

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

Júlio César Lôbo Sampaio

**REGIONALIZAÇÃO BAYESIANA DE VAZÕES MÍNIMAS EM CONDIÇÕES DE
NÃO ESTACIONARIEDADE**

Belo Horizonte
2022

Júlio César Lôbo Sampaio

**REGIONALIZAÇÃO BAYESIANA DE VAZÕES MÍNIMAS EM CONDIÇÕES DE
NÃO ESTACIONARIEDADE**

Dissertação/tese 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/Doutor em Saneamento, Meio Ambiente e Recursos Hídricos.

Área de concentração: Recursos Hídricos

Linha de pesquisa: Modelagem de Processos Hidrológicos

Orientador: Veber Afonso Figueiredo Costa

Belo Horizonte
2022

S192r

Sampaio, Júlio César Lôbo.

Regionalização bayesiana de vazões mínimas em condições de não estacionariedade [recurso eletrônico] / Júlio César Lôbo Sampaio. - 2021. 1 recurso online (xiii, 155 f. : il., color.) : pdf.

Orientador: Veber Afonso Figueiredo Costa.

Mestrado (dissertação) - Universidade Federal de Minas Gerais, Escola de Engenharia.

Apêndices: f. 147-154.

Bibliografia: f. 132-146.

Exigências do sistema: Adobe Acrobat Reader.

1. Engenharia sanitária - Teses. 2. Recursos hídricos - Desenvolvimento - Teses. 3. Teoria bayesiana de decisão estatística - Teses. 4. Secas - 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]
COLEGIADO DO CURSO DE GRADUAÇÃO / PÓS-GRADUAÇÃO EM [SANEAMENTO, MEIO AMBIENTE E
RECURSOS HÍDRICOS]

FOLHA DE APROVAÇÃO

["REGIONALIZAÇÃO BAYESIANA DE VAZÕES MÍNIMAS EM CONDIÇÕES DE NÃO
ESTACIONARIEDADE"]

[Júlio César Lôbo Sampaio]

Dissertação de Mestrado] defendida e aprovada, no dia [05 de março de 2021], pela Banca Examinadora designada pelo [Colegiado do Programa de Pós-Graduação **EM SANEAMENTO, MEIO AMBIENTE E RECURSOS HÍDRICOS**] da Universidade Federal de Minas Gerais constituída pelos seguintes professores:

[Prof. Carlos Henrique Ribeiro Lima] - **Membro Externo**]

[UNB]

[Prof. Francisco Eustáquio Oliveira e Silva] - **Membro Interno**]

[UFMG]

[Prof. Luiz Rafael Palmier - **Membro Interno**]

[UFMG]

[Prof. Veber Afonso Figueiredo Costai - **Orientador**]

[UFMG]

APROVADA PELO COLEGIADO DO PPG SMARH

Sonaly Cristina Rezende Borges de Lima - Coordenadora

Belo Horizonte, 05 de março de 2021.



Documento assinado eletronicamente por **Carlos Henrique Ribeiro Lima, Usuário Externo**, em 16/03/2021, às 09:49, 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 **Veber Afonso Figueiredo Costa, Professor do Magistério Superior**, em 07/05/2021, às 15:06, 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 **Francisco Eustaquio Oliveira e Silva, Professor do Magistério Superior**, em 07/05/2021, às 16:28, 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 **0596298** e o código CRC **6D0FDB16**.

AGRADECIMENTOS

Agradeço ao professor Veber Costa, pela orientação e dedicação nesses últimos dois anos. Em especial, sua visão crítica e expertise foram imprescindíveis para o desenvolvimento da pesquisa e para minha formação acadêmica.

Aos professores do departamento EHR, que contribuíram com as aulas e trocas de conhecimento.

À fundação CAPES, pelo apoio financeiro durante o desenvolvimento desta dissertação.

Aos meus pais, Marcos e Isabel, e meu irmão Caio, pelo amor incondicional e constante incentivo desde sempre.

Aos meus amigos do PPG-SMARH, pelos momentos de descontração e pelo companheirismo durante essa jornada.

À Universidade Federal de Minas Gerais, por ter me abrigado com tanto carinho e emoção desde 2013.

RESUMO

A água tem se tornado cada vez mais um ativo de interesse para a sociedade devido ao aumento da demanda hídrica. A gestão de recursos hídricos, por meio de estabelecimentos de outorga, torna-se uma ferramenta importante para evitar conflitos e garantir o uso sustentável da água. O sistema de monitoramento de vazões é limitado a poucos postos de medição e, portanto, hidrólogos frequentemente recorrem a técnicas de regionalização para fazer estimativas em bacias não monitoradas. Nesse contexto, estimativas regionais de vazões mínimas de referência, indicadores para outorgas, e sua incerteza preditiva tornam-se objeto de interesse.

Este estudo visou desenvolver um modelo hierárquico Bayesiano para regionalização de vazões mínimas condicionado à covariáveis temporais, no caso, à temperatura da superfície do mar (TSM) e avaliar os benefícios dessa abordagem frente a um modelo estacionário de referência. Partindo de um modelo estacionário mais simples, sob a hipótese de independência espacial dos parâmetros e de independência condicional das observações, a representação da variabilidade espacial do modelo foi inicialmente capturada por meio de covariáveis espaciais. Posteriormente, foram investigados os possíveis ganhos na predição por meio do aprofundamento da descrição espacial sob os dados e sob o processo. Em seguida, a TSM foi introduzida no nível do processo do modelo sob a forma de um índice climático customizado, inferido a partir de um campo de valores de TSM. Foram elaborados modelos que avaliassem a inclusão do índice climático customizado nos parâmetros da distribuição de probabilidade empregada.

O modelo foi aplicado na bacia do rio Itajaí-Açu (SC) e na bacia do rio Doce (MG/ES). Os resultados mostraram que o modelo não estacionário apresentou melhor desempenho, em termos do critério DIC, do que o modelo estacionário de referência e que os quantis estimados (como a $Q_{7,10}$) são fortemente influenciados pela variabilidade climática. Além disso, uma descrição mais complexa da dependência espacial do processo traz benefícios para a predição em regiões densamente monitoradas, enquanto a interdependência nas observações dos dados, quando considerada, pode também trazer benefícios para a predição.

Palavras-chave: Modelo hierárquico Bayesiano, regionalização, vazão mínima, não estacionariedade

ABSTRACT

Water has increasingly become an asset of interest to society due to the increase in water demand. The management of water resources, by means of grant establishments, becomes an important tool to avoid conflicts and guarantee the sustainable use of water. The streamflow monitoring system is limited to a few measurement gages and, therefore, hydrologists often use regionalization techniques to make estimates in unmonitored basins. In this context, regional estimates of low flows, grant indicators, and their predictive uncertainty become object of interest.

This study aimed to develop a Bayesian hierarchical model for regionalization of low flows conditioned to temporal covariables, in this case, the sea surface temperature (SST) and to evaluate the benefits of this approach compared to a stationary reference model. Starting from a simpler stationary model, under the hypothesis of spatial independence of the parameters and conditional independence of the observations, the representation of the spatial variability of the model was initially captured by means of spatial covariates. Subsequently, possible gains in prediction were investigated by further describing the spatial description under the data and under the process. Then, the SST was introduced at the process level of the model in the form of a customized climate index, inferred from a field of SST values. Models were developed to assess the inclusion of the customized climate index in the parameters of the probability distribution employed.

The model was applied in the Itajaí-Açu river basin (SC) and in the Doce river basin (MG / ES). The results showed that the non-stationary model performed better, in terms of the DIC criterion, than the reference stationary model and that the estimated quantiles (such as $Q_{7,10}$) are strongly influenced by climatic variability. In addition, a more complex description of the spatial dependence of the process brings benefits to prediction in densely monitored regions, while interdependence in data observations, when considered, can also bring benefits to prediction.

Keywords: Bayesian hierarchical model, regionalization, low flow, non-stationarity

LISTA DE FIGURAS

Figura 3.1 - Fluxo de trabalho para a regionalização por métodos regressivos	20
Figura 3.2 - Efeito do tamanho da bacia hidrográfica (esquerda) e do aninhamento de bacias (direita). Pesos da Krigagem estão em vermelho e os pontos pretos juntos aos rótulos x0 em vermelho indicam os centroides das bacias hidrográficas.....	28
Figura 3.3 - Metodologia da AFR baseada na estimação de vazão e da AFR tradicional.	30
Figura 3.4 - Consequências do sobreajuste e superajuste de um modelo no erro preditivo dos dados de validação e calibração.....	33
Figura 3.5 - Funções de correlação cruzada ajustadas globalmente e para bacias aninhadas e não aninhadas para dados de vazões máximas (T&S: função de correlação cruzada de Tasker e Stedinger, 1989).....	38
Figura 3.6 - Tentativas de inferência de tendências para diferentes tamanhos de amostras de um processo estocástico hipotético.....	39
Figura 3.7 - Três índices utilizados para capturar a variabilidade do ENSO	41
Figura 4.1 - Mapa de localização das bacias em estudo.....	44
Figura 4.2 - Representação esquemática do mecanismo forçador das ondas de calor marítima do oeste do Atlântico Sul.....	47
Figura 4.3 - Esquema do modelo M3/M4 onde são apresentados os níveis dos dados e do processo.	61
Figura 4.4 - Correlação da TSM do mês de janeiro entre na coordenada 0°, 120°O e as 9.646 coordenadas do banco de dados.....	64
Figura 4.5 - ACF da série mensal de TSM na coordenada 0°, 120°O. Linhas tracejadas vermelhas indicam o intervalo com 95% de confiança	65
Figura 5.1 - Localização dos postos fluviométricos da bacia do rio Doce. Os rótulos dos postos fluviométricos correspondem ao ID de cada posto	73
Figura 5.2 - Disponibilidade de dados da bacia do rio Doce. Anos em preto indicam o registro de vazão nula. A escala de cores das séries representa valores normalizados da Q7 mínima anual ($Zx = x - \mu\sigma$).....	74
Figura 5.3 - Localização dos postos fluviométricos da bacia do rio Itajaí-Açu. Os rótulos dos postos fluviométricos correspondem ao ID de cada posto	75
Figura 5.4 - Disponibilidade de dados na bacia do rio Itajaí-Açu. . A escala de cores das séries representa valores normalizados da Q7 mínima anual ($Zx = x - \mu\sigma$)	76
Figura 5.5 - Diagrama dos quocientes de momentos-L dos postos fluviométricos de calibração. O polígono que representa a distribuição LP3 abrange os possíveis pares ordenados da distribuição com o módulo da assimetria inferior a 1,414 (GRIFFIS; STEDINGER, 2007).....	77
Figura 5.6 - Resultado do teste de aderência PPCC com 5% de significância. H0 e H1 indicam a hipótese nula e a hipótese alternativa do teste. SS é uma abreviação para “Sem Solução”, utilizada para quando o método dos momentos-L falhou em gerar estimativas dos parâmetros da distribuição	78
Figura 5.7 - Medidas de aderência das distribuições candidatas aos postos fluviométricos....	80
Figura 5.8 - Distribuição a posteriori dos parâmetros das regressões dos modelos M1-M4 na bacia do rio Doce.....	85
Figura 5.9 - Distribuição a posteriori dos parâmetros das regressões dos modelos M1-M4 na bacia do rio Itajaí-Açu	86
Figura 5.10 - Regressão dos parâmetros de posição e escala da distribuição Gum+ pelo modelo M1. Linha tracejada representa o intervalo de 95% de credibilidade preditivo. As	

estimativas pontuais e o intervalo de 95% de credibilidade desses parâmetros são representadas, respectivamente, por pontos e linhas verticais.....	86
Figura 5.11 - Dependência espacial dos parâmetros da distribuição Gum+ pelos modelos M2 e M4.....	88
Figura 5.12 - Semivariograma empírico das medianas dos resíduos das regressões do logaritmo dos parâmetros de posição e escala do modelo M1	88
Figura 5.13 - Resíduos do parâmetro (a) μ e (b) σ do modelo M4 na bacia do rio Doce.....	89
Figura 5.14 - Resíduos do parâmetro (a) μ e (b) σ do modelo M2 na bacia do rio Doce.....	90
Figura 5.15 - Resíduos do parâmetro (a) μ e (b) σ do modelo M4 na bacia do rio Itajaí-Açu.	91
Figura 5.16 - Resíduos do parâmetro (a) μ e (b) σ do modelo M2 na bacia do rio Itajaí-Açu.	91
Figura 5.17 - Dependência espacial das observações da bacia do rio Doce. As linhas representam a correlação espacial com um modelo exponencial (M3), painel à esquerda, e um modelo exponencial potência (M4), painel à direita. Apenas pares de postos fluviométricos com pelo menos 10 observações concomitantes são plotados	92
Figura 5.18 - Dependência espacial das observações da bacia do rio Itajaí-Açu. As linhas representam a correlação espacial com um modelo exponencial (M3), painel à esquerda, e um modelo exponencial potência (M4), painel à direita. Apenas para pares com pelo menos 10 observações concomitantes.....	93
Figura 5.19 - Parâmetros estimados nos postos de calibração da bacia do rio Doce. Pontos e barras de erro representam, respectivamente, a estimativa pontual e seu intervalo de credibilidade de 95%	96
Figura 5.20 - Parâmetros estimados nos postos de calibração da bacia do rio Itajaí-Açu. Pontos e barras de erro representam, respectivamente, a estimativa pontual e seu intervalo de credibilidade de 95%	97
Figura 5.21 - Curva de quantis em três postos de calibração da bacia do rio Doce. Linhas contínuas e pontilhadas representam, respectivamente, as estimativas pontuais e o intervalo de predição com 95% de credibilidade dos modelos. No cabeçalho dos painéis está código e o ID (entre parêntese) dos postos.....	98
Figura 5.22 - Curva de quantis em três postos de calibração da bacia do rio Itajaí-Açu. Linhas contínuas e pontilhadas representam, respectivamente, as estimativas pontuais e o intervalo de predição com 95% de credibilidade dos modelos. No cabeçalho dos painéis está código e o ID (entre parêntese) dos postos.....	99
Figura 5.23 - Curva de quantis dos postos de validação na bacia do rio Doce. Linhas contínuas e pontilhadas representam, respectivamente, as estimativas pontuais e o intervalo de predição com 95% de credibilidade. No cabeçalho dos painéis está código e o ID (entre parêntese) dos postos.....	101
Figura 5.24 - Curva de quantis dos postos de validação na bacia do rio Doce. Linhas contínuas e pontilhadas representam, respectivamente, as estimativas pontuais e o intervalo de predição com 95% de credibilidade. No cabeçalho dos painéis está código e o ID (entre parêntese) dos postos.....	103
Figura 5.25 - Correlação de Spearman (ρ) entre a primeira componente principal das séries de vazões mínimas da bacia do rio Doce e a malha TSM no Oceano Atlântico sul. A bacia do rio Doce está simbolizada por um ponto preto. Apenas células da malha nas quais ρ foi significativamente diferente de 0 (valor- $p \leq 0,05$) são exibidas. Datum horizontal WGS 84	108
Figura 5.26 - Correlação de Spearman (ρ) entre a primeira componente principal das séries de vazões mínimas da bacia do rio Itajaí-Açu e a malha de TSM no oceano Pacífico. A bacia do rio Itajaí-Açu está simbolizada por um ponto preto. Apenas células da malha nas quais ρ foi significativamente diferente de 0 (valor- $p \leq 0,05$) são exibidas. As regiões relacionada a quatro índices do ENSO - Niño 1+2, 3, 4 e 3.4 - também é apresentada. Datum horizontal WGS 84	109

Figura 5.27 - Mediana dos pesos (ξ_p) atribuídos à malha de TSM em escala reduzida ($10^\circ \times 10^\circ$) no modelo M7 aplicado à bacia do rio Doce. Datum horizontal WGS 84	112
Figura 5.28 - Índice climático customizado (η_n) para a bacia do rio Doce. Traços vermelhos indicam anos nos quais pelo menos 95% de $\pi(\eta_n y)$ foi negativa ou positiva	113
Figura 5.29 - Localização dos centroides dos postos de validação e calibração e a mediana do par de coordenadas inferidas (\hat{x}_0, \hat{y}_0) do sinal do índice climático customizado. A cor dos postos indica a correlação entre as observações e o índice climático customizado. Eixo das coordenadas com origem no centroide da bacia do rio Doce.	115
Figura 5.30 - Dependência espacial da influência do índice climático customizado no parâmetro μ da bacia do rio Doce. Mapa com a mediana de γ_μ (a) e as distribuições a priori e a posteriori do decaimento (b) e do sill (c) da função de decaimento de γ_μ	116
Figura 5.31 - Dependência espacial da influência do índice climático customizado no parâmetro σ da bacia do rio Doce. Mapa com a mediana de γ_σ (a) e as distribuições a priori e a posteriori do decaimento (b) e do sill (c) da função de decaimento de γ_σ	116
Figura 5.32 - Quociente entre as estimativas pontuais da $Q_{7,10}$ dos modelos não estacionário (M7) e estacionário (M2) nos anos com observação. Estimativas dos postos 49 a 57 são obtidas a partir das predições dos modelos. Os anos de 2014 a 2019 não foram utilizados na inferência.	118
Figura 5.33 - Mediana dos pesos (ξ_p) atribuídos à malha de TSM em escala reduzida ($10^\circ \times 10^\circ$) no modelo M5 aplicado à bacia do rio Itajaí-Açu. Datum horizontal WGS 84	120
Figura 5.34 - Índice climático customizado (η_n) para a bacia do rio Itajaí-Açu. Traços vermelhos indicam anos nos quais pelo menos 95% de $\pi(\eta_n y)$ foi negativa ou positiva.....	121
Figura 5.35 - Localização da mediana do par de coordenadas inferidas (\hat{x}_0, \hat{y}_0) do sinal do índice climático customizado. A cor dos postos indica a correlação entre as observações e o índice climático customizado. Eixo das coordenadas com origem no centroide da bacia	122
Figura 5.36 - Dependência espacial da influência do índice climático customizado na bacia do rio Itajaí-Açu. Mapa com a mediana de γ_μ (a) e as distribuições a priori e a posteriori do decaimento (b) e do sill (c) da função de decaimento de γ_μ	123
Figura 5.37 - Quociente entre as estimativas pontuais da $Q_{7,10}$ dos modelos não estacionário (M5) e estacionário (M2) nos anos com observação. Estimativas dos postos 15 e 16 são obtidas a partir das predições dos modelos. O ano indicado no gráfico corresponde ao ano das observações de out-dez.	124
Figura 5.38 - Distribuição a posteriori dos parâmetros das regressões dos modelos M2 e M5 na bacia do rio Itajaí-Açu	125
Figura A.0.1 - Caracterização da bacia do rio Doce.....	147
Figura A.0.2 - Caracterização da bacia do rio Itajaí-Açu.....	148

LISTA DE TABELAS

Tabela 3.1 - Hipóteses dos métodos OLS, WLS e GLS.....	21
Tabela 3.2 - Bacias destacadas em cada um dos Componentes Principais (CPs) sinal (+) indicam correlações positivas e (-) negativas, escala temporal dos picos espectrais das séries temporais dos CPs e índices de teleconexão que apresentam correlação significativa com as series temporais dos CPs onde sinal (+) indicam correlações positivas e (-) negativas.....	42
Tabela 4.1 - Cobertura e uso do solo na bacia do rio Doce nos anos 1985 e 2018	45
Tabela 4.2 - Cobertura e uso do solo na bacia do rio Itajaí-Açu nos anos 1985 e 2018	48
Tabela 4.3 - Caracterização das classes de produção dos aquíferos e os valores adotados.....	51
Tabela 4.4 - Distribuições de probabilidades candidatas	53
Tabela 4.5 - Características dos MHBs avaliados	57
Tabela 4.6 - Distribuições a priori dos hiper-parâmetros dos MHBs desenvolvidos.....	68
Tabela 5.1 - Correlação das covariáveis candidatas com os parâmetros das distribuições Gumbel (Gum+), Log-Normal 2P (LN2) e Gama(Ga) na bacia do rio Doce	81
Tabela 5.2 - Correlação das covariáveis candidatas com os parâmetros das distribuições Gumbel (Gum+), Log-Normal 2P (LN2) e Gama(Ga) na bacia do rio Itajaí-Açu.....	81
Tabela 5.3 - Resultados do DIC para a escolha da covariável	83
Tabela 5.4 - Resultados do DIC para os modelos M1 a M4 nos postos de calibração.....	99
Tabela 5.5 - Componentes principais das séries de vazões mínimas anuais de sete dias de duração da bacia do rio Doce e Itajaí-Açu	105
Tabela 5.6 - Modelos não estacionários M5, M6 e M7.....	110
Tabela 5.7 - DIC no modelo estacionário de referência, M2, e nos modelos não estacionários, M5 a M7, na bacia do rio Doce	111
Tabela 5.8 - DIC no modelo estacionário de referência, M2, e nos modelos não estacionários, M5 a M7, na bacia do rio Itajaí-Açu	119
Tabela B.1 - Postos fluviométricos da bacia do rio Doce	149
Tabela B.2 - Postos fluviométricos da bacia do rio Itajaí-Açu.....	154

