

UNIVERSIDADE FEDERAL DE MINAS GERAIS
Escola de Engenharia
Programa de Pós-Graduação em Saneamento, Meio Ambiente e Recursos Hídricos

Gabriel Farias Neiva

**ESTIMAÇÃO DE VAZÕES EM LOCAIS NÃO MONITORADOS A PARTIR DO EMPREGO
CONJUNTO DE REGRESSÃO FUNCIONAL MÚLTIPLA E KRIGAGEM TOPOLÓGICA**

Belo Horizonte
2024

Gabriel Farias Neiva

**ESTIMAÇÃO DE VAZÕES EM LOCAIS NÃO MONITORADOS A PARTIR DO EMPREGO
CONJUNTO DE REGRESSÃO FUNCIONAL MÚLTIPLA E KRIGAGEM TOPOLÓGICA**

Dissertação apresentada ao Programa de Pós-graduação em Saneamento, Meio Ambiente e Recursos Hídricos da Universidade Federal de Minas Gerais, como requisito parcial à obtenção do título de Mestre em Saneamento, Meio Ambiente e Recursos Hídricos.

Área de concentração: Hidráulica e Recursos Hídricos

Linha de pesquisa: Modelagem de Processos Hidrológicos

Orientador: Veber Afonso Figueiredo Costa

Belo Horizonte
2024

N417e	Neiva, Gabriel Farias. Estimação de vazões em locais não monitorados a partir do emprego conjunto de regressão funcional múltipla e krigagem topológica [recurso eletrônico] / Gabriel Farias Neiva. – 2024. 1 recurso online (133 f. : il., color.) : pdf. Orientador: Veber Afonso Figueiredo Costa. Dissertação (mestrado) – Universidade Federal de Minas Gerais, Escola de Engenharia. Apêndices: f. 101-133. Bibliografia: f. 95-100. 1. Engenharia sanitária – Teses. 2. Recursos hídricos – Teses. 3. Hidrologia – Teses. 4. Previsão hidrológica – Teses. I. Costa, Veber Afonso Figueiredo. II. Universidade Federal de Minas Gerais. Escola de Engenharia. III. Título.
	CDU: 628(043)



UNIVERSIDADE FEDERAL DE MINAS GERAIS
ESCOLA DE ENGENHARIA
PROGRAMA DE PÓS-GRADUAÇÃO EM SANEAMENTO, MEIO AMBIENTE E RECURSOS HÍDRICOS

FOLHA DE APROVAÇÃO

"Estimação de Vazões Em Locais Não Monitorados A Partir do Emprego Conjunto de Regressão Funcional
Múltipla e Krigagem Topológica"

GABRIEL FARIAS NEIVA

Dissertação defendida e aprovada pela banca examinadora constituída pelos Senhores:

Prof. Veber Afonso Figueiredo Costa

Prof. Eber José de Andrade Pinto

Profa Viviane Borda Pinheiro Rocha

Prof. Dirceu Silveira Reis Junior

Aprovada pelo Colegiado do PG SMARH

Versão Final aprovada por

Profa. Priscilla Macedo Moura

Prof. Veber Afonso Figueiredo Costa

Coordenadora

Orientador

Belo Horizonte, 26 de março de 2024.



Documento assinado eletronicamente por **Veber Afonso Figueiredo Costa, Professor do Magistério Superior**, em 26/03/2024, às 12:01, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Dirceu Silveira Reis Junior, Usuário Externo**, em 26/03/2024, às 12:01, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Eber Jose de Andrade Pinto, Professor do Magistério Superior**, em 26/03/2024, às 12:02, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Viviane Borda pinheiro Rocha, Professora do Magistério Superior**, em 16/04/2024, às 09:14, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Eduardo Coutinho de Paula, Coordenador(a) de curso de pós-graduação**, em 05/02/2025, às 13:04, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



A autenticidade deste documento pode ser conferida no site https://sei.ufmg.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador **3132413** e o código CRC **7FA953FF**.

AGRADECIMENTOS

Gostaria de agradecer primeiramente a Deus por me dar forças em todos os momentos.

Aos meus pais, Caio e Benedita, e aos meus irmãos, Matheus e Arthur, por todo apoio e incentivo concedido durante essa jornada.

À Graciele pelo carinho e compreensão durante a elaboração deste trabalho.

Ao meu orientador, professor Veber Costa, pela dedicação, paciência e apoio durante a execução desta dissertação.

Aos professores e funcionários do EHR.

À FAPEMIG, pela bolsa de estudos concedida.

RESUMO

Estimativas de vazões são necessárias a diversas aplicações que compreendem o planejamento e gerenciamento de recursos hídricos, como dimensionamento de obras hidráulicas, concessão de outorgas, projetos de sistemas de abastecimento de água, irrigação e navegação, estudos para lançamento de efluentes e manutenção de vazões ecológicas. Com alguma frequência, tais estimativas são necessárias em locais desprovidos de monitoramento sistemático de vazões, o que torna necessário o emprego de métodos de predição.

Diante deste cenário, neste estudo é proposto um método para estimação de vazões em locais sem monitoramento fluviométrico com base em curvas de permanência regionalizadas e interpolação espacial via Krigagem Topológica. A regionalização das curvas de permanência foi realizada por meio da regressão funcional múltipla. Diferentemente das regressões tradicionais, tal técnica permite estimar as curvas de permanência de maneira integral e com uma única regressão, além de permitir a identificação de métodos de predição com base em características físicas e climáticas das bacias hidrográficas. Uma vez obtida a curva de permanência no local de interesse, foi realizada a transferência de informação hidrológica de estações próximas por meio do método de interpolação espacial não linear conhecido como Top-Kriging e, em seguida, foram geradas as séries sintéticas de vazões diárias.

O método foi aplicado na bacia hidrográfica do Rio São Francisco no estado de Minas Gerais. Os resultados indicaram que a abordagem proposta obteve melhor desempenho nas bacias hidrográficas com características semelhantes aos atributos médios empregados na construção do modelo. Ademais, a qualidade do ajuste das séries sintéticas de vazões e, por consequência, a capacidade de reprodução das vazões máximas anuais, vazões médias mensais e vazões mínimas estiveram diretamente relacionadas com a eficiência do modelo regional das curvas de permanência.

Palavras-chave: Regressão funcional múltipla, Top-Kriging, Regionalização, Curvas de permanência.

ABSTRACT

Streamflow estimates are necessary for a variety of uses in water resources planning and management, such as the design of hydraulic structures, the granting of water uses, the design of water supply systems, irrigation and navigation, and studies for pollutant dilution and ecologic flows. Frequently, such estimates are sought at ungauged locations, which, in turn, would require the identification of prediction models.

In view of this problem, a method for estimating streamflow at ungauged locations from regionalized flow duration curves (FDCs) and topological kriging is developed in this study. Regional FDCs are obtained with functional multiple regression, which, as opposed to traditional regression approaches, allows estimating the quantiles of FDC with a single model identified from physical and climatic characteristics of the catchments. Once the regional FDCs are obtained, the nonlinear spatial interpolation termed Top-Kriging is utilized for transferring information regarding the durations from nearby sites, which allows the generation of synthetic daily streamflow series.

The proposed approach was applied in the São Francisco River catchment, located in the Brazilian state of Minas Gerais. Overall, the performance of the method was appropriate in those catchments that did not considerably deviate from the average values of the covariates utilized during model calibration. Furthermore, the goodness-of-fit of the synthetic streamflow series and, accordingly, the model ability of reproducing the observed annual daily maxima, annual daily minima, and monthly averages were directly related to the suitable estimation of the regional flow duration curves.

Keywords: Functional Multiple Regression, Top-Kriging, Regionalization, Flow Duration Curves

LISTA DE FIGURAS

Figura 1 – Métodos de regionalização baseados em similaridade	19
Figura 2 – Grupo de variáveis que controlam o escoamento	24
Figura 3 – Grade de pontos de uma rede de drenagem utilizada no método da Krigagem Topológica	26
Figura 4 – Efeito da área de drenagem (esquerda) e do aninhamento das bacias hidrográficas (centro e direita) na ponderação pela Krigagem Topológica	27
Figura 5 - Representação da série de vazões diárias (esquerda) pela sua respectiva curva de permanência (direita).....	30
Figura 6 - Aplicação da análise funcional de dados no estudo de crescimento de 10 pessoas. Cada curva está associada a um indivíduo e os pontos retratam às medidas realizadas durante 18 anos	37
Figura 7 - Bacia hidrográfica do Rio São Francisco no estado de Minas Gerais	47
Figura 8 – Zonas climáticas da bacia hidrográfica do Rio São Francisco no estado de Minas Gerais	49
Figura 9 – Precipitação anual acumulada da bacia hidrográfica do Rio São Francisco no estado de Minas Gerais.....	50
Figura 10 – Uso e ocupação do solo na bacia hidrográfica do Rio São Francisco no estado de Minas Gerais	52
Figura 11 – Mapa altimétrico da bacia hidrográfica do Rio São Francisco no estado de Minas Gerais	53
Figura 12 - Mapa de aquíferos presentes na bacia hidrográfica do Rio São Francisco no estado de Minas Gerais.....	55
Figura 13 - Sub-bacias do Rio São Francisco em Minas Gerais e localização dos postos fluviométricos utilizados na pesquisa.....	58
Figura 14 – Representação da curva de permanência empírica em formato funcional	59
Figura 15 - Método de interpolação espacial não linear proposto por Hughes e Smakhtin (1996)	64
Figura 16 – Média da validação cruzada generalizada para um intervalo de valores de λ	70
Figura 17 – FDC da estação 40040000 suavizada para três valores distintos de λ ..	70

Figura 18 – Ramos superior (esquerda) e inferior (direita) da FDC da estação 40040000 suavizada para os valores de λ iguais a 10^{-5} , 10^{-6} e 10^{-7}	71
Figura 19 - Métricas de desempenho do processo de calibração	75
Figura 20 - Desempenho da regressão funcional múltipla para a estação Pirapama (41600000)	76
Figura 21 - Desempenho da regressão funcional múltipla para a estação Juvenília (45260000)	76
Figura 22 - Coeficientes da regressão funcional múltipla.....	78
Figura 23 - Métricas de desempenho do processo de validação	80
Figura 24 - Desempenho da regressão funcional múltipla para a estação Jaguaruna Jusante (40300001)	81
Figura 25 - Desempenho da regressão funcional múltipla para a estação Lagoa das Pedras (45210000).....	82
Figura 26 - Métricas de desempenho das séries de vazões simuladas	86
Figura 27 - Desempenho da série de vazões simuladas na estação Jaguaruna Jusante (40300001)	87
Figura 28 - Desempenho da série de vazões simuladas na estação Lagoa das Pedras (45210000)	87

LISTA DE TABELAS

Tabela 1 – Índices de similaridade utilizados pela literatura	23
Tabela 2 – Classes de uso e ocupação do solo na bacia hidrográfica do Rio São Francisco em MG	51
Tabela 3 – Estações fluviométricas utilizadas no estudo	56
Tabela 4 – Permanências utilizadas na construção das FDCs funcionais	60
Tabela 5 – Estatísticas amostrais das estações fluviométricas utilizadas no estudo	68
Tabela 6 – Características descritivas das estações fluviométricas utilizadas no estudo	72
Tabela 7 – Índices de desempenho da FMR para uma variável dependente	74
Tabela 8 - Q_{MLT} observada versus Q_{MLT} estimada utilizando 3, 4, 5 e 6 estações vizinhas na interpolação espacial.....	84
Tabela 9 - Erros percentuais absolutos médios do desempenho do Top-Kriging	84
Tabela 10 - Erros percentuais absolutos médios para as vazões Q_7 , médias mensais e máximas anuais	89

LISTA DE ABREVIATURAS E SIGLAS

ANA – Agência Nacional de Águas

AFDC - Annual Flow Duration Curve – Curva de Permanência Anual

CBHSF - Comitê da Bacia Hidrográfica do Rio São Francisco

CPRM - Companhia de Pesquisa De Recursos Minerais

FDA - Functional Data Analysis – Análise Funcional de Dados

FDC - Flow Duration Curve – Curva de Permanência de Longo Termo

FMR – Functional Multiple Regression – Regressão Funcional Múltipla

IAHS - International Association of Hydrological Sciences

INMET – Instituto Nacional de Meteorologia

IDW - Inverse Distance Weighted – Ponderação Pelo Inverso da Distância

PSBI - Physiographical-Space Based Interpolation - Interpolação Espacial Baseada no Espaço Fisiográfico

Q_{MLT} - Vazão Média de Longo Termo

Q_7 - Vazão Média Mínima de 7 dias de Duração

SISEMA – Sistema Estadual de Meio Ambiente e Recursos Hídricos

TND - Total Negative Deviation - Desvio Total Negativo

SUMÁRIO

1	INTRODUÇÃO	14
2	OBJETIVOS	17
2.1	Objetivo Geral	17
2.2	Objetivos Específicos	17
3	REVISÃO DA LITERATURA	18
3.1	REGIONALIZAÇÃO HIDROLÓGICA	18
3.1.1	Métodos baseados em similaridade	18
3.1.2	Top-Kriging	24
3.2	CURVA DE PERMANÊNCIA	29
3.2.1	Definição	29
3.2.2	Histórico de utilização das curvas de permanência	30
3.2.3	Construção das curvas de permanência	31
3.2.4	Regionalização das curvas de permanência	32
3.3	ANÁLISE FUNCIONAL DE DADOS	36
3.3.1	Definição	36
3.3.2	Representação dos dados funcionais	38
3.3.3	Suavização dos dados	41
3.3.4	Histórico de utilização da FDA no contexto hidrológico	43
3.3.5	Regressão funcional múltipla	44
4	METODOLOGIA / MATERIAL E MÉTODOS	46
4.1	ÁREA DE ESTUDO	46
4.1.1	Clima	47
4.1.2	Uso e Cobertura do Solo e Relevo	51
4.1.3	Hidrogeologia	54
4.2	ESTAÇÕES FLUVIOMÉTRICAS UTILIZADAS NO ESTUDO	56
4.3	SUAVIZAÇÃO DAS CURVAS DE PERMANÊNCIA	59
4.4	ELABORAÇÃO DO MODELO REGIONAL DAS CURVAS DE PERMANÊNCIA	61
4.5	ESTIMAÇÃO DAS SÉRIES DE VAZÕES DIÁRIAS SINTÉTICAS	63
4.6	AVALIAÇÃO DA METODOLOGIA EMPREGADA	66
5	RESULTADOS E DISCUSSÃO	68
5.1	ELABORAÇÃO DAS CURVAS DE PERMANÊNCIA FUNCIONAIS	68
5.1.1	Definição das estações fluviométricas utilizadas no estudo	68
5.1.2	Representação das curvas de permanência empíricas em termos funcionais	69

5.2	REGIONALIZAÇÃO DAS CURVAS DE PERMANÊNCIA DE LONGO TERMO FUNCIONAIS.....	72
5.2.1	Características físicas, climáticas e morfológicas utilizadas.....	72
5.2.2	Calibração e avaliação de desempenho do modelo regional de curvas de permanência	73
5.2.3	Análise das variáveis dependentes do modelo regional.....	77
5.2.4	Análise de desempenho do processo de validação do modelo regional ..	79
5.3	GERAÇÃO DAS SÉRIES SINTÉTICAS DE VAZÕES DIÁRIAS.....	83
5.3.1	Aplicação da interpolação espacial via Krigagem Topológica	83
5.3.2	Análise das séries sintéticas de vazões diárias estimadas	85
5.3.3	Análise das vazões de interesse prático estimadas	88
6	CONCLUSÕES	91
7	RECOMENDAÇÕES.....	94
	REFERÊNCIAS.....	95
	APÊNDICE A – AVALIAÇÃO DA INSERÇÃO DAS COVARIÁVEIS NA CONSTRUÇÃO DO MODELO REGIONAL.....	101
	APÊNDICE B – CURVAS DE PERMANÊNCIA REGIONALIZADAS PELA FMR (ETAPA DE CALIBRAÇÃO)	102
	APÊNDICE C – CURVAS DE PERMANÊNCIA REGIONALIZADAS PELA FMR (ETAPA DE VALIDAÇÃO).....	119
	APÊNDICE D – PESOS UTILIZADOS NA GERAÇÃO DAS SÉRIES DE VAZÕES DIÁRIAS	126
	APÊNDICE E – SÉRIES DE VAZÕES DIÁRIAS SIMULADAS	127

1 INTRODUÇÃO

Estimativas de vazões em bacias desprovidas de monitoramento sistemático são um dos principais problemas enfrentados pela Engenharia de Recursos Hídricos. Profissionais da área têm se empenhado em propor soluções para a resolução de tal impasse, como, por exemplo, a iniciativa da Associação Internacional de Ciências Hidrológicas (IAHS - *International Association of Hydrological Sciences*) intitulada “Década de Predições em Bacias Não Monitoradas, 2003-2012” (*Decade on Predictions in Ungauged Basins, 2003-2012*), que teve como objetivo direcionar a comunidade científica para uma melhor compreensão dos processos hidrológicos e das incertezas associadas aos modelos de predição. Tal iniciativa foi continuada na década seguinte (2013 a 2022) e foi denominada “Panta Rhei – Tudo Flui” (*Panta Rhei—Everything Flows*). Seu objetivo foi pesquisar sobre as mudanças na hidrologia e na sociedade, melhorando a interpretação sobre a variação dos processos hidrológicos em conexão com as ações antrópicas observadas (MONTANARI *et al.*, 2013).

A compreensão do regime de vazões em uma determinada região é necessária a diversas aplicações relacionadas ao planejamento e ao gerenciamento de recursos hídricos. O estudo de vazões máximas torna-se necessário em diversas obras de engenharia, como extravasores de barragens, condutos, canais, bueiros, desvios de cursos de água e controle de cheias (PINHEIRO, 2009; TUCCI, 2009). O conhecimento da magnitude e frequência de vazões mínimas é importante para o planejamento e projeto de sistemas de abastecimento de água, reservatórios de armazenamento, alocação de cargas de resíduos, manutenção da quantidade e qualidade da água para irrigação, recreação e conservação do ecossistema (MATONSE E KROLL, 2013; OUARDA *et al.*, 2018; SMAKHTIN, 2001).

Algumas vazões de referência podem ser obtidas das chamadas curvas de permanências. Tais curvas são amplamente utilizadas por engenheiros e hidrólogos em todo o mundo em inúmeras aplicações relacionadas ao gerenciamento de recursos hídricos (CASTELLARIN *et al.*, 2006). As curvas de permanência descrevem a relação entre magnitude e frequência das vazões de um curso d’água. Por meio delas, pode-se associar cada vazão ao percentual do tempo que a mesma foi igualada ou superada em um determinado período de observação (COSTA *et al.*, 2011).

A construção das referidas curvas está condicionada à disponibilidade de dados fluviométricos na região de interesse. Frequentemente, contudo, é necessário realizar estimativas de vazões de referência em locais com pouco ou nenhum monitoramento hidrológico. Para contornar esse problema, em geral, recorre-se a técnicas de regionalização hidrológica, a qual consiste na transposição de informações de uma região monitorada para outra sem monitoramento sistemático.

Usualmente, modelos de regionalização tradicionais compreendem a estimação de um quantil específico (por exemplo, a $Q_{7,10}$) ou um número limitados deles. Outra técnica utilizada consiste em estimar vazões de referência por meio de uma curva de permanência ou curva de quantis regionalizada construída a partir de locais monitorados. No entanto, apenas informações parciais podem ser obtidas dessas abordagens devido ao foco em um determinado quantil. Por exemplo, uma curva de permanência regional que possibilida estimar a Q_{95} não pode ser diretamente empregada para se obter outros quantis de vazões mínimas, como a $Q_{7,10}$ (REQUENA *et al.*, 2018a).

Diante desse cenário, neste trabalho é proposto um método para a estimação de vazões mediante regionalização da curva de permanência, por meio da regressão funcional múltipla (REQUENA *et al.*, 2018b), e interpolação espacial não linear, via Krigagem Topológica. Diferentemente das regressões tradicionais, a regressão funcional múltipla permite a construção das referidas curvas de maneira integral, na qual todos os quantis estão conectados, e com uma única regressão. Uma vez obtida a curva de permanência regional, é possível realizar a transferência de informação hidrológica de estações próximas por meio de interpolação espacial não linear. Esse método tem como vantagem a obtenção da série de vazões diárias para o local de destino. Portanto, a análise não fica restrita a apenas um quantil específico, uma vez que qualquer vazão de referência pode facilmente ser obtida a partir da série estimada regionalmente.

O referido método foi aplicado na bacia hidrográfica do Rio São Francisco em Minas Gerais. Tal escolha se deu devido à disponibilidade de dados fluviométricos que viabilizassem o desenvolvimento da regionalização hidrológica das curvas de

permanência de longo termo, bem como a possibilidade de avaliar o comportamento do modelo em regiões com características físicas e climáticas distintas.

A dissertação em questão está dividida em 7 capítulos, além deste designado à introdução. O Capítulo 2 menciona os objetivos geral e específicos do estudo realizado. O Capítulo 3 apresenta a revisão de literatura, que compreende: (1) métodos de regionalização hidrológica baseada em similaridade, assim como a descrição da Krigagem Topológica, (2) curvas de permanência, histórico de utilização e métodos de regionalização, e (3) os principais conceitos da Análise Funcional de Dados, suas aplicações na hidrologia e a construção da Regressão Funcional Múltipla. No Capítulo 4 é descrita a metodologia proposta e as características da bacia hidrográfica avaliada. No Capítulo 5 são apresentados os resultados e a discussão da execução da metodologia. Por fim, os Capítulos 6 e 7 descrevem as conclusões do trabalho e as recomendações para estudos posteriores.

2 OBJETIVOS

2.1 Objetivo Geral

Desenvolver um método para estimar séries de vazão em locais não monitorados, com base em curvas de permanência de longo termo regionalizadas e princípios de interpolação espacial, de maneira a se obter, a partir de um único modelo, estimativas de diferentes vazões.

2.2 Objetivos Específicos

- Analisar o emprego da regressão funcional múltipla na estimação de curvas de permanência de longo termo em locais não monitorados;
- Avaliar a influência de cada covariável na construção do modelo regionalizado; e
- Avaliar a utilização da Krigagem Topológica na estimação de variáveis hidrológicas e na concepção de séries de vazões em conjunto com curvas de permanência regionais;

3 REVISÃO DA LITERATURA

3.1 REGIONALIZAÇÃO HIDROLÓGICA

O processo de regionalização pode ser entendido como a transferência de informação hidrológica de uma bacia hidrográfica para outra. Tal processo pode ser adequado se as bacias forem consideradas semelhantes, mas está sujeito a erros se não o forem (BLÖSCHL E SIVAPALAN, 1995).

Segundo Tucci (2009), a regionalização hidrológica compreende um conjunto de ferramentas que visam extrair ao máximo as informações existentes de bacias hidrográficas, com o objetivo de se estimar as variáveis hidrológicas em locais com pouca ou nenhuma informação disponível. Ainda segundo esse autor, a regionalização pode ser utilizada para melhor explorar as amostras pontuais e, por consequência, refinar as estimativas das variáveis, verificar a consistência da série de dados hidrológicos e identificar a ausência de postos de observação.

Conforme Tucci (2009), são passíveis de regionalização:

- Funções estatísticas de variáveis hidrológicas, como curvas de probabilidade de vazões máximas, médias e mínimas;
- Funções específicas que relacionam variáveis, tais como curvas de regularização, curvas de infiltração e curvas de permanência; e
- Parâmetros de modelos hidrológicos.

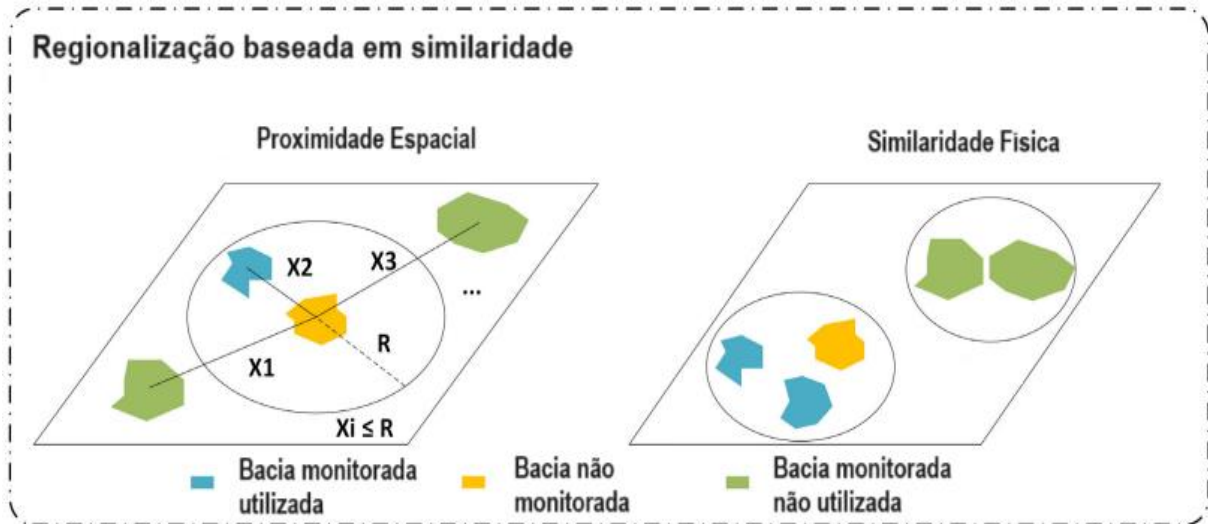
Os métodos de regionalização hidrológica podem ser classificados em três categorias: similaridade, regressão e assinaturas hidrológicas (DASGUPTA, 2023; GUO *et al.*, 2021). Devido ao foco do trabalho em questão, a revisão irá se limitar aos métodos de regionalização por similaridade.

3.1.1 Métodos baseados em similaridade

A regionalização baseada em similaridade presume que bacias hidrográficas com alto grau de similaridade se comportam de forma semelhante. Os métodos que compreendem tais princípios podem ser categorizados em (i) proximidade espacial e (ii) similaridade física. O primeiro pode ser subdividido em (a) proximidade simples e (b) interpolação espacial, em geral pelo método da Ponderação pelo Inverso da

Distância (IDW - *Inverse Distance Weighted*) ou pelo método da Krigagem (Kriging) (DASGUPTA *et al.*, 2023; GUO *et al.*, 2021). A Figura 1 ilustra os conceitos que compreendem tais métodos.

Figura 1 – Métodos de regionalização baseados em similaridade



Fonte: Guo *et al.* (2021) (Modificado)

A abordagem de proximidade espacial considera que bacias hidrográficas próximas umas das outras devem apresentar comportamento semelhante, uma vez que as respostas hidrológicas das bacias geralmente variam de forma gradual e suave no espaço. Por outro lado, a abordagem que considera a similaridade física parte do mesmo princípio, no entanto, como o próprio termo sugere, a similaridade está amparada nas características físicas das bacias analisadas (DASGUPTA *et al.*, 2023; OUDIN *et al.*, 2008).

3.1.1.1 Proximidade Espacial

A aplicação do método da proximidade simples pressupõe que bacias hidrográficas contidas em um espaço geográfico, definido por um raio específico, apresentam comportamento semelhante e a diferença nos valores dos seus parâmetros, de um dado modelo hidrológico, são resultados de fatores aleatórios (RAZAVI E COULIBALY, 2013). A distância Euclidiana, comumente utilizada para definir a distância geográfica entre duas bacias distintas, é descrita pela Equação 1 (HE *et al.*, 2011).

$$d_{t,d} = \sqrt{(X_t - X_d)^2 + (Y_t - Y_d)^2} \quad (1)$$

na qual X_t, Y_t e X_d, Y_d são as coordenadas geográficas dos centroides das bacias não monitorada e monitorada, respectivamente. Os parâmetros do modelo hidrológico da bacia hidrográfica alvo podem ser obtidos de uma bacia ou de um conjunto de bacias monitoradas mais próximas. Esse processo pode ser realizado de duas formas distintas:

- Média dos parâmetros: primeiramente, um conjunto de parâmetros regionais é computado a partir da média dos parâmetros das bacias hidrográficas doadoras (monitoradas). Em seguida, os parâmetros obtidos são aplicados no modelo hidrológico para a bacia alvo (não monitorada); e
- Média dos resultados: o modelo hidrológico é executado na bacia não monitorada utilizando cada conjunto de parâmetros das bacias doadoras para, posteriormente, ser calculada a média dos resultados, sendo considerada como a resposta hidrológica da bacia de interesse.

A ponderação pelo inverso da distância (IDW) considera um conjunto de pontos conhecidos distribuídos em torno de um ponto desconhecido, cujo valor é calculado pela média ponderada dos valores dos pontos conhecidos (SWAIN E PATRA, 2017). Tratando-se de estudos envolvendo bacias hidrográficas, o método baseia-se no inverso da distância espacial entre centroides de bacias hidrográficas. A distância espacial é calculada por meio das coordenadas geográficas dos centroides das bacias analisadas, em geral, expressa pela medida da distância Euclidiana definida na Equação 1. Logo, a variável de interesse pode ser estimada pela Equação 2 (KANISHKA E ELDHO, 2020).

$$P_j = \sum_{i=1}^n W_i p_i \quad (2)$$

na qual P_j é a variável de interesse da bacia alvo, p_i são os valores observados para bacias monitoradas, n é o número de bacias monitoradas utilizadas na regionalização e W_i é uma função peso de cada bacia doadora, calculada conforme a Equação 3.

$$W_i = \frac{d_i^{-2}}{\sum_{i=1}^n d_i^{-2}} \quad (3)$$

na qual d_i é a distância entre o centroide da bacia doadora e o centroide da bacia de interesse.

O método da Krigagem (Kriging) é um método geoestatístico de interpolação espacial, mais sofisticado que a abordagem IDW, no qual o valor de uma variável espacial não medida é estimada como uma média ponderada dos valores observados das regiões adjacentes (SKØIEN *et al.*, 2006; SWAIN E PATRA, 2017). A similaridade espacial dos dados é representada pela variância entre pares de pontos em função da sua distância Euclidiana. Consequentemente, a interpolação por krigagem é tida como um dos melhores estimadores não enviesados, isso é, um método de interpolação em que o viés esperado é zero e o erro esperado pela krigagem é minimizado (SKØIEN *et al.*, 2006).

O valor de uma variável desconhecida $\hat{z}(x_0)$ sobre um ponto x_0 pode ser estimado como a média ponderada definida pela Equação 4.

$$\hat{z}(x_0) = \sum_{i=1}^n \lambda_i z(x_i) \quad (4)$$

em que λ_i é o peso da interpolação do valor medido no ponto x_i e n é o número de postos observados utilizados no processo de interpolação. Os valores de λ_i podem ser estimados resolvendo o sistema composto pelas Equações 5 e 6.

$$\sum_{j=1}^n \lambda_j \gamma_{ij} - \lambda_i \sigma_i^2 + \mu = \gamma_{0j} \quad i = 1, \dots, n \quad (5)$$

$$\sum_{j=1}^n \lambda_j = 1 \quad (6)$$

em que γ_{ij} é a semivariância entre duas medidas i e j , obtida a partir de um modelo teórico de semivariograma, μ é o multiplicador de Lagrange, σ_i^2 representa o erro de

medição ou a incerteza da medida i e o termo γ_{0j} é a semivariância entre o ponto zero e uma medida i (SKØIEN *et al.*, 2006).

Samuel *et al.* (2011) compararam seis técnicas de regionalização de vazões para bacias hidrográficas localizadas na província de Ontario, Canadá. Os métodos compreenderam: IDW, Krigagem, Proximidade Simples, Regressão Múltipla Linear, Similaridade Física (citada adiante) e uma combinação entre o primeiro e o último método citados. As abordagens que demonstraram melhor desempenho foram a Krigagem, o IDW e o método que combina esse último com a Similaridade Física. Analisando o método combinado, este revelou-se superior aos outros dois métodos de proximidade espacial. Os resultados também sugerem que a proximidade espacial entre as bacias analisadas se provou mais relevante para a construção do modelo do que as características físicas das bacias (SAMUEL *et al.*, 2011)

Swain e Patra (2017) analisaram as vazões estimadas pelos métodos de proximidade espacial. O estudo foi realizado em bacias hidrográficas da Índia, as quais foram classificadas em regiões homogêneas pelo algoritmo de Redes Neurais Self-Organizing Map (SOM). Conforme os autores, os métodos Kriging e IDW apresentaram resultados próximos e superiores aos da regionalização pelo método da proximidade simples. Os autores ressaltam que os métodos de regionalização são mais efetivos quando aplicados em bacia hidrográficas classificadas.

Kanishka e Eldho (2020) realizaram um estudo de estimativas de vazões, combinando métodos de classificação de bacias e métodos de regionalização (Proximidade Simples, IDW e Similaridade Física) dos parâmetros do modelo SWAT. O desempenho da Similaridade Física foi superior aos outros dois métodos de regionalização por proximidade espacial. Segundo os autores, caso seja utilizada uma extensa quantidade de atributos no processo de classificação das bacias, as técnicas de proximidade espacial não conseguem atingir resultados tão eficientes quanto a técnica de Similaridade Física.

3.1.1.2 Similaridade Física

A abordagem da similaridade física fundamenta-se na transferência de informações hidrológicas de bacias monitoradas que são semelhantes às bacias não monitoradas

em termos de características físicas, tais como: propriedades e uso e cobertura dos solos, área e topografia, entre outros. A sua utilização baseia-se no fato de que bacias como características semelhantes deveriam se comportar de maneira semelhante. (OUDIN *et al.*, 2008). Outro aspecto importante desse método é considerar que bacias hidrográficas distantes entre si ainda podem apresentar comportamento semelhante (YANG *et al.*, 2018).

O grau de similaridade entre duas bacias pode ser medido com base em um índice calculado a partir das características físicas medidas. Quanto menor for o índice calculado entre a bacia monitorada e a bacia de interesse, mais semelhantes elas serão consideradas. A literatura apresenta algumas formas de cálculo de tal índice, a Tabela 1 apresenta um resumo das equações utilizadas em alguns estudos de regionalização por similaridade física.

Tabela 1 – Índices de similaridade utilizados pela literatura

Autores	Equação Utilizada
Burn e Boorman (1993)	$IS = \sum_{i=1}^k \frac{ x_{a,i} - x_{b,i} }{\Delta x_i}$
He <i>et al.</i> (2011)	$IS = \sqrt{\sum_{i=1}^k \omega_i \left(\frac{x_{a,i} - x_{b,i}}{\sigma_i} \right)^2}$
Kay <i>et al.</i> (2011)	$IS = \sqrt{\sum_{i=1}^k \left(\frac{x_{a,i} - x_{b,i}}{\sigma_i} \right)^2}$
Li <i>et al.</i> (2014)	$IS = \sqrt{\sum_{i=1}^k \omega_i \left(\frac{x_{a,i} - x_{b,i}}{\sigma_i} \right)^2}$
Kanishka e Eldho (2020)	$IS = \sqrt{\sum_{i=1}^k \omega_i \left(\frac{x_{a,i} - x_{b,i}}{\sigma_i} \right)^2}$

Nota: *IS* denota o índice de similaridade, *x* corresponde à característica física em análise, ω representa um parâmetro de ponderação, σ é o desvio padrão das características físicas, *a* e *b* representam as bacias monitoradas e não monitoradas, respectivamente.

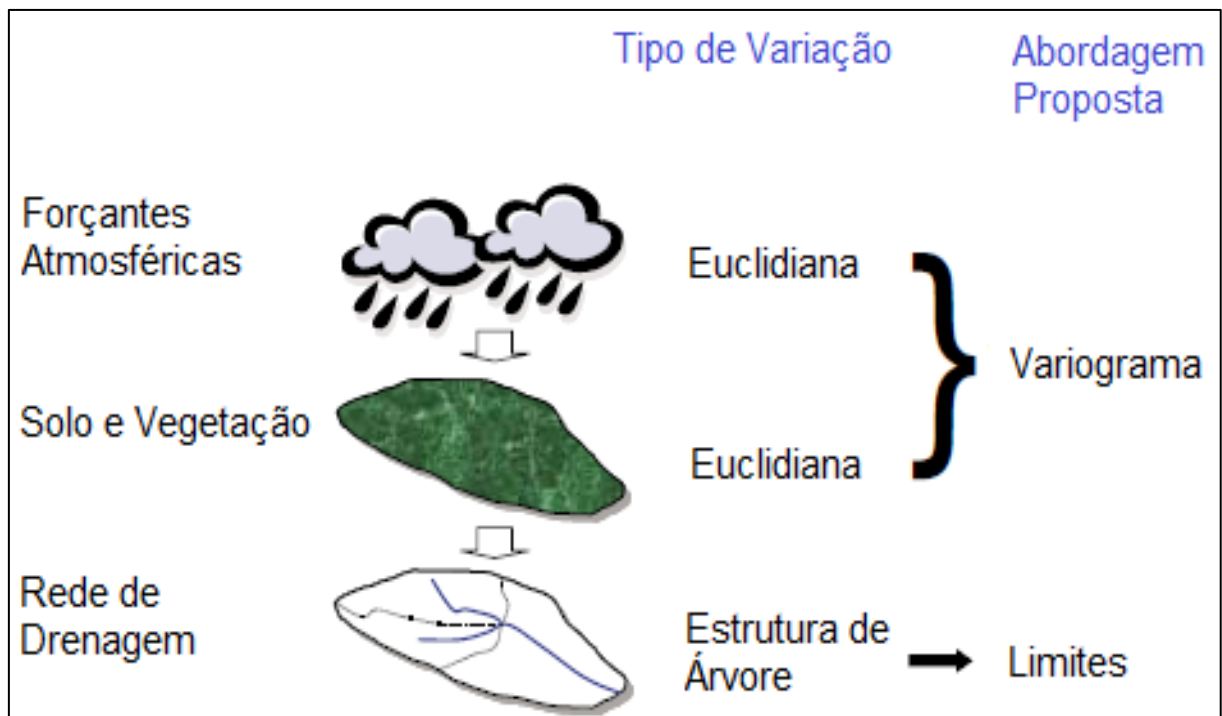
Uma vez definidas as bacias com comportamento semelhante, todo o conjunto de parâmetros do modelo calibrado pode ser transferido da bacia hidrográfica doadora

mais próxima para a bacia de interesse. Outra abordagem consiste em agrupar um conjunto de bacias monitoradas, limitadas por uma distância determinada, e definir os parâmetros do modelo da bacia de interesse com base em uma média ponderada das bacias classificadas.

3.1.2 Top-Kriging

O método de interpolação espacial Top-Kriging, ou Krigagem Topológica, combina dois grupos de variáveis que controlam as vazões em cursos d'água (Figura 2). O primeiro grupo é composto por variáveis que são contínuas no espaço, que estão relacionadas à geração de escoamento local, como a precipitação, a evapotranspiração e as características do solo. Tal processo de geração de escoamento é tido como pontual, isto é, pressupõe-se que a geração de escoamento pode existir em qualquer ponto da bacia hidrográfica. Variáveis dessa natureza podem ser caracterizadas por distâncias Euclidianas e são representadas por variogramas (SKØIEN *et al.*, 2006).

Figura 2 – Grupo de variáveis que controlam o escoamento



Fonte: Adaptado de Skøien *et al.* (2006)

O segundo grupo de variáveis está relacionado à organização da rede de drenagem da bacia e são afetadas pela hierarquia das bacias hidrográficas aninhadas. Podem ser citadas, por exemplo, a vazão média anual, as características dos eventos de cheia e de estiagem, turbidez e temperatura dos cursos d'água (SKØIEN *et al.*, 2006). As variáveis dessa categoria estão definidas apenas para pontos da rede de drenagem analisada e não podem ser representadas pela variação Euclidiana. De outro modo, precisam ser representadas pela variação do tipo de estrutura de árvore.

Na Krigagem Topológica, as medições não são valores pontuais, mas são definidas sobre uma área de influência não nula. Uma variável pontual $z(x)$ pode ser calculada sobre uma área A por meio da Equação 7:

$$z(A_i) = \frac{1}{|A_i|} \int_{A_i} z(x) dx \quad (7)$$

na qual A_i é o suporte espacial de $z(A_i)$. Para variáveis de vazões, A_i é a área de superfície da bacia hidrográfica que drena para um ponto x_i (LAAHA *et al.* 2013).

A predição espacial de uma variável localizada em um ponto x_0 de um curso d'água com área de captação A_0 a partir de amostras $z(A_1), z(A_2), \dots, z(A_n)$, é dada pela Equação 8:

$$\hat{z}(A_0) = \sum_{i=1}^n \lambda_i z(A_i) \quad (8)$$

na qual λ_i são os pesos da interpolação das variáveis observadas das bacias A_i e n é o número de amostras utilizadas no processo de interpolação. Os pesos λ_i podem ser obtidos resolvendo o sistema composto pelas Equações 5 e 6, definidos no Tópico 3.1.1.1.

Assumindo a existência de um variograma pontual γ_p , o valor do variograma ou a semivariância entre duas variáveis observadas de áreas A_i e A_j pode ser obtido pela Equação 9.

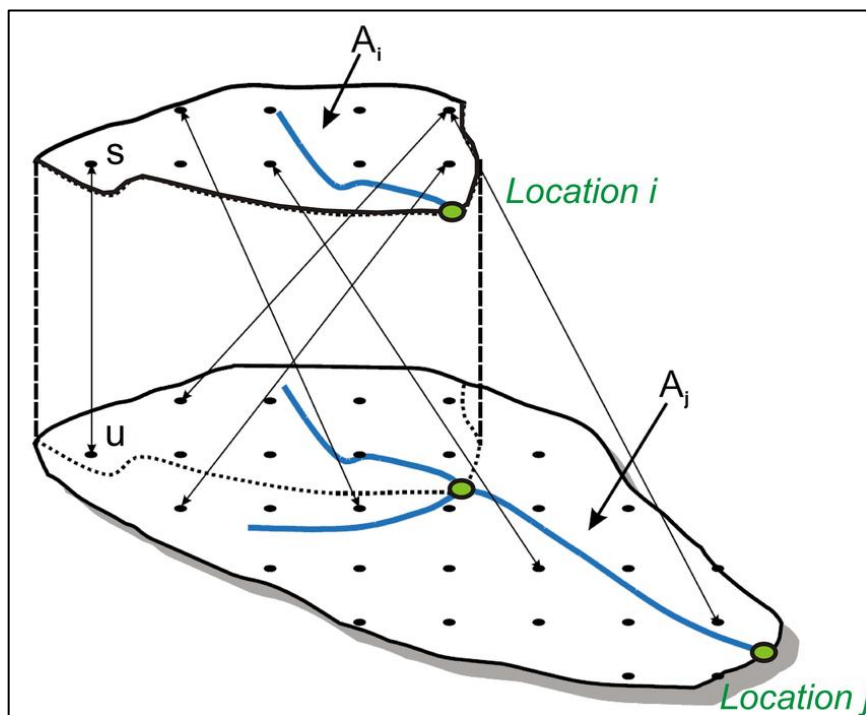
$$\gamma_{ij} = 0,5 \times \text{Var}\left(z(A_i) - z(A_j)\right) = \frac{1}{|A_i||A_j|} \int_{A_j} \int_{A_i} \gamma_p(s - u) ds du$$

$$- 0,5 \times \left[\frac{1}{|A_i|^2} \int_{A_i} \int_{A_i} \gamma_p(s - u) ds du + \frac{1}{|A_j|^2} \int_{A_j} \int_{A_j} \gamma_p(s - u) ds du \right] \quad (9)$$

na qual s e u são vetores de posição dentro de cada área utilizada no processo de integração. A primeira parte da Equação 9 integra a semivariância entre as duas áreas das bacias hidrográficas. A segunda parte da equação, por sua vez, subtrai a semivariância integrada dentro das bacias analisadas. Cabe ressaltar que valor de γ_{ij} aumenta com a distância e diminui com a área de captação das bacias.

A integração descrita pela Equação 9, na maioria dos casos, não possui solução analítica, sendo necessário substituí-la por somas e a área de influência é discretizada por uma grade de pontos. A Figura 3 ilustra essa abordagem para duas bacias hidrográficas aninhadas, discretizadas por uma grade de pontos. As setas indicam os vetores de separação entre os pontos discretizados das áreas A_i e A_j .

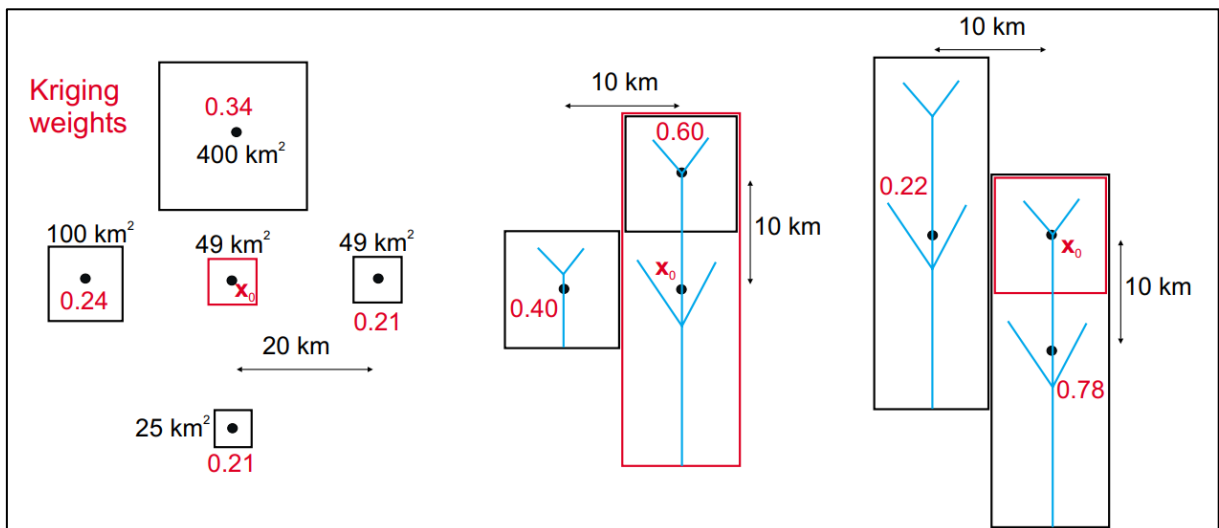
Figura 3 – Grade de pontos de uma rede de drenagem utilizada no método da Krigagem Topológica



Fonte: Laaha *et al.* (2013)

A vantagem da utilização da Krigagem Topológica em relação aos métodos de Krigagem convencionais pode ser visualizada na Figura 4. A imagem à esquerda apresenta a influência da área de drenagem na definição dos valores dos pesos de interpolação, descritos na Equação 8. As quatro bacias utilizadas no processo de regionalização apresentam a mesma distância espacial em relação à bacia não monitorada. Percebe-se que o método atribui maior peso à maior área de drenagem, pois apresenta a medida menos enviesada em relação à média. Outro ponto observado reside no fato de que, embora a bacia mais à direita (e mais próxima à maior bacia monitorada) possua área de drenagem superior à da bacia localizada na parte inferior, ambas receberam o mesmo peso na regionalização. Tal fato ocorre pois a krigagem reduz os pesos das amostras agrupadas (SKØIEN *et al.*, 2006).

Figura 4 – Efeito da área de drenagem (esquerda) e do aninhamento das bacias hidrográficas (centro e direita) na ponderação pela Krigagem Topológica



Fonte: Skøien *et al.* (2006)

O efeito do aninhamento das bacias hidrográficas pode ser visualizado nas imagens central e direita da Figura 4. A distância entre o centroide da bacia de interesse e das bacias com monitoramento é a mesma para ambas as figuras. Na figura central, as bacias doadoras apresentam mesma área de drenagem, entretanto, a bacia que forma uma sub-bacia da área de interesse recebe maior peso na regionalização. O caso oposto é observado na ilustração à direita. A região de interesse compreende uma sub-bacia de um dos locais monitorados. Logo, nesse caso, tal região apresentará maior peso final, ainda que as bacias analisadas apresentem a mesma área de

drenagem e distância em relação ao centroide da região de interesse (SKØIEN *et al.*, 2006).

Castiglioni *et al.* (2011) realizaram um estudo de comparação da estimação da Q_{355} por meio de duas técnicas distintas, o Top-Kriging e o método de Interpolação Espacial Baseado no Espaço Fisiográfico (PSBI - *Physiographical-Space Based Interpolation*). Conforme os autores, as duas abordagens apresentaram desempenhos relativamente próximos, e podem ser entendidos como métodos complementares: a Krigagem Topológica apresentou melhor desempenho de predição nos cursos d'água principais enquanto o PSBI se destacou em bacias de cabeceira.

Laaha *et al.* (2013) utilizaram o Top-Kriging em conjunto com um modelo de regressão para estimar a temperatura média de rios localizados em território Austríaco. A incorporação da Krigagem Topológica no modelo possibilitou a correção de vieses regionais e proporcionou uma redução do RMSE em torno de 20% na estimativa das temperaturas. Ademais, O Top-Kriging aprimorou o ajuste local do modelo, tanto para os rios principais quanto para seus afluentes, utilizando as informações topográficas da rede de drenagem (LAAHA *et al.*, 2013).

Archfield *et al.* (2013) compararam o desempenho do Top-Kriging, PSBI e do método de regressão de Mínimos Quadrados Generalizado (GLS - *Generalized Least Square*) na estimativa de quantis de cheia de tempos de retorno de 10, 50, 100 e 500 anos. De acordo com os autores, o desempenho da Krigagem Topológica foi substancialmente superior às abordagens PSBI e GLS, principalmente para bacias hidrográficas de grandes dimensões. O estudo também mostrou que a união dos métodos Top-Kriging e PSBI apresentou uma melhora marginal das estimativas dos quantis de cheia em relação à utilização exclusiva da Krigagem Topológica.

Pugliese *et al.* (2014) utilizaram a Krigagem Topológica na estimativa de curvas de permanência de longo termo. O Top-Kriging foi utilizado para regionalizar o Desvio Total Negativo (TND - *Total Negative Deviation*), índice que foi utilizado para caracterizar as curvas de permanência adimensionalizadas pela vazão média anual e precipitação média anual. Conforme os autores, a Krigagem Topológica revelou-se uma abordagem confiável para a predição de curvas de permanência de longo termo. O método conseguiu representar de forma otimizada o regime de vazões mínimas

quando em comparação com outros modelos de regionalização empregados na mesma região de estudo.

Almeida *et al.* (2021) aplicaram a abordagem proposta por Pugliese *et al.* (2014) na predição de curvas de permanência de bacias localizadas no estado do Paraná. De acordo com os autores, o método apresentou bom desempenho na predição das curvas de permanência devido aos valores medianos elevados dos coeficientes NSE e LNSE. Os autores destacam ainda que em bacias hidrográficas urbanas, com intervenção humana elevada, o modelo apresentou dificuldade em representar as curvas empíricas.

3.2 CURVA DE PERMANÊNCIA

3.2.1 Definição

A curva de permanência expressa a porcentagem do tempo em que uma determinada vazão foi igualada ou excedida durante um período histórico considerado, representando a relação entre a magnitude e a frequência da referida variável (VOGEL e FENNESSEY, 1994). A curva de permanência também pode ser entendida como o complemento da função acumulada de probabilidade da descarga diária, onde cada valor de vazão Q corresponde a uma probabilidade de excedência p (VOGEL e FENNESSEY, 1994).

