicas para regionalização de dados fluviométricos, utilizando-se de diferentes ferramentas matemáticas. Ao longo dos anos novas tecnologias e métodos surgem, diferentes métodos se unem em novos métodos mais complexos e, por isso, é muito difícil encontrar mais de um autor que classifique as metodologias para regionalização de dados hidrológicos da mesma forma.

Rao et al. (2008) classificam as diferentes abordagens para regionalização como (i) método de resíduos, (ii) análise de correlação canônica, (iii) região de influência, (iv) abordagem hierárquica e sua extensão à abordagem pela região de influência e (v) análise de cluster. Ming Li et al. (2010) citam três metodologias típicas: (i) regressão paramétrica, (ii) vizinho mais próximo e (iii) similaridade hidrológica. Estas classificações são direcionadas à metodologia utilizada para definir se uma ou mais bacias de referência são adequadas, do ponto de vista hidrológico, para terem seus dados regionalizados para a bacia em estudo, ou seja, se as bacias de referência apresentam um comportamento hidrológico com um certo nível de similaridade ao comportamento que se acredita que a bacia em estudo apresente.

Castellarin et al. (2004) e Hashmi e Shamseldin (2014) classificam as técnicas de regionalização de forma mais ampla, como: (i) estatísticas e (ii) paramétricas. Esta classificação, diferentemente da primeira, foca nas ferramentas matemáticas utilizadas para trabalhar e representar os dados regionalizados, se estes serão aproximados de uma função probabilística ou apenas manipulados matematicamente através de parâmetros determinísticos.

Tendo em vista a dificuldade de se definir classificações cartesianas às diferentes metodologias de regionalização existentes, o presente capítulo irá abordar alguns exemplos de metodologias encontradas na literatura à medida que apresenta algumas das ferramentas mais utilizadas em diferentes abordagens. Muitas vezes uma única metodologia utiliza diversas técnicas e ferramentas diferentes para alcançar o resultado.

5.1 REGRESSÃO PARAMÉTRICA

Diversas metodologias para regionalização de dados hidrológicos utilizam da ferramenta matemática de regressões lineares. De acordo com Abdulla e Lettenmaier (1997) a utilização de regressões múltiplas é o método mais utilizado para estimativa de parâmetros regionalizados em aplicações hidrológicas, especialmente para transferência de dados fluviométricos de bacias que possuem estações de medição fluviométrica para bacias que não possuem tais equipamentos. A equação (1) apresenta um modelo geral de regressão múltipla.

$$Y = \beta_1 X_1 + \beta_2 X_2 \dots + \beta_n X_n \quad (1)$$

Na qual Y é uma variável dependente (normalmente a variável que se deseja conhecer), X_1, X_2, \dots, X_n são variáveis independentes (normalmente são dados característicos das bacias) e $\beta_1, \beta_2, \dots, \beta_n$ são parâmetros desconhecidos, que serão determinados através do método de mínimos quadrados utilizando valores de X_1, X_2, \dots, X_n e Y conhecidos.

Um exemplo de modelo de regionalização que utiliza esse tipo de metodologia é o desenvolvido por Abdulla e Lettenmaier (1997) para regionalização de parâmetros hidrológicos estimados através de equações que utilizam dados hidrológicos em macroescala. Para isso, os pesquisadores coletaram dados observados de vazão, precipitação e temperatura em bacias monitoradas e confrontaram com dados climatológicos, morfológicos e propriedades do solo dos mesmos locais. A partir daí foram determinadas quais destas últimas variáveis possuíam potencial para explicar as primeiras (variáveis independentes) e assim foram construídas as equações paramétricas através de regressões múltiplas. Por fim o modelo foi utilizado para estimar dados fluviométricos em bacias não monitoradas.

5.2 ANÁLISE DE AGRUPAMENTO DE DADOS (“CLUSTER ANALYSIS”)

Uma das principais premissas da regionalização de dados fluviométricos é a de que os dados de uma bacia de referência, em conjunto com o modelo de regionalização adotado, irão se transformar em informações hidrológicas úteis para a bacia em estudo (SIVAPALAN et al., 2003). Na literatura encontramos diversas metodologias para a seleção das bacias de referência. De acordo com Linsley et al. (1975) as bacias de referência devem possuir características hidrológicas as mais similares possíveis daquela que se estuda, no entanto, Grigg (1967) e Mosley (1981) elucidam que muitas vezes essa definição é feita de maneira subjetiva. A análise de agrupamento de dados ou “*Cluster analysis*” é uma das metodologias utilizadas com o objetivo de diminuir a subjetividade na seleção ou separação de bacias que possuem similaridade de características hidrológicas.

Rao et al. (2008) fazem uma extensa revisão sobre diversos algoritmos para agrupamento de dados, classificando-os em “*hard clustering*” (fronteira rígida) onde as definições dos agrupamentos seguem a lógica booleana (uma bacia pertence “1” ou não pertence “0” ao agrupamento) ou “*fuzzy clustering*” (fronteira difusa) onde os agrupamentos seguem a lógica *fuzzy* (uma bacia pode se assemelhar em 70% com o agrupamento “a” e 30% com o agrupamento “b”, por exemplo).

5.2.1 Agrupamento de dados de fronteira rígida

Dentro da abordagem de fronteira rígida, encontramos os algoritmos hierárquicos e particionais. Nos algoritmos hierárquicos uma bacia hidrográfica, por exemplo, poderia pertencer a um grupo de bacias muito similares, que por sua vez faz parte de um grupo maior de bacias um pouco menos similares e assim por diante, descrevendo um dendrograma (Figura 9), conforme utilizado por Mosley (1981) para tentar delimitar regiões hidrológicamente homogêneas na Nova Zelândia. Como parâmetros para o agrupamento de dados ele analisou a distribuição de pontos criados pela vazão específica e o coeficiente de variação instantânea de descargas de diversas estações fluviométricas espalhadas pelo país.

Já nos algoritmos particionais o resultado final é um conjunto de agrupamentos onde cada bacia pertence a apenas um dos grupos identificados, Burn (1989) utilizou da metodologia de agrupamento de dados “*K-means*” para identificar regiões hidrológicamente homogêneas na região sul de Manitoba, Canadá. Como parâmetros para o agrupamento de dados ele analisou a distribuição de pontos criados pela vazão específica, o coeficiente de variação instantânea de

descargas e a localização geográfica (latitude e longitude) de diversas estações fluviométricas espalhadas pela região.

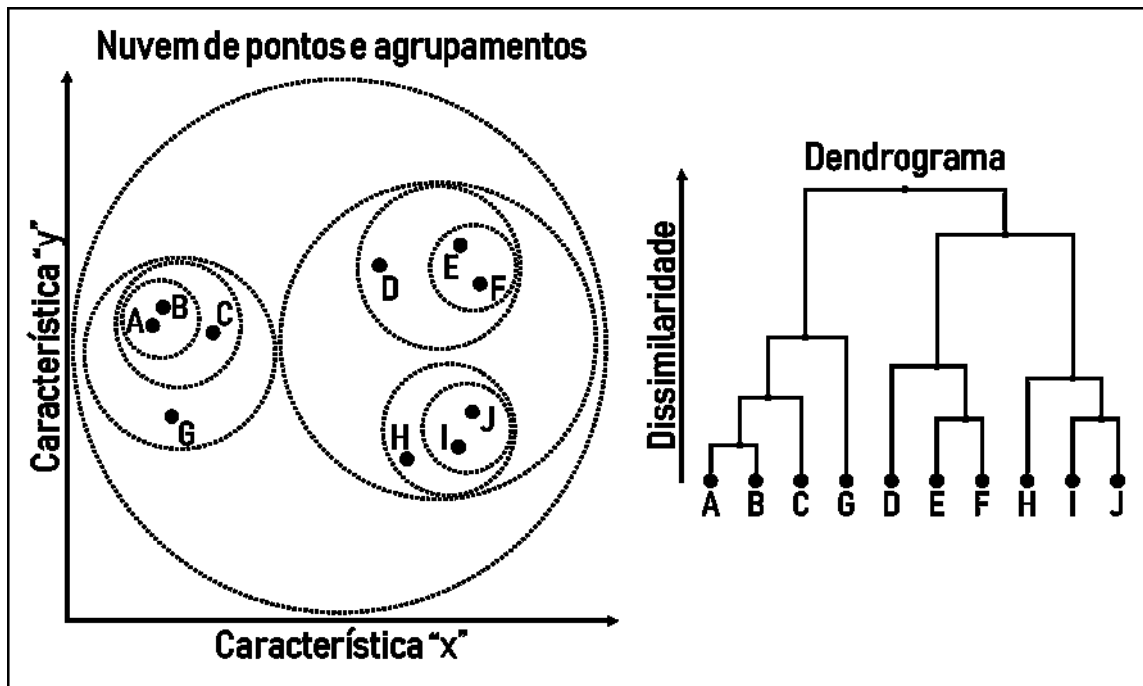


Figura 9 – Representação esquemática de um dendrograma. (Fonte: o autor)

A Figura 10 traz um exemplo genérico de como são definidos os agrupamentos utilizando o método “*K-means*”, onde são determinados “*K*” grupos. O objetivo da metodologia é diminuir ao máximo a distância total entre os pontos dentro dos grupos e os centroides dos respectivos grupos, no entanto, também é necessário definir um valor máximo para “*K*” de modo que o agrupamento tenha significado (uma vez que, no limite, seriam criados um grupo para cada ponto e a distância total seria zero).

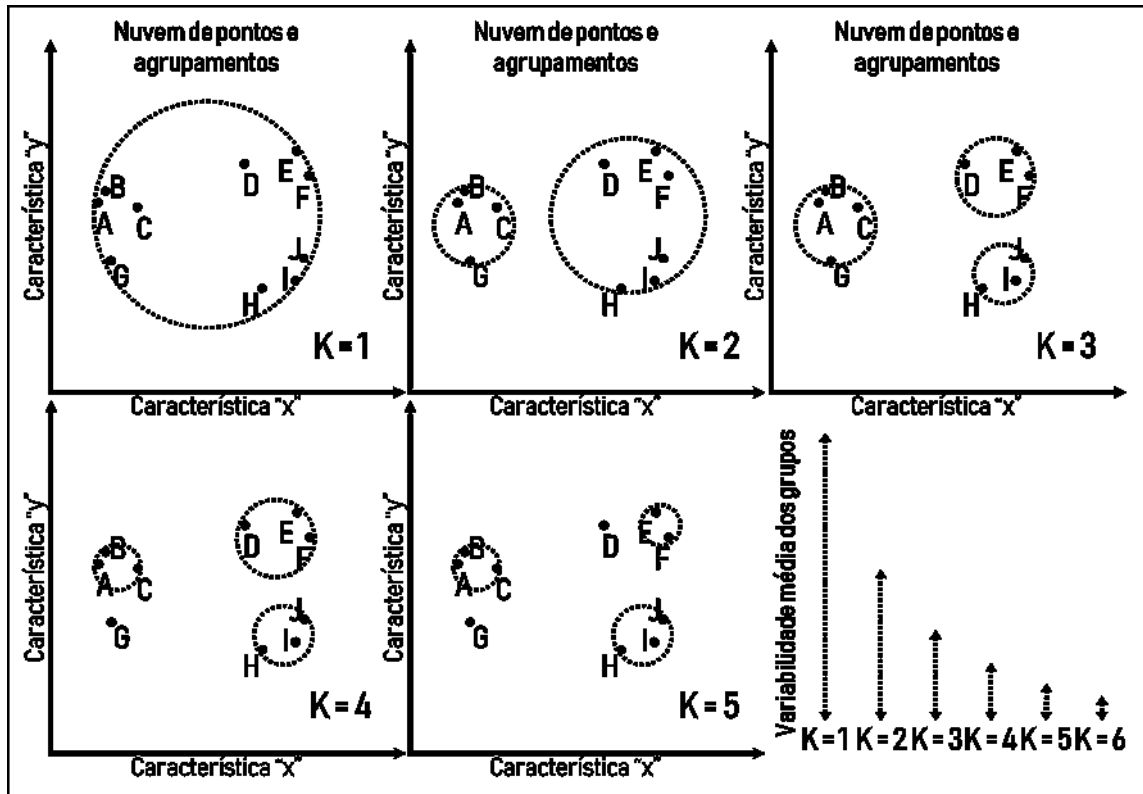


Figura 10 – Exemplo genérico de uma das etapas do algoritmo de “*k-means*” para agrupamento de dados. (Fonte: o autor).

5.2.2 Agrupamento de dados de fronteira difusa

É possível que uma dada bacia apresente características de um ou mais agrupamentos definidos através da metodologia de fronteira rígida, de fato, ao tentar identificar as regiões hidrologicamente homogêneas da Nova Zelândia, Mosley (1981) teve dificuldade em definir fronteiras rígidas para os resultados encontrados. A abordagem de fronteira difusa, tende a representar agrupamentos de bacias hidrológicas consideradas homogêneas de maneira mais fiel a realidade (RAO et al., 2008), pois esta metodologia permite que uma bacia seja identificada como parcialmente pertencente a um ou mais agrupamentos.

Assim como a abordagem de fronteira rígida, existem diversos algoritmos para agrupamento de dados de fronteira difusa. Entre eles o método “*Fuzzy c-Means*” (FCM) pode ser comparado ao método “*k-means*” para fronteira rígida. Rao et al. (2008) apresentaram os resultados obtidos com a utilização de ambas as abordagens para identificar regiões hidrologicamente homogêneas no estado de Indiana nos EUA.

5.2.3 Redes neurais aplicadas para agrupamento de dados

Rao et al. (2008) também utilizam o algoritmo de Kohonen, de mapas auto organizáveis (*Self-Organizing Maps – SOM*) em conjunto com técnicas de agrupamento de dados de fronteiras rígida e difusa na tentativa de agrupar de forma mais precisa e com o número ótimo de clusters as bacias do estado de Indiana, nos EUA.

Mapas auto organizáveis constituem uma técnica de agrupamento de dados que utilizam redes neurais para organizar dados com quantidades elevadas de atributos (muitas dimensões) em um plano bidimensional (conhecido como camada de Kohonen), que é descrito por um conjunto de neurônios. Em algumas aplicações estes neurônios podem descrever diretamente os clusters formados pelo agrupamento de dados ou, como foi feito no trabalho apresentado por Rao et al. (2008), essa rede pode passar por um pós-processamento, onde são utilizados algoritmos de agrupamento de fronteira rígida ou de fronteira difusa, para se chegar ao resultado final, com a divisão dos dados em grupos homogêneos e de mais fácil visualização.

5.3 MODELO DE ÍNDICES

A partir do estudo de diferentes metodologias para regionalização de dados hidrológicos, Li et al. (2010) propuseram uma nova metodologia para predição de dados em bacias sem estações hidrométricas. O modelo de índices (*the index model*) busca unir as vantagens de três diferentes metodologias de regionalização: (i) regressão linear; (ii) vizinho mais próximo (*nearest neighbor*) e (iii) similaridade hidrológica.

De acordo com Li et al. (2010) a regionalização utilizando apenas o método de regressão linear, apesar de apresentar bons resultados para pequenas áreas, muitas vezes não consegue explicar bem os processos em áreas maiores, exigindo a utilização de modelos não lineares.

O modelo de vizinho mais próximo não se preocupa em estabelecer relações específicas entre variáveis hidrológicas observadas, e apenas utiliza os dados da estação mais próxima transportando-os para a bacia que pretende-se estudar, essa característica do modelo de vizinho mais próximo evita alguns erros de modelagem da região hidrológicamente homogênea, mas acaba limitando a sua utilização em locais que possuem um maior disponibilidade de estações.

O modelo de similaridade hidrológica utiliza do conceito de “proximidade hidrológica” que consiste em calcular a “distância” Euclidiana entre o local estudado e a bacia de referência em termos hidrológicos (normalmente área da bacia, pluviosidade média, e características de escoamento do solo), mas acaba ignorando características intrínsecas à localização geográfica.

O modelo de índices procura incorporar a forma básica de modelagem utilizada pela regressão linear, no entanto, permitindo a ocorrência de relações não paramétricas com a introdução de uma função suave que é definida exclusivamente pelos dados de entrada. Essa característica admite o relaxamento das relações monotônicas impostas pela regressão linear e permite que o modelo passe a explicar relações não lineares entre variáveis.

Li et al. (2010) testaram a eficácia do modelo de índices aproximando curvas de permanência de vazões para 227 bacias monitoradas na Austrália. Como conclusão do estudo, foi identificado que o modelo de índices obtém resultados melhores se comparado aos modelos de regressão linear, vizinho mais próximo e similaridade hidrológica. Destes três, o modelo de regressão linear foi o que chegou mais próximo dos resultados do modelo de índices.

5.4 GEOMETRIA HIDRÁULICA E FRACTAL

Pereira et al. (2016) propuseram a utilização conjunta das metodologias de regionalização utilizando geometria hidráulica (GH) e geometria fractal (GFR) para modelar o hidrograma unitário instantâneo geomorfológico (HUIG) de estações selecionadas na bacia do rio Ijuí, no Estado do Rio Grande do Sul. A metodologia propõe a modelagem matemática, estatística e probabilística das bacias utilizando suas características fisiográficas e aplica as Leis de Horton para determinar: (i) o HUIG-C, calibrado através da velocidade; (ii) o HUIG-L que utiliza uma função linear resultante de regressão múltipla envolvendo os coeficientes de GH, GFR e de dados de precipitação média anual; e o HUIG-P que utiliza uma função potencial resultante da inserção das mesmas variáveis.

O resultado encontrado por Pereira et al. (2016) demonstra que a metodologia pode ser aplicada para previsão de picos de cheia, dada a previsão de precipitação, no entanto, os erros obtidos com relação aos valores absolutos de vazão previstos e medidos são da ordem de 30%.

5.5 PROGRAMAÇÃO DE EXPRESSÃO DE GENES

Hashmi e Shamseldin (2014) utilizaram da técnica de programação de expressão de genes para aperfeiçoar equações paramétricas com o objetivo de prever as seguintes características de uma curva de permanência de vazões: (i) Q50, vazão com tempo de permanência de 50%; (ii) Qn10, vazão com 10% de permanência, normalizada a partir da divisão pela Q50; (iii) Qn80, vazão com 80% de permanência, também normalizada através da Q50 e (iv) Qn95, vazão com 95% de permanência, também normalizada.

Os pesquisadores disponibilizaram 26 variáveis de entrada que poderiam ser utilizadas pelo modelo para compor as equações paramétricas, como por exemplo: área de drenagem, perímetro da bacia, precipitação média anual, permeabilidade do solo, declividade média etc., no entanto, esperava-se que o algoritmo genético selecionasse apenas aquelas que realmente contribuiriam efetivamente para determinar as variáveis.

O estudo utilizou 12 estações hidrométricas nas proximidades de Auckland, Nova Zelândia, sendo 6 delas para treinamento do algoritmo genético e 6 para validação dos resultados. Das 26 variáveis disponibilizadas, os pesquisadores restringiram a escolha do modelo apenas na determinação da Q50, que ficou limitada a utilizar a área de drenagem (representada pela letra “A” nas equações em km²) e a precipitação média anual (representada pela sigla “MAP” nas equações em mm), resultando na equação (2).

$$Q50 = \frac{A}{69,22} \left[1 - \frac{156,4}{MAP} \right] \quad (2)$$

$$Qn10 = \frac{\sqrt{Tall}}{Perm} + \sqrt{C + GRAV + \sqrt{(PRAW - MPORD) \cdot Perm}} \quad (3)$$

$$Qn80 = \frac{DSLO \cdot RR \cdot \sqrt{RR + E}}{9,84 \cdot MAP} + \sqrt{\frac{El \cdot PRD}{7,09 - \frac{Perm}{2,66}}} \quad (4)$$

$$Qn95 = \frac{MAP - PET}{d} + \frac{El}{MPORS} - 34,22 \cdot Tall \cdot \frac{As}{Perm} \quad (5)$$

Para as demais variáveis: Qn10, Qn80 e Qn95, o modelo selecionou apenas seis características para determinar a Qn10 (equação (3), onde: **Tall** é o percentual de vegetação alta (%); **Perm** é a permeabilidade do solo (E-06 m/s); **C** é o coeficiente de compacidade da bacia (adimensional); **GRAV** é o percentual de cascalho na superfície do solo (%); **PRAW** é a umidade disponível no solo (mm); e **MPORD** é a macroporosidade do solo entre 0,6 e 0,9 metros de profundidade (%)). sete características para a Qn80 e oito características para determinar a Qn95 (equações (4) e (5), onde: **DSLO** é a profundidade onde ocorre a permeabilidade lenta, abaixo de 4 mm/h (m); **RR** é a declividade média da bacia (adimensional); **E** é a elevação média da bacia (m); **El**

é a raiz quadrada da área da bacia dividida pelo comprimento da bacia (adimensional); **PRD** é a profundidade potencial das raízes no solo (m); **PET** é a evapotranspiração potencial média anual (mm); **d** é a maior distância entre o exutório / ponto onde a estação está instalada, até a borda da bacia (m); **MPORS** é a macroporosidade do solo entre 0 e 0,6 metros de profundidade (%); e **As** é o aspecto da bacia, sendo o Norte = 1 e Sul = a 0 (adimensional).). Encontrando resultados de R^2 entre 0.95 e 0.72 comparando os valores observados e simulados para as bacias.

É importante destacar, no entanto, que as equações obtidas pelo modelo são resultado do treinamento apenas nas bacias próximas a Auckland e não devem ser aplicadas para outras regiões.

5.6 INTERPOLAÇÃO EM SISTEMAS DE INFORMAÇÕES GEOGRÁFICAS (GIS)

Chaves et al. (2002) sugeriram uma metodologia mista de regressões paramétricas e informações geográficas para a regionalização de bacias hidrográficas, identificando valores para vazões em diversos pontos de uma dada bacia que possui ao menos um posto fluviométrico.

A metodologia vislumbra quatro diferentes situações que podem ocorrer em uma bacia que possui um ou mais postos fluviométricos e que se deseja conhecer a vazão em um ponto que não possui medição: (i) Ponto de interesse a montante do posto fluviométrico; (ii) Ponto de interesse entre dois postos fluviométricos; (iii) Ponto de interesse a jusante do posto fluviométrico; e (iv) Ponto de interesse em um canal afluente cuja a foz está entre dois postos fluviométricos. A Figura 11 ilustra as situações.

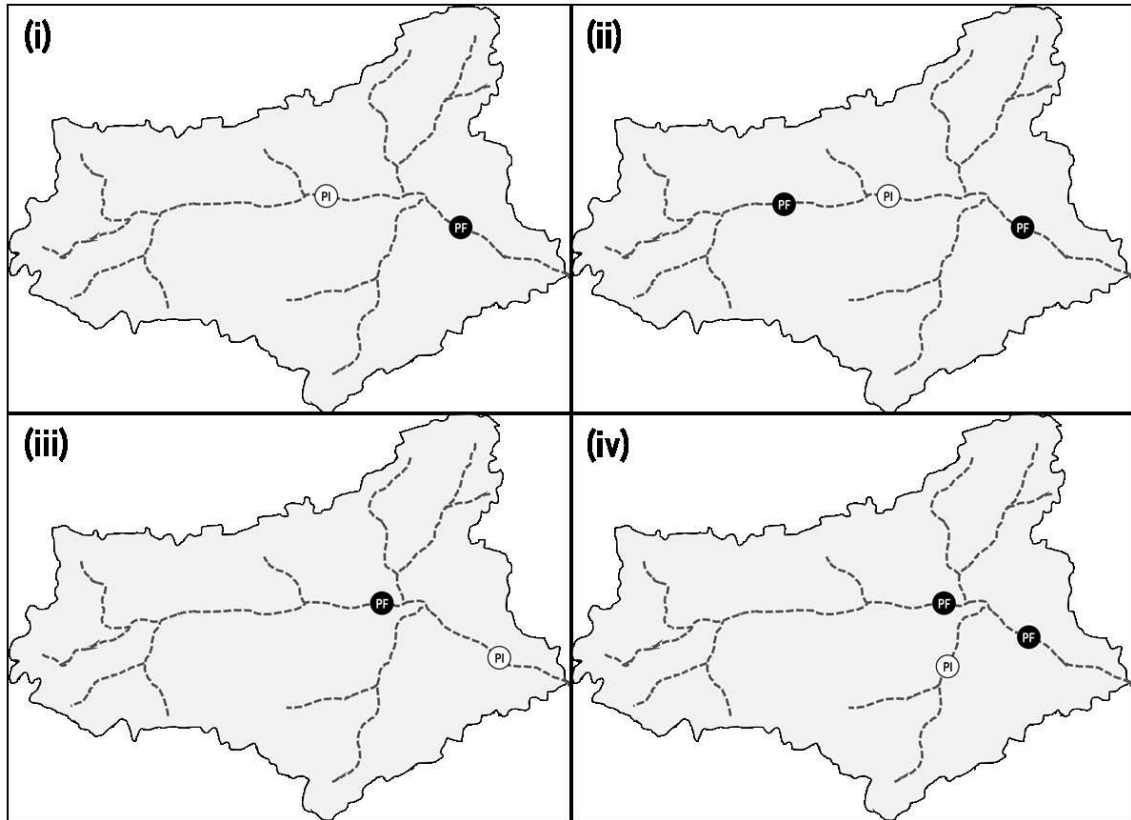


Figura 11 - Diferentes casos abordados pela metodologia de interpolação em GIS. Pontos em preto: postos fluviométricos; Pontos em branco: pontos de interesse. (Fonte: o autor)

Diferentes equações são aplicadas para cada caso, buscando valer-se ao máximo das informações geográficas disponíveis para auxiliar na precisão do resultado. Para isso, nos casos com mais de um posto fluviométrico na bacia, além de utilizar relações de área de drenagem entre os pontos, a metodologia propõe a utilização de pesos proporcionais à distância entre o ponto de interesse e os postos, quanto mais próximo do ponto de interesse, maior a influência do posto fluviométrico.