LISTA DE ABREVIATURAS E SIGLAS

- ACF – Função de autocorrelação
- AR 1 – Processo autorregressivo de primeira ordem
- AFL – Análise de frequência Local
- AFR – Análise de frequência regional
- AIC – Critério de informação de Akaike
- ANA – Agência Nacional de Águas
- CCA – Análise de Correlação Canônica
- CPRM – Companhia de Pesquisa de Recursos Minerais – Serviço Geológico do Brasil
- DIC – *Deviance Information Criterion*
- ENSO – El Niño Oscilação Sul
- FDC – Curva de permanência
- Ga – Distribuição Gama
- GEV – Distribuição Generalizada de Valores Extremos
- GLS – *Generalized Least Square*
- GUM+ – Distribuição Gumbel com assimetria positiva
- GUM- – Distribuição Gumbel com assimetria negativa
- HMC – Hamiltonian Monte Carlo
- LN2 – Distribuição Log-normal de 2 parâmetros
- LN3 – Distribuição Log-normal de 3 parâmetros
- LP3 – Distribuição Log-Perarson tipo III
- MCMC – Markov Chain Monte Carlo
- MHB – Modelo Hierárquico Bayesiano
- ONI – Índice Oceano Niño
- Q₇ – Vazão mínima de sete dias de duração
- Q_{7,10} – Vazão mínima de sete dias de duração com dez anos de tempo de retorno
- Q₉₅ – Vazão superada em 95% do período de observações
- P3 – Distribuição Pearson tipo III
- PSBI – Interpolação Espacial Baseada em Espaço Fisiográfico
- RSBFA – Análise de Frequência Regional Baseada em Vazões
- SNIRH - Sistema Nacional de Informação em Recursos Hídricos
- SOI - Índice de Oscilação Sul
- TME - Exportação de Umidade Tropical

TR – Tempo de retorno

TSM - Temperatura da Superfície do Mar

W2P – Distribuição Weibull de 2 parâmetros

W3P – Distribuição Weibull de 3 parâmetros

WLS – *Weighted Least Square*

ZCAS - Zona de Convergência do Atlântico Sul

SUMÁRIO

1	INTRODUÇÃO	15
2	OBJETIVOS	19
2.1	OBJETIVO GERAL	19
2.2	OBJETIVOS ESPECÍFICOS	19
3	REVISÃO DA LITERATURA	20
3.1	REGIONALIZAÇÃO DE VAZÕES MÍNIMAS	20
3.1.1	Métodos regressivos	20
3.1.2	Métodos gráficos	23
3.1.3	Modelos fisicamente embasados	25
3.1.4	Outras abordagens para regionalização de vazões mínimas	26
3.2	INFERÊNCIA BAYESIANA	30
3.2.1	Fundamentos	30
3.2.2	Conhecimento <i>a priori</i>	31
3.2.3	Tipos de combinação: no pooling, full pooling e partial pooling	33
3.3	MODELO HIERÁRQUICO BAYESIANO	34
3.3.1	Dependência espacial	35
3.3.2	Dependência temporal	38
4	MATERIAL E MÉTODOS	44
4.1	DESCRIÇÃO DA ÁREA DE ESTUDO	44
4.1.1	Bacia do rio Doce	44
4.1.2	Bacia do rio Itajaí-Açu	47
4.2	AQUISIÇÃO E TRATAMENTO DE DADOS	49
4.3	MODELO DE DISTRIBUIÇÃO DE PROBABILIDADE	51
4.4	MODELO HIERÁRQUICO BAYESIANO	56
4.4.1	Modelo estacionário	56
4.4.2	Modelo não estacionário	62
4.4.3	Distribuição a priori	67
4.4.4	Inferência	68
4.4.5	Comparação entre modelos	69
4.4.6	Estimação e predição	71
5	RESULTADOS E DISCUSSÃO	73
5.1	AValiação dos dados	73
5.2	DISTRIBUIÇÃO DE PROBABILIDADE CANDIDATAS	76
5.3	MODELO ESTACIONÁRIO	82
5.3.1	Escolha da covariável	82
5.3.2	Ajuste dos modelos	83
5.4	TELECONEXÃO	104
5.5	MODELO NÃO ESTACIONÁRIO	110
5.5.1	Bacia do rio Doce	110
5.5.2	Bacia do rio Itajaí-Açu	118
6	CONCLUSÕES	125
7	RECOMENDAÇÕES	130
	REFERÊNCIAS	132

1 INTRODUÇÃO

Á água tem se tornado cada vez mais um ativo de interesse para a sociedade devido ao aumento da demanda hídrica (AGÊNCIA NACIONAL DE ÁGUAS E SANEAMENTO BÁSICO, 2019). Problemas de gestão normalmente crescem durante períodos de vazões mínimas, sendo eles de grande interesse face ao aumento da demanda hídrica decorrente do crescimento populacional (MISHRA; SINGH; KONAPALA, 2016). Além da importância para a gestão de abastecimento de água, o conhecimento sobre a magnitude e frequência de vazões mínimas também é importante para a manutenção da quantidade e qualidade da água para irrigação, recreação e conservação do ecossistema, e para o planejamento de reservatórios e disposição de águas residuais (SMAKHTIN, 2001).

Face à importância dos múltiplos usos da água, o gestor de recursos hídricos deve conhecer o regime de estiagem a fim de evitar conflitos e garantir o atendimento da demanda e o uso sustentável da água. O regime de estiagem pode ser caracterizado por índices como a vazão mínima de sete dias de duração com dez anos de tempo de retorno ($Q_{7,10}$), percentis da curva de permanência (e.g. Q_{95} – vazão superada em 95% do período de observações) e outros índices descritos em Smakhtin (2001).

Estimativas regionais de quantis de vazões mínimas, como a Q_7 , podem ser determinados pela análise de frequência regional de vazões mínimas. A análise de frequência regional (AFR) ou regionalização de variáveis hidrológicas é uma técnica empregada com duas finalidades: (i) estimar quantis em locais não monitorados, mas com características físicas conhecidas e (ii) agregar dados de estações próximas à estação de interesse para aumentar a série histórica, e, assim, “trocar tempo por espaço” com o propósito de melhorar a confiabilidade dos quantis estimados.

Na AFR tradicional de vazões mínimas, destacam-se dois métodos: métodos regressivos e *index-flood*. Os métodos regressivos são os mais utilizados por hidrólogos na AFR de vazões mínimas (KIM; LEE, 2010; OUARDA; CHARRON; ST-HILAIRE, 2008). Eles consistem em modelar as respostas de uma variável dependente a variáveis explicativas, como características físicas de uma bacia hidrográfica (e.g. área, declividade). Métodos de estimação para regressão linear podem ser subdivididos em três categorias: Método dos Mínimos Quadrados Ordinários – OLS (*Ordinary Least Square*), Método dos Mínimos Quadrados Ponderados – WLS

(*Weighted Least Square*) e Método dos Mínimos Quadrados Generalizados – GLS (*Generalized Least Square*). Apesar de suas particularidades, todos esses métodos preceituam a delimitação de regiões homogêneas e não possibilitam a quantificação total da incerteza preditiva. Nesse *framework*, as incertezas são representadas pelo intervalo de predição da regressão e ignoram a incerteza herdada pela estimativa pontual do índice de vazão mínimas, já que não são observações de fato.

O *index flood*, proposto por Dalrymple (1960) e originalmente utilizado no contexto de vazões máximas, foi expandido para as vazões mínimas por Pilon (1990). Apesar de referenciar a cheia (*flood*), o método pode ser aplicável a qualquer variável (NAGHETTINI; PINTO, 2017). O *index flood* baseia-se em hipóteses restritivas (HOSKING; WALLIS, 1997): (i) invariância temporal; (ii) independência espacial e serial; (iii) as observações em todos os postos são identicamente distribuídas; e (iv) as distribuições dos postos é a mesma a menos de um fator de escala. Renard (2011) questiona a validade das hipóteses do método, uma vez que, segundo o autor, algumas são dificilmente observadas. Além da redução amostral para respeitar o critério de região homogênea e as hipóteses restritivas, segundo Thorarinsdottir *et al.* (2018), o método não considera a incerteza paramétrica na regressão e na curva de quantis. As múltiplas etapas do método dificultam a quantificação da incerteza na medida em que ela é propagada para as etapas sucessivas.

Como alternativa aos métodos de regionalização supracitados, surge como candidata a abordagem pelo paradigma Bayesiano. Modelos Hierárquicos Bayesianos (MHBs) constituem uma alternativa a métodos usuais, pois (i) permitem quantificar a incerteza preditiva de maneira trivial; (ii) dispensam a delimitação de regiões homogêneas, etapa que reduz o número de amostras disponíveis e introduz subjetividade à inferência, e conseqüentemente tornam-se uma opção atrativa para bacias com heterogeneidade espacial; (iii) têm flexibilidade para incorporar estruturas complexas de dependência nos dados e nos parâmetros do modelo, de forma a relaxar algumas hipóteses restritivas (e.g., independência condicional dos dados e independência dos resíduos do modelo) e facilitar a introdução da não estacionariedade por meio de covariáveis temporais; e (iv) a inferência em múltiplas camadas é realizada simultaneamente, o que permite a interação entre os níveis do modelo.

As vantagens supracitadas dos MHBs, assim como o avanço tecnológico das últimas décadas, atraíram diversas pesquisas relacionadas à regionalização de precipitações máximas diárias (BRACKEN *et al.*, 2016; COOLEY *et al.*, 2007; RENARD, 2011), vazões máximas diárias (LIMA *et al.*, 2016; SAMPAIO; COSTA, 2021; WU, YENAN *et al.*, 2018; WU, YUN BIAO; XUE; LIU, 2019; YAN; MORADKHANI, 2014) e de vazões mínimas (AHN; PALMER; STEINSCHNEIDER, 2017; KIM; LEE, 2010). Avanços têm sido feitos na descrição da dependência espacial da regionalização (RENARD, 2011), em termos da dependência dos resíduos do modelo, assim como da dependência entre as observações concomitantes, assim como na incorporação de covariáveis temporais. Esse último, têm sido utilizado para incorporar a influência de fenômenos climáticos de larga escala, usualmente representado por índices climáticos, nos processos hidrológicos.

Na regionalização de vazões mínimas, métodos Bayesianos foram utilizados duas vezes, até o conhecimento do autor. Kim e Lee (2010) utilizaram uma regressão Bayesiana para estimar quantis de vazão mínima com sete dias de duração na Coréia do Sul. Ahn *et al.* (2017) utilizaram um MHB que incorporou a dependência espacial dos resíduos do modelo, assim como uma covariável temporal relacionada a Temperatura da Superfície do Mar (TSM) para fazer previsões regionais de vazões mínimas. A influência da dependência entre observações concomitantes na regionalização de vazões mínimas ainda carece de investigação, assim como a influência da dependência espacial dos resíduos do modelo em estudos regionais de vazões mínimas. Modelos de regionalização não estacionários

Este trabalho explora a abordagem Bayesiana na regionalização de vazões mínimas por meio de modelos hierárquicos. A predição da regionalização é extremamente importante devido à sua utilidade para estimar quantis em locais não monitorados e, portanto, serão abordados aspectos espaço-temporais que visem melhorar as estimativas e predições do modelo. Duas bacias hidrográficas – bacia do rio Doce e do rio Itajaí-Açu – são utilizadas para investigar a influência de premissas básicas da regionalização sob métodos tradicionais, como a dependência espacial entre os resíduos do modelo e a dependência entre observações concomitantes nas predições dos modelos. A não estacionariedade neste estudo não abrange modelos de tendência temporal ou de previsões de mudança climática, ela na verdade condiciona a inferência à observações de covariáveis temporais, as quais representam

fenômenos climáticos de larga escala. A literatura reporta relação entre a TSM e o processo hidrológico influente nas duas bacias analisadas, por isso, ela é utilizada como covariável temporal na construção de modelos não estacionários. A componente não estacionária dos modelos não estacionários desenvolvidos seguiram Ahn *et al.* (2017). Este estudo, além de seu caráter investigativo a respeito das premissas básicas de dependência espacial e da avaliação da inclusão de covariáveis temporais, traz uma análise regional importante para as bacias hidrográficas abrangidas.

2 OBJETIVOS

2.1 Objetivo geral

Desenvolver um modelo Bayesiano Hierárquico para estimar quantis de vazões mínimas em condição de não estacionariedade e quantificar a incerteza preditiva correspondente.

2.2 Objetivos específicos

- Identificar covariáveis espaciais e temporais que auxiliem a explicar o processo estocástico em análise;
- Avaliar os efeitos da incorporação da dependência espacial na estimação dos parâmetros e na incerteza preditiva;
- Avaliar a influência da interdependência espacial entre as observações de vazão mínima na inferência estatística; e
- Avaliar o desempenho de um modelo não estacionário que considere a variabilidade climática.

3 REVISÃO DA LITERATURA

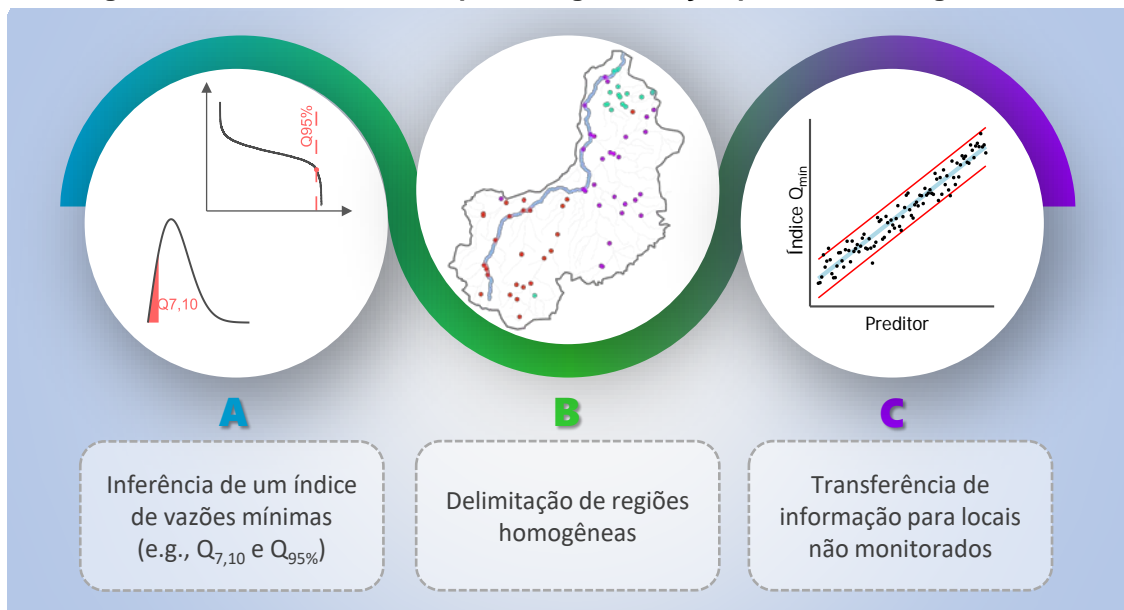
3.1 Regionalização de vazões mínimas

Os métodos de estimação de vazões mínimas em bacias não monitoradas podem ser divididos em três categorias, a saber: métodos regressivos regionais, métodos gráficos e modelos fisicamente embasados (MISHRA; SINGH; KONAPALA, 2016).

3.1.1 Métodos regressivos

Métodos regressivos, em geral, podem ser descritos por três etapas: inferência do índice de vazões mínimas (como a $Q_{7,10}$ e a Q_{95}) em locais monitorados, delimitação de regiões homogêneas e transferência de informações dos postos monitorados para o local de interesse onde não há observações (MISHRA; SINGH; KONAPALA, 2016). Esse procedimento é esquematizado na Figura 3.1.

Figura 3.1 - Fluxo de trabalho para a regionalização por métodos regressivos



Na etapa A da Figura 3.1, pela abordagem usual de regionalização por regressão linear, a inferência do índice de vazão mínima é estimada pontualmente. Para a estimativa da $Q_{7,10}$, utiliza-se usualmente o método dos momentos convencionais, método dos momentos-L ou método da máxima verossimilhança (mais detalhes no item 0).

A delimitação de regiões homogêneas, etapa B da Figura 3.1.

, é uma etapa que introduz julgamento subjetivo na análise (FARSADNIA *et al.*, 2014). Existe na literatura uma ampla variedade de métodos propostos para delimitar regiões homogêneas, baseados em análises multivariadas de *clusters* com dados físicos e climáticos das bacias monitoradas e nas estatísticas das observações. Alguns dos métodos de delimitação de *clusters* utilizados em métodos regressivos são a Região de Influência – ROI (BURN, 1988, 1990), Análise de Correlação Canônica – CCA (OUARDA *et al.*, 2001; TSAKIRIS; NALBANTIS; CAVADIAS, 2011), agrupamento *fuzzy c-means* (GOMES; BLANCO; PESSOA, 2018; RAO; SRINIVAS, 2006) e agrupamento *k-means* (KIM; LEE, 2010).

A função de transferência de informações para locais não monitorados pode ser realizada por meio de regressões lineares, cujos parâmetros são usualmente estimados pelos métodos OLS, WLS e GLS. As principais diferenças nos métodos são resumidas na Tabela 3.1. Neste trabalho, os métodos foram abordados brevemente, mas para mais detalhes o leitor é remetido à Luiz (2013).

Tabela 3.1 - Hipóteses dos métodos OLS, WLS e GLS

Hipóteses	OLS	WLS	GLS
Resíduos são homoscedásticos	✓	✓	✗
Resíduos são independentes	✓	✗	✗

Em um modelo de regressão linear simples, a relação entre uma variável de resposta Y para uma dada variável explicativa X pode ser expressa pela Eq. 3.1.

$$Y = \beta_0 + \beta_1 X + \varepsilon \quad (3.1)$$

na qual β_0 e β_1 são os coeficientes do modelo e ε denota os erros ou resíduos do modelo.

No método OLS, a estimativa dos parâmetros da regressão é realizada de tal forma que os valores estimados minimizem uma função objetivo, representada pelo somatório do quadrado dos resíduos. A quantificação da incerteza por meio de intervalos de confiança por esse método possui como premissa as hipóteses de independência e homoscedasticidade dos resíduos (COSTA, 2017). No entanto, a

regionalização de variáveis hidrológicas frequentemente viola essas hipóteses, cenário no qual o desempenho do método OLS decai (STEDINGER; TASKER, 1985).

Eletrobrás (1985) publicou uma metodologia para regionalização de vazões mínimas, máximas e médias. Nesse estudo, vazões mínimas são regionalizadas por meio de regressões lineares, cujos parâmetros são estimadas pelo método OLS pela inferência frequentista. A partir da publicação desse estudo, muitos trabalhos (e.g., LISBOA *et al.*, 2019; MATOS *et al.*, 2020; MELLO; PESSOA; SANTANA, 2020; RIBEIRO; MARQUES; SILVA, 2005) o referenciam e vêm utilizando o método OLS no *framework* de regionalização de vazões mínimas. De fato, a regionalização de vazões mínimas pela abordagem da regressão com o método OLS aparenta ser majoritária entre os hidrólogos.

O método WLS, diferentemente do método OLS, designa pesos diferentes para as observações. Tasker (1980) propôs algumas alterações ao método convencional WLS para aplicações em hidrologia, visando introduzir na inferência a desigualdade no números de observações das séries. O método modificado separa a variância das estimativas em duas componentes: variância dos erros do modelo e a variância dos erros de amostragem. O método ainda assume que os erros do modelo são independentes, no entanto, não assume que os resíduos sejam homoscedásticos.

O método GLS de Stedinger e Tasker (1985) oferece um modelo mais contundente para trabalhar dados de vazão. Ele incorpora os mesmos avanços do método WLS em relação ao método OLS e, além disso, leva em consideração a dependência cruzada em observações concomitantes. Stedinger e Tasker (1985) mostraram que o desempenho do método GLS supera substancialmente o método WLS quando as amostras tem correlação cruzada maior ou igual a 0,6; e que ambos os métodos GLS e WLS proporcionam grandes melhorias nas estimativas dos parâmetros dos modelos regressivos se comparados ao método OLS. O método WLS, assim como o método GLS, possuem como dificuldade a estimação da matriz de covariância dos resíduos (STEDINGER; TASKER, 1985).

Apesar dos promissores resultados encontrados por Stedinger e Tasker (1985), a utilização dos métodos WLS (e.g., LUIZ, S.; FERNANDES; JUNIOR, 2013; RIES III;

FRIESZ, 2000) e GLS (e.g., LUIZ, S. F., 2013; VOGEL; KROLL, 1990) na regionalização de vazões mínimas é bastante inferior ao método OLS.

A regionalização de vazões mínimas baseada em métodos regressivos por meio do paradigma Bayesiano também é possível, contudo, até o conhecimento do autor, só um trabalho explorou essa área. Kim e Lee (2010) utilizaram o estimador OLS na regionalização da Q_7 na Coréia pela abordagem frequentista e Bayesiana. Os resultados dos autores mostraram que o modelo Bayesiano reduziu as incertezas na regionalização.

3.1.2 Métodos gráficos

Métodos regressivos fazem estimativas de uma única estatística de vazões mínimas. Em contrapartida, métodos gráficos permitem a estimativa de um conjunto de estatísticas similares por meio de funções contínuas (MISHRA; SINGH; KONAPALA, 2016) (e.g., curva de permanência e curva de quantis). Devido ao foco deste trabalho, a revisão irá se ater a métodos de estimação de curva de quantis.

Naghattini e Pinto (2017) discorrem sobre dois métodos que se enquadram na classificação acima: regionalização dos parâmetros da distribuição de probabilidade e o *index flood*. O primeiro remete fortemente ao método regressivo. Nesse esquema, ajusta-se uma mesma distribuição de probabilidade aos postos monitorados de uma mesma região homogênea e regionaliza-se seus parâmetros por métodos regressivos. Ao contrário da metodologia tradicional por métodos regressivos, na qual se regionaliza uma vazão mínima de referência (normalmente o produto final de interesse), essa abordagem concede ao analista a possibilidade de estimar quantis para qualquer tempo de retorno.

O *index flood* é amplamente utilizado em estudos de vazões máximas, no entanto, seu uso ainda é incipiente na análise de vazões mínimas comparado aos métodos regressivos e ele é restrito a poucos trabalhos (e.g., CPRM, 2001; SHI *et al.*, 2010; SHIAU; WU, 2009).