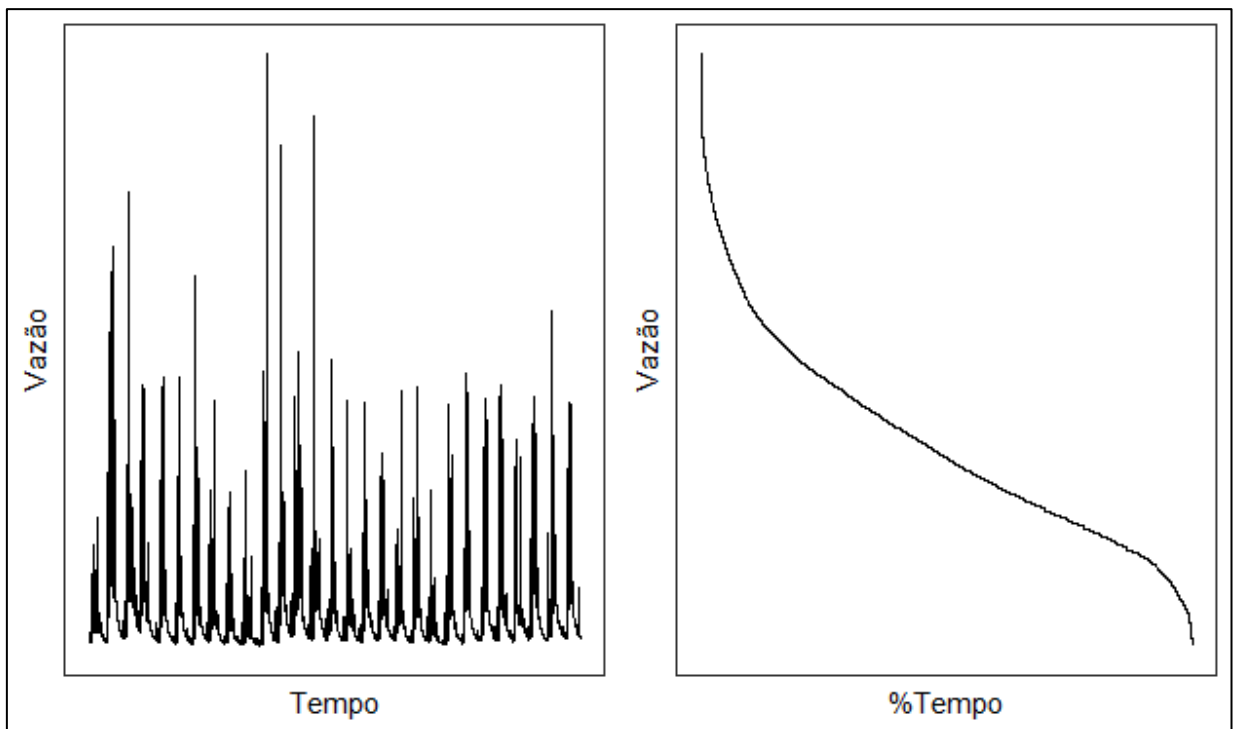
A curva de permanência é uma das ferramentas mais utilizadas na hidrologia. Isso se deve à sua capacidade de condensar uma riqueza de informações sobre a variabilidade do escoamento em uma única imagem (VOGEL E FENNESSY, 1995). A sua forma é única para uma dada bacia e é determinada pela interação das características climáticas, da forma e morfologia da bacia, da cobertura vegetal e das características do solo, que juntos controlam os componentes do escoamento (CASTELLARIN *et al.*, 2013).

A sua utilização está associada a diversas aplicações que compreendem a engenharia de recursos hídricos, as quais incluem: estudos de vazões mínimas e ecológicas (PUMO 2013; SMAKHTIN, 2001), mudanças climáticas (MULLER E THOMPSON, 2016; PUMO *et al.*, 2016), estudos de sedimentos (SINHA *et al.*, 2019), avaliação do

impacto do uso do solo (BROWN *et al.* 2013) e calibração de parâmetros de modelos hidrológicos (COSTA *et al.*, 2014; PINHEIRO E NAGHETTINI, 2013).

A Figura 5 apresenta a forma geral de uma curva de permanência construída a partir de uma série temporal de medições fluviométricas.

Figura 5 - Representação da série de vazões diárias (esquerda) pela sua respectiva curva de permanência (direita)



3.2.2 Histórico de utilização das curvas de permanência

Conforme Foster (1934, *apud* Vogel e Fennessey 1994), a primeira utilização das curvas de permanência é atribuída a Clemens Herschel, entorno de 1880. Vogel e Fennessey (1994) apresentaram um breve histórico de utilização de curvas de permanência, realizados durante a primeira metade do século XX em solo norte-americano. De acordo com os autores, os trabalhos mais destacados sobre construção, análise e interpretação de curvas de permanência foram publicados antes do emprego dos computadores. Dentre os documentos apresentados, podem ser citados os trabalhos de Mitchell (1957), no desenvolvimento de métodos para estimar curvas de permanência em locais monitorados, parcialmente monitorados e sem monitoramento hidrológico; e Cross e Bernhagen (1949) e Saville *et al.* (1933), na

construção das referidas curvas no estado de Ohio e Carolina do Norte, respectivamente.

Searcy (1959) apresentou um plano para construção, análise a aplicação de curvas de permanência de longo termo em projetos de recursos hídricos. Conforme o autor, as curvas de permanência são ferramentas importantes no estudo do regime hidrológico de bacias hidrográficas e na comparação entre bacias distintas. Ademais, o autor destaca a possibilidade da construção das curvas de permanência anuais, método referido a Saville e Watson (1933) e a Barrows (1943).

Yokoo e Sivapalan (2011) analisaram a variabilidade das curvas de permanência mediante combinação de diferentes características físicas e climáticas de bacias hidrográficas. Segundo os autores, as curvas de permanência podem ser decompostas em três partes, cada uma com características e origem distintas: (1) parte superior, governada pelo escoamento superficial durante os eventos de precipitação e um fator de perda por infiltração, (2) parte intermediária, controlada pelo armazenamento da bacia, sendo influenciada pelo clima e geologia do local, (3) parte inferior, definida pelo escoamento de base durante os períodos de seca, somado às perdas por evapotranspiração.

Recentemente, Leong e Yokoo (2021) realizaram uma revisão de estudos associados às curvas de permanência, entre os anos de 2000 e 2020. Destacam-se os estudos que relacionaram as características físicas e climáticas na construção das curvas de permanência. Segundo os autores, são necessários estudos futuros que incluam a ação antrópica na construção das curvas de permanência, uma vez que tais ações estão cada vez mais presentes no regime hidrológico.

3.2.3 Construção das curvas de permanência

As curvas de permanência podem ser construídas de duas formas distintas. A primeira considera as vazões da série histórica completa disponível e são denominadas curvas de permanência de longo termo, denotada pela sigla em inglês FDC (*flow duration curve*). Tais curvas são sensíveis a períodos atípicos de seca e de cheias e a sua interpretação está restrita ao período de dados utilizado na sua construção. Dessa

forma, recomenda-se que o intervalo de tempo empregado na sua montagem seja longo o suficiente para produzir a uma distribuição limite para as vazões.

A segunda abordagem consiste em construir as referidas curvas para cada ano civil ou hidrológico e são denotadas por AFDC (do inglês, *annual flow duration curve*). Essa abordagem permite a construção de uma AFDC mediana, que representa a probabilidade de excedência das vazões médias diárias em um ano civil ou hidrológico típico, não sendo afetado por períodos extremos úmidos ou secos. Ademais, a AFDC mediana permite a construção de intervalos de confiança para os quantis anuais obtidos e atribuição de períodos de retorno específicos associados às curvas anuais.

Embora as AFDCs demonstrem ser menos sensíveis a períodos de extremos hidrológicos, as curvas de permanência de longo termo ainda são copiosamente utilizadas em estudos associados à engenharia de recursos hídricos. Além do mais, diferentemente das curvas de permanência anuais, as FDCs podem ser utilizadas no preenchimento de falhas e na geração de séries sintéticas de vazões em bacias sem monitoramento sistemático, por intermédio de um modelo regional (CASTELLARIN *et al.*, 2004).

Para a construção das curvas de permanência, inicialmente ordena-se as vazões observadas de maneira decrescente. Em seguida, plota-se cada vazão observada versus sua probabilidade de excedência correspondente. Em geral, utiliza-se a posição de plotagem de Weibull para obter tal probabilidade. Conforme Fenessey e Vogel (1990), a posição de plotagem de Weibull é dada pela Equação (10):

$$p_i = \frac{i}{n + 1} \quad (10)$$

na qual i expressa o índice das vazões ordenadas de forma crescente e n corresponde ao número total de dados do período observado.

3.2.4 Regionalização das curvas de permanência

De acordo com Castellarin *et al.* (2004) os métodos de regionalização das curvas de permanência podem ser classificados em três categorias: estatísticos, paramétricos e gráficos.

3.2.4.1 Método Estatístico

O método de regionalização estatístico considera as FDCs como o complemento da função acumulada de probabilidades (FAP). Os procedimentos de regionalização utilizados por essa abordagem consistem nos seguintes passos (CASTELLARIN *et al.*, 2004):

- Definição de uma distribuição de probabilidades adequada para ajustar a curva de permanência de uma região específica;
- Estimação dos parâmetros da distribuição escolhida para todas as bacias hidrográficas monitoradas; e
- Construção de um modelo regional, empregando a regressão múltipla para relacionar os parâmetros da distribuição com variáveis físicas e climáticas dos locais conhecidos.

Vogel e Fennessey (1990) utilizaram a distribuição lognormal de dois parâmetros para construir um modelo regional de FDCs para o estado de Massachusetts, Estados Unidos. Conforme os autores, tal distribuição é adequada para a regionalização da cauda inferior das curvas de permanência e a aplicaram sobre o intervalo $0,5 \leq p \leq 0,9$. A área de drenagem e o desnível máximo da bacia foram utilizados como variáveis explicativas dos parâmetros da distribuição lognormal. Ademais, os autores também calcularam os intervalos de confiança aproximados para as bacias hidrográficas monitoradas, baseando-se na variância dos modelos de regressão para os dois parâmetros da distribuição.

Costa *et al.* (2020) regionalizaram curvas de permanências para 13 rios intermitentes e efêmeros localizados nos estados do Ceará e do Piauí. As distribuições candidatas analisadas foram a Burr XXI estendida, Kappa, Wakeby e a Gama generalizada, sendo que a primeira distribuição foi a escolhida. Os parâmetros do modelo regional foram obtidos utilizando a técnica da Regressão Polinomial Evolucionária (EPR - Evolutionary Polynomial Regression), a qual mostrou-se capaz de relacionar as variáveis explicativas aos parâmetros do modelo, assim com a sua variação regional.

3.2.4.2 Método Paramétrico

Essa abordagem de regionalização consiste em representar as curvas de permanência por meio de expressões analíticas. Os parâmetros do modelo regional são obtidos relacionando-os com as características físicas e/ou climáticas das bacias monitoradas, de forma semelhante aos parâmetros da distribuição de probabilidades descrita no método estatístico. Nesse método, são comumente utilizadas equações dos tipos polinomiais e exponenciais no processo de regionalização.

Quimpo *et al.* (1983) utilizaram essa abordagem para estimar FDCs em bacias não monitoradas nas Filipinas. Os autores utilizaram uma equação exponencial de dois parâmetros dada pela Equação 11:

$$Q_p = Q_A \exp(cp) \quad (11)$$

na qual Q_p indica a vazão de um dia específico associado a uma permanência p e Q_A e c são os parâmetros do modelo. Para a construção do modelo regional, Q_A foi estimada com base na área de drenagem e c foi obtido por meio de um mapa de contornos desenvolvido para todo o arquipélago das Filipinas.

Mimikou e Kaemaki (1985) aplicaram a regionalização paramétrica na construção de curvas de permanência para a região nordeste da Grécia. Os autores utilizaram a seguinte equação polinomial de terceiro grau na concepção do modelo regional:

$$Q_p = a - bp + cp^2 - dp^3 \quad (12)$$

na qual os parâmetros da equação a, b, c e d , são não negativos. As características físicas e climáticas das bacias monitoradas empregadas na construção do modelo foram a área de drenagem, precipitação média anual, declividade equivalente da bacia e o comprimento do rio principal.

Franchini e Suppo (1996) empregaram a abordagem paramétrica na regionalização do ramo inferior da curva de permanência ($p \geq 0,3$) para bacias localizadas na região centro-sul da Itália. Os autores empregaram a Equação 13 na concepção do modelo regional.

$$Q_p = c + a(1 - p)^b \quad (13)$$

na qual os parâmetros a, b e c são positivos. O parâmetro b está associado às características físicas das bacias hidrográficas monitoradas e controla a concavidade das curvas de permanência.

3.2.4.3 Método Gráfico

A abordagem gráfica baseia-se em dois princípios básicos. O primeiro consiste em obter as curvas de permanência adimensionais por meio da divisão das curvas citadas por um índice de vazão específico, como, por exemplo, a vazão média de longo termo. O segundo passo compreende a obtenção de uma curva regional adimensional por meio da média das curvas empíricas adimensionalizadas das bacias monitoradas. De acordo com Shu e Ouarda (2012), a principal diferença entre a abordagem gráfica das outras duas citadas (estatística e paramétrica) reside no fato de que nenhuma estimativa de distribuição de probabilidades ou do formato das FDCs precisa ser realizada.

Smakhtin *et al.* (1997) utilizaram o método gráfico na regionalização de FDCs de bacias localizadas na região centro-oeste da África do Sul. Os autores assumiram como homogênea toda a região de estudo, baseados nas semelhanças intensas entre as FDCs adimensionais. Entretanto, Castellarin *et al.* (2004) sugerem a aplicação da abordagem em bacias hidrográficas heterogêneas, contando que as bacias hidrográficas sejam avaliadas em grupos com características e regimes hidrológicos semelhantes.

Requena *et al.* (2018b) sugeriram a utilização o conceito de Análise Funcional de Dados na construção completa das FDCs, não sendo necessário o emprego de equações analíticas específicas que limitam o formato das curvas de permanências. Os autores construíram as FDCs funcionais para 109 estações fluviométricas localizadas na província de Quebec, Canadá. O modelo regional foi obtido por intermédio da Regressão Funcional Múltipla, que possibilita associar uma variável dependente, em termos funcionais, às características físicas e climáticas das bacias de interesse. O presente trabalho aplicou os conceitos da Análise Funcional de Dados,

assim como da Regressão Funcional Múltipla na concepção da regionalização hidrológica. Os seus principais fundamentos serão descritos nos tópicos seguintes.

3.3 ANÁLISE FUNCIONAL DE DADOS

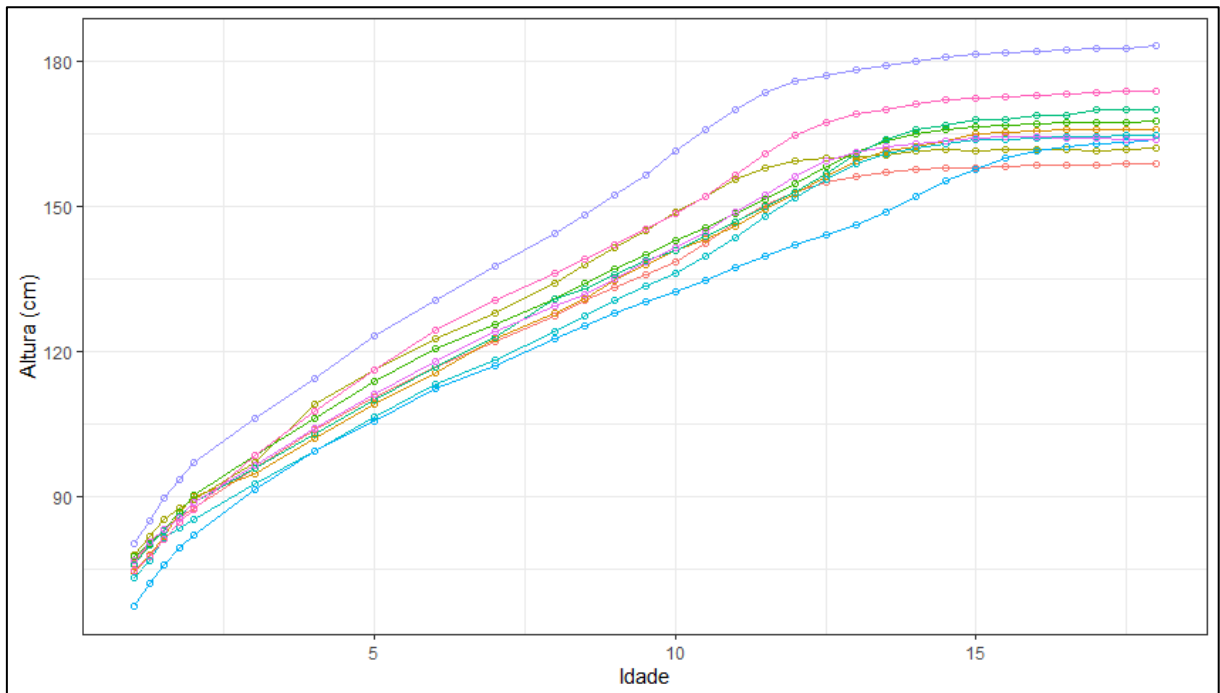
3.3.1 Definição

Dados funcionais estão se tornando cada vez mais recorrentes nos mais diversos campos da sociedade, o que torna necessário o desenvolvimento de ferramentas estatísticas que possibilitem analisar e interpretar esse tipo de informação (CHEBANA *et al.*, 2012). O conceito de análise funcional de dados foi inicialmente proposto por Ramsay, em 1982 e, posteriormente, difundido pelo livro “Functional Data Analysis” escrito por Ramsay e Silverman em 2005 (MASSELOT *et al.*, 2016). Conforme Muller (2011), essa abordagem refere-se à análise estatística de dados que abrange funções ou superfícies aleatórias, nas quais cada função é interpretada como um único elemento.

Essencialmente, a análise funcional de dados, denominada FDA (do inglês, *Functional Data Analysis*), tem como princípio expressar os dados de uma série de medições como funções, diferentemente das abordagens tradicionais que os tratam como escalares. Dessa forma, qualquer conjunto de amostras medidas, no tempo ou espaço, de uma determinada variável que apresente comportamento suave, pode ser analisado do ponto de vista da FDA. Por exemplo, a Figura 6 demonstra um estudo de medição da altura de dez pessoas tomadas em um conjunto de 31 idades durante 18 anos.

Observa-se que a discretização temporal das medidas de altura não é uniforme, ou seja, a estatura foi aferida em intervalos desigualmente espaçados, consistindo em medidas trimestrais até um ano de idade, anuais entre dois e oito anos de idade e semestrais de nove anos em diante. Embora cada registro seja definido como um valor discreto, percebe-se uma variação suave da altura, evidenciada conforme a frequência de medidas aumenta. Logo, o conjunto de medidas de estaturas aferidas pode, então, ser analisado e interpretado como um dado funcional.

Figura 6 - Aplicação da análise funcional de dados no estudo de crescimento de 10 pessoas. Cada curva está associada à um indivíduo e os pontos retratam as medidas realizadas durante 18 anos



Fonte: Ramsay e Silverman (2005)

De acordo com Ramsay e Dalzell (1991), existem quatro razões práticas para se analisar os dados por um aspecto funcional. São elas:

- A frequência das observações funcionais em contextos aplicados está aumentando, à medida que a coleta automatizada de dados se torna mais acessível aos pesquisadores. Ademais, procedimentos de suavização e interpolação podem produzir representações funcionais de conjuntos finitos de observações;
- Alguns problemas de modelagem são mais realistas quando pensados em termos funcionais, ainda que apenas algumas observações estejam disponíveis;
- Os objetivos de uma análise podem ser de natureza funcional, como estimar uma função completa, suas derivadas ou valores de características funcionais dispondo-se apenas de dados finitos; e
- Levar em conta certas características, como a suavidade para dados multivariados decorrentes de processos funcionais, pode ter implicações importantes nos resultados das análises.

Os objetivos da análise funcional de dados são essencialmente os mesmos de qualquer outro ramo da estatística. Alguns deles incluem (Ramsay e Silverman, 2005):

- Representar os dados de maneira a facilitar análises futuras;
- Exibir os dados de forma que várias características relevantes sejam destacadas;
- Estudar padrões e variações entre os dados em análise;
- Explicar a variação em um resultado ou variável dependente utilizando a informação de entrada ou da variável independente; e
- Comparar dois ou mais conjuntos de dados com respeito a certos tipos de variações em que dois conjuntos de dados podem conter diferentes conjuntos de replicações das mesmas funções ou diferentes funções para um conjunto comum de replicações.

3.3.2 Representação dos dados funcionais

A primeira etapa da análise funcional de dados consiste em representar os dados originais em termos funcionais. Embora a função alvo deva apresentar comportamento suave, os dados originais $y = (y_1, \dots, y_n)$ podem apresentar erros observacionais ou ruídos oriundos do processo de medição. Dessa forma, pode-se escrever:

$$y_j = x(t_j) + \varepsilon_j \quad (14)$$

na qual $x(t_j)$ é uma função suave e ε_j pode ser definido como ruído ou erro de medição e contribui para a não suavidade dos dados. Um procedimento comum de suavização consiste em representar os dados medidos como um somatório de funções de base. Um sistema de funções de base consiste em um conjunto de funções conhecidas ϕ_k , matematicamente independentes entre si, e que possuem a propriedade de representar qualquer função por meio de uma combinação linear de um número suficientemente grande dessas funções (Ramsay e Silverman, 2005). O processo é representado por uma expansão linear, conforme a Equação 15:

$$x(t) = \sum_{k=1}^K c_k \phi_k \quad (15)$$

sendo K o número de funções de base utilizadas na construção do modelo e c_k os coeficientes escalares das respectivas funções. A quantidade de bases K determina o grau de suavidade da curva obtida (Ramsay e Silverman, 2005).

O sistema de funções de base escolhido deve apresentar características que estejam relacionadas com as funções a serem estimadas. Isso facilita a obtenção de uma aproximação da função alvo, reduzindo o número de funções de base K . Existem diversas classes de funções de base passíveis de escolha em função do tipo de dados observado. No contexto da FDA destacam-se as bases de Fourier e do tipo B-Splines.

3.3.2.1 Bases de Fourier

As funções de base de Fourier, uma das bases mais conhecidas e utilizadas em análises estatísticas, são definidas pelas Equações 16, 17 e 18:

$$\phi_0(t) = 1 \quad (16)$$

$$\phi_{2r-1}(t) = \sin r\omega t \quad (17)$$

$$\phi_{2r}(t) = \cos r\omega t \quad (18)$$

Para $r \geq 1$. Essa base é do tipo periódica e o parâmetro ω determina o período $T = 2\pi/\omega$ (Ramsay e Silverman, 2005). A expansão da série de Fourier é dada pela Equação 19.

$$x(t) = c_0 + c_1 \sin \omega t + c_2 \cos \omega t + c_3 \sin 2\omega t + c_4 \cos 2\omega t + \dots \quad (19)$$

Na qual c_0 é o intercepto da função e $c_1, \dots, c_n$ são os coeficientes das bases de Fourier. Se os valores de t_j são igualmente espaçados em um intervalo definido T e o período é igual ao comprimento desse intervalo, então, as bases são ditas ortogonais, ou seja, o produto entre as matrizes de bases $\Phi' \Phi$ é uma matriz diagonal.

As séries de Fourier são aplicadas fundamentalmente em funções consideradas estáveis, isto é, funções que não apresentam características locais marcantes e que a sua curvatura tende a ser da mesma ordem em todos os lugares. Elas geralmente produzem expansões que são suaves uniformemente. Entretanto, são inadequadas

para dados que apresentem descontinuidade na função ou em derivadas de menor ordem (Ramsay e Silverman, 2005).

3.3.2.2 Bases B-Splines

Diferentemente das funções de base Fourier, as funções Splines são utilizadas na aproximação de dados com comportamento não periódico. Essas funções combinam a velocidade do cálculo computacional de funções polinomiais com a flexibilidade substancialmente maior de adaptação, frequentemente atingida com uma quantidade menor de bases (Ramsay e Silverman, 2005).

O primeiro passo da construção de uma Spline é dividir o intervalo onde a função está definida em L subintervalos. Os pontos internos são denominados nós (τ_l) e são definidos conforme a quantidade de subintervalos especificados, $l = 1, \dots, L - 1$. Em cada intervalo, a função Spline é definida por um polinômio de ordem m . A ordem de um polinômio é dada pelo número de constantes necessárias para defini-lo, sendo uma unidade a mais que o seu grau, que é o valor da maior potência contida no polinômio.

As Splines também garantem que a transição entre polinômios adjacentes seja suave nos nós que os limitam, isso é, os valores dos polinômios, bem como as suas derivadas de ordem até $m - 2$, devem coincidir no ponto analisado.

A construção do sistema de bases Splines é definida por funções $\phi_k(t)$ que satisfazem as seguintes propriedades:

- Cada função $\phi_k(t)$ é uma spline definida por uma ordem m e uma sequência de nós τ ;
- Qualquer combinação linear dessas funções de base também é uma spline; e
- Qualquer spline definida por m e τ pode ser descrita como uma combinação linear dessas funções de base.

O modelo mais popular que cumpre os requisitos citados é o sistema de bases B-Splines, desenvolvido por De Boor (1978). Uma Spline de ordem m é positiva apenas em m intervalos adjacentes, conferindo um suporte compacto e garantindo uma eficiência de cálculo.

O número de parâmetros necessários para definir uma Spline, onde cada junção corresponde a um nó, é igual à ordem da função acrescida do número de nós internos $L - 1$. Dessa forma, considerando um sistema B-Splines de ordem m , pode-se reescrever a Equação 15 como:

$$x(t) = \sum_{k=1}^{m+L-1} c_k B_k(t, \tau) \quad (20)$$

3.3.3 Suavização dos dados

O objetivo da suavização dos dados consiste em representar um conjunto de observações discretas $y = y_1, \dots, y_n$ por um modelo $y_j = x(t_j) + \varepsilon_j$. A função $x(t)$ é descrito por uma expansão linear de funções de base dada pela Equação 21.

$$x(t) = \sum_{k=1}^K c_k \phi_k = \mathbf{c}' \boldsymbol{\phi} \quad (21)$$

sendo $\mathbf{c}'$ o vetor de coeficientes c_k , $\boldsymbol{\phi}$ é o vetor de bases e K é a quantidade de coeficientes e bases compreendidos nos respectivos vetores. Dessa forma, é possível aproximar funções de base conhecidas aos dados originais minimizando a soma do quadrado dos erros, conforme Equação 22:

$$SQE = \sum_{j=1}^n [y_j - \sum_{k=1}^K c_k \phi_k(t_j)]^2 \quad (22)$$

Em termos matriciais, tem-se:

$$SQE = (\mathbf{y} - \boldsymbol{\Phi} \mathbf{c})' (\mathbf{y} - \boldsymbol{\Phi} \mathbf{c}) \quad (23)$$

Entretanto, a aproximação pelo método da soma do quadrado dos erros pode não fornecer uma boa aproximação dos dados originais, uma vez que os sistemas de base utilizados precisam ter as mesmas características essenciais dos processos que geram os dados medidos. Por esse motivo, na análise de dados funcionais é comum que se utilize o método de penalização da não suavidade para aproximar os dados discretos por uma função. Tal método mantém as vantagens da suavização por mínimos quadrados, mas evita algumas das suas limitações. Ademais, essa técnica também se adapta a uma diversidade de problemas que compreendem a análise funcional de dados.

Uma maneira de quantificar a noção de não suavidade de uma curva é utilizar o quadrado da segunda derivada da função x em t , $[D^2x(t)]^2$, também conhecida como curvatura da função em t . Logo, uma medida natural da não suavidade de uma função x é a integral do quadrado da sua segunda derivada, conforme Equação 24.

$$PEN_2(x) = \int [D^2x(t)]^2 dt \quad (24)$$

Funções que apresentam alta variabilidade produzirão elevados valores de $PEN_2(x)$, pois as suas segundas derivadas serão altas em pelo menos parte do intervalo de interesse. Amparado pelo conceito de penalidade, é possível reescrever o critério de mínimos quadrados, definido na Equação 22, adicionando um termo de penalização da não suavidade da função. Sendo assim, a soma dos quadrados dos erros penalizada é expressa pela Equação 25.

$$SQEPEN_\lambda = \sum_{j=1}^T (y_j - x_i(t_j))^2 + \lambda \int [D^2x_i(t)]^2 dt \quad (25)$$

onde $y_j, j = 1, \dots, T$ são os dados observados utilizados para ajustar a função suavizada $x_i(t_j)$. De acordo com Ramsay e Silverman (2005), o elemento λ é um parâmetro de suavização que mede o equilíbrio entre os dados e a função ajustada, medida pela soma dos quadrados dos erros no primeiro termo, e a variabilidade da função x , quantificada por $\int [D^2x(t)]^2 dt$ no segundo termo.

Desse modo, conforme λ aumenta, a penalização da não suavidade de funções não lineares também aumenta por intermédio do termo $\int [D^2x(t)]^2 dt$ e, por conseguinte, a soma do quadrado dos erros penalizada dará maior enfoque na suavidade de x e menor relevância para o ajuste aos dados. Sendo assim, à medida que $\lambda \rightarrow \infty$, a função ajustada x deverá se aproximar da regressão linear padrão dos dados observados.

Em contrapartida, conforme λ diminui, a curva tende a se tornar mais variável, visto que haverá menos penalização para a sua não suavidade. Portanto, ao passo que $\lambda \rightarrow 0$, a curva ajustada x se aproximará da interpolação dos dados, satisfazendo $x_i(t_j) = y_j$ para todo j . Entretanto, inclusive nesse caso limite, a curva interpolada não

é arbitrariamente variável. Em vez disso, tal curva será a função mais suave, duas vezes diferenciável, que se ajusta adequadamente aos dados.

3.3.4 Histórico de utilização da FDA no contexto hidrológico

A utilização da FDA no contexto hidrológico é recente e vem aumentando ao longo dos anos. Por exemplo, Chebana *et al.* (2012) utilizaram o conceito de análise de frequência no estudo de vazões máximas empregando a FDA na construção de hidrogramas. De acordo com os autores, a abordagem utilizando dados funcionais forneceu uma ferramenta eficaz e eficiente na estimativa do risco associados a eventos extremos de cheia. Os autores recomendam utilizar a análise funcional de dados em conjunto com a análise multivariada, com o objetivo de obter uma melhor compreensão dos eventos de cheia.

Masselot *et. al* (2016) realizaram a previsão de vazões com base em dados de precipitações de uma estação localizada na província de Quebec, Canadá. A FDA foi utilizada na construção dos hidrogramas mediante Modelos Lineares Funcionais. Os resultados obtidos foram confrontados com a abordagem de Redes Neurais Artificiais. O modelo construído pela FDA foi capaz de prever a forma geral do hidrograma analisado, enquanto o modelo de Redes Neurais apresentou melhor comportamento em padrões de curtos prazo, como eventos de cheia. De acordo com os autores, o modelo baseado em FDA necessitaria de uma quantidade considerável de funções de base para atingir os picos registrados no hidrograma medido.

Ternynck *et al.* (2016) utilizaram o conceito da análise funcional de dados na classificação de hidrogramas realizada por meio do algoritmo K-means. Para tanto, os autores utilizaram a FDA na representação da série temporal de vazões para um conjunto de estações localizadas na província de Quebec, Canadá. Conforme os autores, a abordagem funcional utilizada pode fornecer estimativas melhores do que as obtidas pelo método de classificação multidimensional. Ademais, o método funcional possui a vantagem de ser automático, ou seja, não necessita da predefinição de índices hidrológicos.

Requena *et al.* (2018) propuseram a regionalização da curva de permanência por meio da Regressão Funcional Múltipla - essa técnica será descrita com mais detalhes no

Tópico 3.3.5. Tal abordagem permite estimar todos os quantis das FDCs por meio de uma única regressão. De acordo com os autores, o uso de tal método possui a vantagem de fornecer informações sobre a influência de cada variável explicativa sobre os percentis da curva regionalizada. Os autores também mencionam que essa técnica pode ser recomendada quando se deseja estimar um grande número de quantis.

Mazelan e Suhaila (2023) utilizaram a Análise Funcional de Dados para converter dados pontuais de precipitações em curvas funcionais. Conforme os autores, as curvas funcionais obtidas podem ser analisadas para qualquer intervalo de tempo. Ademais, informações como a variabilidade do período chuvoso podem ser analisadas mediante a derivada da função de base utilizada no processo de suavização. Os autores também ressaltaram que, em algumas regiões, a representação da variabilidade do padrão das precipitações necessita de mais funções de base do que em outras.

3.3.5 Regressão funcional múltipla

A regressão linear múltipla é um dos mecanismos mais importantes no contexto de regionalização hidrológica. Essa ferramenta permite correlacionar características físicas, climáticas e morfológicas de uma bacia a uma variável dependente de interesse, como, por exemplo, parâmetros de distribuições de probabilidade, quantis de curvas de permanências e de curvas de frequência de vazões.

Caso a variável independente Y varie linearmente com as variáveis dependentes x_n , pode-se definir um modelo geral conforme a Equação 26 (NAGHETTINI e PINTO, 2007).

$$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n + \varepsilon \quad (26)$$

sendo Y a variável dependente, $x_1, \dots, x_n$ as variáveis independentes, $\beta_1, \dots, \beta_n$ os coeficientes do modelo e ε o resíduo da regressão.

Em alguns casos, pode-se realizar a transformação das variáveis com o intuito de linearizar relações não lineares. Tal transformação é realizável por anamorfose

logarítmica. Um dos modelos linearizáveis é o do tipo multiplicativo, dado pela Equação 27.

$$Y = \beta_0 X_1^{\beta_1} X_2^{\beta_2} \varepsilon \quad (27)$$

Realizando a transformação logarítmica, temos:

$$\ln Y = \ln \beta_0 + \beta_1 \ln X_1 + \beta_2 \ln X_2 + \ln \varepsilon \quad (28)$$

Na estimativa de curvas de permanência é usual descrever a Equação 28 da seguinte forma:

$$\ln Q_i^p = \ln \beta_0 + \beta_1 \ln X_{i1} + \beta_2 \ln X_{i2} + \dots + \beta_k \ln X_{ik} + \varepsilon_i \quad (29)$$

na qual Q_i^p é um quantil da FDC associado a uma permanência p de uma estação $i = 1, \dots, n$; $X_{i1}, \dots, X_{ik}$ são os descritores da bacia (características físicas, climáticas e morfológicas); $\beta_1, \dots, \beta_k$ são os coeficientes do modelo e ε_i é o resíduo da regressão.

O conceito de análise funcional de dados pode ser estendido à regressão linear múltipla e, a partir disso, obter uma regressão de formato funcional. Considerando as equações descritas no Tópico 3.3.2, pode-se reescrever a Equação 29 em termos funcionais. Dessa forma, a Regressão Funcional Múltipla (FMR – *Functional Multiple Regression*) pode ser definida pela Equação 30 (REQUENA *et al.*, 2018b):

$$\ln Q_i(p) = \ln \beta_0(p) + \beta_1(p) \ln X_{i1} + \beta_2(p) \ln X_{i2} + \dots + \beta_k(p) \ln X_{ik} + \varepsilon_i(p) \quad (30)$$

na qual $Q_i(p)$ são os quantis da FDC associados às permanências p de uma estação $i = 1, \dots, n$; $X_{i1}, \dots, X_{ik}$ são os descritores da bacia (características físicas, climáticas e morfológicas); $\beta_1(p), \dots, \beta_k(p)$ são os coeficientes funcionais do modelo e $\varepsilon_i(p)$ é o termo residual em função de p .

Note que, diferentemente da Equação 29, que os descritores $\beta_1, \dots, \beta_k$ da FMR são funções definidas no intervalo de p , em oposição à Regressão Linear Múltipla, onde os coeficientes são escalares. Essa característica permite que seja possível avaliar a influência de cada variável dependente na construção do modelo de regressão.

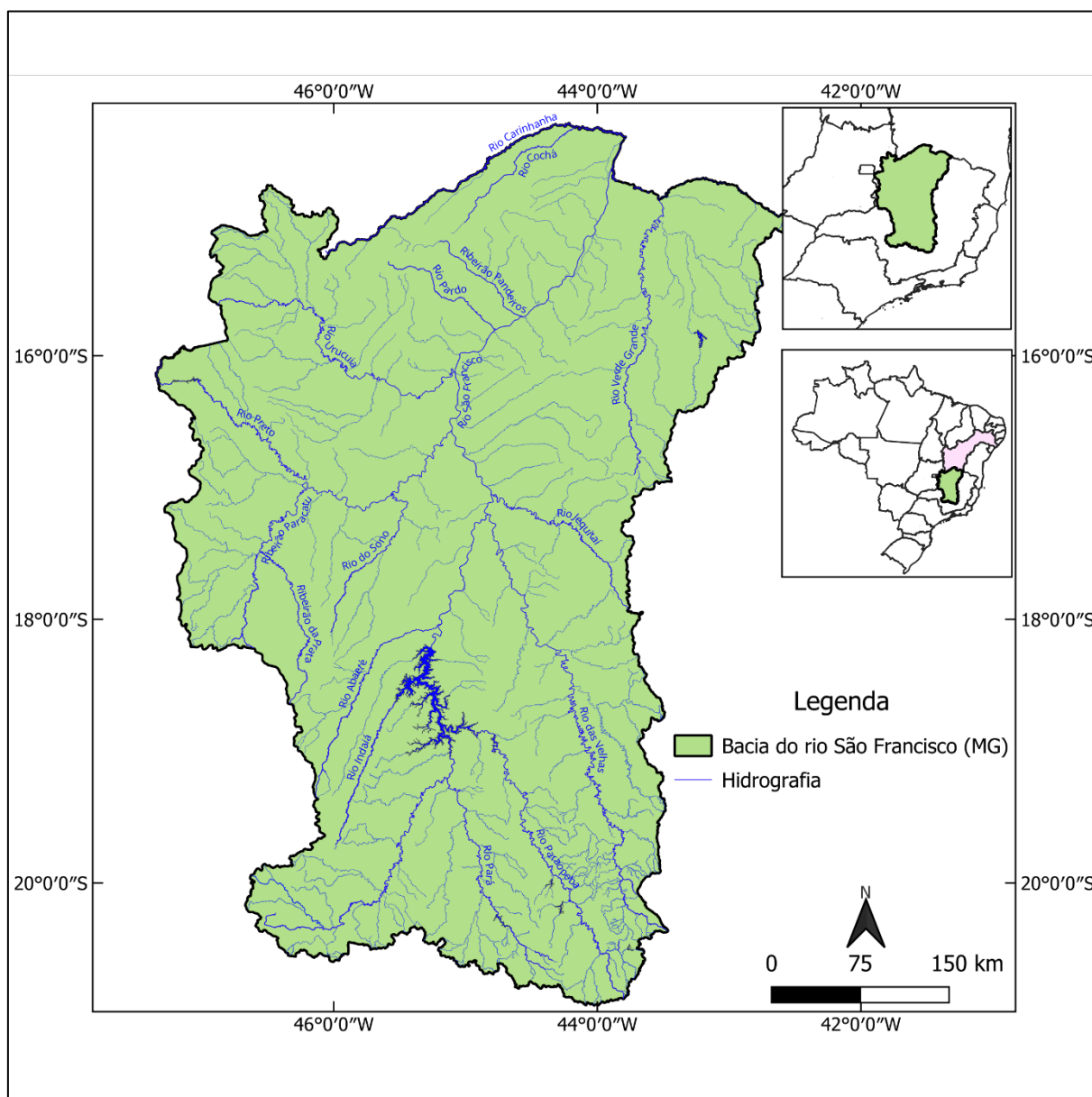
4 METODOLOGIA / MATERIAL E MÉTODOS

4.1 ÁREA DE ESTUDO

A área escolhida para o desenvolvimento da metodologia foi a região da bacia hidrográfica do Rio São Francisco contida no estado de Minas Gerais (Figura 7). Essa escolha foi condicionada à disponibilidade de dados fluviométricos que permitissem o desenvolvimento da regionalização hidrológica das curvas de permanência de longo termo e à importância socioeconômica da bacia em termos estaduais e nacionais. Outro ponto importante para essa escolha foi a heterogeneidade apresentada, tanto física quanto climática, pelas sub-bacias que compõem a região em questão, uma vez que se pretendeu averiguar a capacidade de predição do modelo em uma região com características diversas.

A região do estudo em questão está inserida na região noroeste do estado de Minas Gerais e possui uma área de drenagem de aproximadamente 234000 km². Essa região engloba importantes cidades mineiras, como Belo Horizonte, Betim, Contagem, Divinópolis, Montes Claros, Ribeirão das Neves, Santa Luzia e Sete Lagoas. A população residente estimada para a região é de aproximadamente 8 milhões de habitantes distribuída em 246 municípios, onde 30% desse percentual está contido na capital do estado, Belo Horizonte (CPRM, 2018).

Figura 7 - Bacia hidrográfica do Rio São Francisco no estado de Minas Gerais



O Rio São Francisco nasce na Serra da Canastra, na região Oeste do estado mineiro, e recebe contribuições de importantes cursos de água ao longo do seu percurso, como os rios das Velhas, Abaeté, Paraopeba, Paracatu, Urucua, Jequitai, Verde Grande e Caririnha (CPRM, 2018). Ademais, o curso do Rio São Francisco ainda conta com a usina Hidrelétrica de Três Marias, uma das mais importantes do estado.

4.1.1 Clima

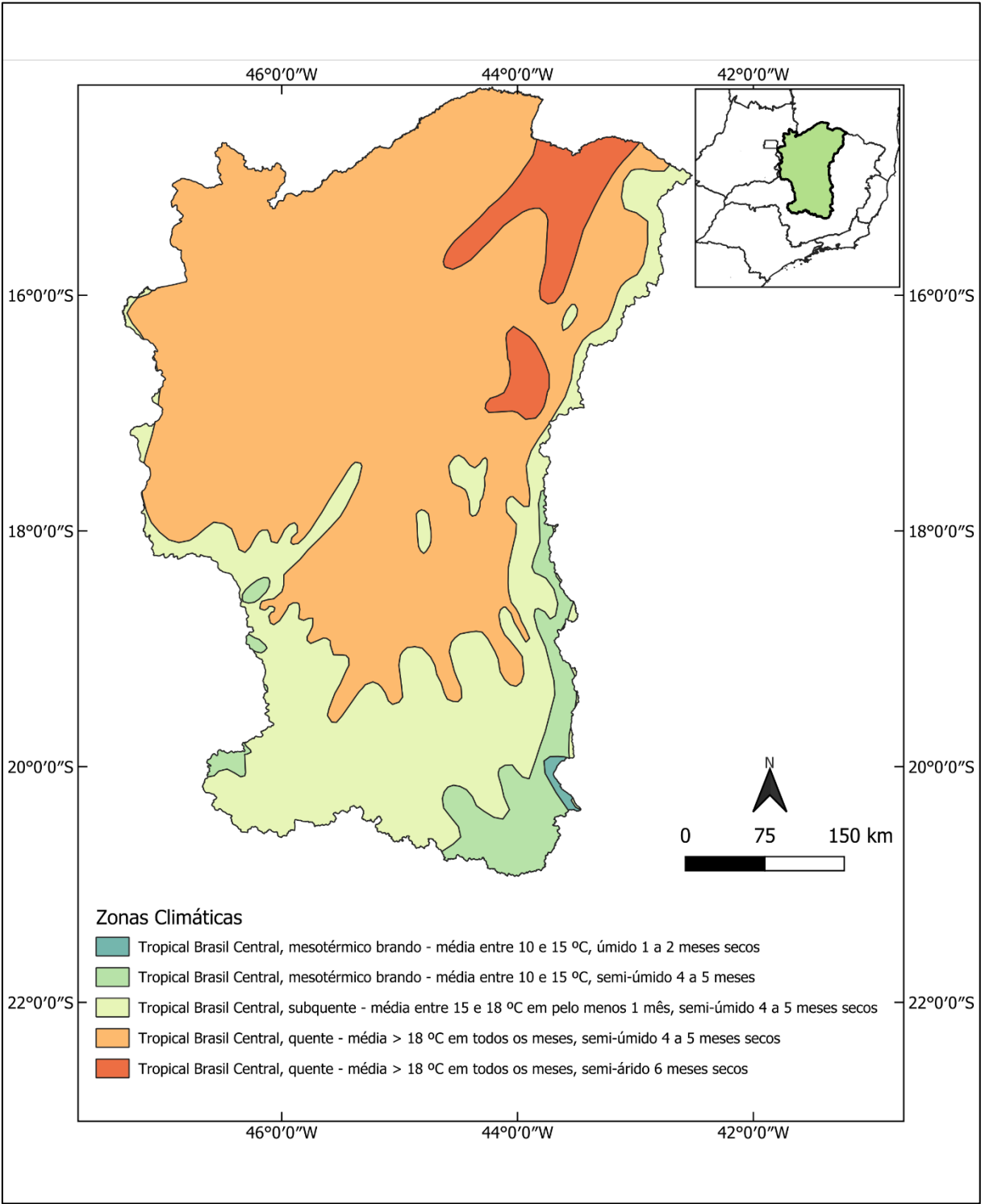
A região de estudo compreende 5 zonas climáticas tropicais, conforme pode ser visualizado na Figura 8. Grande parte da bacia hidrográfica do Rio São Francisco, em

Minas Gerais, apresenta o clima Brasil Central quente, semiúmido, possuindo cerca de 4 a 5 meses secos ao ano. Essa zona climática costuma apresentar temperaturas médias acima de 18°. O segundo clima mais presente na região de estudo é o subquente semiúmido, com médias de temperaturas variando entre 15 e 18 °C. Essa zona contempla a Serra da Canastra, região onde nasce o Rio São Francisco. Também são observados, em menor escala, os climas mesotérmicos brandos, úmido, a sudeste, e semiúmido, tanto a sudeste quanto a sudoeste da área de estudo. Ademais, cabe ressaltar a presença de uma parcela semiárida, de clima quente, localizada à nordeste da região, próxima à divisa com o estado da Bahia.

Com base em dados interpolados das Normais Climatológicas do Brasil de 1981-2010 (INMET), as temperaturas máximas médias variam entre 28 °C, registradas na parte sul da bacia hidrográfica, e 32 °C, onde está inserida a região semiárida da região em questão. Assim como as máximas, as temperaturas médias mínimas são maiores para a região semiárida, com médias próximas aos 20 °C, e menores para algumas localidades nas partes sul e sudeste da região de interesse, apresentando médias em torno de 15 °C.

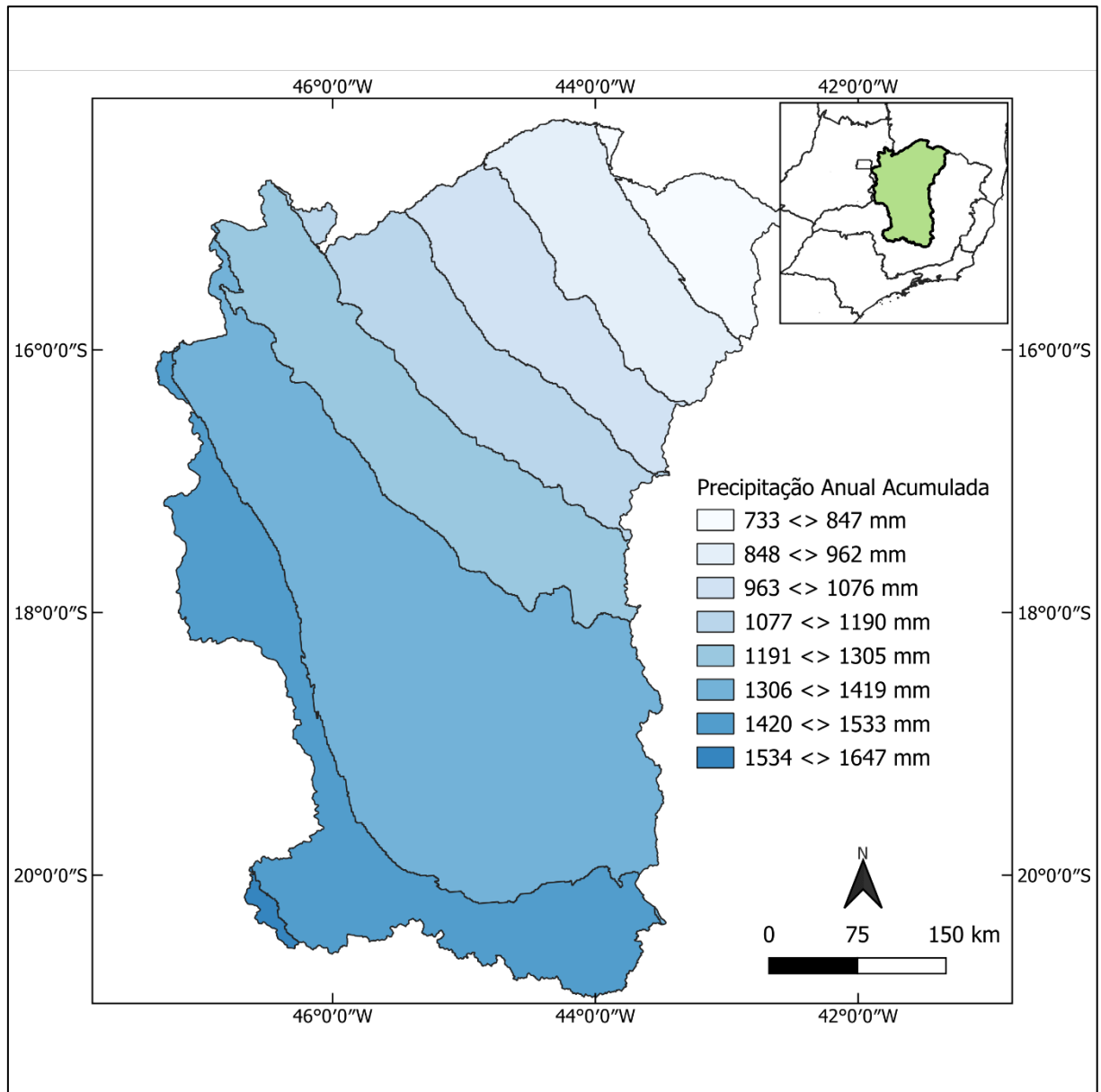
Em relação aos totais anuais de precipitação, baseando-se nos dados obtidos pelas Normais Climatológicas do Brasil de 1961-1990 (SISEMA, 2018), os volumes acumulados anualmente apresentam grande variabilidade dentro da área de estudo (Figura 9). Os menores valores acumulados são registrados na parcela da região onde está inserido o clima semiárido, com precipitação anual acumulada variando entre 733 e 847 mm. Já na parte sul, região onde se encontra a nascente do Rio São Francisco, são registradas precipitações anuais acumuladas próximas aos 1700 mm.

Figura 8 – Zonas climáticas da bacia hidrográfica do Rio São Francisco no estado de Minas Gerais



Fonte: SISEMA (2018a)

Figura 9 – Precipitação anual acumulada da bacia hidrográfica do Rio São Francisco no estado de Minas Gerais



Fonte: SISEMA (2018b)

4.1.2 Uso e Cobertura do Solo e Relevo

Devido ao grande território coberto pela bacia hidrográfica do Rio São Francisco, vários tipos de usos e cobertura são observados na região de interesse, conforme a Figura 10. De acordo com as estimativas realizadas pelo Projeto MapBiomias (2022), a agropecuária consiste na principal classe de uso e cobertura dos solos, presente em mais de 70% do território da bacia em estudo (Tabela 2). Desse montante, cerca de 27% são destinados à silvicultura e 34% ao cultivo de cana de açúcar. Áreas não vegetadas correspondem ao segundo uso de solo mais presente na região em questão, correspondendo a aproximadamente 16% da área total, onde as áreas urbanizadas representam mais de 14% desse total. Além disso, também são verificadas florestas, formação natural não florestal e corpos d'água como usos e cobertura dos solos da área de estudo.