De acordo com Chaves et al. (2002) a metodologia proposta apresentou resultados mais assertivos do que os resultados obtidos para os mesmos locais utilizando a metodologia tradicional de razões entre áreas de drenagem.

6. PROPOSTA DE ALGORITMO PARA REGIONALIZAÇÃO DE DADOS FLUVIOMÉTRICOS EM ESCALA LOCAL

Diante das diversas opções de metodologias para regionalização de dados fluviométricos e da complexidade de implantação de cada uma delas, passando pelas fases de desenvolvimento, correção de erros e validação, o presente trabalho propõe inicialmente o teste de uma metodologia simplificada, limitando o número de variáveis hidrológicas analisadas com o

objetivo de criar um modelo robusto, e que possa posteriormente ser comparado a outros modelos mais complexos, desta forma será possível identificar os benefícios agregados com a introdução de mais variáveis aos cálculos de regionalização.

Como ponto de partida para a definição do algoritmo e da metodologia que será implementada, define-se os limites nos quais pretende-se atuar e as informações disponíveis dentro da área de atuação. No presente trabalho, propõe-se um algoritmo capaz de estimar dados de vazão em cursos d'água selecionados dentro do território brasileiro. Para isso serão utilizados dados hidrológicos disponibilizados pela Agência Nacional de Águas (ANA) com o objetivo de gerar séries históricas e curvas de permanência de vazões sintéticas para os locais estudados.

Os dados serão coletados através do servidor da agência em formato “xml”, onde estarão disponíveis as seguintes informações de cada bacia monitorada (ANA, 2019b):

a) Informações disponíveis para estações fluviométricas:

- I. Latitude e longitude da estação;
- II. Área de drenagem da estação;
- III. Altitude da estação;
- IV. Tipo de estação (Fluviométrica / Pluviométrica);
- V. Código da estação;
- VI. Código da bacia;
- VII. Código da sub-bacia;
- VIII. Nome do município;
- IX. Nome do estado;
- X. Nome da estação;
- XI. Nome do rio;
- XII. Data das medições disponíveis;
- XIII. Dados de vazão diários;
- XIV. Dados de nível diários;
- XV. Nível de consistência dos dados disponíveis.

b) Informações disponíveis para estações pluviométricas:

- I. Latitude e longitude da estação;
- II. Altitude da estação;
- III. Tipo de estação (Fluviométrica / Pluviométrica);
- IV. Código da estação;

- V. Código da bacia;
- VI. Código da sub-bacia;
- VII. Nome do município;
- VIII. Nome do estado;
- IX. Nome da estação;
- X. Data das medições disponíveis;
- XI. Dados de chuva diários;
- XII. Nível de consistência dos dados disponíveis.

Observando a Figura 8, fica claro que também é necessário determinar as informações disponíveis da bacia que se pretende estudar. Quanto mais informações forem disponíveis, mais opções de metodologias de regionalização de dados poderão ser aplicadas. Por outro lado, quanto mais informações forem necessárias como “entrada” para o software, mais limitada será sua aplicação, pois nem sempre será possível obter todos os dados necessários. Como o objetivo definido para o algoritmo envolve a aplicação em todo o território brasileiro e utilizando uma metodologia simplificada, optou-se por limitar as informações disponíveis da bacia estudada em:

- I. Latitude e longitude do local que se deseja estudar; e
- II. Área da bacia formada até o local que se deseja estudar.

Desta forma, define-se que o algoritmo deve ser capaz de efetuar os cálculos de regionalização para um dado local, considerando apenas estas duas informações de entrada. Essa definição impossibilita a utilização de metodologias que utilizem dados hidrológicos como parâmetro de entrada, uma vez que estes dados não são conhecidos (dados de morfologia da bacia, características de absorção de água no solo etc.) e, portanto, os dados pluviométricos disponíveis não são utilizados nesta análise. A seguir são apresentadas as considerações feitas a respeito de cada uma das metodologias estudadas, analisando a sua aplicabilidade para o objetivo desejado:

- a) Regressão paramétrica: possui aplicabilidade estabelecendo-se uma relação entre a área de drenagem e outras variáveis hidrológicas desconhecidas, calibrando-se o modelo com dados conhecidos da região do entorno do local de estudo (utilizando, por exemplo, conjuntamente com a metodologia do vizinho mais próximo);
- b) Análise de agrupamento de dados: não possui aplicabilidade direta. No entanto, a metodologia poderia ser aplicada se utilizada em conjunto com técnicas de vizinho mais

próximo, uma vez que a única informação “hidrológica” disponível como dado de entrada da bacia de interesse é a área de drenagem, não sendo suficiente para inferir a homogeneidade do local de estudo com as demais estações analisadas. Neste caso, a metodologia seria aplicada para definir qual o grupo (cluster) mais próximo do local de estudo, ou até mesmo o cluster onde o local de estudo está inserido (caso a metodologia de agrupamento leve em consideração as informações geográficas das estações);

- c) Modelo de índices: Assim como a regressão paramétrica, o modelo de índices também possui aplicabilidade estabelecendo-se uma relação entre a área de drenagem e outras variáveis hidrológicas desconhecidas, calibrando-se o modelo com dados conhecidos da região do entorno do local de estudo. No entanto, devido a limitação de informações, não se espera que o modelo performe melhor do que a metodologia de regressão paramétrica convencional;
- d) Geometria hidráulica fractal: não possui aplicabilidade direta, pois a metodologia depende de características morfológicas das bacias, que não estão sendo fornecidas neste estudo;
- e) Programação de expressão de genes: Assim como o modelo de índices, não existe um impedimento para a utilização da metodologia, no entanto, devido a limitação de informações, não se espera que o modelo performe melhor do que a metodologia de regressão paramétrica convencional;
- f) Interpolação em SIG: possui aplicabilidade sem a necessidade de dados hidrológicos adicionais.

A partir das considerações realizadas, propõe-se o algoritmo descrito na Figura 12 e detalhado em seguida.

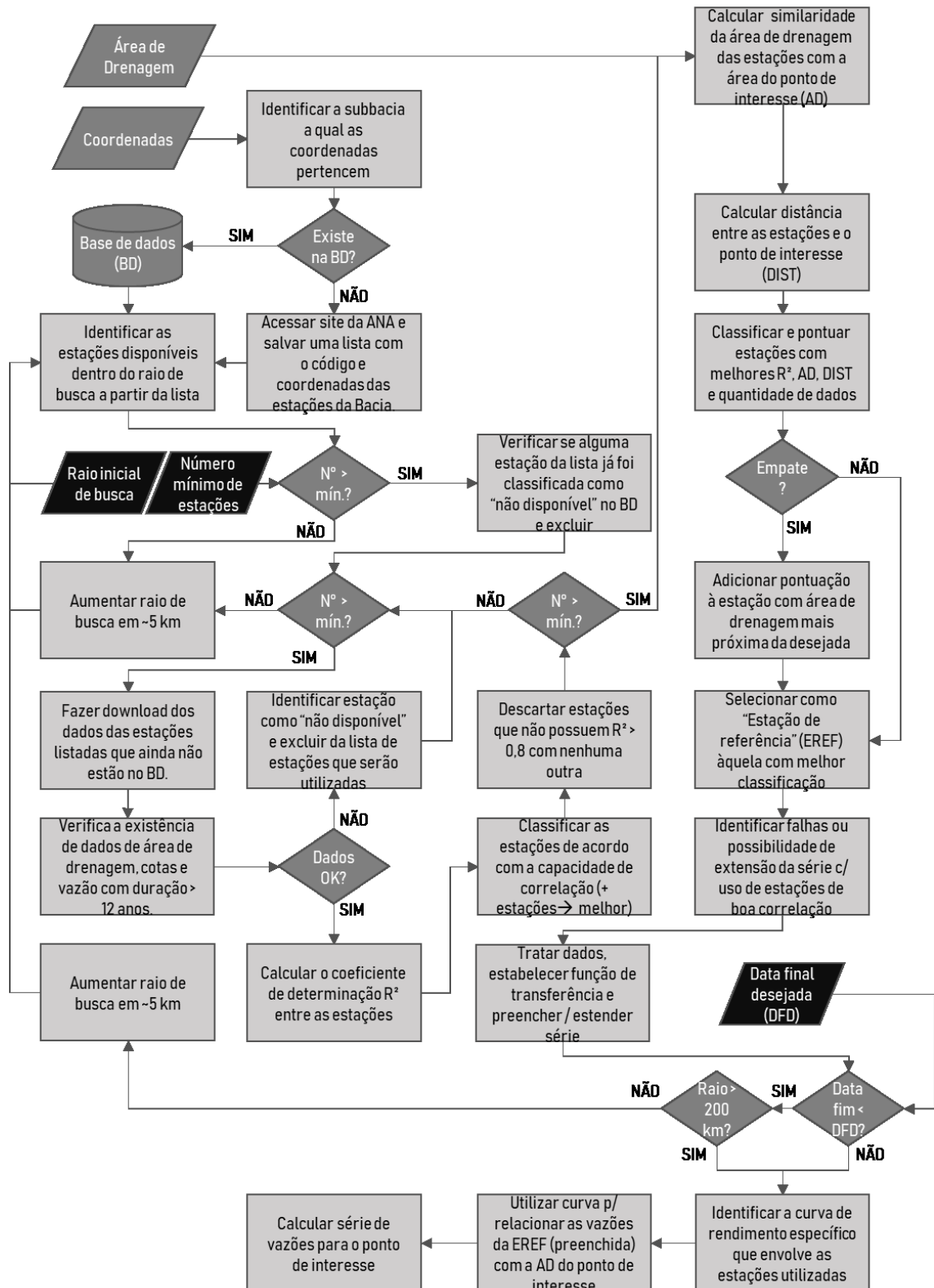


Figura 12 - Fluxograma do algoritmo proposto. (Fonte: o autor)

O algoritmo tem início com a entrada de dados de coordenadas e área de drenagem do local que se pretende estudar. A partir dos dados de coordenadas, identifica-se através de uma comparação com um mapa das poligonais que descrevem as sub-bacias brasileiras, pré-carregado no software, em qual delas se encontra o ponto inserido, essa análise é feita através de algoritmos de “ponto em polígono”. A Figura 13 ilustra o processo:

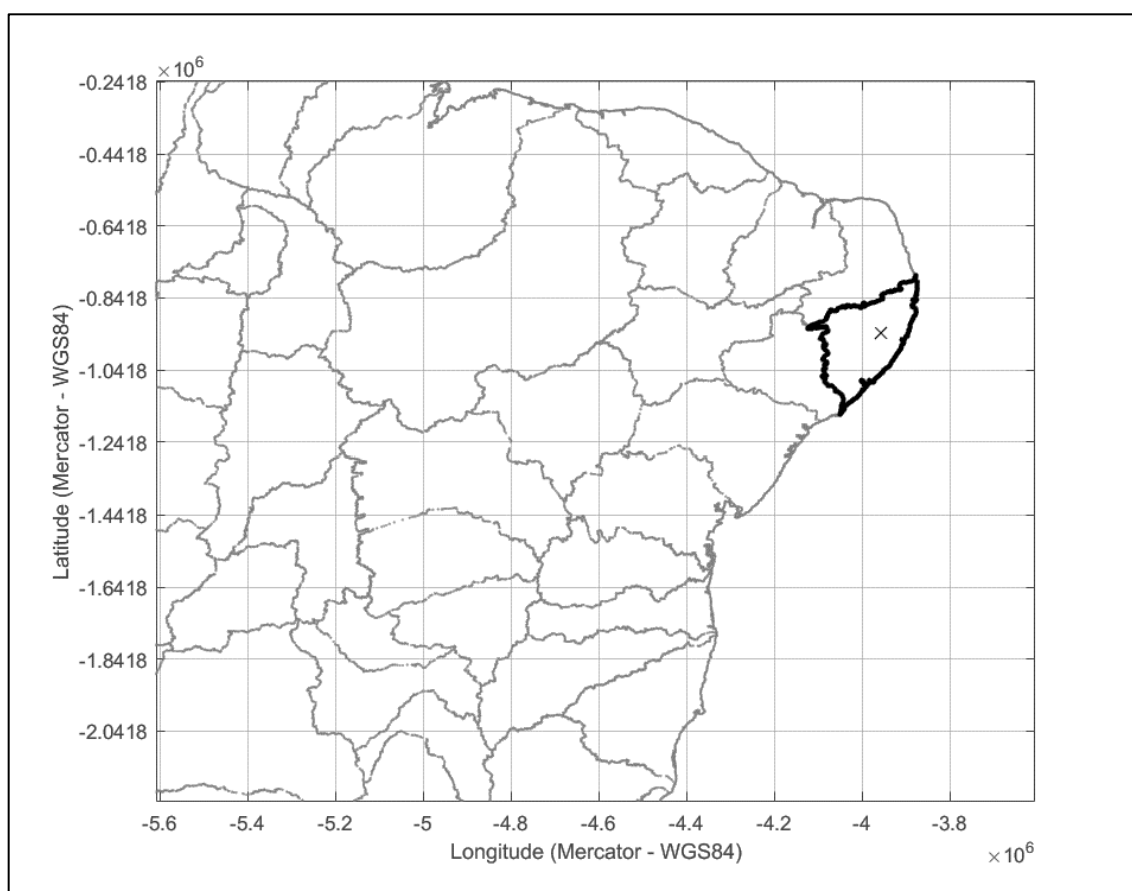


Figura 13 - Definido o local de estudo (x), o software identifica a sub-bacia a qual o ponto pertence (no exemplo da figura a sub-bacia 39 – “Paraíba e outros”). (Fonte: o autor)

Definida a sub-bacia a qual o ponto pertence, é possível limitar o esforço computacional para buscar estações hidrométricas potencialmente “doadoras” de dados, apenas dentro da área da sub-bacia. Para isso é solicitada de forma automática através de comunicação via internet com o serviço de dados da ANA uma lista de todas as estações hidrométricas pertencentes à sub-bacia em estudo. Para realizar esta requisição, o software utiliza o código da sub-bacia, que é padronizado pela ANA (ex.: 56 para a sub-bacia do Rio Doce, ou 61 para a sub-bacia do Rio Grande etc.). Antes, porém, é verificado se a sub-bacia em questão já existe no banco de dados do software, a fim de evitar a comunicação desnecessária com o banco de dados online da ANA.

A partir da lista de estações pertencentes à sub-bacia e definido um raio de busca inicial (adotado como 30 km), o software irá filtrar apenas as estações dentro deste raio, como possíveis “doadoras” de dados, comparando as coordenadas do local de interesse com o local das estações cadastrado pela ANA. Caso não sejam encontradas no mínimo quatro estações, este raio de busca é ampliado em 5 km e assim sucessivamente, até que se tenham quatro estações. Caso o raio de busca ultrapasse 1.000 km o software identifica que o local desejado não possui dados suficientes para que seja realizado o estudo e notifica o usuário com uma mensagem de erro, interrompendo a rotina.

Após conseguir um total de quatro estações para realizar os estudos, o software verifica na base de dados se alguma das estações selecionadas já foi identificada em análises anteriores como sendo uma estação que não possui dados suficientes, ou por algum motivo não poderá ser utilizada na análise. Caso exista alguma estação nessa situação, ela é excluída da análise. Se a exclusão limitar o número de estações em menos de quatro, o raio de busca é novamente ampliado em 5 km e o processo de seleção descrito no parágrafo anterior é reiniciado até que se tenha quatro estações novamente.

A partir da listagem de estações selecionadas, o software verifica quais delas já estão na base de dados (já foram utilizadas em análises passadas) e as restantes entram em uma fila para que tenham seus dados importados automaticamente através do serviço de dados da ANA. O processo de importação é composto pelas seguintes etapas: (i) são importados os dados de vazão diários das estações; (ii) são importados os dados de cotas diárias medidas nas estações; (iii) são importados os dados de área de drenagem das estações; (iv) os dados de cotas e vazões (para uma mesma data) são confrontados e é ajustada uma ou mais curvas-chave para a estação; (v) dados com desvio superior a 5% da(s) curva(s)-chave são eliminados da análise, esse valor foi escolhido por representar a faixa inferior de erros obtidos com a metodologia de medição de vazão por curva-chave, de acordo com Santos et al. (2001), no entanto, pode-se estudar o ajuste deste filtro em futuros trabalhos (Figura 14); (vi) estações com menos de doze anos de dados históricos são eliminadas e passam a integrar a lista de estações com dados insuficientes, para que nas próximas análises do software, não seja gasto tempo fazendo download dos dados (essa medida foi adotada após verificar através de experimentação que, estações com menor volume de dados normalmente apresentam erros muito elevados, provavelmente por não terem passado por muitas análises de consistência); e (vii) as demais estações têm seus dados salvos no banco de dados. Novamente, caso o número de estações “em estudo” fique abaixo de quatro, o raio de busca é ampliado em 5 km e o processo é reiniciado.

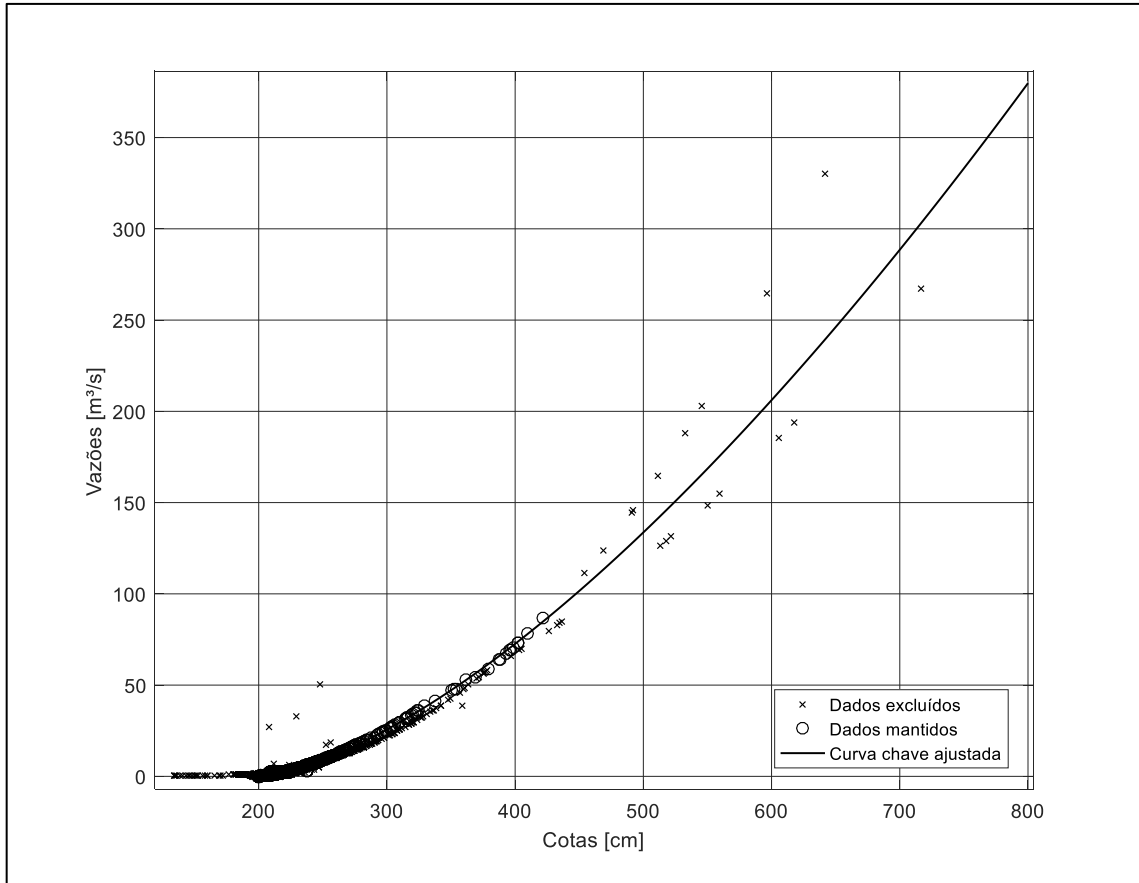


Figura 14 - Filtro de dados através do ajuste da curva-chave. (Fonte: o autor)

Com os dados das estações armazenados, são geradas séries de vazões médias mensais a partir dos dados diários e avaliadas as correlações entre as estações em estudo (Figura 15). Caso a correlação entre uma estação e outra obtiver um coeficiente de determinação (R^2) superior a 80% ambas as estações recebem uma pontuação de qualidade (Pontuação de Qualidade de Correlação - PQC). Caso uma estação não possua nenhuma PQC, ela é retirada da análise. Novamente, caso o número de estações “em estudo” fique abaixo de quatro, o raio de busca é ampliado em 5 km e o processo é reiniciado.

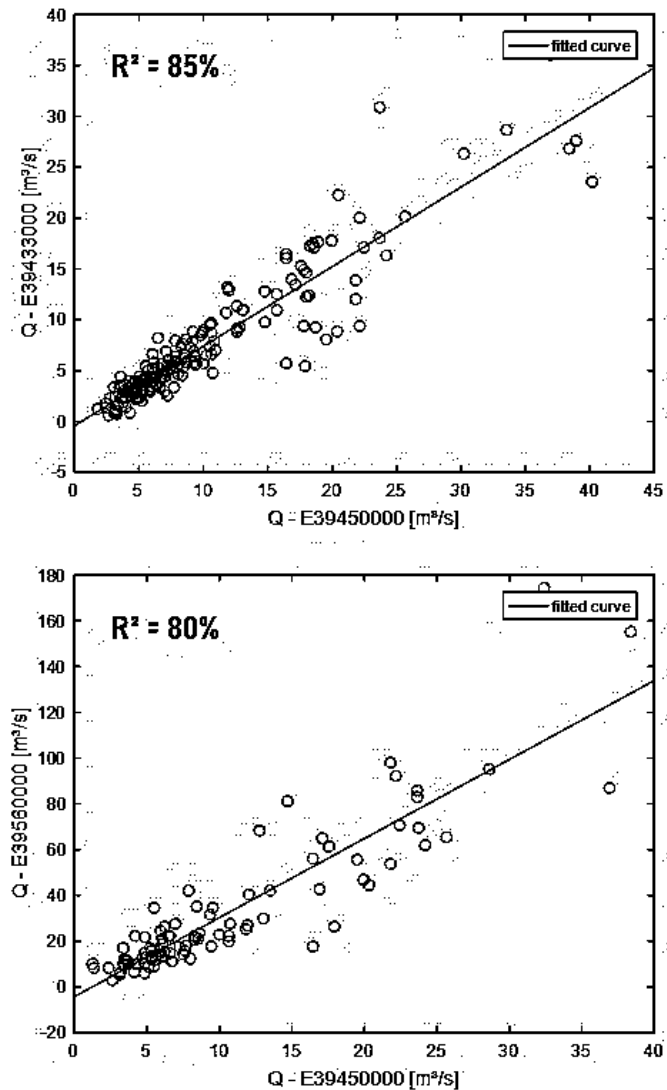


Figura 15 – Exemplos de correlações entre dados de estações fluviométricas. (Fonte: o autor)

Finalmente, com as estações selecionadas deve-se estabelecer qual delas será a estação de referência para “doação” de dados e quais delas servirão para preenchimento de falhas e extensão de série. Para tal, é estabelecido um ranking de pontuação a fim de determinar a qualidade das estações para “doação” de dados e aquela com melhor ranking é selecionada para ser a estação de referência.