O *index flood* baseia-se na teórica da invariância de escala, a qual assume que a distribuição de probabilidade de todos os postos é idêntica em uma região homogênea, com excessão de um fator de escala. Hosking e Wallis (1997)

promoveram um grande avanço na AFR com a implementação do método dos momentos-L no *index flood*. Hosking e Wallis (1997) construíram estatísticas para tornar algumas etapas metodológicas da AFR menos subjetivas e implementaram o estimador de parâmetros utilizando o método dos momentos-L, estimador robusto e menos sensível à *outliers* (NAGHETTINI, 2017a). Naghettini e Pinto (2017) resumem o *index flood* às seguintes etapas:

1. Procurar inconsistências nos dados;
2. Adimensionalização das séries;
3. Definição da distribuição empírica local;
4. Delimitação de regiões homogêneas;
5. Definição da curva de frequência regional;
6. Análise de regressão; e
7. Estimação de quantis.

Hosking e Wallis (1997) propuseram que a delimitação de regiões homogêneas fosse seguida por dois passos: primeiro, um pré agrupamento com base nas características físicas da bacia, e em seguida, o ajuste do agrupamento com base nos quociente de assimetria-L e curtose-L das séries observadas. Os autores também propuseram uma medida de heterogeneidade para quantificar e classificar a homogeneidade de uma região delimitada. Contudo, mesmo embasando-se nas estatísticas das séries, essa etapa continua sendo uma fonte de subjetividade conforme discutida na seção 3.2. Hosking e Wallis (1997) sugerem ajustes para reduzir a heterogeneidade ou melhorar a coerência física das regiões, entre eles, excluir um ou mais postos da análise. Essa abordagem é problemática pois a heterogeneidade decorre de uma fonte de informação válida e excluí-la da análise não a torna menos real e tem potencial de enviesar as previsões (da mesma forma como mantê-la sob a hipótese de regiões homogêneas).

Renard (2011) aborda várias limitações do método:

1. *A delimitação de regiões homogêneas, onde a hipótese de invariância de escala seja válida, é longe de ser evidente;*
2. *A hipótese de invariância de escala pode ser simplesmente muito restritiva em alguns casos;*
3. *Na maioria dos casos, o index flood é definido como a média ou mediana das observações locais, mas as justificativas físicas por trás dessa seleção não são claras;*
4. *A distribuição regional é estimada juntando os dados locais adimensionalizados e tratando-os como se fossem independentes, o que raramente é o caso; e*
5. *Métodos comuns de regressão, como o OLS, podem ser estatisticamente ineficientes pois os valores adimensionalizados são estatísticas (ao contrário de observações). Logo, as regressões são afetadas pelos erros, cujos valores e propriedades podem ser espacialmente dependentes;*
6. *Os pontos acima dificultam a quantificação total da incerteza das estimativas (Renard, 2011, p. 1)*

3.1.3 Modelos fisicamente embasados

A estimação de vazões de referência em bacias não monitoradas também pode ser estimada por modelos fisicamente embasados. A vantagem dessa abordagem é que se obtém uma série histórica de vazões (condicionada a existência de dados pluviométricos e fluviométricos próximos ao local de interesse), ao invés de apenas uma estatística de interesse. Ao mesmo tempo, o modelo tenta representar diferentes regimes hidrológicos – cheias e estiagens –, o que introduz maiores incertezas quando comparado à outras abordagens que visam apenas a estatística de interesse.

Souza (2011) utilizou o modelo SWAT calibrado com uma estação fluviométrica para estimar a Q_{95} nas sub-bacias contidas na área de drenagem da estação de calibração. A validação dos resultados previstos foi realizada por meio de campanhas de medição em campo nas sub-bacias não monitoradas. A validação do modelo não foi contundente, uma vez que as vazões medidas em campo ocorreram em períodos de vazões médias (SOUZA, 2011).

England e Hisdal (2009) compararam as predições do índice de vazão mínima Q_c (vazão mínima comum) por meio de método regressivo e de modelo chuva-vazão (HBV) em dezenas de bacias na Noruega. As estimativas do método regressivo foram mais precisas do que as do modelo HBV em relação ao índice Q_c . England e Hisdal (2009) atribuem as maiores incertezas pela abordagem do modelo hidrológico às etapas adicionais dessa metodologia e da modelagem de períodos que não são característicos de vazões mínimas.

3.1.4 Outras abordagens para regionalização de vazões mínimas

3.1.4.1 Interpolação espacial

Chaves *et al.* (2002) propuseram uma metodologia simples e robusta baseada em interpolação espacial para a estimação de vazões mínimas de referência em locais não monitorados. A metodologia é fundamentada no método de razão de área de drenagem, proposto Stedinger *et al.* (1993) e representado pela Eq. 3.2.

$$Q_y = \frac{A_y}{A_x} Q_x \quad (3.2)$$

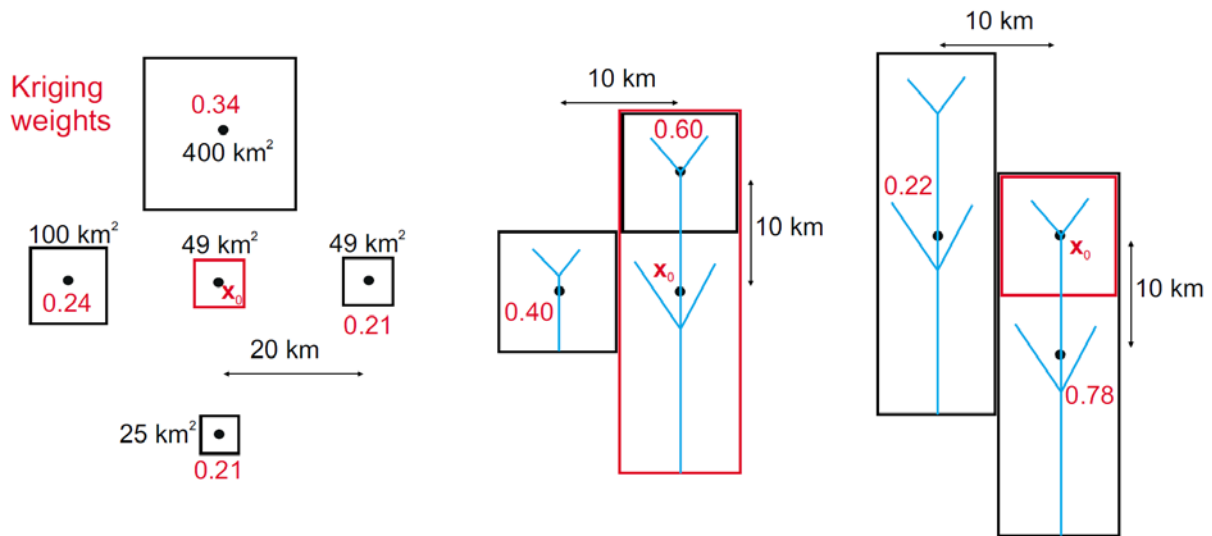
na qual Q_y , Q_x , A_y e A_x são, respectivamente, a vazão a ser estimada no local y , a vazão de referência no posto fluviométrico vizinho x e a área de drenagem de y e x .

A metodologia de Chaves *et al.* (2002) propõe diferentes fórmulas, em função da disposição hidrográfica do local de interesse e dos postos fluviométricos usados na interpolação, que ponderam a proporcionalidade com a área de drenagem com a distância dos postos ao local de interesse. O Método foi comparado com a regressão simples estimada pelo método OLS em uma bacia hidrográfica e obteve melhores métricas de desempenho (Nash-Sutcliffe e erro médio). O método explicitamente assume que toda variabilidade é explicada pela área de drenagem e, segundo o autor, dispensa a delimitação de regiões homogêneas. Outros fatores como clima, topografia, solo e geologia influenciam as vazões mínimas (SMAKHTIN, 2001), logo, é de se esperar que o desempenho do método seja melhor em locais com características similares aos postos usados na interpolação (i.e., locais que compartilhem certo grau de homogeneidade). Possivelmente em função disso, Chaves *et al.* (2002) indica que o método deve ser aplicado em locais com área, fisiografia e clima similares.

O Top-kriging (SKØIEN; MERZ; BLÖSCHL, 2006) é um método de interpolação que incorpora a dependência espacial de dois tipos de variáveis: variáveis contínuas no espaço, como chuva e característica do solo (variação euclidiana), e a síntese hidrológica por trás da rede de drenagem (variação do tipo estrutura de árvore). Ao invés de utilizar variogramas diretamente, o Top-kriging utiliza uma ponderação de pares de variogramas que consideram a natureza aninhada das bacias hidrográficas. A Figura 3.2 ilustra a interface entre os pesos designados aos postos fluviométricos em relação à estrutura hidrográfica e ao tamanho das bacias.

O método Interpolação Espacial Baseada em Espaço Fisiográfico (PSBI) (CHOKMANI; OUARDA, 2004) interpola a variável de interesse no espaço fisiográfico no lugar do espaço geográfico. O espaço fisiográfico é construído com base na análise de componentes principais (ACP) ou análise de correlação canônica (ACC). No estudo original de Chokmani e Ouarda (2004), as melhores estimativas de quantis de vazões máximas foram geradas a partir do espaço fisiográfico construído pela ACC.

Figura 3.2 - Efeito do tamanho da bacia hidrográfica (esquerda) e do aninhamento de bacias (direita). Pesos da Krigagem estão em vermelho e os pontos pretos juntos aos rótulos x_0 em vermelho indicam os centroides das bacias hidrográficas.



Fonte: Chokmani e Ouarda (2004)

O Top-kriging e o PSBI, assim como o método de Chaves, a princípio não requerem a delimitação de regiões homogêneas. Os dois primeiros métodos de interpolação citados fornecem uma descrição mais detalhada da variabilidade espacial, incorporando outras covariáveis além da área de drenagem.

Castiglioni *et al.* (2009) avaliou a regionalização da vazão associada a uma duração de 355 dias pelo PSBI, por técnicas determinísticas convencionais (Inverso do Quadrado da Distância, Polígonos de Thiessen) e geoestatísticas (Krigagem Universal e Krigagem Ordinária) e comparou os resultados com o método regressivo. Os resultados dos autores sugerem que métodos geoestatísticos e o método regressivo têm melhor desempenho em relação a métodos determinísticos e que o desempenho da Krigagem universal é superior ao do método regressivo.

Castiglioni *et al.* (2011) comparou as estimativas do Top-kriging e o PSBI pela krigagem universal na regionalização da Q_{355} . Segundo os autores, o desempenho do Top-kriging foi marginalmente superior ao do PSBI, mas os métodos devem ser vistos como complementares. O Top-kriging teve melhor desempenho na predição ao longo dos principais cursos d'água e o PSBI na cabeceira da bacia.

3.1.4.2 Análise de Frequência Regional Baseada em Vazões

Requena *et al.* (2017) propuseram uma metodologia inovadora para regionalização de vazões, denominada Análise de Frequência Regional Baseada em Vazões (RSBFA, do acrônimo em inglês *Regional Streamflow-Based Frequency Analysis*).

Introduzida primeiramente para regionalizar vazões máximas, a metodologia foi estendida para vazões mínimas por Requena *et al.* (2018). A metodologia consiste em regionalizar curvas de permanência por meio de regressão funcional múltipla, gerar séries sintéticas de vazão e, a partir delas, realizar a análise de frequência local para estimar os quantis de interesse em uma seção não monitorada. O fluxo de trabalho da nova metodologia e sua diferença em relação aos métodos usuais de regressão é esquematizada na Figura 3.3. A grande vantagem da nova abordagem proposta consiste em estimar uma série de vazões diárias, a partir das quais se pode inferir qualquer estatística do local não monitorado, como percentis e a probabilidade de não superação de vazões máximas, médias e mínimas. A técnica não exige delimitação de regiões homogêneas, no entanto, Requena *et al.* (2017) levantam a hipótese que sua utilização poderia melhorar as estimativas do modelo em locais não monitorados.

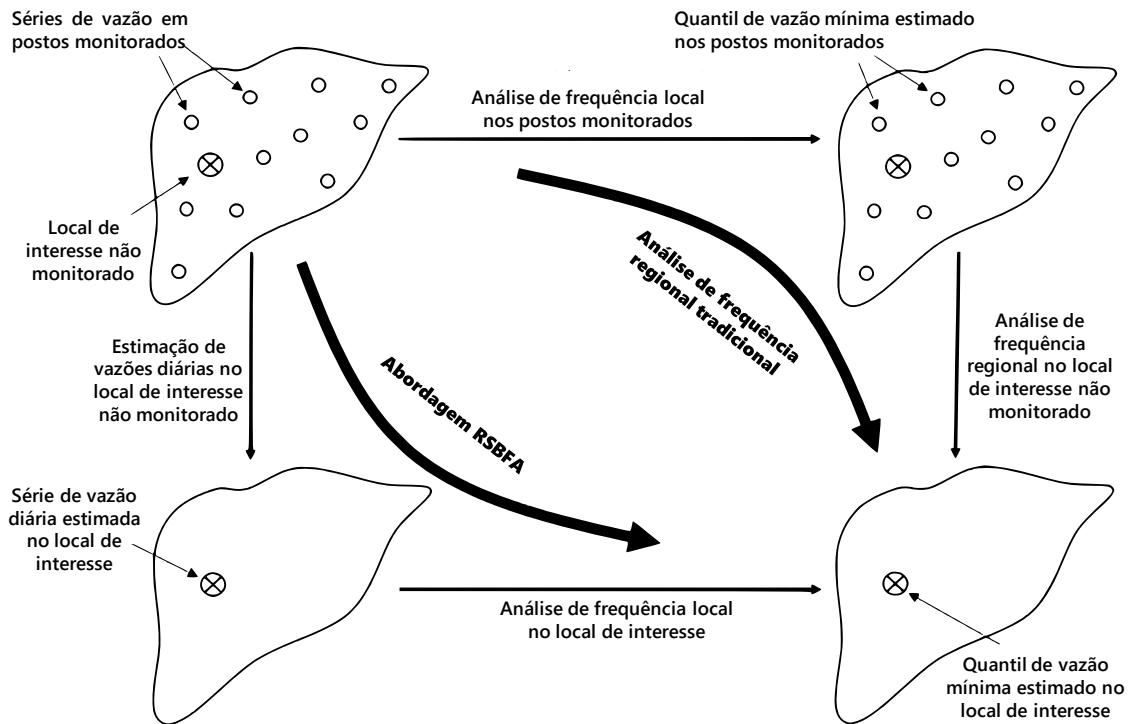
Um aspecto importante do método concerne à forma como as séries de vazões sintéticas são geradas. Seja $Q_{m,t}^D$ a vazão no local de interesse no dia m do ano t , para $m \in \{1,2, \dots, 365\}$, e $Q_{m,t}^j$, de maneira análoga, a vazão no posto vizinho j , para $j \in \{1,2, \dots, N\}$; FDC a função de transformação de vazão em probabilidade de não superação pela curva de permanência e FDC^{-1} a sua inversa; e por fim, w_j o peso estimado para o posto j , a estimação da vazão no local não monitorado é calculada pela Eq. 3.3.

$$Q_{m,t}^D = \frac{\sum_{j=1}^N w_j FDC_D^{-1}[FDC_j(Q_{m,t}^j)]}{\sum_{j=1}^N w_j} \quad (3.3)$$

O método, portanto, depende da disponibilidade temporal de dados nos postos vizinhos, uma vez que a série histórica é utilizada no processo. Na presença de falhas

em todos os postos fluviométricos para um determinado período de um ano, o método não consegue estimar a vazão no local de interesse nesse período.

Figura 3.3 - Metodologia da AFR baseada na estimação de vazão e da AFR tradicional.



Fonte: adaptado de Requena et al. (2018)

3.1.4.3 Modelo Hierárquico Bayesiano

Modelos Hierárquicos Bayesianos (MHB) também oferecem um caminho para realizar a análise de frequência regional de vazões mínimas. Essa abordagem será apresentada em mais detalhes na seção 3.3, após a introdução à inferência Bayesiana.

3.2 Inferência Bayesiana

3.2.1 Fundamentos

A abordagem Bayesiana permite a introdução formal de distribuições de probabilidade aos parâmetros do modelo e atualização do conhecimento prévio sobre o fenômeno em estudo com base na amostra. Nessa conjuntura, um analista pode introduzir ao modelo conhecimentos de sua experiência profissional ou da literatura por meio da especificação das distribuições *a priori* dos parâmetros, $p(\theta)$. Na inferência, o conhecimento do analista é atualizado com a informação agregada pelos dados, por meio da função de verossimilhança $f(x|\theta)$, e obtém-se a distribuição *a posteriori* dos parâmetros, $\pi(\theta|x)$, representada pela Eq 3.4.

$$\pi(\boldsymbol{\theta}|x) = \frac{f(x|\boldsymbol{\theta})p(\boldsymbol{\theta})}{f(x)} \propto f(x|\boldsymbol{\theta})p(\boldsymbol{\theta}) \quad (3.4)$$

na qual $f(x)$ representa a distribuição preditiva *a priori*, que atua como uma constante para que a integral de $\pi(\boldsymbol{\theta}|x)$ seja igual a 1.

Se a distribuição *a priori* dos parâmetros de um modelo Bayesiano não é informativa, então a distribuição *a posteriori* é dominada pela função de verossimilhança e a estimativa do modelo converge para a estimativa do método da máxima verossimilhança da abordagem frequentista.

Na Eq. 3.4, a distribuição *a posteriori* é decomposta na probabilidade condicional de x dado $\boldsymbol{\theta}$ vezes a probabilidade de $\boldsymbol{\theta}$ acontecer. Essa decomposição não necessariamente necessita ser restrita a dois termos (Eq. 3.5).

$$\pi(\boldsymbol{\theta}|x, \boldsymbol{\alpha}, \boldsymbol{\beta}) \propto f(x|\boldsymbol{\theta}, \boldsymbol{\alpha}, \boldsymbol{\beta}) f(\boldsymbol{\theta}|\boldsymbol{\alpha}) f(\boldsymbol{\alpha}|\boldsymbol{\beta}) p(\boldsymbol{\beta}) \quad (3.5)$$

Esse tipo de abordagem em vários níveis é a base dos Modelos Hierárquicos Bayesianos (MHB).

3.2.2 Conhecimento *a priori*

A distribuição *a priori* pode ser informativa, pouco informativa ou não informativa. A distribuição *a priori* não informativa, também chamada de difusa ou vaga, visa não introduzir na inferência nenhuma informação externa aos dados, que se traduz a um dos lemas encontrados em diversas referências Bayesianas: *let the data speak for themselves* (deixe os dados falarem por si). No entanto, em alguns casos, distribuições *a priori* difusas podem ser, na verdade, muito informativas. Essas situações podem ocorrer, por exemplo, com *prioris* difusas estabelecidas em parâmetros transformados (LEMOINE, 2019; SEAMAN; SEAMAN; STAMEY, 2012) e a adoção de distribuição gama inversa para parâmetros cujos valores se situem próximas de zero, intervalo no qual essa distribuição é informativa (GELMAN, 2006).

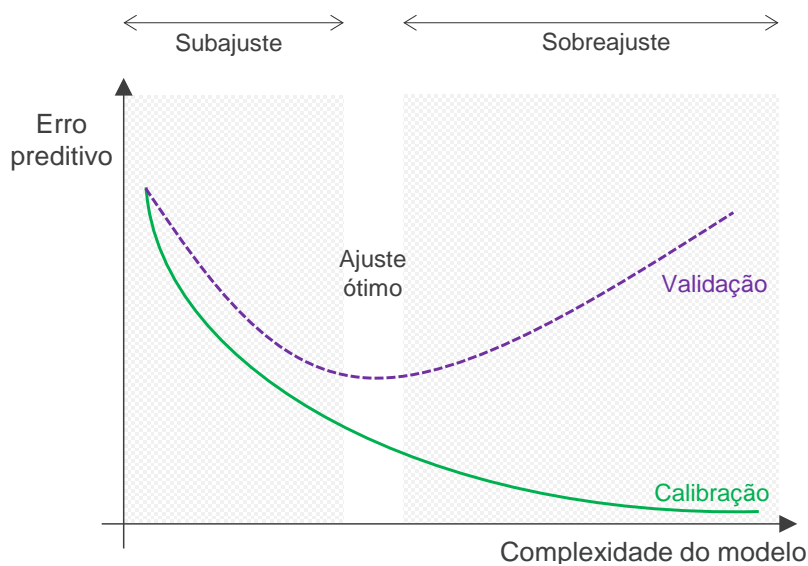
A distribuição *a priori* pouco informativa tem o intuito de restringir a distribuição *a posteriori* para valores razoáveis a partir do conhecimento prévio do analista, mas sem tentar capturar todo o conhecimento do analista sobre o fenômeno (GELMAN *et al.*,

2013). Em contrapartida, as distribuições *a priori* informativas, como o nome sugere, introduzem na inferência um conhecimento prévio sobre o fenômeno com bastante precisão, o que se traduz em distribuições *a priori* com variância reduzida.

No contexto de análise de frequência, Kuczera (1999) propôs a utilização de informações de regionalização na determinação da distribuição *a priori* informativa para o parâmetro de forma na análise de frequência local, metodologia essa adotada por Silva *et al.* (2016) para determinar o parâmetro de forma da distribuição Generalizada de Pareto (GP) na análise de frequência local (AFL) de vazão máxima; Martins e Stedinger (2000) utilizaram uma distribuição *a priori* (chamada pelos autores de *priori* geofísica), determinada pela distribuição beta, em que limitada o parâmetro de forma da distribuição Generalizada de Valores Extremos (GEV) para vazão máxima anual de acordo com valores mais prováveis segundo a literatura; Madsen e Rosbjerg (1997) chegaram em uma distribuição *a priori* informativa com base no método dos mínimos quadrados generalizados (GLS) para os parâmetros da distribuição Generalizada de Pareto (GP); Renard (2011) utilizou uma distribuição normal com média zero e desvio padrão 0,3 para representar a distribuição *priori* do parâmetro de forma da GEV para a precipitação máxima diária.