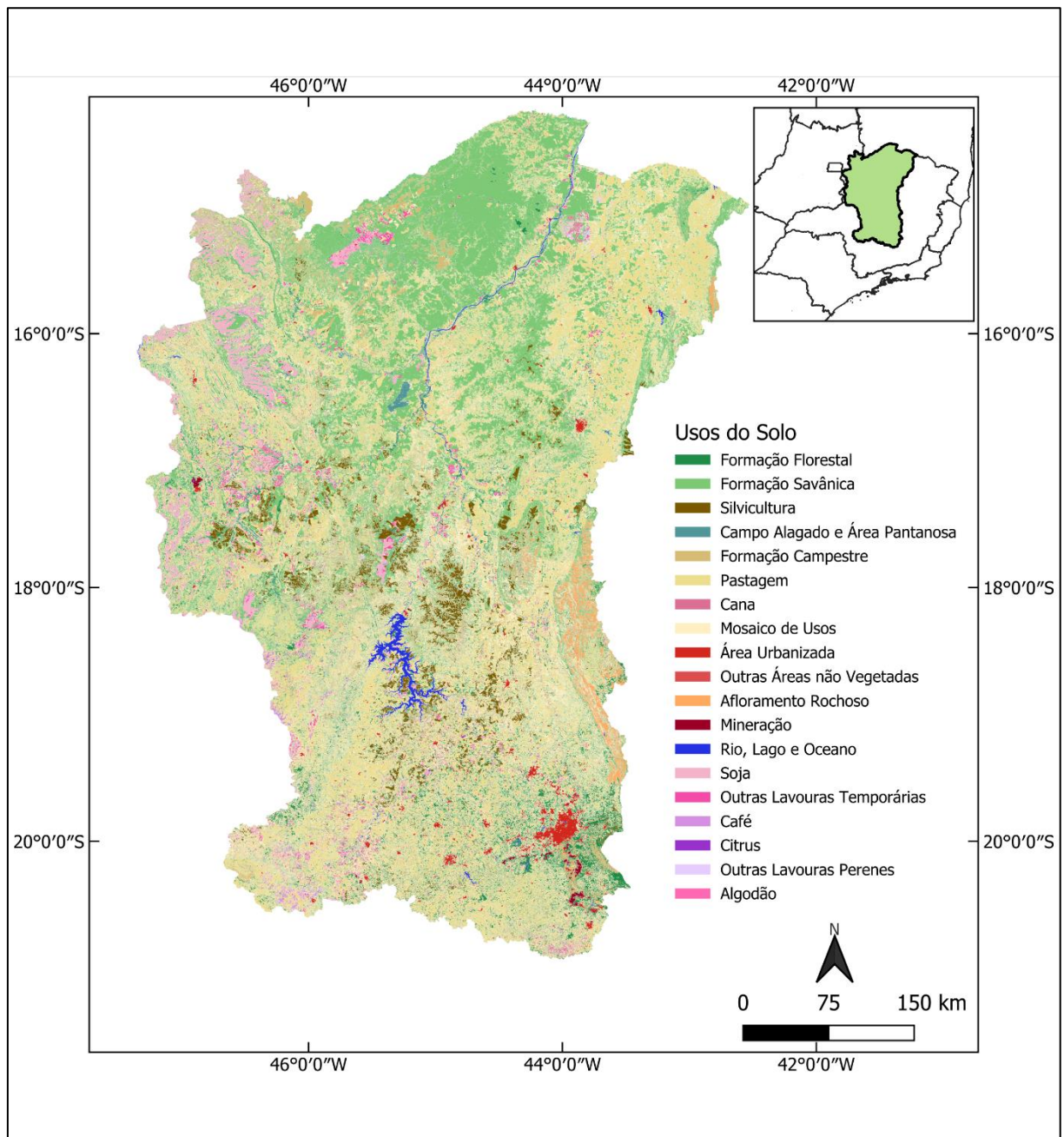
Tabela 2 – Classes de uso e ocupação do solo na bacia hidrográfica do Rio São Francisco em MG

Classe	Área Coberta (%)
Agropecuária	71,226
Área não vegetada	16,275
Floresta	7,356
Formação Natural não Florestal	4,991
Corpos D'água	0,150

Fonte: Projeto MapBiomias (2022)

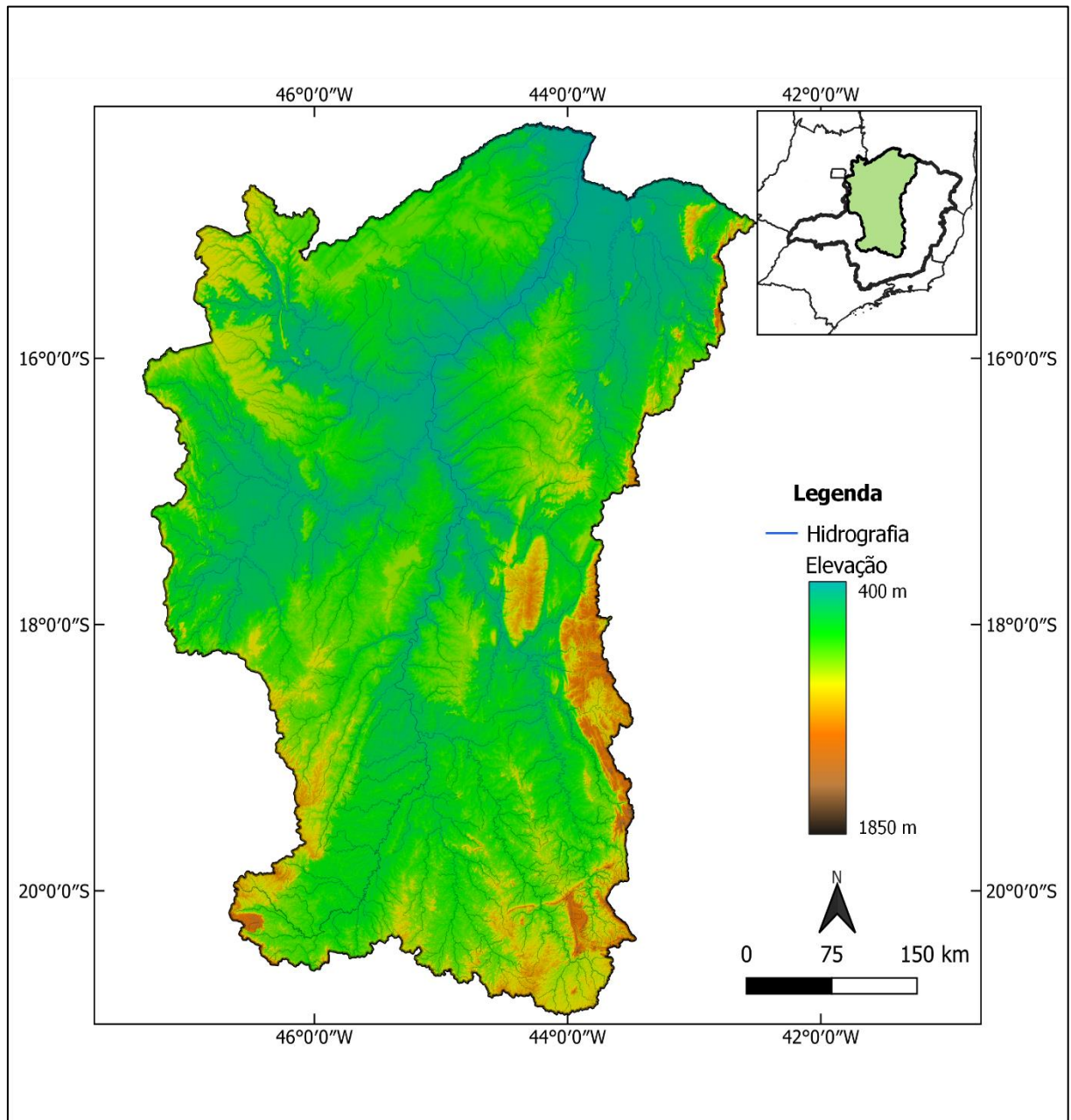
Com relação ao relevo da bacia hidrográfica analisada, a altimetria média da região apresenta grande variabilidade espacial, conforme pode ser visualizado na Figura 11. As superfícies localizadas ao sul da bacia, assim como a área que compreende a nascente do Rio São Francisco, apresentam elevações superiores a 1400 metros. O relevo decresce substancialmente ao longo da bacia atingindo elevações próximas a 400 metros no limite entre os estados de Minas Gerais e da Bahia. Dessa forma, é possível relacionar a variabilidade espacial das precipitações, citadas no tópico anterior, com a altimetria da bacia hidrográfica em foco, visto que tais características aparentam apresentar uma certa relação.

Figura 10 – Uso e ocupação do solo na bacia hidrográfica do Rio São Francisco no estado de Minas Gerais



Fonte: Projeto MapBiomias (2022)

Figura 11 – Mapa altimétrico da bacia hidrográfica do Rio São Francisco no estado de Minas Gerais



Fonte: NASA JPL (2013)

4.1.3 Hidrogeologia

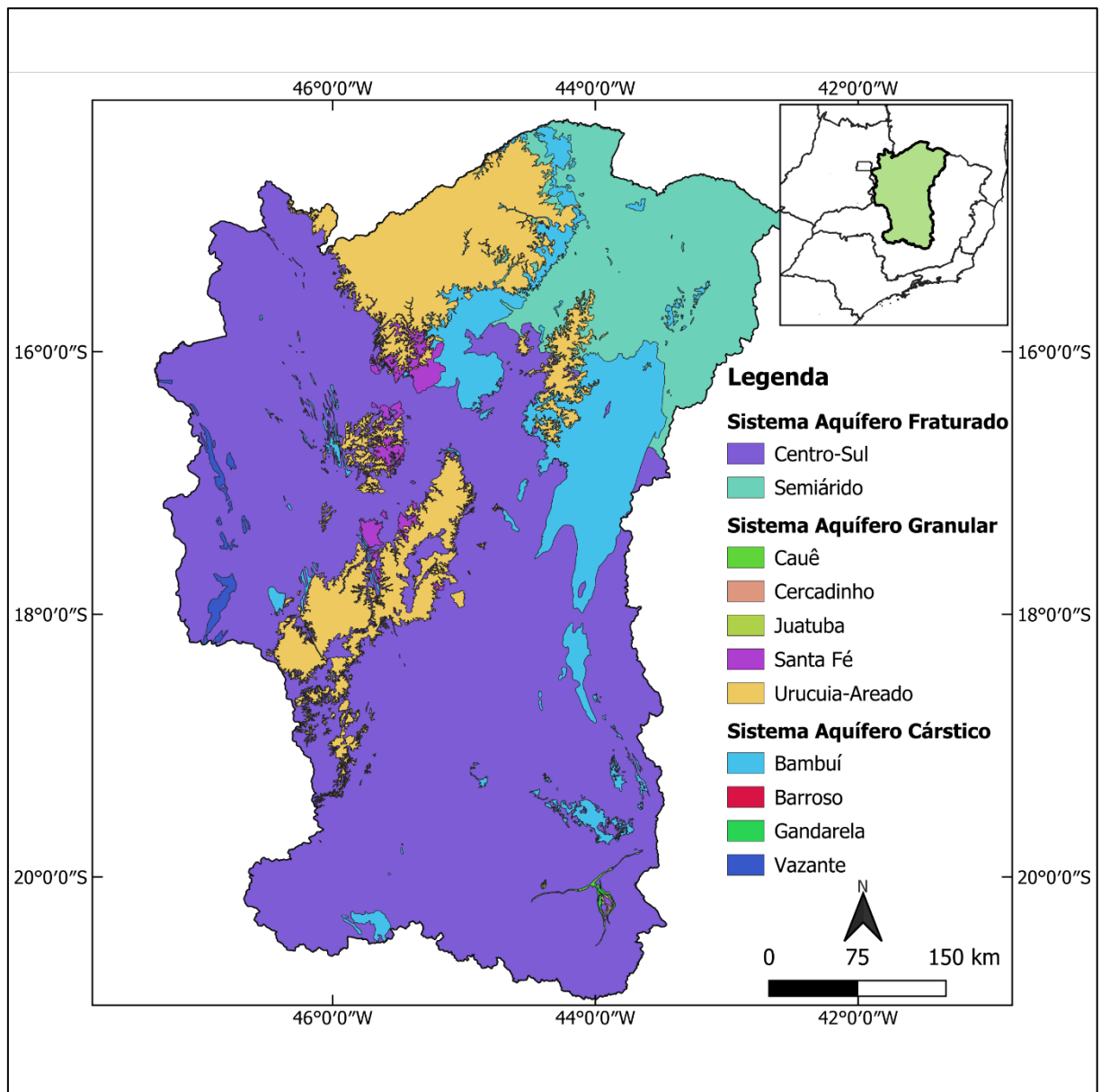
A porção mineira da bacia hidrográfica do rio São Francisco apresenta três sistemas aquíferos principais (CBHSF, 2016). São eles:

- Fraturado: “Associado às rochas, quer cristalinas de natureza metamórfica (xistos, migmatitos, granulitos, gnaisses), metassedimentar (quartzitos, metapelitos, entre outras) e ígnea (granitóides, rochas vulcânicas), quer sedimentares consolidadas (arenitos, conglomerados, siltitos, argilitos). Estes meios, em geral, impermeáveis ou de muito reduzida permeabilidade podem apresentar fraturação que permite a circulação da água e a individualização de aquíferos (CBHSF, 2016).”
- Cárstico: “Formado em rochas carbonatadas (calcários, calcarenitos, dolomitos, mármore) onde a circulação da água se faz nas discontinuidades com origem na dissolução do carbonato pela água. Apesar de alguma heterogeneidade dos meios cársticos, as suas produtividades são muito superiores às registradas nos meios rochosos fraturados ou mesmo nos granulares (CBHSF, 2016).”
- Granular (Poroso): “Formado por rochas sedimentares detríticas pouco ou não consolidadas. A circulação de água é feita nos poros entre os grãos, sendo que nas situações em que a presença da argila é reduzida, podem apresentar elevada permeabilidade e interesse aquífero (CBHSF, 2016).”

O domínio hidrogeológico fraturado ocupa cerca de 75% da área da bacia hidrográfica em estudo, dos quais 63% correspondem ao sistema aquífero fraturado Centro-sul e 12% ao semiárido. A unidade hidrogeológica granular ocupa aproximadamente 15 % do território em questão. A maior parte desse montante é abrangida pelo aquífero Urucuia-areado, sistema que apresenta uma das maiores produtividades dos aquíferos observados. Por fim, 10% da bacia hidrográfica apresentam os sistemas aquíferos cársticos Bambuí, Vazante, Gandarela e Barroso.

Os domínios hidrogeológicos verificados na porção mineira da bacia hidrográfica do rio São Francisco são apresentados na Figura 12.

Figura 12 - Mapa de aquíferos presentes na bacia hidrográfica do Rio São Francisco no estado de Minas Gerais



Fonte: ANA (2013)

4.2 ESTAÇÕES FLUVIOMÉTRICAS UTILIZADAS NO ESTUDO

Uma vez definida a área de estudos, prosseguiu-se para a seleção e análise da base de dados das bacias hidrográficas inseridas na região de interesse. Tal informação foi obtida do trabalho de Regionalização de vazões mínimas sazonais do Estado de Minas Gerais (EHR, 2021). Foram obtidas ao todo 51 estações fluviométricas distribuídas ao longo de toda a bacia hidrográfica do Rio São Francisco, das quais 36 foram destinadas à calibração do modelo e 15 foram reservadas para o processo de validação. A seleção das estações para esta finalidade ocorreu de forma aleatória.

As informações relevantes às estações utilizadas na concepção desse trabalho estão dispostas na Tabela 3.

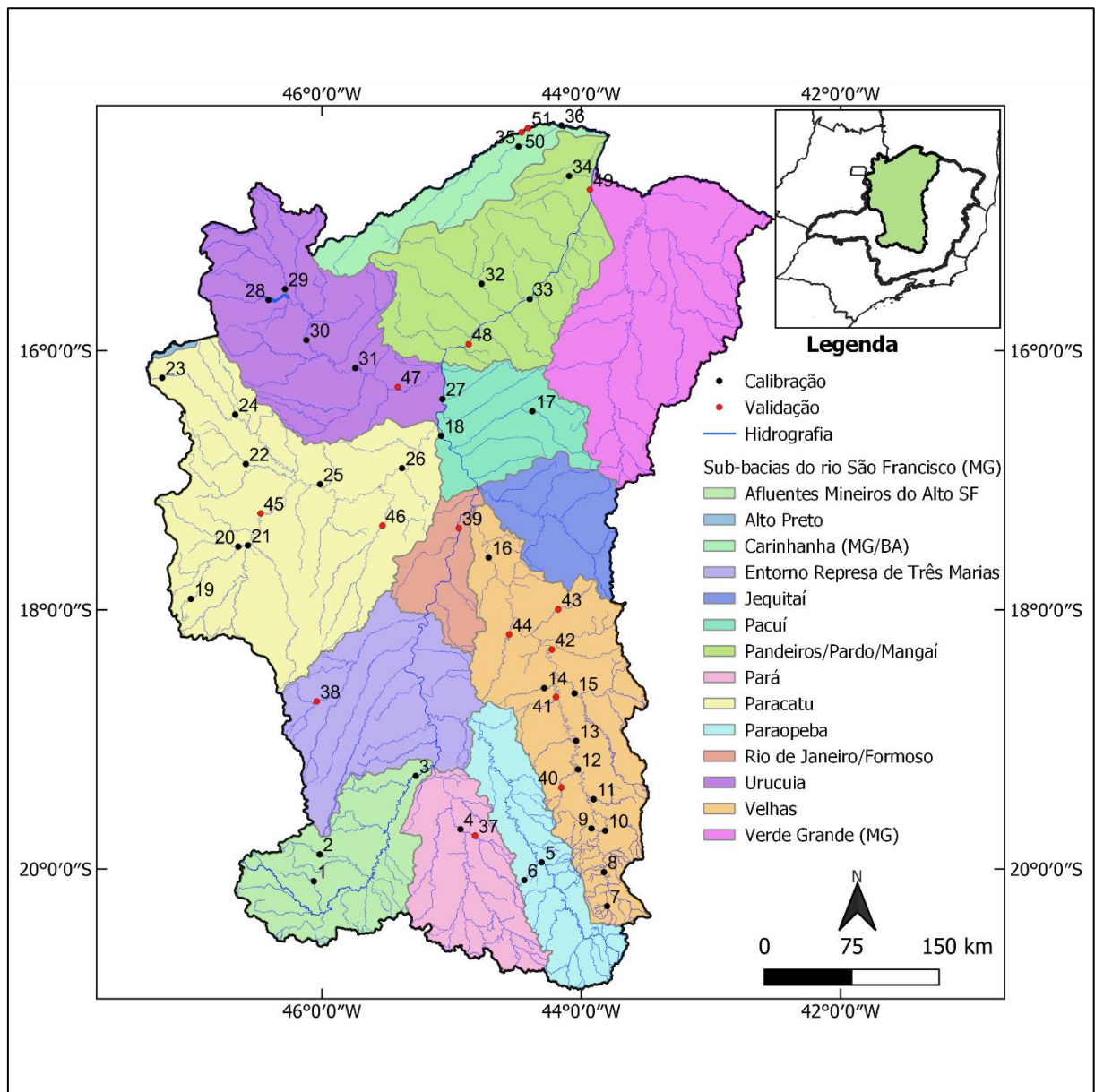
Tabela 3 – Estações fluviométricas utilizadas no estudo

ID	Código	Estação Fluviométrica	Curso de Água	Latitude	Longitude	Período de Dados
1	40040000	Fazenda Ajudas	Rio Ajudas	-20,095	-46,062	06/1939 - 10/2019
2	40060001	Tapiraí Jusante	Rio Da Perdição	-19,888	-46,017	07/1975 - 10/2019
3	40100000	Porto Das Andorinhas	Rio São Francisco	-19,281	-45,275	09/1956 - 09/2019
4	40330000	Velho Da Taipa	Rio Pará	-19,694	-44,931	07/1938 - 10/2019
5	40800001	Ponte Nova Do Paraopeba	Rio Paraopeba	-19,949	-44,305	01/1938 - 10/2019
6	40810800	Fazenda Pasto Grande	Ribeirão Serra Azul	-20,087	-44,438	06/1978 - 07/2019
7	41180000	Itabirito Linigrafo	Ribeirão Mata Porcos	-20,288	-43,800	06/1956 - 10/2019
8	41199998	Honório Bicalho Montante	Rio Das Velhas	-20,025	-43,824	04/1971 - 10/2019
9	41250000	Vespasiano	Ribeirão Da Mata	-19,688	-43,921	09/1938 - 10/2019
10	41260000	Pinhões	Rio Das Velhas	-19,705	-43,815	10/1975 - 09/2019
11	41380000	Ponte Preta	Rio Jabuticatubas	-19,462	-43,904	08/1941 - 10/2019
12	41410000	Jequitibá	Rio Das Velhas	-19,231	-44,025	07/1965 - 10/2019
13	41600000	Pirapama	Rio Das Velhas	-19,011	-44,038	11/1968 - 10/2019
14	41685000	Ponte Do Picão	Ribeirão Do Picão	-18,604	-44,285	07/1965 - 10/2019
15	41780002	Presidente Juscelino Jusante	Rio Paraúna	-18,645	-44,051	08/1976 - 10/2019
16	41990000	Várzea Da Palma	Rio Das Velhas	-17,596	-44,713	07/1938 - 10/2019
17	42187000	Ponte Dos Ciganos	Rio Pacuí	-16,467	-44,376	10/1973 - 12/2018
18	42210000	Cachoeira Da Manteiga	Rio São Francisco	-16,657	-45,081	01/1959 - 10/2019
19	42250000	Fazenda Limoeiro	Rio Claro	-17,915	-47,011	07/1973 - 10/2019
20	42257000	Barra Do Escurinho	Ribeirão Escurinho	-17,513	-46,646	10/1974 - 06/2018

ID	Código	Estação Fluviométrica	Curso de Água	Latitude	Longitude	Período de Dados
21	42290000	Ponte Da Br-040 - Paracatu	Rio Paracatu	-17,503	-46,571	08/1965 - 10/2019
22	42435000	Fazenda Barra Da Égua	Ribeirão Barra Da Égua	-16,874	-46,587	11/1974 - 09/2017
23	42460000	Fazenda Limeira	Rio Preto	-16,209	-47,233	05/1972 - 12/2015
24	42545500	Fazenda O Resfriado	Ribeirão Roncador	-16,493	-46,669	12/1972 - 04/2019
25	42690001	Porto Da Extrema	Rio Paracatu	-17,030	-46,014	04/1955 - 10/2019
26	42980000	Porto Alegre	Rio Paracatu	-16,907	-45,383	09/1963 - 09/2019
27	43200000	São Romão	Rio São Francisco	-16,373	-45,070	08/1952 - 10/2019
28	43250002	Buritis Jusante	Rio Urucuia	-15,608	-46,412	11/1976 - 10/2019
29	43300000	Fazenda Carvalho	Rio São Domingos	-15,524	-46,284	06/1973 - 10/2019
30	43429998	Arinos Montante	Rio Urucuia	-15,918	-46,119	08/1977 - 10/2019
31	43670000	Vila Urucuia	Rio Urucuia	-16,133	-45,742	04/1967 - 07/2019
32	44250000	Usina Do Pandeiros Montante	Rio Pandeiros	-15,483	-44,768	08/1973 - 07/2019
33	44290002	Pedras De Maria Da Cruz	Rio São Francisco	-15,601	-44,397	09/1972 - 07/2019
34	44540000	Fazenda Bom Retiro	Rio Japoré	-14,651	-44,095	10/1974 - 07/2011
35	45220000	Capitânea	Rio Cochá	-14,424	-44,483	11/1967 - 10/2019
36	45260000	Juvenília	Rio Carinhanha	-14,260	-44,152	04/1964 - 05/2019
37	40300001	Jaguaruna Jusante	Rio São João	-19,744	-44,818	01/1977 - 10/2019
38	41050000	Major Porto	Ribeirão Areado	-18,706	-46,040	08/1968 - 10/2019
39	41135000	Pirapora Barreiro	Rio São Francisco	-17,369	-44,943	06/1968 - 08/2019
40	41440005	Represa Jusante	Ribeirão Jequitibá	-19,371	-44,153	06/1979 - 08/2019
41	41650002	Ponte Do Licínio Jusante	Rio Das Velhas	-18,673	-44,194	10/1976 - 10/2019
42	41818000	Santo Hipólito	Rio Das Velhas	-18,306	-44,226	06/1938 - 10/2019
43	41890000	Estação De Curimatai	Rio Curimataí	-17,996	-44,178	08/1965 - 10/2019
44	41940000	Ponte Do Bicudo	Rio Do Bicudo	-18,191	-44,556	07/1965 - 06/2019
45	42395000	Santa Rosa	Rio Paracatu	-17,255	-46,473	05/1955 - 09/2019
46	42850000	Cachoeira Das Almas	Rio Do Sono	-17,351	-45,533	10/1973 - 10/2019
47	43880000	Santo Inácio	Rio Urucuia	-16,281	-45,414	03/1965 - 10/2019
48	44200000	São Francisco	Rio São Francisco	-15,949	-44,868	10/1924 - 10/2019
49	44500000	Manga	Rio São Francisco	-14,757	-43,932	11/1932 - 10/2019
50	45131000	São Gonçalo	Rio Carinhanha	-14,314	-44,459	06/1948 - 10/2019
51	45210000	Lagoa Das Pedras	Rio Carinhanha	-14,281	-44,410	05/1969 - 10/2019

A Figura 13 apresenta a divisão das sub-bacias do Rio São Francisco em Minas Gerais e a localização de todas as estações utilizadas nesse estudo.

Figura 13 - Sub-bacias do Rio São Francisco em Minas Gerais e localização dos postos fluviométricos utilizados na pesquisa



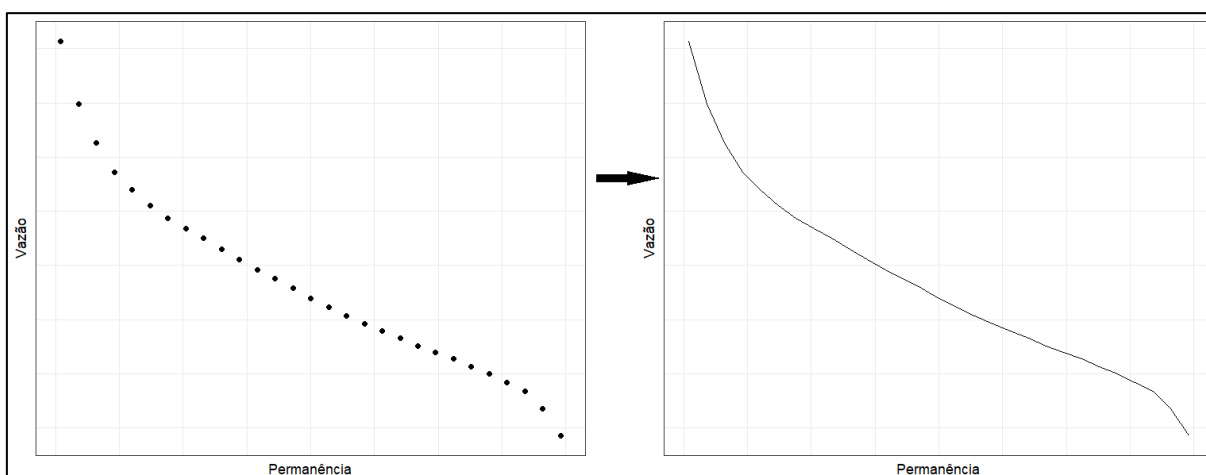
Fonte: ANA (2015)

As curvas de permanência expressam a porcentagem do tempo que determinada vazão foi igualada ou excedida durante o intervalo utilizado na sua construção. O ideal é que se disponha de um período longo o suficiente para que seja possível representar de forma eficaz o comportamento hidrológico do curso d'água em questão. Dessa forma, foi estabelecido um período mínimo de 30 anos de medições sem falhas e comum a todas as estações, para a construção das FDCs empíricas.

4.3 SUAVIZAÇÃO DAS CURVAS DE PERMANÊNCIA

A construção das curvas de permanência de longo termo empíricas seguiu o procedimento indicado no Tópico 3.2.3. As vazões de cada estação foram ordenadas de maneira decrescente e foi atribuída a cada uma sua correspondente probabilidade de excedência, conforme a posição de plotagem de Weibull, dada pela Equação 1. Concebidas as FDCs empíricas, necessita-se representar os dados em formato funcional, conforme ilustra a Figura 14. Tal procedimento é realizado expressando as FDCs empíricas como uma soma de funções de base conhecidas, conforme representado pela Equação 15.

Figura 14 – Representação da curva de permanência empírica em formato funcional



Observando o comportamento das curvas de permanência, optou-se por utilizar o sistema de bases B-Splines de ordem cúbica no processo de suavização dos dados, as mesmas funções utilizadas no estudo elaborado por Requena *et al.* (2018b).

Conforme descrito no Tópico 3.3.2.2, os sistemas de bases do tipo B-Splines pressupõem a definição dos nós que serão utilizados na suavização dos dados. Nesse estudo, os nós serão correspondentes às permanências das FDCs empíricas, definidas no intervalo $p(0,1)$.

Para ganhar flexibilidade com as funções B-Splines, é necessário estabelecer uma quantidade suficiente de nós internos que permitam garantir a aderência da função aos dados observados (RAMSAY E SILVERMAN, 2005). Requena *et al.* (2018b), sugerem utilizar não mais do que 50 nós na concepção das curvas funcionais.

Ademais, a distribuição dos nós deve priorizar as regiões onde a função expressa as variações mais complexas, concentrando menos pontos onde a curva apresenta comportamento linear (RAMSAY E SILVERMAN, 2005).

Tendo em vista as considerações citadas, foi estabelecido um total de 45 permanências como nós do sistema de bases B-splines, priorizando os ramos superior e inferior das curvas de permanência. O conjunto de nós utilizados está descrito na Tabela 4.

Tabela 4 – Permanências utilizadas na construção das FDCs funcionais

Quantidade de nós	Permanências
45	0.1, 0.3, 0.5, 0.7, 0.9, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10
	15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85
	90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 99.1, 99.3, 99.5, 99.7, 99.9

A escolha do parâmetro de suavização λ é uma importante etapa do processo de suavização e pode ser realizada de duas formas distintas. A primeira considera a escolha do parâmetro como livre e particular a cada conjunto de dados em estudo. Tal processo consiste em utilizar vários parâmetros de suavização para, assim, selecionar o que melhor se adequa aos dados medidos (SILVA, 2019). A segunda efetua a escolha do parâmetro λ de maneira automática, geralmente, utilizando a validação cruzada generalizada. A escolha de λ é realizada minimizando a Equação 31.

$$VCG(\lambda) = \left(\frac{n}{n-df(\lambda)} \right) \left(\frac{\sum_{j=1}^n (y_j - \hat{x}(t_j))^2}{n-df(\lambda)} \right) \quad (31)$$

na qual $df(\lambda)$ são os graus de liberdade da curva suavizada n vezes. No entanto, o método automático deve ser empregado somente como um guia, ao invés de uma regra fixa (ZIN *et al.*, 2020). No presente trabalho, o método automático foi utilizado de forma a refinar a escolha do melhor parâmetro que se adequa aos dados. Inicialmente, foi definido um intervalo de valores possíveis de λ para que, em seguida, com base no critério da validação cruzada e do ajuste visual das curvas, fosse escolhido o parâmetro mais adequado aos valores medidos.

4.4 ELABORAÇÃO DO MODELO REGIONAL DAS CURVAS DE PERMANÊNCIA

A próxima etapa da metodologia compreendeu a construção do modelo regional das FDCs. Para esse propósito, foi utilizada a regressão funcional múltipla para regionalizar as curvas de permanência de longo termo, a qual relaciona as permanências da curva regional às características físicas, climáticas e morfológicas das bacias hidrográficas utilizadas no processo de calibração.

Na presente dissertação, foi utilizada a Regressão Funcional Múltipla (FMR) (Equação 30), proposta por Requena *et al.* (2018b), para regionalizar as FDCs. A FMR correlaciona a resposta funcional com variáveis independentes e coeficientes funcionais, conforme a Equação 32 (REQUENA *et al.*, 2018b):

$$y_i(t) = \beta_0(t) + \sum_{j=1}^q z_{ij}\beta_j(t) + \varepsilon_i(t); i = 1, \dots, n \quad (32)$$

na qual $y_i(t)$ é a resposta funcional, n é o tamanho da amostra, z_{ij} são as variáveis independentes, $\beta_0(t)$ é o intercepto da função, $\beta_j(t)$ são os coeficientes funcionais e $\varepsilon_i(t)$ são os resíduos funcionais da regressão. Na presente dissertação, $y_i(t)$ são as FDCs funcionais e z_{ij} são as características físicas, climáticas e morfológicas das bacias hidrográficas.

Reescrevendo a Equação 32 em termos matriciais, tem-se:

$$\mathbf{y}(t) = \mathbf{Z}\boldsymbol{\beta}(t) + \boldsymbol{\varepsilon}(t) \quad (33)$$

Os coeficientes funcionais podem ser obtidos minimizando a soma do quadrado dos erros, conforme a Equação 34 (REQUENA *et al.*, 2018b).

$$SQE(\boldsymbol{\beta}) = \int (\mathbf{C}\boldsymbol{\Phi}(t) - \mathbf{Z}\mathbf{B}\boldsymbol{\Theta}(t))^T (\mathbf{C}\boldsymbol{\Phi}(t) - \mathbf{Z}\mathbf{B}\boldsymbol{\Theta}(t)) dt \quad (34)$$

onde $\boldsymbol{\Phi}(t)$ e $\boldsymbol{\Theta}(t)$ são os vetores contendo as funções de base linearmente independentes, $\mathbf{C}$ é a matriz $n \times K_y$ e $\mathbf{B}$ é a matriz $n \times K_\beta$ de coeficientes, sendo K_y e K_β o número das respectivas funções de base. Adicionando o termo de penalização

da não suavidade, descrito no Tópico 3.3.3, temos a soma dos quadrados dos erros penalizada, expressa pela Equação 35 (REQUENA *et al.*, 2018b).

$$SQEPEN_{\lambda}(\beta) = \int (\mathbf{C}\Phi(t) - \mathbf{ZB}\Theta(t))^T (\mathbf{C}\Phi(t) - \mathbf{ZB}\Theta(t)) dt + \lambda \int (\mathbf{LB}\Theta(t))^T (\mathbf{LB}\Theta(t)) dt \quad (35)$$

onde $\mathbf{L}$ é um operador diferencial linear geralmente definido como a segunda derivada de β . Com o auxílio da Equação 35, foi possível construir o modelo regional da Regressão Funcional Múltipla. Tal procedimento compreendeu a inserção de cada variável, uma por vez, seguida da avaliação visual do ajuste das curvas e de algumas métricas de desempenho, descritas nas seções seguintes. O acréscimo de uma variável dependente só foi considerado caso existisse melhora expressiva em cada um dos pontos citados. Tal decisão foi considerada visando um modelo parcimonioso e com maior eficiência computacional.

O processo de regionalização considerou as seguintes características descritivas das bacias como variáveis dependentes do modelo: a área de drenagem (A), o comprimento do rio principal (L), a produção do aquífero (Q_s), a precipitação média anual (P), a densidade de drenagem (DD) e o coeficiente de compactidade (Kc). As 5 primeiras covariáveis foram obtidas do trabalho de Regionalização de vazões mínimas sazonais do Estado de Minas Gerais (EHR/2021). As características morfológicas das bacias hidrográficas, densidade de drenagem e coeficiente de compactidade, foram obtidas pelas Equações 36 e 37, respectivamente.

$$DD = \frac{L_t}{A} \quad (36)$$

$$Kc = \frac{0,282P}{\sqrt{A}} \quad (37)$$

na qual, L_t corresponde ao comprimento total de rede de drenagem, A é a área de drenagem e P refere-se ao perímetro da bacia hidrográfica.

Após a construção do modelo regional, com o intuito de garantir a monotonicidade intrínseca da curva de permanência, o modelo de suavização “loess” foi aplicado nas

FDCs calibradas. Tal técnica permite encontrar uma curva suavizada por meio de uma regressão de mínimos quadrados ponderada, que é aplicada a um conjunto de pontos próximos de um intervalo previamente determinado (MARIANI E BASU, 2014). Para mais informações, o leitor pode consultar Cleveland (1979).

Obtida a equação funcional regional, foi possível avaliar a influência de cada variável dependente, isso é, cada característica descritiva das bacias hidrográficas, na concepção do modelo regional. Tal análise foi realizada considerando os coeficientes funcionais da FMR relacionados às suas respectivas características descritivas. O processo de avaliação foi realizado por meio da análise gráfica de cada ramo das curvas de permanência de longo termo regionais.

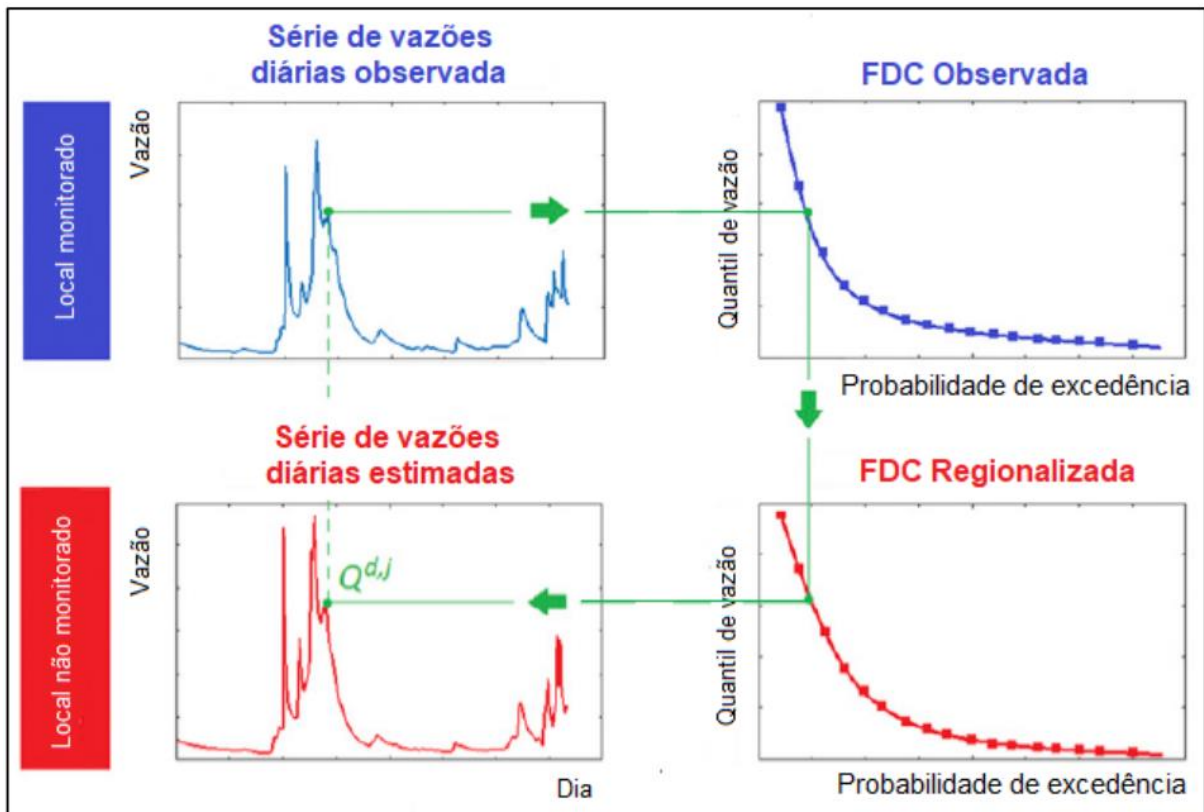
A etapa seguinte consistiu na avaliação do comportamento das equações regionais na estimativa das curvas de permanência de longo termo para as estações não utilizadas na construção do modelo. Esse processo objetiva examinar a capacidade preditiva do modelo, considerando as relações regionais obtidas na etapa de calibração.

Com esse propósito, o comportamento das curvas regionais obtidas no processo de validação foi avaliado por meio das mesmas métricas de desempenho utilizadas na etapa de calibração. A análise gráfica do ajuste visual também foi realizada para as estações não compreendidas na construção do modelo regional.

4.5 ESTIMAÇÃO DAS SÉRIES DE VAZÕES DIÁRIAS SINTÉTICAS

A estimação de séries de vazões diárias sintéticas é comumente efetuada por meio do método proposto por Hughes e Smakhtin (1996) (por exemplo, REQUENA *et al*, 2017; REQUENA *et al*, 2018a; REQUENA *et al*, 2018b; SHU E OUARDA, 2012), ilustrada pela Figura 15. Essa abordagem admite que a probabilidade de excedência da vazão em um determinado dia é a mesma para os locais de destino e origem.

Figura 15 - Método de interpolação espacial não linear proposto por Hughes e Smakhtin (1996)



Fonte: Adaptado de Requena *et al.* (2017)

O método citado pode ser implementado considerando as seguintes etapas: (i) inicialmente, a vazão média diária de um dia específico d de uma estação fluviométrica monitorada é obtida por meio da sua respectiva série histórica de vazões observadas; (ii) a probabilidade de excedência p da vazão média diária da bacia medida é extraída da sua respectiva curva de permanência de longo termo empírica; em seguida (iii) a probabilidade de excedência obtida na etapa anterior é identificada na FDC regionalizada para a bacia hidrográfica de interesse, assim como o seu quantil de vazão média diária correspondente. Pontos intermediários da FDC são obtidos por interpolação linear entre duas permanências conhecidas; por fim, (iv) assume-se que a descarga obtida pela curva regional corresponde à vazão média da bacia hidrográfica não monitorada naquele dia específico. Realizando tal processo para cada dia presente no intervalo preestabelecido, é possível obter a série de vazões médias diárias para um conjunto de regiões de interesse.

Caso sejam consideradas n postos monitorados no processo de interpolação espacial, a vazão diária será a média ponderada das descargas obtidas, conforme a Equação 38.

$$Q_d = \frac{\sum_{j=1}^n \lambda_j Q_{d,j}}{\sum_{j=1}^n \lambda_j} \quad (38)$$

Em estudos de transferência de séries de vazões, a identificação das bacias doadoras é essencial para se obter estimativas adequadas para as regiões de interesse. Em geral, é preferível utilizar mais de uma região monitorada. Entretanto, o emprego de muitas bacias monitoradas não é indicado (REQUENA *et al.*, 2017). São comumente utilizadas entre 3 e 5 bacias doadoras no processo de interpolação espacial (por exemplo, HUGHES E SMAKHTIN, 1996; REQUENA *et al.*, 2017; REQUENA *et al.*, 2018a; REQUENA *et al.*, 2018b; SHU E OUARDA, 2012) e a sua seleção, na maior parte das vezes, é definida pela proximidade espacial entre as bacias hidrográficas de interesse. Hughes e Smakhtin (1996) sugerem que a seleção das regiões doadoras pode considerar os locais que estão presentes em um mesmo rio ou em seus afluentes, mas distâncias elevadas podem influir negativamente no resultado da predição.

Tendo em vista os aspectos citados, o presente trabalho avaliou a utilização da abordagem proposta por Hughes e Smakhtin (1996) e conjunto com a Krigagem Topológica na estimação das séries de vazões médias diárias para as estações de validação, definidas no Tópico 4.2. O Top-Kriging foi utilizado para definir os pesos de interpolação λ_j da Equação 38. Para tanto, a variável utilizada na predição espacial do método foi a vazão média de longo termo (Q_{MLT}) das estações de calibração.

Seguindo as indicações encontradas na literatura, foram avaliadas as utilizações de 3, 4, 5 e 6 estações monitoradas na geração das séries sintéticas de vazões. A avaliação da abordagem foi efetuada seguindo as métricas de desempenho descritas no tópico seguinte. Por fim, dispondo das séries de vazões estimadas, foi avaliada a capacidade de predição do modelo das vazões médias mínimas de 7 dias de duração (Q_7), vazões médias mensais e vazões médias diárias máximas anuais.

Todas as etapas da presente dissertação foram desenvolvidas em linguagem R por meio do software RStudio (versão 2023.12.1-402).

4.6 AVALIAÇÃO DA METODOLOGIA EMPREGADA

A eficiência da abordagem proposta foi avaliada com o auxílio de métricas de desempenho. Seguindo Requena *et al.* (2018b) e Costa *et al.* (2020), foram consideradas as seguintes métricas: coeficiente de determinação (R^2), raiz quadrada da média dos erros quadráticos (RMSE), raiz do quadrado médio do erro da predição (RMSPE), critério de eficiência de Nash-Sutcliffe (NSE), coeficiente de persistência (CP), critério de eficiência de Kling-Gupta (KGE), viés percentual (PBIAS), erro médio absoluto (MAE), cujas equações estão definidas nas Eqs. 39 a 46.

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (39)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (40)$$

$$RMSPE = \sqrt{\frac{1}{n} \sum_{i=1}^n \left(\frac{y_i - \hat{y}_i}{y_i} \right)^2} 100\% \quad (41)$$

$$NSE = 1 - \frac{\sum_{i=1}^n (\hat{y}_i - y_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (42)$$

$$CP = 1 - \frac{\sum_{i=2}^n (\hat{y}_i - y_i)^2}{\sum_{i=1}^{n-1} (y_{i+1} - y_i)^2} \quad (43)$$

$$KGE = 1 - \sqrt{(r - 1)^2 + \left(\frac{\sigma_{sim}}{\sigma_{obs}} - 1 \right)^2 + \left(\frac{\mu_{sim}}{\mu_{obs}} - 1 \right)^2} \quad (44)$$

$$PBIAS = 100\% \frac{\sum_{i=1}^n (\hat{y}_i - y_i)}{\sum_{i=1}^n y_i} \quad (45)$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (46)$$

nas quais n é o tamanho da amostra, y_i é o i -ésimo valor observado, $\hat{y}_i$ é o i -ésimo valor estimado, $\bar{y}$ é a média dos valores observados. Com relação ao KGE, r

corresponde ao coeficiente de correlação linear, σ_{sim} e σ_{obs} são os desvios padrão estimados e observados, μ_{sim} e μ_{obs} são as médias dos valores simulados e observados, respectivamente.

A eficiência da Krigagem topológica, assim como a reprodução das vazões destacadas (médias mínimas de 7 dias de duração, médias mensais e máximas anuais) foram apreciadas por meio do erro percentual absoluto médio (MAPE) dado pela Equação 47.

$$MAPE = \frac{100\%}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right| \quad (47)$$

sendo n o tamanho da amostra, y_i o i -ésimo valor observado, $\hat{y}_i$ o i -ésimo valor estimado.

5 RESULTADOS E DISCUSSÃO

5.1 ELABORAÇÃO DAS CURVAS DE PERMANÊNCIA FUNCIONAIS

5.1.1 Definição das estações fluviométricas utilizadas no estudo

A fim de representar com certa precisão o comportamento hidrológico das bacias hidrográficas utilizadas no presente estudo, foi estabelecido um período mínimo de 30 anos de medições sem falhas para a construção das curvas de permanência de longo termo. O período de dados que atendeu a esse critério compreendeu o intervalo de outubro de 1979 e setembro de 2009. A Tabela 5 apresenta algumas estatísticas amostrais das vazões médias diárias, no intervalo citado, das estações utilizadas na concepção desse trabalho.

Tabela 5 – Estatísticas amostrais das estações fluviométricas utilizadas no estudo

Código	Curso de Água	Média (m³/s)	Mediana (m³/s)	Máxima (m³/s)	Mínima (m³/s)
40040000	Rio Ajudas	5,65	3,30	91,40	0,71
40060001	Rio Da Perdição	12,21	9,00	207,73	3,33
40100000	Rio São Francisco	242,88	148,70	2901,45	33,54
40330000	Rio Pará	109,65	73,21	1523,92	19,74
40800001	Rio Paraopeba	86,03	60,08	2193,84	11,88
40810800	Ribeirão Serra Azul	0,79	0,53	30,14	0,11
41180000	Ribeirão Mata Porcos	7,74	5,80	350,65	2,28
41199998	Rio Das Velhas	30,58	22,06	470,59	8,92
41250000	Ribeirão Da Mata	8,99	5,16	293,77	0,72
41260000	Rio Das Velhas	64,08	42,64	1062,08	15,34
41380000	Rio Jabuticatubas	6,72	3,42	247,84	0,81
41410000	Rio Das Velhas	100,93	64,36	1450,71	24,99
41600000	Rio Das Velhas	112,09	73,59	1425,37	26,98
41685000	Ribeirão Do Picão	3,21	1,37	61,54	0,01
41780002	Rio Paraúna	75,79	31,45	1779,17	5,81
41990000	Rio Das Velhas	316,78	162,06	3965,83	41,91
42187000	Rio Pacuí	5,56	3,57	184,94	0,64
42210000	Rio São Francisco	1267,56	919,02	9861,73	352,22
42250000	Rio Claro	8,00	5,11	130,69	0,48
42257000	Ribeirão Escurinho	29,06	17,69	187,87	2,15
42290000	Rio Paracatu	104,31	55,95	867,11	8,10
42435000	Ribeirão Barra Da Égua	16,98	9,43	145,78	0,51

Código	Curso de Água	Média (m³/s)	Mediana (m³/s)	Máxima (m³/s)	Mínima (m³/s)
42460000	Rio Preto	57,64	47,41	618,10	1,77
42545500	Ribeirão Roncador	7,82	3,83	481,95	0,74
42690001	Rio Paracatu	354,04	211,90	2601,19	34,80
42980000	Rio Paracatu	485,61	278,81	3297,97	36,77
43200000	Rio São Francisco	1723,84	1190,81	10289,58	378,15
43250002	Rio Urucuia	39,33	17,64	559,04	2,94
43300000	Rio São Domingos	47,33	27,94	478,20	3,21
43429998	Rio Urucuia	139,56	76,10	919,16	8,99
43670000	Rio Urucuia	198,52	114,62	1366,52	11,18
44250000	Rio Pandeiros	22,76	18,98	153,40	8,07
44290002	Rio São Francisco	2077,20	1416,54	12969,98	514,67
44540000	Rio Japoré	1,90	1,33	39,24	0,35
45220000	Rio Cochá	11,73	10,21	87,51	4,62
45260000	Rio Carinhanha	139,75	124,45	632,51	72,52
40300001	Rio São João	20,94	13,56	454,73	2,02
41050000	Ribeirão Areado	20,41	9,35	557,29	2,11
41135000	Rio São Francisco	872,91	704,80	5648,92	296,60
41440005	Ribeirão Jequitibá	3,10	2,09	62,39	0,65
41650002	Rio Das Velhas	133,62	84,21	1709,45	28,31
41818000	Rio Das Velhas	205,21	116,35	2161,56	37,10
41890000	Rio Curimataí	18,62	4,33	658,43	0,40
41940000	Rio Do Bicudo	18,94	6,34	452,35	0,10
42395000	Rio Paracatu	169,34	95,36	1105,98	18,78
42850000	Rio Do Sono	55,24	28,10	1248,44	3,56
43880000	Rio Urucuia	244,89	140,66	1570,44	13,72
44200000	Rio São Francisco	1992,06	1347,49	16738,43	428,33
44500000	Rio São Francisco	2007,07	1351,58	10797,65	433,49
45131000	Rio Carinhanha	67,70	60,49	239,44	32,32
45210000	Rio Carinhanha	128,77	115,41	594,62	66,93

5.1.2 Representação das curvas de permanência empíricas em termos funcionais

A primeira etapa do método da Análise Funcional de Dados consiste em expressar os dados pontuais de uma série de medições em termo funcionais. Para tanto, as curvas de permanência empíricas das 36 FDCs utilizadas no processo de calibração foram expressas em termos funcionais, considerando um sistema de funções de base do tipo B-Splines, no qual os seus respectivos nós foram definidos na Tabela 4 do Tópico 4.3.