O ranking avalia as seguintes características das estações e estabelece as seguintes pontuações: (i) PQC (número de estações com cuja a correlação atingiu um coeficiente de determinação superior a 80%): cada ponto equivale a um ponto no ranking, limitado a três pontos; (ii) quantidade de dados disponível na estação: até 15 anos, um ponto, entre 15 e 25 anos, dois pontos e acima de 25 anos, três pontos; (iii) diferença entre áreas de drenagem (área do local que se deseja estudar e área da estação): menor ou igual a 20%, seis pontos, entre 20% e 50%,

quatro pontos, entre 50% e 90%, dois pontos e superior ou igual a 90%, um ponto; e (iv) distância entre a estação e o ponto de interesse: até 10 km, seis pontos, até 20 km, quatro pontos e acima de 20 km, 2 pontos. Caso ocorra empate, a estação com menor diferença entre áreas de drenagem será escolhida como estação de referência. A Tabela 2 resume o sistema de classificação.

Tabela 2 – Sistema de pontuação para classificação das estações e escolha da estação de referência. (Fonte: o autor)

Parâmetro avaliado	Peso atribuído	Avaliações possíveis	Pontuação final
Número de correlações com $R^2 > 0,8$	1	Se = 1 → 1 Se = 2 → 2 Se \geq 3 → 3	1 a 3
Número de anos com dados fluviométricos	1	Se \leq 15 → 1 Se \leq 25 → 2 Se > 25 → 3	1 a 3
diferença de área de drenagem com relação ao local estudado	2 + critério de desempate	Se \geq 90% → 1 Se entre 20% e 50% → 2 Se \leq 20% → 3	2 a 6
Distância entre a estação e o local estudado	2	Se > 20 km → 1 Se entre 10 e 20 km → 2 Se \leq 10 km → 3	2 a 6
Totais	-	-	6 a 18

Definida a estação de referência, verifica-se a possibilidade de preenchimento de falhas e extensão da série de vazões. Para tal, seleciona-se a estação que possui melhor correlação com a estação de referência, define-se através de análise de regressão linear uma equação paramétrica que relaciona os dados de vazão das estações com as áreas de drenagem das mesmas e utilizando a equação, transfere-se os dados de vazão da estação para a referência. Caso ao final o processo ainda exista a possibilidade de preenchimento de falhas ou extensão da série, o processo é repetido com a próxima estação com melhor correlação, resultando em uma estação com a série de dados a mais completa possível.

Por fim, utilizando-se os dados de todas as estações selecionadas, cria-se uma curva de rendimento específico da região. A curva de rendimento específico é a representação gráfica das vazões médias de cada estação fluviométrica selecionada dividida pela área de drenagem da bacia hidrográfica daquela estação (vazão específica), os valores obtidos são dispostos em um gráfico (Figura 16) onde o eixo das abscissas é representado pelas áreas de drenagem e o eixo das ordenadas é representado pelas vazões específicas.

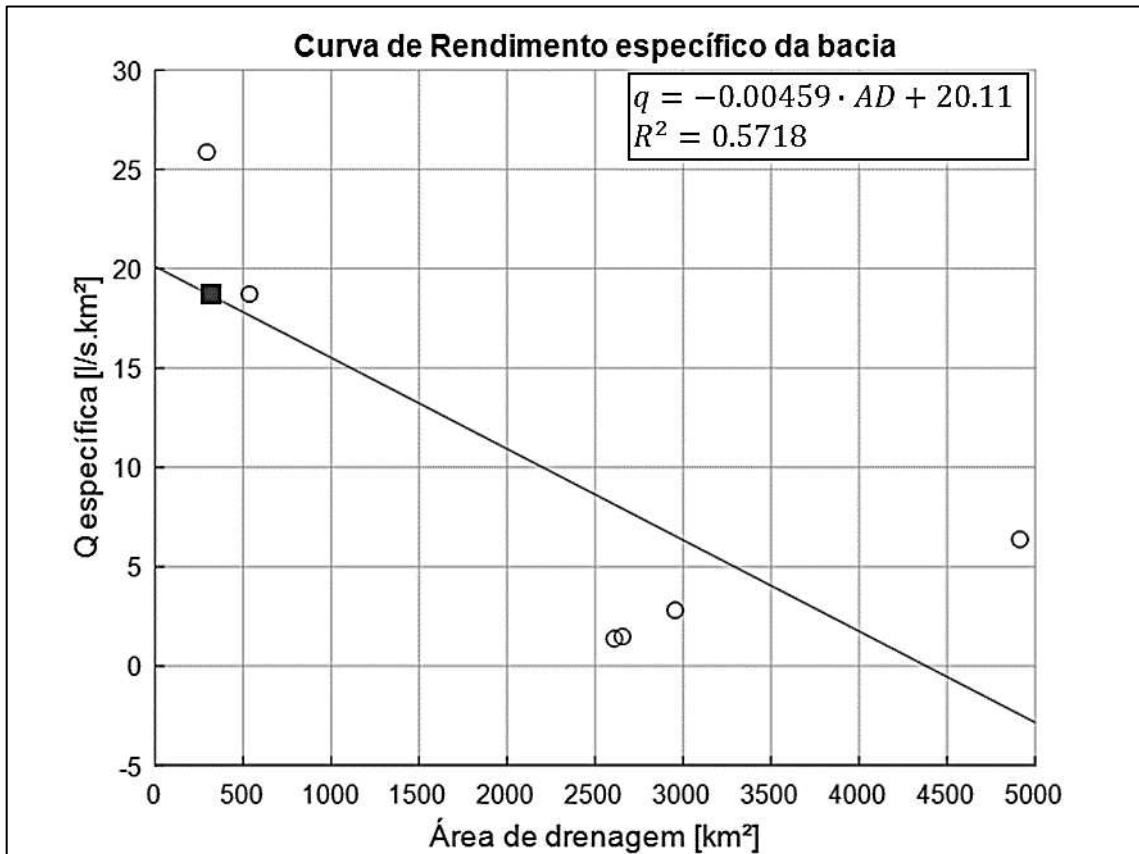


Figura 16 – Curva estimada de rendimento específico para a bacia na qual o local de estudo está inserido. Círculos: Dados das estações; Quadrado: previsão para um local genérico na bacia. (Fonte: o autor)

A curva de rendimentos específicos de uma região, no contexto do modelo proposto, tem como objetivo capturar o comportamento de diminuição da vazão por km² de uma região com o aumento da área de drenagem. Desta forma, é possível estabelecer uma relação de vazões entre a estação de referência e o local de estudo que não seja apenas diretamente proporcional à área de drenagem, mas que reflita o comportamento de vazão específica da região para a área de drenagem estudada.

$$Q_{loc_i} = Q_{ref_i} \times \frac{AD_{loc}}{AD_{ref}} \times \frac{q_{loc}}{q_{ref}} \quad (6)$$

A partir da curva, conforme explicado no parágrafo anterior, define-se o rendimento específico para o local estudado (Figura 16) e em seguida se estabelece a relação entre o rendimento específico do local e da estação de referência. Juntamente com a relação entre as áreas de drenagem dos locais, este valor é inserido em uma equação de regressão linear para estimativa dos valores de vazões do ponto de interesse (Eq. (6) – onde Q_{loc_i} é a vazão resultante para o dia “i” no local em estudo (m^3/s); Q_{ref_i} é a vazão medida na estação de referência para o dia “i” (m^3/s); AD_{loc} é a área de drenagem do local em estudo (km^2); AD_{ref} é a área de drenagem da estação de referência (km^2); q_{loc} é a vazão específica do local em estudo, encontrada conforme a Figura 16 ($m^3/s/km^2$) e q_{ref} é a vazão específica calculada para a estação de referência ($m^3/s/km^2$)). Detalhes do código em MATLAB® podem ser observados no Anexo I.

7. VERIFICAÇÃO DOS RESULTADOS E EFICÁCIA DO ALGORITMO DE REGIONALIZAÇÃO

A fim de se verificar a precisão da estimativa de descarga para um local qualquer utilizando o algoritmo proposto, foram realizadas 271 simulações em um total de 271 locais onde existem estações fluviométricas. Para tal, uma lista com os dados de coordenadas, áreas de drenagem e códigos das estações foi utilizada como entrada em uma versão ligeiramente modificada do software. Nesta versão, o primeiro passo do algoritmo é excluir os dados da estação correspondente ao local em estudo do banco de dados e, em seguida, para evitar o download dos dados novamente, o código da estação é inserido na lista de estações com dados insuficientes. A partir daí o algoritmo segue normalmente, conforme explicado no capítulo anterior, sem utilizar os dados da estação existente originalmente no mesmo local. Ao fim do processo, o código da estação é retirado da lista de estações com dados insuficientes e os dados reais da estação são comparados aos dados estimados pelo algoritmo para o mesmo local e área de drenagem.

No presente trabalho, a aplicação das séries estimadas de vazões é voltada para o cálculo hidroenergético de empreendimentos hidrelétricos, com o objetivo de determinar a Garantia Física de tais empreendimentos, portanto, a fim reduzir a complexidade da análise e possíveis erros no processo de análise dos resultados, principalmente devido a deslocamentos temporais das descargas líquidas, por exemplo: estuda-se uma estação fluviométrica na parte baixa de uma

bacia a jusante de outra estação fluviométrica que está sendo utilizada como estação de referência para o local estudado.

Na estação de montante – referência – ocorre um pico de vazão no dia “t” devido a uma pancada de chuva. No local em estudo este pico ocorreu apenas um dia depois (t + 1), por estar a jusante, mas por conta de limitações do algoritmo, este pico será representado no mesmo dia “t” para ambas as estações. Com essa situação, ao comparar os dados da estação real e dados estimados pelo software para a estação em um mesmo dia o erro provavelmente será elevado. Para evitar esse tipo de inconsistência, os erros são analisados comparando-se as curvas de permanência de vazões reais e estimadas com uma resolução temporal de 1%. As curvas de permanência são obtidas organizando-se os dados de vazão históricos em ordem decrescente, o menor (e último) valor da série representa a vazão com permanência de aproximadamente 100%, ou seja, 100% do tempo amostrado é observada uma vazão maior ou igual na série. Já o maior (e primeiro) valor da série representa a vazão com permanência de aproximadamente 1%, ou seja, aproximadamente 1% do tempo amostrado é observada uma vazão maior ou igual na série.

$$RMSE (\%) = \sqrt{\frac{1}{n} \cdot \sum_{i=1}^n \frac{q_{o_i} - q_{s_i}}{q_{o_i}}} \times 100 \quad (7)$$

$$R^2 (\%) = 1 - \frac{\sum_{i=1}^n (q_{o_i} - q_{s_i})^2}{\sum_{i=1}^n (q_{o_i} - \bar{q}_o)^2} \times 100 \quad (8)$$

As figuras de mérito utilizadas na análise dos resultados são: (i) Raiz do erro médio quadrado normalizado – a fim de verificar os erros encontrados entre o valor medido e estimado (RMSE – do inglês: *Root Mean Square Error*); e (ii) Coeficiente de determinação (R^2) – a fim de verificar se o comportamento e forma dos hidrogramas foi bem representada pelo modelo, se comparado com o real. O RMSE normalizado é calculado conforme apresentado na equação (7) e o R^2 é calculado conforme a equação (8), onde: n é o número de dados comparados, q_{o_i} é a i ésima vazão observada na curva de permanência (m^3/s) e q_{s_i} é a i ésima vazão sintética na curva de permanência (m^3/s), \bar{q}_o é a média dos dados observados (m^3/s).

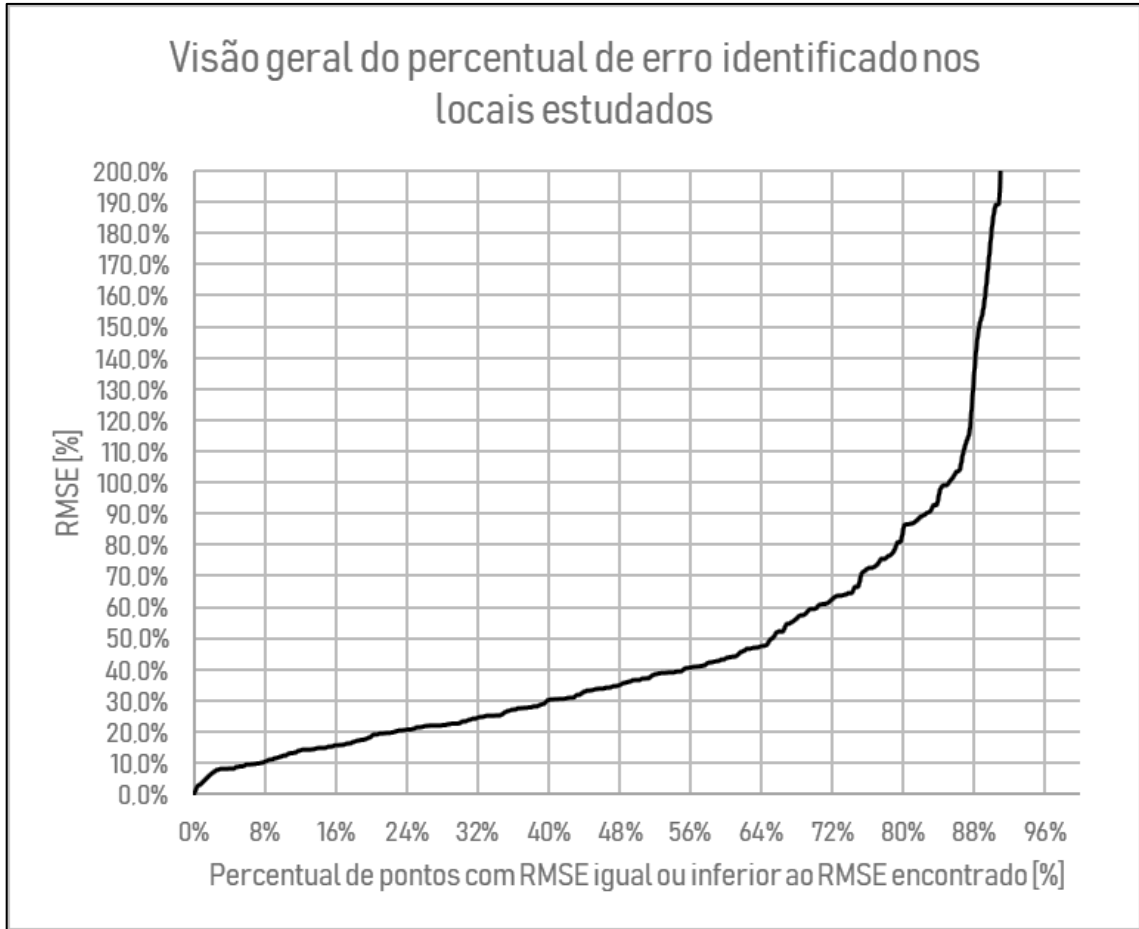
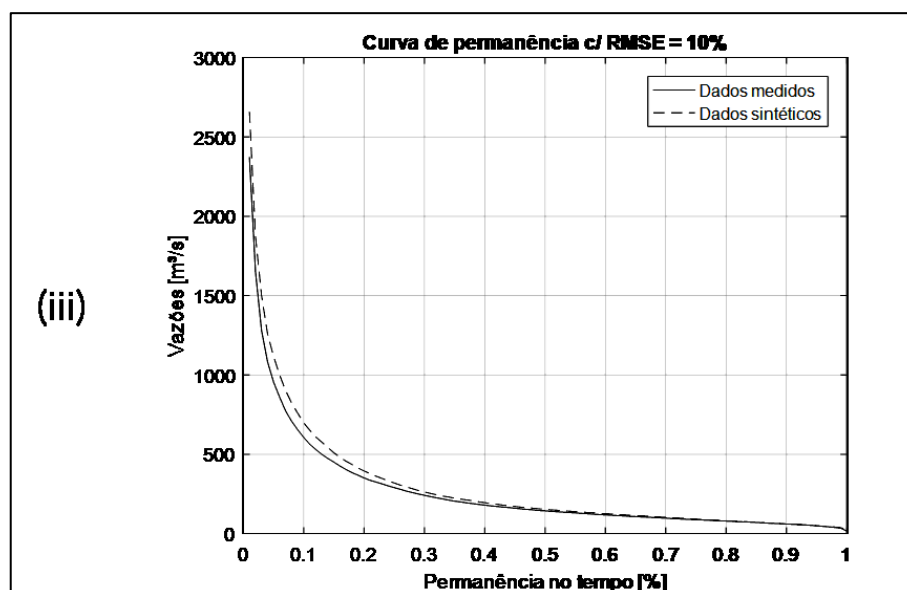
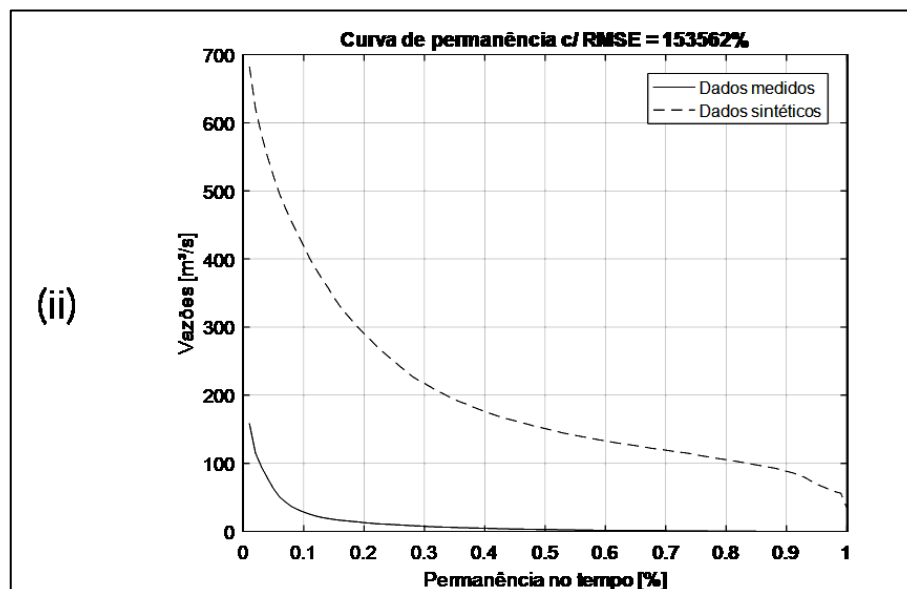
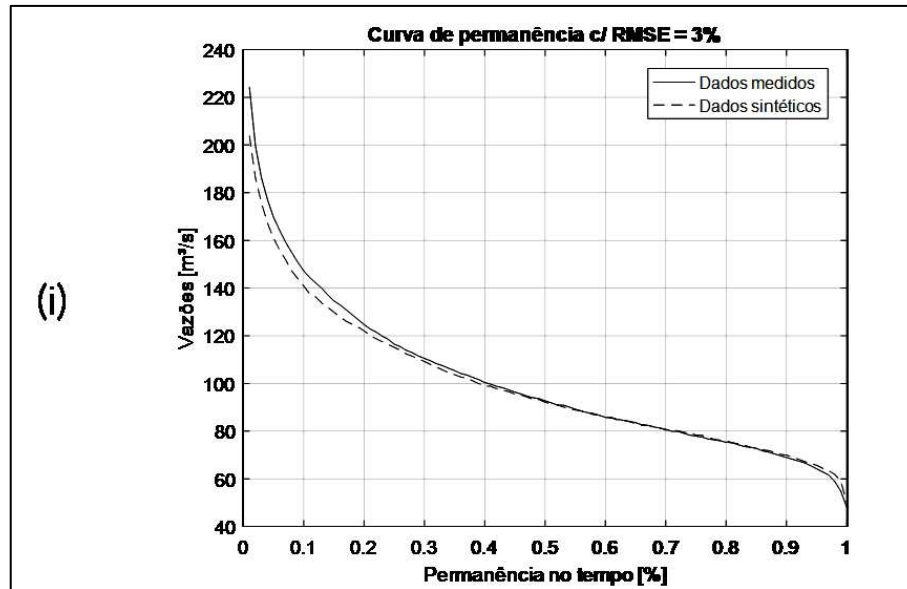
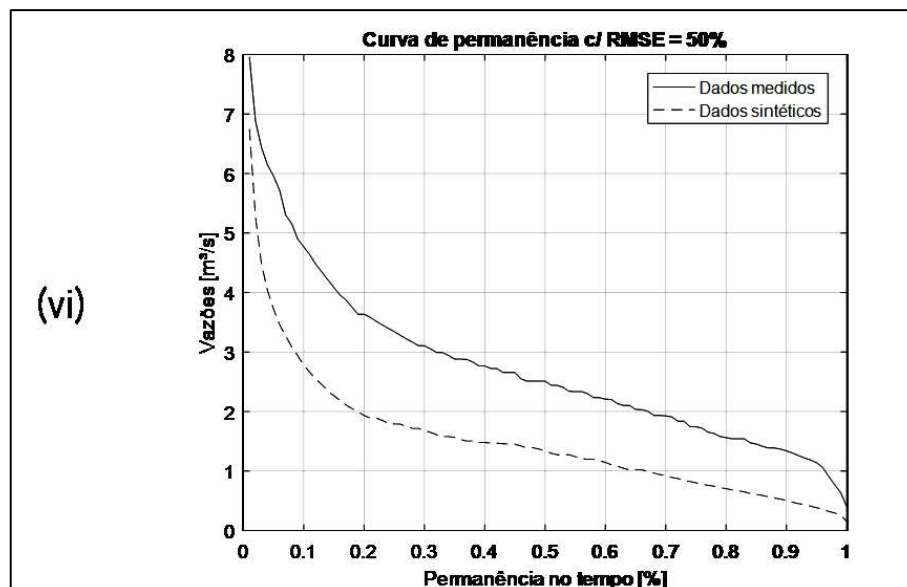
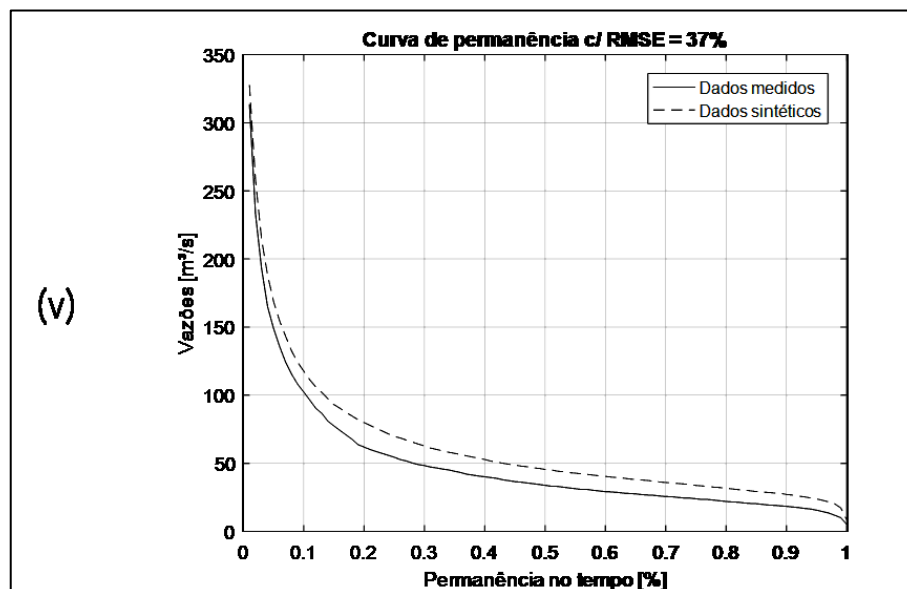
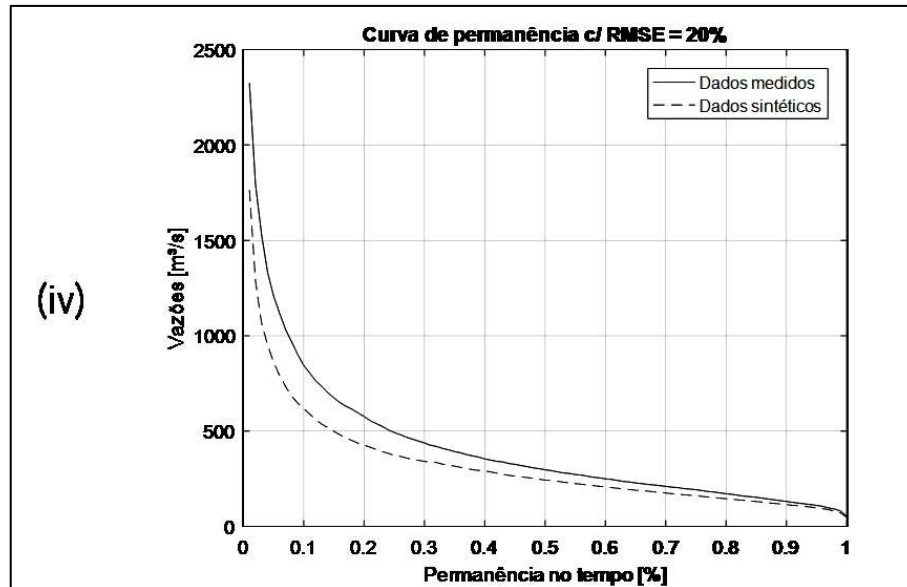


Figura 17 – Distribuição acumulada dos erros: visão geral (Fonte: o autor)

A primeira análise realizada, tem como objetivo obter uma visão geral da ordem de grandeza dos erros obtidos com a metodologia. Para isso foram comparadas as 271 curvas de permanência de vazões das estações com suas respectivas versões sintéticas (geradas a partir do modelo). Para cada curva, se obteve um valor de RMSE normalizado e R^2 .