Simpson *et al.* (2017) desenvolveram a *Penalized Complexity Prior Distribution* (PCPD), uma distribuição *a priori* que penaliza o comportamento da cauda das distribuições quando essa se afasta de um modelo padrão mais simples, logo, atuando na prevenção de sobreajuste (Figura 3.4). Opitz *et al.* (2018) utilizaram a PCPD no parâmetro de forma da distribuição Generalizada de Pareto para os dados de precipitação máxima diária nos Países Baixos; Costa e Sampaio (no prelo) também recorreram à PCPD para restringir o parâmetro de forma do modelo misto Gama-Pareto para as vazões máximas em duas bacias nos Estados Unidos.

Figura 3.4 - Consequências do subajuste e superajuste de um modelo no erro preditivo dos dados de validação e calibração



A informalidade no estabelecimento do conhecimento *a priori* é uma grande crítica da vertente Bayesiana (KUCZERA, 1982). Sob a perspectiva dos críticos, os resultados ficam suscetíveis ao bom senso do analista entre uma escolha objetiva e fundamentalmente embasada e uma escolha subjetiva, com potencial de gerar resultados enviesados e tendenciosos.

3.2.3 Tipos de combinação: no pooling, full pooling e partial pooling

Em esquemas de MBH, três abordagens podem ser consideradas na agregação de dados espaciais na inferência estatística: *no pooling*, *full pooling* e *partial pooling* (DEVINENI *et al.*, 2013). O *no pooling* considera que cada parâmetro do modelo distributivo é modelado localmente, o que permite capturar a variabilidade espacial (equivalente à AFL). O *full pooling* estabelece que todos os locais possuem o mesmo coeficiente, o que possibilita aumentar o tamanho da amostra e reduzir a variância do modelo (equivalente à AFR clássica). Porém, ele intrinsecamente considera que há homogeneidade espacial. Por fim, o *partial pooling* é um meio termo entre as outras abordagens. Ele assume que os parâmetros locais dispersam em torno de um valor médio regional. Essa consideração preserva particularidades locais enquanto que se reduz o número de parâmetros independentes estimados e conseqüentemente a incerteza do modelo. Logo, o *partial pooling* dispensa a delimitação de regiões homogêneas.

Lima *et al.* (2016) construíram um MBH para regionalização de vazões máximas com a distribuição Generalizada de Valores Extremos (GEV), onde o *partial pooling* foi utilizado para modelar o parâmetro de forma de GEV e a escalabilidade com a área para modelar os parâmetros de posição e escala. O modelo obteve desempenho superior ao *index flood* e foi capaz de reduzir as incertezas na estimação de quantis. Com o mesmo esquema de Lima *et al.* (2016), Wu *et al.* (2018) obtiveram resultados semelhantes. Sampaio e Costa (2021), a partir do MHB de Lima *et al.* (2016), avaliaram a introdução da dependência espacial dos resíduos das regressões dos parâmetros da GEV nas estimativas e predições do MHB. Esses autores verificaram que a abordagem avaliada diminui o viés e a incerteza das estimativas em locais não monitorados em regiões mais densamente monitoradas. Por outro lado, o modelo de referência de Lima *et al.* (2016) foi superior nas predições em locais escarços em monitoramento.

No contexto de não estacionariedade, Sun *et al.* (2014) utilizaram o *partial pooling* para modelar a declividade da tendência linear temporal do parâmetro de posição da GEV em seu MBH de regionalização de vazões máximas.

3.3 Modelo Hierárquico Bayesiano

MHBs são geralmente compostos por três camadas: distribuição *a priori*, processo e dados (COOLEY *et al.*, 2012). Considere um MHB com o vetor de parâmetros $\boldsymbol{\psi}$, tal que $\boldsymbol{\psi} = (\boldsymbol{\psi}_1, \boldsymbol{\psi}_2)$, onde $\boldsymbol{\psi}_1$ são os parâmetros na camada dos dados (i.e., os parâmetros da distribuição de probabilidades) e $\boldsymbol{\psi}_2$ os parâmetros na camada do processo (i.e., dos modelos que descrevem a variabilidade espacial dos parâmetros da distribuição de probabilidades). A distribuição *a posteriori* dos parâmetros do modelo representada pela amostra de dados \boldsymbol{x} , $\pi(\boldsymbol{\psi}|\boldsymbol{x})$, pode ser definida por:

$$\pi(\boldsymbol{\psi}|\boldsymbol{x}) \propto \pi(\boldsymbol{x}|\boldsymbol{\psi}_1) \pi(\boldsymbol{\psi}_1|\boldsymbol{\psi}_2) \pi(\boldsymbol{\psi}_2) \quad (3.6)$$

Em modelos complexos, a solução analítica da Eq. 3.6 pode ser inviável e deve-se recorrer às soluções numéricas. Com o advento de métodos baseados em simulações de Monte Carlo via Cadeias de Markov (MCMC, do acrônimo em inglês para *Markov chain Monte Carlo*) (desenvolvidos entre a década de 1950 e 1970), houve uma grande expansão na popularidade da abordagem de métodos Bayesianos, pois se tornou possível estimar a solução aproximada da Eq. 3.6 (REICH; GHOSH, 2019). Os

algoritmos mais populares baseados em MCMC são Metropolis-Hastings, Gibbs sampling e Hamiltonian Monte Carlo (HMC), os quais estão implementados em *softwares* como OpenBUGS¹, JAGS², Stan³ e NIMBLE⁴ com integração à linguagem R.

Processos espaço-temporais frequentemente manifestam dependência e heterogeneidade, complexidades que requerem modelos flexíveis para serem trabalhadas, tais como Modelos Hierárquicos Bayesianos (MHBs) (HAINING; LI, 2020). A heterogeneidade é facilmente acomodada no modelo por meio do *partial pooling*. Em um modelo regional, essa abordagem assume que os parâmetros locais dispersam em torno de um valor médio regional. Essa consideração preserva particularidades locais enquanto se reduz o número de parâmetros independentes estimados e, conseqüentemente, a incerteza do modelo. A utilização do *partial pooling* na inferência permite dispensar a delimitação de regiões homogêneas, uma etapa que introduz julgamento subjetivo na análise (FARSADNIA *et al.*, 2014) e que é fundamental em outras técnicas de regionalização de vazões mínimas.

As dependências espacial e temporal serão abordadas nas seções 3.3.1 e 3.3.1, respectivamente.

3.3.1 Dependência espacial

Na inferência de modelos regionais, dois tipos de dependência espacial são usualmente ignoradas: a dependência dos dados e a dos resíduos do modelo. Em poucas palavras, a primeira refere-se a dependência espacial em uma curta escala de tempo, relacionada às observações concomitantes de uma série histórica, e a segunda a dependência espacial climática (longa escala de tempo), relacionada ao comportamento médio do processo no espaço.

Eventos de vazões mínimas, em geral, têm grande extensão espacial, logo, a hipótese de independência entre observações pode não ser válida (LAAHA *et al.*, 2013). Nesse caso, modelos que não contabilizam a correlação espacial entre observações podem

¹ <http://www.openbugs.net/>

² <https://mcmc-jags.sourceforge.io/>

³ <https://mc-stan.org/>

⁴ <https://r-nimble.org/>

gerar estimativas cuja incerteza é subestimada devido à inflação do número efetivo de graus de liberdade dos dados. Em termos de predição, isso resultaria na subestimação do intervalo de credibilidade em locais não monitorados, além de poder resultar em viés nas estimativas (RENARD, 2011). Em relação à hipótese de independência espacial dos resíduos dos parâmetros, por sua vez, Kjeldsen e Jones (2007, 2009) mostram que ela pode não ser realística. Na verdade, modelos com piores capacidades preditivas exibem resíduos mais espacialmente correlacionados e, mesmo com a adição de novas covariáveis físicas ou climáticas, uma parcela da correlação dos resíduos se mantém devido ao erro estrutural da regressão linear (KJELDSEN; JONES, 2009). Por fim, como demonstrado por Renard (2011) e Kjeldsen e Jones (2007, 2009), a dependência espacial dos erros de regressão podem impactar a precisão e acurácia das estimativas da distribuição regional dos parâmetros nos postos utilizados na AFR.

A dependência espacial pode ser introduzida no MHB em duas instâncias: no nível do processo e no nível dos dados. No nível do processo, a abordagem mais simples consiste em utilizar as características físicas das bacias hidrográficas (e.g., área de drenagem, elevação) como descritoras da variação espacial dos parâmetros da distribuição marginal por meio de regressões lineares. Pode-se explicitamente utilizar a latitude e longitude como covariáveis para exprimir a variabilidade espacial dos parâmetros, no entanto, tal abordagem nem sempre é capaz de representar a complexidade por trás da distribuição espacial dos parâmetros (COOLEY *et al.*, 2012).

A segunda abordagem parte da hipótese de que os parâmetros da distribuição marginal fazem parte de um processo latente Gaussiano. Nessa abordagem, os resíduos são espacialmente dependentes e uma estrutura de dependência espacial é explicitamente prescrita por meio da covariância. Davison *et al.* (2012) observam que a abordagem por meio do processo latente permite capturar variações locais complexas que modelos de superfície de tendência não conseguem reproduzir. A estrutura de covariância (e.g., semivariograma exponencial, esférico, matérn) utilizada no processo latente Gaussiano deve ser definida pelo analista. O semivariograma exponencial é o modelo mais adotado em MHBs para regionalização, principalmente em estudos sobre precipitação máxima diária (e.g., BRACKEN *et al.*, 2016; COOLEY; NYCHKA; NAVEAU, 2007; DAVISON; PADOAN; RIBATET, 2012; DYRRDAL *et al.*, 2015; GARCÍA *et al.*, 2018), mas também foi utilizado em estudos de extremos

máximos de vazão (YAN; MORADKHANI, 2014, 2016) e extremos mínimos de vazão (AHN; PALMER; STEINSCHNEIDER, 2017).

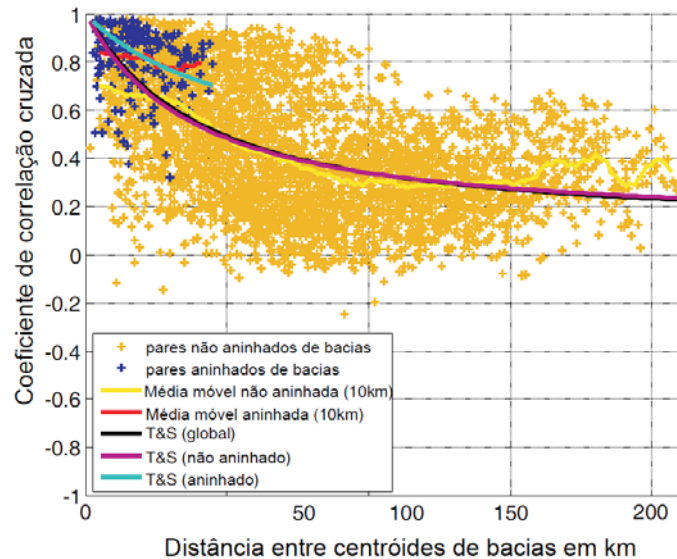
Sozinha, a segunda abordagem assume que as observações são independentes no espaço condicionadas aos parâmetros da distribuição de probabilidade, i.e., toda a variabilidade espacial é capturada pelos parâmetros, que eventualmente se propaga para as observações. No entanto, tal hipótese é questionável quando se trata de variáveis climáticas (COOLEY *et al.*, 2012). A dependência espacial entre observações concomitantes de vazões é, na verdade, reconhecida na literatura. Para tanto, Stedinger e Tasker (1985a) avocam as vantagens do método GLS na estimação de parâmetros de modelos regressivos, e Hosking e Wallis (1997), dentro do *framework* do *index flood* com momentos-L, reconhecem que é comum observações serem positivamente correlacionadas e sugerem que esse efeito aumente a variabilidade da curva de quantis e introduza, em pequena quantidade, vies nas estimativas.

Por fim, a dependência espacial pode ser mais profundamente explicada no nível dos dados do MHBs, considerando a correlação cruzada de observações concomitantes. Cooley *et al.* (2012) sugerem duas formas de introduzir a dependência no nível dos dados, a primeira por meio do uso de cópulas e a segunda por meio de processos *max-stable*. O uso de cópulas vêm sendo empregado para representar a dependência das observações em MHBs. Bracken *et al.* (2016) e Renard (2011) utilizaram cópulas elípticas (e.g., *t-Student* e Gaussiana) para modelar a distribuição conjunta de postos pluviométricos.

Apesar da dependência espacial da chuva ser intuitivamente bem representada em termos de distância euclidiana por meio de semivariogramas, o uso desses para vazões é questionável, devido a estrutura complexa por trás da dependência imposta pela rede de drenagem. A hierarquia da rede de drenagem parece influenciar tanto a dependência espacial dos resíduos dos modelos quanto dos dados. Costa e Fernandes (2021), por exemplo, relatam que as previsões do modelo regressivo para curvas de permanência foram prejudicadas por não considerar a hierarquia da rede de drenagem. Eles sugerem que a inclusão dessa estrutura de dependência nos resíduos e nas áreas incrementais de bacias aninhadas na inferência poderiam melhorar as previsões. No nível das observações, Guse *et al.* (2009) mostram que a

correlação cruzada entre pares de observações de vazões máximas de bacias não aninhadas e aninhadas são diferentes e sugerem que modelos distintos de correlação cruzada devem ser utilizados para essas duas configurações (Figura 3.5).

Figura 3.5 - Funções de correlação cruzada ajustadas globalmente e para bacias aninhadas e não aninhadas para dados de vazões máximas (T&S: função de correlação cruzada de Tasker e Stedinger, 1989).



Fonte: adaptado de Guse et al., 2009

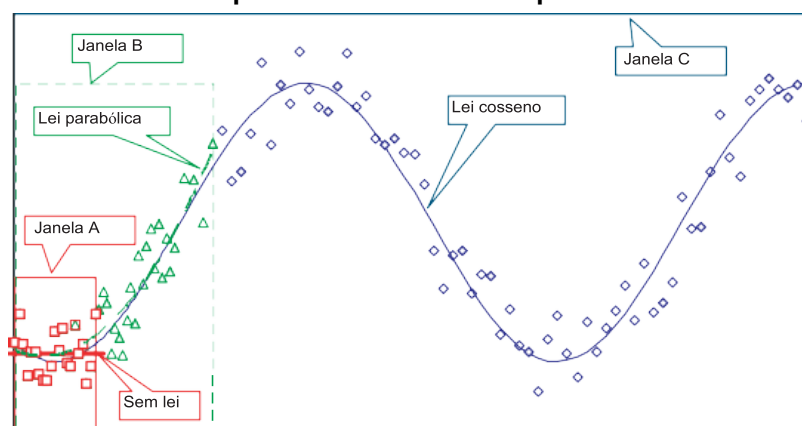
3.3.2 Dependência temporal

Nas últimas décadas, a comunidade científica de recursos hídricos vem crescentemente abordando a não estacionariedade. Milly *et al.* (2008) criaram um debate com o trabalho “*Stationarity is dead: whither water management?*”, no qual alegam que atividades antrópicas estão interferido substancialmente no clima do planeta e alterando os processos físicos do ciclo hidrológico. Milly *et al.* (2008) sugerem que um substituto para a hipótese de estacionariedade seja desenvolvido. Muitas críticas (e.g. LINS; COHN, 2011; MATALAS, 2012; MONTANARI; KOUTSOYIANNIS, 2014) foram direcionadas à declaração de Milly *et al.* (2008).

Detectar a não-estacionariedade pode ser complicado por uma série de fatores (LINS; COHN, 2011): séries curtas facilmente exibem tendências em decorrência de variações multidecenais ou seculares; a significância estatística é fortemente dependente da escolha do teste de hipótese; o início e o final da tendência, escolhidos pelo analista, afetam diretamente a magnitude e a direção da tendência, e realizações de processos de memória longa (estacionários) podem exibir características similares

àquelas de processos não-estacionários no caso de janelas de observação reduzidas (MONTANARI; KOUTSOYIANNIS, 2014). A grande crítica em relação a não estacionariedade consiste na inferência de tendências (que por si só já são de difícil detecção, conforme discutido anteriormente) sem embasamento físico sobre a mudança ocorrida no processo do mundo real. Por exemplo, Koutsoyiannis (2006) observou que propagações de tendências lineares não confiáveis de modelos determinísticos (i.e. sem evidência de persistência da tendência histórica), no horizonte de previsão, podem gerar valores irrealistas (no caso dos autores, vazões negativas). A tentativa de inferir tendências a partir de uma amostra curta de dados (cenário comum em séries de variáveis hidrológicas), ignorando-se o processo hidrológico que lhe conduz, pode originar modelos com excelente ajuste à amostra e às expectativas do analista, mas que falham em representar a população oriunda do processo hidrológico, que deveria ser o objeto de interesse (Figura 3.6).

Figura 3.6 - Tentativas de inferência de tendências para diferentes tamanhos de amostras de um processo estocástico hipotético



Fonte: adaptado de Koutsoyannis (2006)

Lins e Cohn (2011) defendem que, na ausência de informações, a melhor alternativa seria aceitar a ignorância sobre o processo e manter a hipótese de estacionariedade. Por outro lado, quando há informações sólidas ou conhecimento de fatores potencialmente associados à mudança no processo, modelos não estacionários tornam-se uma opção viável (KOUTSOYIANNIS, 2006; SERINALDI; KILSBY; LOMBARDO, 2018). Stedinger e Griffis (2011), por exemplo, argumentam que é justificável o uso de índices climáticos como variável explicativa para modelar a variabilidade climática. O ponto crítico, ainda segundo os autores, é distinguir um sinal suficientemente forte para incluí-lo na análise e trazer melhorias. Outros exemplos de mudanças que podem implicar não estacionariedade são urbanização da bacia

hidrográfica (LINS; COHN, 2011) e implementação de estruturas hidráulicas (KOUTSOYIANNIS; MONTANARI, 2015).

A dependência temporal em MHBs é comumente introduzida no nível do processo, por meio de métodos regressivos nos quais covariáveis temporais são utilizadas como preditoras dos parâmetros da distribuição de probabilidade marginal (e.g., MEHMOOD *et al.*, 2019; SUN *et al.*, 2014). A seguir são mostrados alguns exemplos de covariáveis temporais utilizadas em modelos não estacionários de variáveis hidrológicas.

3.3.2.1 Ano hidrológico das medições

A utilização do ano hidrológico das observações (i.e. um índice de tempo) é a abordagem mais simples de modelos não estacionários encontrados na literatura. Normalmente se estabelece um modelo de regressão linear entre os parâmetros da distribuição de probabilidades dos dados e um índice de tempo (e.g., BADJANA *et al.*, 2017; BARTIKO; CHAFFE; BONUMÁ, 2017; DO VALE MOREIRA; NAGHETTINI; ELEUTÉRIO, 2016; SUN *et al.*, 2015; WANG; XUAN, 2020). Os estudos que seguem essa abordagem geralmente visam identificar a existência de tendências temporais, as quais, por vezes, são atribuídas a um agente de mudança do processo físico gerador da variável aleatória em análise (e.g., mudança no uso do solo, fenômenos climáticos de larga escala). As atribuições de mudanças temporais nas séries de vazões a agentes são fortemente criticadas por Merz *et al.* (2012) em alguns casos. Esses autores, em uma revisão de estudos de tendências de vazões, concluíram que muitas dessas atribuições são baseadas em raciocínio qualitativo ou mesmo especulação.

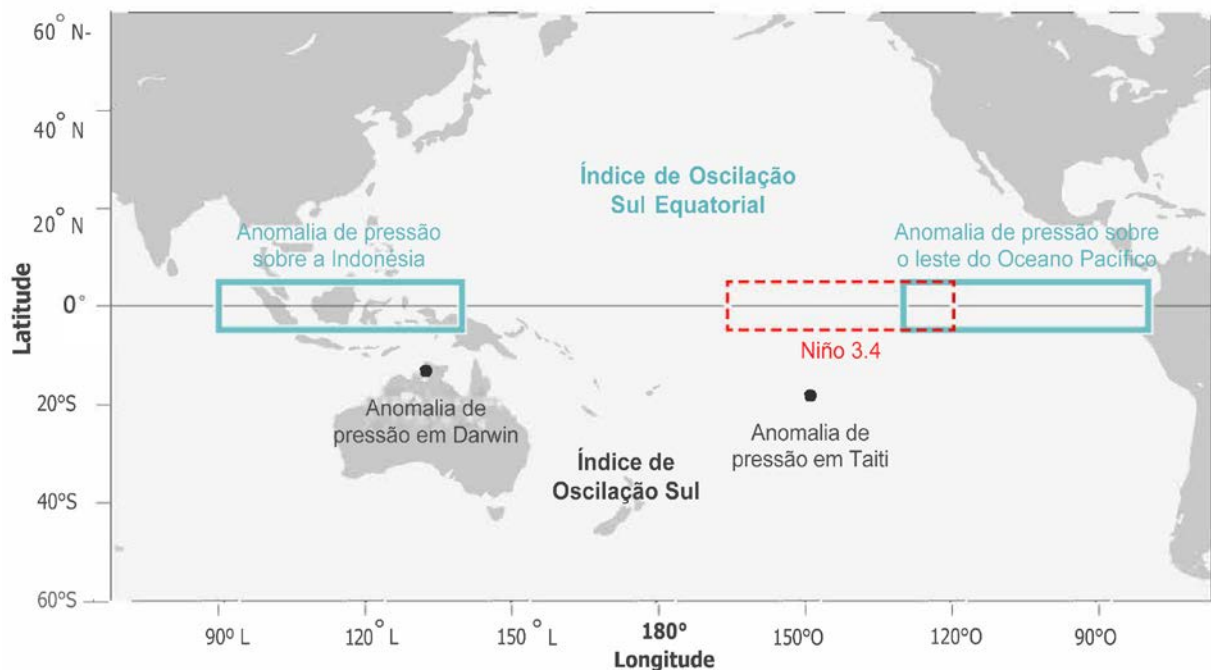
3.3.2.2 Índices climáticos

Índices climáticos são elaborados para resumir um fenômeno climático em uma única série temporal. Os índices climáticos usualmente trazem anomalias de uma variável com base em períodos de referência, e seus valores anuais usualmente passam por um filtro de média móvel. Um dos fenômenos climáticos mais populares é o El Niño Oscilação Sul (ENSO, do acrônimo em inglês *El Niño Southern Oscillation*), associado a mudança periódica da temperatura do oceano pacífico, ora mais quente (El Niño) e ora mais frio (La Niña) do que a temperatura média (CAI *et al.*, 2020). O ENSO está

relacionado a variações climáticas em todo o globo, cujas consequências se manifestam em termos socioeconômicos e ambientais (CAI *et al.*, 2020). O ENSO, por exemplo, pode ser descrito por diversos índices, dentre eles o Niño 3.4, Índice de Oscilação Sul (SOI) e SOI Equatorial (Figura 3.7). Os índices podem representar anomalias de temperatura da superfície do mar (Niño 3.4), de pressão atmosférica (SOI), dentre outras variáveis.