Posteriormente, efetuou-se a suavização dos dados utilizando a soma dos quadrados dos erros penalizadas ($SQEPEN_\lambda$) definida na Equação 25. O comportamento do parâmetro de suavização λ para o intervalo predefinido ($10^{-1}, 10^{-10}$) pode ser visualizado na Figura 16.

Figura 16 – Média da validação cruzada generalizada para um intervalo de valores de λ

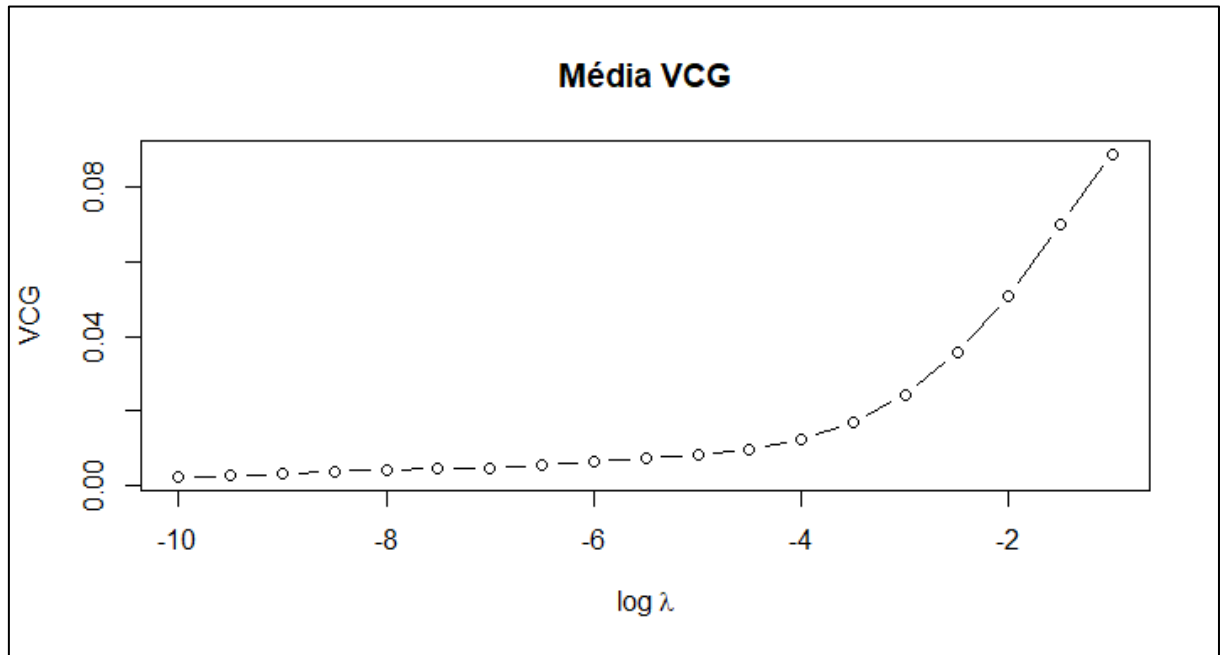


Figura 17 – FDC da estação 40040000 suavizada para três valores distintos de λ

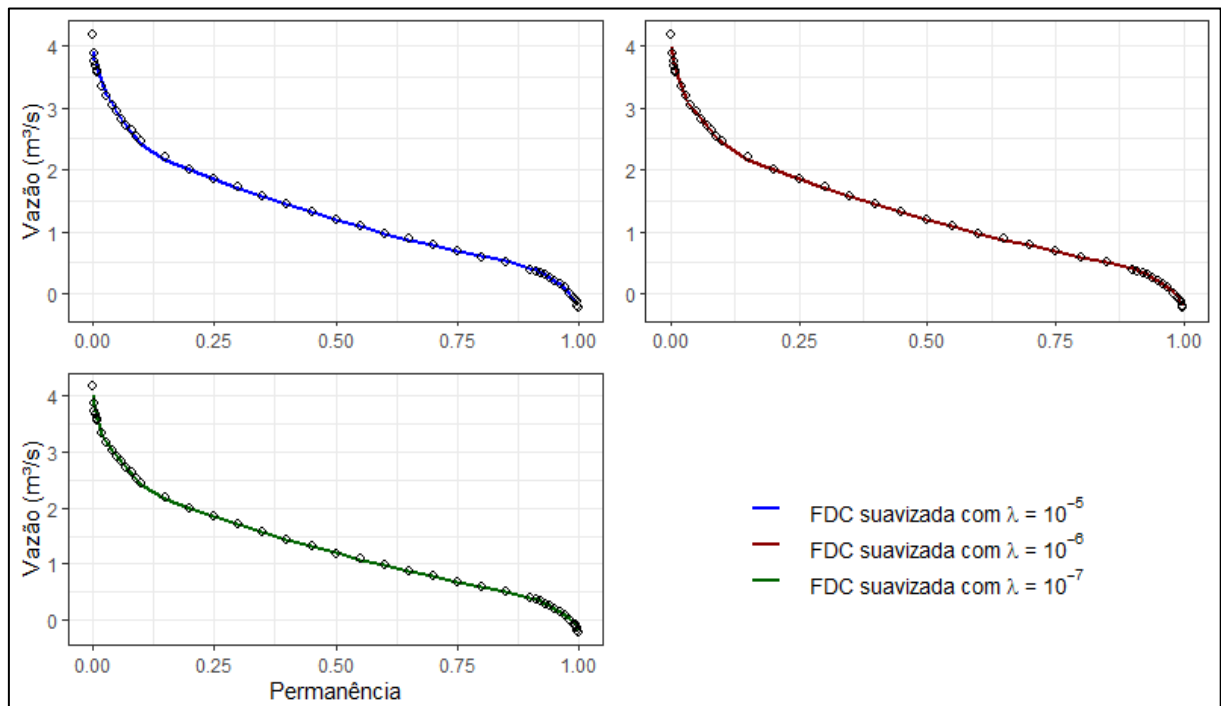
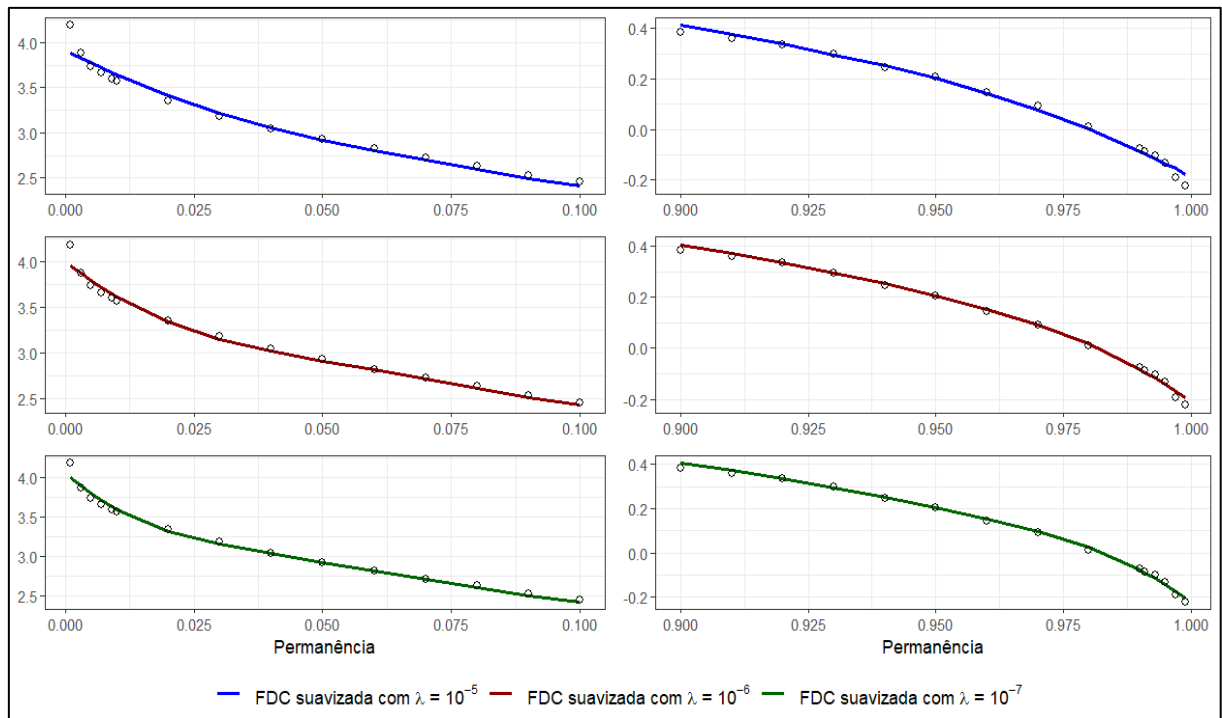


Figura 18 – Ramos superior (esquerda) e inferior (direita) da FDC da estação 40040000 suavizada para os valores de λ iguais a 10^{-5} , 10^{-6} e 10^{-7}



Percebe-se um decaimento mais acentuado do valor da validação cruzada generalizada (VCG) entre o intervalo compreendido pelos valores de 10^{-1} e 10^{-5} . A partir deste último, nota-se que curva da VCG tende à estabilidade. Para auxiliar na escolha do melhor parâmetro, as Figuras 17 e 18 ilustram a FDC suavizada para a Estação 40040000 utilizando os valores 10^{-5} , 10^{-6} e 10^{-7} .

Conforme pode ser observado, a suavização utilizando $\lambda = 10^{-5}$ expõe uma certa limitação no que concerne à representação das caudas das curvas de permanência. Em contrapartida, a utilização de $\lambda = 10^{-7}$ indica que a suavização começa a se apoiar na interpolação dos dados medidos, o que não é indicado devido à possibilidade de modelagem de ruídos inerentes aos processos de mensuração de tais dados. Desse modo, o parâmetro de suavização foi definido como $\lambda = 10^{-6}$, uma vez que aparenta ser a melhor escolha baseando-se nas análises visuais de aderência aos dados.

5.2 REGIONALIZAÇÃO DAS CURVAS DE PERMANÊNCIA DE LONGO TERMO FUNCIONAIS

5.2.1 Características físicas, climáticas e morfológicas utilizadas

As características físicas, climáticas e morfológicas das 36 bacias hidrográficas, utilizadas como variáveis dependentes na concepção da Regressão Funcional Múltipla, podem ser visualizadas na Tabela 6.

Tabela 6 – Características descritivas das estações fluviométricas utilizadas no estudo

ID	Código	Curso de Água	A (km ²)	L (km)	Q _s (m ³ /h)	P (mm)	DD (junções/km ²)	Kc
1	40040000	Rio Ajudas	243,30	61,54	13,62	1666,60	0,52	1,90
2	40060001	Rio Da Perdição	571,08	63,15	9,85	1694,54	0,64	1,82
3	40100000	Rio São Francisco	13861,34	456,65	14,51	1519,82	0,61	2,01
4	40330000	Rio Pará	7350,60	232,00	5,53	1443,89	0,47	2,04
5	40800001	Rio Paraopeba	5689,82	221,19	6,06	1508,95	0,40	1,94
6	40810800	Ribeirão Serra Azul	54,53	11,57	5,86	1499,80	0,39	1,62
7	41180000	Ribeirão Mata Porcos	378,88	50,32	13,14	1544,68	0,37	2,01
8	41199998	Rio Das Velhas	1542,25	93,02	10,76	1567,20	0,40	1,61
9	41250000	Ribeirão Da Mata	703,50	72,12	18,48	1407,68	0,37	1,77
10	41260000	Rio Das Velhas	3716,79	164,75	11,82	1529,48	0,40	2,03
11	41380000	Rio Jabuticatubas	564,02	68,70	17,94	1391,31	0,39	1,74
12	41410000	Rio Das Velhas	7070,95	305,86	16,23	1453,61	0,43	1,81
13	41600000	Rio Das Velhas	8049,50	380,07	16,52	1432,73	0,44	1,91
14	41685000	Ribeirão Do Picão	539,71	67,27	17,50	1200,64	2,17	1,89
15	41780002	Rio Paraúna	4007,06	269,91	18,46	1324,90	0,47	2,13
16	41990000	Rio Das Velhas	26536,09	746,19	17,53	1289,49	0,70	2,18
17	42187000	Rio Pacuí	1317,30	114,09	34,88	1028,54	1,71	1,98
18	42210000	Rio São Francisco	106663,52	889,54	15,68	1334,43	0,76	2,14
19	42250000	Rio Claro	465,11	51,14	15,84	1500,00	1,37	1,56
20	42257000	Ribeirão Escurinho	2043,32	143,70	16,65	1417,81	1,89	1,82
21	42290000	Rio Paracatu	7746,53	188,60	15,61	1442,81	1,44	1,87
22	42435000	Ribeirão Barra Da Égua	1619,87	81,70	9,39	1381,48	1,89	1,95
23	42460000	Rio Preto	3947,15	192,27	16,58	1573,37	1,29	1,94
24	42545500	Ribeirão Roncador	686,87	120,66	17,50	1501,73	1,08	2,72
25	42690001	Rio Paracatu	30076,62	366,94	14,78	1420,90	1,46	2,26
26	42980000	Rio Paracatu	41454,49	481,51	14,16	1374,85	1,48	2,25
27	43200000	Rio São Francisco	153642,28	929,82	15,24	1337,27	0,90	2,38

ID	Código	Curso de Água	A (km ²)	L (km)	Q _s (m ³ /h)	P (mm)	DD (junções/km ²)	Kc
28	43250002	Rio Urucuia	2952,60	135,54	17,50	1501,11	1,12	1,71
29	43300000	Rio São Domingos	3185,30	157,18	17,50	1325,36	1,39	2,17
30	43429998	Rio Urucuia	11678,87	271,61	19,82	1322,14	1,32	2,05
31	43670000	Rio Urucuia	18611,94	388,77	21,54	1289,59	1,44	1,98
32	44250000	Rio Pandeiros	3248,52	113,65	88,27	1076,70	2,36	1,94
33	44290002	Rio São Francisco	194021,88	1067,98	17,61	1303,43	0,99	2,29
34	44540000	Rio Japoré	677,32	74,92	7,37	910,78	2,39	1,97
35	45220000	Rio Cochá	2381,12	136,96	86,80	1062,80	2,69	2,13
36	45260000	Rio Carinhanha	16406,49	384,06	76,88	1081,38	2,43	1,96
37	40300001	Rio São João	1560,34	136,14	5,51	1443,12	0,45	2,14
38	41050000	Ribeirão Areado	1194,25	72,48	10,00	1500,00	1,28	1,83
39	41135000	Rio São Francisco	62003,39	771,38	12,30	1423,92	0,68	2,21
40	41440005	Ribeirão Jequitibá	238,27	40,61	37,44	1333,09	0,62	1,73
41	41650002	Rio Das Velhas	10676,48	500,84	16,76	1382,13	0,51	2,00
42	41818000	Rio Das Velhas	16691,78	601,23	17,26	1346,33	0,55	2,12
43	41890000	Rio Curimataí	1437,40	87,73	20,92	1186,71	1,08	1,96
44	41940000	Rio Do Bicudo	2079,82	124,93	17,50	1201,76	1,52	1,61
45	42395000	Rio Paracatu	12799,47	246,79	13,65	1429,47	1,61	2,00
46	42850000	Rio Do Sono	4401,96	185,29	8,02	1300,16	1,33	1,89
47	43880000	Rio Urucuia	23889,02	444,33	20,64	1260,36	1,47	1,98
48	44200000	Rio São Francisco	183539,31	995,71	16,08	1318,29	0,96	2,29
49	44500000	Rio São Francisco	202576,14	1185,08	17,62	1285,82	1,02	2,38
50	45131000	Rio Carinhanha	6022,96	341,99	85,56	1133,72	2,24	2,54
51	45210000	Rio Carinhanha	12579,21	318,18	82,24	1101,07	2,42	1,97

5.2.2 Calibração e avaliação de desempenho do modelo regional de curvas de permanência

Após a construção das curvas empíricas suavizadas, foi identificado o modelo regional utilizando a FMR. A Regressão Funcional Múltipla relaciona um conjunto de permanência pré-definidas, definidas no Tópico 4.3, às características físicas, climáticas e morfológicas das 36 bacias hidrográficas selecionadas para este fim. O processo para a definição das variáveis dependentes utilizadas no modelo regional compreendeu a inserção de cada característica descritiva, seguida da avaliação da avaliação comportamental do modelo mediante análise das suas métricas de desempenho.

O desempenho da utilização de uma única covariável explicativa na construção do modelo é apresentado na Tabela 7.

Tabela 7 – Índices de desempenho da FMR para uma variável dependente

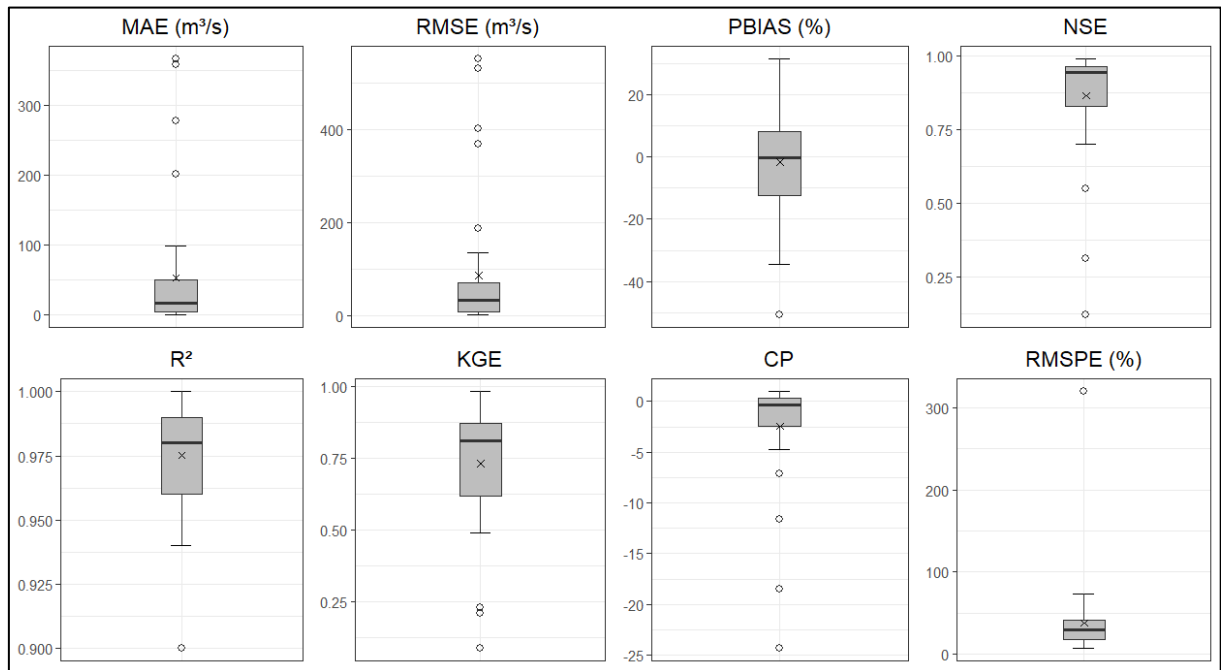
Variáveis		MAE	RMSE	PBIAS	NSE	R ²	KGE	CP	RMSPE
A	(Média)	57.37	91.02	10.54	-0.37	0.97	0.30	-34.46	74.62
	(Mediana)	28.70	47.33	-16.50	0.83	0.98	0.58	-0.93	34.62
L	(Média)	132.75	187.91	16.14	-0.60	0.97	0.17	-38.60	96.96
	(Mediana)	28.77	44.34	-18,60	0.71	0.97	0.46	-4.53	44.32
Q _s	(Média)	331,00	482.72	269.28	-75.48	0.96	-3.56	-997.09	551.56
	(Mediana)	72.25	110.40	-9,60	-0.38	0.97	-0.27	-78.32	95.27
P	(Média)	330.24	481.94	281,10	-95,80	0.96	-3.64	-1214.54	535.22
	(Mediana)	80.58	122.55	-6.15	-0.46	0.97	-0.27	-79.69	93,20
DD	(Média)	329.84	480.83	292.92	-105.44	0.96	-3.82	-1297.27	525.31
	(Mediana)	71.43	109.56	-14.10	-0.41	0.97	-0.27	-78.71	88.26
Kc	(Média)	326.69	475.86	268.16	-72.17	0.96	-3.48	-1244.32	579.78
	(Mediana)	75.44	118.07	-22.35	0.05	0.97	-0.07	-47.37	81.37

Nota-se que a área de drenagem é a variável explicativa que apresenta os melhores índices de desempenho, tanto em termos médios quanto medianos. O MAE, RMSE e RMSPE apresentam valores significativamente menores em relação às outras variáveis explicativas. O NSE apresentado pelo modelo que utiliza a área de drenagem é de 0,83, em termos medianos, sendo bem superior do que a maioria das covariáveis analisadas. A segunda variável explicativa que apresentou melhor desempenho foi o comprimento do talvegue principal. Entretanto, tal variável demonstrou-se inferior à área de drenagem em praticamente todos os índices analisados, portanto, essa última foi escolhida como a primeira variável explicativa a integrar o modelo regional.

Análises semelhantes foram realizadas para os modelos que compreendiam a utilização de duas, três e quatro variáveis explicativas. Tal procedimento se encontra disponível no Apêndice A deste estudo. O modelo que apresentou melhor desempenho foi definido com três variáveis explicativas, a saber: a área de drenagem, precipitação média anual e a densidade de drenagem. A inserção da produção do aquífero (Q_s), do comprimento do rio principal (L) e do coeficiente de compacidade (Kc) não representou ganhos significativos ao modelo, portanto, essas características foram descartadas na construção do modelo regional final.

A Figura 19 sintetiza a qualidade do ajuste do modelo regional calibrado para a 36 estações selecionadas para este fim.

Figura 19 - Métricas de desempenho do processo de calibração



Em relação às métricas o valor do MAE e do RMSE variaram entre 366,47 e 0,36 m³/s e 0,89 e 552,69 m³/s, respectivamente, e os valores mais elevados foram observados nas bacias que apresentavam as maiores áreas de drenagem. Em contrapartida, os maiores valores do RMSPE foram verificados nas bacias cujas áreas possuem dimensões menores do que a média do conjunto de regiões utilizadas no estudo. Por sua vez, o critério de eficiência de Nash-Sutcliffe (NSE) apresentou desempenho superior a 92% para 23 estações analisadas. As menores eficiências, medidas por tal critério, foram observadas nas estações Capitânea (45220000) e Tapiraí Jusante (40060001), que apresentaram os valores de 0,12 e 0,31, respectivamente.

Em termos medianos, a abordagem utilizada praticamente não indicou a presença de viés no modelo calibrado. Os valores obtidos para o coeficiente de determinação (R^2) foram elevados e acima de 95% para 30 bacias hidrográficas. Ainda considerando os valores medianos, o coeficiente de persistência (CP) apresentou o valor de -0,35 e o critério de eficiência de Kling-Gupta (KGE) se aproximou da unidade, valor tido como ideal para tal índice, e, assim como o NSE, os menores valores, 0,21 e 0,09, foram verificados nas estações Capitânea (45220000) e Tapiraí Jusante (40060001), respectivamente. De maneira geral, em termos médios, os resultados atingidos para a regionalização das FDCs se aproximaram dos obtidos por Requena *et al.* (2018b).

Nas Figuras 20 e 21 são apresentadas as FDCs regionais da etapa de calibração das Estações Pirapama (41600000) e Juvenília (45260000), respectivamente, em contraste com as curvas observadas. Os gráficos das demais estações podem ser visualizados no Apêndice B.

Figura 20 - Desempenho da regressão funcional múltipla para a estação Pirapama (41600000)

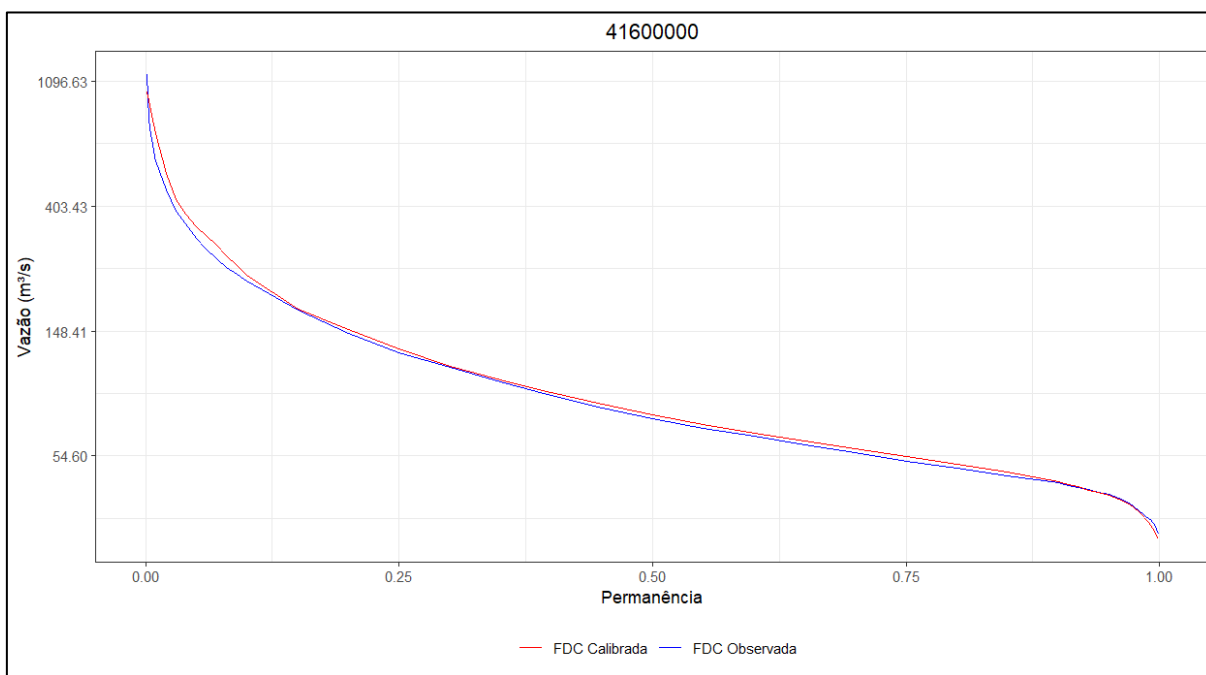
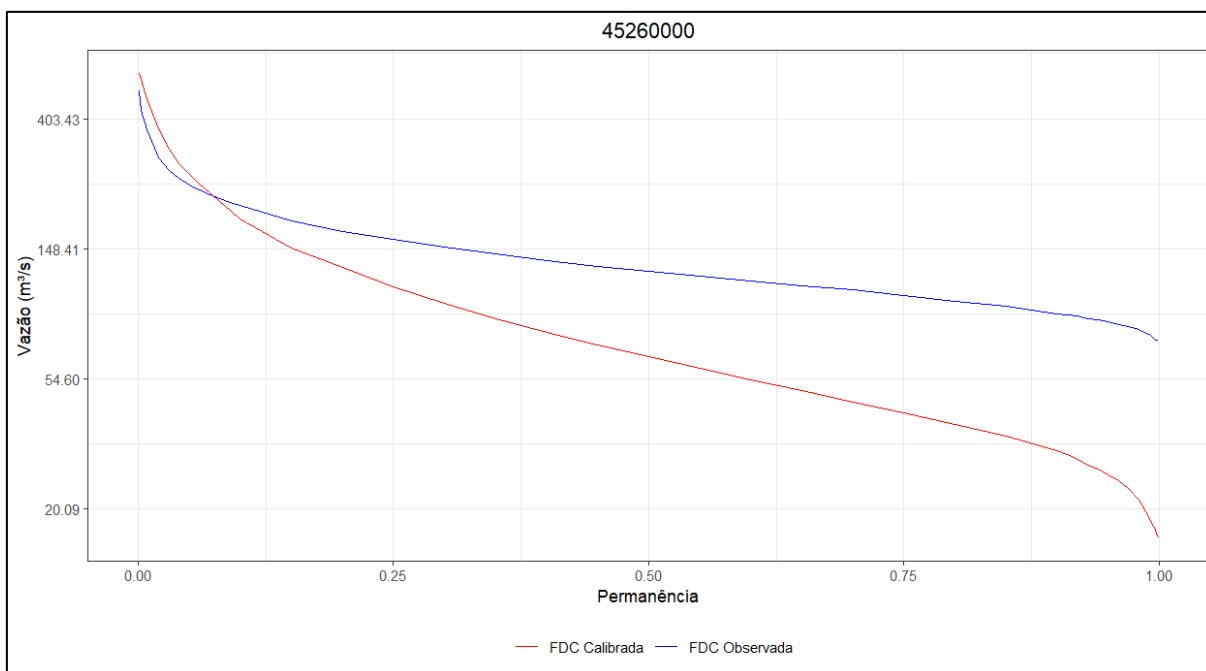


Figura 21 - Desempenho da regressão funcional múltipla para a estação Juvenília (45260000)



O modelo calibrado apresentou melhor desempenho para a estação Pirapama (41600000), conseguindo reproduzir de forma eficiente tanto o ramo inferior quanto o superior da curva de permanência empírica. Tal desempenho pode estar associado às características descritivas da estação em questão, visto que as mesmas apresentam valores próximos às médias das características físicas utilizadas no processo de calibração. Para efeito de comparação, as estações Porto das Andorinhas (40100000), Velho da Taipa (40330000) e Jequitibá (4141000) também apresentaram desempenhos superiores à média observada e, assim como observado para a estação Pirapama (41600000), as estações referidas também possuem características físicas semelhantes às médias observadas.

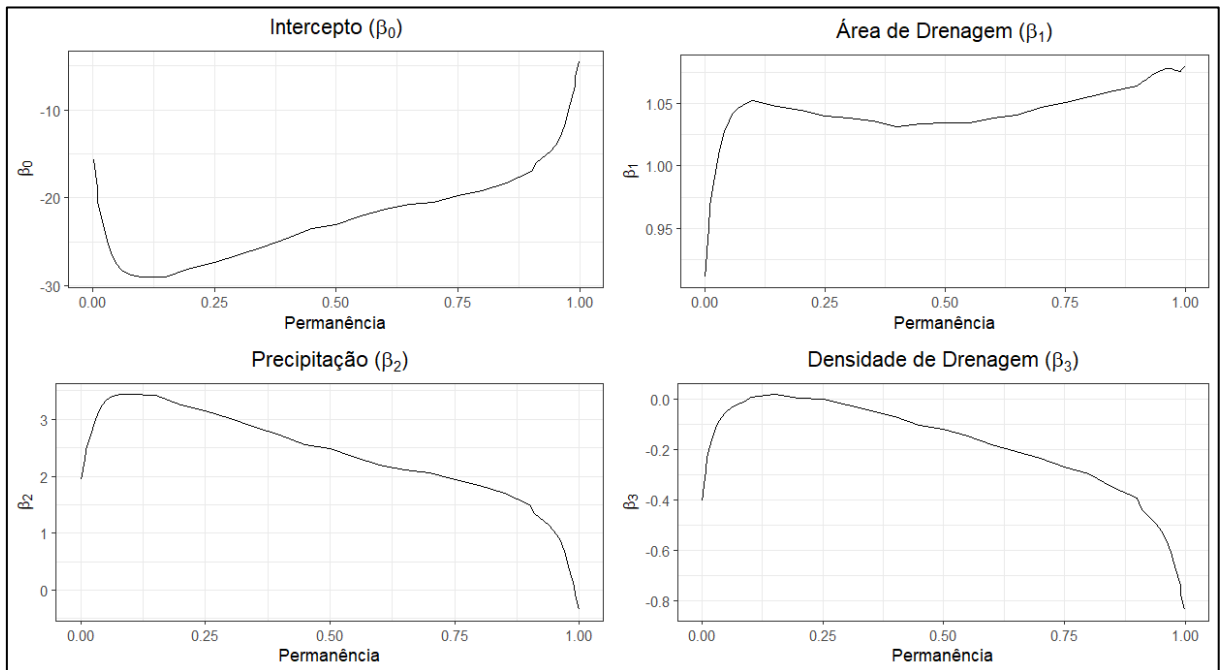
Situação oposta é constatada na estação Juvenília (45260000), na qual a curva de permanência calibrada não manifestou boa aderência à FDC empírica. Analisando as particularidades da bacia hidrográfica representada pela estação citada, percebe-se que a região apresenta a densidade de drenagem muito superior às outras bacias utilizadas no processo de calibração. Comportamentos semelhantes foram verificados nas estações Capitânea (45220000) e Ponte do Picão (41685000), que também apresentam densidades de drenagem elevadas. No entanto, a curva de permanência calibrada para a estação 41685000 superestimou as vazões médias diárias nos ramos médio e inferior da FDC de longo termo, situação oposta à observada para as estações 45260000 e 45220000. Ademais, a sub-bacia Carinhonha (MG/BA), onde encontram-se localizadas as estações Juvenília e Capitânea, drena o Aquífero Urucuia. Tal sistema apresenta elevada produtividade e contribui para cerca de 90% das vazões dos cursos de água onde estão situadas as estações 45260000 e 45220000 (VIEIRA, 2021). O aspecto citado pode ter contribuído para a subestimação das vazões médias diárias das curvas de permanência de longo termo das estações supracitadas.

5.2.3 Análise das variáveis dependentes do modelo regional

Conforme citado na etapa metodológica, uma das vantagens da utilização da regionalização pela regressão funcional múltipla reside no fato de poder avaliar o comportamento das características físicas, climáticas e morfológicas na construção da curva de permanência regional. Tendo definido as características descritivas relevantes à construção do modelo final, pode-se analisar a influência de cada variável

dependente por meio dos coeficientes funcionais que as acompanham, conforme descrito no Tópico 3.3.5 (Equação 30). O comportamento de tais coeficientes, associados às suas respectivas características descritivas, podem ser visualizados na Figura 22.

Figura 22 - Coeficientes da regressão funcional múltipla



A área de drenagem, retratada pelo coeficiente β_1 , exerce influência positiva e constante em praticamente toda a curva de permanência regional, estando em concordância com os resultados obtidos por Requena *et al.* (2018b). Conforme os autores, esse desempenho é coerente com a literatura, uma vez que a área de drenagem é tida como uma das características descritivas mais importantes para a caracterização da variabilidade do escoamento em bacias hidrográficas.

Avaliando a precipitação média anual, associada ao coeficiente β_2 , nota-se que essa variável exerce grande influência para as menores permanências e, consequentemente, para as vazões mais elevadas. Para permanência acima de 20%, a influência da precipitação decresce constantemente até um ponto próximo à permanência de 85% onde, a partir desse momento, a inclinação da curva torna-se mais acentuada. O resultado obtido é compreensível, visto que os eventos de precipitação contribuem mais intensamente para os eventos de cheia das bacias

hidrográficas, e menos para o período de recessão, nos quais o escoamento é definido principalmente pelo escoamento de base.

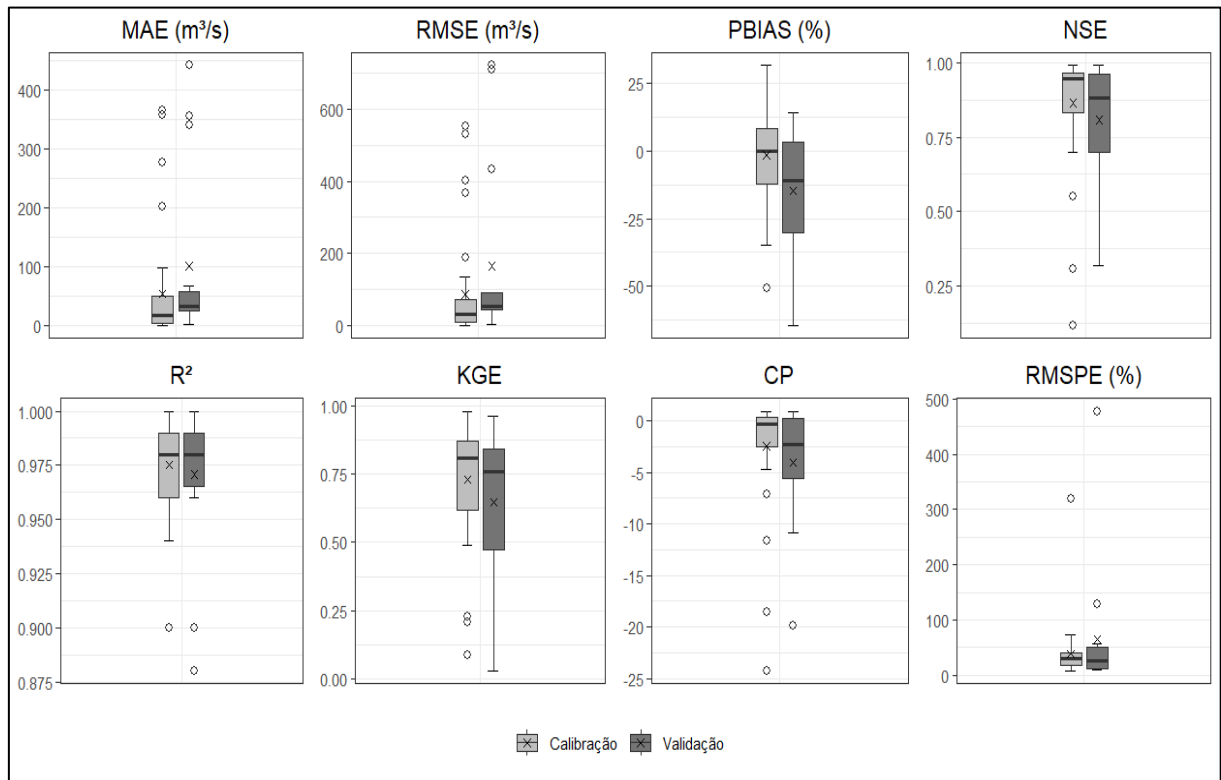
Diferentemente dos resultados observados para a área de drenagem e precipitação média anual, a densidade de drenagem exerceu influência negativa em praticamente todo o comprimento das FDCs calibradas. Inicialmente, a atuação dessa característica decaiu até a permanência de 10%, onde aparenta estabilizar em zero, exercendo influência nula no modelo regional. Na parte central da FDC, a partir da permanência de 25%, a atuação da densidade de drenagem aumenta forma linear e negativa ao longo da curva regional e, em oposição à precipitação média anual, a partir da permanência de 85%, a influência de tal variável torna-se mais acentuada. Esse comportamento auxilia na compreensão da aderência do modelo regional às curvas empíricas, posto que, conforme citado no tópico anterior, as estações que apresentaram os piores desempenhos dispunham das maiores densidades de drenagem das bacias hidrográficas avaliadas. Logo, para as referidas estações a curva calibrada subestimou as vazões medidas para praticamente todo o comprimento da FDC.

5.2.4 Análise de desempenho do processo de validação do modelo regional

A etapa de validação do modelo consiste em avaliar a capacidade preditiva do modelo. Para tanto, a predição das curvas de permanência regionais foi realizada para as 15 estações fluviométricas destinadas a esse fim. As características descritivas dessas estações foram utilizadas na equação regional obtida pela regressão funcional múltipla para que, então, fossem derivadas as curvas regionais das estações de validação. Os resultados obtidos foram comparados das FDCs empíricas, construídas empregando o mesmo período de 30 anos de medições utilizados na etapa de calibração.

A Figura 23 condensa os índices de desempenho obtidos para as 15 estações utilizadas no processo de validação.

Figura 23 - Métricas de desempenho do processo de validação



O desempenho do modelo regional é ligeiramente inferior ao observado para as estações utilizadas na etapa de calibração. Em termos de valores médios, o MAE e o RMSE apresentaram valores mais elevados, variando entre 1,28 e 442,18 m³/s e 3,50 e 721,34 m³/s, respectivamente. Assim como na etapa de calibração, os maiores valores foram verificados para as estações que apresentaram as maiores áreas de drenagem. O índice RMSPE apresentou desempenho mediano ligeiramente superior ao processo de calibração, evoluindo de 30,03 para 26,58, ainda que a estação Ponte do Bicudo (41940000) tenha apresentado resultado consideravelmente superior às demais estações (478,10).

Em relação ao critério de eficiência de Nash-Sutcliffe (NSE), 7 estações apresentaram desempenho superior a 92%. O pior desempenho foi obtido pela estação de Curimataí (41890000), onde o modelo regional alcançou apenas 32% para a referida métrica. Diferentemente da etapa de calibração, o modelo regional empregado no processo de validação indicou a presença de viés negativo nos resultados atingidos, apontando que o modelo tende a subestimar as vazões das FDCs empíricas. O coeficiente de determinação (R^2) permaneceu elevado, tendo registrado valores superiores a 95%

para 13 estações de validação. Por fim, tanto o coeficiente de persistência (CP) quanto o critério de eficiência de Kling-Gupta (KGE) apresentaram desempenhos ligeiramente inferiores à etapa de calibração.

As Figuras 24 e 25 apresentam a aderência do modelo regional às curvas empíricas das estações Jaguaruna Jusante (40300001) e Lagoa das Pedras (45210000), respectivamente. O comportamento das demais FDCs regionais da etapa de validação é apresentado no Apêndice C.

Figura 24 - Desempenho da regressão funcional múltipla para a estação Jaguaruna Jusante (40300001)

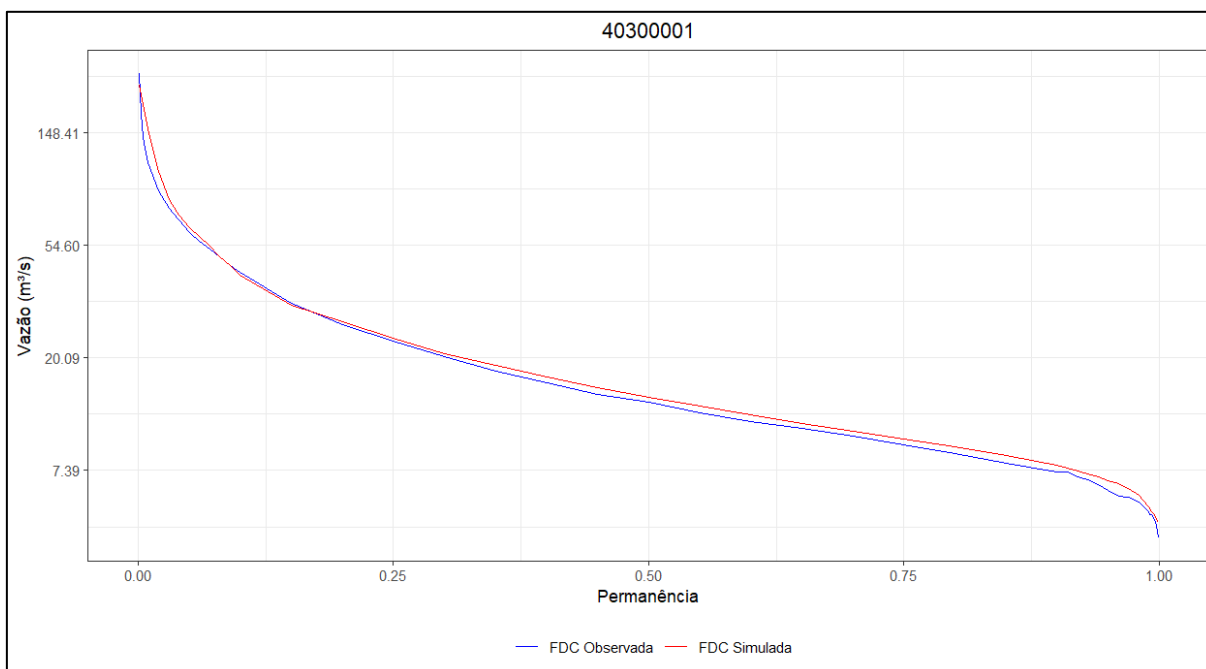
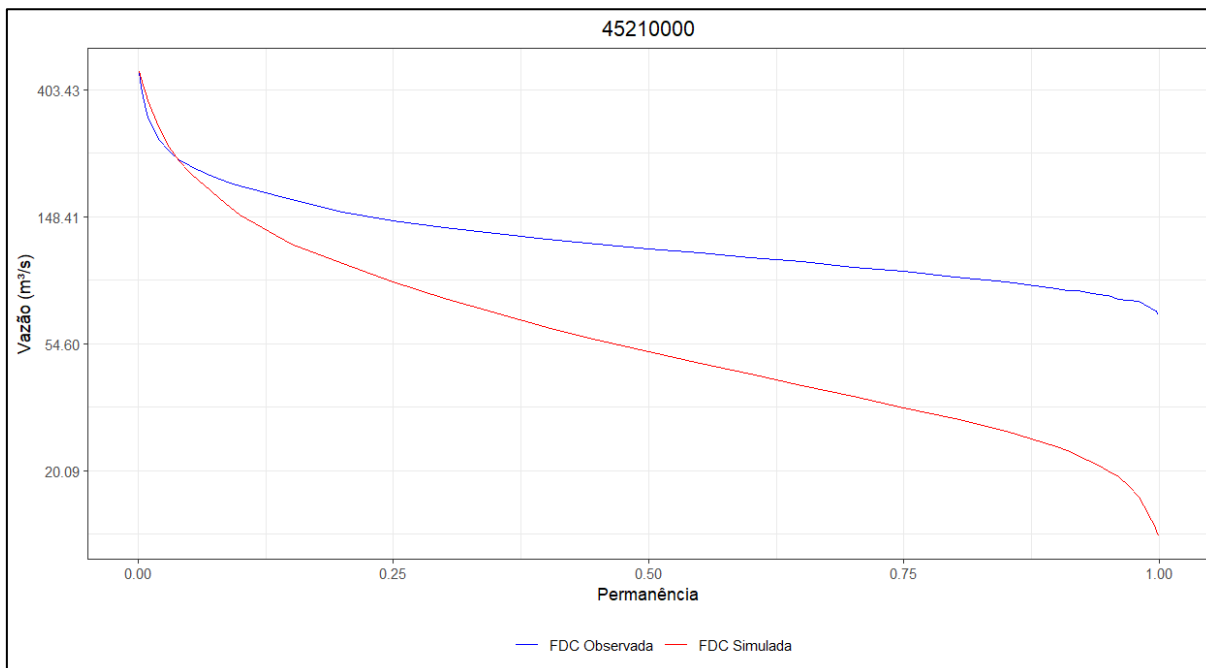


Figura 25 - Desempenho da regressão funcional múltipla para a estação Lagoa das Pedras (45210000)



Analisando as Figuras 24 e 25, percebe-se que o modelo apresenta desempenhos distintos para o conjunto de estações analisadas. A melhor performance foi atingida na estação Jaguaruna Jusante (40300001), onde o modelo conseguiu reproduzir cada ramo da curva de permanência empírica. Assim como descrito na etapa de calibração, tal desempenho pode estar relacionado às características físicas da bacia hidrográfica analisada, dado que a mesma apresenta atributos semelhantes à média utilizada na construção do modelo.

A aderência do modelo regional na estação Lagoa das Pedras (45210000) não foi adequada, isso é, o modelo não conseguiu reproduzir o comportamento da FDC empírica, subestimando praticamente todas as vazões medidas. Comportamento semelhante foi averiguado para a estação São Gonçalo (45131000), localizada próxima à última estação citada. Analisando as características descritivas das bacias hidrográficas abrangidas por tais estações, nota-se que ambas apresentam elevados valores para o atributo densidade de drenagem. Conforme citado anteriormente, tal atributo contribui negativamente para a variação das vazões no modelo regional desse estudo, o que pode ter contribuído para a subestimação das vazões médias diárias das FDCs das estações referidas.

Além do aspecto mencionado, cabe ressaltar que assim como as estações Juvenília (45260000) e Capitânea (45220000), as estações Lagoa das Pedras (45210000) e São Gonçalo (45131000) também estão localizadas na sub-bacia Carinhonha (MG/BA), sobre o Sistema Aquífero Urucuia. Tal resultado expõe a limitação do modelo empregado em reproduzir a variação das FDCs empíricas de bacias hidrográficas localizadas sobre um Sistema Aquífero tão produtivo. Ademais, o modelo apresentou limitações na reprodução da cauda inferior das curvas de permanência. Entretanto, outros modelos de regionalização também apresentaram tal limitação (por exemplo, CASTELLARIN *et al.*, 2018; COSTA, 2011; PUGLIESE *et al.*, 2016).

Tendo em vista os aspectos citados, a inclusão de algumas variáveis explicativas associadas ao escoamento de base no modelo regional, bem como a delimitação de regiões homogêneas (REQUENA *et al.*, 2018b), poderiam aprimorar a capacidade preditiva da abordagem proposta, visto que os melhores desempenhos foram observados em bacias com características próximas aos valores médios utilizados. Além disso, considerando a limitação do modelo em reproduzir o ramo inferior das curvas de permanência de longo termo, a construção da FMR em seções podia refinar os resultados obtidos.

5.3 GERAÇÃO DAS SÉRIES SINTÉTICAS DE VAZÕES DIÁRIAS

5.3.1 Aplicação da interpolação espacial via Krigagem Topológica

As séries sintéticas de vazões médias diárias foram obtidas mediante aplicação da metodologia proposta por Hughes e Smakhtin (1996). A obtenção dos pesos utilizados na ponderação das vazões diárias foi realizada utilizando a interpolação espacial não linear via Krigagem Topológica. Para esse fim, foram utilizadas as vazões médias de longo termo (Q_{MLT}) das 36 estações de calibração para estimar as Q_{MLTs} das 15 estações destinada à validação. Tal procedimento considerou o emprego de 3, 4, 5 e 6 bacias hidrográficas vizinhas no processo de interpolação.

Os resultados estimados das vazões médias de longo termo assim como os respectivos erros percentuais absolutos médios (MAPE) estão exibidos nas Tabelas 8 e 9, respectivamente.