Observando a Figura 17, verifica-se que apenas 8% dos locais obtiveram erros médios abaixo de 10%, aproximadamente 23% dos locais obtiveram erros médios abaixo de 20%, 50% dos locais obtiveram erros de até 37%, e a partir daí os erros começam a aumentar de forma mais rápida, com 80% dos locais apresentando erros de até 81% e 95% dos locais apresentando erros de até 569%, indicando que em algumas situações a metodologia alcança valores muito diferentes do medido. A Figura 18 apresenta alguns exemplos de erros nas faixas mencionadas.





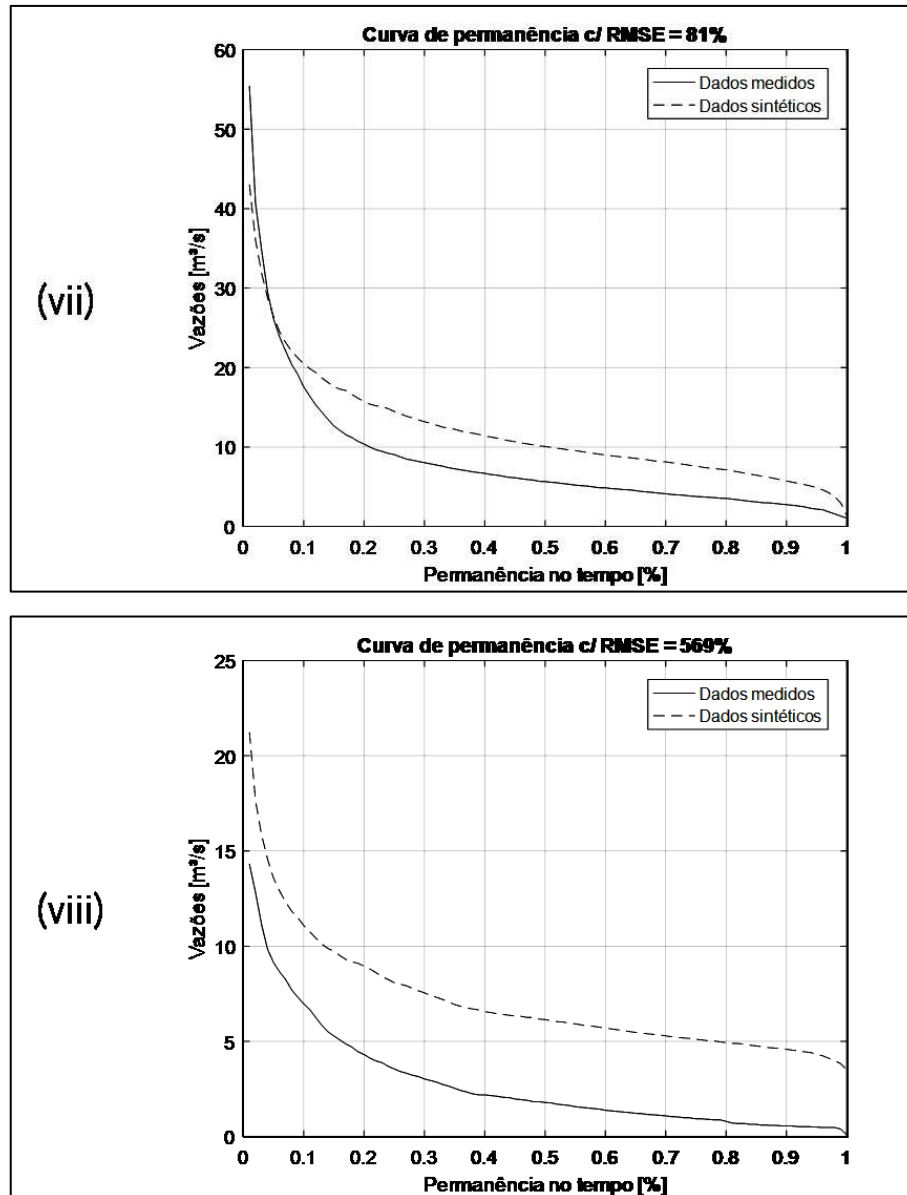


Figura 18 - Resultados de oito estações distintas: (i) melhor resultado da amostra; (ii) pior resultado da amostra; (iii) erro (RMSE) de 10%; (iv) erro de 20%; (v) erro de 37%; (vi) erro de 50%; (vii) erro de 81%; e (viii) erro de 569%. (Fonte: o autor)

A análise realizada para o RMSE foi reproduzida para obtenção do R^2 , a fim de verificar se o comportamento das versões sintéticas e medidas das curvas de permanência se correlacionam. A análise apresentou os seguintes resultados: (i) 41% dos locais obtiveram R^2 superior a 99%; (ii) 77% dos locais obtiveram R^2 superior a 95%; (iii) 87% dos locais obtiveram R^2 superior a 90%; (iv) 90% dos locais obtiveram R^2 superior a 85%; e (v) assim como ocorreu com o RMSE, a partir de certo ponto as correlações pioram significativamente, com 95% dos locais apresentando R^2 superior a 67% e 99% dos locais com apresentando R^2 superior a 35%. A Figura 19 apresenta os resultados encontrados.

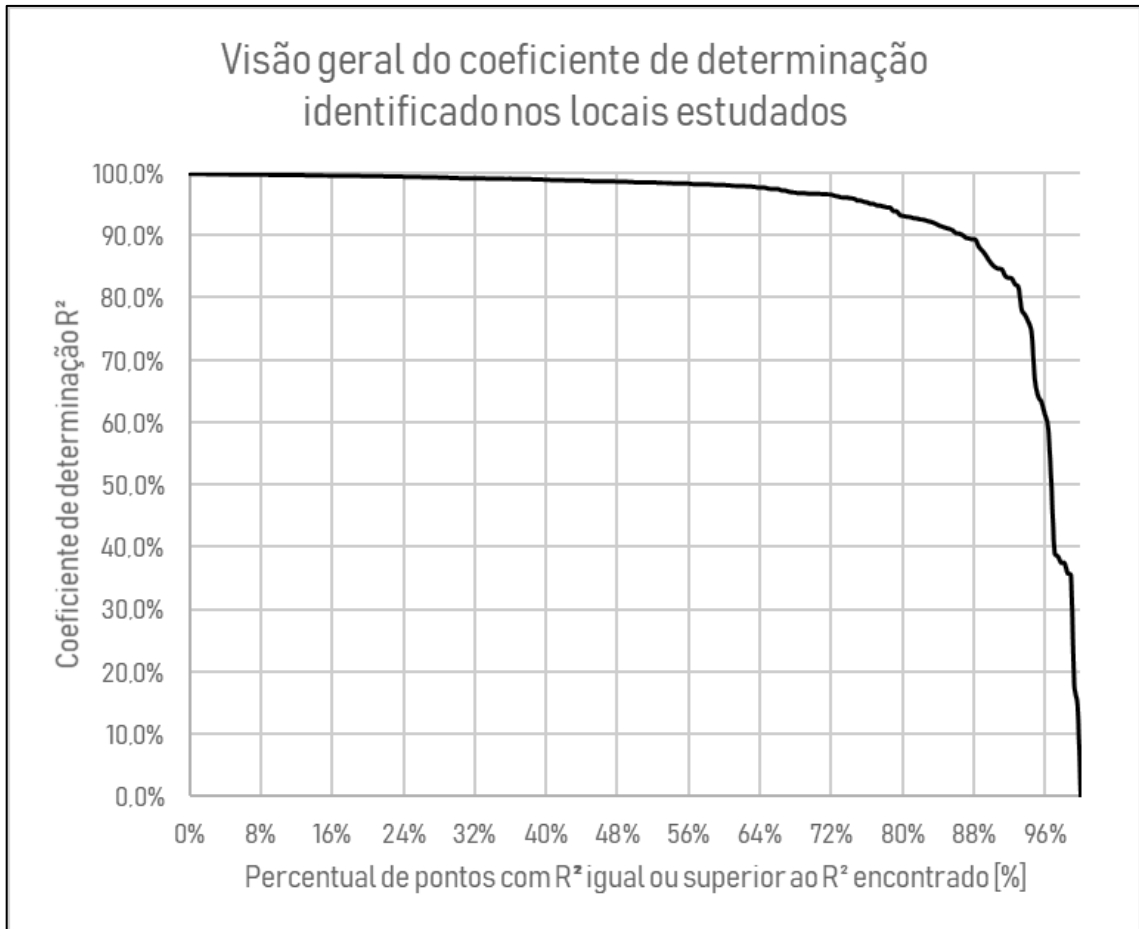
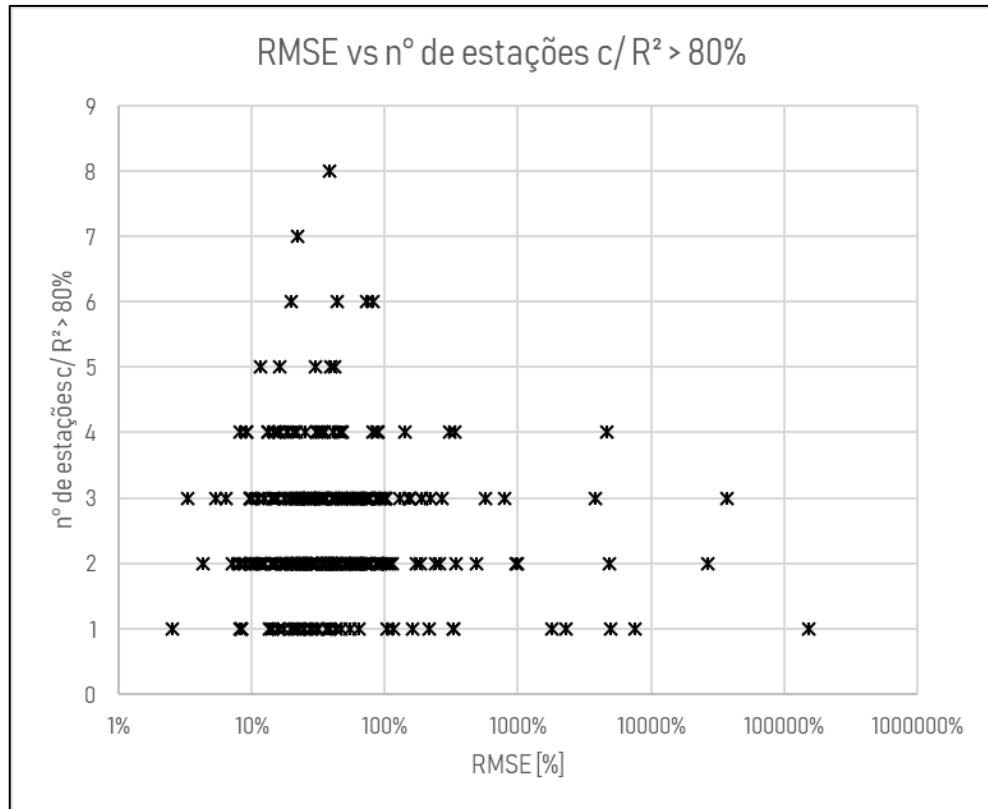


Figura 19 - Distribuição acumulada dos R²: visão geral (Fonte: o autor)

Em seguida, procurou-se identificar se alguma das variáveis utilizadas como critério de pontuação para escolha da estação de referência (Tabela 2) possui influência maior ou menor nos erros encontrados, com o objetivo de identificar possíveis alvos de melhoria na distribuição de pontos e pesos. A primeira variável analisada foi o número de correlações com coeficiente de determinação superior a 80%. Na amostra utilizada, com 271 simulações, foram identificadas a utilização de estações de referência com 1 até 8 estações com R² superior a 80%. No entanto, como pode ser observado na Figura 20, o número de estações com alto nível de correlação, não garante, necessariamente, um resultado com baixos valores de RMSE.



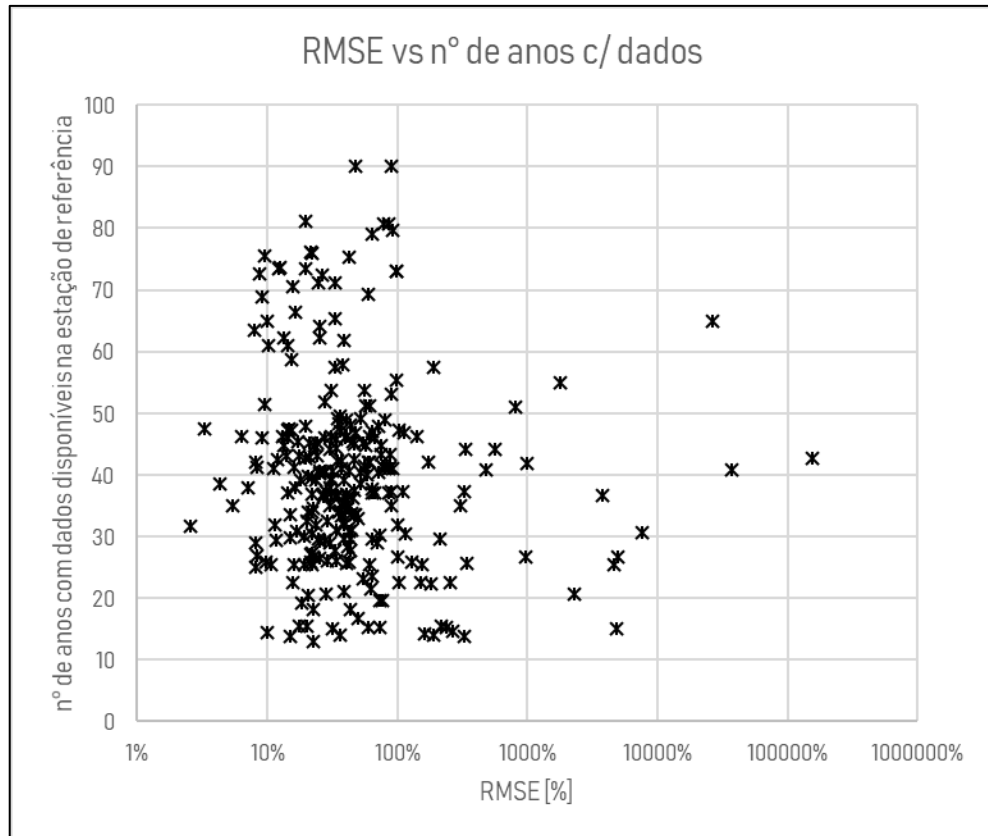
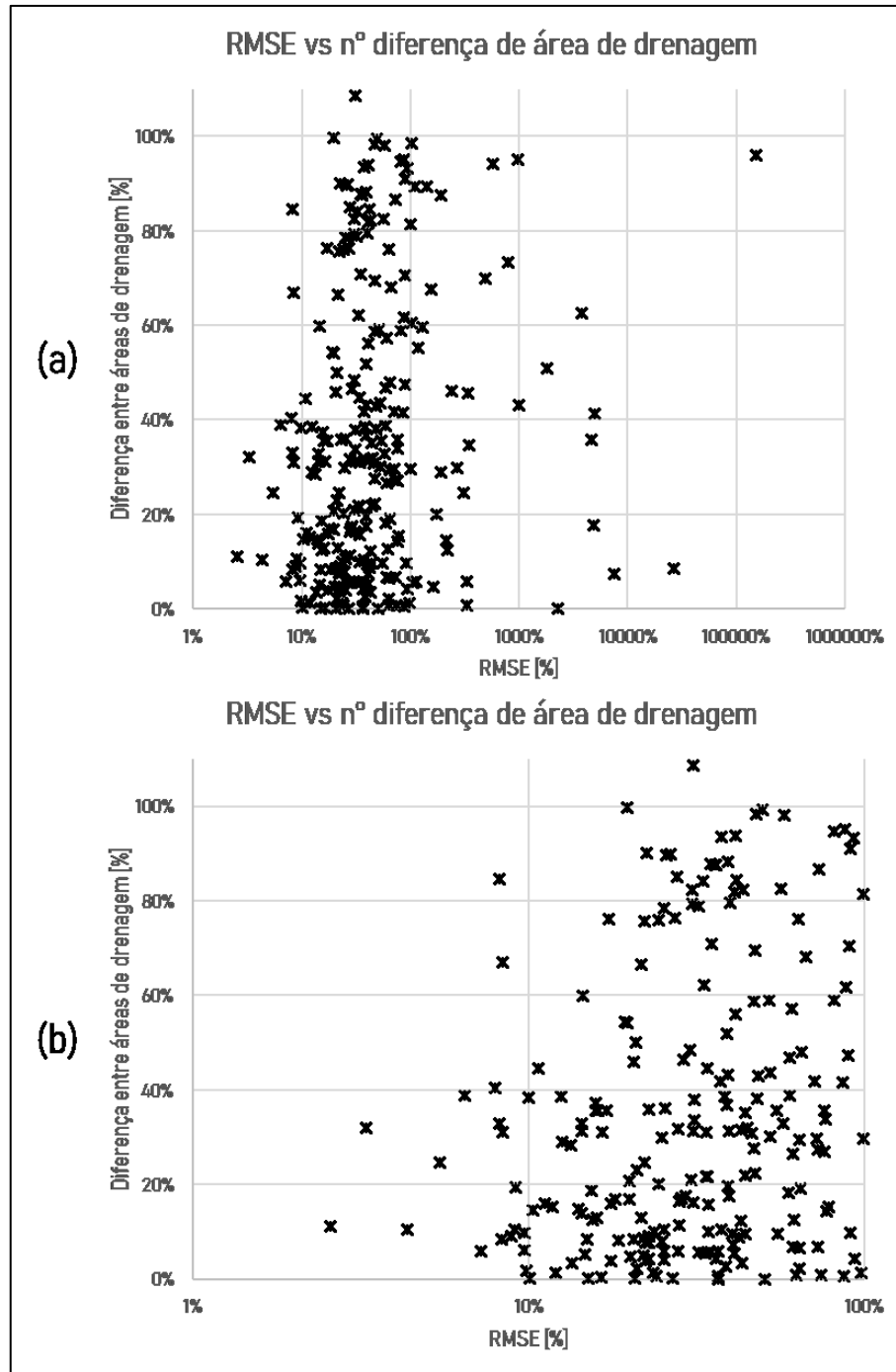


Figura 21 - RMSE vs nº de anos c/ dados disponíveis na estação de referência. (Fonte: o autor)

A terceira variável analisada é a diferença de área de drenagem com relação ao local estudado. Esta variável foi ajustada com o dobro do peso no sistema de qualificação por classificação, exatamente por existir a expectativa de que a variável tenha maior influência na qualidade do modelo. No entanto, analisando a Figura 22 (a), não se consegue estabelecer uma relação clara entre a diferença entre áreas de drenagem e o RMSE. Na Figura 22 (b) verifica-se uma leve tendência de aumento do erro com o aumento da diferença entre as áreas, mas nada conclusivo. Na Figura 22 (c) verifica-se que não existem RMSE na faixa de até 10% com diferença de área de drenagem superior a 85%, sugerindo alguma influência da variável na qualidade do modelo.



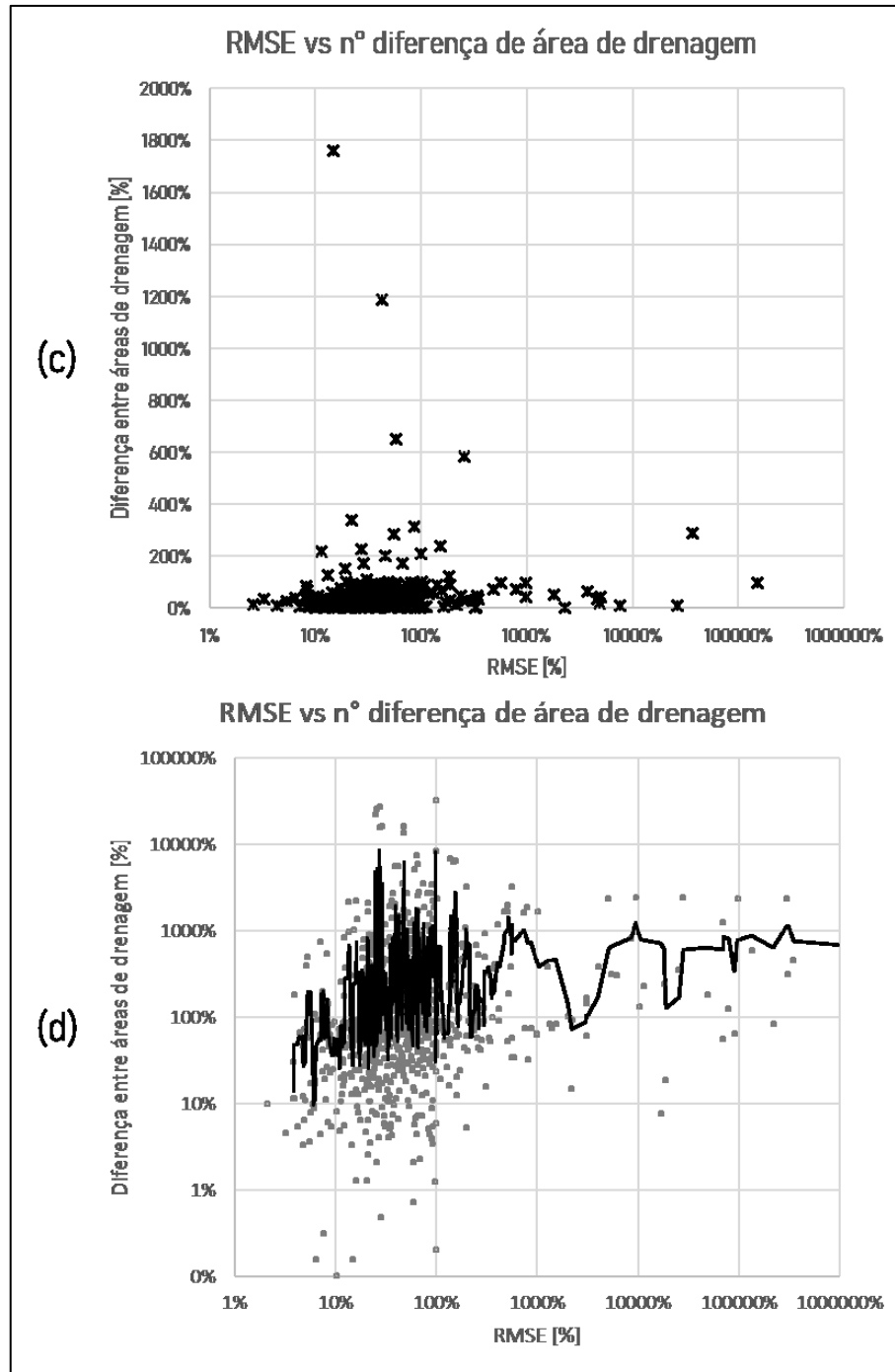


Figura 22 - RMSE versus diferença entre as áreas de drenagem. (Fonte: o autor)

Para excluir a hipótese de que a sensibilidade da análise estaria prejudicada por conta da utilização de múltiplas estações em uma mesma curva de permanência (tanto para preenchimento quanto para extensão das séries) com diferentes áreas de drenagem, realizou-se uma nova análise, separando os resultados de acordo com a estação utilizada como referência para cada período da série histórica. O resultado desta análise pode ser visualizado na Figura 22 (d), em que se percebe uma ligeira tendência de diminuição do RMSE com a diminuição da

diferença entre as áreas de drenagem (uma linha de tendência com a média móvel de 5 períodos foi adicionada para destacar o comportamento). No entanto, ao mesmo tempo que existem erros inferiores a 10% com diferença de áreas de drenagem de menos de 1%, existem erros inferiores a 10% com diferenças de áreas de drenagem superior a 500%, tornando a análise inconclusiva.