Figura 3.7 - Três índices utilizados para capturar a variabilidade do ENSO

Índices do ENSO



Fonte: Adaptado de climate.gov

A teleconexão entre índices climáticos e variáveis hidrológicas foi objetivo de muitos estudos. Capozzoli *et al.* (2017) investigaram padrões de variabilidade da vazão das principais bacias brasileiras em função de índices climáticos por meio de 163 postos fluviométricos. Os resultados dos autores são resumidos na Tabela 3.2, na qual observa-se que todas as grandes bacias brasileiras apresentaram correlação significativa com índices climáticos.

Tabela 3.2 - Bacias destacadas em cada um dos Componentes Principais (CPs) sinal (+) indicam correlações positivas e (-) negativas, escala temporal dos picos espectrais das séries temporais dos CPs e índices de teleconexão que apresentam correlação significativa com as series temporais dos CPs onde sinal (+) indicam correlações positivas e (-) negativas.

CP	Bacias destacadas	Série Temporal do CP	
		Picos espectrais significativos	Correlação significativa com índices de teleconexão
1	Paraná (Oriental) (+), Tocantins (+), Atlântico Leste (+) e São Francisco (+)	Semianual, interanual e decadal	Niño 3.4(+), PDO (+), IOS (-), AMO (-), TSA (-)
2	Uruguai (+), Atlântico Sudeste (+), Paraná (centro-sul) (+), Atlântico Leste (-), alto Paraná (-)Tocantins (-) e São Francisco (-)	Interanual e semianual	Niño 3.4(+), TSA (+), IOD (+), NAO (+), IOS (-), AMM (-)
3	Paraná (central e leste) (+), baixo Paraná (-), Uruguai (-) e Atlântico Sudeste (-)	Interanual, interdecadal e intrasazonal	IOS (+), AMO (+), AAO (+), PDO (-), Niño 3.4(-)
4	Paraná (norte) (+), Amazônica (sul) (+), Tocantins (norte) (+), Atlântico Leste (sul) (-)	Intrasazonal e interanual	TSA (+), PDO (+), AO (+), AAO (+)
5	Uruguai (+), Atlântico Sudeste (leste) (+), Paraná (central-sul) (-), Paraná (norte) (+), São Francisco (-), Atlântico Leste (-)	Semianual e intrasazonal	Niño 3.4(+), TSA (+), PDO (+), AO (+), IOD (+), AMM (-), IOS (-)
6	Alto Paraná (+), Tocantins (-) e Atlântico Leste (sul) (-) e Amazônica (-)	Semianual a intrasazonal	AMO (+), Niño 3.4 (+), IOS (-), TSA (-)

Fonte: adaptado de Capozzoli *et al.* (2017)

Índices climáticos são frequentemente utilizados na análise de frequência não estacionária para descrever a variabilidade temporal em função de mecanismos de larga escala. Bracken *et al.* (2018) utilizaram o Índice Multivariado do ENSO (MEI), Oscilação Decadal do Pacífico (PDO) e Oscilação Multidecadal do Atlântico (AMO) como covariáveis do parâmetro de posição da distribuição Generalizada de Valores Extremos. Ouarda e Charron (2019), por sua vez, utilizaram os índices SOI, PDO e Pacífico Norte Americano (PNA) em regressões associadas aos parâmetros da distribuição GEV na análise de frequência de vazões máximas.

3.3.2.3 Transporte de umidade atmosférico

Steinschneider e Lall (2015) utilizaram o mecanismo persistente de transporte de umidade dos oceanos tropicais, denominado Exportações de Umidade Tropical (*Tropical Moisture Exports – TME*, em inglês), como covariável temporal na AFR de extremos de chuva no Oeste dos Estados Unidos. Essa abordagem requisitou a identificação das TMEs, a partir de vetores tridimensionais de campos de vento, vapor de água e acompanhamento das trajetórias das massas de ar. Os resultados dos autores sugerem que o modelo com a magnitude e frequência das excedências associados a covariável TME tem melhor desempenho do que modelos estacionários.

3.3.2.4 Variáveis de entrada para índices climáticos

A Temperatura da Superfície do Mar (TSM), assim como as alturas geopotenciais e outras variáveis, são a base de muitos índices climáticos. Por exemplo, a TSM é utilizada nos índices Niño 3.4, 1+2, 3 e 4, Oscilação Multi-decadal do Atlântico (AMO), Oscilação Decadal do Pacífico (PDO), entre outros, e as alturas geopotenciais são utilizadas, por exemplo, no índice Oscilação do Atlântico Norte (NAO). É natural, portanto, que essas variáveis sejam candidatas para serem empregadas em modelos não estacionários.

Renard e Lall (2014) propuseram um índice climático customizado a partir de um campo de alturas geopotenciais sobre o Oceano Atlântico norte para a previsão de enchentes no outono de 16 bacias localizadas na França. No modelo elaborado pelos autores, o peso dos elementos dos campos climáticos era inferido junto com o modelo e a influência do índice climático customizada era uniforme em todos os postos.

Ahn *et al.* (2017), em um estudo de previsão de vazões mínimas de verão no nordeste dos Estados Unidos, aplicaram um conceito similar ao de Renard e Lall (2014). Os autores utilizaram um índice climático customizado a partir da TSM no Oceano Atlântico do Norte. No modelo elaborado pelos autores, a influência do índice climático não era uniforme na bacia. Zeng *et al.* (2017) também utilizaram a TSM do nordeste do Oceano Índico e o oeste do Oceano Pacífico na AFL de extremos máximos de vazão e precipitação para um reservatório na China.

4 MATERIAL E MÉTODOS

4.1 Descrição da área de estudo

A área de estudo será composta por duas bacias hidrográficas, a bacia do rio Itajaí-Açu e a bacia do rio Doce (Figura 4.1). A região Sul do Brasil é conhecida por ter sinais fortes de teleconexão com variáveis climáticas (e.g. BARREIRO, 2010), enquanto a região de Minas Gerais não o é. Esse esquema permitirá avaliar o modelo não-estacionários em duas situações distintas de intensidade de sinais de teleconexão.

Figura 4.1 - Mapa de localização das bacias em estudo



Mapas da precipitação média anual⁵, elevação⁶, cobertura e uso do solo⁷, hidrogeologia⁸ e ocorrência de cavernas⁹ podem ser consultados no APÊNDICE A –

4.1.1 Bacia do rio Doce

A bacia do rio Doce, pertencente a região hidrográfica do Atlântico sudeste e drena uma área de aproximadamente 83.400 km², da qual 86% está localizada em Minas

⁵ Projeto Atlas Pluviométrico 1:5.000.000 (CPRM, 2011)

⁶ SRTMGL1 v003 resolução de 30 m (NASA JPL, 2013)

⁷ Levantamento anual da cobertura e uso do solo sintetizado para as classes principais resolução de 30 m (PROJETO MAPBIOMAS, 2020)

⁸ Mapa hidrogeológico do sudeste e do sul do Brasil 1:1.800.000 (IBGE, 2015a, b)

⁹ Mapeamento de cavidades naturais subterrâneas (cavernas) e áreas de sua ocorrência (CECAV, 2018, 2019)

Gerais e 14% no Espírito Santo (ELESBON *et al.*, 2015). Aproximadamente 3,1 milhões de pessoas residem na bacia.

A mudança da cobertura e ocupação da bacia tem origem nos primeiros séculos da colonização portuguesa. Com a descoberta do ouro na região em 1698, as cabeceiras do rio Doce passaram a ser amplamente povoadas (HORA *et al.*, 2012). O movimento migratório gerou demandas alimentícias e foi o catalizador das primeiras atividades agrícolas na região. O desflorestamento foi acentuado a partir do início do século XX com a construção da Estrada de Ferro Vitória-Minas, que potencializou a ocupação do território da bacia (FELIPPE *et al.*, 2016). Originalmente, 90% da bacia era coberta por florestas, porém hoje em dia a maior parte dela está antropizada, principalmente pelo uso agropecuário (IGAM, 2010). Apesar disso, nos últimos anos não houve significativas alterações na porcentagem das macro classes da cobertura da bacia (Tabela 4.1).

Tabela 4.1 - Cobertura e uso do solo na bacia do rio Doce nos anos 1985 e 2018

Classe	1985	2018
Floresta	28,0%	29,2%
Formação Natural não Florestal	1,5%	1,8%
Agropecuária	69,5%	67,7%
Área não vegetada	0,5%	0,8%
Corpos D'água	0,5%	0,6%

Fonte: MapBiomias¹⁰ (2020)

O clima da bacia é heterogêneo em função da sua grande extensão territorial. Segundo a classificação de Köppen realizada por Alvares *et al.* (2013), as altitudes médias e baixas da bacia se enquadram no clima tropical com inverno seco (Aw) e a região de altitudes mais elevadas, onde a temperatura é inferior e a precipitação é maior devido à orografia, contém um mosaico de clima subtropical úmido com inverno seco e verão quente (Cwa) ou temperado (Cwb).

A precipitação média anual e a evapotranspiração real na bacia do rio Doce são, respectivamente, de 1.238 e 827 mm (MMA, 2005 *apud* MMA, 2006). A estimativa da precipitação média anual pela interpolação da isoietas de CPRM (2011) fornece o valor

¹⁰ Projeto MapBiomias - é uma iniciativa multi-institucional para gerar mapas anuais de cobertura e uso do solo a partir de processos de classificação automática aplicada a imagens de satélite. A descrição completa do projeto encontra-se em <http://mapbiomas.org>. A acurácia global da versão 4.1 é de 90,2%.

de 1265 mm. O regime pluviométrico da bacia é composto por dois períodos bem definidos: período seco (abril a setembro) e período chuvoso (outubro a março). No período seco, há atuação de frentes frias com baixo teor de umidade e do Anticiclone Subtropical do Atlântico Sul (ASAS)¹¹, que resultam em um inverno seco na região (CUPOLILLO, 2008). No período chuvoso, chuvas persistentes incidem na bacia devido a associação de sistemas convectivos e sistemas frontais, que originam a Zona de Convergência do Atlântico Sul (ZCAS) (CUPOLILLO, 2008).

A bacia do rio Doce, assim como a região Sudeste do Brasil, enfrentou nos últimos anos um período extenso e severo de estiagem. Uma seca meteorológica incidiu na bacia do rio Doce no período 2013 a 2017 (DE JESUS *et al.*, 2020). A seca de 2014/15, em especial, foi bastante severa e gerou uma grave crise hídrica no sudeste do Brasil com impactos no sistema de abastecimento e geração de energia hidrelétrica (CUNHA *et al.*, 2019). A seca do verão austral de 2014 foi estudada por Coelho *et al.* (2016), que além de realizar um diagnóstico desse evento, identificaram que a temperatura da superfície do mar (TSM) próxima à costa sudeste do Brasil possui correlação negativa forte com a precipitação de janeiro-fevereiro-março na região Sudeste. Em um estudo recente, Rodrigues *et al.* (2019) identificaram o mecanismo responsável pela supressão da Zona de Convergência do Atlântico Sul (ZCAS) que teria atuado não só no verão austral de 2014, como também em outros verões, e resultado em anomalias negativas de precipitação (Figura 4.2). A descrição do mecanismo é apresentada por, Rodrigues *et al.* (2019):

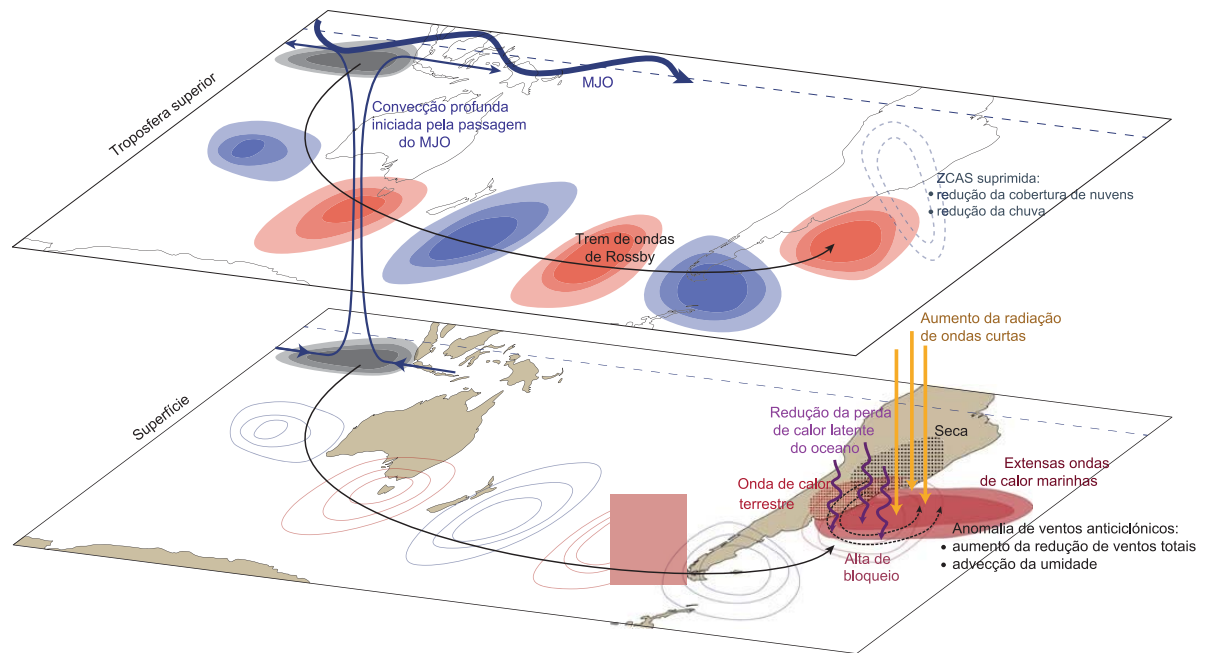
*A convecção tropical profunda, principalmente no Oceano Índico, associada à passagem MJO¹², aciona um trem de ondas Rossby que se estende pelo Pacífico Sul até a ponta da América do Sul e vira em direção ao equador para alcançar a América do Sul subtropical. Aí é estabelecido um anticiclone bloqueador que suprime a ZCAS, reduzindo a cobertura de nuvens e a precipitação. Isso, por sua vez, aumenta a entrada de radiação de ondas curtas no oceano. Ao mesmo tempo, a perda de calor latente do oceano é reduzida devido aos ventos mais fracos e à advecção de umidade aumentada. O calor é retido no oceano, o que leva a extensas OCMs¹³ (RODRIGUES *et al.*, 2019).*

¹¹ Sistema de alta pressão localizado sobre o sul do Oceano Atlântico

¹² Oscilação de Madden e Julian (MJO) – célula de convecção tropical que viaja do oeste para o leste num período de 30 a 60 dias.

¹³ Onda de calor marítima (OCM).

Figura 4.2 - Representação esquemática do mecanismo forçador das ondas de calor marítima do oeste do Atlântico Sul



Fonte: adaptado de Rodrigues et al. (2019)

Na bacia do rio Doce, onde existe um período de estiagem bem definido, o escoamento de base é um fator importante na manutenção do regime de estiagem. Nessa conjuntura, o domínio e a produtividade do aquífero podem influenciar variações espaciais no comportamento de vazões mínimas. A maior parte da bacia está inserida no domínio de aquífero fissural, sendo sua maior parte de produtividade baixa. Uma parcela substancial do domínio fissural, localizada ao longo da longitude de 42º, em especial, possui produtividade alta. Ademais, ainda existe uma parcela significativa do território de depósitos aluvionares no sudoeste da bacia (região do rio Casca e cabeceira do rio Doce), os quais constituem um aquífero poroso de produtividade média.

4.1.2 Bacia do rio Itajaí-Açu

A bacia hidrográfica do rio Itajaí-Açu está localizada no estado de Santa Catarina, na região Sul do Brasil. A bacia do rio Itajaí-Açu faz parte da bacia hidrográfica do Atlântico Sul e drena uma área aproximada de 15.000 km², na qual cerca de um 1,1 milhões de pessoas habitam (COMITÊ DO ITAJAÍ, 2010).

A bacia está inserida no bioma da Mata Atlântica. Em relação ao uso e cobertura do solo da bacia, Vibrans (2003), baseado num compilado de registros históricos, estima

que o percentual de áreas desmatadas na bacia tenha sido de 8% em 1860, subindo para 26% em 1880 e chegando a 50% em 1917. Estimativas de cobertura e uso do solo das últimas décadas (Tabela 4.2) indicam que o desflorestamento se estabilizou e em torno de 67% da bacia era coberta pela classe de floresta em 2018. O segundo uso do solo mais recorrente é o agropastoril.

Tabela 4.2 - Cobertura e uso do solo na bacia do rio Itajaí-Açu nos anos 1985 e 2018

Classe	1985	2018
Floresta	65,6%	67,6%
Formação Natural não Florestal	0,0%	0,0%
Agropecuária	33,3%	29,8%
Área não vegetada	0,9%	2,4%
Corpos D'água	0,2%	0,2%

Fonte: MapBiomias (2020)

Na classificação de Köppen realizada por Alvares *et al.* (2013), o clima da bacia se enquadra nas categorias: subtropical oceânico sem estação seca e com verão quente (Cfa) ou temperado (Cfb).

A distribuição da precipitação na bacia, sob influência da maritimidade e continentalidade, aumenta no sentido oeste-leste, variando de 1500 a 1900 mm (COMITÊ DO ITAJAÍ, 2010). A estimativa de precipitação média anual pela interpolação da isoietas de CPRM (2011) é de 1607 mm. Não há uma estação seca ou mês seco na região. Apesar disso, o regime pluviométrico possui uma estação chuvosa principal no verão (dezembro a março), uma estação chuvosa secundária na primavera (setembro e outubro) e um período de 5 meses menos chuvoso no outono/inverno (abril a agosto) (COMITÊ DO ITAJAÍ, 2010). Na região onde a bacia está inserida, a precipitação está associada principalmente a sistemas frontais originados no Pacífico, ciclones e frentes frias que se desenvolvem na região, complexos convectivos de mesoescala, sistemas ciclônicos e bloqueios atmosféricos (REBOITA *et al.*, 2010).

A bacia do rio Itajaí-Açu está localizada na porção sudeste da América do Sul, região conhecidamente sob influência do ENSO. Diversos estudos relacionam o ENSO aos totais e extremos de precipitação (CAI *et al.*, 2020; FERNANDES; RODRIGUES, 2018; GRIMM; TEDESCHI, 2009; TEDESCHI; GRIMM; CAVALCANTI, 2015, 2016) e extremos máximos de vazão (WARD *et al.*, 2014). Na região na qual a bacia do rio Itajaí-Açu se encontra, o El Niño influencia significativamente as precipitações do

período de novembro a junho e as estiagens são correlacionadas com anomalias frias de TSM no oceano Pacífico (PENALBA; RIVERA, 2016).

O domínio hidrológico da bacia do rio Itajaí-Açu é bem definido. A porção oeste da bacia é constituída por um mosaico de unidades de aquífero poroso de produtividade baixa. Ao leste da confluência do rio Itajaí do Norte com o rio Itajaí-Açu, o regime de aquíferos passa a ser do tipo fissural de baixa produtividade, com exceção de depósitos aluvionares ao longo dos principais cursos d'água.

4.2 Aquisição e tratamento de dados

Os dados de vazão mínima mensal na bacia do rio Doce foram obtidos do estudo de regionalização de vazões mínimas sazonais para o Estado de Minas Gerais (Termo de Execução Descentralizada nº 08/2018/ANA). Para automatizar a elaboração das séries de vazão mínima anual, adotou-se como critério a exclusão de períodos com falhas nas séries de vazão mínima sazonal nos três meses finais da estação de estiagem (julho a setembro) e o primeiro mês da estação chuvosa (outubro). A incorporação do mês de outubro no critério de exclusão foi vista como necessária, pois por vezes há um atraso no início do período chuvoso ou mesmo poucas contribuições que levam a depleção das reservas subterrâneas a se estender além do período de estiagem.

Dados de vazão média diária na bacia do rio Itajaí-Açu foram obtidos por meio da plataforma digital da Agência Nacional de Águas (ANA) (<http://www.snirh.gov.br/hidroweb>). Nesse caso, a vazão média diária era conhecida e isso permitiu que a determinação da vazão mínima anual fosse avaliada anualmente considerando os períodos de falha e a recessão do hidrograma. As amostras de blocos de mínimos dos postos fluviométricos foram organizadas considerando o intervalo outubro-setembro, de forma a reduzir a chance dos mínimos da série sejam dependentes entre si.

Séries de vazões mínimas podem apresentar autocorrelação serial devido a efeitos do abastecimento subterrâneo, regulação de lagos, armazenamento em áreas úmidas e na calha do curso d'água (SALAS *et al.*, 2019). Em função dessa característica da variável em estudo, as séries de Q_7 mínima anual dos postos fluviométricos restantes foram submetidas a uma análise de independência por meio da função de

autocorrelação (ACF) em busca de correlação significativa de memória curta (i.e. dependência serial). Os postos fluviométricos com dependência serial não foram utilizados para nenhuma inferência ou análise, uma vez as observações não são independentes, o que firma um conflito teórico advindo da má representação da probabilidade associada às observações. Em termos práticos, a utilização desses postos também pode atrapalhar a identificação de uma estrutura temporal entre a covariável temporal estudada (a ser discutido na seção 4.4.2.1) e as vazões mínimas, uma vez que a série já possui um componente não estacionário ligado à memória do processo.