Tabela 8 - Q_{MLT} observada versus Q_{MLT} estimada utilizando 3, 4, 5 e 6 estações vizinhas na interpolação espacial

Estações	Q_{MLT} Observada	Q_{MLT} Estimada 3 Estações	Q_{MLT} Estimada 4 Estações	Q_{MLT} Estimada 5 Estações	Q_{MLT} Estimada 6 Estações
40300001	21,00	22,80	22,65	22,35	22,31
41050000	20,47	13,56	13,75	15,23	15,08
41135000	874,48	720,46	815,20	802,64	817,80
41440005	3,10	3,22	3,13	3,31	3,24
41650002	133,85	142,23	151,43	147,19	155,61
41818000	205,50	235,94	243,24	235,54	228,94
41890000	18,66	16,55	16,26	17,39	16,25
41940000	19,00	18,37	18,03	20,45	21,68
42395000	170,03	162,73	166,05	162,47	160,59
42850000	55,39	50,32	50,39	50,59	51,98
43880000	245,66	282,11	289,12	285,97	285,33
44200000	1998,13	2000,27	2006,77	2013,94	2020,41
44500000	2013,43	2217,64	2235,11	2254,55	2269,70
45131000	67,76	49,14	52,11	54,38	54,90
45210000	128,87	105,29	108,53	112,61	114,00

Tabela 9 - Erros percentuais absolutos médios do desempenho do Top-Kriging

Estações	MAPE (%)			
	3 Estações	4 Estações	5 Estações	6 Estações
40300001	8,545	7,845	6,434	6,237
41050000	33,777	32,825	25,626	26,362
41135000	17,613	6,778	8,215	6,481
41440005	3,727	1,049	6,816	4,400
41650002	6,265	13,137	9,972	16,258
41818000	14,813	18,369	14,621	11,407
41890000	11,302	12,876	6,782	12,939
41940000	3,335	5,095	7,596	14,082
42395000	4,289	2,339	4,443	5,549
42850000	9,142	9,023	8,654	6,156
43880000	14,837	17,691	16,409	16,147
44200000	0,107	0,432	0,791	1,115
44500000	10,142	11,010	11,975	12,728
45131000	27,478	23,088	19,742	18,978
45210000	18,296	15,785	12,613	11,535
Média	12,244	11,823	10,713	11,358
Desvio Padrão	9,264	8,842	6,349	6,580

A análise da Tabela 9 demonstra que o desempenho dos valores estimados oscila para uma mesma estação à medida que mais estações vizinhas são utilizadas na Krigagem Topológica. No entanto, em termos médios, os resultados do MAPE apresentaram uma ligeira redução. Também pode-se observar que a utilização de mais do que 5 estações vizinhas no processo de interpolação espacial não ocasionou melhora significativa do desempenho do modelo, conforme avaliado pelo erro percentual absoluto médio. Esse resultado está de acordo com a literatura, visto que alguns autores sugerem a utilização de 3 a 5 estações vizinhas nessa etapa da modelagem (por exemplo, HUGHES E SMAKHTIN, 1996; REQUENA *et al*, 2017; SHU E OUARDA, 2012).

Em relação ao desempenho individual de cada estação, pode-se inferir que o Top-Kriging demonstrou melhor desempenho para as estações fluviométricas localizadas no curso do rio principal da bacia hidrográfica analisada. A estação Major Porto (41050000) apresentou a pior desempenho do conjunto avaliado. Tal resultado pode estar condicionado à ausência de uma densa rede de estações monitoradas ao seu entorno. Apesar disso, a performance da Krigagem Topológica foi avaliada de forma positiva e adequada aos fins desse estudo, uma vez que, em termos médios, os resultados de predição da Q_{MLT} se aproximaram dos valores observados. Os resultados atingidos corroboram com os obtidos em trabalhos anteriores (por exemplo, ARCHFIELD *et al.*, 2013; PUGLIESE *et al.*, 2014) e, dessa forma, indica o Top-Kriging como uma ferramenta a ser considerada em estudos de predição hidrológica.

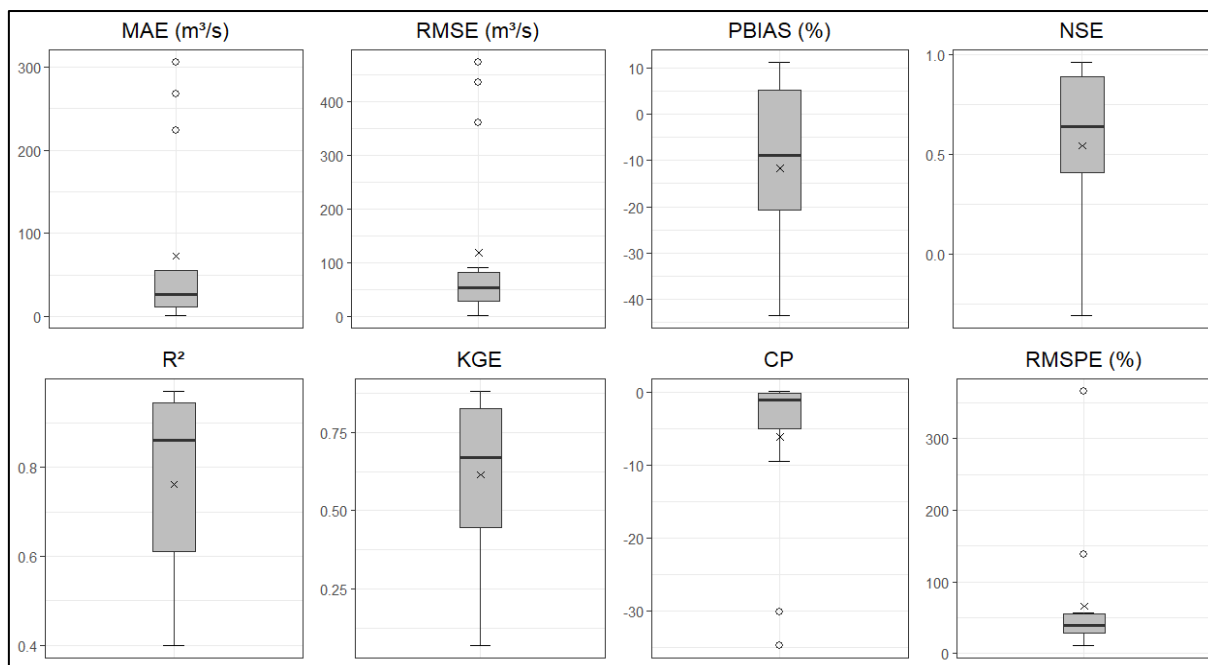
Considerando os aspectos citados, optou-se pela utilização de apenas 5 estações monitoradas no processo de interpolação espacial não linear realizado pelo Top-Kriging. Os pesos utilizados como ponderação das vazões diárias estão descritos no Apêndice D.

5.3.2 Análise das séries sintéticas de vazões diárias estimadas

Obtidas curvas de permanências regionais e os pesos utilizados na ponderação das vazões médias diárias, prosseguiu-se para a geração das séries sintéticas de vazões, utilizando a metodologia proposta por Hughes e Smakhtin (1996). O método proposto foi empregado nas mesmas 15 estações de validação utilizadas na regionalização das curvas de permanências de longo termo.

A Figura 26 expressa as métricas de desempenho do método citado para as estações consideradas sem monitoramento.

Figura 26 - Métricas de desempenho das séries de vazões simuladas



O índice MAE variou entre 1,01 (estação 41440005) e 268,15 m³/s (estação 44500000). Considerando os índices RMSE e RMSPE, os melhores desempenhos foram apresentados pelas estações 41440005 (2,14 m³/s) e 44200000 (11,79), respectivamente. Os piores desempenhos para as referidas métricas foram reservados às estações 44500000 (474,12 m³/s) e 41940000 (365,51). Assim como observado na etapa de validação das FDCs regionais, as séries sintéticas indicaram a presença de viés negativo nos resultados atingidos.

No que concerne o coeficiente de eficiência de Nash-Sutcliffe (NSE), apenas 6 estações apresentaram resultados acima de 75%. Os piores desempenho foram registrados nas estações São Gonçalo (45131000) e Lagoa das Pedras (45210000) com valores próximos de -23% e -31%, respectivamente. Os valores do coeficiente de determinação (R^2) foram superiores a 90% em apenas 7 das 15 estações utilizadas. Os piores resultados para tal métrica foram observados nas bacias hidrográficas com as menores áreas de drenagem. O coeficiente de persistência (CP) e o critério de eficiência de Kling-Gupta (KGE) apresentaram desempenhos equiparáveis à etapa de regionalização das curvas de permanência de longo termo.

As Figuras 27 e 28 apresentam as séries de vazões diárias simuladas e observadas para as estações fluviométricas Jaguaruna Jusante (40300001) e Lagoa das Pedras (45210000), respectivamente. As séries sintéticas das demais estações encontram-se disponíveis no Apêndice E.

Figura 27 - Desempenho da série de vazões simuladas na estação Jaguaruna Jusante (40300001)

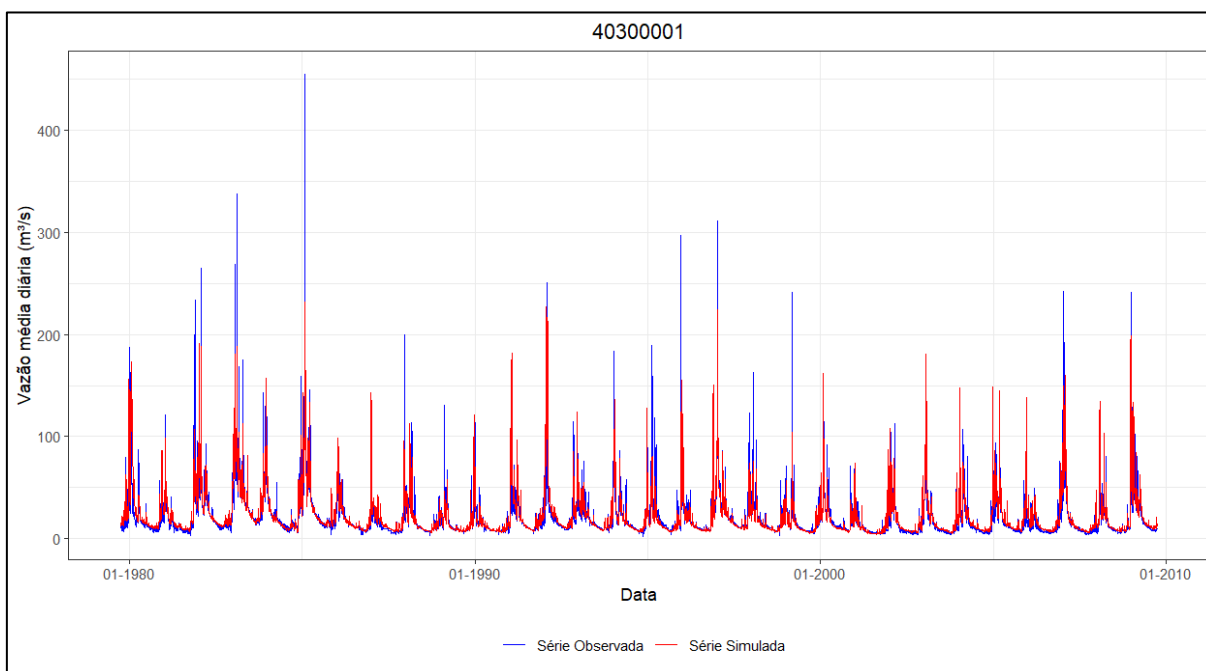
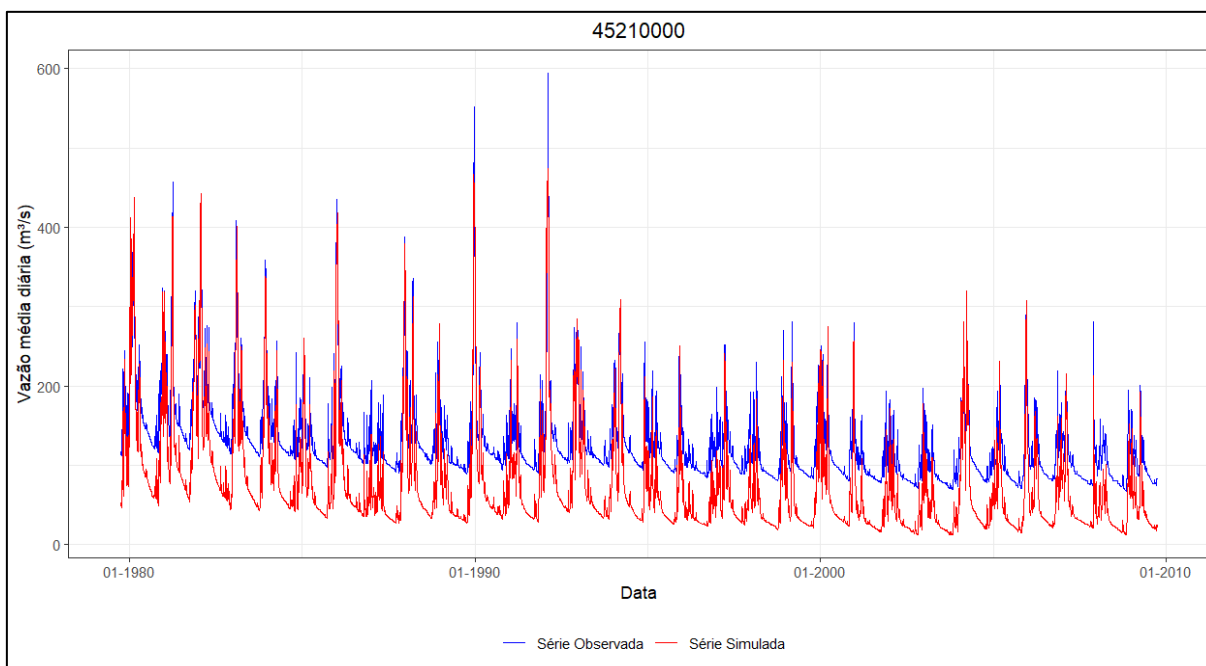


Figura 28 - Desempenho da série de vazões simuladas na estação Lagoa das Pedras (45210000)



Como pode ser visto nas Figuras 27 e 28, o modelo apresentou capacidade preditiva variável. Para a estação Jaguaruna Jusante (40300001), o modelo apresentou bom desempenho, conseguindo reproduzir a variabilidade das vazões em quase sua totalidade, com exceção de algumas vazões de pico. O mesmo não pode ser dito para a estação Lagoa das Pedras (45210000), pois a mesma exibiu um dos piores desempenhos do modelo regional. Com efeito, o modelo não conseguiu reproduzir a série de vazões observadas, com exceção de algumas vazões de pico, subestimando a maior parte dos valores medidos. Comportamento semelhante foi observado para a estação São Gonçalo (45131000), o que já era esperado, visto que ambas estações apresentaram desempenhos similares, no que se refere à FDC regional. Além disso, os postos citados apresentam características semelhantes e estão inseridas em locais próximos, com condições climáticas semelhantes e sobre o Sistema Aquífero Urucuia, o que pode ter contribuído para os resultados observados.

Com base nas séries sintéticas de vazões, bem como nos valores dos índices de desempenho do modelo, é notório que a qualidade do ajuste das séries simuladas está diretamente relacionada com o desempenho do modelo regional das FDCs, isso é, os melhores resultados das séries de vazões médias diárias foram obtidos nas estações em que a curva de permanência regional melhor se aproximou da curva de permanência empírica. Dessa forma, como a estação Jaguaruna Jusante (40300001) apresentou um dos melhores desempenhos no que se refere ao ajuste da FDC regional, a qualidade do ajuste da sua respectiva série simulada também foi elevada. Esses resultados corroboram aqueles obtidos por Requena *et al.* (2018b), dado que os melhores desempenhos obtidos pelos autores tanto para as FDCs regionais quanto para as séries simuladas estão diretamente relacionados.

5.3.3 Análise das vazões de interesse prático estimadas

Por fim, procurou-se avaliar a capacidade do modelo em estimar algumas vazões de interesse prático, utilizando como base as séries de vazões simuladas obtidas no tópico anterior. Neste estudo, foram avaliadas as vazões médias mínimas com 7 dias de duração (Q_7), vazões médias mensais e vazões máximas anuais.

A Tabela 10 condensa os resultados obtidos para o erro percentual absoluto médio (MAPE) das 15 estações fluviométricas destinadas à validação dos dados.

Tabela 10 - Erros percentuais absolutos médios para as vazões Q_7 , médias mensais e máximas anuais

Estações	Q_7	MAPE (%)	
		Médias Mensais	Máximas Anuais
40300001	17,301	6,564	28,145
41050000	28,078	12,941	71,503
41135000	39,855	27,995	28,672
41440005	30,084	23,427	31,061
41650002	21,200	7,773	17,213
41818000	47,158	14,973	17,206
41890000	239,009	71,039	79,414
41940000	953,765	60,542	66,468
42395000	11,806	5,748	16,066
42850000	78,998	21,246	67,765
43880000	66,349	18,262	21,329
44200000	5,579	7,397	9,396
44500000	5,766	10,032	18,618
45131000	71,211	46,951	15,885
45210000	70,458	45,242	9,527
Média	112,441	25,342	33,218

É possível perceber que, em geral, as estações fluviométricas que obtiveram os melhores desempenhos nas etapas anteriores apresentaram as melhores estimativas das vazões analisadas. Para as vazões médias mínimas de 7 dias de duração (Q_7), os melhores desempenhos foram observados para as estações São Francisco (44200000) e Manga (44500000), que possuem as maiores áreas de drenagem e as vazões médias mais elevadas. Em contrapartida, os piores resultados foram destinados às estações de Curimataí (41890000) e Ponte do Bicudo (41940000) que, ao contrário das estações São Francisco (44200000) e Manga (44500000), apresentam áreas e vazões médias reduzidas, em comparação com os valores médios das regiões utilizadas. Dessa forma, é possível concluir que além da qualidade do ajuste das séries simuladas, as características descritivas citadas também contribuem para a eficiência da predição de vazões mínimas.

Em relação às vazões médias mensais, a discrepância entre as vazões simuladas e observadas foi atenuada. No entanto, os melhores desempenhos se restringiram às estações que apresentaram os melhores resultados para a regionalização das vazões médias mínimas de 7 dias de duração. Por fim, no que concerne às vazões máximas

anuais, o modelo apresentou desempenho mediano, e o melhor desempenho foi, mais uma vez, registrado pela estação São Francisco (44200000), mas inferior às vazões mínimas e médias mensais. Tal resultado pode estar relacionada com a limitação do modelo regional em reproduzir as vazões de pico, situação observada em praticamente toda a região de estudo. É importante observar que, apesar das séries simuladas das estações São Gonçalo (45131000) e Lagoa das Pedras (45210000) terem subestimado a maior parte das vazões observadas, o desempenho das referidas estações na predição de máximos anuais foi elevado.

6 CONCLUSÕES

A presente dissertação teve como objetivo a elaboração de um método que possibilitasse a estimação de séries de vazão em regiões sem monitoramento fluviométrico, utilizando em conjunto curvas de permanência de longo termo (FDC) regionalizadas e princípios de interpolação espacial não linear. O trabalho dividiu-se em duas etapas principais, sendo a primeira destinada à regionalização das curvas citadas e a segunda reservada à geração das séries sintéticas de vazões diárias.

O método proposto foi implementado na porção mineira da bacia hidrográfica do Rio São Francisco. Essa escolha se deu devido à disponibilidade de dados fluviométricos medidos para que fosse possível desenvolver a regionalização hidrológica das curvas de permanência de longo termo de maneira eficiente, uma vez que concepção de tais curvas carece de uma quantidade suficiente de vazões observadas na representação do comportamento hidrológico das bacias avaliadas. Outro importante ponto foi a intenção de avaliar o comportamento do modelo em regiões com características físicas e climáticas distintas, dado que a utilização do modelo em questão não prevê a delimitação de regiões homogêneas.

A regionalização das curvas de permanência considerou a técnica utilizada por Requena *et al.* (2018b), denominada Regressão Funcional Múltipla (FMR), que consiste em regionalizar as FDCs utilizando os princípios da Análise Funcional de Dados (FDA), relacionando um conjunto de quantis de permanências predeterminadas às características físicas, climáticas e morfológicas da região de estudo. Tal abordagem tem como fundamento expressar os dados pontuais de uma série de medições como funções, concebidas por uma soma de funções de base conhecidas. Dessa forma, as curvas de permanência empíricas são tratadas como uma única função e os seus respectivos quantis podem ser estimados mediante uma única regressão, otimizando todo o processo de regionalização.

A abordagem regional FMR foi aplicada a 51 estações fluviométricas, obtidas do trabalho de Regionalização de vazões mínimas sazonais do Estado de Minas Gerais (EHR/2021), das quais 36 foram utilizadas na calibração e 15 foram destinadas à validação do modelo proposto. As curvas de permanência empíricas foram construídas para um período de 30 anos de observações (outubro de 1979 a setembro

de 2009) e suas respectivas curvas funcionais foram definidas para um conjunto de 45 permanências utilizando o sistema de bases que mais se adequou às características funcionais das FDCs, as funções do tipo B-Splines.

Em seguida foram avaliadas as características descritivas das bacias hidrográficas que apresentaram a melhor correlação aos quantis de permanências das FDCs funcionais, acrescentando cada variável dependente, uma por vez, ao modelo calibrado. O modelo final compreendeu a utilização de três variáveis dependentes, sendo a área de drenagem, precipitação média anual e densidade de drenagem. O modelo calibrado apresentou valores medianos reduzidos para os índices MAE (17,57) RMSE (32,57) e RMSPE (30,03) e os valores mais elevados para tais métricas foram registrados nas bacias que apresentavam as maiores áreas de drenagem. Ademais, o viés mediano foi praticamente zero e o critério de eficiência de Nash-Sutcliffe (NSE) foi superior à 92% em 23 das bacias analisadas. As métricas obtidas no respectivo estudo se assemelharam em grande parte às alcançadas por Requena *et al.* (2018b). Em geral, o modelo apresentou limitações na regionalização das FDCs de bacias hidrográficas com características distintas aos atributos médios observados, principalmente as que apresentam elevados valores para a densidade de drenagem, situação verificada pelo mal desempenho obtido nas estações localizadas sobre o Sistema Aquífero Urucuia.

A utilização da FMR permitiu avaliar a influência das variáveis dependentes na construção do modelo calibrado. A área de drenagem apresentou influência positiva de constante ao longo de toda curva, atestando a importância dessa característica nos processos de regionalização. A precipitação média anual demonstrou grande influência para as permanências menores, reduzindo sua atuação à medida que as mesmas aumentavam. Por fim, a densidade de drenagem atuou de forma a reduzir as vazões estimadas ao longo de praticamente todo o comprimento das FDCs estimadas.

A validação do modelo regional obteve resultados ligeiramente inferiores aos registrados na etapa de calibração. Nessa etapa, o modelo indicou a presença de viés negativo, apontando que o modelo tende a subestimar as vazões observadas. Além disso, apenas 7 estações apresentaram desempenho superior à 92% para o NSE.

Essa fase confirmou que os melhores desempenhos se resumiram às estações que apresentavam características descritivas semelhantes às médias observadas. Os resultados obtidos também indicaram que a inclusão de novas características relacionadas ao escoamento de base no modelo regional, bem como a construção da FMR em segmentos, poderia aprimorar a capacidade preditiva da abordagem proposta, tendo em vista o mal desempenho do modelo em reproduzir o ramo inferior das FDCs. Ademais, a definição de regiões homogêneas poderia refinar os resultados alcançados. De forma geral, o emprego da FMR na regionalização de curvas de permanência apresentou resultados promissores.

A etapa seguinte do projeto consistiu na geração das séries sintéticas de vazões das estações utilizadas na validação do modelo. Para tanto, foi utilizada a metodologia proposta por Hughes e Smakhtin (1996). Os pesos empregados na ponderação das vazões diárias foram obtidos por meio da Krigagem Topológica. Tal método mostrou-se uma ferramenta eficiente no que se refere à interpolação espacial de características hidrológicas. Os resultados obtidos para a interpolação das vazões médias de longo termo se aproximaram dos valores observados, e os melhores resultados foram averiguados nas estações localizadas no curso d'água principal da bacia hidrográfica do Rio São Francisco.

Os pesos de interpolação foram extraídos do Top-Kriging e utilizados na obtenção das séries simuladas de vazões. Alguns índices avaliados foram equiparáveis à etapa de regionalização das FDCs, entretanto, apenas 6 estações fluviométricas apresentaram NSE maiores do que 75%. De maneira geral, os melhores desempenhos foram registrados nas estações em que a curva de permanência regional aderiu de forma eficiente à FDC empírica, constatando a estrita relação entre a qualidade do ajuste da curva regional e a série de vazões simuladas. Ademais, o modelo apresentou limitações em reproduzir as vazões de pico dos hidrogramas observados.

A última etapa do estudo consistiu em avaliar a capacidade do modelo em reproduzir algumas vazões de interesse prático. Foi constatado que o modelo proposto demonstra melhor desempenho na predição de vazões médias mensais e os melhores resultados foram alcançados pelas estações que apresentaram as maiores áreas de drenagem e maiores vazões médias observadas.

7 RECOMENDAÇÕES

Com base nas conclusões apresentadas, pode-se concluir que os objetivos definidos na presente dissertação foram alcançados. No entanto, a fim de aprimorar os resultados obtidos, as seguintes recomendações são sugeridas para estudos futuros:

- Avaliar a utilização do modelo em regiões homogêneas delimitadas. Apesar da aplicação do método proposto não carecer de delimitação de regiões homogêneas, tal procedimento pode fornecer resultados mais eficientes, visto que o modelo apresentou melhor desempenho em estações com características similares aos valores médios;
- Avaliar a adição de novas estações fluviométricas, sobretudo nas sub-bacias onde foram utilizadas poucas redes de monitoramento fluviométrico, sob a perspectiva da delimitação de regiões homogêneas;
- Empregar outras características físicas, climáticas e morfológicas na construção das curvas de permanência de longo termo regionais, bem como utilizar um novo conjunto de permanência na concepção da FMR; e
- Considerar a construção da Regressão Funcional Múltipla em segmentos, ou seja, avaliar o emprego de características físicas distintas na concepção do ramo superior, intermediário e inferior das FDCs regionais, garantindo a continuidade da função na interseção dos trechos.

REFERÊNCIAS

- ALMEIDA, Arlan S. et al. Regionalização de curvas de permanência no estado do paran . Simp sio Brasileiro de Recursos H dricos, v. 24, p. 1–10, 2021.
- ANA – AG NCIA NACIONAL DE  GUAS. Sistemas aqu feros. 2013. Dispon vel em: <https://metadados.snirh.gov.br/geonetwork/srv/api/records/3ec60e4f-85ea-4ba7-a90c-734b57594f90>
- ANA – AG NCIA NACIONAL DE  GUAS. Sub-bacias hidrogr ficas de n vel 3. 2015. Dispon vel em: <https://siga.cbhsaofrancisco.org.br/web-plan.html>
- ARCHFIELD, Stacey A. et al. Topological and canonical kriging for design flood prediction in ungauged catchments: an improvement over a traditional regional regression approach? **Hydrology and Earth System Sciences**, v. 17, n. 4, p. 1575-1588, 2013.
- BL SCHL, G nter; SIVAPALAN, Murugesu. Scale issues in hydrological modelling: a review. **Hydrological processes**, v. 9, n. 3-4, p. 251-290, 1995.
- BROWN, Alice E.; WESTERN, Andrew W.; MCMAHON, Thomas A.; ZHANG, Lu. Impact of forest cover changes on annual streamflow and flow duration curves. **Journal of Hydrology**, v. 483, p. 39-50, 2013
- BURN, Donald H.; BOORMAN, David B. Estimation of hydrological parameters at ungauged catchments. **Journal of Hydrology**, v. 143, n. 3-4, p. 429-454, 1993.
- CASTELLARIN, Attilio; GALEATI, Giorgio; BRANDIMARTE, Luigia; MONTANARI, Alberto; BRATH, Armando. Regional flow-duration curves: reliability for ungauged basins. **Advances In Water Resources**, v. 27, n. 10, p. 953-965, 2004.
- CASTELLARIN, Attilio; CAMORANI, Giorgio; BRATH, Armando. Predicting annual and long-term flow-duration curves in ungauged basins. **Advances in Water Resources**, v. 30, n. 4, p. 937-953, 2007.
- CASTELLARIN, Attilio et al. Prediction of flow duration curves in ungauged basins. **Runoff prediction in ungauged basins: Synthesis across processes, places and scales**, p. 135-162, 2013.
- CASTELLARIN, A.; PERSIANO, S.; PUGLIESE, A.; ALOE, A.; SK IEN, J. O.; PISTOCCHI, A. Prediction of streamflow regimes over large geographical areas: interpolated flow–duration curves for the Danube region. **Hydrological Sciences Journal**, v. 63, n. 6, p. 845-861, 2018.
- CASTIGLIONI, S.; CASTELLARIN, A.; MONTANARI, A.; SK IEN, J. O.; LAAHA, G.; BL SCHL, G. Smooth regional estimation of low-flow indices: physiographical space based interpolation and top-kriging. **Hydrology And Earth System Sciences**, v. 15, n. 3, p. 715-727, 4 mar. 2011. Copernicus GmbH.
- CBHSF – COMIT  DA BACIA HIDROGR FICA DO RIO S O FRANCISCO. Resumo executivo do plano de recursos h dricos da bacia hidrogr fica do rio S o

Francisco. Alagoas, 2016. Disponível em: <https://cbhsaofrancisco.org.br/plano-de-recursos-hidricos-da-bacia-hidrografica-do-rio-sao-francisco/>

CHEBANA, Fateh; DABO-NIANG, Sophie; OUARDA, Taha BMJ. Exploratory functional flood frequency analysis and outlier detection. **Water Resources Research**, v. 48, n. 4, 2012.

CLEVELAND, William S. Robust locally weighted regression and smoothing scatterplots. **Journal of the American statistical association**, v. 74, n. 368, p. 829-836, 1979.

COSTA, V. A. F. Modelos Regionais de Curvas de Permanência de Vazões para Rios Perenes, Intermitentes e Efêmeros, e seu Emprego na Calibração Indireta de Modelos de Simulação Hidrológica. Universidade Federal de Minas Gerais, 2011.

COSTA, V. A. F.; FERNANDES, W. S.; NAGHETTINI, M. Modelos regionais de curvas de permanência de vazões e sua utilização na calibração de modelos de simulação hidrológica. XIX Simpósio Brasileiro de Recursos Hídricos, 2011.

COSTA, Veber; FERNANDES, Wilson; NAGHETTINI, Mauro. Regional models of flow-duration curves of perennial and intermittent streams and their use for calibrating the parameters of a rainfall-runoff model. **Hydrological Sciences Journal**, v. 59, n. 2, p. 262-277, 2014.

COSTA, Veber; FERNANDES, Wilson; STARICK, Ângelo. Identifying regional models for flow duration curves with evolutionary polynomial regression: Application for intermittent streams. **Journal of Hydrologic Engineering**, v. 25, n. 1, p. 04019059, 2020.

CPRM - COMPANHIA DE PESQUISA DE RECURSOS MINERAIS. Atlas geoquímico da bacia do Rio São Francisco. Belo Horizonte, 2018. Disponível em: <https://rigeo.cprm.gov.br/jspui/handle/doc/20939>

DASGUPTA, Rijurekha; DAS, Subhasish; BANERJEE, Gourab; MAZUMDAR, Asis. Revisit hydrological modeling in ungauged catchments comparing regionalization, satellite observations, and machine learning approaches. **HydroResearch**, 2023.

DE BOOR, Carl. **A practical guide to splines**. New York: springer-verlag, 1978.

EHR – DEPARTAMENTO DE ENGENHARIA HIDRÁULICA E RECURSOS HÍDRICOS DA UNIVERSIDADE FEDERAL DE MINAS GERAIS. Regionalização de vazões mínimas sazonais do Estado de Minas Gerais, 2021.

FRANCHINI, Marco; SUPPO, Mauro. Regional analysis of flow duration curves for a limestone region. **Water resources management**, v. 10, p. 199-218, 1996.

FENNESSEY, Neil; VOGEL, Richard M. Regional flow-duration curves for ungauged sites in Massachusetts. **Journal of Water Resources Planning and Management**, v. 116, n. 4, p. 530-549, 1990.

GUO, Yuhan; ZHANG, Yongqiang; ZHANG, Lu; WANG, Zhonggen. Regionalization of hydrological modeling for predicting streamflow in ungauged catchments: A

comprehensive review. **Wiley Interdisciplinary Reviews: Water**, v. 8, n. 1, p. e1487, 2021.

HE, Y.; BÁRDOSSY, A.; ZEHE, E. A review of regionalisation for continuous streamflow simulation. **Hydrology and Earth System Sciences**, v. 15, n. 11, p. 3539-3553, 2011.

HUGHES, D. A.; SMAKHTIN, V. Daily flow time series patching or extension: a spatial interpolation approach based on flow duration curves. **Hydrological Sciences Journal**, v. 41, n. 6, p. 851-871, 1996.

INMET – INSTITUTO NACIONAL DE METEOROLOGIA. Normais climatológicas do Brasil. Disponível em: <https://portal.inmet.gov.br/normais>

KANISHKA, Ganvir; ELDHO, T. I. Streamflow estimation in ungauged basins using watershed classification and regionalization techniques. **Journal of Earth System Science**, v. 129, p. 1-18, 2020.

KAY, A. L. et al. An investigation of site-similarity approaches to generalisation of a rainfall–runoff model. **Hydrology and Earth System Sciences**, v. 11, n. 1, p. 500-515, 2007.

LAAHA, G., Skøien, J. O., Nobilis, F., e Blöschl, G. Spatial prediction of stream temperatures using top-kriging with an external drift. **Environmental Modeling & Assessment**, v. 18, p. 671-683, 2013.

LEONG, C.; YOKOO, Y. A step toward global-scale applicability and transferability of flow duration curve studies: A flow duration curve review (2000–2020). **Journal of Hydrology**, v. 603, p. 126984, 2021.

Li, Fapeng et al. Runoff predictions in ungauged catchments in southeast Tibetan Plateau. **Journal of Hydrology**, v. 511, p. 28-38, 2014.

MARIANI, M. C.; BASU, K. Local regression type methods applied to the study of geophysics and high frequency financial data. **Physica A: Statistical Mechanics and Its Applications**, v. 410, p. 609-622, 2014.

MASSELOT, Pierre et al. Streamflow forecasting using functional regression. **Journal of Hydrology**, v. 538, p. 754-766, 2016.

MATONSE, Adão H.; KROLL, Charles N. Applying hillslope-storage models to improve low flow estimates with limited streamflow data at a watershed scale. **Journal of hydrology**, v. 494, p. 20-31, 2013.

MAZELAN, N. A.; SUHAILA, J. Exploring rainfall variabilities using statistical functional data analysis. In: **IOP Conference Series: Earth and Environmental Science**. IOP Publishing, 2023. p. 012007.

MIMIKOU, Maria; KAEMAKI, Stella. Regionalization of flow duration characteristics. **Journal of Hydrology**, v. 82, n. 1-2, p. 77-91, 1985.

MONTANARI, Alberto et al. “Panta Rhei—everything flows”: change in hydrology and society—the IAHS scientific decade 2013–2022. **Hydrological sciences journal**, v. 58, n. 6, p. 1256-1275, 2013.

MÜLLER, Marc F.; THOMPSON, Sally E. Comparing statistical and process-based flow duration curve models in ungauged basins and changing rain regimes. **Hydrology and Earth System Sciences**, v. 20, n. 2, p. 669-683, 2016.

NASA JPL. NASA Shuttle Radar Topography Mission Global 1 arc second, 2013. Disponível em: <https://lpdaac.usgs.gov/products/srtmgl1v003/>

OUARDA, T. B.; CHARRON, C.; HUNDECHA, Y.; ST-HILAIRE, A.; CHEBANA, F. Introduction of the GAM model for regional low-flow frequency analysis at ungauged basins and comparison with commonly used approaches. **Environmental modelling & software**, v. 109, p. 256-271, 2018.

LOUDIN, L., Andréassian, V., Perrin, C., Michel, C., e Le Moine, N. Spatial proximity, physical similarity, regression and ungauged catchments: A comparison of regionalization approaches based on 913 French catchments. **Water resources research**, v. 44, n. 3, 2008.

NAGHETTINI, Mauro; PINTO, Éber José de Andrade. **Hidrologia estatística**. CPRM – Companhia de Pesquisa de Recursos Minerais, p. 552, 2007.

PINHEIRO, Viviane Borda. Calibração de um modelo chuva-vazão em bacias sem monitoramento fluviométrico a partir de curvas de permanência sintéticas. Universidade Federal de Minas Gerais, 2009.

PINHEIRO, Viviane Borda; NAGHETTINI, Mauro. Calibration of the Parameters of a Rainfall-Runoff Model in Ungauged Basins Using Synthetic Flow Duration Curves as Estimated by Regional Analysis. **Journal of Hydrologic Engineering**, v. 18, n. 12, p. 1617-1626, 2013.

PROJETO MAPBIOMAS. Mapbiomas, 2022. Disponível em: <https://brasil.mapbiomas.org/colecoes-mapbiomas/>.

PUGLIESE, Alessio; CASTELLARIN, Attilio; BRATH, Armando. Geostatistical prediction of flow–duration curves in an index-flow framework. **Hydrology and Earth System Sciences**, v. 18, n. 9, p. 3801-3816, 2014.

PUGLIESE, Alessi; FARMER, William H.; CASTELLARIN, Attilio; ARCHFIELD, Stacey. A.; VOGEL, Richard. M. Regional flow duration curves: Geostatistical techniques versus multivariate regression. **Advances in water resources**, v. 96, p. 11-22, 2016.

PUMO, Dario; NOTO, Leonardo Valerio; VIOLA, Francesco. Ecohydrological modelling of flow duration curve in Mediterranean river basins. **Advances in Water Resources**, v. 52, p. 314-327, 2013.

PUMO, Dario; CARACCIOLO, Domenico; VIOLA, Francesco; NOTO, Leonardo V. Climate change effects on the hydrological regime of small non-perennial river basins. **Science of the Total Environment**, v. 542, p. 76-92, 2016.

QUIMPO, Rafael G.; ALEJANDRINO, Angel A.; MCNALLY, Thomas A. Regionalized flow duration for Philippines. **Journal of Water Resources Planning and Management**, v. 109, n. 4, p. 320-330, 1983.

RAMSAY, James O.; DALZELL, C. J. Some tools for functional data analysis. **Journal of the Royal Statistical Society Series B: Statistical Methodology**, v. 53, n. 3, p. 539-561, 1991.

RAMSAY, J. O.; SILVERMAN, B. W. Functional Data Analysis. 2 ed. New York: Springer, 2005.

RAZAVI, Tara; COULIBALY, Paulin. Streamflow prediction in ungauged basins: review of regionalization methods. **Journal of hydrologic engineering**, v. 18, n. 8, p. 958-975, 2013.

REQUENA, Ana I.; OUARDA, Taha B. M. J.; CHEBANA, Fateh. Flood frequency analysis at ungauged sites based on regionally estimated streamflows. **Journal of Hydrometeorology**, v. 18, n. 9, p. 2521-2539, 2017.

REQUENA, Ana I.; OUARDA, Taha B.M.J.; CHEBANA, Fateh. Low-flow frequency analysis at ungauged sites based on regionally estimated streamflows. **Journal of Hydrology**, v. 563, p. 523-532, 2018a.

REQUENA, Ana I.; CHEBANA, Fateh; OUARDA, Taha BMJ. A functional framework for flow-duration-curve and daily streamflow estimation at ungauged sites. **Advances in Water Resources**, v. 113, p. 328-340, 2018b.

SAMUEL, Jos; COULIBALY, Paulin; METCALFE, Robert A. Estimation of continuous streamflow in Ontario ungauged basins: Comparison of regionalization methods. **Journal of Hydrologic Engineering**, v. 16, n. 5, p. 447-459, 2011.

SAVILLE, Thorndike; WATSON, John Dargan. An Investigation of the flow-duration characteristics of North Carolina Streams. **Eos, Transactions American Geophysical Union**, v. 14, n. 1, p. 406-425, 1933.

SEARCY, J. K. Flow duration curves. In: Manual of Hydrology, part. 2. Low – flow techniques. USGS Supply Paper. Washington: US Government Printing Office, p. 33, 1959.

SINHA, Rajiv; GUPTA, Alok; MISHRA, Kanchan; TRIPATHI, Shivam; NEPAL, Santosh; WAHID, S.M.; SWARNKAR, Somil. Basin-scale hydrology and sediment dynamics of the Kosi river in the Himalayan foreland. **Journal of Hydrology**, v. 570, p. 156-166, 2019.

SISEMA – SISTEMA ESTADUAL DE MEIO AMBIENTE E RECURSOS HÍDRICOS. Zonas Climáticas, 2018a. Disponível em:
<https://idesisema.meioambiente.mg.gov.br/geonetwork/srv/por/catalog.search#/metadata/b6bfa41c-07ec-4c31-83f4-262fcd628526>

SISEMA – SISTEMA ESTADUAL DE MEIO AMBIENTE E RECURSOS HÍDRICOS. Precipitação média anual, ZEE – MG, 2018b. Disponível em:

<https://idesisema.meioambiente.mg.gov.br/geonetwork/srv/api/records/7fa14753-e577-4f61-b5bf-029c0377a886>

SKØIEN, J.O.; MERZ, R.; BLÖSCHL, G. Top-kriging – geostatistics on stream networks. **Hydrology and Earth System Sciences**, v.10, p. 277–287, 2006.

SMAKHTIN, V. Y.; HUGHES, D. A.; CREUSE-NAUDIN, E. Regionalization of daily flow characteristics in part of the Eastern Cape, South Africa. **Hydrological Sciences Journal**, v. 42, n. 6, p. 919-936, 1997.

SMAKHTIN, V. U. Low flow hydrology: A review. **Journal of Hydrology**, v. 240, n. 3–4, p. 147–186, 2001.

SWAIN, Janaki Ballav; PATRA, Kanhu Charan. Streamflow estimation in ungauged catchments using regionalization techniques. **Journal of Hydrology**, v. 554, p. 420-433, 2017.

TERNYNCK, Camille; ALAYA, Mohamed Ali Ben; CHEBANA, Fateh; DABO-NIANG, Sophie; OUARDA, Taha B. M. J. Streamflow Hydrograph Classification Using Functional Data Analysis. **Journal of Hydrometeorology**, v. 17, n. 1, p. 327-344, 2015.

TUCCI, Carlos. E. M. Hidrologia: ciência e aplicação. 4. ed. Porto Alegre, **UFRGS/ABRH**, p. 943, 2009.

VIEIRA, Myrlla de Souza Batista. Estudo das vazões do sistema aquífero Urucuia em períodos de recessão hídrica. Universidade de Brasília, 2021.

VOGEL, Richard. M.; FENNESSEY, Neil. M. Flow-duration curves I: new interpretation and confidence intervals. **Journal of Water Resources Planning and Management**, v. 120 (4), p. 485-504, 1994.

VOGEL, Richard M.; FENNESSEY, Neil M. Flow duration curves II: a review of applications in water resources planning 1. **JAWRA Journal of the American Water Resources Association**, v. 31, n. 6, p. 1029-1039, 1995.

YANG, Xue et al. Runoff prediction in ungauged catchments in Norway: comparison of regionalization approaches. **Hydrology Research**, v. 49, n. 2, p. 487-505, 2018.

YOKOO, Yoshiyuki; SIVAPALAN, Murugesu. Towards reconstruction of the flow duration curve: development of a conceptual framework with a physical basis. **Hydrology and Earth System Sciences**, v. 15, n. 9, p. 2805-2819, 2011.

ZIN, Muhammad Athif Mat et al. Smoothing and differentiation of kinematic data using functional data analysis approach: an application of automatic and subjective methods. **Applied Sciences**, v. 10, n. 7, p. 2493, 2020.

APÊNDICE A – AVALIAÇÃO DA INSERÇÃO DAS COVARIÁVEIS NA CONSTRUÇÃO DO MODELO REGIONAL

Tabela A.1 - Índices de desempenho da FMR para duas variáveis dependentes

Variáveis		MAE	RMSE	PBIAS	NSE	R ²	KGE	CP	RMSPE
A+L	(Média)	53,62	90,20	10,36	-0,35	0,97	0,32	-32,99	74,12
	(Mediana)	26,47	47,50	-13,65	0,88	0,98	0,66	-1,38	33,36
A+Q _s	(Média)	65,72	103,79	7,93	0,24	0,97	0,44	-11,94	69,12
	(Mediana)	23,98	36,59	-6,65	0,83	0,97	0,60	-1,95	33,54
A+P	(Média)	52,19	81,94	-1,38	0,82	0,97	0,71	-3,70	49,48
	(Mediana)	15,48	34,28	-4,00	0,94	0,98	0,82	0,075	29,04
A+DD	(Média)	73,67	117,10	1,62	0,71	0,97	0,57	-7,28	47,33
	(Mediana)	25,54	40,06	-2,95	0,88	0,98	0,68	-2,54	35,39
A+Kc	(Média)	55,97	90,21	10,06	-0,32	0,97	0,31	-33,34	72,56
	(Mediana)	26,32	42,70	-13,25	0,85	0,98	0,62	-1,02	30,48

Tabela A.2 - Índices de desempenho da FMR para três variáveis dependentes

Variáveis		MAE	RMSE	PBIAS	NSE	R ²	KGE	CP	RMSPE
A+P+L	(Média)	45.45	73.69	-1.46	0.83	0.97	0.71	-3.34	49.43
	(Mediana)	17.25	34.10	-2.25	0.94	0.98	0.81	-0.25	28.94
A+P+Q _s	(Média)	53.38	83.41	-1.25	0.84	0.97	0.72	-3.20	49.47
	(Mediana)	18.28	34.81	-2.80	0.95	0.98	0.83	-0.06	29.84
A+P+DD	(Média)	53.22	85.92	-1.58	0.87	0.98	0.73	-2.45	37.94
	(Mediana)	17.57	32.57	-0.20	0.95	0.98	0.81	-0.35	30.03
A+P+Kc	(Média)	53.11	84.05	-1.32	0.82	0.97	0.71	-3.69	48.82
	(Mediana)	16.36	31.41	-3.05	0.94	0.98	0.83	-0.03	28.05

Tabela A.3 - Índices de desempenho da FMR para quatro variáveis dependentes

Variáveis		MAE	RMSE	PBIAS	NSE	R ²	KGE	CP	RMSPE
A+P+DD+L	(Média)	44.89	72.95	-1.61	0.87	0.97	0.74	-2.38	37.56
	(Mediana)	18.09	37.62	-1.00	0.94	0.98	0.81	-0.44	27.82
A+P+DD+Q _s	(Média)	53.15	85.53	-1.50	0.87	0.97	0.73	-2.33	37.39
	(Mediana)	18.79	32.73	-0.75	0.95	0.98	0.81	-0.37	27.87
A+P+DD+Kc	(Média)	53.92	87.78	-1.52	0.87	0.97	0.73	-2.46	37.70
	(Mediana)	19.55	39.61	1.15	0.94	0.98	0.81	-0.43	25.92

APÊNDICE B – CURVAS DE PERMANÊNCIA REGIONALIZADAS PELA FMR (ETAPA DE CALIBRAÇÃO)

Figura B.1 - Desempenho da regressão funcional múltipla para a estação Fazenda Ajudas (40040000)

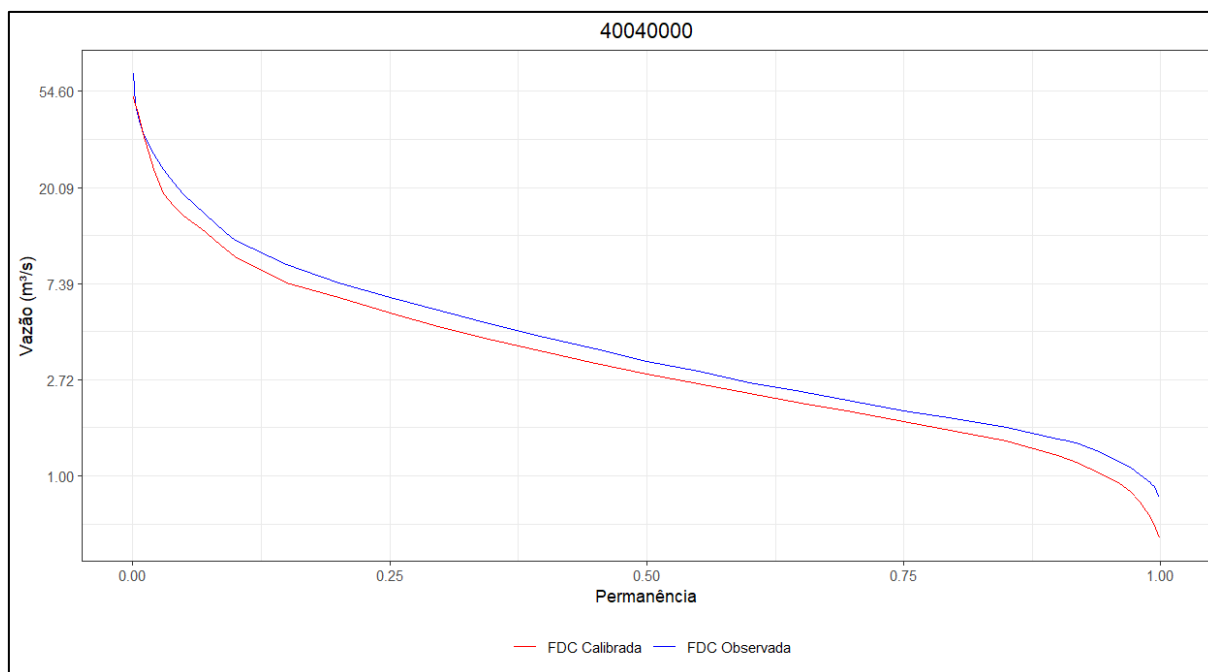


Figura B.2 - Desempenho da regressão funcional múltipla para a estação Tapiraí Jusante (40060001)

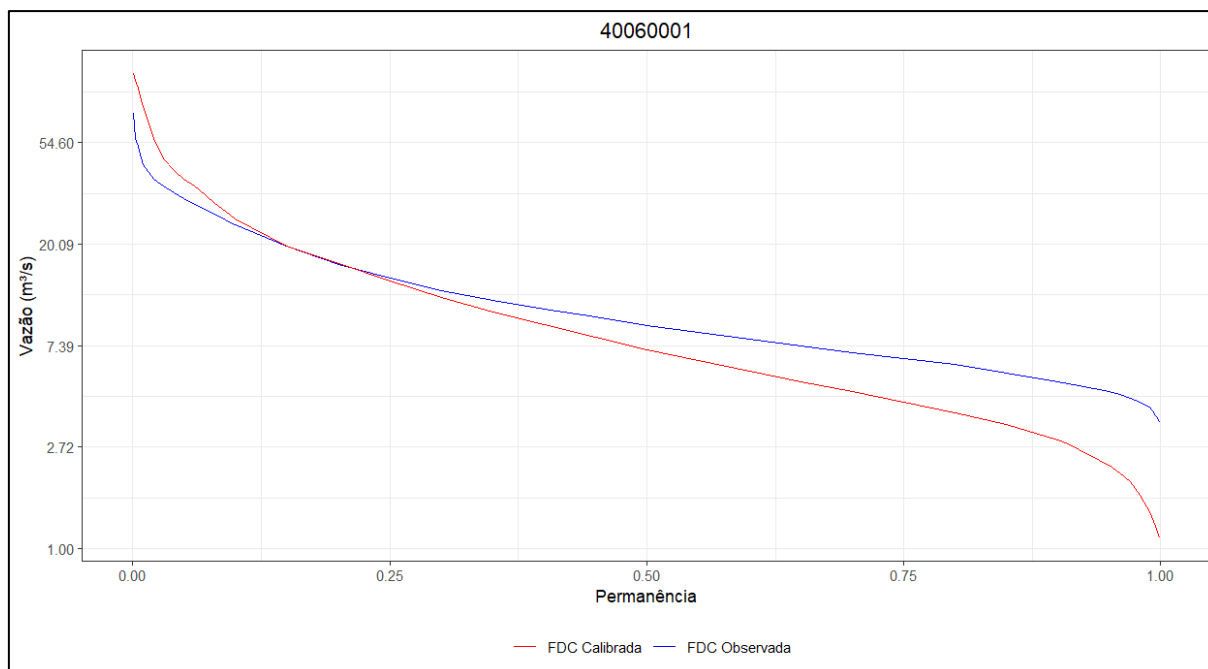


Figura B.3 - Desempenho da regressão funcional múltipla para a estação Porto das Andorinhas (40100000)