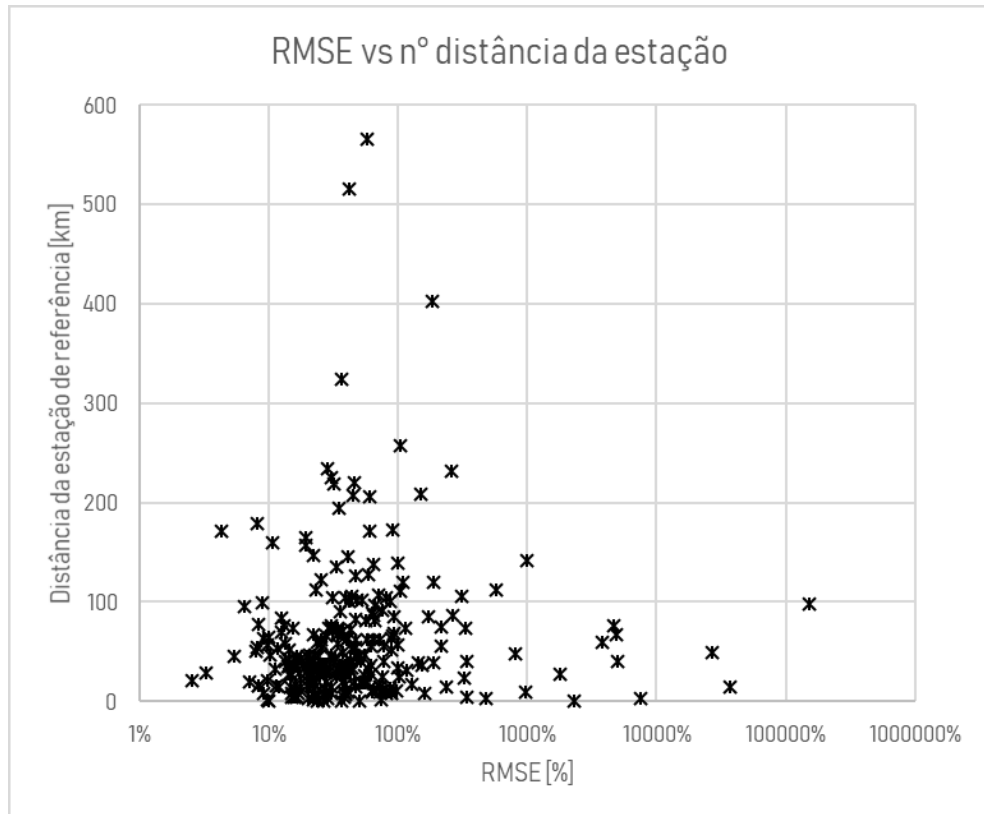


Figura 23 – RMSE versus distância da estação de referência. (Fonte: o autor)

A quarta variável analisada é a distância entre as estações doadoras e o local de estudo. Assim como a diferença de área de drenagem, o peso dado para esta variável no sistema de qualificação por classificação foi o dobro das demais. Essa variável é utilizada com frequência na metodologia de “vizinho mais próximo” para regionalização de dados fluviométricos. Observando a Figura 23, nota-se (assim como ocorreu para as demais variáveis analisadas) que não é possível determinar uma correlação linear entre a diminuição da distância e do RMSE, no entanto, também verifica-se que erros abaixo de 10% só ocorrem para distâncias de no máximo 180 km.

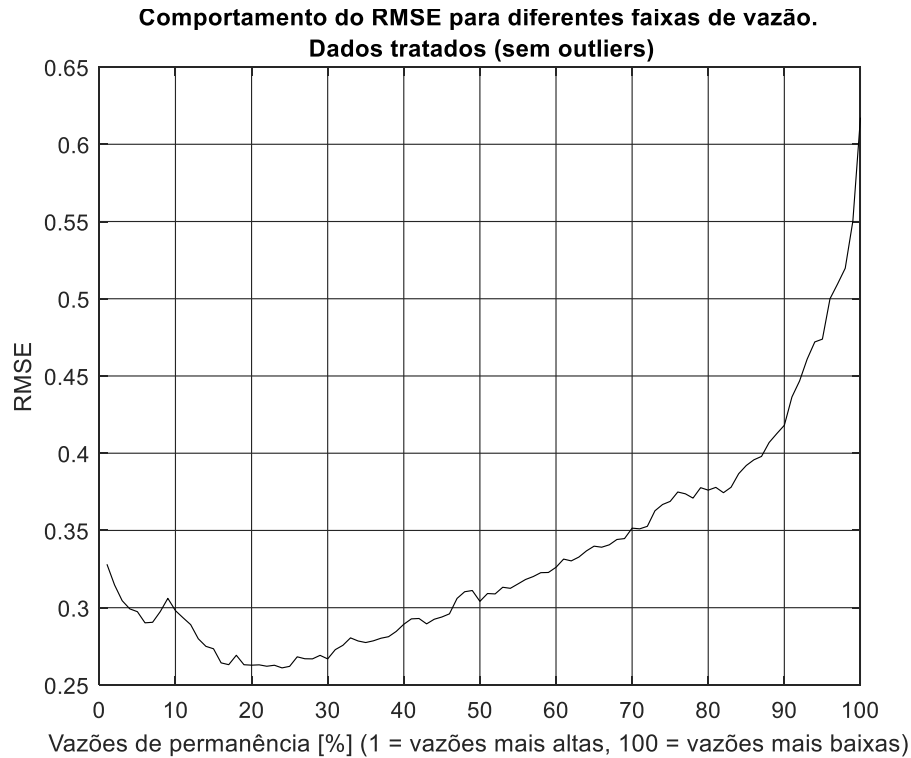


Figura 24 - Comportamento do RMSE para diferentes faixas de vazão. (Fonte: o autor).

Por fim, verifica-se os erros médios por faixa de vazão. Conforme apontam McMillan et al. (2012), para diferentes faixas de vazão, são encontrados diferentes níveis de erro, principalmente para metodologias que utilizam curvas-chave para determinar as descargas. A Figura 24 apresenta os valores médios de RMSE para cada faixa de vazão da curva de permanência de vazões das 271 estações analisadas. Para facilitar a visualização e possibilitar a análise, os dados considerados “*outliers*” foram eliminados. Observando a figura, torna-se evidente que os maiores erros tendem a ocorrer em níveis mais baixos de vazão.

8. CONCLUSÕES

O algoritmo proposto buscou simplificar ao máximo a entrada de dados para produzir uma estimativa de curva de permanência de vazões para um dado local, partindo do princípio da parcimônia. Observando os resultados obtidos com elevados coeficientes de determinação (R^2), fica claro que o modelo conseguiu reproduzir a forma e comportamento das curvas de permanência originais, com elevados valores de R^2 . No entanto, apesar dos bons resultados obtidos para aproximadamente 40% da amostra de 271 postos fluviométricos, foram observados erros elevados entre os dados medidos e estimados nos 60% restantes, o que demonstra que, apesar de existir uma correlação forte entre os dados, há uma distância significativa entre eles que deve ser explicada por uma ou mais variáveis que não foram

consideradas no modelo. Esses erros são mais acentuados em vazões de maior permanência (mais baixas), que normalmente são fortemente dependentes de características do aquífero de cada bacia. Cabe informar que existe a possibilidade de erros nos dados das estações utilizadas e até mesmo nos dados observados das estações que foram selecionadas para estudo, e que parte dos dados utilizados das estações, por serem mais recentes, podem não ter sido consistidos pela ANA e podem apresentar erros de maior magnitude.

Em um primeiro momento, com o objetivo de minimizar os erros encontrados, buscou-se identificar pontos de melhoria no sistema de qualificação por classificação, para a escolha de postos fluviométricos de referência que produzissem melhores resultados. No entanto, nenhuma das variáveis candidatas (quantidade de estações com correlação, n° de anos na amostra, relação entre áreas de drenagem e distância entre as estações) apresentou potencial para reduzir os erros do modelo.

Por existir uma forte correlação entre os dados sintéticos e reais, fica claro que é possível se chegar uma solução para minimizar os erros, introduzindo mais variáveis no modelo, a fim de explicar com maior precisão os fenômenos que impactam a disponibilidade hídrica de cada bacia. A mais forte variável candidata para aperfeiçoamento do modelo é a pluviosidade média anual, por ser uma grandeza com ampla disponibilidade de dados e alta correlação com a vazão, a pluviosidade média anual é uma das variáveis mais utilizadas na bibliografia para regionalização de vazões. No entanto, é importante destacar que ainda existem alguns ajustes que podem ser aplicados ao modelo atual que, teoricamente, também poderiam melhorar de alguma forma os resultados.

Entre os pontos de melhoria, destaca-se a etapa de criação da curva de rendimento específico da bacia, que atualmente é limitada a uma aproximação por polinômio de primeiro grau. Em trabalhos futuros, pretende-se implementar outras formas de aproximação, entre elas a interpolação da vazão específica entre os dois pontos mais próximos (com área imediatamente superior e inferior) do local em estudo.

Outro ponto de melhoria identificado é a possibilidade de utilização da metodologia de interpolação em GIS aplicada à curva de rendimento específico para filtrar a utilização de estações próximas, mas em bacias distintas (paralelas) àquela em análise. Mas para isso, seria necessário introduzir a capacidade de identificação do ponto de interesse em relação a um mapa hidrológico pré-carregado.

Por fim, é importante destacar que, a despeito dos erros identificados, não é descartada a possibilidade de utilização do modelo, uma vez que em aproximadamente 40% dos resultados as diferenças entre os dados simulados e reais não ultrapassam os 30%. Adicionalmente, os erros médios encontrados na Figura 24, apresentam comportamento e magnitude similar ao que foi reportado por McMillan et al. (2012) para medições de vazão através de curvas-chave (que é o método mais utilizado para medição de vazões no Brasil) e, portanto, os erros médios encontrados estariam dentro da faixa de incerteza do método de medição.

REFERÊNCIAS BIBLIOGRÁFICAS

ABDULLA, Fayez A; LETTENMAIER, Dennis P. Development of regional parameter estimation equations for a macroscale hydrologic model. **Journal of Hydrology**, [s. l.], v. 197, n. 1-4, p. 230-257, 1 out. 1997.

AGÊNCIA NACIONAL DE ÁGUAS (Brasília, DF). **Hidroweb**: Baixar Inventário. 2019. Disponível em: <<http://www.snirh.gov.br/hidroweb/publico/apresentacao.jsf>>. Acesso em: 13 julho 2019a.

AGÊNCIA NACIONAL DE ÁGUAS (Brasília, DF). **Orientações para consistência de dados fluviométricos**. 1. ed. Brasília: ANA, SGH, 2012. 19 p.

AGÊNCIA NACIONAL DE ÁGUAS (Brasília, DF). **ServiceANA**. 2019. Disponível em: <<http://telemetriaws1.ana.gov.br/ServiceANA.aspx>>. Acesso em: 13 julho 2019b.

AGÊNCIA NACIONAL DE ENERGIA ELÉTRICA (Brasília, DF). **Bem-vindo à ANEEL**. 2019. Disponível em: <<http://www.aneel.gov.br/a-aneel>>. Acesso em: 13 julho 2019.

AGÊNCIA NACIONAL DE ENERGIA ELÉTRICA (Brasília, DF). **BIG** – Banco de Informações de Geração. 2019. Disponível em: <<http://www2.aneel.gov.br/aplicacoes/capacidadebrasil/capacidadebrasil.cfm>>. Acesso em: 5 junho 2019.

BRASIL. Ministério de Minas e Energia (Brasília, DF). **Portaria nº 101, de 22 de março de 2016**. Define a metodologia de cálculo da garantia física de energia de novos empreendimentos de geração de energia elétrica do Sistema Interligado Nacional – SIN. Brasília: 2016.

BRASIL. Ministério de Minas e Energia (Brasília, DF). **Portaria nº 463, de 3 de dezembro de 2009**. Estabelece a metodologia para o cálculo e revisão dos montantes de garantia física de energia de Usinas Hidrelétricas não despachadas centralizadamente pelo Operador Nacional do Sistema Elétrico - ONS. Brasília: 2009.

BRASIL. Ministério de Minas e Energia (Brasília, DF); CEPEL (Rio de Janeiro, RJ) **Manual de Inventário Hidroelétrico de Bacias Hidrográficas**: edição 2007. Rio de Janeiro: E-papers, 2007. 684 p. ISBN 978-85-7650-137-4.

BRASIL. Ministério de Minas e Energia (Brasília, DF); EMPRESA DE PESQUISA ENERGÉTICA (Rio de Janeiro, RJ). **NOTA TÉCNICA PR 04/18** – Potencial dos Recursos Energéticos no Horizonte 2050. Rio de Janeiro: 2018. 184 p.

BRASIL. Ministério de Minas e Energia (Brasília, DF); EMPRESA DE PESQUISA ENERGÉTICA (Rio de Janeiro, RJ). **Plano Nacional de Energia 2030: Geração Hidrelétrica**. Brasília: Eletrobrás, 2007.

BURN, Donald H. Cluster Analysis as Applied to Regional Flood Frequency. **Journal of Water Resources Planning and Management**, [s. l.], v. 115, n. 5, p. 567-582, set. 1989.

CASTELLARIN, Attilio *et al.* Regional flow-duration curves: reliability for ungauged basins. **Advances in Water Resources**, [s. l.], v. 27, n. 10, p. 953-965, out. 2004.

CCEE. **Regras de Comercialização: Garantia Física**. Versão 2019.1.0. São Paulo: 2019.

CEPEL. **NEWAVE** - Modelo de Planejamento da Operação de Sistemas Hidrotérmicos Interligados de Longo e Médio Prazo. 2019. Disponível em: <
http://www.cepel.br/pt_br/produtos/newave-modelo-de-planejamento-da-operacao-de-sistemas-hidrotermicos-interligados-de-longo-e-medio-prazo.htm>. Acesso em: 10 agosto 2019.

CHAVES, Henrique Marinho Leite *et al.* Regionalização de Vazões Mínimas em Bacias Através de Interpolação em Sistemas de Informação Geográfica. **Revista Brasileira de Recursos Hídricos**. [s. l.]. Vol. 7. n.3 jun/set 2002, pág. 43-551.

GOMES, Júlio; SANTOS, Irani dos. Análise comparativa de campanhas de medição de descarga líquida pelos métodos convencional e acústico. *In*: Simpósio Brasileiro de Recursos Hídricos, 13., 1999, Belo Horizonte. **Anais [...]**. Belo Horizonte: ABRH, 1999.

GRIGG, D. B. Regions, models and classes. *In*: CHORLEY, Richard J.; HAGGETT, Peter. **Models in geography**. 2. ed. Londres: Methuen, 1967.

HASHMI, Muhammad Z.; SHAMSELDIN, Asaad Y. Use of Gene Expression Programming in regionalization of flow duration curve. **Advances in Water Resources**, [s. l.], v. 68, p. 1-12, jun. 2014. DOI <https://doi.org/10.1016/j.advwatres.2014.02.009>. Disponível em: <https://www.sciencedirect.com/science/article/abs/pii/S0309170814000323?via%3Dihub>. Acesso em: 6 out. 2019.

HASHMI, Muhammad Z.; SHAMSELDIN, Asaad Y. Use of Gene Expression Programming in regionalization of flow duration curve. **Advances in Water Resources**, [s. l.], v. 68, p. 1-12, jun. 2014.

HERSCHY, Reginald W. **Streamflow Measurement**. 3. ed. Abingdon: Taylor & Francis, 2009. ISBN 0-203-93139-4.

HUANG, Hening. Estimation of Type A uncertainty of moving-boat ADCP streamflow measurements. In: CONSTANTINESCU, George *et al.* **River Flow 2016**. 1. ed. Londres: CRC Press, 2016. cap. Special Session, p. 696-701. ISBN 978-1-138-02913-2.

HUANG, Hening. Uncertainty Model for In Situ Quality Control of Stationary ADCP Open-Channel Discharge Measurement. **Journal of Hydraulic Engineering**, [s. l.], v. 138, n. 1, p. 4-12, 1 jan. 2012.

KELMAN, Jerson. **Geração de Séries Sintéticas de Vazões**. In: Curso de Engenharia Hidrológica. EPUSP, DAEE, ABRH, São Paulo, 1983. p. 1-32.

LI, Ming *et al.* A new regionalization approach and its application to predict flow duration curve in ungauged basis. **Journal of Hydrology**, [s. l.], v. 389, Issues 1-2, p. 137-145, 28 jul. 2010.

LINSLEY, Ray K.; KOHLER, Max Adam; PAULHUS, Joseph L. H. **Hydrology for Engineers**. 2. ed. Michigan, EUA: McGraw-Hill, 1975. 482 p. ISBN 9780070379671.

MCMILLAN, Hilary *et al.* Benchmarking observational uncertainties for hydrology: rainfall, river discharge and water quality. **Hydrological Processes**, [s. l.], v. 26, p. 4078-4111, 30 abr. 2012.

MINISTÉRIO DO DESENVOLVIMENTO REGIONAL (Brasil). AGÊNCIA NACIONAL DE ÁGUAS. **ServiceANA**. [S. l.], 2019. Disponível em: <http://telemetriaws1.ana.gov.br/ServiceANA.asmx>. Acesso em: 8 out. 2019.

MIRANDA, Adriana de. **Water Architecture in the lands of Syria: the Water Wheels**. Roma: «L'ERMA» di BRETSCHNEIDER, 2007. 376 p., 24 cm. (Studia Archaeologica, 156). Bibliografia: p. 39-58. ISBN 88-8265-433-8.

MOSLEY, M. P. Delimitation of New Zealand hydrologic regions. **Journal of Hydrology**, [s. l.], v. 49, n. 1-2, p. 173-192, jan. 1981.

MUELLER, D. S. Field evaluation of boat-mounted acoustic Doppler instruments used to measure streamflow. *In: IEEE/OES WORKING CONFERENCE ON CURRENT MEASUREMENT TECHNOLOGY*, 7th., 2003, San Diego, CA, USA, USA. **Proceedings** [...]. [S. l.: s. n.], 2003.

OBERG, Kevin; MUELLER, David S. Validation of Streamflow Measurements Made with Acoustic Doppler Current Profilers. **Journal of Hydraulic Engineering**, [s. l.], v. 133, n. 12, p. 1421-1432, 1 dez. 2007.

OLIVEIRA, Edimar J. de et al. Influência da variação da produtividade das usinas hidrelétricas no cálculo da energia firme. **Revista Controle & Automação**, Natal, v. 20, n. 2, p. 247-255, abr. 2009.

OLIVEIRA, Fabiano Salomão de et al. Determinação do período crítico do sistema interligado nacional. In: Seminário nacional de produção e transmissão de energia elétrica, 18., 2005, Curitiba, 16 out. 2005. **Informes técnicos**. [s.l.]: 2005. GPL-08.

PEREIRA, Marco Alésio Figueiredo *et al.* Regionalização com geometria hidráulica e fractal: estudo de caso com hidrograma unitário instantâneo geomorfológico. **Revista Brasileira de Recursos Hídricos**, Porto Alegre, v. 21, n. 2, p. 247-359, abril/jun. 2016. DOI <http://dx.doi.org/10.21168/rbrh.v21n2.p347-359>. Disponível em: <https://www.abrh.org.br/SGCv3/index.php?PUB=1&ID=190&SUMARIO=5190>. Acesso em: 5 out. 2019.

RAO, A. Ramachandra; SRINIVAS, V. V. **Regionalization of Watersheds: An Approach Based on Cluster Analysis**. 1. ed. [S. l.]: Springer Science+Business Media B.V., 2008. 245 p. v. 58. ISBN 978-1-4020-6852-2.

SANTOS, Irani dos et al. **Hidrometria Aplicada**. Curitiba: Instituto de Tecnologia para o Desenvolvimento, 2001. 372 p. ISBN 85-88519-01-1.

SAUER, V. B; MEYER, R. W. **Determination of error in individual discharge measurements**. Norcross: U.S. Geological Survey, 1992.

SIMPSON, Michael R. **Discharge Measurements Using a Broad-Band Acoustic Doppler Current Profiler**. Sacramento: U.S. Geological Survey, 2001.

SIVAPALAN, M. *et al.* IAHS Decade on Predictions in Ungauged Basins (PUB), 2003–2012: Shaping an exciting future for the hydrological sciences. **Hydrological Sciences Journal**, [s. l.], v. 48, n. 6, p. 857-880, dez. 2003.

WAHL, Tony L. Performance Limits of Width-Contracted Flumes. *In*: Conference on hydraulic measurements and experimental methods, 2002, Colorado. **Proceedings** [...]. [S. l.]: EWRI/IAHR, 2002.

ANEXOS

ANEXO A – Código fonte do software em MATLAB®

FUNÇÃO INICIAL:**Contents**

Variáveis de entrada:

Carrega a base de dados:

Determina a Série de descargas do local indicado:

Salva as novas estações encontradas na base de dados:

Emissão do Relatório:

Variáveis não declaradas:

Emissão do Relatório:

Tabela 0:

Tabela 1:

Tabela 2:

Tabela 3:

Tabela 4:

Tabela 5:

Tabela 6:

Figura 1:

Figura 2:

Figura 3:

Figuras da tabela 3

Salva os dados hidrológicos

Salva todas as tabelas em Excel:

`function relatorio_hydrofit(x,y,AD,forceWinner)`**Variáveis de entrada:**Coordenadas X e Y do ponto: Área de drenagem do local (km²): Selecionar estação de referência com Overwrite:**Carrega a base de dados:**`try``load('DB.mat')``catch``DB = struct();``DB. EST_AVA = struct();``DB. EST_NAVA = 0;``end`**Determina a Série de descargas do local indicado:**`[EL,DB] = hydrofit6(x,y,AD,DB,forceWinner);`**Salva as novas estações encontradas na base de dados:**`save('.\Assistent Data\DB.mat','DB');`**Emissão do Relatório:****Variáveis não declaradas:**`fields = fieldnames(EL.ERef.ListaERef);`**Emissão do Relatório:****Tabela 0:**`VariaveisTabela0 = {'Variavel' 'Valor' 'Und'};``Variavel_01T0 = {'Coordenada E (Mercator - WGS84)'; 'Coordenada N (Mercator - WGS84)'; 'Área de Drenagem do Empreendimento'; 'Q7,10'};``Variavel_02T0 = {x;y;AD;sprintf('%1.2f',EL.Q710)};``Variavel_03T0 = {'m'; 'm'; 'km²'; 'm³/s'};``Tabela_0 =``table(Variavel_01T0,Variavel_02T0,Variavel_03T0,'VariableNames',VariaveisTabela0);`**Tabela 1:**`VariaveisTabela1 = {'Estacoes Selecionadas'; 'Latitude'; 'Longitude'; 'Area de drenagem'};``for i = 1:size(fields,1)``latitudes(i,1) = EL.ERef.ListaERef.(fields{i,1}).LAT;``longitudes(i,1) = EL.ERef.ListaERef.(fields{i,1}).LON;``AreasDrenagem(i,1) = EL.ERef.ListaERef.(fields{i,1}).AreaDrenagem;``Distancias(i,1) = EL.ERef.ListaERef.(fields{i,1}).distancia;``end`

```

Tabela_1 =
table(fieldnames(EL.ERef.ListaERef),latitudes,longitudes,AreasDrenagem,'VariableNames',VariaveisTabela1);
Tabela 2:
VariaveisTabela2{1,1} = 'Ano';
ANO = (year(EL.Vazoes(1,2)):year(EL.Vazoes(end,2)))';
for i = 1:size(fields,1)
    VariaveisTabela2{i+1,1} = fields{i,1};
    for j = 1:size(ANO,1)
        MatrizTabela2{j,1} = ANO(j,1);
        MatrizTabela2{j,i+1} =
sum(double(logical(double(EL.ERef.ListaERef.(fields{i,1}).VazoesMensais(:,2) ...
    <= datenum(ANO(j,1),12,31)) .*
double(EL.ERef.ListaERef.(fields{i,1}).VazoesMensais(:,2) > datenum(ANO(j,1)-1,12,31)))));
    end
end
Tabela_2 = cell2table(MatrizTabela2,'VariableNames',VariaveisTabela2);
Tabela 3:
VariaveisTabela3 = {'Estacao';'R2';'Periodo_Utilizado';'Correlacao'};
fields_Vpre = fieldnames(EL.ERef.Vpre);
j = 1;
for i = 1:size(EL.ERef.CTable,1)
    if strcmp(EL.ERef.CTable{i,1}(1,:),EL.ERef.CodEstacao)
        Tabela3{j,1} = EL.ERef.CTable{i,2};
        Tabela3{j,2} = EL.ERef.CTable{i,3};
        try
            LMP = MesesPreenchidos(EL.ERef.Vpre.(Tabela3{j,1}));
            stringTotal = '';
            for m = 1:size(LMP,1)
                if m == 1
                    memo_i = LMP(m,1);
                else
                    if LMP(m,1)-LMP(m-1,1) < 40
                        continue
                    else
                        memo_f = LMP(m-1,1);
                        string = sprintf('%s a %s',datestr(memo_i,'mmm-
yyyy'),datestr(memo_f,'mmm-yyyy'));
                        stringTotal = sprintf('%s\n%s',stringTotal,string);
                        memo_i = LMP(m);
                    end
                end
            end
            if size(LMP,1) == 1
                stringTotal = sprintf('%s',datestr(memo_i,'mmm-yyyy'));
            end
            Tabela3{j,3} = sprintf('%s',stringTotal);
            Tabela3{j,4} = '<Inserir Figura>';
            j = j+1;
        catch
        end
    end
end
Tabela_3 = cell2table(Tabela3,'VariableNames',VariaveisTabela3);
Tabela 4:
VariaveisTabela4 = {'Media' 'Maxima' 'Minima' 'Especific'};
Q_especifica = mean(EL.ERef.VazoesMensais(:,1))*1000/EL.ERef.AreaDrenagem;
Tabela4 = {mean(EL.ERef.VazoesMensais(:,1)), max(EL.ERef.VazoesMensais(:,1)),
min(EL.ERef.VazoesMensais(:,1)), Q_especifica};
Tabela_4 = cell2table(Tabela4,'VariableNames',VariaveisTabela4);
Tabela 5:
VariaveisTabela5 = {'Ano' 'JAN' 'FEV' 'MAR' 'ABR' 'MAI' 'JUN' 'JUL' 'AGO' 'SET' 'OUT' 'NOV'
'DEZ'};
Matriz = SerieVazaoANEEL(EL.ERef);
for i = 2:size(Matriz,1)