A estimação do risco e de quantis de vazões mínimas em séries que exibem autocorreção pode ser conduzida por outras abordagens, e.g., modelos autorregressivos de média móvel (ARMA) e modelos de média móvel (SALAS *et al.*, 2019). Essas abordagens permitem considerar o efeito de dependência serial, a qual traz implicações em um contexto de gestão de recursos hídricos. A persistência temporal observada nessas séries indica que anos secos tem maior probabilidade de serem seguidos por anos secos. Em um ambiente de previsão, essas abordagens podem auxiliar os gestores de recursos hídricos a entrar com medidas preventivas de restrição do uso d'água ou mesmo auxiliar os usuários a se planejarem antecipadamente para um cenário de escassez hídrica.

Adotou-se como critério de no mínimo de 10 anos de dados disponíveis neste estudo de regionalização. Estações fluviométricas com 8 e 9 anos de dados, um pouco abaixo do critério estabelecido, foram incorporadas ao banco de dados com a finalidade de serem utilizadas apenas para validação dos modelos elaborados, uma vez que a inferência com séries curtas está sujeita a maior viés.

A caracterização física da área de contribuição das estações fluviométricas foi realizada por meio dos atributos: área de drenagem, precipitação anual, declividade média, densidade de drenagem e vazão do aquífero. A área de drenagem foi obtida do inventário da ANA. A precipitação média anual foi obtida a partir da interpolação das isoietas de CPRM (2011). A declividade das bacias e a rede de drenagem foram ambas obtidas a partir do MDE SRTM (NASA JPL, 2013) de 30 m. No entanto, para a bacia do rio Itajaí, a densidade foi obtida a partir da BHO. Não foi possível utilizar os dados de drenagem da BHO da bacia do rio Doce devido à multiescala empregada

em sua metodologia – 1:50.000 nas cabeceiras da bacia e 1:100.000 no restante da bacia. Por fim, a produção do aquífero¹⁴ foi determinada a partir de uma ponderação com a área dos aquíferos localizados na área de drenagem dos postos. A produção dos aquíferos é quantificada em termos de faixas de valores, por isso, adotou-se a média dessas faixas para representar as classes de aquífero (Tabela 4.3).

Tabela 4.3 - Caracterização das classes de produção dos aquíferos e os valores adotados

Classe	Faixa de produção (m ³ /h)	Valor adotado (m ³ /h)
Muito baixa	0-3	1,5
Baixa	3-10	6,5
Média	10-40	25
Alta	40-100	70

4.3 Modelo de distribuição de probabilidade

Variáveis aleatórias hidrológicas, como vazão e precipitação, não podem ser infinitamente amostradas, uma vez que dependem do processo natural de geração dessas variáveis. Nesses casos, a amostragem é finita e pode ser vista como realização da variável aleatória. As realizações de uma variável aleatória X , com distribuição $F_X(x)$, podem ser representadas por $x_i \in \{x_1, x_2 \dots, x_N\}$, onde N é a n ésima realização da variável X . É de interesse do especialista conhecer a probabilidade excedência ou não-excedência de quantis específicos para projeto de obras hidráulicas e estabelecimento de vazões de outorga, dentre outros motivos. Para isso, é necessário inferir a distribuição da variável aleatória X a partir de suas x_i realizações.

Muitos estudos tentaram avaliar a melhor função de distribuição de probabilidade para a vazão mínima em várias regiões do globo, no entanto, é improvável que uma única distribuição seja universalmente aceita (SMAKHTIN, 2001). As distribuições mais frequentemente citadas na literatura para vazões mínimas de cursos d'água perenes

¹⁴ Obtido a partir do conjunto de dados de hidrogeologia da região Sul e Sudeste do IBGE (escala 1:1.800.00). Disponível em: <https://www.ibge.gov.br/geociencias/informacoes-ambientais/geologia/15824-hidrogeologia.html?=&t=downloads>. Acesso em: 01 de outubro de 2021.

são GEV, Gumbel, Weibull de 2 e 3 parâmetros (W2P e W3P), LogNormal de 2 e 3 parâmetros (LN2 e LN3), e LogPearson e Pearson de três parâmetros (respectivamente, LP3 e P3) (BHATTI; KROLL; VOGEL, 2019; GRANEMANN; MINE, 2015; SMAKHTIN, 2001).

A escolha do modelo de distribuição de probabilidade constitui uma fonte de incerteza na análise de frequência e pode ser incorporada na inferência por meio da abordagem Bayesiana (LIANG; CHANG; LI, 2012; MERZ, BRUNO; THIEKEN, 2005). Esse procedimento seria demasiadamente custoso, assim como conceber diversos MHB para contemplar as distribuições candidatas e escolher o melhor modelo com base em métricas de ajuste. Por simplicidade, optou-se por estabelecer *a priori* a distribuição de probabilidade a ser empregada no MHB. Ao contrário de estudos de vazões máximas, o foco deste trabalho é voltado para quantis associados a maiores frequências de superação, os quais são menos sensíveis à escolha da distribuição de probabilidade.

Foram avaliadas as distribuições GEV, W2P, W3P, LN2, LN3, P3, LP3 e a distribuição Gumbel com cauda exponencial inferior e superior (respectivamente, GUM⁻ e GUM⁺ – também chamadas, respectivamente, de Gumbel para mínimos e Gumbel para máximos) como candidatas ao modelo de distribuição de probabilidade. Ambos os modelos de distribuição Gumbel possuem a moda igual ao parâmetro de posição, porém a distribuição Gumbel com cauda exponencial inferior tem assimetria negativa e aquela com cauda superior tem assimetria positiva. Apesar do pressuposto comum de que vazões mínimas são negativamente assimétricas, esse fato nem sempre é observado (como em BHATTI; KROLL; VOGEL, 2019; CHEN *et al.*, 2006; CPRM, 2001; LEME; CHAUDHRY, 2005; MELO, 2016; MODARRES, 2008). A parametrização das distribuições avaliadas é apresentada na Tabela 4.4.

Após a proposição das distribuições candidatas, foi necessário estimar seus parâmetros. Decidiu-se seguir essa abordagem inicial sob a ótica frequentista, embora a abordagem Bayesiana também pudesse ser empregada, devido a existência de rotinas de programação que facilitariam o ajuste das distribuições candidatas nos postos fluviométricos. Em um contexto frequentista, a estimação de parâmetros de distribuições de probabilidade é comumente realizada utilizando um dos seguintes

métodos: método dos momentos, método dos momentos-L e método da máxima verossimilhança.

O método dos momentos, apesar de ser facilmente implementado, depende das estimativas de momentos amostrais, que por sua vez, tem grande potencial de serem viesadas em amostras pequenas (como boa parte dos dados disponíveis de vazão) e especialmente para distribuições de três ou mais parâmetros. O método da máxima verossimilhança geralmente estima quantis e parâmetros com as menores variâncias amostrais (NAGHETTINI, 2017a). Apesar disso, assim com o método dos momentos, pode levar a estimativas de pior qualidade com amostras curtas. Além disso, o método requer resolução numérica de equações não-lineares, cujos resultados podem divergir em função do método empregado em sua resolução. Como alternativa, o método dos momentos-L tem, em muitas vezes, mais acurácia do que o método da máxima verossimilhança para amostras pequenas (RAO; HAMED, 2000). Além de envolver a resolução de equações de menor complexidade que o método da máxima verossimilhança, o método é robusto à presença de *outliers*. Em função de suas características e fácil implementação, foi utilizado o método dos momentos-L para estimar os parâmetros. A utilização desse método é facilitada pela existência de funções de ajuste por meio do pacote `lmom` (HOSKING, 2019) do R. Para mais detalhes sobre os métodos de estimação, sugere-se ao leitor consultar Naghettini (2017a) e Rao e Hamed (2000).

A escolha da distribuição de probabilidade foi avaliada por três vertentes: testes de aderência, métricas de ajuste e a análise exploratória de preditores para os parâmetros das distribuições de probabilidade.

Tabela 4.4 - Distribuições de probabilidades candidatas

Distribuição de probabilidade	Função de densidade de probabilidade (f_X)	Domínio dos parâmetros
GEV	$f_X(x) = \frac{1}{\sigma} \left[1 + \xi \left(\frac{x - \mu}{\sigma} \right) \right]^{-1/\xi+1} \exp \left\{ - \left[1 + \xi \left(\frac{x - \mu}{\sigma} \right) \right]^{-1/\xi} \right\}$	Se $\xi = 0$: (EV1 Gumbel) Se $\xi < 0$: (EV2 Fréchet) $x > \mu - \sigma/\xi$

		Se $\xi > 0$: (EV3 Weibull) $x < \mu + \sigma/\xi$
Gum ⁺	$f_X(x) = \frac{1}{\sigma} \exp \left[-\frac{x-\mu}{\sigma} - \exp \left(-\frac{x-\mu}{\sigma} \right) \right]$	$\sigma > 0$
Gum ⁻	$f_X(x) = \frac{1}{\sigma} \exp \left[\frac{x-\mu}{\sigma} - \exp \left(\frac{x-\mu}{\sigma} \right) \right]$	$\sigma > 0$
W2	$f_X(x) = \frac{\xi}{\sigma} \left(\frac{x}{\sigma} \right)^\xi \exp \left[-\left(\frac{x}{\sigma} \right)^\xi \right]$	$x \geq 0, \sigma \geq 0, \xi > 0$
W3	$f_X(x) = \frac{\xi}{\sigma} \left(\frac{x-\mu}{\sigma} \right)^\xi \exp \left[-\left(\frac{x-\mu}{\sigma} \right)^\xi \right]$	$x \geq \mu, \sigma \geq 0, \xi > 0$
LN2	$f_X(x) = \frac{1}{x\sigma\sqrt{2\pi}} \exp \left\{ -\frac{[\ln(x) - \mu]^2}{2\sigma^2} \right\}$	$x \geq 0, \mu > 0, \sigma > 0$
LN3	$f_X(x) = \frac{1}{(x-a)\sigma\sqrt{2\pi}} \exp \left\{ -\frac{[\ln(x-a) - \mu]^2}{2\sigma^2} \right\}$	$x > a, \mu > 0, \sigma > 0$
Ga	$f_X(x) = \frac{1}{\sigma^\xi \Gamma(\xi)} x^{\xi-1} e^{-x/\sigma}$	$x \geq 0, \sigma > 0, \xi > 0$
P3	$f_X(x) = \frac{1}{\sigma \Gamma(\xi)} \left(\frac{x-a}{\sigma} \right)^{\xi-1} e^{-(x-a)/\sigma}$	$x > a, \xi > 0$
LP3	$f_X(x) = \frac{1}{\sigma x \Gamma(\xi)} \left(\frac{\ln(x) - a}{\sigma} \right)^{\xi-1} e^{-(\ln(x)-a)/\sigma}$	$x > a, \xi > 0$

Empregou-se o teste de aderência *Probability Plot Correlation Coefficient* (PPCC). O teste avalia se a distribuição sob investigação seria um modelo plausível para a distribuição da variável em mãos. Para determinar a aderência da distribuição aos dados observados, o teste compara a correlação linear dos quantis empíricos e teóricos em ordem crescente com um valor crítico de referência. A correlação é calculada pela Eq. 4.1.

$$r = \frac{\sum_{m=1}^N (x_m - \bar{x})(w_m - \bar{w})}{\sqrt{\sum_{m=1}^N (x_m - \bar{x})^2 \sum_{m=1}^N (w_m - \bar{w})^2}} \quad (4.1)$$

na qual x_m é o m-ésimo quantil em ordem crescente, w_m são os quantis teóricos calculados por $w_m = F_X^{-1}(1 - q_m)$, onde q_m denota a probabilidade empírica do elemento de ordem m e F_X é a função de densidade acumulada da distribuição candidata.

Esse teste de aderência tem maior poder de teste em comparação com outros testes convencionais, além de possuir equacionamento simples e intuitivo (NAGHETTINI, 2017b). A decisão de rejeitar a hipótese nula (variável segue a distribuição de probabilidade investigada) parte da comparação da estatística de teste com valores críticos (r_{crit}), na qual rejeita-se a hipótese nula quando $r < r_{crit}$. Neste trabalho os valores críticos de teste foram calculados ao invés de consultar aqueles disponíveis na literatura. Esse procedimento foi necessário pois os valores críticos da literatura são fornecidos em função do tamanho da amostra e do parâmetro de forma da distribuição, que nem sempre abrangem os valores de interesse. Seguiu-se a metodologia de Heo *et al.*, (2008) para a estimação de valores críticos das distribuições de probabilidade candidatas correspondentes a 5% de probabilidade do erro tipo I.

Como métricas de desempenho de ajuste, seguindo Costa, Fernandes e Starick (2020), foram apreciadas as métricas: erro médio absoluto (MAE), raiz do erro médio quadrado (RMSE), percentual de viés (PBIAS), média percentual absoluta do erro (MAPE) e raiz do quadrado médio do erro de predição (RMSPE), cujas formulações são apresentadas nas Eqs. 4.2 a 4.6.

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (4.2)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (4.3)$$

$$PBIAS = 100\% \frac{\sum_{i=1}^n (y_i - \hat{y}_i)}{\sum_{i=1}^n y_i} \quad (4.4)$$

$$MAPE = \frac{100\%}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right| \quad (4.5)$$

$$RMSPE = 100\% \sqrt{\frac{1}{n} \sum_{i=1}^n \left(\frac{y_i - \hat{y}_i}{y_i} \right)^2} \quad (4.6)$$

nas quais n é o tamanho da amostra, y_i é o i -ésimo elemento da amostra e \hat{y}_i é o i -ésimo elemento estimado com a mesma probabilidade de não-superação do elemento y_i .

Por fim, desejou-se avaliar sumariamente a existência de preditores para os parâmetros das distribuições, característica importante para a predição de vazões mínimas em locais não monitorados. A escalabilidade da área de drenagem com os parâmetros da distribuição GEV no contexto de vazões máximas (e.g, LIMA *et al.*, 2016; WU, YENAN *et al.*, 2018; WU, YUN BIAO; XUE; LIU, 2019) é um exemplo de característica desejada. Preditores que expliquem a variação nos parâmetros das distribuições irão diminuir a incerteza preditiva e melhorar as estimativas do modelo Bayesiano.

A avaliação da predição dos parâmetros de cada distribuição foi avaliada por meio da correlação com as covariáveis climáticas: área de drenagem (A), precipitação média anual (P), densidade de drenagem (DD), declividade do curso d'água principal (S) e a produção do aquífero (QSUB).

4.4 Modelo hierárquico Bayesiano

4.4.1 Modelo estacionário

Nesta etapa do trabalho serão construídos MHBs aninhados para permitir a avaliação das estimativas frente a diferentes níveis de complexidade e hipóteses. Uma síntese

dos modelos estacionários elaborados é apresentada na Tabela 4.5, na qual a complexidade aumenta do modelo M1 ao M4.

Tabela 4.5 - Características dos MHBs avaliados

Hipótese	M1	M2	M3 e M4
Independência condicional no espaço dos dados	✓	✓	✗
Independência espacial dos parâmetros	✓	✗	✗

4.4.1.1 Modelo M1

Seja $Y[s, t]$ a vazão mínima anual de sete dias de duração observada no posto fluviométrico $s \{s_1, \dots, s_N\}$ no ano $t \{t_1, \dots, t_M\}$. Dentro no nível hierárquico dos dados, tem-se que a distribuição de probabilidade das observações é representada por meio de uma distribuição de probabilidade marginal $p(\cdot)$ com vetor de parâmetros θ de comprimento K . Formalmente,

$$Y[s, t] \sim p(\theta[\tilde{s}])$$

Nessa concepção, as observações são independentes e os parâmetros do modelo variam no espaço, mas não no tempo.

No nível do processo, um modelo regressivo (Eq. 4.7) é instaurado para relacionar a resposta da variável dependente, os parâmetros da distribuição de probabilidade, com preditores físicos e climáticos que visam explicar a variação espacial dos parâmetros.

$$\log(\theta_k[\tilde{s}_i]) = \alpha_{k0} + (\mathbf{x}_k[\tilde{s}_i])^t \boldsymbol{\alpha}_{k1} + \varepsilon_k[\tilde{s}_i], \quad (\varepsilon_k[\tilde{s}_1], \dots, \varepsilon_k[\tilde{s}_N]) \sim N(0, \delta_k) \quad (4.7)$$

na qual k é um dos K parâmetros de θ , α_0 é o intercepto, $\boldsymbol{\alpha}_1$ é o vetor de coeficientes relacionados ao vetor \mathbf{x} de covariáveis climáticas, ε é o resíduo da regressão, δ é o desvio padrão do resíduo, e seguindo a notação de Renard (2011), $\tilde{s}_i, i = 1, \dots, N$, denota os postos do conjunto de dados de calibração.

Os resíduos do modelo são normalmente distribuídos e independentes, isso é, a correlação entre eles não é formalmente estruturada na inferência, mas caso exista, ela deve se manifestar *a posteriori*.

4.4.1.2 Modelo M2

Sob o indício de que a hipótese dos resíduos da regressão do modelo não são espacialmente independentes (KJELDSEN; JONES, 2007, 2009), a dependência espacial dos resíduos da regressão dos parâmetros da distribuição marginal é introduzida no nível do processo do MHB, sob a forma de um campo aleatório Gaussiano.

Os resíduos da regressão são vistos como realizações de um processo Gaussiano, estacionário e isotrópico, com matriz de variância-covariância determinada por variogramas. Assim como outros estudos, e.g. Yan e Moradkhani (2015), Ahn *et al.* (2017) e Sampaio e Costa (2021), utilizou-se o semivariograma exponencial para modelar a matriz de covariância do processo Gaussiano. A função de covariância é definida como:

$$\Sigma_k = \begin{cases} \delta_k^2 \cdot \exp\left(-\frac{1}{\phi_k} \|s_i - s_j\|\right) & , i \neq j \\ \delta_k^2 + \tau_k^2 & , i = j \end{cases} \quad (4.8)$$

na qual δ_k^2 é a variância marginal dos resíduos da regressão do parâmetro k (ou *partial sill*), ϕ_k é o *range*, τ_k é o *nugget* e $\|s_i - s_j\|$ é a matriz de distância Euclidiana entre os centroides das áreas de drenagem dos postos i e j .

A função de covariância (Eq. 4.8) é isotrópica, isso é, suas propriedades independem da direção e a dependência espacial é modelada exclusivamente a partir da distância Euclidiana. Consequentemente, a função de covariância utilizada não considera a estrutura de dependência da rede de drenagem. Ao não incorporar essa característica física do processo, a predição de dois pontos aninhados pode estimar quantis maiores de vazão no posto a montante do que no posto a jusante, o que não seria fisicamente coerente. Apesar dessa limitação, essa estrutura foi utilizada em estudos de vazão máxima (SAMPAIO; COSTA, 2021; YAN; MORADKHANI, 2014) e de vazão mínima (AHN; PALMER; STEINSCHNEIDER, 2017). Embora deficiente no aspecto levantado,

essa abordagem parece ser eficaz em reduzir a incerteza preditiva e o viés das estimativas.

A estrutura do modelo de regressão, portanto, é definida como:

$$\log(\theta_k[\tilde{s}_i]) = \alpha_{k0} + (\mathbf{x}^k[\tilde{s}_i])^t \boldsymbol{\alpha}_k + \varepsilon_k[\tilde{s}_i], \quad (\varepsilon_k[\tilde{s}_1], \dots, \varepsilon_k[\tilde{s}_N]) \sim MVN(\mathbf{0}, \boldsymbol{\Sigma}_k)$$

4.4.1.3 Modelo M3 e M4

Dados de vazão mínima de postos fluviométricos próximos são prováveis de apresentar correlação cruzada devido à natureza do processo. Os principais *inputs* do processo, regime pluviométrico, pedologia e hidrogeologia são conhecidamente variáveis contínuas no espaço que manifestam autocorrelação espacial para qual técnica de geoestatística (e.g. Krigagem) baseadas em semivariogramas são frequentemente utilizadas.

Em um estudo em larga escala pelo continente europeu, Gudmundsson *et al.* (2011) mostram que a correlação média regional em climas úmidos e temperados é elevada para todos os percentis de vazões – ~0.6 a 0.85 para a Q₉₅. Devido à similaridade do clima das bacias deste estudo aos climas citados, acredita-se que a correlação-cruzada também possa ser significativa.

A dependência espacial das observações é introduzida no nível dos dados por meio de uma Cópula Elíptica. A escolha dessa abordagem é apoiada na praticidade da teoria de cópulas, que não depende da distribuição marginal, e devido à capacidade dessa família de cópula descrever processos de grande dimensão (RENARD, 2011). Essa família de cópula possui duas variantes, a Gaussiana e a *t-Student*. Neste estudo optou-se por utilizar a cópula Gaussiana, assim como Bracken *et al.* (2016), devido a sua simplicidade. A comparação de modelos de cópula foge do escopo deste estudo. De fato, a cópula gaussiana utilizada neste estudo pode não representar adequadamente a dependência dos eventos extremos, em parte, pela ausência de dependências nas caudas dessa família. Entretanto, ressalta-se que, assim como Renard (2011), o objetivo deste estudo não está em estimar a probabilidade conjunta de eventos extremos, mas em estimar as vazões mínimas (em especial o quantil com probabilidade de não excedência de 10%) na bacia sob a condição de existência dependência nos dados. Quando o interesse do estudo está na probabilidade conjunta

de baixas probabilidade de eventos extremos, a família de cópulas cogitadas deve ser cuidadosamente avaliada, e.g. Dodangeh *et al.* (2020).