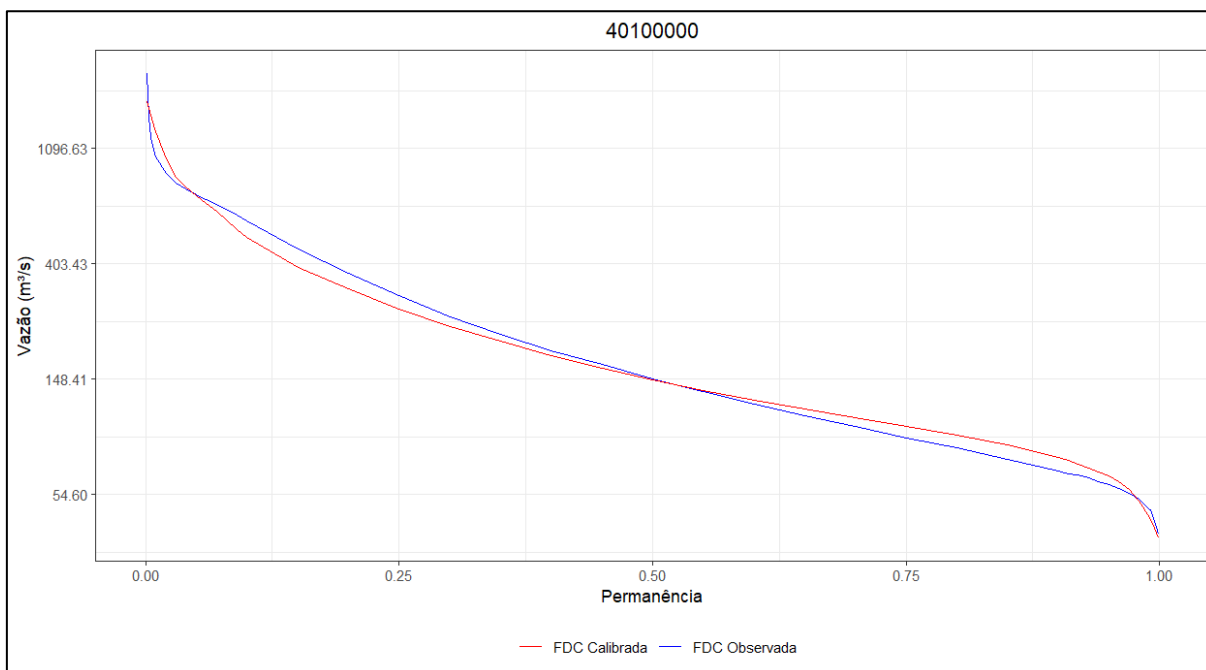


Figura B.4 - Desempenho da regressão funcional múltipla para a estação Velho da Taipa (40330000)

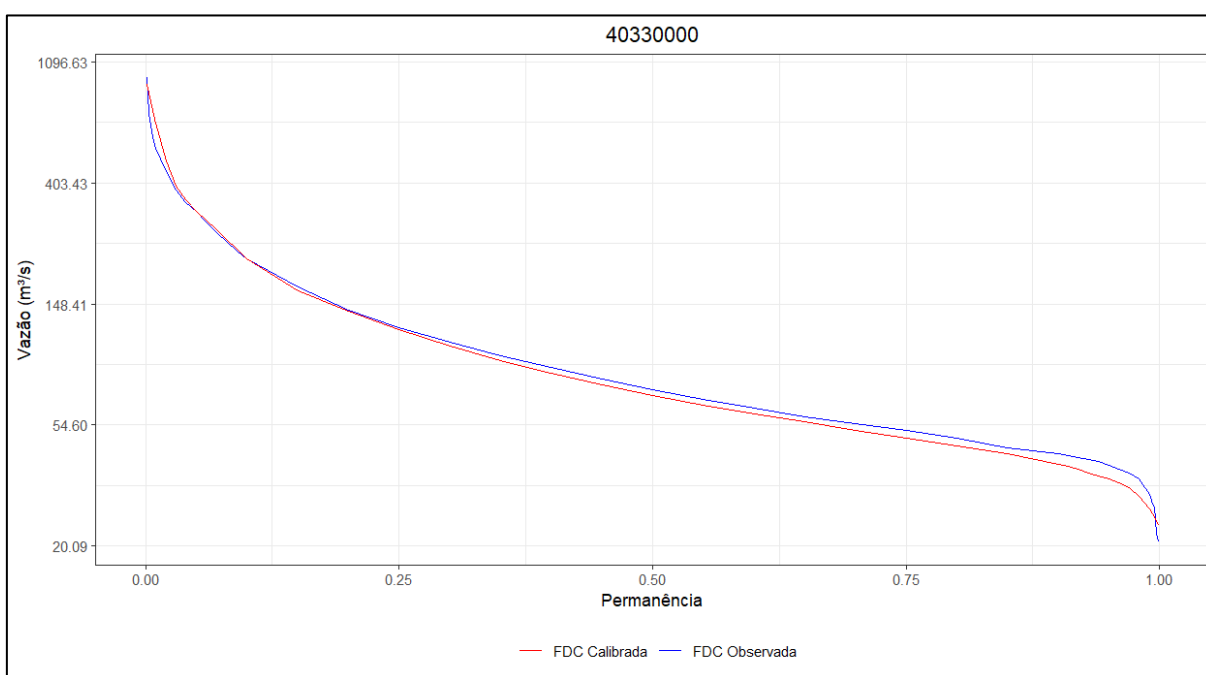


Figura B.5 - Desempenho da regressão funcional múltipla para a estação Ponte Nova do Paraopeba (40800001)

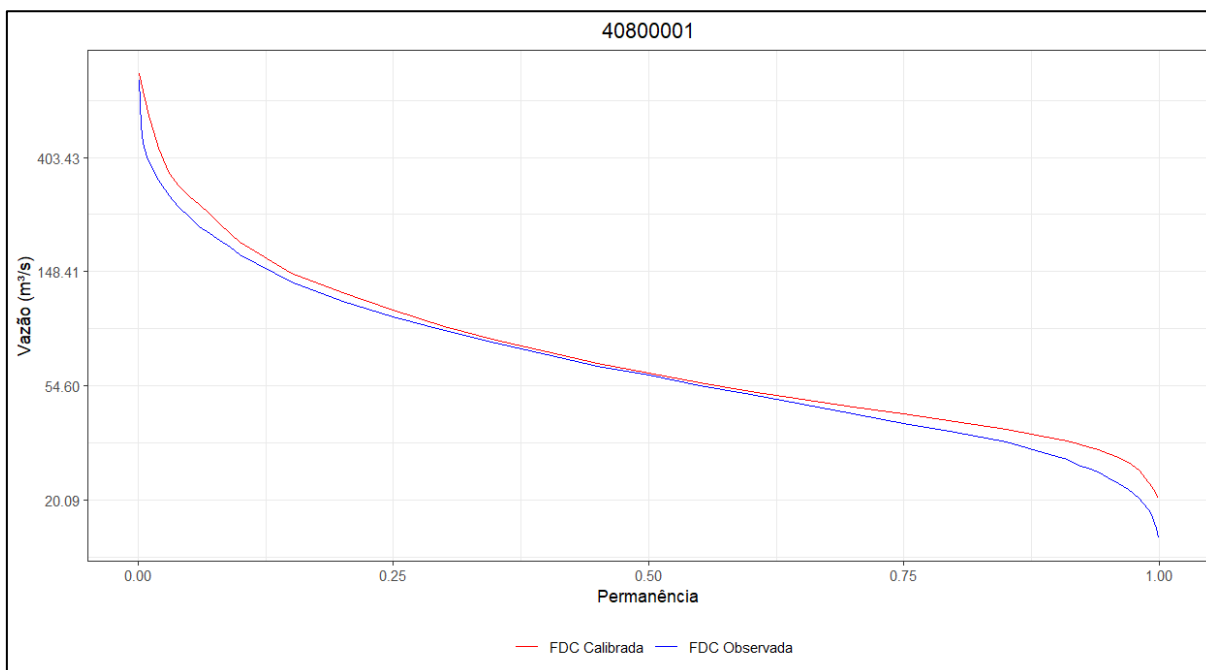


Figura B.6 - Desempenho da regressão funcional múltipla para a estação Fazenda Pasto Grande (40810800)

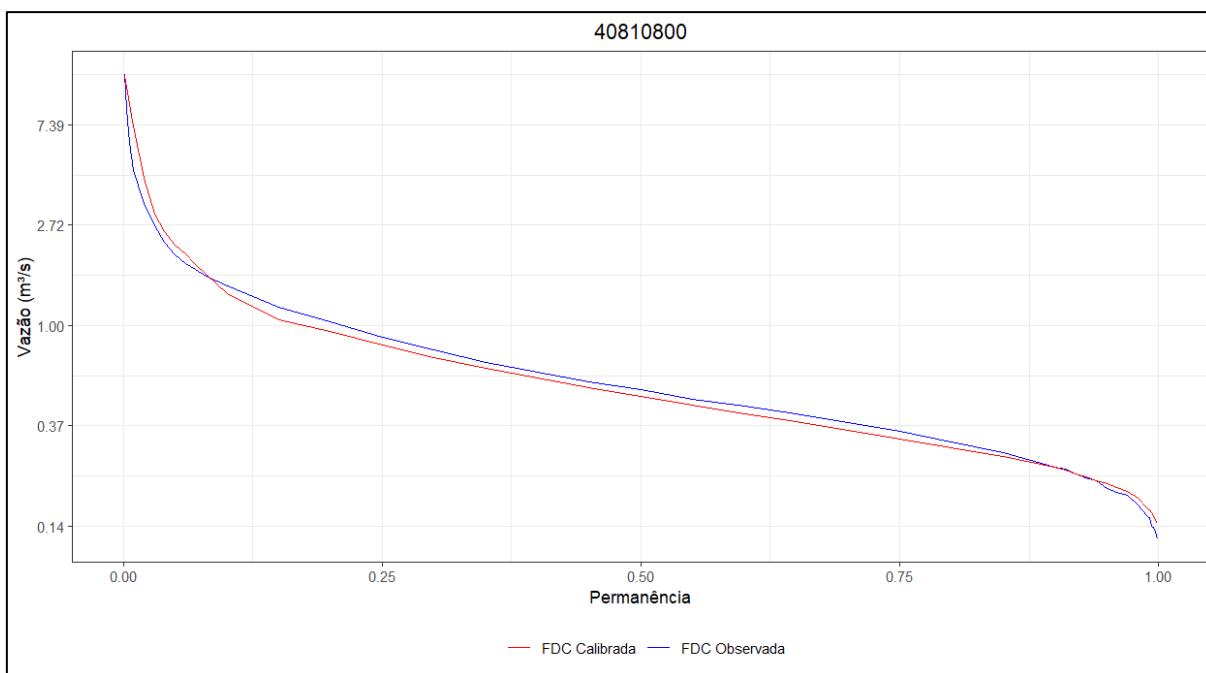


Figura B.7 - Desempenho da regressão funcional múltipla para a estação Itabirito - Linígrafo (41180000)

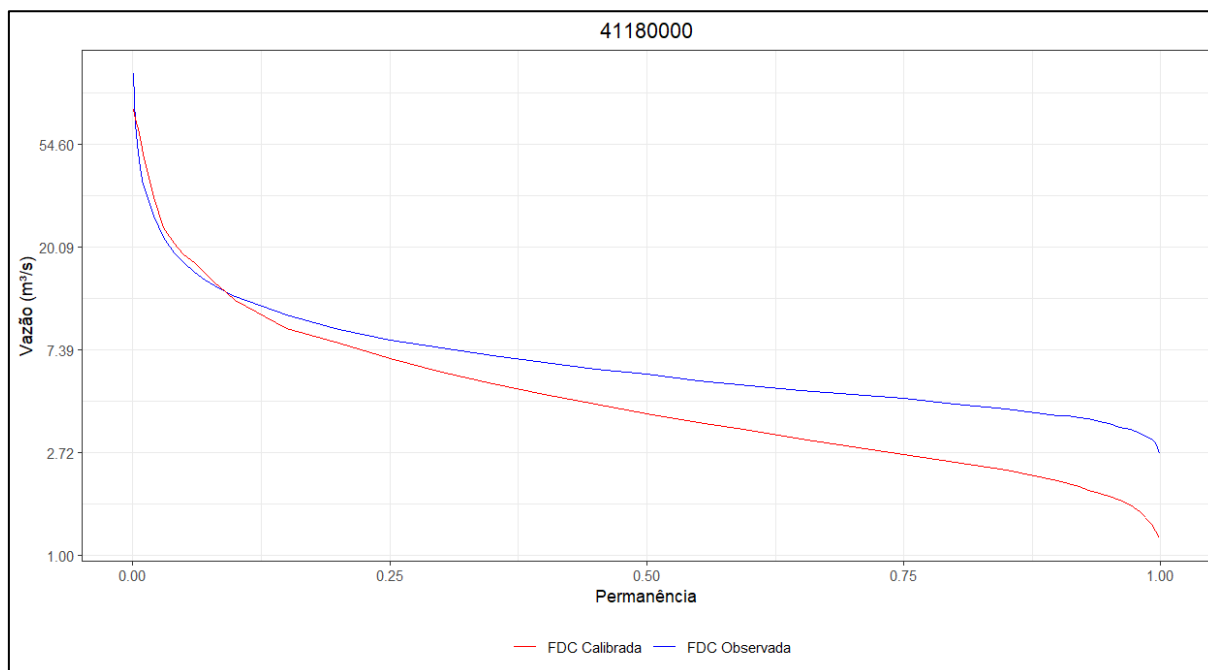


Figura B.8 - Desempenho da regressão funcional múltipla para a estação Honório Bicalho Montante (41199998)

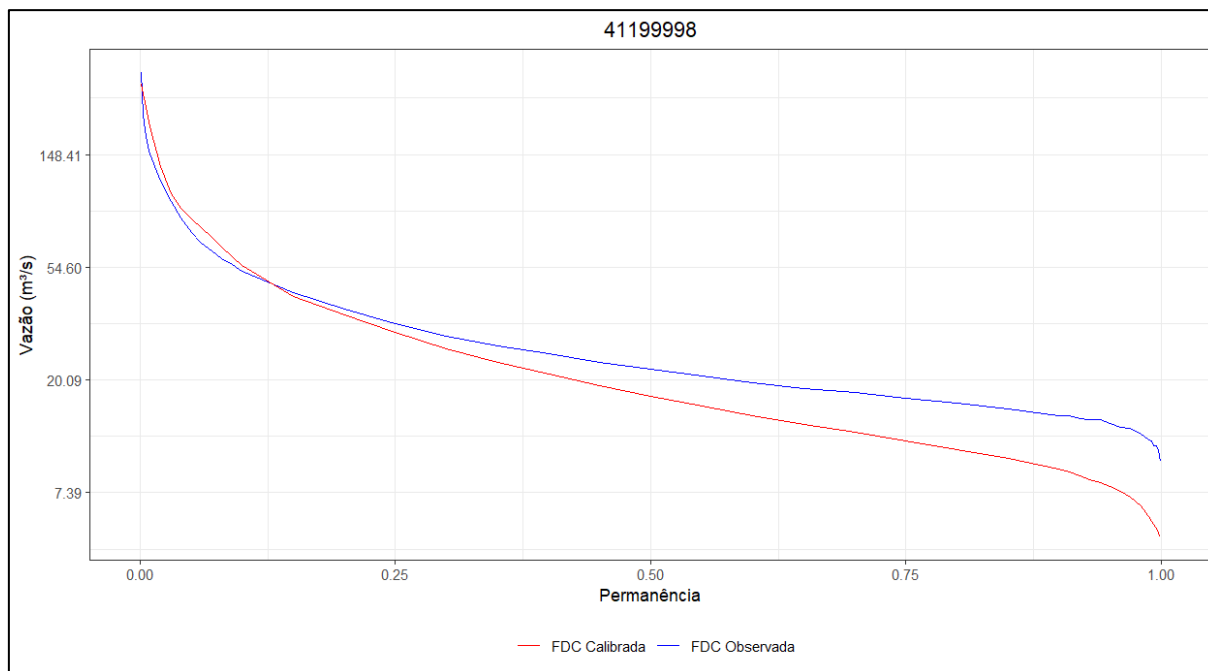


Figura B.9 - Desempenho da regressão funcional múltipla para a estação Vespasiano (41250000)

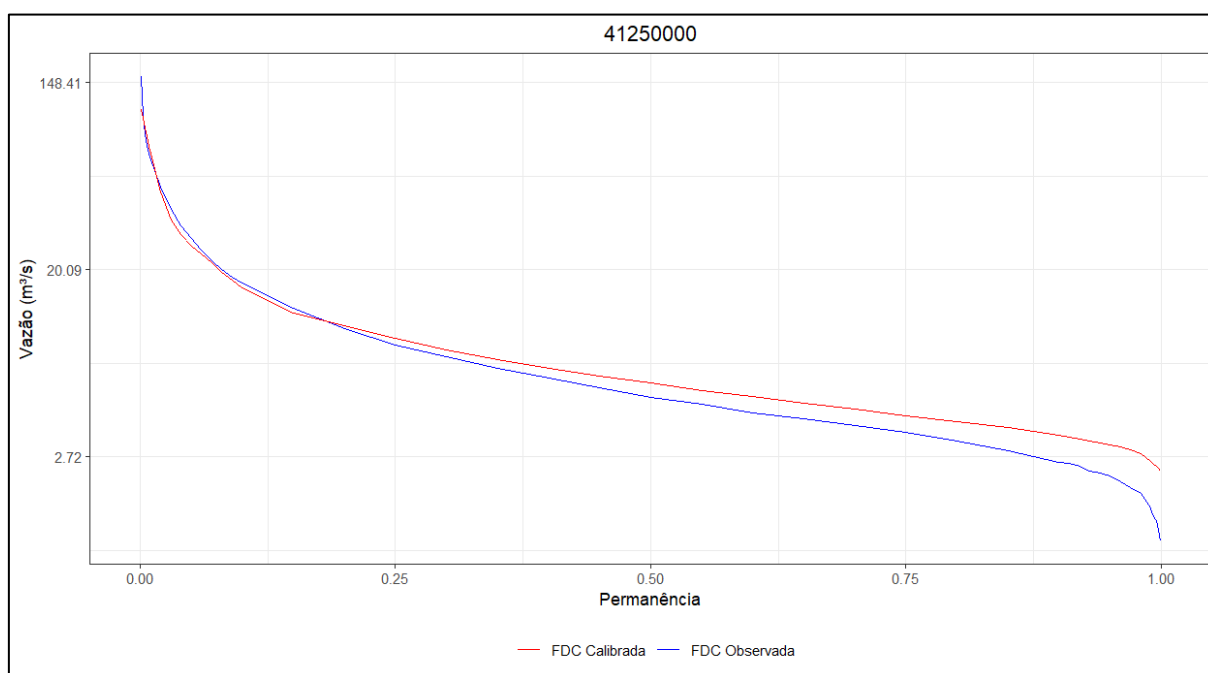


Figura B.10 - Desempenho da regressão funcional múltipla para a estação Pinhões (41260000)

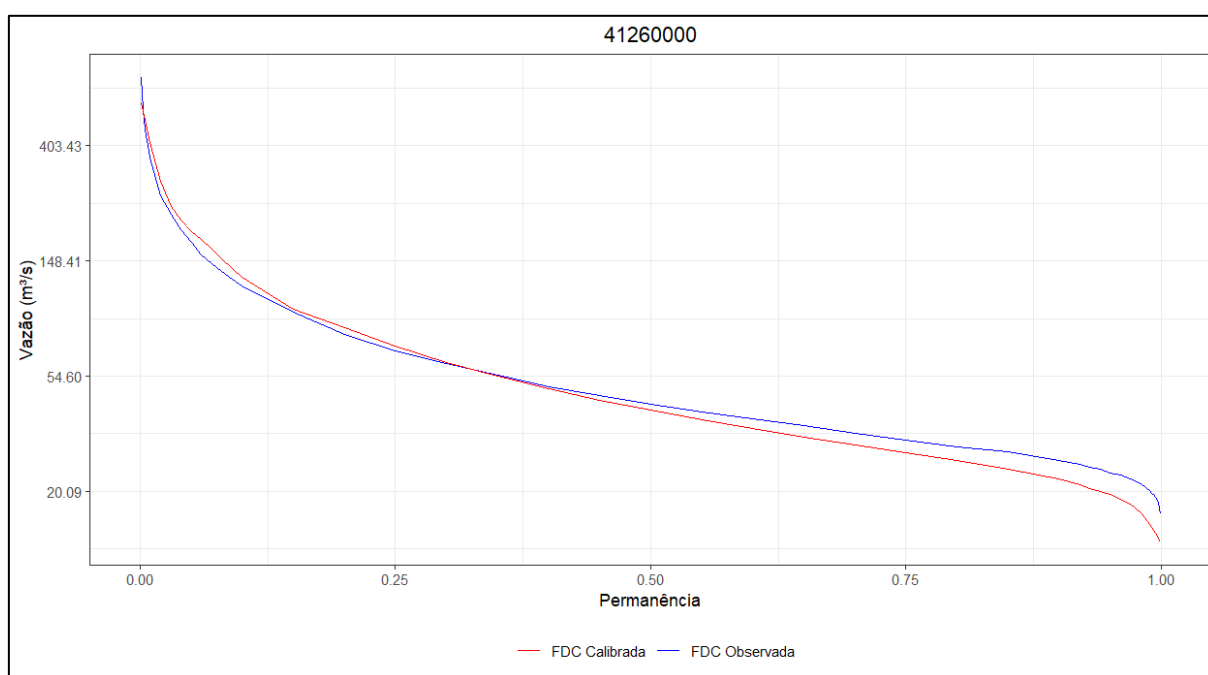


Figura B.11 - Desempenho da regressão funcional múltipla para a estação Ponte Preta (41380000)

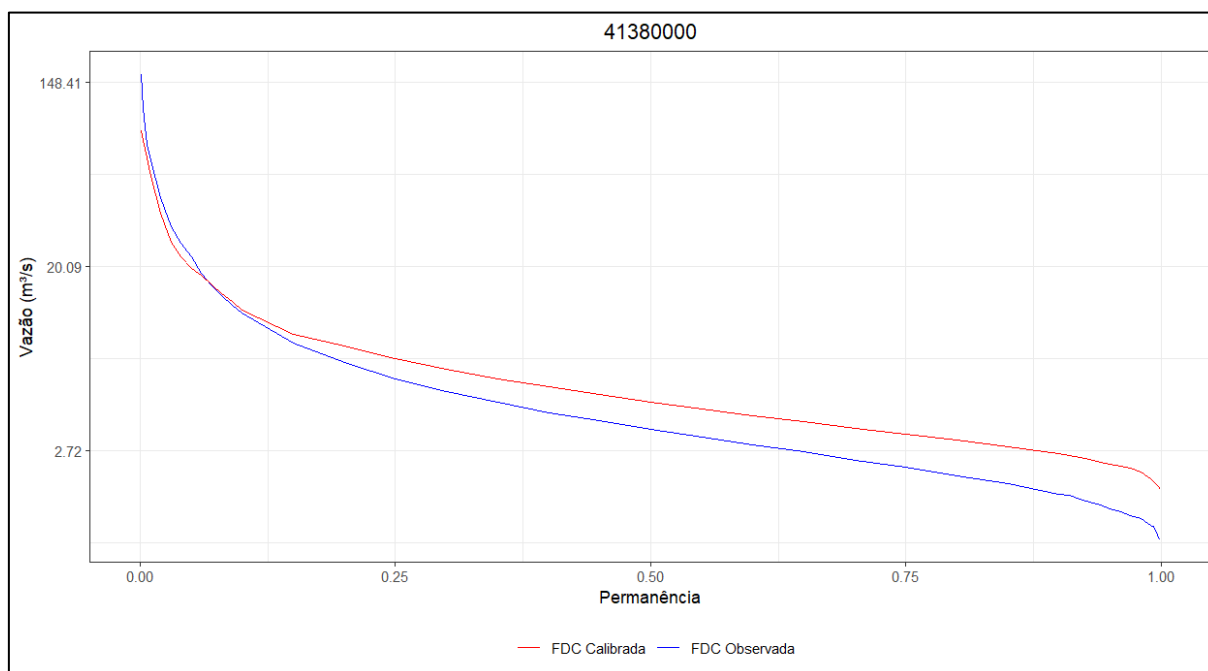


Figura B.12 - Desempenho da regressão funcional múltipla para a estação Jequitibá (41410000)

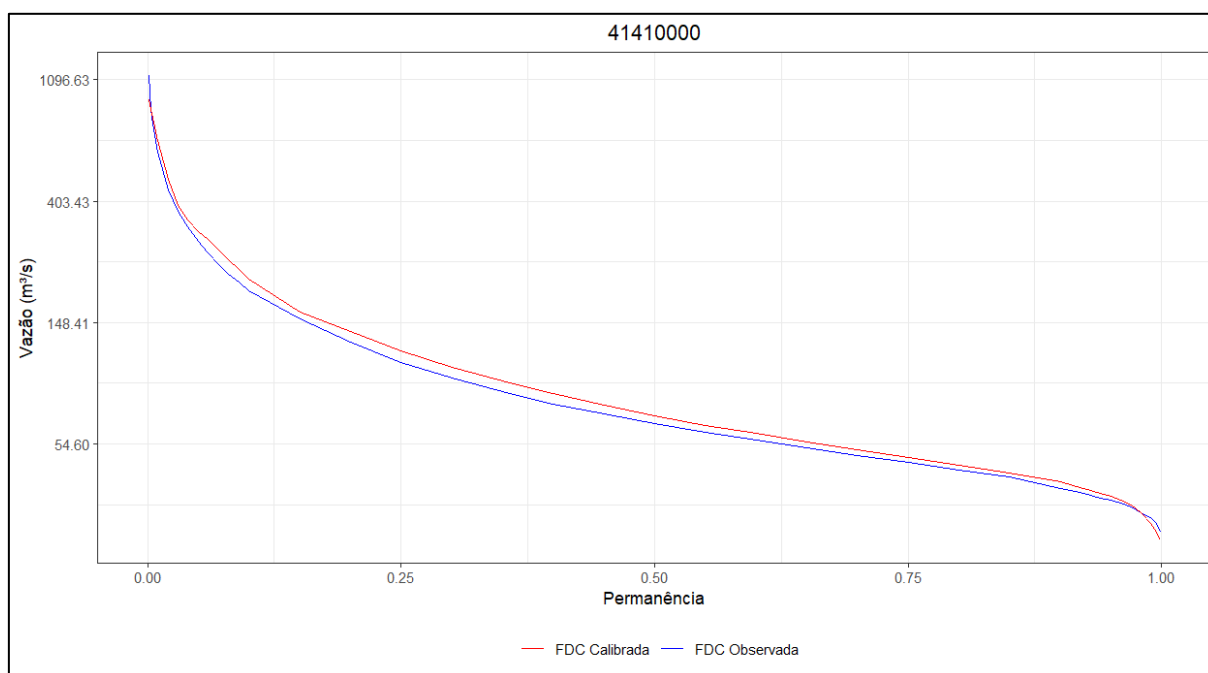


Figura B.13 - Desempenho da regressão funcional múltipla para a estação Ponte do Picão (41685000)

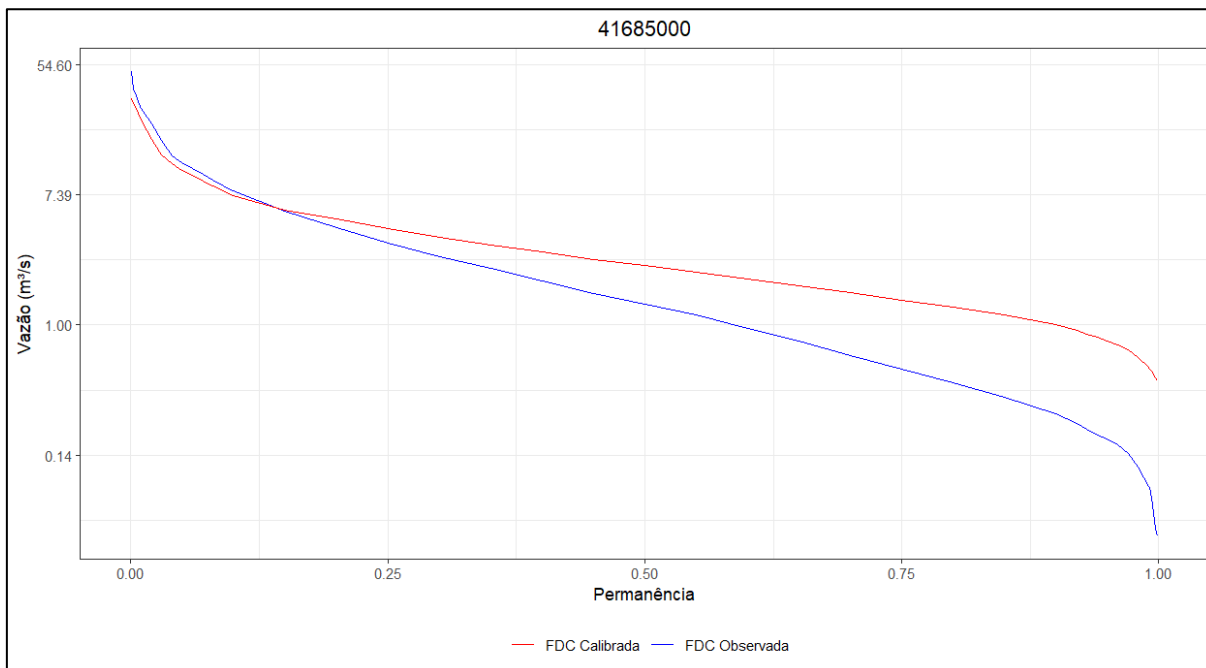


Figura B.14 - Desempenho da regressão funcional múltipla para a estação Presidente Juscelino Jusante (41780002)

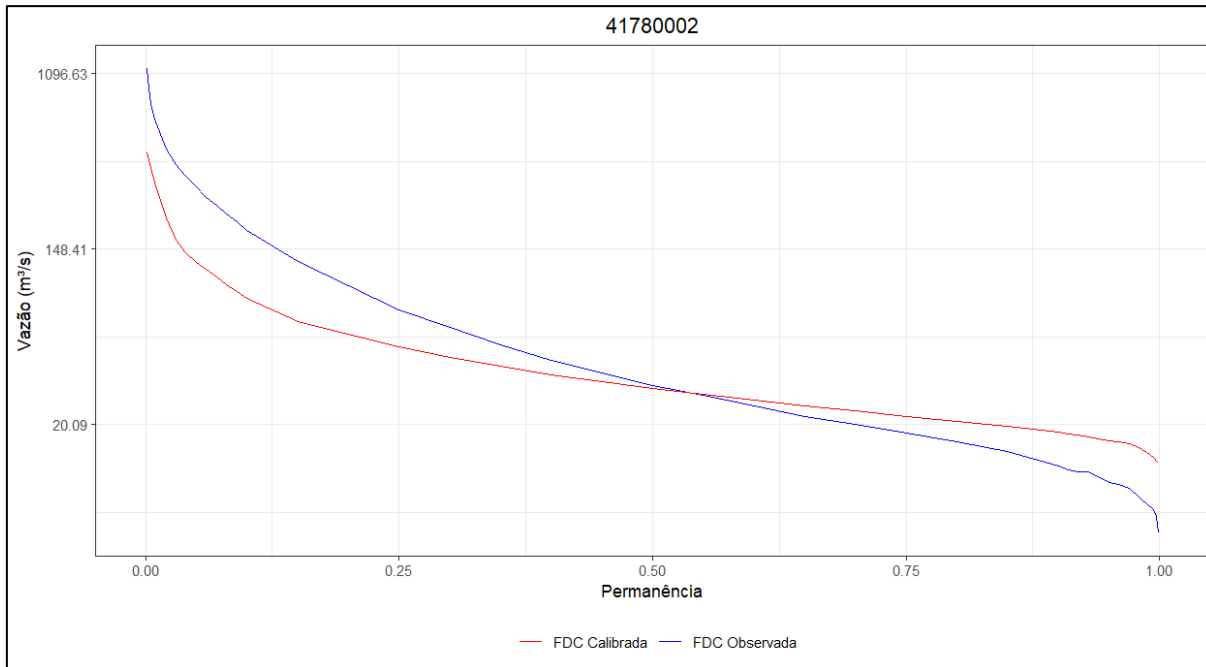


Figura B.15 - Desempenho da regressão funcional múltipla para a estação Várzea da Palma (41990000)

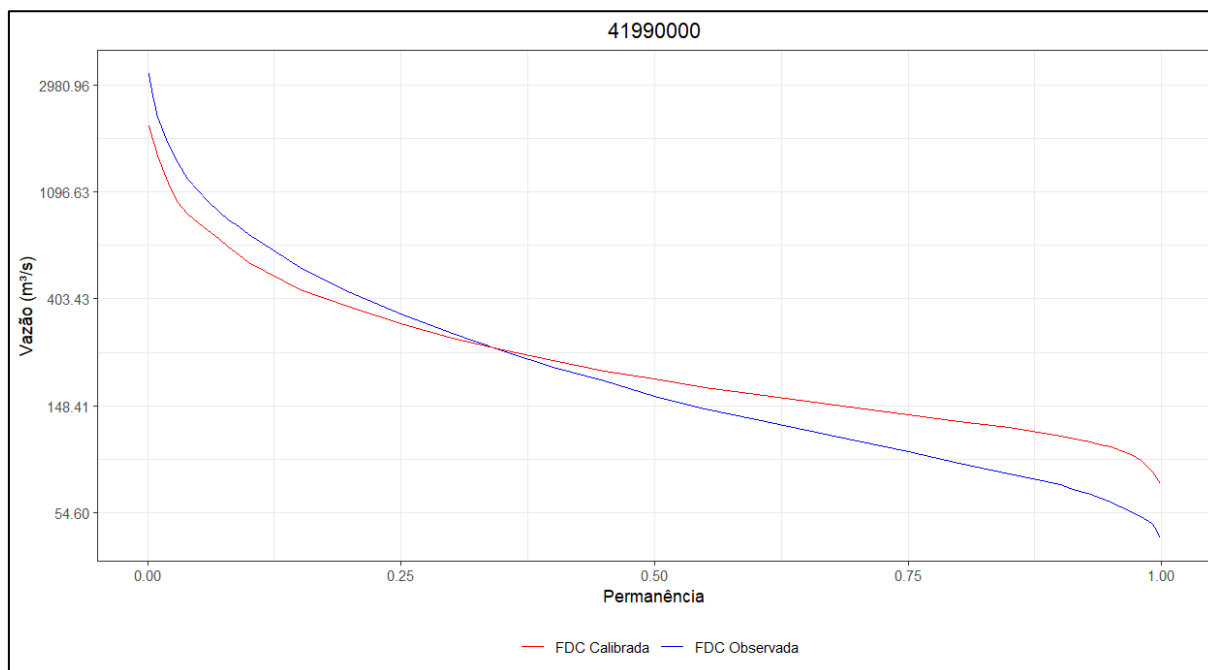


Figura B.16 - Desempenho da regressão funcional múltipla para a estação Ponte dos Ciganos (42187000)

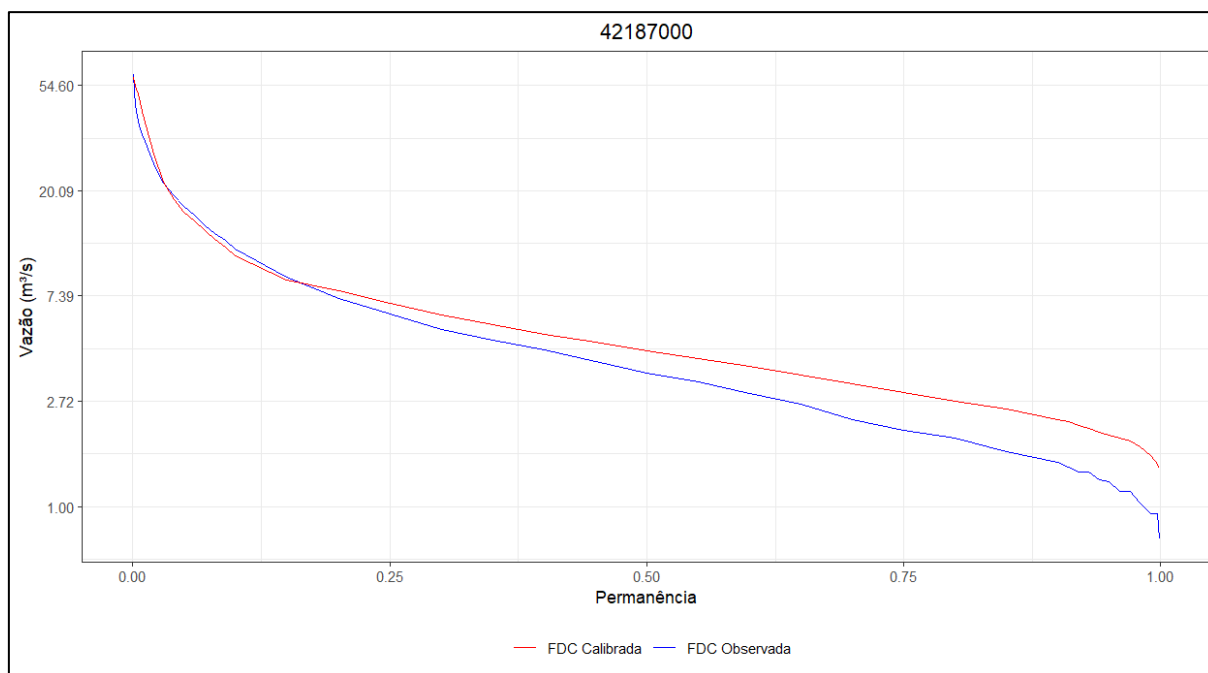


Figura B.17 - Desempenho da regressão funcional múltipla para a estação Cachoeira da Manteiga (42210000)

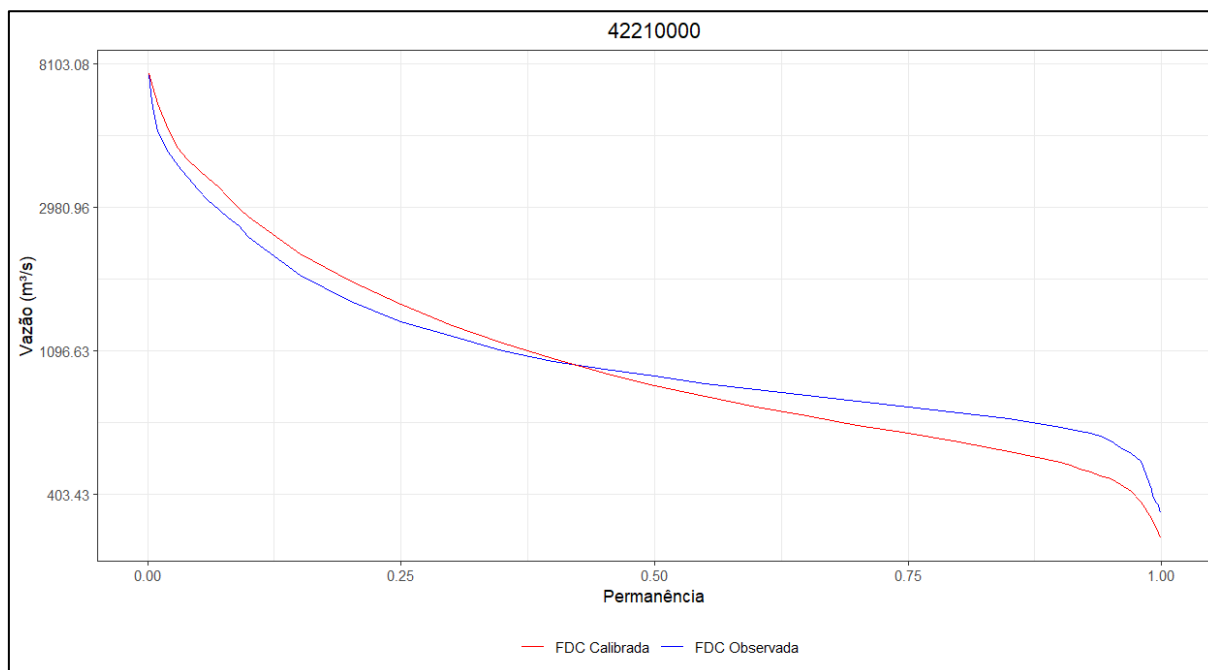


Figura B.18 - Desempenho da regressão funcional múltipla para a estação Fazenda Limoeiro (42250000)

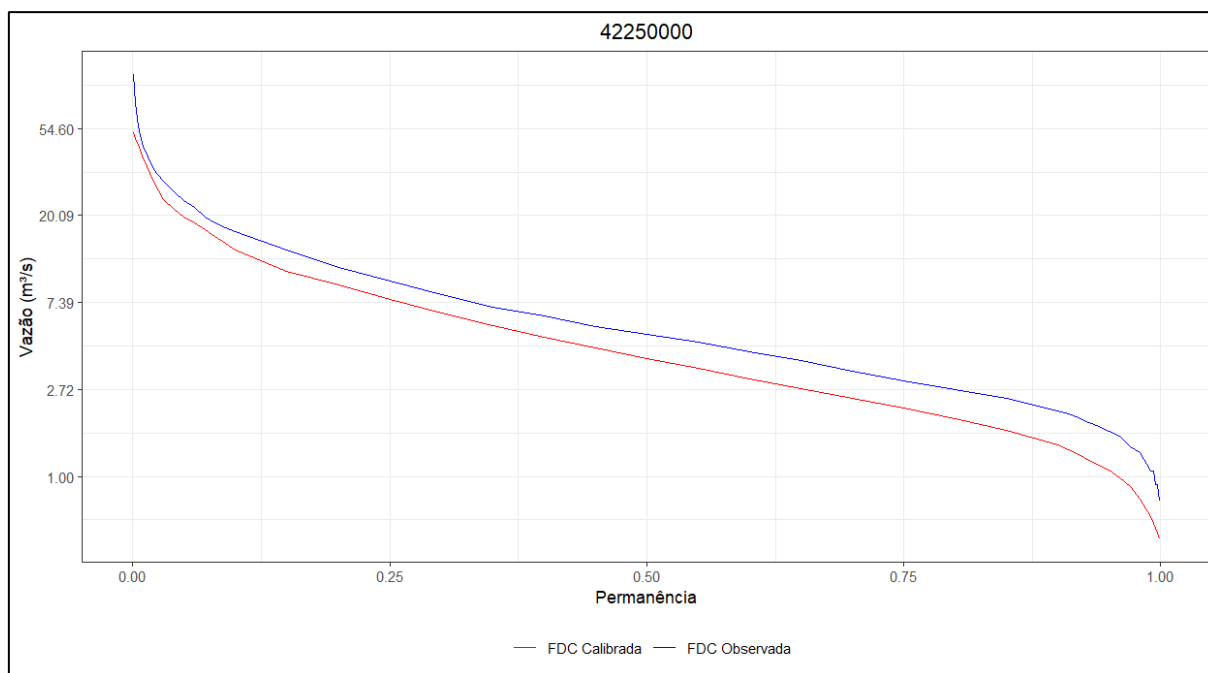


Figura B.19 - Desempenho da regressão funcional múltipla para a estação Barra do Escurinho (42257000)

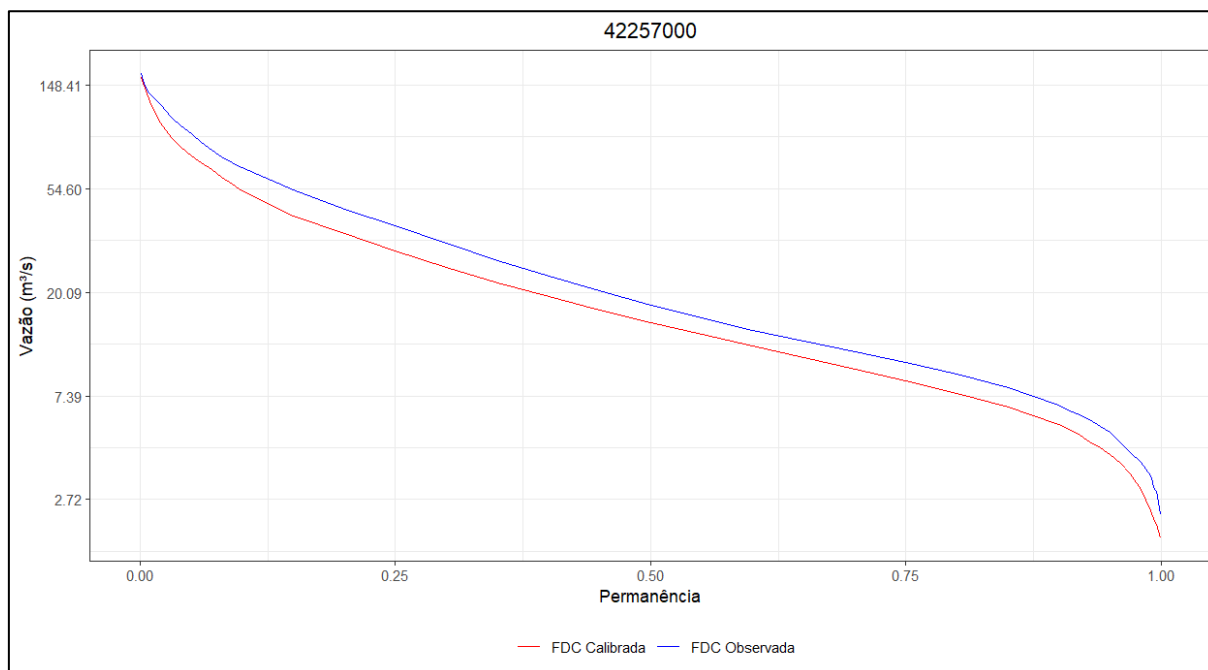


Figura B.20 - Desempenho da regressão funcional múltipla para a estação Ponte Da Br-040 - Paracatu (42290000)

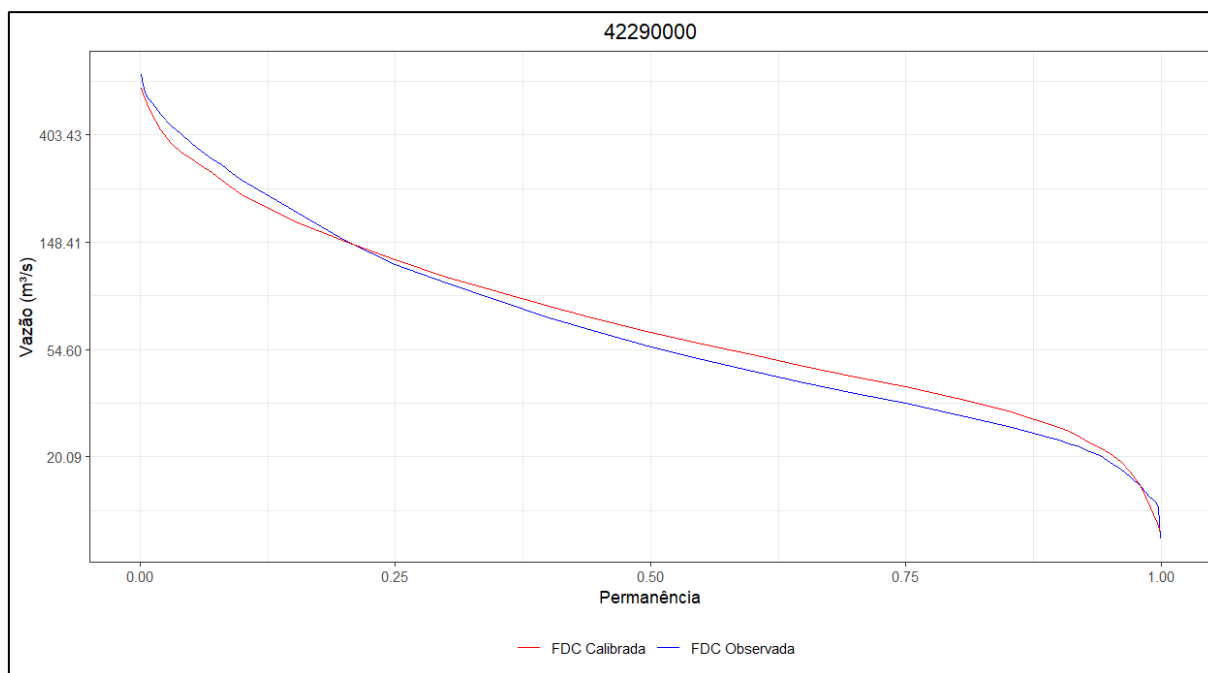


Figura B.21 - Desempenho da regressão funcional múltipla para a estação Fazenda Barra da Égua (42435000)

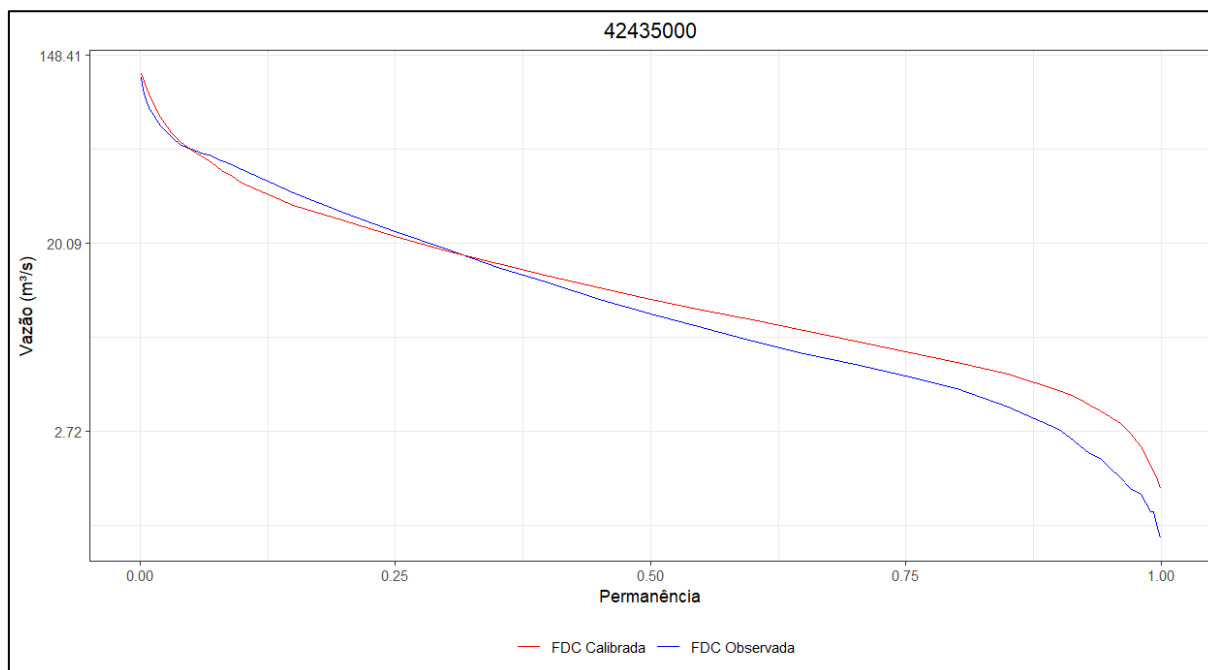


Figura B.22 - Desempenho da regressão funcional múltipla para a estação Fazenda Limeira (42460000)