```

```

    for j = 1:size(Matriz,2)
        Tabela5{i-1,j} = Matriz{i,j};
    end
end
Tabela_5 = cell2table(Tabela5,'VariableNames',VariaveisTabela5);
Tabela 6:
VariaveisTabela6 = {'Ano' 'JAN' 'FEV' 'MAR' 'ABR' 'MAI' 'JUN' 'JUL' 'AGO' 'SET' 'OUT' 'NOV'
'DEZ'};
Matriz = SerieVazaoANEEL(EL);
for i = 2:size(Matriz,1)
    for j = 1:size(Matriz,2)
        Tabela6{i-1,j} = Matriz{i,j};
    end
end
Tabela_6 = cell2table(Tabela6,'VariableNames',VariaveisTabela6);
Figura 1:
figure(1);
[~, PolyBacia] = findBacia(x,y);
% Conversão de valores x,y para geográfica:
mstruct = defaultm('mercator');
mstruct.origin = [0 0 0];
mstruct = defaultm(mstruct);
mstruct.geoid = referenceEllipsoid(7030);
mstruct = defaultm(mstruct);
[lat,lon] = minvtran(mstruct,PolyBacia(:,1),PolyBacia(:,2));
plot(lon,lat,'-k'); % Poligono da Sub-Bacia;
hold on;
plot(longitudes,latitudes,'ro'); % Estações;
% plota os nomes das estações utilizadas:
text(longitudes,latitudes,fieldnames(EL.ERef.ListaERef));
% Conversão de valores x,y para geográfica:
mstruct = defaultm('mercator');
mstruct.origin = [0 0 0];
mstruct = defaultm(mstruct);
mstruct.geoid = referenceEllipsoid(7030);
mstruct = defaultm(mstruct);
[lat2,lon2] = minvtran(mstruct,x,y);
% plota o local em estudo:
plot(lon2,lat2,'xk'); % Local;
text(lon2,lat2,'Local Estudado');
% Carrega mapa hiraulico do Brasil:
load('MHB.mat');
mstruct = defaultm('mercator');
mstruct.origin = [0 0 0];
mstruct = defaultm(mstruct);
mstruct.geoid = referenceEllipsoid(7030);
mstruct = defaultm(mstruct);
[MHB(:,2),MHB(:,1)] = minvtran(mstruct,MHB(:,1),MHB(:,2));
% Cria matriz de recorte do mapa hidraulico:
MR_MHB = MHB(logical(double(MHB(:,1))>=min(lon)-2) .* double(MHB(:,1)<=max(lon)+2) .*
double(MHB(:,2))>=min(lat)-2) .* double(MHB(:,2)<=max(lat)+2)),:);
plot(MR_MHB(:,1),MR_MHB(:,2),'.','MarkerSize',3,'Color','b')
% Identifica a estação mais distante:
d_max = max(Distancias);
[lat_max,lon_max] = minvtran(mstruct,x+d_max+5000,y+d_max+5000);
[lat_min,lon_min] = minvtran(mstruct,x-d_max-5000,y-d_max-5000);
FG1 = gca;
FG1.XLim = [lon_min lon_max];
FG1.YLim = [lat_min lat_max];
FG1.XGrid = 'On';
FG1.YGrid = 'On';
title(FG1,'Localização das estações selecionadas');
ylabel(FG1,'Latitude [°]');
xlabel(FG1,'Longitude [°]');
hold off;

```

```

saveas(FG1, 'Figura_1.emf');
saveas(FG1, 'Figura_1.fig');
Figura 2:
if ~isempty(EL.ERef.DadosCurvaChave)
    fprintf('\nA Figura 2 apresenta a curva-chave média calculada para a estação %s e
    respectivos dados utilizados.\n', EL.ERef.CodEstacao);
    figure(2);

scatter(EL.ERef.DadosCurvaChave.Tratado(:,2), EL.ERef.DadosCurvaChave.Tratado(:,3), 'bo');
hold on;
fields = fieldnames(EL.ERef.CurvaChave);
for i = 1:size(fields,1)
    plot(EL.ERef.CurvaChave.(fields{i,1}));
    legendas{i+1,1} = sprintf('%1.4f*(x-
%1.2f)\ \ ^%1.3f', EL.ERef.CurvaChave.(fields{i,1}).a, EL.ERef.CurvaChave.(fields{i,1}).b, EL.ER
ef.CurvaChave.(fields{i,1}).c);
    hold on;
end
legendas{1,1} = 'Dados utilizados';
legend(legendas);
FG = gca;
FG.XGrid = 'On';
FG.YGrid = 'On';
titulo2 = sprintf('Curva-chave da estação %s', EL.ERef.CodEstacao);
title(FG, titulo2);
xlabel2 = 'Vazão [m³/s]';
ylabel2 = 'Cota [m]';
xlabel(FG, xlabel2);
ylabel(FG, ylabel2);
saveas(FG, 'Figura_2.emf');
end
Figura 3:
figure(size(Tabela3,1)+3);
scatter(EL.MRE(:,1), EL.MRE(:,2), 'bo');
hold on;
scatter(EL.AreaDrenagem, EL.CRE(EL.AreaDrenagem), 'or');
plot(EL.CRE);
FG = gca;
FG.XGrid = 'On';
FG.YGrid = 'On';
title(FG, 'Curva de Rendimento específico da bacia');
ylabel(FG, 'Q específica [l/s.km²]');
xlabel(FG, 'Área de drenagem [km²]');
hold off;
saveas(FG, 'Figura_3.emf');
Figuras da tabela 3
Para cada linha da tabela:
VEA = EL.ERef.ListaERef.(EL.ERef.CodEstacao).VazoesMensais;
for i = 1:size(Tabela3,1)
    figure(i+2);
    VEB = EL.ERef.ListaERef.(Tabela3{i,1}).VazoesMensais;
    % Define os dados correlacionáveis (mesma data):
    [IntData, idVa, idVb] = intersect(VEA(:,2), VEB(:,2));
    A = sortrows(VEA(idVa,:), 2);
    B = sortrows(VEB(idVb,:), 2);
    CurvaR2 = fit(A(:,1), B(:,1), 'poly1');
    plot(CurvaR2);
    hold on;
    scatter(A(:,1), B(:,1), 'go');
    FG = gca;
    a_label = sprintf('Q - %s [m³/s]', EL.ERef.CodEstacao);
    b_label = sprintf('Q - %s [m³/s]', Tabela3{i,1});
    ylabel(FG, b_label);
    xlabel(FG, a_label);
    hold off;
end

```

```
    filenameFG = sprintf('Figura_Tabela_3_%1.0f.emf',i);
    saveas(FG,filenameFG);
end
Salva os dados hidrológicos
save EL EL;
Salva todas as tabelas em Excel:
filename = 'Tabelas.xlsx';
writetable(Tabela_0,filename,'Sheet',1);
writetable(Tabela_1,filename,'Sheet',2);
writetable(Tabela_2,filename,'Sheet',3);
writetable(Tabela_3,filename,'Sheet',4);
writetable(Tabela_4,filename,'Sheet',5);
writetable(Tabela_5,filename,'Sheet',6);
writetable(Tabela_6,filename,'Sheet',7);
fclose('all');
```


FUNÇÃO PRINCIPAL:

Contents

Identifica na Base de dados qual a Bacia o ponto está inserido:

Define um raio e busca encontra estações de referência:

Verifica se as estações encontradas já foram catalogadas como NODATA:

Prepara para baixar os dados das estações selecionadas:

Baixa os dados das estações que ainda não estão no banco de dados:

Cria a lista de estações selecionadas e que possuem dados suficientes:

Verifica se existem estações suficientes para completar a análise:

Filtra estações com baixa qualidade de correlação:

Calcula o Ranking das estações de referência:

Define a estação de referência com base no Ranking:

Preenchimento de série da estação de referência:

Fazer Curva de Rendimentos específicos da região de busca:

Calcular série de vazões para o local do empreendimento.

Identifica na Base de dados qual a Bacia o ponto está inserido:

Define um raio e busca encontra estações de referência:

Verifica se as estações encontradas já foram catalogadas como NODATA:

Prepara para baixar os dados das estações selecionadas:

Baixa os dados das estações que ainda não estão no banco de dados:

Cria a lista de estações selecionadas e que possuem dados suficientes:

Filtra estações com baixa qualidade de correlação:

Calcula o Ranking das estações de referência:

```
function [EstacaoLocal,DB] = hydrofit6(x,y,AreaDrenagem,DB,forceWinner)
```

```
n_estacoes_min = 4;
```

```
AnoFinal = datenum(2018,1,1);
```

```
h = waitbar(0,'calculando série hidrológica...');
```

Identifica na Base de dados qual a Bacia o ponto está inserido:

```
[xDoc,DB] = P1(x,y,DB);
```

```
waitbar(1/10);
```

Define um raio e busca encontra estações de referência:

```
rBusca = 0.3; % 0.3 graus decimais, equivale aprox. 30 km.
```

```
[ListaEstacoes,rBusca,EST_RANG] = P2(rBusca,xDoc,x,y,n_estacoes_min,forceWinner);
```

```
waitbar(2/10);
```

Verifica se as estações encontradas já foram catalogadas como NODATA:

```
[ListaCodEstacoes,rBusca,EST_RANG] =
```

```
P3(ListaEstacoes,DB,rBusca,EST_RANG,y,x,xDoc,n_estacoes_min,forceWinner);
```

```
waitbar(3/10);
```

Prepara para baixar os dados das estações selecionadas:

```
ListaCodEstacoesBaixar = P4(DB,ListaCodEstacoes);
```

```
waitbar(4/10);
```

Baixa os dados das estações que ainda não estão no banco de dados:

```
DB = P5(ListaCodEstacoesBaixar,EST_RANG,DB);
```

```
waitbar(5/10);
```

Cria a lista de estações selecionadas e que possuem dados suficientes:

```
ListaCodEstacoesRef = P6(DB,ListaCodEstacoes);
```

```
waitbar(6/10);
```

Verifica se existem estações suficientes para completar a análise:

```
while size(ListaCodEstacoesRef,1) < n_estacoes_min
```

```
    rBusca = rBusca + 0.05;
```

```
    warning('Estacoes encontradas não são representativas, ampliando área de busca para %1.0f km.',rBusca*100);
```

```
    [ListaEstacoes,rBusca,EST_RANG] = P2(rBusca,xDoc,x,y,n_estacoes_min,forceWinner);
```

```
    [ListaCodEstacoes,rBusca,EST_RANG] =
```

```
P3(ListaEstacoes,DB,rBusca,EST_RANG,y,x,xDoc,n_estacoes_min,forceWinner);
```

```
    ListaCodEstacoesBaixar = P4(DB,ListaCodEstacoes);
```

```
    DB = P5(ListaCodEstacoesBaixar,EST_RANG,DB);
```

```
    ListaCodEstacoesRef = P6(DB,ListaCodEstacoes);
```

```
    if rBusca > 10
```

```
        error('Não foram encontradas estações suficientes, ou as estações encontradas não correlacionam e toda a bacia já foi pesquisada.');
```

```
    end
```

```
end
```

Filtra estações com baixa qualidade de correlação:

```

[EstacoesRef, CTable] = P7(ListaCodEstacoesRef,DB);
ListaCodEstacoesRefQF = fieldnames(EstacoesRef);
Lastyear = ErefLastyear(EstacoesRef);
while size(ListaCodEstacoesRefQF,1)< n_estacoes_min || Lastyear < AnoFinal ||
size(forceWinner,2) > 7 && not(sum(strcmp(ListaCodEstacoesRefQF,forceWinner)))
    rBusca = rBusca + 0.05;
    warning('Estacoes encontradas não correlacionam, ampliando área de busca para %1.0f
km.',rBusca*100);
    [ListaEstacoes,rBusca,EST_RANG] = P2(rBusca,xDoc,x,y,n_estacoes_min,forceWinner);
    [ListaCodEstacoes,rBusca,EST_RANG] =
P3(ListaEstacoes,DB,rBusca,EST_RANG,y,x,xDoc,n_estacoes_min,forceWinner);
    ListaCodEstacoesBaixar = P4(DB,ListaCodEstacoes);
    DB = P5(ListaCodEstacoesBaixar,EST_RANG,DB);
    ListaCodEstacoesRef = P6(DB,ListaCodEstacoes);
    [EstacoesRef, CTable] = P7(ListaCodEstacoesRef,DB);
    Lastyear = ErefLastyear(EstacoesRef);
    ListaCodEstacoesRefQF = fieldnames(EstacoesRef);
    if rBusca > 2 && size(ListaCodEstacoesRefQF,1)>= n_estacoes_min
        break
    elseif rBusca > 10
        error('Não foram encontradas estações suficientes, ou as estações encontradas não
correlacionam e toda a bacia já foi pesquisada.');
```

end

```

waitbar(7/10);
Calcula o Ranking das estações de referência:
[EstacoesRef, Winner] = P8(EstacoesRef,AreaDrenagem,forceWinner,x,y);
waitbar(8/10);
Define a estação de referência com base no Ranking:
Eref = EstacoesRef.(Winner);
Eref.CodEstacao = Winner;
Preenchimento de série da estação de referência:
Eref = CompletSerie2(CTable,Winner,EstacoesRef,Eref);
Eref = SerieVazMensal(Eref);
while max(Eref.Vazoes(:,2)) < AnoFinal
    rBusca = rBusca + 0.05;
    warning('Estacoes encontradas não são atualizadas, ampliando área de busca para %1.0f
km.',rBusca*100);
    [ListaEstacoes,rBusca,EST_RANG] = P2(rBusca,xDoc,x,y,n_estacoes_min,forceWinner);
    [ListaCodEstacoes,rBusca,EST_RANG] =
P3(ListaEstacoes,DB,rBusca,EST_RANG,y,x,xDoc,n_estacoes_min,forceWinner);
    ListaCodEstacoesBaixar = P4(DB,ListaCodEstacoes);
    DB = P5(ListaCodEstacoesBaixar,EST_RANG,DB);
    ListaCodEstacoesRef = P6(DB,ListaCodEstacoes);
    [EstacoesRef, CTable] = P7(ListaCodEstacoesRef,DB);
    [EstacoesRef, Winner] = P8(EstacoesRef,AreaDrenagem,forceWinner,x,y);
    Eref = EstacoesRef.(Winner);
    Eref.CodEstacao = Winner;
    Eref = CompletSerie2(CTable,Winner,EstacoesRef,Eref);
    Eref = SerieVazMensal(Eref);
    if rBusca > 2
        break
    end
end
waitbar(9/10);
Fazer Curva de Rendimentos específicos da região de busca:
ListaEstacoesRef = fieldnames(EstacoesRef);
for i = 1:size(ListaEstacoesRef,1)
    MRE(i,1) = EstacoesRef.(ListaEstacoesRef{i,1}).AreaDrenagem; % Área de Drenagem em km².
    MRE(i,2) = (mean(EstacoesRef.(ListaEstacoesRef{i,1}).Vazoes(:,1))*1000)...
/EstacoesRef.(ListaEstacoesRef{i,1}).AreaDrenagem; % Rendimento Específico em
l/s/km².
end
if std(MRE(:,2)) > 3 % Caso tenha muita divergência entre os valores
    [~,idx] = min(MRE(:,2));
```

```

    CRE = fit([MRE(idx,1);0],[MRE(idx,2);MRE(idx,2)], 'poly1');
else % Caso normal
    CRE = fit(MRE(:,1),MRE(:,2), 'poly1');
end
if CRE.p1 > 0 % caso a curva seja crescente
    [~,idx] = min(abs(MRE(:,1)-AreaDrenagem));
    CRE = fit([MRE(idx,1);0],[MRE(idx,2);MRE(idx,2)], 'poly1');
end
if abs((ERef.AreaDrenagem-AreaDrenagem)/AreaDrenagem) < 0.1 && ERef.distancia < 10000 %
caso a estação de referência esteja muito próxima
    [~,idx] = min(abs(MRE(:,1)-AreaDrenagem));
    CRE = fit([MRE(idx,1);0],[MRE(idx,2);MRE(idx,2)], 'poly1');
end
for i = 1:size(ListaEstacoesRef,1)
    if strcmp(forceWinner,ListasEstacoesRef{i,1})
        idx = i; % Item vinculado: MRE e ListasEstacoesRef
        CRE = fit([MRE(idx,1);0],[MRE(idx,2);MRE(idx,2)], 'poly1');
    end
end
% Definir de acordo com a curva de rendimento específico o fator de ajuste
% para transposição de vazões da estação de referência para o local
% estudado.
qLocal = CRE(AreaDrenagem);
qERef = (mean(ERef.Vazoes(:,1))*1000)/ERef.AreaDrenagem;
if qLocal < 0 || qLocal < min(MRE(:,2)) % caso a vazão específica fique negativa ou menor
que o mínimo.
    CRE = fit(MRE(:,1),MRE(:,2), 'power1');
    qLocal = CRE(AreaDrenagem);
    if AreaDrenagem > 1.05*max(MRE(:,1)) && qLocal > MRE(MRE(:,1)==max(MRE(:,1)),2)
        qLocal = MRE(MRE(:,1)==max(MRE(:,1)),2);
        CRE = fit([max(MRE(:,1));0],[qLocal;qLocal], 'poly1');
    end
end
FARE = qLocal/qERef;
Calcular série de vazões para o local do empreendimento.
EstacaoLocal.ERef = ERef;
EstacaoLocal.Vazoes(:,1) =
max(0,ERef.Vazoes(:,1).*(AreaDrenagem/ERef.AreaDrenagem).*(FARE));
EstacaoLocal.Vazoes(:,2) = ERef.Vazoes(:,2);
if sum(EstacaoLocal.Vazoes(:,1) < 0)
    warning('Valores de vazão negativos!');
    EstacaoLocal.Vazoes(EstacaoLocal.Vazoes(:,1)<0,1) = 0;
end
EstacaoLocal.AreaDrenagem = AreaDrenagem;
EstacaoLocal.ERef.ListaERef = EstacoesRef;
EstacaoLocal.Q710 = CompQ710(EstacaoLocal);
EstacaoLocal.permanencia = permanenciaQ2(EstacaoLocal.Vazoes(:,1));
EstacaoLocal.CRE = CRE;
EstacaoLocal.MRE = MRE;
waitbar(10/10);
close(h);
end
Identifica na Base de dados qual a Bacia o ponto está inserido:
function [xDoc,DB] = P1(x,y,DB)
CodBacia = findBacia(x,y);
% Verifica se já foram mapeadas as estações desta bacia:
if isfield(DB,'EBACIAS') && isfield(DB.EBACIAS,(sprintf('E%d',CodBacia)))
    % Caso positivo, utilizar arquivo que lista as estações:
    xDoc = DB.EBACIAS.(sprintf('E%d',CodBacia));
else
    % Caso negativo, baixar arquivo do site da ANA e inserir na DB:
    xDoc = ListEst(CodBacia);
    DB.EBACIAS.(sprintf('E%d',CodBacia)) = xDoc;
end

```

```

end
Define um raio e busca encontra estações de referência:
function [ListaEstacoes,rBusca,EST_RANG] = P2(rBusca,xDoc,x,y,n_estacoes_min,forceWinner)
ListaEstacoes = [];
while size(ListaEstacoes,1)< n_estacoes_min || size(forceWinner,2) > 7 &&
not(sum(strcmp(ListaEstacoes,forceWinner)))
    [EST_RANG] = findAvaliabeEST2(xDoc,x,y,rBusca);
    ListaEstacoes = fieldnames(EST_RANG);
    if size(ListaEstacoes,1)< n_estacoes_min
        rBusca = rBusca + 0.05;
        warning('Ampliando raio de busca por estações para %1.0fkm',rBusca*100);
        if rBusca > 10
            error('Não foram encontradas estações suficientes, ou as estações encontradas
não correlacionam e toda a bacia já foi pesquisada.');
```

end

```

    end
end
end
Verifica se as estações encontradas já foram catalogadas como NODATA:
function [ListaCodEstacoes,rBusca,EST_RANG] =
P3(ListaEstacoes,DB,rBusca,EST_RANG,y,x,xDoc,n_estacoes_min,forceWinner)
% Cria uma lista com o código das estações encontradas:
ListaCodEstacoes = zeros(size(ListaEstacoes,1),1);
for p = 1:size(ListaEstacoes,1)
    ListaCodEstacoes(p,1) = str2double(ListaEstacoes{p,1}(1,2:end));
end
% Filtra as estações já catalogadas como NODATA:
ListaCodEstacoes = setdiff(ListaCodEstacoes,DB.EST_NAVA);
% se sobrar menos de duas estações, o raio de busca deverá ser expandido:
while isempty(ListaCodEstacoes) || size(ListaCodEstacoes,1)< n_estacoes_min ||
size(forceWinner,2) > 7 && not(sum(strcmp(ListaEstacoes,forceWinner)))
    rBusca = rBusca + 0.05;
    warning('Estacoes encontradas não possuem dados, ampliando área de busca para %1.0f
km.',rBusca*100);
    [EST_RANG] = findAvaliabeEST2(xDoc,x,y,rBusca);
    ListaEstacoes = fieldnames(EST_RANG);
    ListaCodEstacoes = zeros(size(ListaEstacoes,1),1);
    for p = 1:size(ListaEstacoes,1)
        ListaCodEstacoes(p,1) = str2double(ListaEstacoes{p,1}(1,2:end));
    end
    ListaCodEstacoes = setdiff(ListaCodEstacoes,DB.EST_NAVA);
    if rBusca > 10
        error('Não foram encontradas estações suficientes, ou as estações encontradas não
correlacionam e toda a bacia já foi pesquisada.');
```

end

```

end
end
end
Prepara para baixar os dados das estações selecionadas:
function ListaCodEstacoesBaixar = P4(DB,ListaCodEstacoes)
% Verifica se as estações selecionadas já existem no banco de dados:
ListaEstacoesDB = fieldnames(DB.EST_AVA);
if isempty(ListaEstacoesDB)
    ListaCodEstacoesDB = [];
else
    ListaCodEstacoesDB = zeros(size(ListaEstacoesDB,1),1);
    for p = 1:size(ListaEstacoesDB,1)
        ListaCodEstacoesDB(p,1) = str2double(ListaEstacoesDB{p,1}(1,2:end));
    end
end
end
ListaCodEstacoesDB = intersect(ListaCodEstacoesDB,ListaCodEstacoes);
ListaCodEstacoesBaixar = setdiff(ListaCodEstacoes,ListaCodEstacoesDB);
end
Baixa os dados das estações que ainda não estão no banco de dados:
function DB = P5(ListaCodEstacoesBaixar,EST_RANG,DB)
if ~isempty(ListaCodEstacoesBaixar)
```

```

for i = 1:size(ListaCodEstacoesBaixar,1)
    lat_e = EST_RANG.(sprintf('%d',ListaCodEstacoesBaixar(i,1))).LAT;
    lon_e = EST_RANG.(sprintf('%d',ListaCodEstacoesBaixar(i,1))).LON;
    CodEstacao = ListaCodEstacoesBaixar(i,1);
    [EST, NODATA] = ANAgetdata(CodEstacao,lat_e,lon_e);
    if ~isempty(NODATA)
        DB.EST_NAVA = [DB.EST_NAVA;NODATA];
    else
        DB.EST_AVA.(sprintf('%d',ListaCodEstacoesBaixar(i,1))) =
EST.(sprintf('%d',ListaCodEstacoesBaixar(i,1)));
    end
end
end
end
Cria a lista de estações selecionadas e que possuem dados suficientes:
function ListaCodEstacoesRef = P6(DB,ListaCodEstacoes)
ListaEstacoesDB = fieldnames(DB.EST_AVA);
ListaCodEstacoesDB = zeros(size(ListaEstacoesDB,1),1);
for p = 1:size(ListaEstacoesDB,1)
    ListaCodEstacoesDB(p,1) = str2double(ListaEstacoesDB{p,1}(1,2:end));
end
ListaCodEstacoesRef = intersect(ListaCodEstacoesDB,ListaCodEstacoes);
end
Filtra estações com baixa qualidade de correlação:
function [EstacoesRef, CTable] = P7(ListaCodEstacoesRef,DB)
for p = 1:size(ListaCodEstacoesRef,1)
    EstacoesRef.(sprintf('%d',ListaCodEstacoesRef(p,1))) =
DB.EST_AVA.(sprintf('%d',ListaCodEstacoesRef(p,1)));
end
[EstacoesRef, CTable] = Qualifilter3(EstacoesRef);
end
Calcula o Ranking das estações de referência:
function [EstacoesRef, Winner] = P8(EstacoesRef,AreaDrenagem,forceWinner,x,y)
fields = fieldnames(EstacoesRef);
for p = 1:size(fields)
    lat_e = EstacoesRef.(fields{p,1}).LAT;
    lon_e = EstacoesRef.(fields{p,1}).LON;
    EstacoesRef.(fields{p,1}).distancia = CalcDist(x,y,lat_e,lon_e);
end
[EstacoesRef, Winner] = RankStation(EstacoesRef,AreaDrenagem,forceWinner);
end