A distribuição conjunta de observações é modelada por uma cópula elíptica Gaussiana C , tal que o nível dos dados dos modelos M3 e M4 pode ser definido por:

$$(Y[s_1, t], \dots, Y[s_N, t]) \sim C(\boldsymbol{\Sigma}_d, \{\boldsymbol{\theta}[s_1], \dots, \boldsymbol{\theta}[s_N]\})$$

$$Y[s, t] \sim p(\boldsymbol{\theta}[s])$$

A dependência espacial da cópula Gaussiana é caracterizada pela matriz de dependência entre os pares de observações, $\boldsymbol{\Sigma}_d$, chamada de dependograma. A maioria dos estudos (e.g. BRACKEN *et al.*, 2016a; BRUNNER; FURRER; FAVRE, 2019; RENARD, 2011) constroem o dependograma a partir de funções exponenciais ou de misturas de exponenciais, principalmente devido a sua simplicidade.

Neste estudo decidiu-se investigar o uso de outro modelo de dependência espacial, para verificar se um modelo mais simples, como o exponencial, é o suficiente para descrever a dependência espacial dos dados. Apesar de ser uma comparação simples e limitada a 2 modelos, o modelo de dependência dos dados não foi comparado em outros estudos de AFR Bayesiana e, conforme expresso por Renard (2011), ele merece estudos adicionais. Por isso, o modelo M3 foi construído com o modelo exponencial simples (Eq. 4.9) e o modelo M4 com o modelo exponencial potência (Eq. 4.10). Esse modelo introduz maior flexibilidade em relação ao modelo exponencial simples devido à adição de um novo parâmetro.

$$\boldsymbol{\Sigma}_d = \exp\left(-\frac{1}{\eta} \|s_i - s_j\|\right) \quad (4.9)$$

$$\boldsymbol{\Sigma}_d = \exp\left(-\left(\frac{1}{\eta} \|s_i - s_j\|\right)^p\right) \quad (4.10)$$

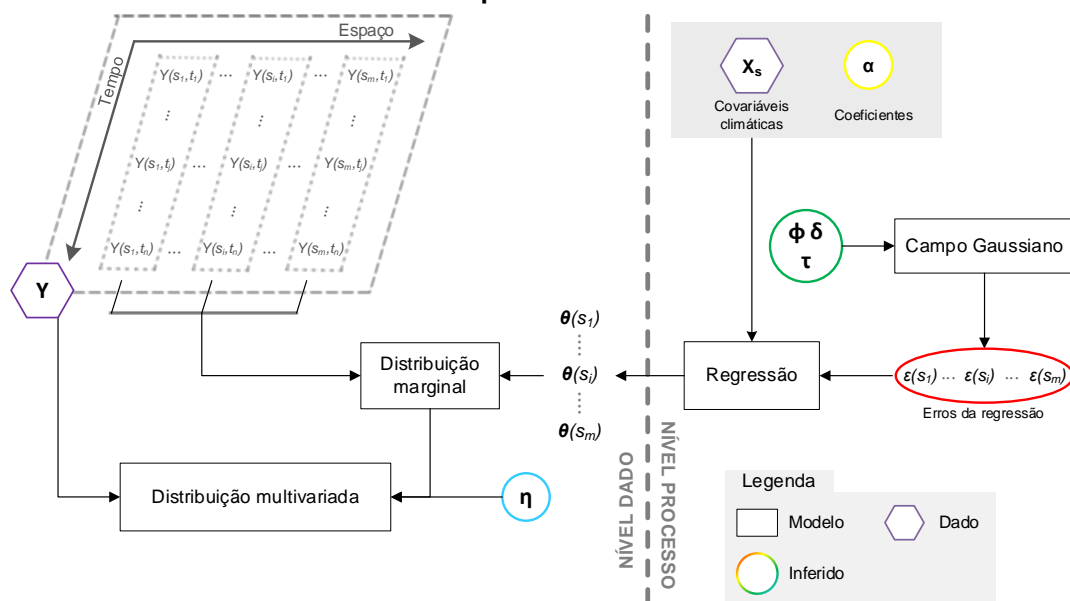
na qual η é o *range* da cópula, $\|s_i - s_j\|$ é a matriz de distância euclidiana entre os centroides das áreas de drenagem dos postos i e j , e p é um parâmetro de forma definido no intervalo $(0,2]$, para o qual há uma família de variogramas válidos

(BANERJEE; CARLIN; GELFAND, 2014). O modelo exponencial e gaussiano são casos particulares do modelo exponencial potência, para $p = 1$ e $p = 2$, respectivamente.

O modelo de dependência dos dados empregado é baseado na distância Euclidiana entre os centroides das áreas de drenagem. Novamente, essa concepção não considera a estrutura da rede de drenagem, por isso, ela pode ser inadequada para discernir os diferentes graus de dependência entre bacias aninhadas e não-aninhadas. Essa limitação do modelo é reconhecida e, de fato, uma melhor descrição espacial que considere a estrutura de dependência imposta pela rede de drenagem poderia ser mais adequada para quantificar o efeito da hipótese de dependência dos dados. Apesar disso, por efeitos de simplicidade e sob a ciência de sua limitação, prossegue-se com o uso dessa abordagem.

Os modelo M3 e M4 são esquematizados na Figura 4.3.

Figura 4.3 - Esquema do modelo M3/M4 onde são apresentados os níveis dos dados e do processo.



Fonte: adaptado de Renard (2011)

A presença de falhas torna o uso da cópula complexo ou desafiador em termos computacionais. Por esse motivo, Renard (2011) optou pela escolha de postos pluviométricos de forma que a janela temporal de dados de calibração não houvesse falhas. Essa abordagem, como o próprio autor afirma, não é uma solução prática aceitável, uma vez que a exclusão de postos pode reduzir a acurácia preditiva do

modelo. Bracken *et al.* (2016), por outro lado, optaram por dividir os dados em dois conjuntos – dados completos e dados com falhas, de forma que a verossimilhança da cópula foi computada apenas com os dados completos. Embora esses autores afirmem que isso não afetou fortemente as estimativas do parâmetro da cópula, por questões de clareza nos resultados, optou-se por utilizar os dados na íntegra neste estudo. Esse procedimento exigiu que a covariância da cópula fosse calculada anualmente com dimensões variáveis. Para o número de postos avaliados, menor que 50 em ambas bacias, essa abordagem aumentou o esforço computacional, mas a duração da inferência do modelo se manteve em valores razoáveis.

4.4.2 Modelo não estacionário

A observação da Q_7 mínima anual pode estar condicionada a covariáveis temporais, que neste trabalho, são abordadas por meio da temperatura da superfície do mar (TSM). Optou-se por essa abordagem ao invés de impor um modelo de tendência temporal (i.e., tempo como covariável) pois julga-se que não há fundamento para supor que possíveis tendências retratadas pelos dados irão persistir no futuro e pois amostras curtas podem manifestar tendências que na verdade compõe a própria variabilidade amostral do processo. Posto isso, a não estacionariedade deste trabalho deve ser vista exclusivamente como um condicionamento a uma covariável medida no tempo. Em outras palavras, não há suposição de tendências temporais ou previsão para cenários futuros de mudanças climáticas.

Na concepção proposta, a TSM explica parte da variabilidade temporal das observações de vazões mínima. Conceitualmente, essa abordagem atribui a não estacionariedade do processo a uma única classe de variável (ainda que a TSM de múltiplas localidades sejam utilizadas como covariáveis). No entanto, acredita-se que outros agentes também expliquem parte da variabilidade anual das realizações de vazões mínimas nas bacias em estudo. Na bacia do rio Doce, por exemplo, parte da precipitação anual se deve a ZCAS, cuja intensidade e posição variam anualmente. Logo caso houvesse um índice relacionado a esse fenômeno (semelhante a Steinschneider e Lall (2015)), ele seria um potencial candidato a covariável temporal. Na medida em que mais dados e mais conhecimento sobre a interação de variáveis hidrológicas e aspectos da atmosfera e da bacia se tornarem disponíveis, outras

covariáveis poderiam ser consideradas. Para este trabalho, focou-se apenas em na TSM, a qual tem o suporte da literatura nas duas bacias de estudo.

Na bacia do rio Doce, Rodrigues *et al.* (2019) identificaram um mecanismo durante o verão austral comum em muitas secas no sudeste do Brasil. Esse mecanismo está associado ao aumento da temperatura no Oceano Atlântico do sul, próximo da costa brasileira. A busca de covariáveis de TSM para essa bacia foi limitada ao Oceano Atlântico do sul. Na bacia do rio Itajá-Açu, diversas referências relatam a influência do ENSO no regime pluviométrico da região onde a bacia está inserida. Nessa bacia, a busca de covariáveis de TSM foi limitada ao Oceano Pacífico.

A TSM é utilizada no cálculo de diversos índices climáticos, como o Niño 3.4, que visam quantificar fenômenos atmosféricos-oceânicos de larga escala. A vantagem de se utilizar diretamente a TSM ao invés de índices climáticos reside na possibilidade de identificar sinais mais fortes de teleconexão. Índices climáticos podem não apresentar sinais fortes com a hidrologia de uma determinada região, uma vez que eles são desenvolvidos para maximizar a variância explicada do clima em macro escala (AHN; PALMER; STEINSCHNEIDER, 2017). Ademais, a utilização de índices climáticos impossibilita utilizar a função de ligação log-log, uma vez que os índices são normalizados (i.e. há valores negativos).

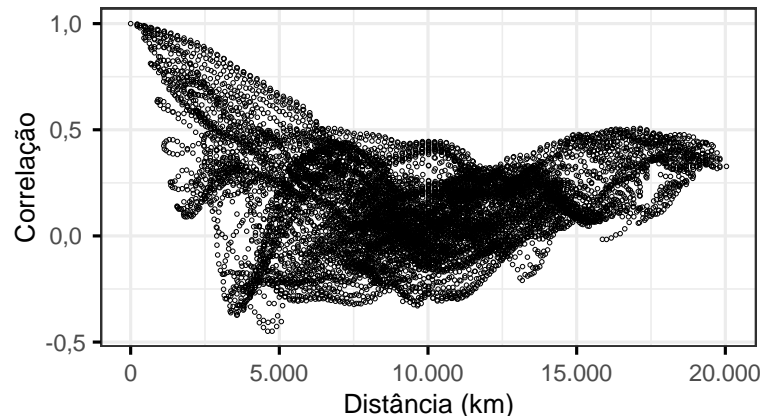
4.4.2.1 Avaliação de teleconexões

Dados de TSM estão disponíveis em escala global com resolução espacial de $2^{\circ} \times 2^{\circ}$ e no período de 1854 a 2020, em escala mensal (HUANG *et al.*, 2017). Esses dados podem ser manipulados em ambiente R por meio do pacote *ncdf4* (PIERCE, 2019) do R. Esse enorme repertório torna a escolha da covariável climática mais complicada por dois motivos. Primeiro, é computacionalmente demandante construir um modelo com milhares de covariáveis temporais. Segundo, o enorme conjunto de dados de TSM (9.646 séries) torna mais provável a aparição de correlações espúrias entre os dados de TSM e as vazões mínimas da bacia de estudo.

Para contornar o primeiro desafio descrito acima, seguindo Renard e Lall (2014), Zeng *et al.* (2017) e Ahn *et al.* (2017), diminuiu-se a resolução da informação original dos dados de TSM. As resoluções adotadas nos estudos citados foram, respectivamente,

10° x 10°, 7° x 22° e 12° x 12°. Os dados disponíveis com resolução de 2° x 2° foi transformado para uma escala de 10° x 10° (~1110 km x 1110 km na linha do Equador) a partir de uma média aritmética. A redução na escala dos dados não compromete substancialmente o sinal dos dados originais, uma vez que TSM exhibe forte correlação espacial (Figura 4.4).

Figura 4.4 - Correlação da TSM do mês de janeiro entre na coordenada 0°, 120°O e as 9.646 coordenadas do banco de dados

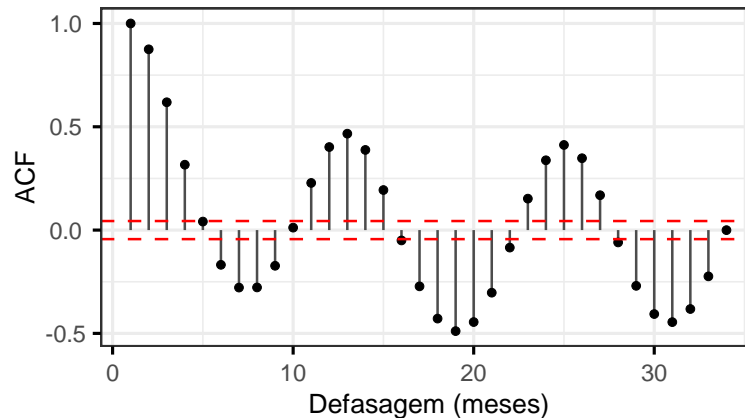


O segundo desafio levantado refere-se à possibilidade de existência de correlações espúrias entre dados de TSM e as vazões mínimas. O modelo não estacionário proposto é concebido sob a perspectiva de que a imputação de covariáveis temporais ao processo estocástico deva ser embasada (raciocínio dedutivo), ao invés de um modelo exclusivamente conduzido pelos dados. Essa concepção é importante para manter umnexo entre a estatística e o processo físico. Frente ao princípio exposto, a área de abrangência da TSM em estudo foi restrita a:

- Oceano Atlântico do sul, no polígono delimitado pelas longitudes 60°O e 0° e as latitudes 10°S e 50°S, para a inferência do modelo regional da bacia do rio Doce, baseando-se no mecanismo descrito por Rodrigues *et al.* (2019); e
- Oceano Pacífico, no polígono delimitado pelas longitudes 160°E e 80°O e as latitudes 5°N e 5° S. O polígono proposto abrange as áreas utilizadas no cálculo dos índices climáticos Niño 1+2, 3, 4, 3.4 e Índice Oceano Niño (ONI, do acrônimo em inglês *Oceanic Niño Index*) comumente utilizados para representar o ENSO. A área do cálculo do ONI é coincidente com a área do Niño 3.4.

Os dados de TSM do banco de dados utilizados são disponibilizados em escala mensal. Neste estudo, optou-se por realizar uma análise exploratória desses dados a fim de escolher o mês com sinal mais forte com as séries de vazões mínimas. É provável que mais de um mês da TSM possa responder pela variabilidade das séries de vazão mínima, no entanto, a fim de simplificar a abordagem proposta, selecionou-se apenas o mês com sinal mais forte. Apesar de utilizar um único mês, a persistência observada nos dados de TSM, exemplificada pela Figura 4.5, sugere que eventuais sinais identificados com o mês escolhido também existam, em parte, nos meses adjacentes.

Figura 4.5 - ACF da série mensal de TSM na coordenada 0°, 120°O. Linhas tracejadas vermelhas indicam o intervalo com 95% de confiança



Para auxiliar na escolha do mês de TSM a ser utilizado no modelo, recorreu-se a análise de componentes principais (ACP) para facilitar a identificação de sinais entre a TSM e as séries de vazões mínimas. Dessa maneira, as múltiplas séries de vazões mínimas de cada bacia são reduzidas às componentes principais. Em seguida, é computada a correlação de Spearman entre as componentes principais das séries de vazões mínimas e as séries mensais de TSM. A correlação de Spearman foi utilizada pois ela não necessita que as variáveis sejam normalmente distribuídas (dados mensais de TSM não são normalmente distribuídos).

A janela temporal selecionada para a ACP na bacia do rio Doce foi 1974-2013 e na bacia do rio Itajaí-Açu foi 1979/80-2018/19, de forma que as séries de vazão possuíssem 40 observações. Idealmente séries longas são ideias para capturar a relação entre as variáveis, mas um comprometimento entre número de estações e tamanho das séries foi necessário devido ao *trade-off* entre esses elementos. Na bacia

do rio Doce (Itajaí-Açu), 9 (5) postos fluviométricos tinham dados disponíveis para a janela temporal adotada.

4.4.2.2 Modelo

Após a identificação do mês com sinal mais forte entre a TSM e as vazões mínimas dentro das regiões do oceano com indícios de correlação com o regime hidrológico, segue-se para a introdução dessa covariável no modelo não estacionário.

A dependência temporal é introduzida no modelo no nível do processo, que é visto como a soma de duas componentes, a estacionária (η_e) e a não estacionária (η_n):

$$\log(\theta_k[\tilde{s}_i]) = \eta_{e,k}[\tilde{s}_i] + \gamma_k[\tilde{s}_i]\eta_n$$

na qual γ representa a componente espacial do índice customizado η_n . η_e depende do modelo estacionário utilizado como referência para construir o modelo não estacionário.

O índice climático customizado é derivado de uma ponderação dos dados de TSM, cujos pesos são inferidos juntos com o modelo para tentar capturar o sinal mais forte com os postos fluviométricos em estudo de cada bacia. O cálculo do índice climático é obtido pela Eq. 4.11.

$$\eta_n = \sum_p C_{t,p} \times \xi_p \quad (4.11)$$

na qual $C_{t,p}$ é o preditor climático, no ano t e local p , e ξ são os pesos atribuídos à malha de preditores climáticos.

Os preditores, $C_{t,p}$, são centralizados subtraindo-se as médias de cada p local para facilitar a convergência do algoritmo e interpretação da variável preditora, tal que em um ano típico a componente não estacionária seja nula. Assim como na componente estacionária, utiliza-se a função de ligação logarítmica entre a variável dependente e o preditor para manter a consistência na regressão proposta, tal que $C_{t,p} = \log(TSM_{t,p}) - \overline{\log(TSM_p)}$.

Seguindo Ahn *et al.* (2017), o índice climático customizado é permitido variar espacialmente nas bacias de estudo, ao invés de se assumir que essa quantidade é constante em toda a bacia. Esse expediente permite flexibilizar a resposta da bacia ao índice climático assinalando regiões com sinal mais/menos forte com o índice climático. Neste estudo, similar a Ahn *et al.* (2017), o campo espacial de resposta do índice climático foi modelado a partir de uma função de decaimento exponencial. Assume-se que o sinal do índice é máximo em um par de coordenadas inferidas na bacia (x_0, y_0) e que ele decai espacialmente sob a forma de uma função exponencial:

$$\gamma_k[\tilde{s}_i] = \delta_{t,k}^2 \exp(-\phi_{t,k} \|\tilde{s}_i - s_0\|)$$

na qual $\delta_{t,k}^2$ é o *sill* e ϕ_t é o decaimento ($\frac{1}{\phi_{t,k}}$ é chamado de *range*) da função exponencial.

As coordenadas (x_0, y_0) são quantidades a serem inferidas pelo modelo e não são conhecidas *a priori*.

4.4.3 Distribuição a priori

No nível das distribuições *a priori*, foram consideradas distribuições pouco informativas para a maioria dos parâmetros, principalmente com base em Bracken *et al.* (2016), Lima *et al.* (2016) e Sampaio e Costa (2021). Para os parâmetros ϕ_k , no entanto, Banerjee *et al.* (2014) e Davison *et al.* (2012) recomendam que sejam eliciadas distribuições mais informativas para o *range* do semivariograma por problemas de identifiabilidade. Por isso, para esses parâmetros a distribuição Gama com parâmetro de forma e taxa iguais a 5 e 0,1, respectivamente, foi eliciada de forma a penalizar distâncias pequenas onde não se há informação sobre o processo e restringir o intervalo de valores possíveis, inspirados nos estudos de Najafi e Moradkhani (2013) e Yan e Moradkhani (2014). Assim como Ahn *et al.* (2017), utilizou-se distribuições *a priori* informativas para pesos dos índices climáticos para manter o significado desse parâmetro. A Tabela 4.6 mostra as distribuições *a priori* empregadas para os hiper-parâmetros dos modelos desenvolvidos.

Tabela 4.6 - Distribuições a priori dos hiper-parâmetros dos MHBs desenvolvidos

Hiper parâmetro	Distribuição <i>a priori</i>
Intercepto da regressão dos parâmetros de $p(\cdot)$	$\alpha_{k0} \sim N(0, 10)$
Coefficiente angular da regressão dos parâmetros de $p(\cdot)$	$\alpha_{k1} \sim N(1, 1)$
Variância da regressão dos parâmetros de $p(\cdot)$	$\delta_k^2 \sim N(0, 1)$
<i>Nugget</i> do semivariograma do processo gaussiano	$\tau_k^2 \sim N(0, 1)$
<i>Range</i> do semivariograma do processo gaussiano	$\phi_k \sim Ga(5, 0,1)$
<i>Range</i> da copula	$\eta \sim Ga(0,01, 0,01)$
Parâmetro de forma da função exponencial potência da estrutura de dependência espacial da copula	$p \sim U(0, 2)$
Pesos do campo de valor de TSM	$\xi_p \sim U(-1, 1)$
<i>Sill</i> da função exponencial da estrutura de dependência espacial do índice climático customizado	$\delta_{t,k}^2 \sim N(0, 100)$
Decaimento da função exponencial da estrutura de dependência espacial do índice climático customizado	$\phi_{t,k} \sim U(0, 1)$

Nota: $N(a, b)$ denota uma distribuição normal com média a e desvio padrão b ; $Ga(a, b)$ denota uma distribuição Gama com parâmetro de forma a e parâmetro de taxa b ; $U(a, b)$ denota uma distribuição Uniforme entre a e b .

4.4.4 Inferência

A distribuição *a posteriori* dos modelos propostos é de difícil solução analítica e, portanto, recorreu-se à exploração da referida distribuição por meio do amostrador No-U-Turn (NUTS), uma variação do algoritmo HMC MCMC (HOFFMAN; GELMAN, 2014). O algoritmo foi implementado no *software* livre Stan, acessado pelo pacote

*RStan*¹⁵ (STAN DEVELOPMENT TEAM, 2020) do R. A distribuição *a posteriori* foi estimada por meio quatro cadeias com 2.500 iterações, cada qual com *warmup*¹⁶ de 1.250 iterações, totalizando 5.000 realizações para compor a distribuição *a posteriori*. A convergência do modelo foi avaliada por meio da estatística \hat{R} e do número efetivo do tamanho das amostras (*ESS*, do inglês *Effective Sample Size*), como recomendado por Vehtari *et al.* (2019). Seguindo as recomendações dos autores, a convergência foi verificada para o critério $\hat{R} \leq 1,01$ e $ESS \geq 100$ por cadeia. Foi utilizado nesta pesquisa um computador pessoal com processador Intel i7-8565U de 1.80 GHz e uma memória RAM de 8 Gb.