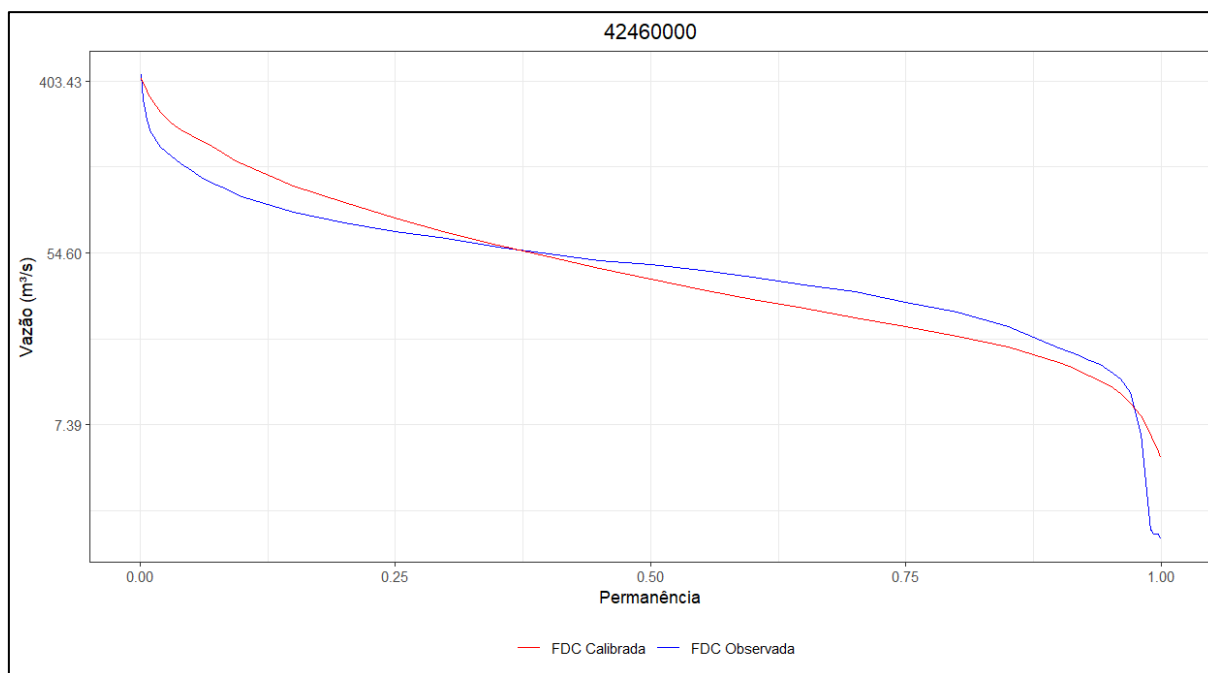


Figura B.23 - Desempenho da regressão funcional múltipla para a estação Fazenda O Resfriado (42545500)

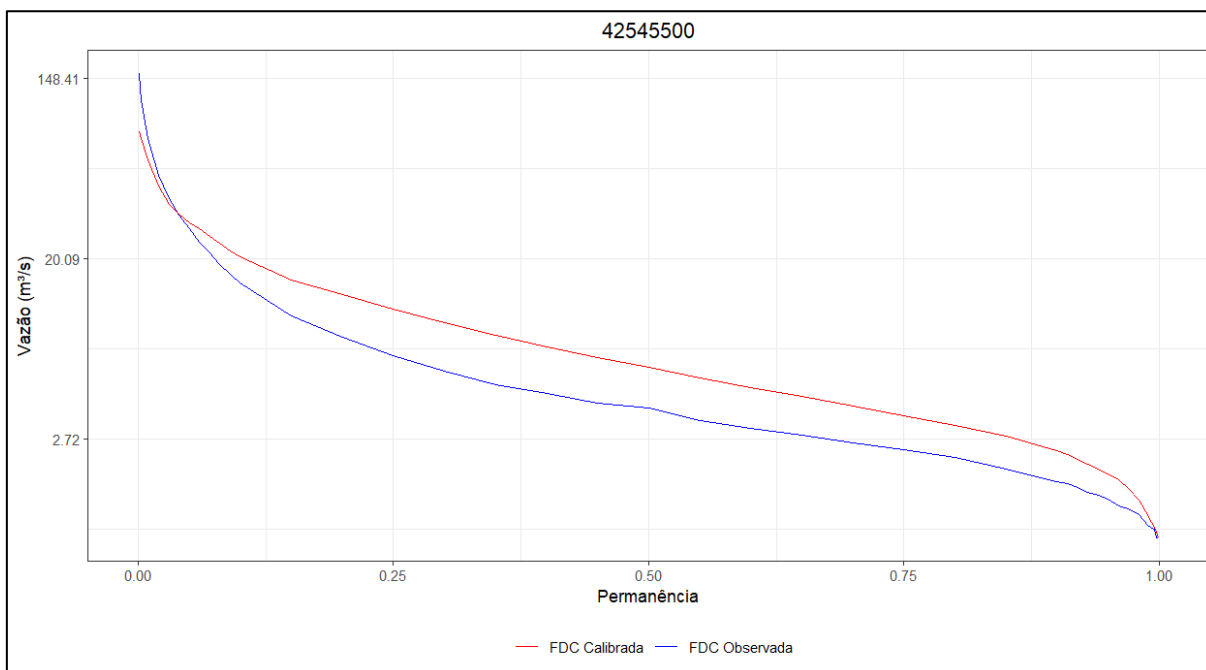


Figura B.24 - Desempenho da regressão funcional múltipla para a estação Porto da Extrema (42690001)

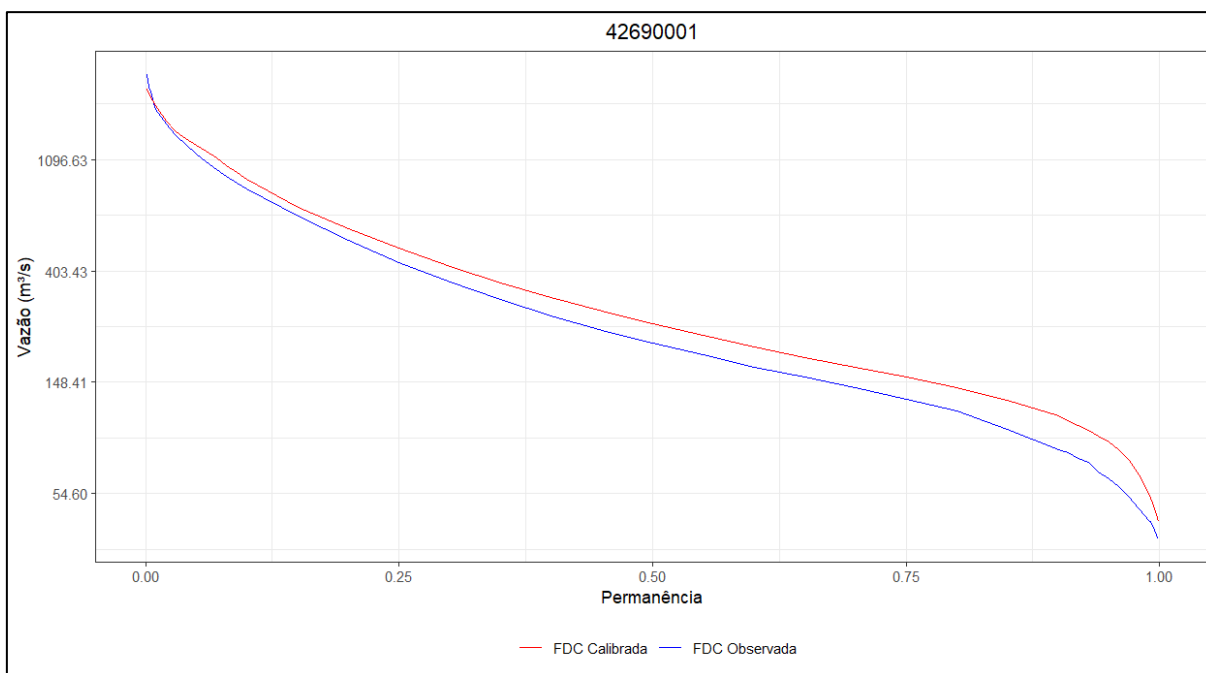


Figura B.25 - Desempenho da regressão funcional múltipla para a estação Porto Alegre (42980000)

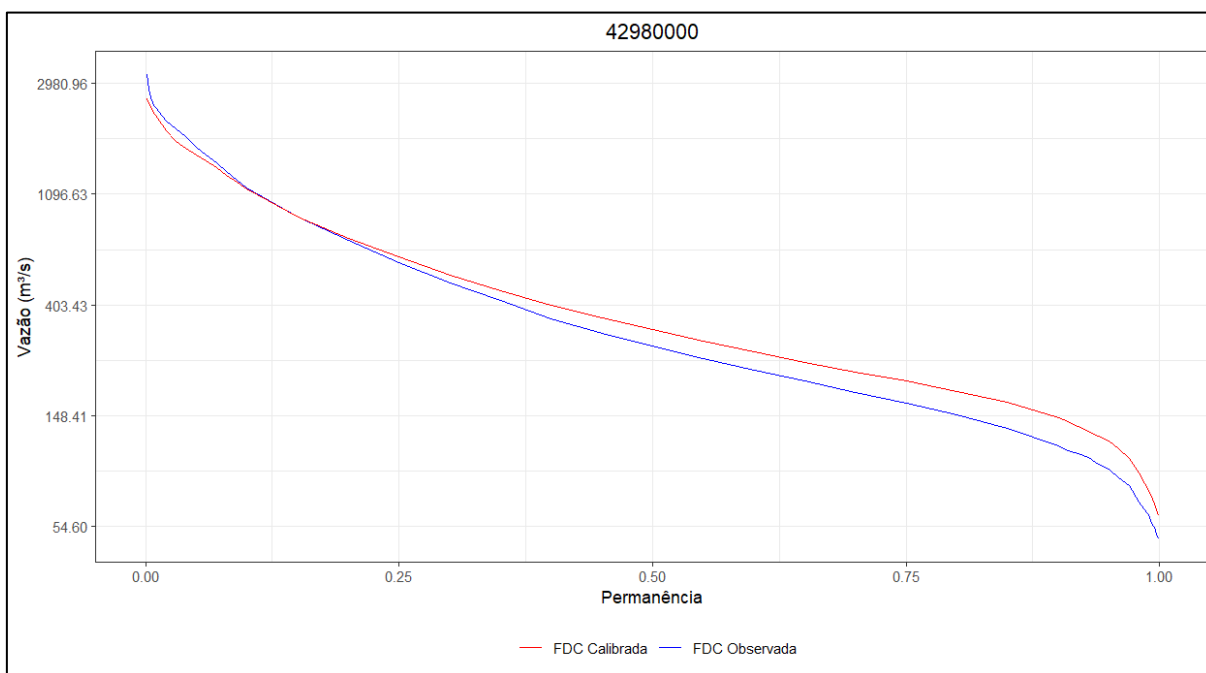


Figura B.26 - Desempenho da regressão funcional múltipla para a estação São Romão (43200000)

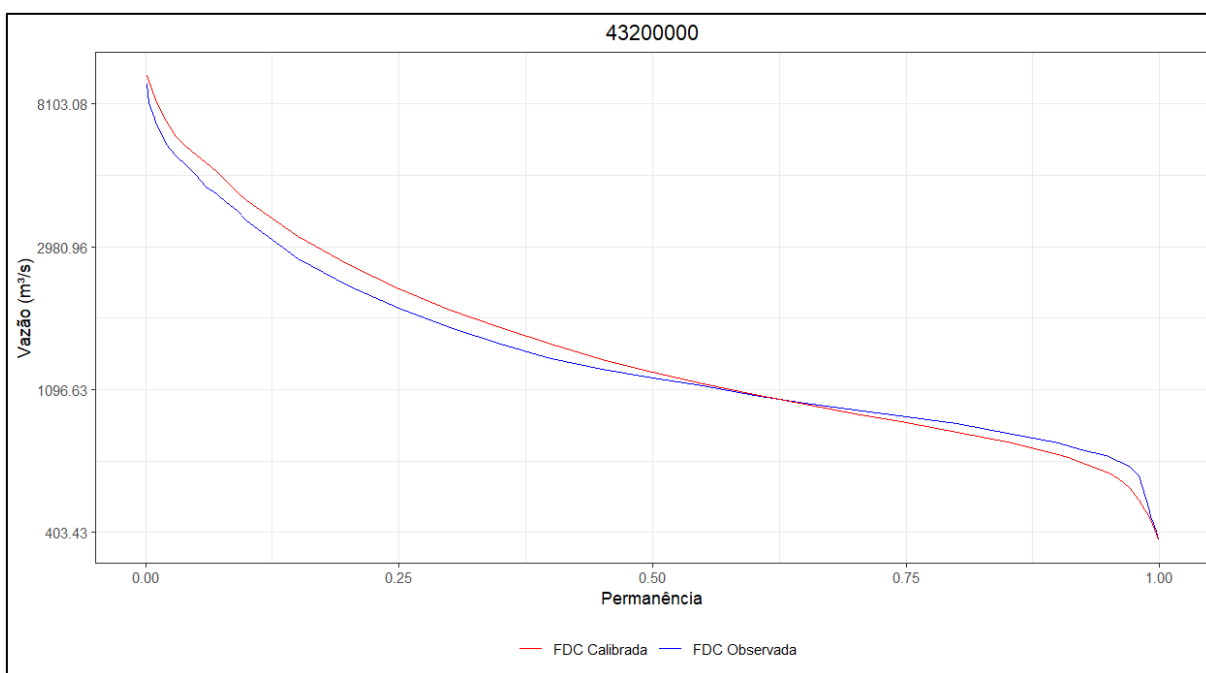


Figura B.27 - Desempenho da regressão funcional múltipla para a estação Buritis Jusante (43250002)

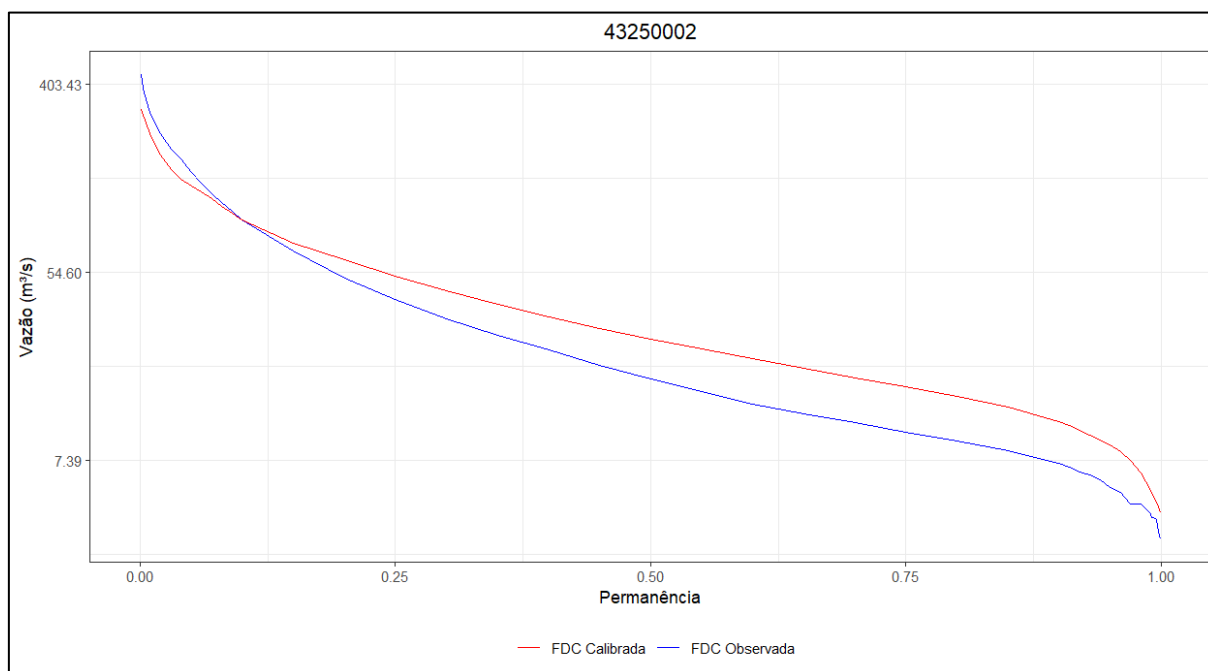


Figura B.28 - Desempenho da regressão funcional múltipla para a estação Fazenda Carvalho (43300000)

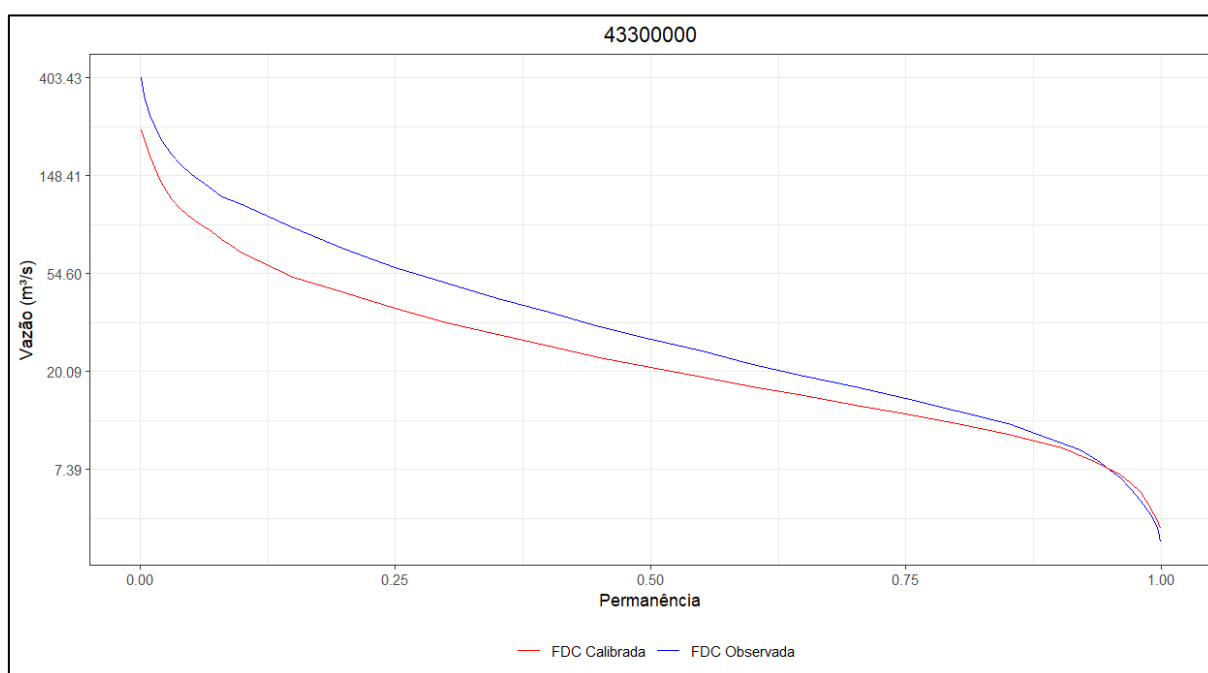


Figura B.29 - Desempenho da regressão funcional múltipla para a estação Arinos Montante (43429998)

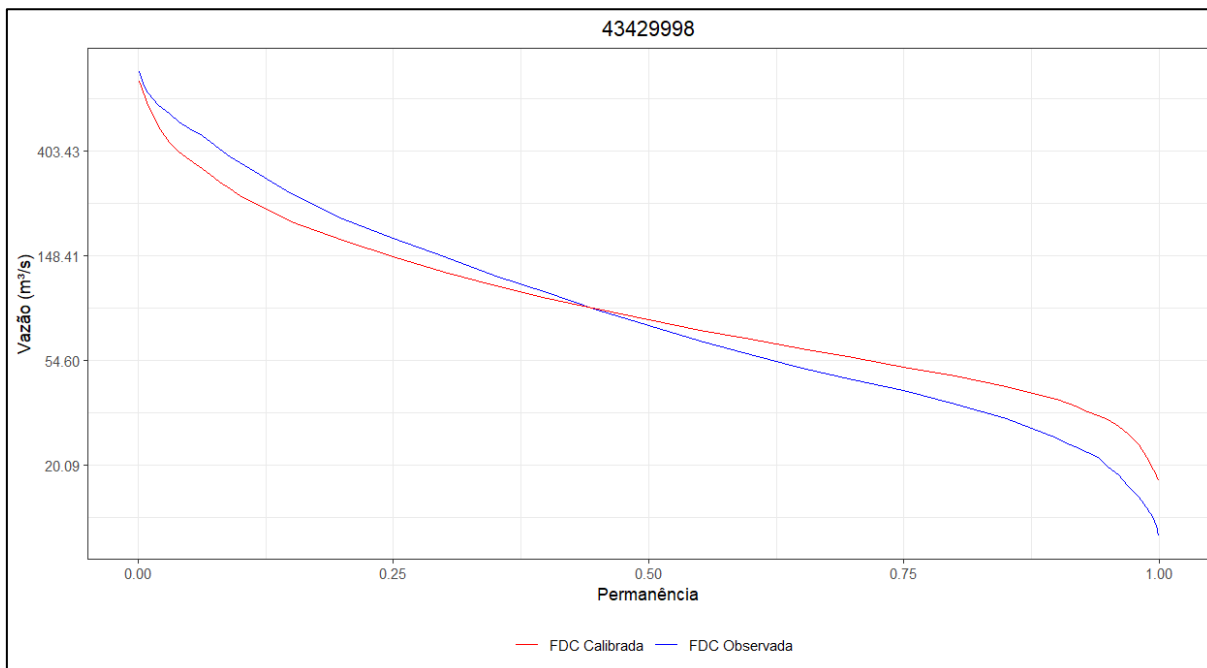


Figura B.30 - Desempenho da regressão funcional múltipla para a estação Vila Urucuia (43670000)

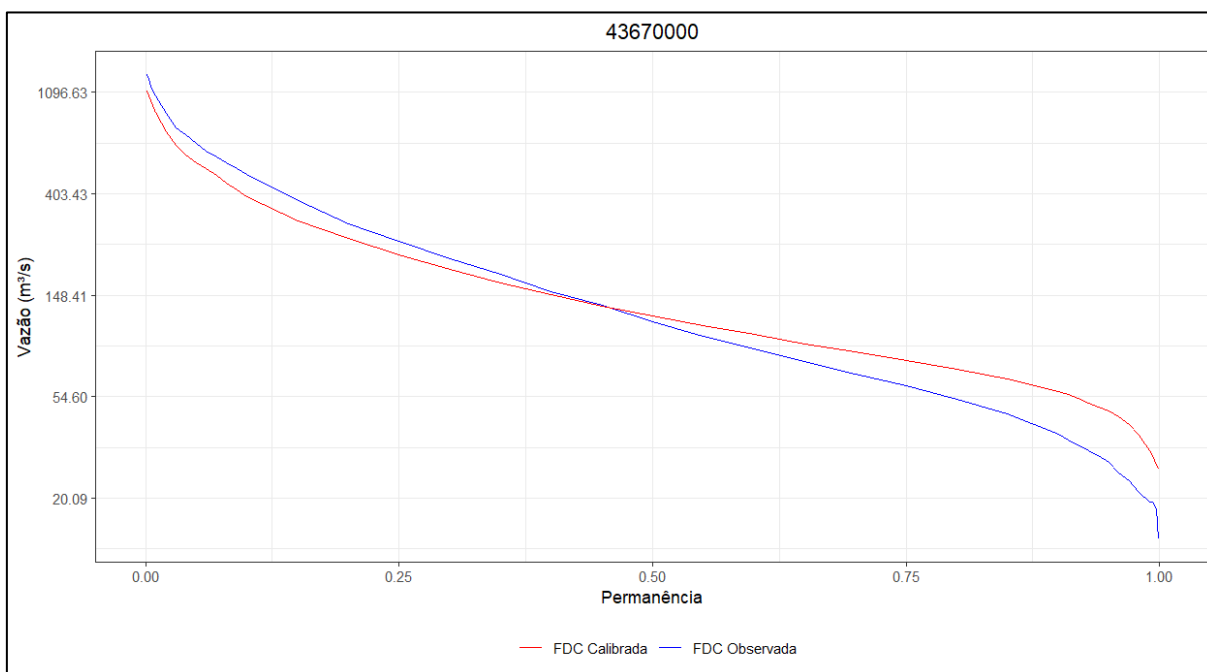


Figura B.31 - Desempenho da regressão funcional múltipla para a estação Usina do Pandeiros Montante (44250000)

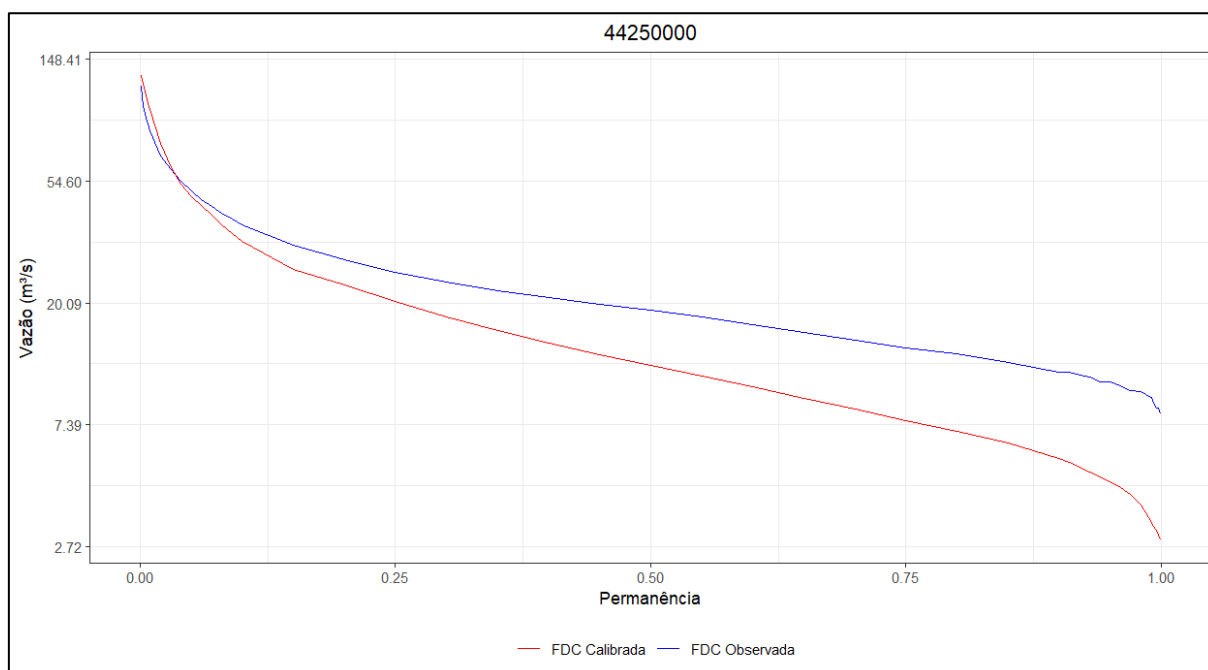


Figura B.32 - Desempenho da regressão funcional múltipla para a estação Pedras de Maria da Cruz (44290002)

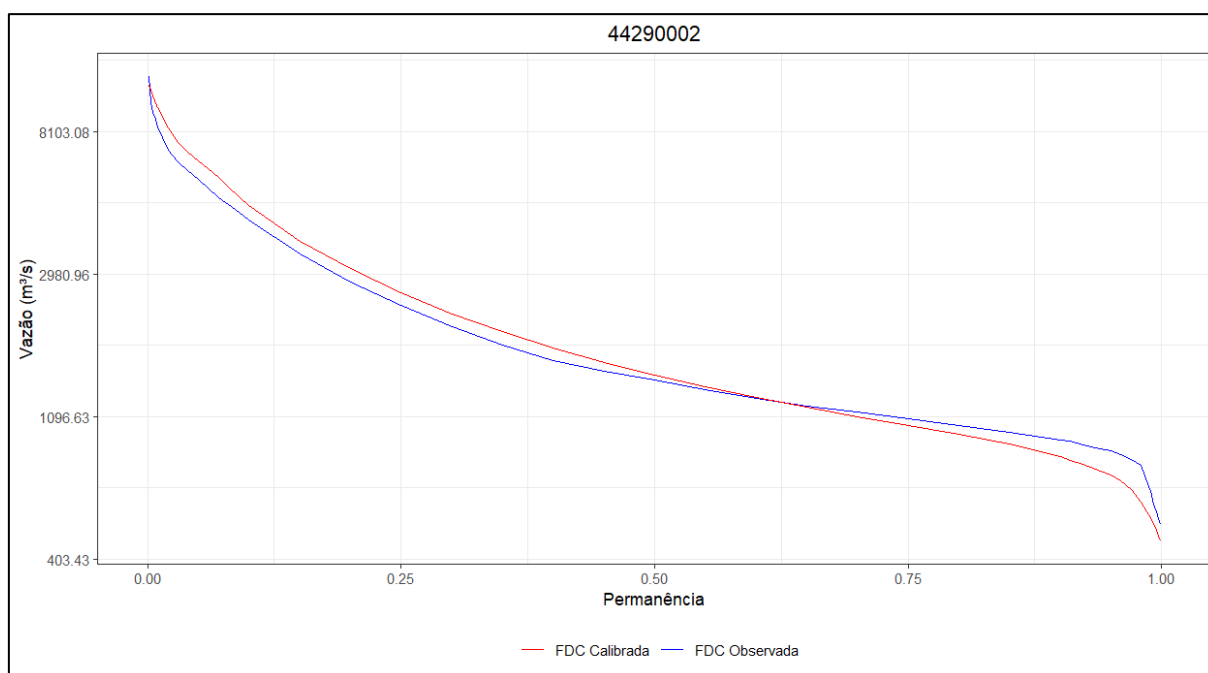


Figura B.33 - Desempenho da regressão funcional múltipla para a estação Fazenda Bom Retiro (44540000)

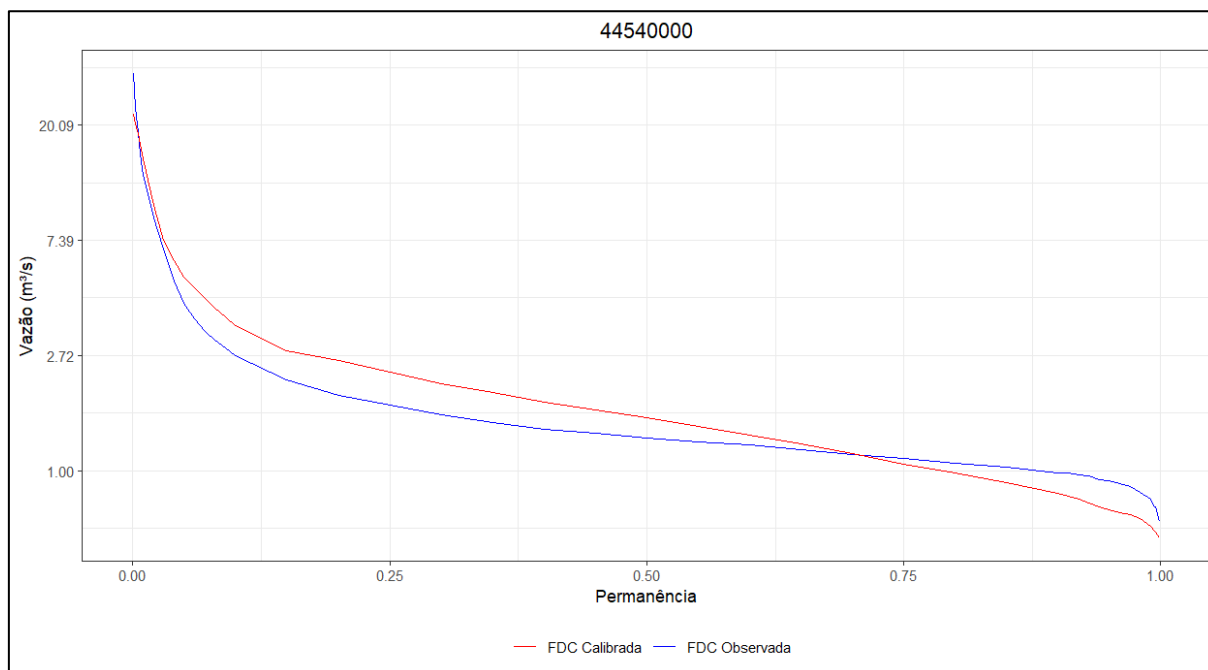
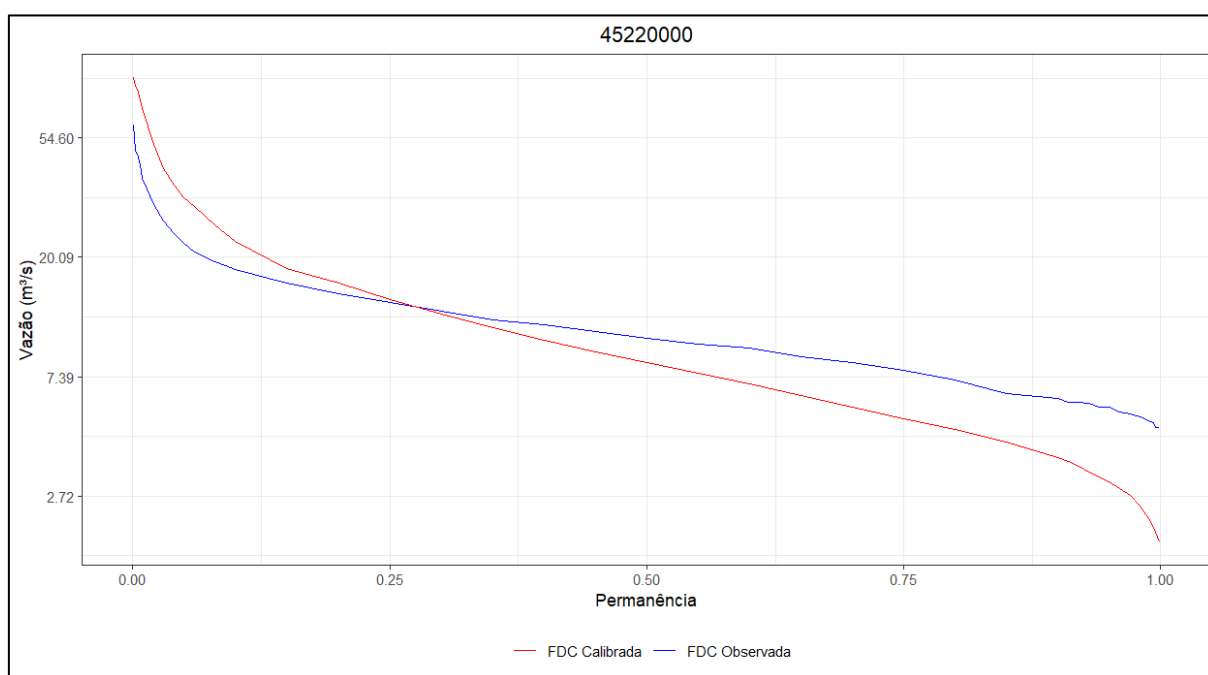


Figura B.34 - Desempenho da regressão funcional múltipla para a estação Capitânea (45220000)



APÊNDICE C – CURVAS DE PERMANÊNCIA REGIONALIZADAS PELA FMR (ETAPA DE VALIDAÇÃO)

Figura C.1 - Desempenho da regressão funcional múltipla para a estação Major Porto (41050000)

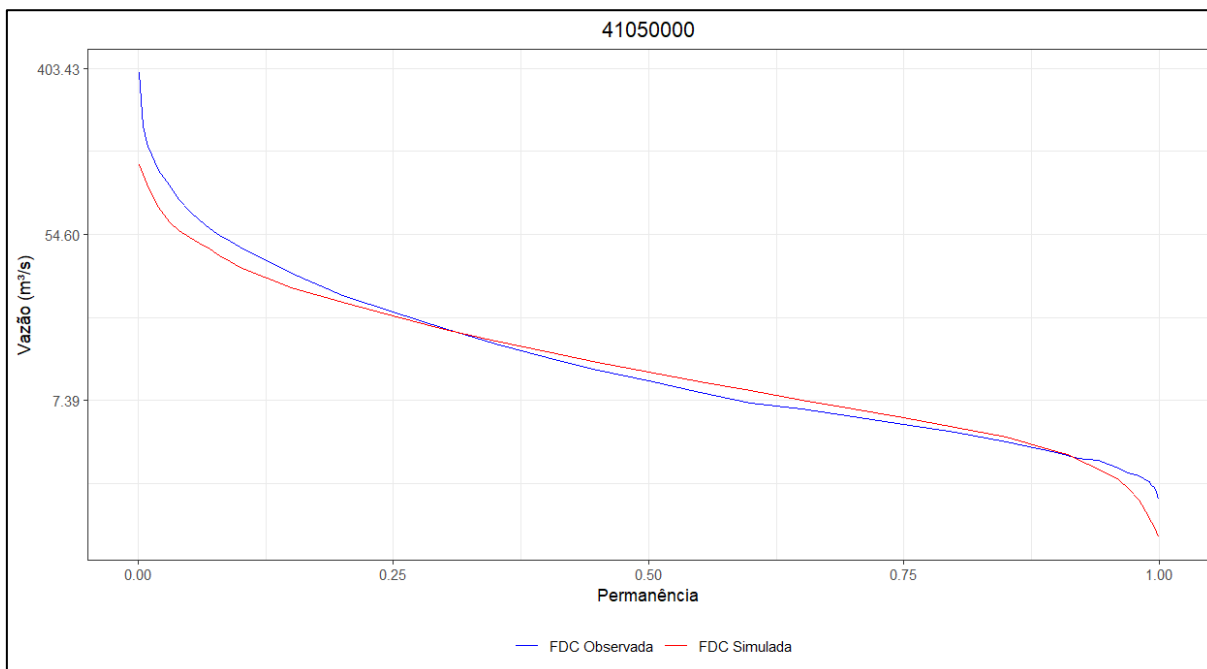


Figura C.2 - Desempenho da regressão funcional múltipla para a estação Pirapora Barreiro (41135000)

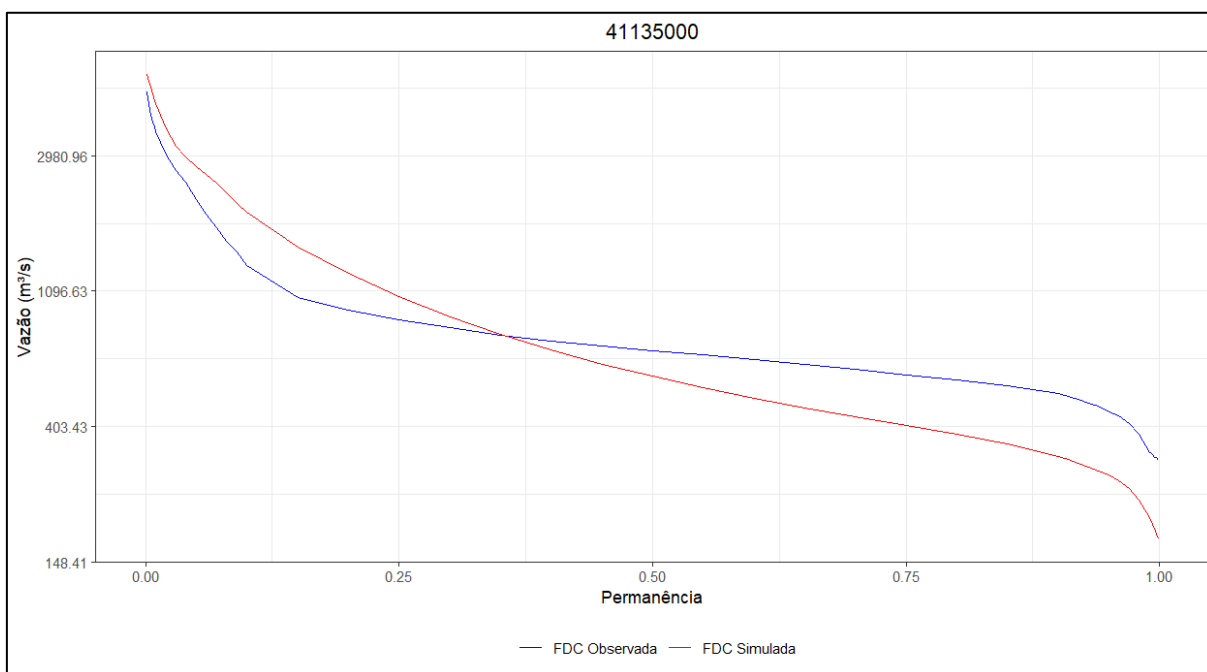


Figura C.3 - Desempenho da regressão funcional múltipla para a estação Represa Jusante (41440005)

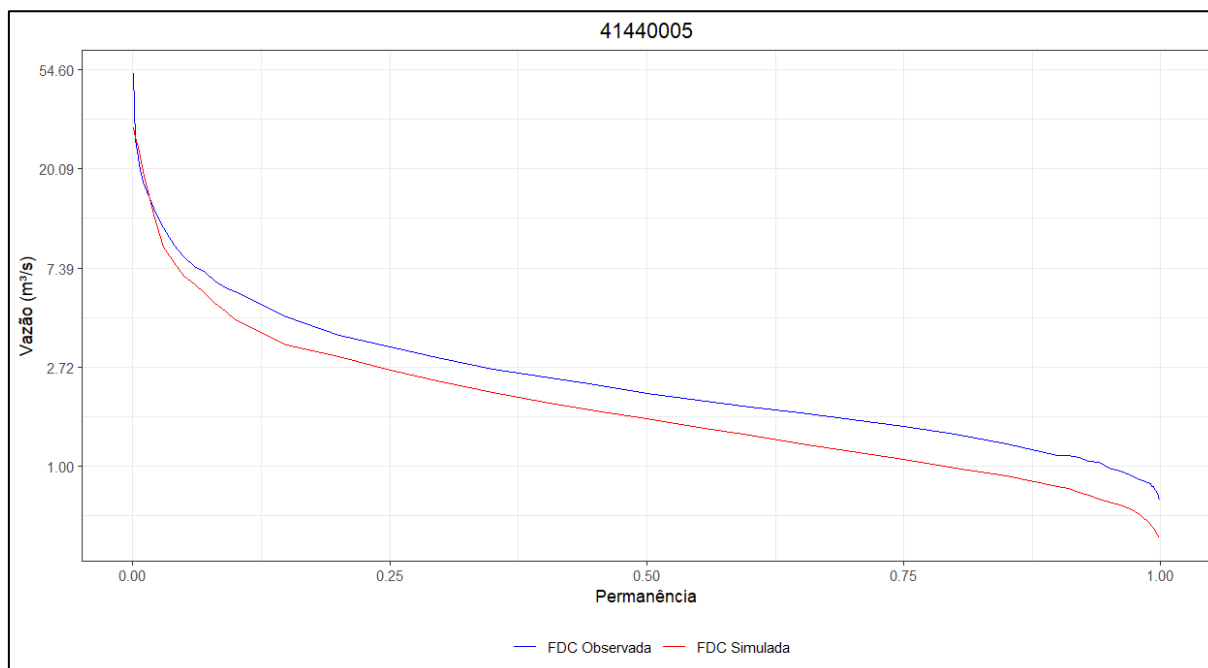


Figura C.4 - Desempenho da regressão funcional múltipla para a estação Ponte do Licínio Jusante (41650002)

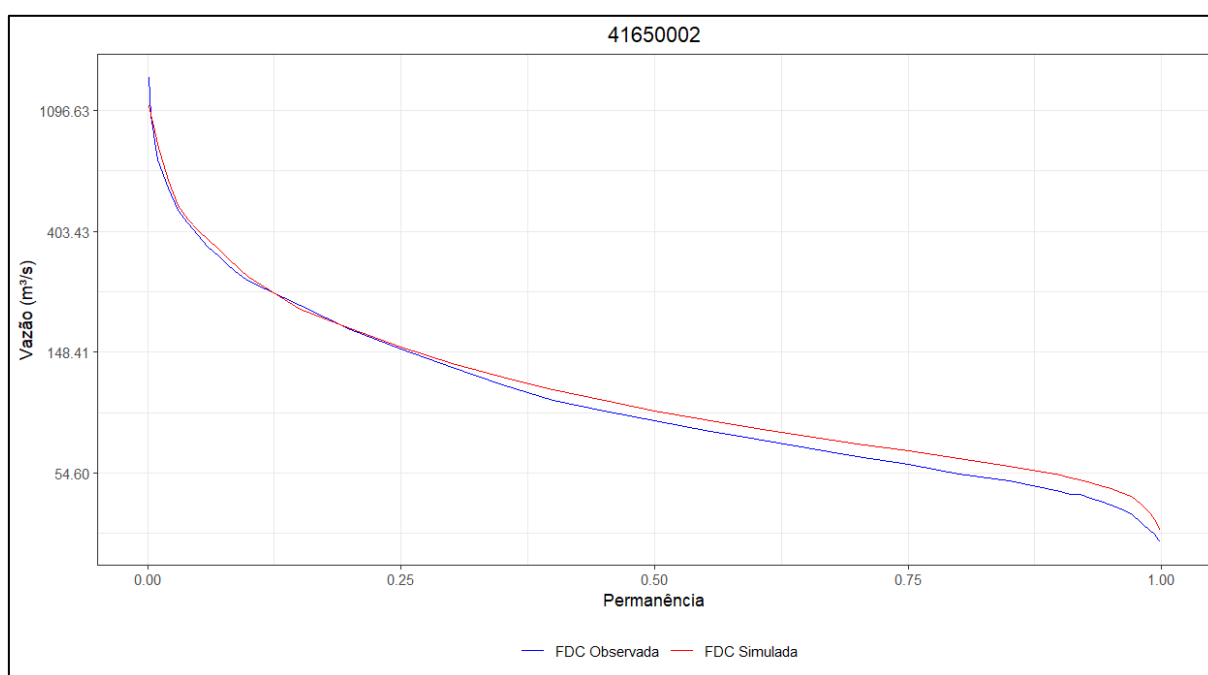


Figura C.5 - Desempenho da regressão funcional múltipla para a estação Santo Hipólito (41818000)

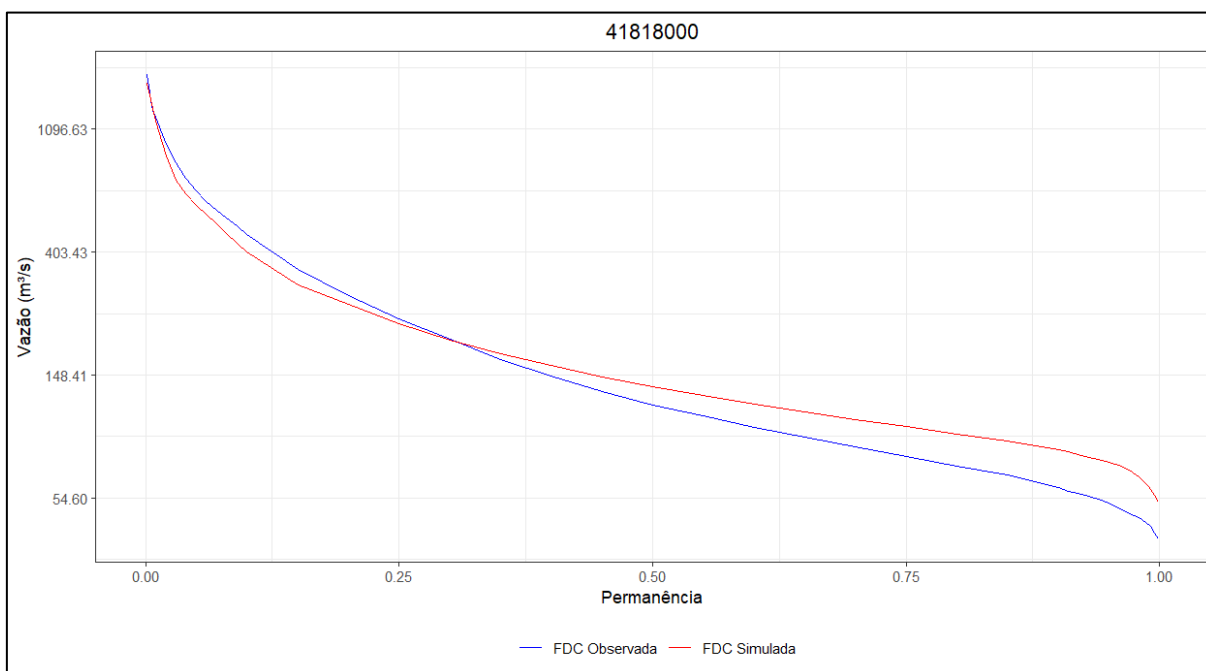


Figura C.6 - Desempenho da regressão funcional múltipla para a estação de Curimataí (41890000)

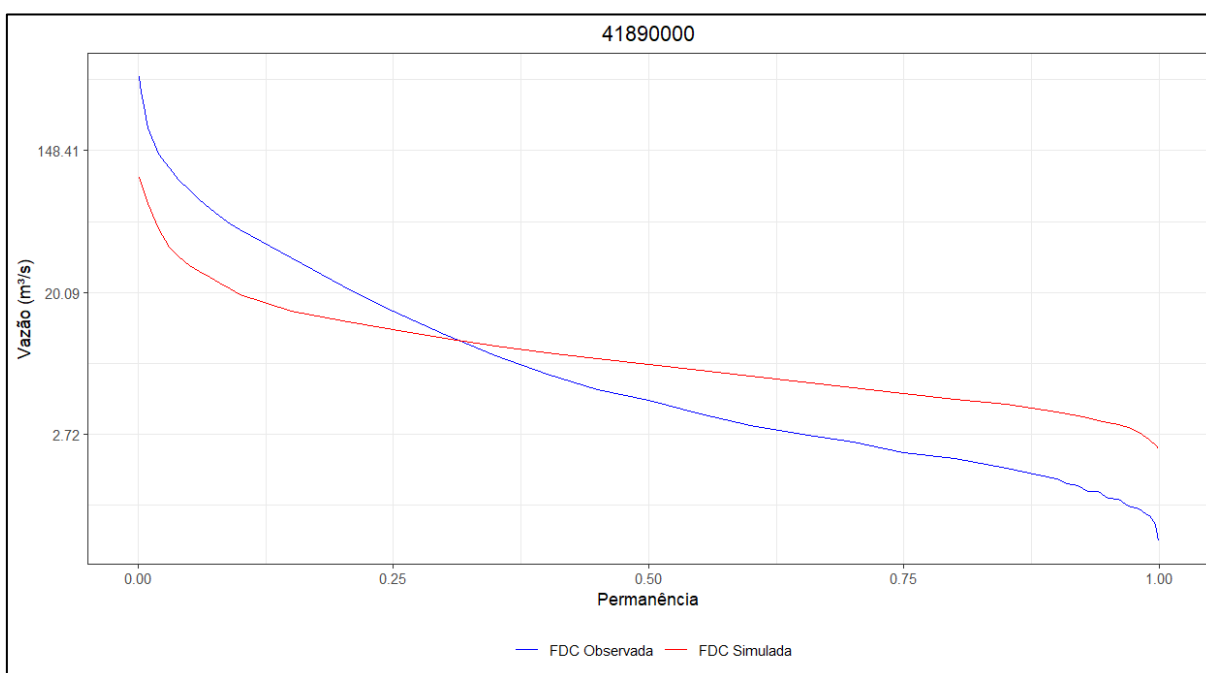


Figura C.7 - Desempenho da regressão funcional múltipla para a estação Ponte do Bicudo (41940000)