```

FUNÇÕES AUXILIARES:

Contents

Converte os valores de x e y para latitude e longitude:

```
function dist = CalcDist(x,y,lat_e,lon_e)
```

Converte os valores de x e y para latitude e longitude:

```
mstruct = defaultm('mercator');
mstruct.origin = [0 0 0];
mstruct = defaultm(mstruct);
mstruct.geoid = referenceEllipsoid(7030);
mstruct = defaultm(mstruct);
[xe,ye] = mfwdtran(mstruct,lat_e,lon_e);
dist = sqrt((xe-x)^2+(ye-y)^2);
```

Contents

Verifica quantos anos de vazões foram medidos:

Classifica o Vetor de vazões de acordo com o ano:

Define o vetor de eventos de média móvel de 7 dias de vazões para cada ano:

Log Pearson tipo 3

Transforma a amostra em log:

Estabelece a média da amostra:

Estabelece o desvio padrão da amostra:

Estabelece o fator de frequência (K10) para 10 anos de recorrência:

Estabelece a magnitude do evento (Q710):

```
function Q710 = CompQ710(Estacao)
```

Verifica quantos anos de vazões foram medidos:

```
[Yi,~,~] = datevec(Estacao.Vazoes(1,2));
[Yf,~,~] = datevec(Estacao.Vazoes(end,2));
```

Classifica o Vetor de vazões de acordo com o ano:

```
[Ys,~,~] = datevec(Estacao.Vazoes(:,2));
```

```
Estacao.Vazoes(:,3) = Ys;
```

Define o vetor de eventos de média móvel de 7 dias de vazões para cada ano:

```
Min7dias = zeros(Yf-Yi+1,1);
```

```
j = 1;
```

```
for i = Yi:Yf
```

```
    % Seleciona apenas anos com mais de 70% de dados coletados:
```

```
    if size(Estacao.Vazoes(Estacao.Vazoes(:,3)==i,1),1) > 365*0.7
```

```
        Med7dias = movmean(Estacao.Vazoes(Estacao.Vazoes(:,3)==i,1),7);
```

```
        Med7dias = Med7dias(1:end-6,1);
```

```
        % Escolhe o menor evento da série para compor o vetor de Mínimas:
```

```
        Min7dias(j,1) = min(Med7dias);
```

```
        j = j+1;
```

```
    end
```

```
end
```

```
Min7dias(Min7dias(:,1)<=0,:) = [];
```

```
Amostra = Min7dias;
```

Log Pearson tipo 3

Transforma a amostra em log:

```
Min7dias = reallog(Min7dias);
```

Estabelece a média da amostra:

```
MuAmostra = mean(Min7dias);
```

Estabelece o desvio padrão da amostra:

```
dpAmostra = std(Min7dias);
```

Estabelece o fator de frequência (K10) para 10 anos de recorrência:

```
y = skewness(Min7dias);
```

```
K10 = -0.0001878775*y^5 + 0.0011011289*y^4 - 0.0006418722*y^3 - 0.0502790786*y^2 +
0.1073961052*y + 1.2823993035;
```

Estabelece a magnitude do evento (Q710):

```
Q710 = MuAmostra - K10*dpAmostra;
```

```
Q710 = exp(Q710);
```

Contents

Verifica se as estações com correlação adequadas podem contribuir com dados faltantes na estação de Referência

Cria Estação Sintética de Referência com série de dados completada (completar falhas com média):

```
function ERef = CompletSerie2(CTable,Winner,Estacoes,ERef)
```

Verifica se as estações com correlação adequadas podem contribuir com dados faltantes na estação de Referência

```

j = 1;
ERef.CTable = CTable;
for i = 1:size(CTable,1)
    if CTable{i,1} == Winner
        CTable_ref{j,1} = CTable{i,1};
        CTable_ref{j,2} = CTable{i,2};
        CTable_ref{j,3} = CTable{i,3};
        j = j+1;
    end
end
% Preencher em ordem:
% 1. Selecionar a estação com melhor correlação;
CTable_ref = sortrows(CTable_ref,3);
for i = 1:size(CTable_ref,1)
    Epre = Estacoes.(CTable_ref{end-i+1,2});
    NomeEstacao = CTable_ref{end-i+1,2};
    ERef = preencher2(Epre,ERef,NomeEstacao);
end
% 5. Após fazer os preenchimentos, verificar tudo novamente para a próxima
% estação com melhor correlação.
Cria Estação Sintética de Referência com série de dados completada (completar falhas com
média):
VmedMes = QmedMensal(ERef);
DataIni = min(ERef.Vazoes(:,2));
DataFim = max(ERef.Vazoes(:,2));
nVazoes = zeros(DataFim-DataIni+1,2);
% Define a primeira linha do novo vetor de vazões:
nVazoes(1,2) = DataIni;
Idx = find(ERef.Vazoes(:,2)==DataIni);
if ~isempty(Idx)
    nVazoes(1,1) = ERef.Vazoes(Idx(1,1),1);
else
    [~,M] = datevec(nVazoes(1,2));
    nVazoes(1,1) = VmedMes(M,1);
    ERef.Vmed = nVazoes(1,:);
end
ERef.Vmed = [];
for i = 2:size(nVazoes,1)
    nVazoes(i,2) = nVazoes(i-1,2)+1;
    Idx = find(ERef.Vazoes(:,2)==nVazoes(i,2));
    if ~isempty(Idx)
        nVazoes(i,1) = ERef.Vazoes(Idx(1,1),1);
    else
        [~,M] = datevec(nVazoes(i,2));
        nVazoes(i,1) = VmedMes(M,1);
        ERef.Vmed = [ERef.Vmed;nVazoes(i,:)];
    end
end
ERef.Vazoes = nVazoes;
http://www.mathworks.com/products/matlab/
function LastYear = ErefLastyear(EstacoesRef)
fields = fieldnames(EstacoesRef);
LastYear = 0;
for i = 1:size(fields,1)
    LastYear_i = max(EstacoesRef.(fields{i,1}).Vazoes(:,2));
    if LastYear_i > LastYear
        LastYear = LastYear_i;
    end
end

```

Contents
Declara e atribui as variáveis de entrada:
Invoca os dados de Vazão:
Invoca os dados de Cotas:

```

Invoca os dados de Área de Drenagem:
Interpreta os dados dos arquivos xml:
Verifica e filtra os dados que se encaixam na curva-chave da estação:
Verifica se a série tem representatividade estatística:
Insere as variáveis encontradas na variável de saída:
Deleta os arquivos de apoio utilizados:
function [EST, NODATA] = ANAgetdata(CodEstacao,lat_e,lon_e)
% EST é Struct com os seguintes campos:
% EST.EXXXXXXX - Código da Estação;
% EST.EXXXXXXX.Vazoes - Matriz [Vazões diárias, Data];
% EST.EXXXXXXX.VazoesMensais - Matriz [Vazões mensais, Data];
% EST.EXXXXXXX.AreaDrenagem - Área de Drenagem da estação;
% EST.EXXXXXXX.LAT - Latitude da estação em graus decimais (WGS84);
% EST.EXXXXXXX.LON - Longitude da estação em graus decimais (WGS84);
Declara e atribui as variáveis de entrada:
EST.(sprintf('E%d',CodEstacao)).Vazoes = [];
EST.(sprintf('E%d',CodEstacao)).AreaDrenagem = [];
EST.(sprintf('E%d',CodEstacao)).LAT = lat_e;
EST.(sprintf('E%d',CodEstacao)).LON = lon_e;
NODATA = [];
Invoca os dados de Vazão:
xmlVazoes = v_import_webANA2(CodEstacao);
if strcmp(xmlVazoes,'none')
    warning('Estação %1.0f Não possui dados de vazão!',CodEstacao);
    NODATA = CodEstacao;
    warning('off');
    delete(xmlVazoes);
    warning('on');
    return
end
Invoca os dados de Cotas:
xmlCotas = c_import_webANA2(CodEstacao);
if strcmp(xmlCotas(1,:), 'none') && strcmp(xmlCotas(2,:), 'none')
    warning('Estação %1.0f Não possui dados de cotas!',CodEstacao);
    NODATA = CodEstacao;
    warning('off');
    delete(xmlVazoes);
    delete(xmlCotas(1,:));
    delete(xmlCotas(2,:));
    warning('on');
    return
end
Invoca os dados de Área de Drenagem:
xmlArea = d_import_webANA2(CodEstacao);
if strcmp(xmlArea,'none')
    warning('Estação %1.0f Não possui área de drenagem!',CodEstacao);
    NODATA = CodEstacao;
    warning('off');
    delete(xmlVazoes);
    delete(xmlCotas(1,:));
    delete(xmlCotas(2,:));
    delete(xmlArea);
    warning('on');
    return
end
Interpreta os dados dos arquivos xml:
Vazões:
[Vazoes,leitura] = intxmlV2(xmlVazoes);
trys = 0;
while strcmp('fail',leitura)
    pause(10)
    xmlVazoes = v_import_webANA2(CodEstacao);
    if strcmp(xmlVazoes,'none')
        warning('Estação %1.0f Não possui dados de vazão!',CodEstacao);
        NODATA = CodEstacao;

```



```

        warning('off');
        try
            delete(xmlVazoes);
        catch
        end
        warning('on');
        return
    end
[Vazoes, leitura] = intxmlV2(xmlVazoes);
trys = trys + 1;
if trys > 5
    warning('Estação %1.0f Não possui dados de cotas!', CodEstacao);
    NODATA = CodEstacao;
    warning('off');
    try
        delete(xmlVazoes);
    catch
    end
    warning('on');
    return
end
end
% Cotas 1:
[Cotas1, leitura] = intxmlC(xmlCotas(1,:));
trys = 0;
while strcmp('fail', leitura)
    pause(10)
    xmlCotas = c_import_webANA2(CodEstacao);
    if strcmp(xmlCotas(1,:), 'none') && strcmp(xmlCotas(2,:), 'none')
        warning('Estação %1.0f Não possui dados de cotas!', CodEstacao);
        NODATA = CodEstacao;
        warning('off');
        try
            delete(xmlVazoes);
        catch
        end
        try
            delete(xmlCotas(1,:));
            delete(xmlCotas(2,:));
        catch
        end
        warning('on');
        return
    end
[Cotas1, leitura] = intxmlC(xmlCotas(1,:));
trys = trys + 1;
if trys > 5
    warning('Estação %1.0f Não possui dados de cotas!', CodEstacao);
    NODATA = CodEstacao;
    warning('off');
    try
        delete(xmlVazoes);
    catch
    end
    try
        delete(xmlCotas(1,:));
        delete(xmlCotas(2,:));
    catch
    end
    warning('on');
    return
end
end
% Cotas 2:
[Cotas2, leitura] = intxmlC(xmlCotas(2,:));

```

```

trys = 0;
while strcmp('fail',leitura)
    pause(10)
    xmlCotas = c_import_webANA2(CodEstacao);
    if strcmp(xmlCotas(1,:), 'none') && strcmp(xmlCotas(2,:), 'none')
        warning('Estação %1.0f Não possui dados de cotas!', CodEstacao);
        NODATA = CodEstacao;
        warning('off');
        delete(xmlVazoes);
        delete(xmlCotas(1,:));
        delete(xmlCotas(2,:));
        warning('on');
        return
    end
    [Cotas1,leitura] = intxmlC(xmlCotas(2,:));
    trys = trys + 1;
    if trys > 5
        break
    end
end
AD = intxmlA(xmlArea);
if isnan(AD) || AD < 10 || size(Vazoes,1) < (12*12*31)
    warning('Estação %1.0f Não possui área de drenagem ou não é representativa! (AD = %1.0f, serie size [yr] = %1.0f)', CodEstacao, AD, size(Vazoes,1)/(12*31));
    NODATA = CodEstacao;
    warning('off');
    delete(xmlVazoes);
    delete(xmlCotas(1,:));
    delete(xmlCotas(2,:));
    delete(xmlArea);
    warning('on');
    return
end
Verifica e filtra os dados que se encaixam na curva-chave da estação:
[SerieVazoes, CurvData, CurvFun] = KeyCurvFilter5(Cotas1, Cotas2, Vazoes);
Verifica se a série tem representatividade estatística:
if size(SerieVazoes,1) < (12*12*31) || size(SerieVazoes,1) <= (15*12*31) &&
max(SerieVazoes(:,2)) < datenum(1970,1,1)
    warning('Estação %1.0f Não possui mínimo de 12 anos de dados consistidos! [%1.0f]', CodEstacao, floor(size(SerieVazoes,1)/(365)));
    NODATA = CodEstacao;
    warning('off');
    delete(xmlVazoes);
    delete(xmlCotas(1,:));
    delete(xmlCotas(2,:));
    delete(xmlArea);
    warning('on');
    return
end
Insera as variáveis encontradas na variável de saída:
EST.(sprintf('E%d', CodEstacao)).Vazoes = SerieVazoes;
EST.(sprintf('E%d', CodEstacao)).AreaDrenagem = AD;
EST.(sprintf('E%d', CodEstacao)) = SerieVazMensal(EST.(sprintf('E%d', CodEstacao)));
EST.(sprintf('E%d', CodEstacao)).CurvaChave = CurvFun;
EST.(sprintf('E%d', CodEstacao)).DadosCurvaChave = CurvData;
Deleta os arquivos de apoio utilizados:
warning('off');
delete(xmlVazoes);
delete(xmlCotas(1,:));
delete(xmlCotas(2,:));
delete(xmlArea);
warning('on');
Contents
Faz a comunicação com o site da ANA
function xDoc = ListEst(CodBacia)

```

Faz a comunicação com o site da ANA

```

filename = sprintf('ESTACOES%d.xml',CodBacia);
Codi = CodBacia*1000000;
Codf = Codi+999999;
url =
sprintf('http://telemetriaws1.ana.gov.br/ServiceANA.asmx/HidroInventario?codEstDE=%d&codEst
ATE=%d&tpEst=1&nmEst=&nmRio=&codSubBacia=&codBacia=&nmMunicipio=&nmEstado=&sgResp=&sgOper=&
telemetrica=',Codi,Codf);
outfilename = 0;
while outfilename == 0;
    try
        outfilename = websave(filename,url);
    catch
        warning('\nFalha na conexão com servidor da ANA, tentando novamente em 10
segundos...\n');
        pause(10);
    end
end
xDoc = xmlread(outfilename);
delete(outfilename);
Contents
Nível de consistência 2: Consistido
Pausa para não sobrecarregar o servidor:
Nível de consistência 1: Bruto
function outfilename = c_import_webANA2(codEstacao)
Nível de consistência 2: Consistido
filename = sprintf('c2%d.xml',codEstacao);
dataFim = datestr(now(),'dd/mm/yyyy');
url =
sprintf('http://telemetriaws1.ana.gov.br/ServiceANA.asmx/HidroSerieHistorica?codEstacao=%d&
dataInicio=01/01/1900&dataFim=%s&tipoDados=1&nivelConsistencia=2',codEstacao,dataFim);
try
    outfilename(2,:) = websave(filename,url);
catch
    pause(10)
    try
        outfilename(2,:) = websave(filename,url);
    catch
        warning('Falha ao baixar os dados de cotas consistidos da estação %d',codEstacao)
        outfilename(2,:) = 'none';
    end
end
Pausa para não sobrecarregar o servidor:
pause(2);
Nível de consistência 1: Bruto
filename = sprintf('c1%d.xml',codEstacao);
url =
sprintf('http://telemetriaws1.ana.gov.br/ServiceANA.asmx/HidroSerieHistorica?codEstacao=%d&
dataInicio=01/01/1900&dataFim=%s&tipoDados=1&nivelConsistencia=1',codEstacao,dataFim);
try
    outfilename(1,:) = websave(filename,url);
catch
    pause(10)
    try
        outfilename(1,:) = websave(filename,url);
    catch
        warning('Falha ao baixar os dados de cotas brutos da estação %d',codEstacao)
        outfilename(1,:) = 'none';
    end
end
end
http://www.mathworks.com/products/matlab/
function outfilename = d_import_webANA2(codEstacao)
filename = sprintf('d%d.xml',codEstacao);
url =
sprintf('http://telemetriaws1.ana.gov.br/ServiceANA.asmx/HidroInventario?codEstDE=%d&codEst

```

```

ATE=%d&tpEst=1&nmEst=&nmRio=&codSubBacia=&codBacia=&nmMunicipio=&nmEstado=&sgResp=&sgOper=&
telemetrica=',codEstacao,codEstacao);
try
    outfile = websave(filename,url);
catch
    pause(10)
    try
        outfile = websave(filename,url);
    catch
        warning('Falha ao baixar os dados de Área de Drenagem da estação %d',codEstacao)
        outfile = 'none';
    end
end

```

Contents

Converte os valores de x e y para latitude e longitude:

Identifica as estações

```
function [EST_AVA] = findAvaliabeEST2(xDoc,x,y,rBusca)
```

Converte os valores de x e y para latitude e longitude:

```
mstruct = defaultm('mercator');
```

```
mstruct.origin = [0 0 0];
```

```
mstruct = defaultm(mstruct);
```

```
mstruct.geoid = referenceEllipsoid(7030);
```

```
mstruct = defaultm(mstruct);
```

```
[lat,lon] = minvtran(mstruct,x,y);
```

Identifica as estações

```
nEstacoes(1,1) = 0;
```

```
dist(1,1) = 0;
```

```
long_est(1,1) = 0;
```

```
lat_est(1,1) = 0;
```

```
size = xDoc.item(0).item(3).item(1).getLength;
```

```
j = 1;
```

```
k = 1;
```

```
for i = 1:2:size
```

```
    confSerie = sprintf('Table%d',j);
```

```
    try
```

```
        if strcmp(xDoc.item(0).item(3).item(1).item(i).getAttribute('diffgr:id'),confSerie)
```

```
            DateIni =
```

```
            datenum(char(xDoc.item(0).item(3).item(1).item(i).item(79).item(0).getData),'yyyy-mm-dd
            HH:MM:SS');
```

```
            lon_e =
```

```
            str2double(char(xDoc.item(0).item(3).item(1).item(i).item(43).item(0).getData));
```

```
            lat_e =
```

```
            str2double(char(xDoc.item(0).item(3).item(1).item(i).item(41).item(0).getData));
```

```
            if sqrt((lon_e-lon)^2+(lat_e-lat)^2) <= rBusca
```

```
                CodEstacao =
```

```
                str2double(char(xDoc.item(0).item(3).item(1).item(i).item(35).item(0).getData));
```

```
                nEstacoes(k,1) = CodEstacao;
```

```
                dist(k,1) = sqrt((lon_e-lon)^2+(lat_e-lat)^2);
```

```
                long_est(k,1) = lon_e;
```

```
                lat_est(k,1) = lat_e;
```

```
                k = k+1;
```

```
            end
```

```
        else
```

```
        end
```

```
    catch
```

```
    end
```

```
    j = j+1;
```

```
end
```

```
EST_AVA = struct();
```

```
for i = 1:k-1
```

```
    field = sprintf('E%d',nEstacoes(i,1));
```

```
    EST_AVA.(field).LAT = lat_est(i,1);
```

```
    EST_AVA.(field).LON = long_est(i,1);
```

```
end
```

<http://www.mathworks.com/products/matlab/>

```

function [CodBacia, PolyBacia] = findBacia(x,y)
load('SBMAPBR.mat');
SBMAP = SBMAPBR;
CodBacia = 0;
PolyBacia = [0 0];
CodBacias = unique(SBMAP(:,3));
for i = 1:size(CodBacias,1)
    BaciaSelected = SBMAP(SBMAP(:,3)==CodBacias(i,1),:);
    if inpolygon(x,y,BaciaSelected(:,1),BaciaSelected(:,2))
        CodBacia = CodBacias(i,1);
        PolyBacia = BaciaSelected;
        break
    end
end
if CodBacia == 0
    warning('\nBacia não identificada!\n');
end
http://www.mathworks.com/products/matlab/
function ADrenagem = intxmlA(filename)
xDoc = xmlread(filename);
try
    if
        strcmp(char(xDoc.item(0).item(1).item(1).item(1).item(1).item(1).item(1).item(1).item(47).get
        Attribute('name')), 'AreaDrenagem')
            ADrenagem =
                str2double(char(xDoc.item(0).item(3).item(1).item(1).item(47).item(0).getData));
        elseif
            strcmp(char(xDoc.item(0).item(1).item(1).item(1).item(1).item(1).item(1).item(1).get
            Attribute('name')), 'Vazio')
                ADrenagem = NaN;
            return
        else
            ADrenagem = NaN;
        end
    end
catch
    ADrenagem = NaN;
end
Contents
Lê o arquivo xml e armazena em uma variável:
Identifica o nº de meses com medição de vazão:
Faz a leitura dos dados:
Faz a média das 3 medições diárias:
function [SerieHistorica,leitura] = intxmlC(filename)
Lê o arquivo xml e armazena em uma variável:
leitura = 'fail';
while strcmp(leitura,'fail')
    try
        xDoc = xmlread(filename);
        leitura = 'OK';
    catch
        warning('Erro de leitura no arquivo %s',filename);
        SerieHistorica = [];
        return
    end
end
Identifica o nº de meses com medição de vazão:
sizexDoc = xDoc.item(0).item(3).item(1).getLength;
if sizexDoc < 10
    SerieHistorica = [];
    return
end
Faz a leitura dos dados:
SerieHistorica(:,1) - Dados de Vazão; SerieHistorica(:,2) - Nivel de Consistencia;
SerieHistorica(:,3) - Data;
SerieHistorica = zeros(sizexDoc*31,3);

```

```

p = 1;
u = 1;
for i = 1:2:sizeXDoc
    confSerie = sprintf('SerieHistorica%d',p);
    try
        if strcmp(xDoc.item(0).item(3).item(1).item(i).getAttribute('diffgr:id'),confSerie)
            % SerieHistorica(:,3) - Data;
            [Y,M,~,~,~,~] =
datevec(char(xDoc.item(0).item(3).item(1).item(i).item(5).item(0).getData), 'yyyy-mm-dd
HH:MM:SS');
            for j = 29:2:89
                CotaAtual = char(xDoc.item(0).item(3).item(1).item(i).item(j).getNodeName);
                diaAtual = str2double(CotaAtual(1,5:end));
                if strcmp(CotaAtual(1,1:4), 'Cota') && size(CotaAtual,2)<7
                    try
                        % SerieHistorica(:,1) - Dados de Vazão;
                        SerieHistorica(u,1) =
str2double(char(xDoc.item(0).item(3).item(1).item(i).item(j).item(0).getData));
                        % SerieHistorica(:,2) - Nivel de Consistencia;
                        % Variável descartada para economia de processamento.
                        % SerieHistorica(:,3) - Data;
                        SerieHistorica(u,3) = datenum(Y,M,diaAtual);
                        u = u+1;
                    catch
                        end
                    end
                end
            end
        end
    catch
        end
    end
    p = p+1;
end
SerieHistorica = SerieHistorica(SerieHistorica(:,3)~=0,:);
SerieHistorica = sortrows(SerieHistorica,3);
Faz a média das 3 medições diárias:
dias = unique(SerieHistorica(:,3));
SerieHistorica2 = zeros(size(dias,1),3);
for i = 1:size(dias,1)
    idx = find(SerieHistorica(:,3)==dias(i,1));
    SerieHistorica2(i,1) = mean(SerieHistorica(idx,1));
    SerieHistorica2(i,2) = mean(SerieHistorica(idx,2));
    SerieHistorica2(i,3) = dias(i,1);
end
clear SerieHistorica
SerieHistorica = SerieHistorica2;
SerieHistorica = sortrows(SerieHistorica,3);
Contents
Lê o arquivo xml e armazena em uma variável:
Identifica o nº de meses com medição de vazão:
Faz a leitura dos dados:
function [SerieHistorica,leitura] = intxmlV2(filename)
Lê o arquivo xml e armazena em uma variável:
leitura = 'fail';
while strcmp(leitura,'fail')
    try
        xDoc = xmlread(filename);
        leitura = 'OK';
    catch
        warning('Erro de leitura no arquivo %s',filename);
        SerieHistorica = [];
        return
    end
end
end
Identifica o nº de meses com medição de vazão:
size = xDoc.item(0).item(3).item(1).getLength;

```

```

if size < 10
    SerieHistorica = [];
    return
end
Faz a leitura dos dados:
SerieHistorica(:,1) - Dados de Vazão; SerieHistorica(:,2) - Nivel de Consistencia;
SerieHistorica(:,3) - Data;
SerieHistorica = zeros(size*31,3);
p = 1;
u = 1;
for i = 1:2:size
    confSerie = sprintf('SerieHistorica%d',p);
    try
        if strcmp(xDoc.item(0).item(3).item(1).item(i).getAttribute('diffgr:id'),confSerie)
            % SerieHistorica(:,3) - Data;
            data_i =
datenum(char(xDoc.item(0).item(3).item(1).item(i).item(5).item(0).getData), 'yyyy-mm-dd
HH:MM:SS');
            % SerieHistorica(:,2) - Nivel de Consistencia;
            nvc =
str2double(char(xDoc.item(0).item(3).item(1).item(i).item(3).item(0).getData));
            for j = 31:2:91
                try
                    % SerieHistorica(:,1) - Dados de Vazão;
                    SerieHistorica(u,1) =
str2double(char(xDoc.item(0).item(3).item(1).item(i).item(j).item(0).getData));
                    % SerieHistorica(:,2) - Nivel de Consistencia;
                    SerieHistorica(u,2) = nvc;
                    % SerieHistorica(:,3) - Data;
                    SerieHistorica(u,3) = data_i+(0.5*j-14.5)-1;
                    u = u+1;
                catch
                    end
            end
        end
    catch
    end
    p = p+1;
end
SerieHistorica = SerieHistorica(SerieHistorica(:,3)~=0,:);
SerieHistorica = sortrows(SerieHistorica,3);
http://www.mathworks.com/products/matlab/
function outfile = v_import_webANA2(codEstacao)
filename = sprintf('v%d.xml',codEstacao);
dataFim = datestr(now(),'dd/mm/yyyy');
url =
sprintf('http://telemetriaws1.ana.gov.br/ServiceANA.asmx/HidroSerieHistorica?codEstacao=%d&
dataInicio=01/01/1900&dataFim=%s&tipoDados=3&nivelConsistencia=2',codEstacao,dataFim);
try
    outfile = websave(filename,url);
catch
    pause(10)
    try
        outfile = websave(filename,url);
    catch
        warning('Falha ao baixar os dados de vazão da estação %d',codEstacao)
        outfile = 'none';
    end
end
Contents
Estabelece o nível de erro aceito:
Cria a tabela de dados Cota x Vazão
Determina o número de amostras de dados:
function [nSerieVazoes,FitTable,fitObj] = KeyCurvFilter5(Cotas1,Cotas2,Vazoes)
Estabelece o nível de erro aceito:

```

```

err = 0.05; % 5% de erro em relação à curva fit.
Cria a tabela de dados Cota x Vazão
warning('off');
try
    % Limpa dados duplicados de cotas e vazões:
    Cotas1(:,3) = floor(Cotas1(:,3));
    Cotas2(:,3) = floor(Cotas2(:,3));
    Vazoes(:,3) = floor(Vazoes(:,3));
    Cotas1 = uniqueRows(Cotas1,3);
    Cotas2 = uniqueRows(Cotas2,3);
    Vazoes = uniqueRows(Vazoes,3);
    % Cria a tabela de dados:
    Tab1 = zeros(size(Vazoes,1),5);
    for i = 1:size(Vazoes,1)
        Tab1(i,1) = Vazoes(i,3);
        if ~isempty(Cotas1(Cotas1(:,3)==Tab1(i,1),1))
            Tab1(i,2) = Cotas1(Cotas1(:,3)==Tab1(i,1),1);
        elseif ~isempty(Cotas2(Cotas2(:,3)==Tab1(i,1),1))
            Tab1(i,2) = Cotas2(Cotas2(:,3)==Tab1(i,1),1);
        else
            Tab1(i,2) = 0;
        end
        Tab1(i,3) = Vazoes(i,1);
        Tab1(i,4) = Vazoes(i,2);
    end
    Tab2 = Tab1(Tab1(:,2)~=0,:);
    Tab2 = Tab2(Tab2(:,3)>0,:);
Determina o número de amostras de dados:
    if size(Tab2,1)/360 > 30
        ndA = 3*200;
    elseif size(Tab2,1)/360 > 20
        ndA = 2*200;
    elseif size(Tab2,1)/360 > 10
        ndA = 200;
    else
        ndA = 150;
    end
    ndA = min([ndA size(Tab2,1)]);
    nA = floor(size(Tab2,1)/ndA);
    iA = ones(nA,2);
    for i = 1:size(iA,1)
        iA(i,1) = (i-1)*ndA+1;
        iA(i,2) = i*ndA;
    end
    Tab4 = zeros(size(Tab2,1),size(nA,1));
    Tab5 = zeros(size(Tab2,1),size(nA,1));
    for i = 1:nA
        amostra = Tab2(iA(i,1):iA(i,2),[2 3]);
        % cria a função da Curva-chave:
        type = fittype('a*(x-b)^c');
        CUV = [];
        tentativa = 1;
        while isempty(CUV) && tentativa <= 5
            try
                CUV = fit(amostra(:,1),amostra(:,2),type,'StartPoint',[0 50 1.5]);
            catch
                pause(1)
                amostra = sortrows(amostra,2);
                amostra = amostra(2:end-1,:);
            try
                CUV = fit(amostra(:,1),amostra(:,2),type);
            catch
                pause(2.5)
            try
                amostra = sortrows(amostra,1);

```



```

        for j = 11:size(amostra,1)
            amostra(j,3) = abs(A(j,2)-mean(A(j-10:j,2)))/mean(A(j-10:j,2));
        end
        amostra = amostra(amostra(:,3)<0.5,[1 2]);
        CUV = fit(amostra(:,1),amostra(:,2),type);
    catch
        try
            amostra = uniqueRows(amostra,2);
            CUV = fit(amostra(:,1),amostra(:,2),type);
        catch
            end
        end
    end
    end
    tentativa = tentativa +1;
end
Tab4(:,i) = abs(Tab2(:,3)-CUV(Tab2(:,2)))./Tab2(:,3);
Tab5(:,i) = Tab4(:,i)<err; % <10% de erro, então pertence ao grupo i.
fitObj.(sprintf('CUV%1.0f',i)) = CUV;
end
% Determina qual o fit selecionado para cada dado:
[Tab3(:,1),Tab3(:,2)] = min(Tab4,[],2);
Tab3 = Tab3(Tab3(:,1)<err,:);
% Determina se o fit é utilizado:
fields = fieldnames(fitObj);
ift = 1;
for i = 1:size(fields,1)
    nfit = sum(Tab3(:,2)==i);
    if nfit < 0.05*size(Tab2,1)
        fitObj = rmfield(fitObj,fields{i,1});
    else
        idxfit(ift,1) = i;
        ift = ift+1;
    end
end
end
Tab6 = logical(sum(Tab5(:,idxfit),2)); % Seleciona os dados com erro inferior ao
aceitavel.
Tab7 = Tab2(Tab6,:);
FitTable.Bruto = Tab1;
FitTable.Tratado = Tab7;
nSerieVazoes = [Tab7(:,3) Tab7(:,1)];
catch
    nSerieVazoes = Vazoes(Vazoes(:,2)==2,[1 3]);
    FitTable = [];
    fitObj = [];
end
warning('on');
Contents
Datas para comparação:
Contagem de dias preenchidos por mês
define os meses que foram preenchidos:
function LMP = MesesPreenchidos(VpreE)
data_inicial = VpreE(1,2);
data_final = VpreE(end,2);
Datas para comparação:
[Y,M,~] = datevec(data_inicial);
for i = 1:months(data_inicial,data_final)
    Vmeses(i,1) = datenum(Y,M,1);
    M = M+1;
    if M == 13
        M = 1;
        Y = Y+1;
    end
end
end
Contagem de dias preenchidos por mês

```

```

[Y2,M2,~] = datevec(VpreE(:,2));
VpreE(:,3) = datenum(Y2,M2,ones(size(Y2,1),1));
Vcontador = Vmeses;
Vcontador(:,2) = 0;
for i = 1:size(VpreE,1)
    for j = 1:size(Vmeses,1)
        if VpreE(i,3) == Vmeses(j,1)
            Vcontador(j,2) = Vcontador(j,2) + 1;
        end
    end
end
end
define os meses que foram preenchidos:
LMP = Vcontador(Vcontador(:,2)>=10,1);
http://www.mathworks.com/products/matlab/
function VmedMes = QmedMensal(ERef)

for M = 1:12
    j = 1;
    for i = 1:size(ERef.Vazoes,1)
        [~,Mi] = datevec(ERef.Vazoes(i,2));
        if M == Mi
            MQ(j,M) = ERef.Vazoes(i,1);
            j = j+1;
        end
    end
end
end
VmedMes = mean(MQ,1)';
Contents
Identifica as estações com correlação superior a 80%:
Criar Ranking de estações dando pontuação:
function [Estacoes, CTable] = Qualifilter3(Estacoes)
% Estacoes é struct:
% Estacoes.EXXXXXXXX - Código da Estação;
% Estacoes.EXXXXXXXX.Vazoes - Série de Vazões e datas;
Identifica as estações com correlação superior a 80%:
Identifica os códigos das estações:
fields = fieldnames(Estacoes);
% Declara variáveis:
QuaCorr = zeros(size(fields,1),size(fields,1));
CTable{1,1} = '';
k = 1;
% Para cada estação...
for i = 1:size(fields,1)
    % Comparada com cada uma das outras estações...
    for j = 1:size(fields,1)
        % Identificar datas comuns de medição entre duas estações:
        [IntData,idVi,idVj] =
intersect(Estacoes.(fields{i,1}).VazoesMensais(:,2),Estacoes.(fields{j,1}).VazoesMensais(:,
2));
        % se existirem mais de 5 anos de dados em comum...
        if size(IntData,1)>5*12
            % cria um vetor com os dados de vazão da estação "i"
            V1 = Estacoes.(fields{i,1}).VazoesMensais;
            % cria outro vetor com os dados de vazão da estação "j"
            V2 = Estacoes.(fields{j,1}).VazoesMensais;
            % cria um vetor com os dados de vazão "i" que tem datas em
            % comum com a estação "j".
            A = sortrows(V1(idVi,:),2);
            % cria um vetor com os dados de vazão "j" que tem datas em
            % comum com a estação "i".
            B = sortrows(V2(idVj,:),2);
            % Avalia a correlação entre os dados com datas em comum.
            [~,Gof] = fit(A(:,1),B(:,1),'poly1');
            % Defini o valor de r² da correlação.
            QuaCorr(i,j) = Gof.rsquare;
        end
    end
end
end

```

```

% Se o valor de r² >= 0.8, a correlação é boa e é inserida na
% tabela de Correlações:
% [CodEstacao01 CodEstacao02 r²]
if Gof.rsquare >= 0.8 && i~=j
    CTable{k,1} = fields{i,1};
    CTable{k,2} = fields{j,1};
    CTable{k,3} = Gof.rsquare;
    k = k+1;
end
end
end
end
Criar Ranking de estações dando pontuação:
Define correlações de qualidade como aquelas com R² > 0.8
MCQ = double(QuaCorr>=0.8);
% Vetor de correlações de qualidade por estação.
VCQ = sum(MCQ)-1;
sVCQ = unique(VCQ);
% Define o Ranking de qualidade e elimina aquelas que não atendem o mínimo:
j = 1;
for i = 1:size(fields,1)
    for r1 = 1:size(sVCQ,2)
        if VCQ(1,i)==sVCQ(1,r1)
            if sVCQ(1,r1)==0
                fields2del(j,1) = i;
                j = j+1;
            else
                Estacoes.(fields{i,1}).rank(1,1) = r1;
            end
        end
    end
end
end
for i = 1:j-1
    Estacoes = rmfield(Estacoes,fields{fields2del(i,1),1});
end
Contents
Aplicar critérios de desempate se necessário:
function [Estacoes, Winner] = RankStation(Estacoes,AreaDrenagem,forceWinner)
fields = fieldnames(Estacoes);
for i = 1:size(fields,1)
    % Ranking de Qualidade (já calculado) - peso 1 - (maior melhor)
    Estacoes.(fields{i,1}).rank(1,2) = 1;
    % Ranking de quantidade de dados - peso 1 - (maior melhor)
    Estacoes.(fields{i,1}).rank(2,1) = size(Estacoes.(fields{i,1}).Vazoes,1);
    Estacoes.(fields{i,1}).rank(2,2) = 1;
    % Ranking de Area de drenagem - peso 2 - (próximo de 1 melhor)
    Estacoes.(fields{i,1}).rank(3,1) = AreaDrenagem/Estacoes.(fields{i,1}).AreaDrenagem;
    Estacoes.(fields{i,1}).rank(3,2) = 2;
    % Ranking de distância - peso 2 - (menor melhor)
    Estacoes.(fields{i,1}).rank(4,1) = Estacoes.(fields{i,1}).distancia;
    Estacoes.(fields{i,1}).rank(4,2) = 2;
    % Classificação de pontos para os Rankings:
    % Qualidade:      1 = 1 pt. | 2 = 2 pt. | (>=3) = 3 pt.
    if Estacoes.(fields{i,1}).rank(1,1) >= 3
        Estacoes.(fields{i,1}).rank(1,3) = 3;
    elseif Estacoes.(fields{i,1}).rank(1,1) == 2
        Estacoes.(fields{i,1}).rank(1,3) = 2;
    else
        Estacoes.(fields{i,1}).rank(1,3) = 1;
    end
    % Quantidade:      (<=180) = 1 pt. | 181-300 = 2 pt. | (>300) = 3 pt.
    if Estacoes.(fields{i,1}).rank(2,1) > 25*365 % mais de 25 anos
        Estacoes.(fields{i,1}).rank(2,3) = 3;
    elseif Estacoes.(fields{i,1}).rank(2,1) > 15*365 % mais de 15 anos
        Estacoes.(fields{i,1}).rank(2,3) = 2;

```

```

else
    Estacoes.(fields{i,1}).rank(2,3) = 1; % menos de 15 anos
end
% Area de Drenagem: (<0.5 ou >1.5) = 1 pt. | (0.5-0.8 e 1.2-1.5) = 2 pt. | 0.8-1.2
= 3 pt.
if Estacoes.(fields{i,1}).rank(3,1) >= 0.8 && Estacoes.(fields{i,1}).rank(3,1) <= 1.2
    Estacoes.(fields{i,1}).rank(3,3) = 3;
elseif Estacoes.(fields{i,1}).rank(3,1) >= 0.5 && Estacoes.(fields{i,1}).rank(3,1) <=
1.5
    Estacoes.(fields{i,1}).rank(3,3) = 2;
else
    if Estacoes.(fields{i,1}).rank(3,1) <= 0.1 || Estacoes.(fields{i,1}).rank(3,1) >=
10
        Estacoes.(fields{i,1}).rank(3,3) = 0.5;
    else
        Estacoes.(fields{i,1}).rank(3,3) = 1;
    end
end
% Distância: >0.2 = 1 pt. | (>0.1 e <= 0.2) = 2 pt. | (<=0.1) = 3 pt.
if Estacoes.(fields{i,1}).rank(4,1) <= 10000 % 10 km
    Estacoes.(fields{i,1}).rank(4,3) = 3;
elseif Estacoes.(fields{i,1}).rank(4,1) <= 20000 % 20 km
    Estacoes.(fields{i,1}).rank(4,3) = 2;
else
    Estacoes.(fields{i,1}).rank(4,3) = 1;
end
% Calcula a pontuação final de cada rank com base nos pesos.
Estacoes.(fields{i,1}).rank(:,4) =
Estacoes.(fields{i,1}).rank(:,2).*Estacoes.(fields{i,1}).rank(:,3);
% Calcula a ranking final de cada estação de acordo com a soma das
% pontuações
Estacoes.(fields{i,1}).PTRank = sum(Estacoes.(fields{i,1}).rank(:,4));
% Determina a estação vencedora por by-pass de pontuação:
if strcmp(fields{i,1},forceWinner)
    Estacoes.(fields{i,1}).PTRank = 99;
end
VPTRANK(i,1) = Estacoes.(fields{i,1}).PTRank;
end
Aplicar critérios de desempate se necessário:
nEmpate = size(VPTRANK(VPTRANK(:,1)==max(VPTRANK),:),1);
j = 1;
if nEmpate > 1
    % Selecionar Estações que empataram
    for i = 1:size(fields,1)
        if Estacoes.(fields{i,1}).PTRank == max(VPTRANK)
            % Criar critério de desempate com base no tamanho da Bacia
            fieldsSelect{j,1} = fields{i,1};
            fieldsSelect{j,2} = abs(Estacoes.(fields{i,1}).rank(3,1)-1);
            j = j+1;
        end
    end
    % Adicionar pontuação de desempate para Estação com melhor tamanho de
    % Bacia
    fieldsSelect = sortrows(fieldsSelect,2);
    Winner = fieldsSelect{1,1};
    Estacoes.(Winner).PTRank = Estacoes.(Winner).PTRank+1;
else
    for i = 1:size(fields,1)
        if Estacoes.(fields{i,1}).PTRank == max(VPTRANK)
            Winner = fields{i,1};
        end
    end
end
Contents
Cria o Vetor de Vazões Mansais:

```

Garante que a estação de referência está com a série de vazões ordenada:
Classifica em um novo vetor as vazões diárias agora com data ano e mês:

```
function Estacao = SerieVazMensal(Estacao)
```

Cria o Vetor de Vazões Mensais:

```
Estacao.VazoesMensais = [];
```

Garante que a estação de referência está com a série de vazões ordenada:

```
Estacao.Vazoes = sortrows(Estacao.Vazoes,2);
```

Classifica em um novo vetor as vazões diárias agora com data ano e mês:

```
for i = 1:size(Estacao.Vazoes,1)
    [Y,M] = datevec(Estacao.Vazoes(i,2));
    VazoesFlag(i,1) = Estacao.Vazoes(i,1);
    VazoesFlag(i,2) = datenum(Y,M,1);
end
Meses = unique(VazoesFlag(:,2));
for i = 1:size(Meses,1)
    if size(VazoesFlag(VazoesFlag(:,2)==Meses(i,1),1),1)>0.7*30
        VazoesMensais(i,1) = mean(VazoesFlag(VazoesFlag(:,2)==Meses(i,1),1));
        VazoesMensais(i,2) = Meses(i,1);
    end
end
VazoesMensais(VazoesMensais(:,1)==0,:) = [];
Estacao.VazoesMensais = VazoesMensais;
```

<http://www.mathworks.com/products/matlab/>

```
function Matriz = SerieVazaoANEEL(EstacaoLocal)
```

```
Matriz = {'',1,2,3,4,5,6,7,8,9,10,11,12};
```

```
dini = min(EstacaoLocal.Vazoes(:,2));
```

```
dfim = max(EstacaoLocal.Vazoes(:,2));
```

```
[Yi,~,~,~,~,~] = datevec(dini);
```

```
[Yf,~,~,~,~,~] = datevec(dfim);
```

```
j = 2;
```

```
for i = Yi:Yf
```

```
    Matriz{j,1} = i;
```

```
    j = j+1;
```

```
end
```

```
for i = 2:size(Matriz,1)
```

```
    for j = 2:size(Matriz,2)
```

```
        data_i = datenum(Matriz{i,1},Matriz{1,j},1);
```

```
        [Y,M,~] = datevec(data_i);
```

```
        data_f = datenum(Y,M+1,1)-1;
```

```
        Matriz{i,j} =
```

```
mean(EstacaoLocal.Vazoes(logical(double(EstacaoLocal.Vazoes(:,2))>data_i) ...
    .* double(EstacaoLocal.Vazoes(:,2)<=data_f)),1));
```

```
    end
```

```
end
```

<http://www.mathworks.com/products/matlab/>

```
function FunPerm = permanenciaQ2(Vazoes)
```

```
p = (.01:0.01:1)';
```

```
while size(Vazoes,1)<=size(p,1)*2
```

```
    Vazoes = [Vazoes;Vazoes];
```

```
end
```

```
Vazoes = sort(Vazoes,'descend');
```

```
pc = floor(p*size(Vazoes,1));
```

```
if pc(1,1) == 0
```

```
    pc(1,1) = 1;
```

```
end
```

```
Y = Vazoes(pc,1);
```

```
FunPerm = [p Y];
```

Contents

1. Definir equação para preenchimento de falhas e extensão;

```
function Eref = preencher2(Epre,Eref,NomeEstacao)
```

1. Definir equação para preenchimento de falhas e extensão;

```
[~,idVi,idVj] = intersect(Eref.Vazoes(:,2),Epre.Vazoes(:,2));
```

```
V1 = Eref.Vazoes;
```

```
V2 = Epre.Vazoes;
```

```

A = sortrows(V1(idVi,:),2);
B = sortrows(V2(idVj,:),2);
[fObj,gof] = fit(B(:,1),A(:,1),'poly1');
tol = 1;
while gof.rsquare < 0.95
    Asyn = fObj.p1.*B(:,1)+fObj.p2;
    Aerr = abs(A(:,1)-Asyn)./A(:,1);
    [fObj,gof] = fit(B(Aerr < tol,1),A(Aerr < tol,1),'poly1');
    tol = tol - 0.05;
end
% 3. Identificar falhas e verificar se estação pode contribuir com
% preenchimento;
% 4. Identificar possibilidade de contribuição com extensão;
[~,idVi] = setdiff(Epre.Vazoes(:,2),Eref.Vazoes(:,2));
for i = 1:size(idVi,1)
    Vpre(i,1) = fObj.p1*(Epre.Vazoes(idVi(i,1),1))+fObj.p2;
    Vpre(i,2) = Epre.Vazoes(idVi(i,1),2);
end
try
    Eref.Vazoes = sortrows([Eref.Vazoes;Vpre],2);
    Eref.Vpre.(NomeEstacao) = Vpre;
    Eref.VpreFunc.(NomeEstacao) = fObj;
    Eref.VreGof.(NomeEstacao) = gof;
catch
end

```

Contents

Organiza a matriz A de forma crescente, considerando os valores da column;
 Seleciona o índice dos valores únicos da column da matriz A:
 Cria a matriz B, uma versão da Matriz A, mas apenas com valores únicos na column:
 Sobre o código:

```

function B = uniqueRows(A,column)
% uniqueRows limpa Matrizes que possam possuir valores duplicados.
%
% A = [1 5 3
%      2 5 4
%      3 1 7]
%
% B = uniqueRows(A,2)
%
% B = [3 1 7
%      1 5 3]
%
%See also sortrows, unique

```

Organiza a matriz A de forma crescente, considerando os valores da column;
 A = sortrows(A,column);
 Seleciona o índice dos valores únicos da column da matriz A:
 [~,ia,~] = unique(A(:,column));
 Cria a matriz B, uma versão da Matriz A, mas apenas com valores únicos na column:
 B = A(ia,:);