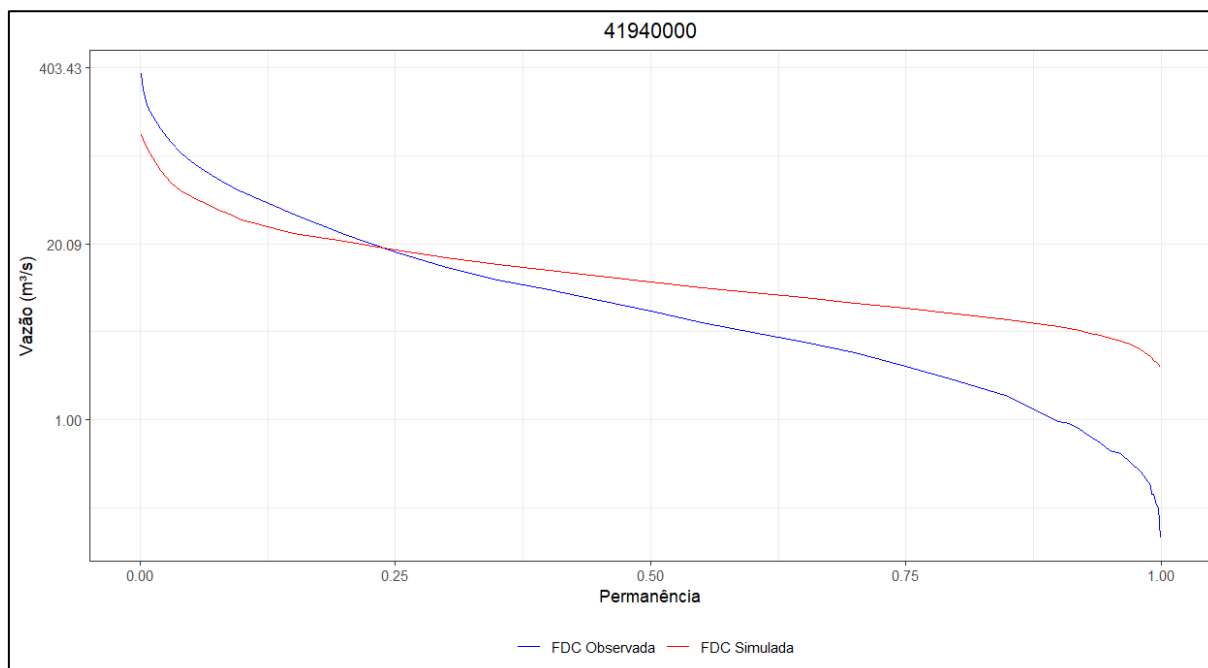


Figura C.8 - Desempenho da regressão funcional múltipla para a estação Santa Rosa (42395000)

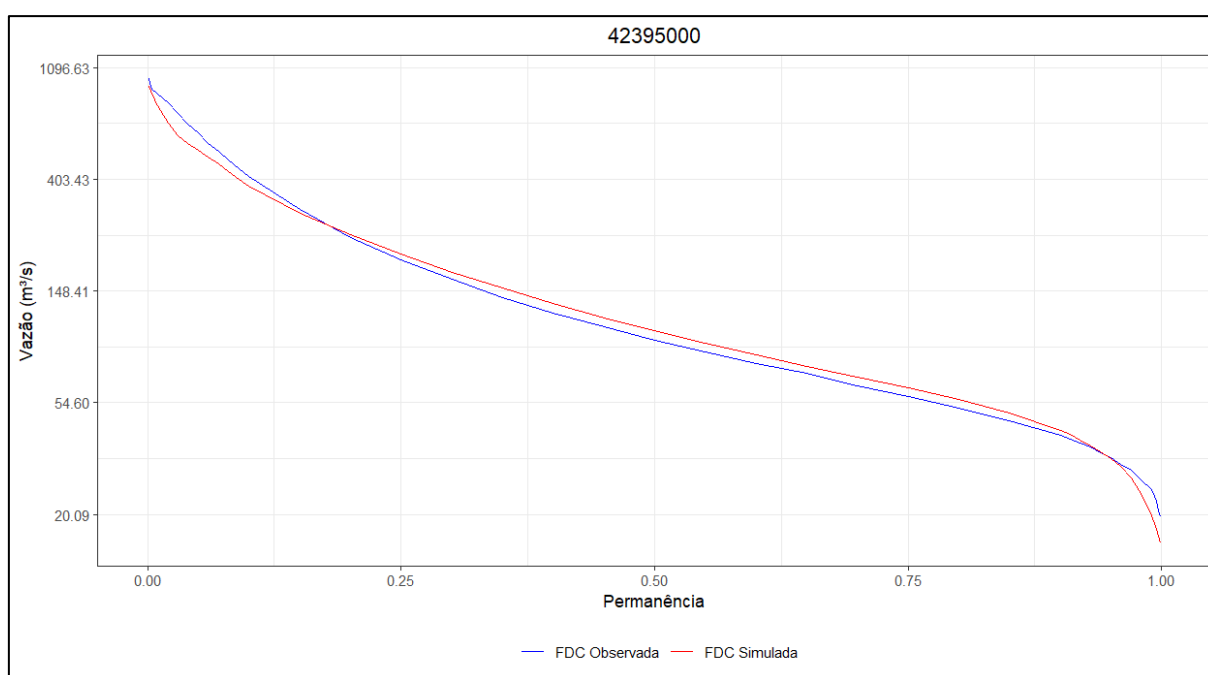


Figura C.9 - Desempenho da regressão funcional múltipla para a estação Cachoeira das Almas (42850000)

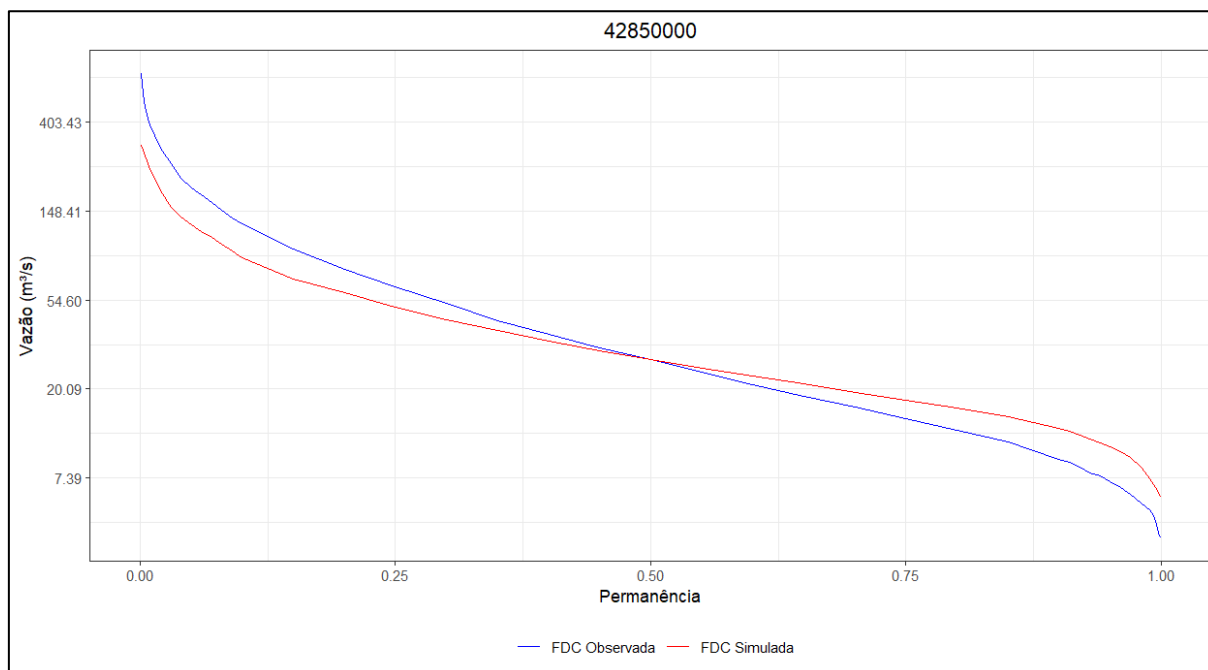


Figura C.10 - Desempenho da regressão funcional múltipla para a estação Santo Inácio (43880000)

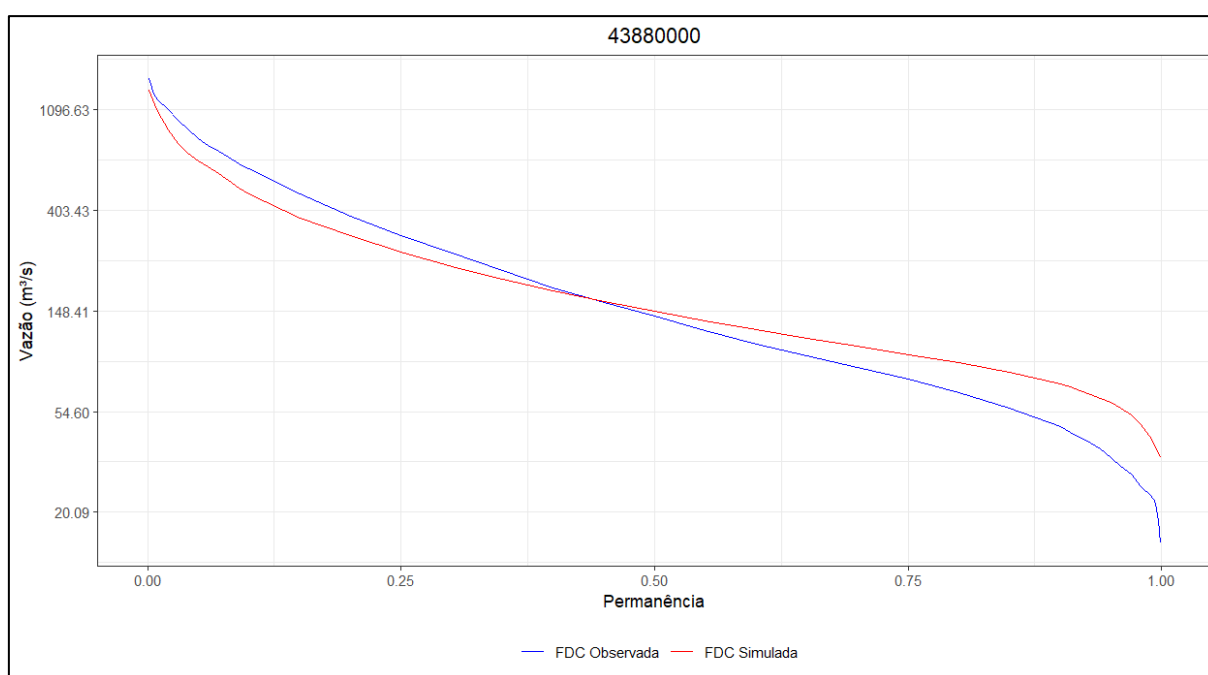


Figura C.11- Desempenho da regressão funcional múltipla para a estação São Francisco (44200000)

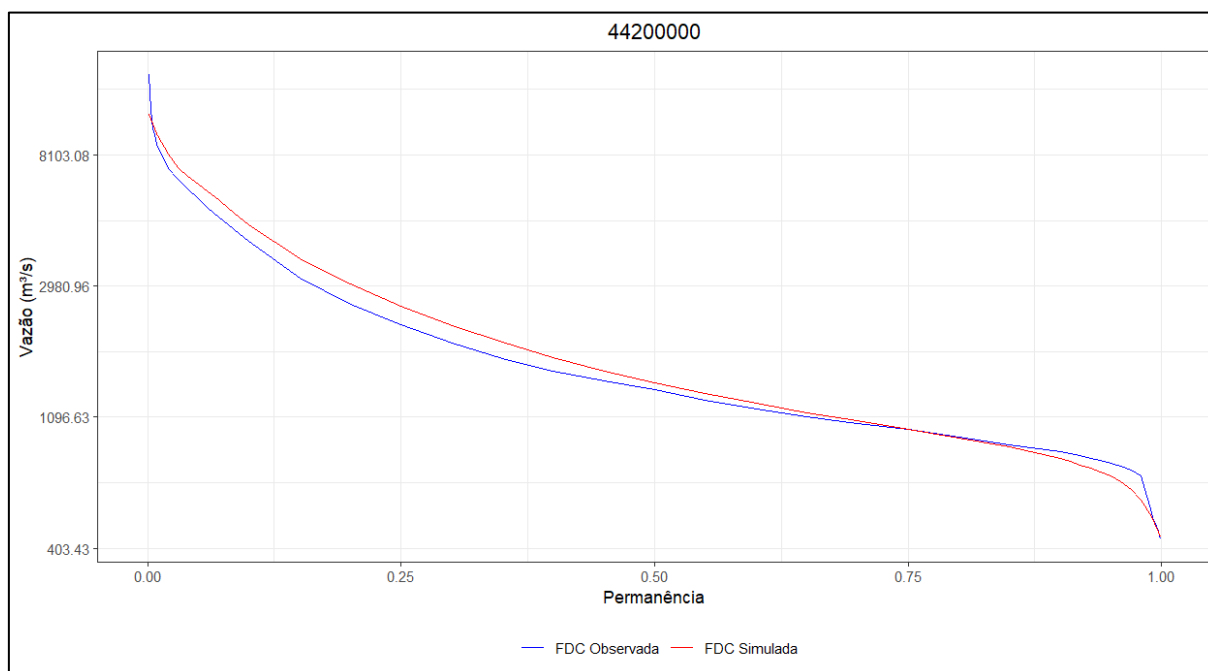


Figura C.12 - Desempenho da regressão funcional múltipla para a estação Manga (44500000)

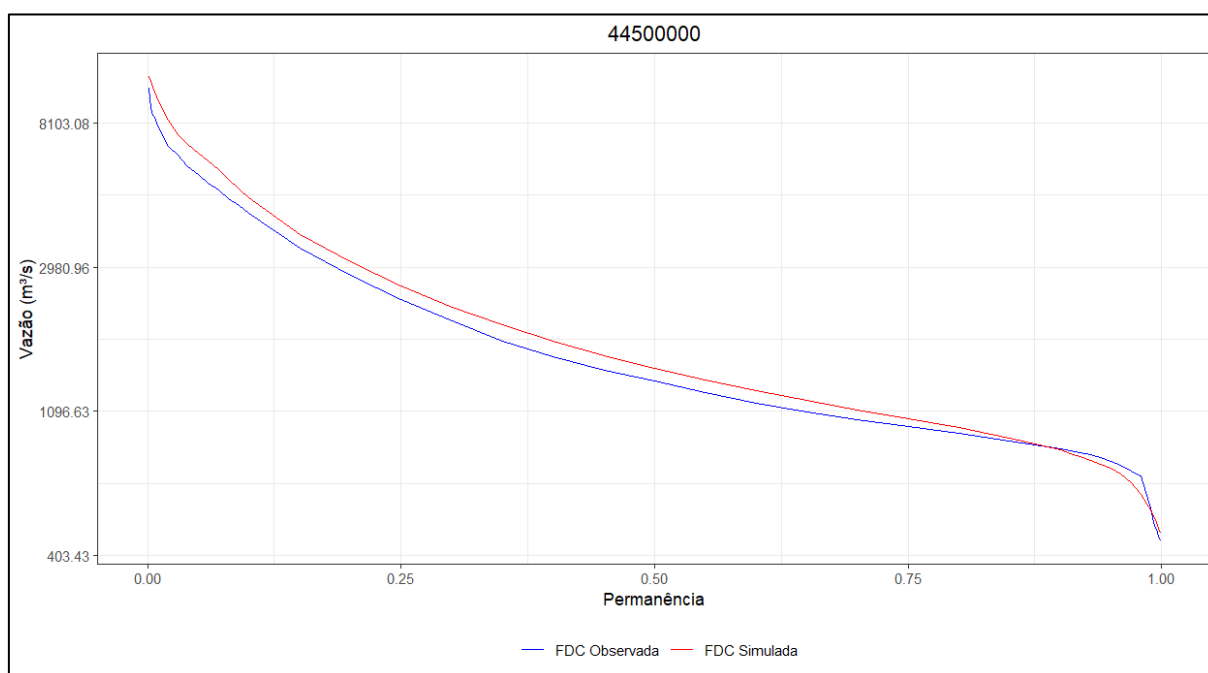
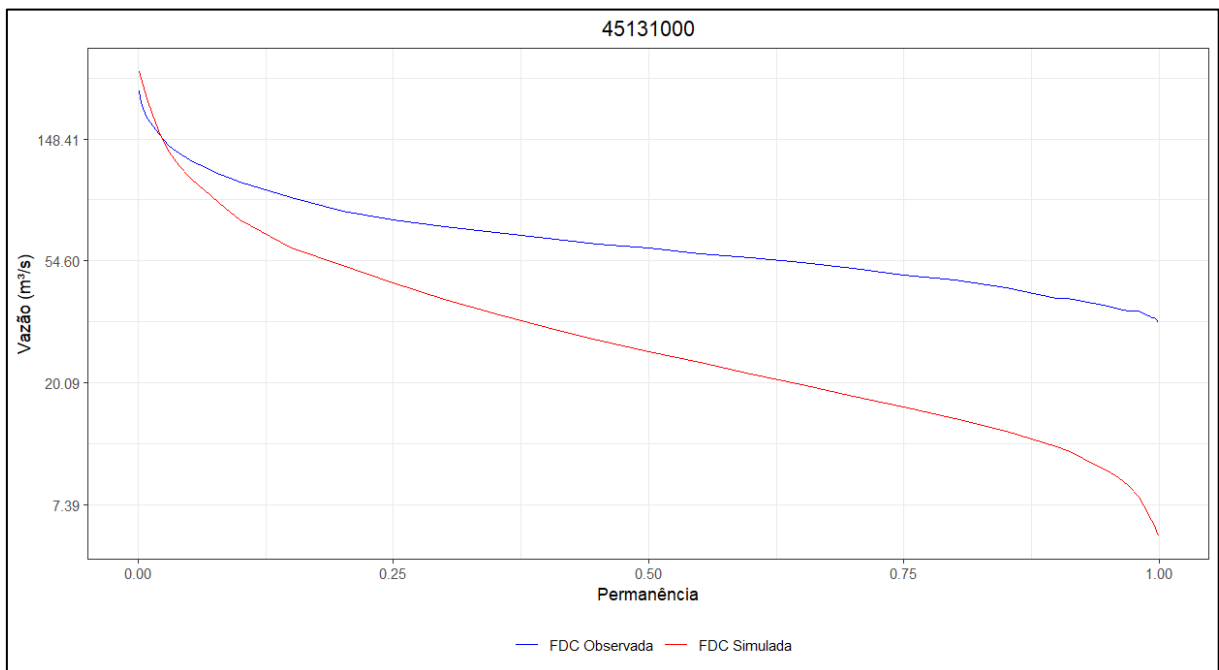


Figura C.13 - Desempenho da regressão funcional múltipla para a estação São Gonçalo (45131000)



[illegible]

Nota: Campos com “*” correspondem a valores positivos, mas inferiores à 10^{-2} .

APÊNDICE E – SÉRIES DE VAZÕES DIÁRIAS SIMULADAS

Figura D.1 - Desempenho da série de vazões simuladas na estação Major Porto (41050000)

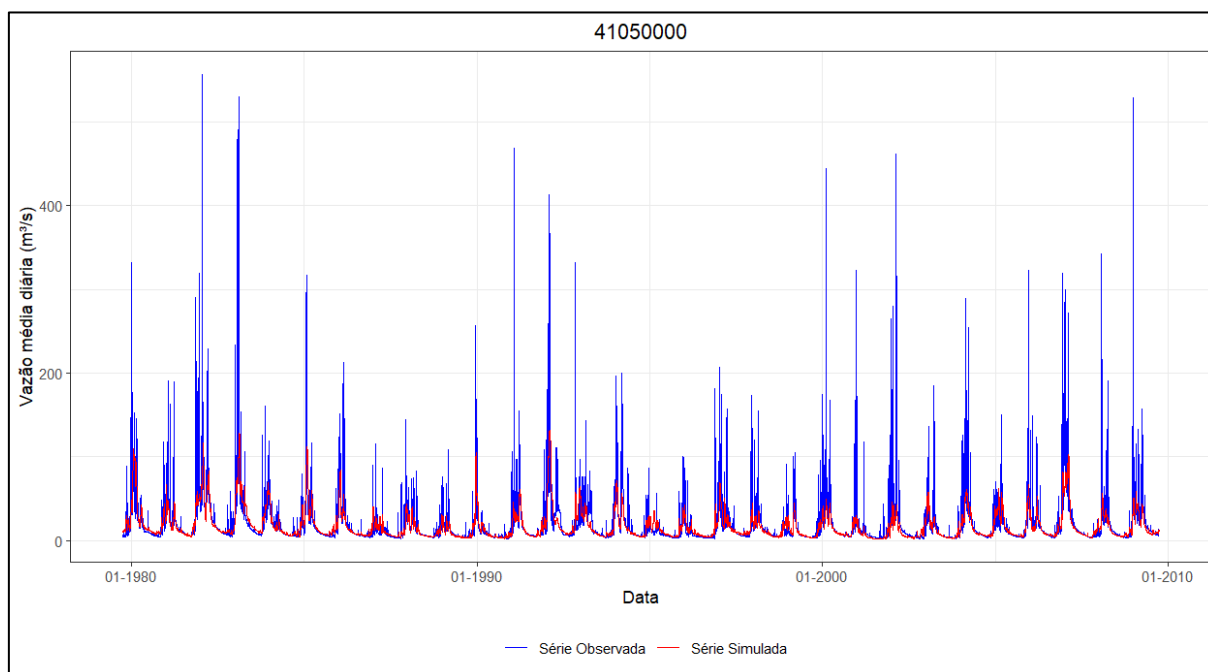


Figura D.2 - Desempenho da série de vazões simuladas na estação Pirapora Barreiro (41135000)

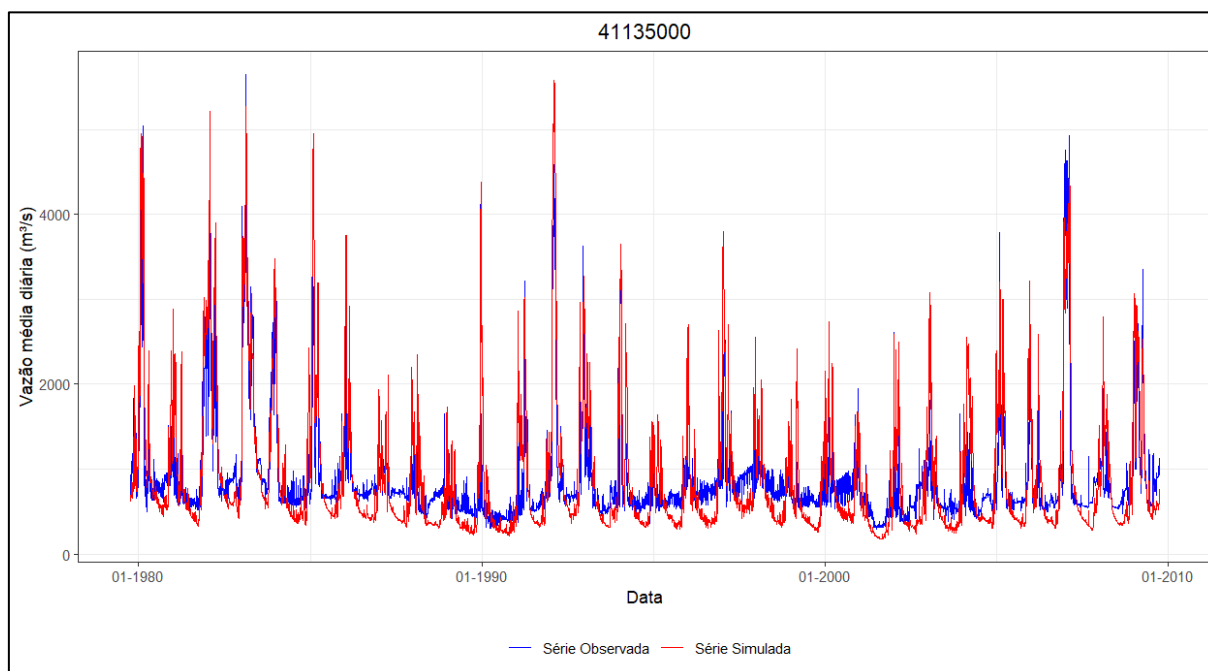


Figura D.3 - Desempenho da série de vazões simuladas na estação Represa Jusante (41440005)

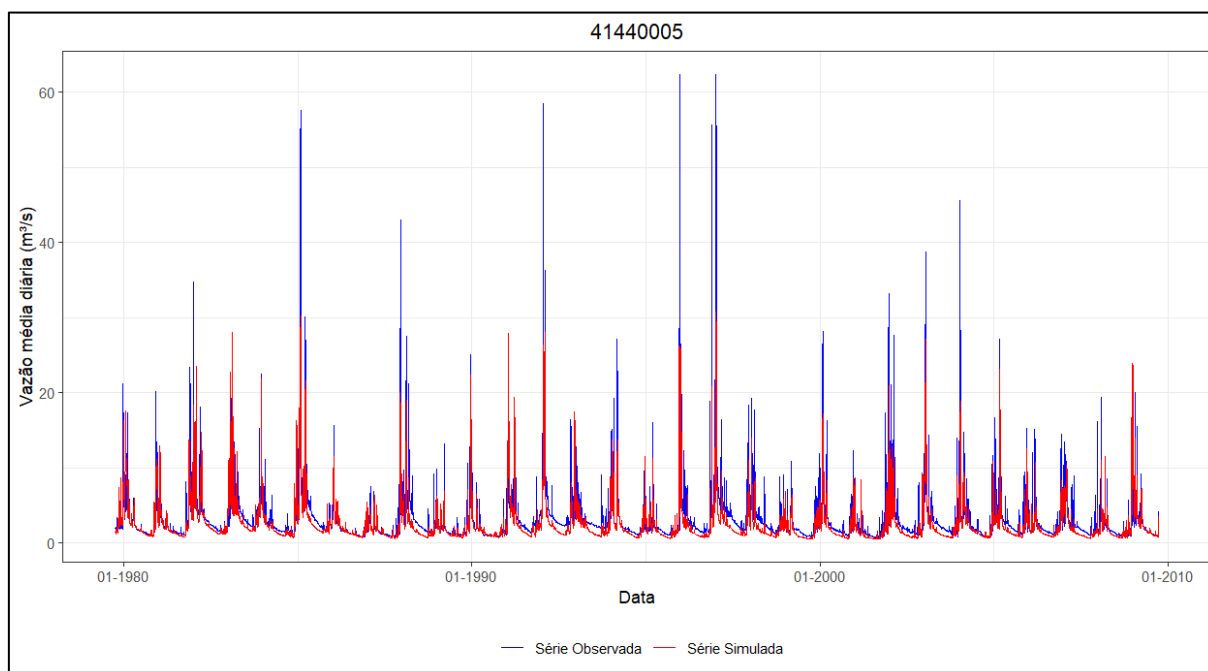


Figura D.4 - Desempenho da série de vazões simuladas na estação Ponte do Licínio Jusante (41650002)

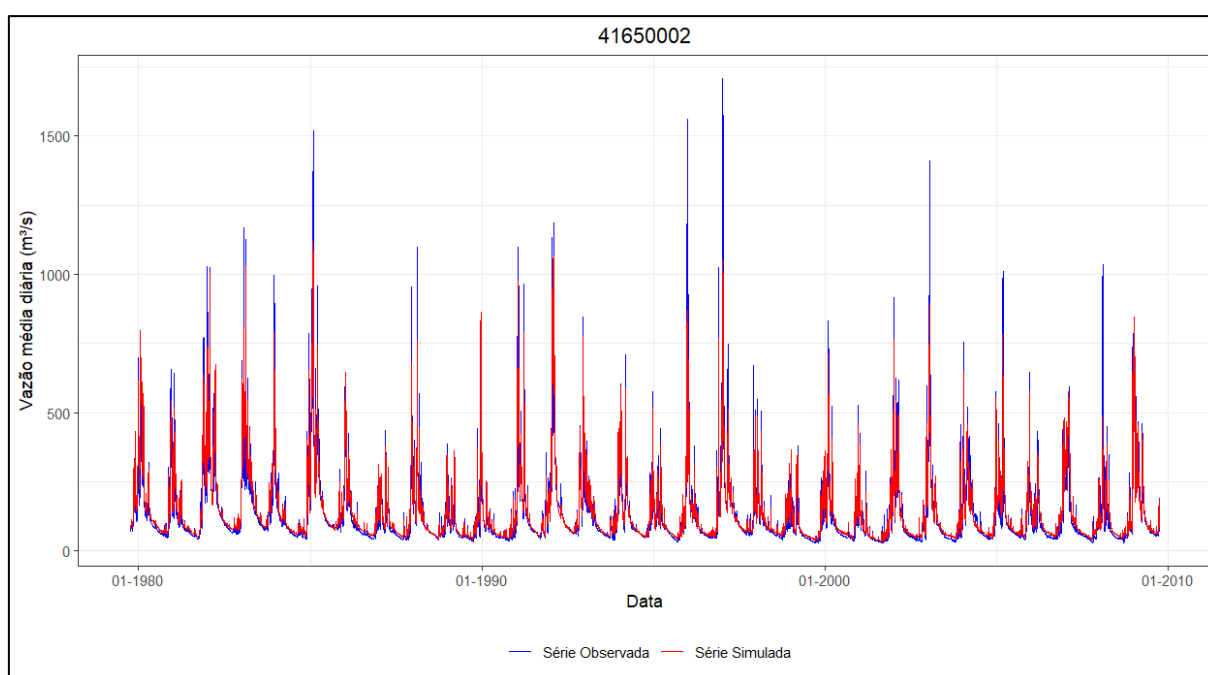


Figura D.5 - Desempenho da série de vazões simuladas na estação Santo Hipólito (41818000)

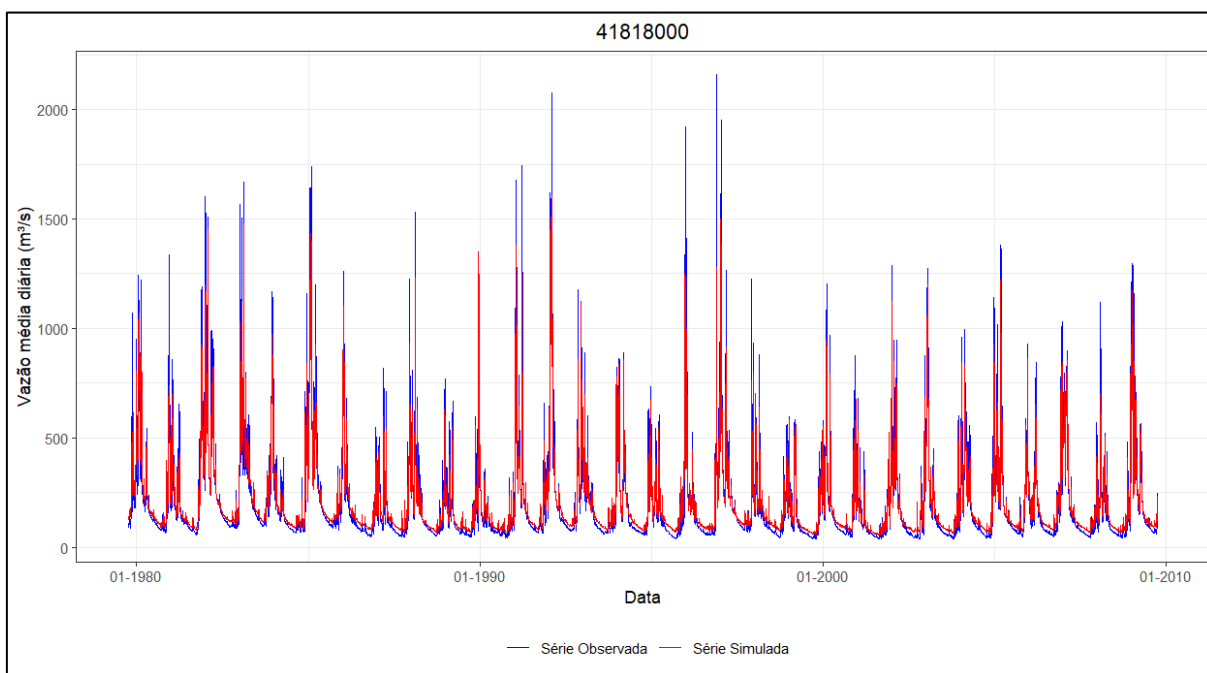


Figura D.6 - Desempenho da série de vazões simuladas na estação de Curimataí (41890000)

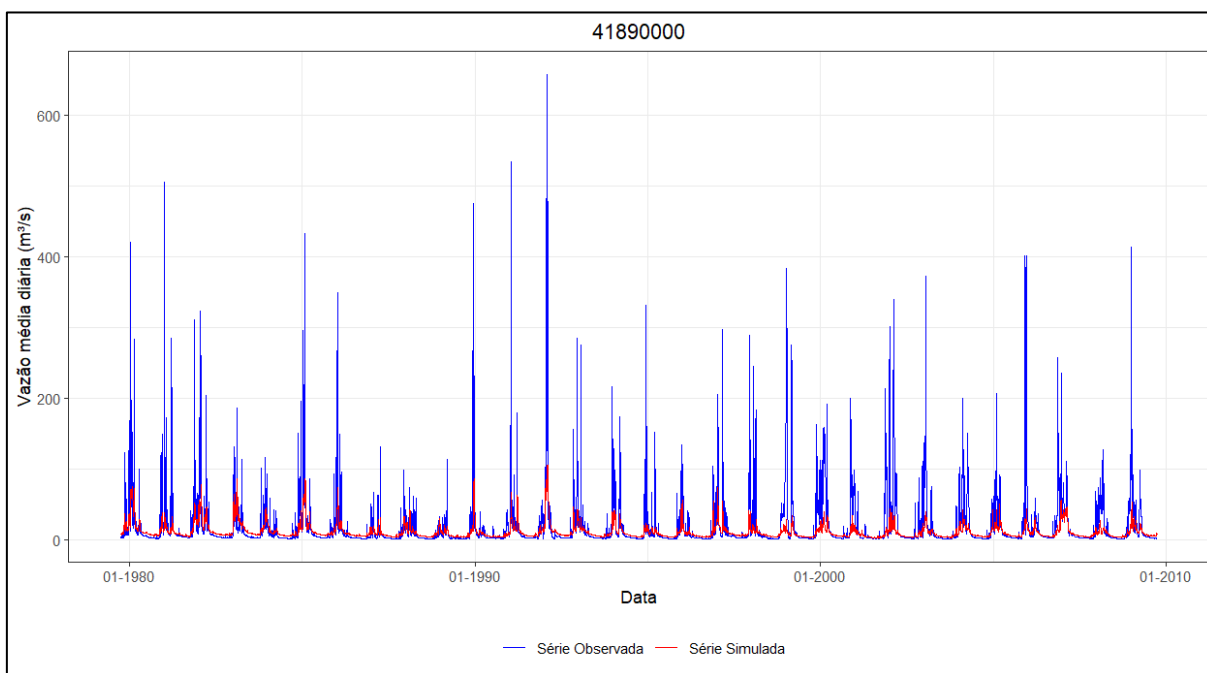


Figura D.7 - Desempenho da série de vazões simuladas na estação Ponte do Bicudo (41940000)

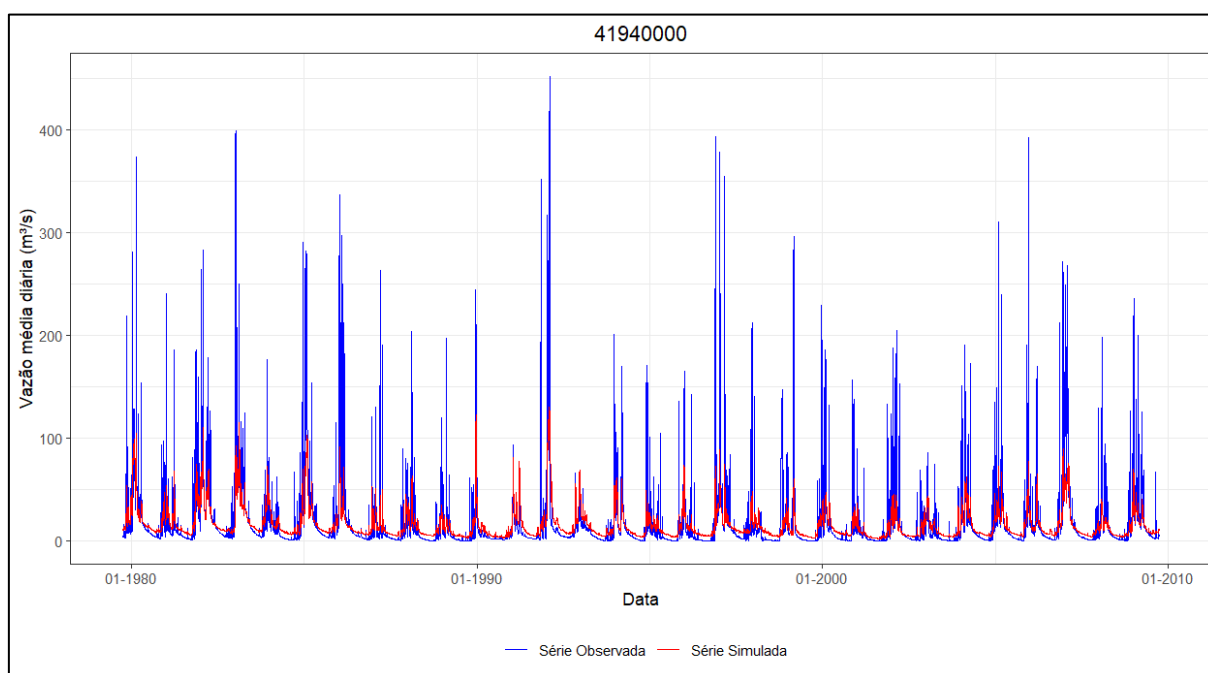


Figura D.8 - Desempenho da série de vazões simuladas na estação Santa Rosa (42395000)

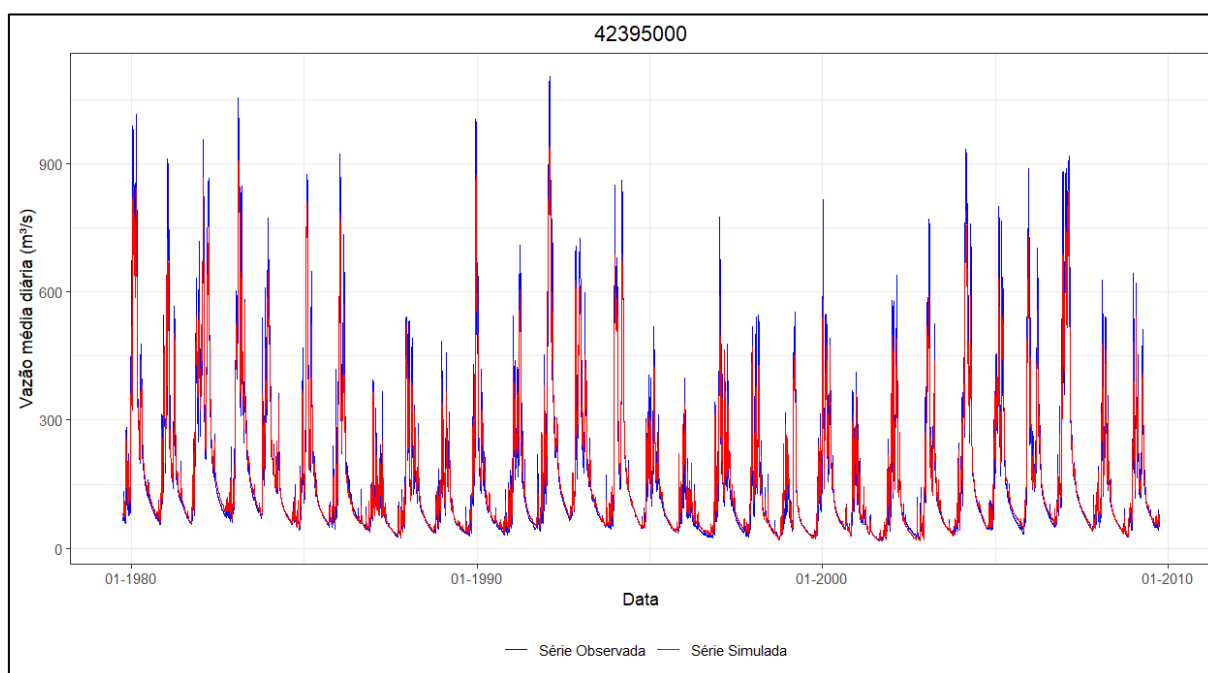


Figura D.9 - Desempenho da série de vazões simuladas na estação Cachoeira das Almas (42850000)

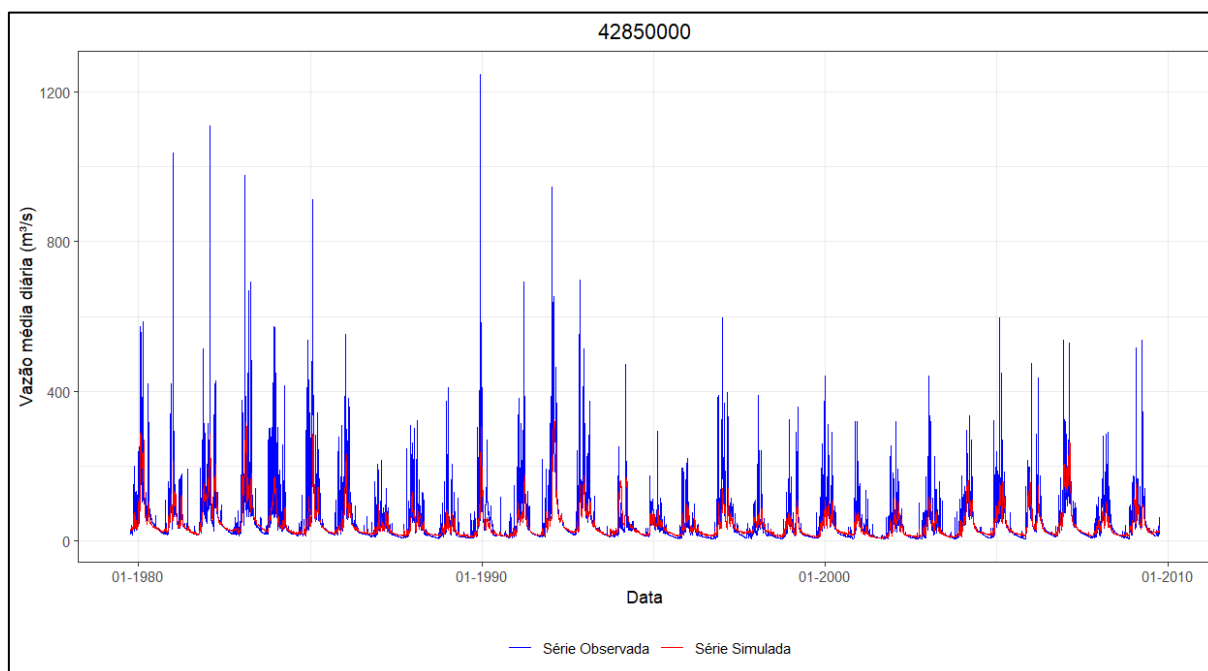


Figura D.10 - Desempenho da série de vazões simuladas na estação Santo Inácio (43880000)

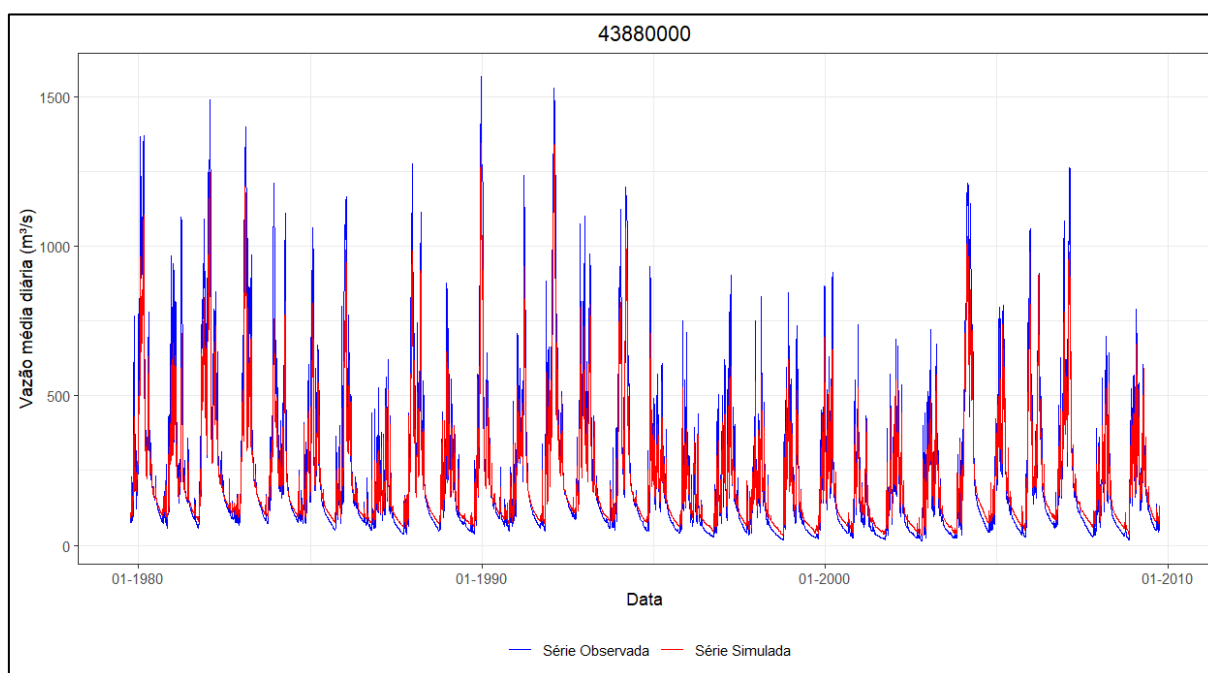


Figura D.11 - Desempenho da série de vazões simuladas na estação São Francisco (44200000)

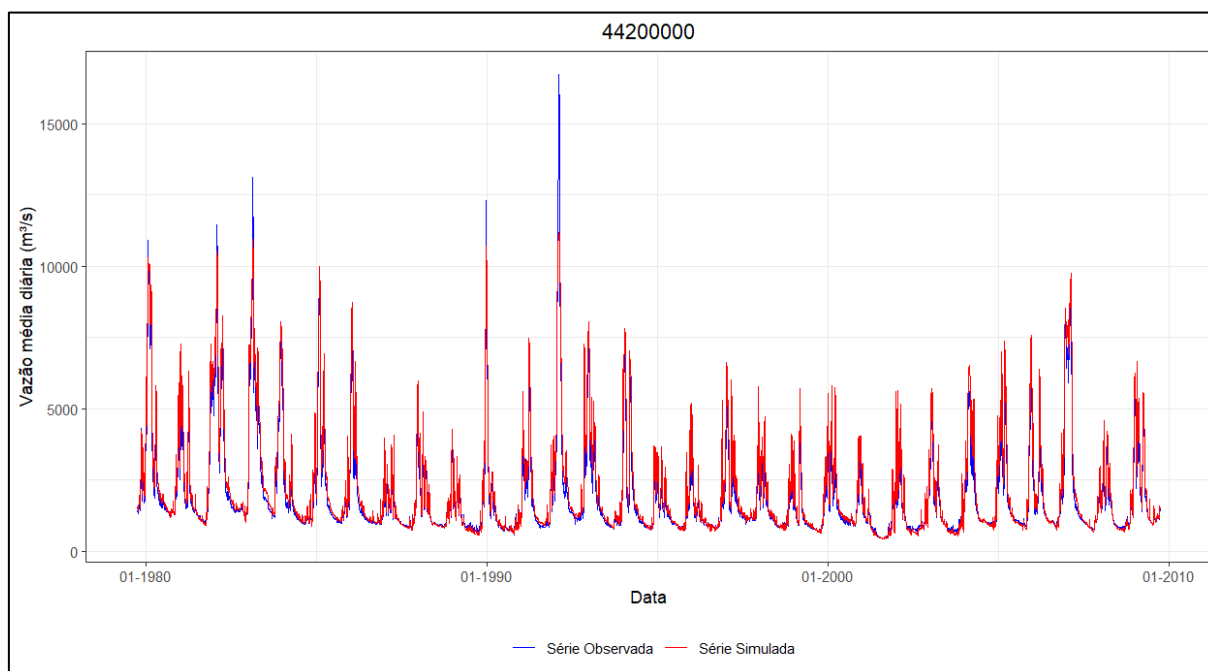


Figura D.12 - Desempenho da série de vazões simuladas na estação Manga (44500000)

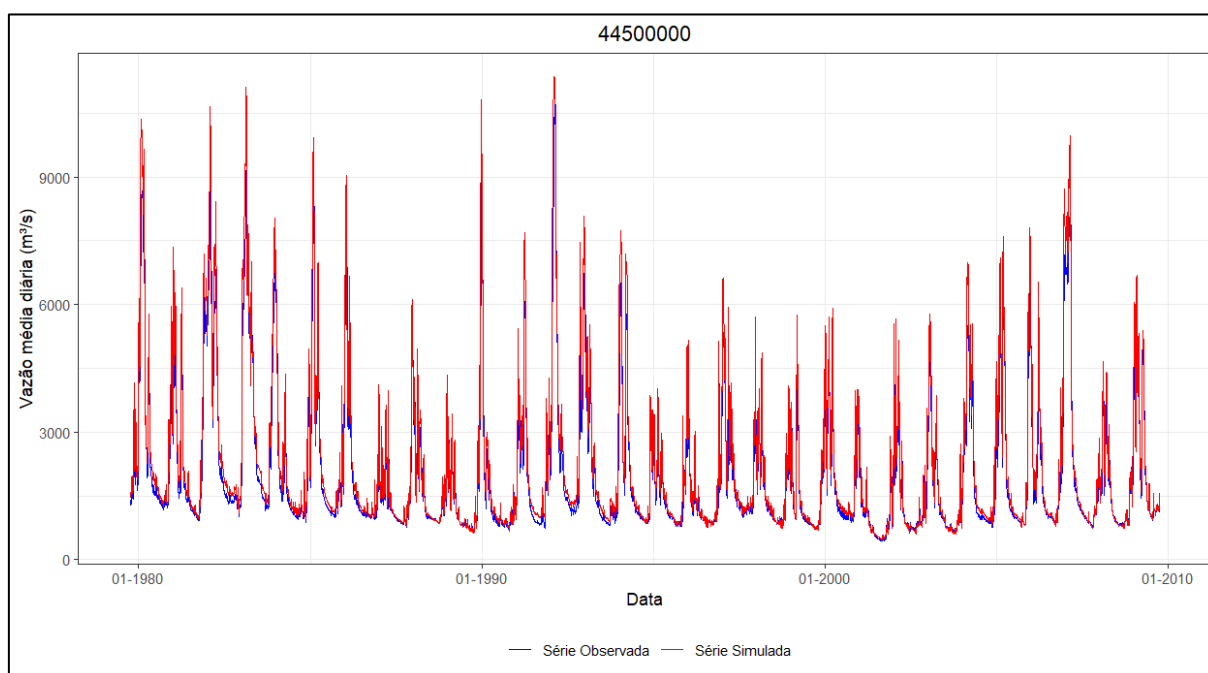


Figura D.13 - Desempenho da série de vazões simuladas na estação São Gonçalo (45131000)