Alguns postos fluviométricos não foram utilizados na inferência dos modelos (i.e. calibração) para ser utilizados como postos de validação. O desempenho dos modelos e a influência de suas hipóteses nos resultados serão exploradas por meio da predição nesses postos. O detalhamento dos postos de validação e calibração pode ser consultado nas Tabela B.1 e Tabela B.2 do Apêndice B.

4.4.5 Comparação entre modelos

Nos modelos apresentados acima, as covariáveis físicas ou climáticas foram genericamente apresentadas. A escolha das covariáveis é discutida a seguir.

Primeiramente, selecionou-se para análise apenas as covariáveis com correlação significativa (*valor-p* < 0,05) com os parâmetros da distribuição de probabilidade selecionada, etapa descrita na seção 0. As covariáveis candidatas foram utilizadas para elaborar MHBs sob a forma do modelo M1. Em seguida, a seleção das covariáveis é realizada por meio da comparação dos modelos gerados.

A comparação de modelos é usualmente realizada por métodos de validação cruzada ou critérios de informação. Dentre os critérios de informação usados no contexto Bayesiano, o *Deviance Information Criterion* (DIC) (SPIEGELHALTER *et al.*, 2002) é bastante popular. O DIC é amplamente usado no contexto Bayesiano (e.g., SUN *et al.*, 2014; YAN; MORADKHANI, 2015; ZENG *et al.*, 2017) e pode ser visto como uma

¹⁵ Versão 2.21.2

¹⁶ As amostras do *Warmup* são utilizadas para adaptar os parâmetros de amostragem, não sendo parte da distribuição *a posteriori* visada.

versão Bayesiana do clássico *Akaike Information Criterion* (AIC). O DIC é representado pela soma da *deviance* média (\bar{D}), que representa a acurácia preditiva do modelo, e do número efetivo de parâmetros (p_D), fator penalizante de modelos mais complexos. Menores valores de DIC são preferidos. A formulação do DIC é apresentada nas equações a seguir.

$$D(\boldsymbol{\theta}) = -2 \log(f(\mathbf{y}|\boldsymbol{\theta}))$$

$$\bar{D} = E[D]$$

$$p_D = \bar{D} - D(\mathbf{y}|\bar{\boldsymbol{\theta}})$$

$$DIC = \bar{D} + p_D$$

O DIC, apesar de ser implementado com uma medida de penalização para modelos mais complexos (p_D), ainda é passível a indicar modelos sobreajustados (SPIEGELHALTER *et al.*, 2014). Outros métodos de comparação entre modelos foram cogitados, como o WAIC (WATANABE, 2010), PSIS-LOO (VEHTARI; GELMAN; GABRY, 2017) e validação cruzada *k-fold*. Os dois primeiros métodos, implementados no pacote *loo* (VEHTARI *et al.*, 2020) do R, indicaram problemas no diagnóstico de suas funções, por isso não foram avaliados. A validação cruzada *k-fold*, embora computacionalmente menos exigente do que a validação *leave-one-out*, não foi viável devido ao grande tempo de processamento. No final, assim como muitos estudos recentes, utilizou-se o DIC para comparar os MHBs.

Após a seleção das covariáveis por meio do DIC, foram elaborados os modelos M1 a M4. Esses modelos foram avaliados em termos de suas premissas para investigar aspectos de dependência espacial dos dados e do processo. No final, eles foram comparados entre si por meio do DIC e de suas predições nos postos de validação para a seleção de um modelo estacionário de referência, a partir do qual foi desenvolvido o modelo não estacionário.

O modelo não estacionário apresentado na seção 4.4.2 tem uma componente que varia temporalmente, representada pelo índice climático customizado, que é introduzida nos parâmetros $\boldsymbol{\theta}$ da distribuição de probabilidade dos dados. Permitir que

todos os parâmetros do vetor θ variem no tempo pode levar a um modelo mais complexo no qual o seu desempenho pode não ser necessariamente o melhor. O DIC foi empregado para avaliar, dentre os arranjos possíveis com modelos estacionários e não estacionários, aquele com melhor desempenho.

4.4.6 Estimação e predição

Em um posto monitorado \tilde{s} , a estimativa dos parâmetros da distribuição de probabilidades pode ser diretamente aproximada a partir das amostras do MCMC dos hiper-parâmetros (e.g., α_{k0} , α_{k1} , δ_k) e dos resíduos inferidos neste posto ($\varepsilon_k[\tilde{s}]$).

A predição em um local não monitorado \hat{s} , sob o *framework* do modelo M1, exige uma etapa adicional em relação às estimativas em postos de calibração. Em \hat{s} , o modelo não foi ajustado, portanto, não há resíduos das regressões dos K parâmetros da distribuição $p(\cdot)$. Uma vez que os resíduos são espacialmente independentes, eles podem ser amostrados, tal que $\varepsilon_k[\hat{s}] \sim N(0, \delta_k)$.

Sob os modelos M2 a M4, os resíduos manifestam dependência espacial, e portanto, a localização de \hat{s} é um parâmetro de entrada para a predição. Seguindo Renard (2011), e denominando o vetor de parâmetros do variograma de \mathbf{v} , os resíduos das regressões em \hat{s} , condicionados às contrapartes dos postos \tilde{s} (como também a $\tilde{\mathbf{y}}, \mathbf{x}$, e \mathbf{v}), podem ser expressos como:

$$p(\varepsilon_k[\hat{s}] | \varepsilon_k[\tilde{s}_1], \dots, \varepsilon_k[\tilde{s}_N], \mathbf{v}_k, \tilde{\mathbf{y}}, \mathbf{x}) = N(\mu_{k,conditional}, \sigma_{k,conditional}) \quad (4.12)$$

Para estimar os parâmetros $\mu_{k,conditional}$ e $\sigma_{k,conditional}$, é necessário: (i) a matriz $N \times N$ de variância-covariância dos erros da regressão, $\mathbf{\Omega}_k$, estimada a partir dos dados de calibração do modelo, (ii) o vetor $1 \times N$ de covariância entre os resíduos do local e o dos postos, $\mathbf{\Lambda}_k$, e (iii) a variância marginal dos resíduos da regressão σ_k^2 . Essas quantidades são combinadas em uma nova matriz de variância-covariância, $\mathbf{\Gamma}_k$ (RENARD, 2011):

$$\mathbf{\Gamma}_k = \begin{pmatrix} \sigma_k^2 & \mathbf{\Lambda}_k \\ \mathbf{\Lambda}_k^t & \mathbf{\Omega}_k \end{pmatrix}$$

Os parâmetros da distribuição Gaussiana da Eq. 4.12 podem ser calculados por (RENARD, 2011):

$$\mu_{k,conditional} = \mathbf{\Lambda}_k \mathbf{\Omega}_k^{-1} (\varepsilon_k[\tilde{s}_1], \dots, \varepsilon_k[\tilde{s}_N])^t \quad (4.13)$$

$$\sigma_{k,conditional}^2 = \sigma_k^2 - \mathbf{\Lambda}_k \mathbf{\Omega}_k^{-1} \mathbf{\Lambda}_k^t \quad (4.14)$$

Uma vez determinadas essas quantidades, os resíduos em \hat{s} são amostrados, tal que $\varepsilon_k[\hat{s}] \sim N(\mu_{k,conditional}, \sigma_{k,conditional})$.

Após a amostragem dos resíduos, o restante do procedimento é similar ao dos postos de calibração. É válido notar que os parâmetros da cópula não são utilizados para a predição.

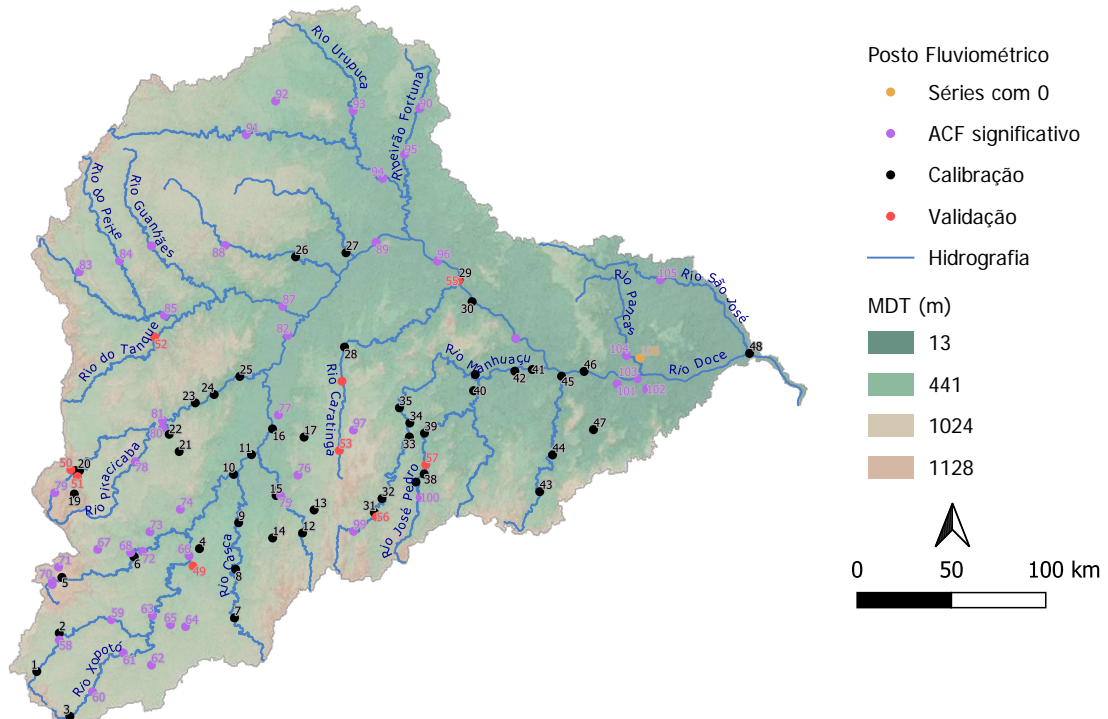
5 RESULTADOS E DISCUSSÃO

5.1 Avaliação dos dados

Os dados das estações fluviométricas das bacias do rio Doce e Itajaí-açu são descritos a seguir. Para mais detalhes como código, nome, tamanho da série de dados, identificado (ID) adotado para cada posto, entre outros, consultar as tabelas do APÊNDICE B – .

Ao todo, 106 estações fluviométricas foram consideradas no estudo da bacia do rio Doce. A disposição espacial dessas estações é ilustrada na Figura 5.1 e a disponibilidade temporal de dados na Figura 5.2.

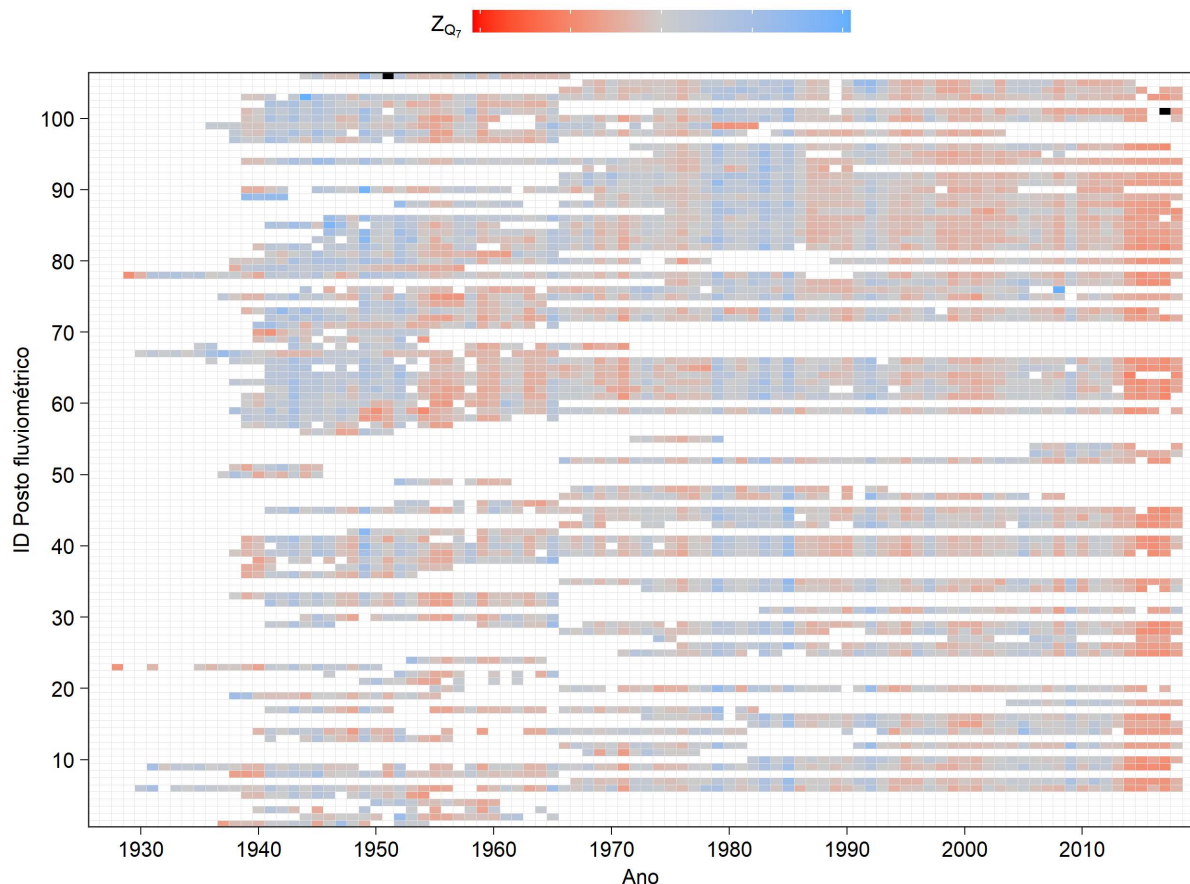
Figura 5.1 - Localização dos postos fluviométricos da bacia do rio Doce. Os rótulos dos postos fluviométricos correspondem ao ID de cada posto



A função de autocorrelação (ACF) indicou uma dependência serial muito forte em uma grande parcela dos postos da bacia do rio Doce (Figura 5.1). A estatística da ACF nessa bacia foi influenciada pela excepcional seca dos últimos anos (ver Figura 5.2). Os dados de 2014 a 2018 foram excluídos das séries de Q₇ mínima anual da bacia do rio Doce por causa da excepcionalidade desse evento que alterou significativamente as estatísticas da série, como a memória do processo. Admite-se que outra abordagem é necessária para analisar adequadamente eventos extremos de vazões

mínimas plurianuais, contudo, esse estudo visa avaliar o modelo proposto em um cenário genérico com foco em quantis pequenos associados à vazões de outorga, e por isso, prosseguiu-se com a exclusão desse período dos dados da bacia do rio Doce. Após a exclusão, a dependência serial nas séries de vazões mínimas caiu substancialmente, de 37 postos sem correlação serial de *lag* 1 significativa, obteve-se 57 após esse procedimento. É possível notar que a maioria dos postos de calibração ficou concentrada no eixo sul da bacia. A retirada dos dados de 2014 a 2018 dos postos da bacia do rio Doce influenciou na inferência do padrão de teleconexão com a covariável temporal estudada, uma vez que, o período deletado está associado à covariável temporal adotada (COELHO *et al.*, 2016).

Figura 5.2 - Disponibilidade de dados da bacia do rio Doce. Anos em preto indicam o registro de vazão nula. A escala de cores das séries representa valores normalizados da Q7 mínima anual ($Z_x = \frac{x-\mu}{\sigma}$)



A estação 56995000-Cachoeira do Oito, localizada no Espírito Santo, não foi utilizada neste estudo devido a presença de vazões nulas. Apesar de não ser um impedimento à realização da análise de frequência de vazões mínimas, e.g. Stedinger, Vogel e Foufoula-Georgiou (1993), a presença de valores nulos impõe maior complexidade ao

tema e, portanto, decidiu-se remover essas estações. A estação 56993551-Jusante Córrego da Piaba também registrou vazão nula, mas essa observação foi excluída por estar no período de 2014-2018. Essas duas estações com registros de vazões nulas possuem áreas de drenagem de, respectivamente, 1160 e 898 km². Trata-se, portanto, de duas bacias de tamanho médio. Acredita-se que as vazões nulas são imputadas a períodos com menores alturas pluviométricas, associados à baixa produção do aquífero fissural dominante nas duas bacias e à demanda hídrica para irrigação.

Na bacia do rio Itajaí-Açu, 20 estações fluviométricas foram consideradas nesse estudo. A localização dessas estações e a disponibilidade de dados são ilustradas na Figura 5.3 e Figura 5.4, respectivamente. Nessa bacia, a ACF indicou correlação serial de *lag* 1 significativa em 5 dos 20 postos fluviométricos analisados.

Figura 5.3 - Localização dos postos fluviométricos da bacia do rio Itajaí-Açu. Os rótulos dos postos fluviométricos correspondem ao ID de cada posto

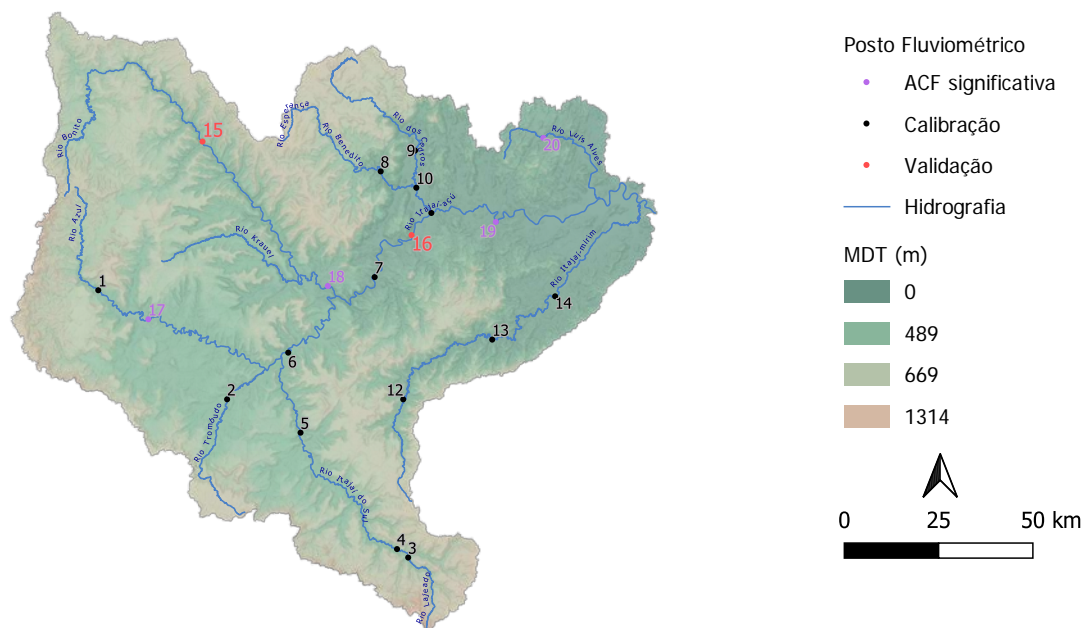
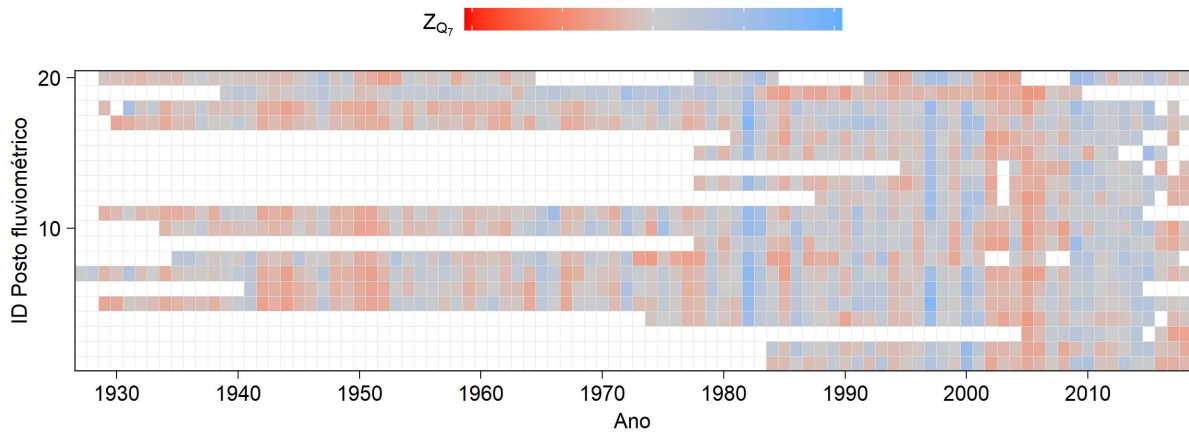


Figura 5.4 - Disponibilidade de dados na bacia do rio Itajaí-Açu. . A escala de cores das séries representa valores normalizados da Q7 mínima anual ($Z_x = \frac{x-\mu}{\sigma}$)



A porcentagem de estações com dependência serial no rio Doce foi superior em relação à bacia do rio Itajaí-Açu. Em parte, isso pode estar associado às diferenças no uso e cobertura do solo entre as bacias. A atividade agropecuária na bacia do rio Doce é mais intensa e, além disso, nessa bacia há muitos pequenos empreendimentos de aproveitamento energético que, em conjunto, podem estar influenciando esse comportamento. Essas observações vão de acordo com Dudley *et al.* (2020), que em um estudo com milhares de estações fluviométricas nos Estados Unidos, observaram que séries de vazões mínimas de bacias agrícolas e regularizadas apresentavam maior porcentagem de dependência serial do que bacias de referência (bacias com influência antrópica mínima).

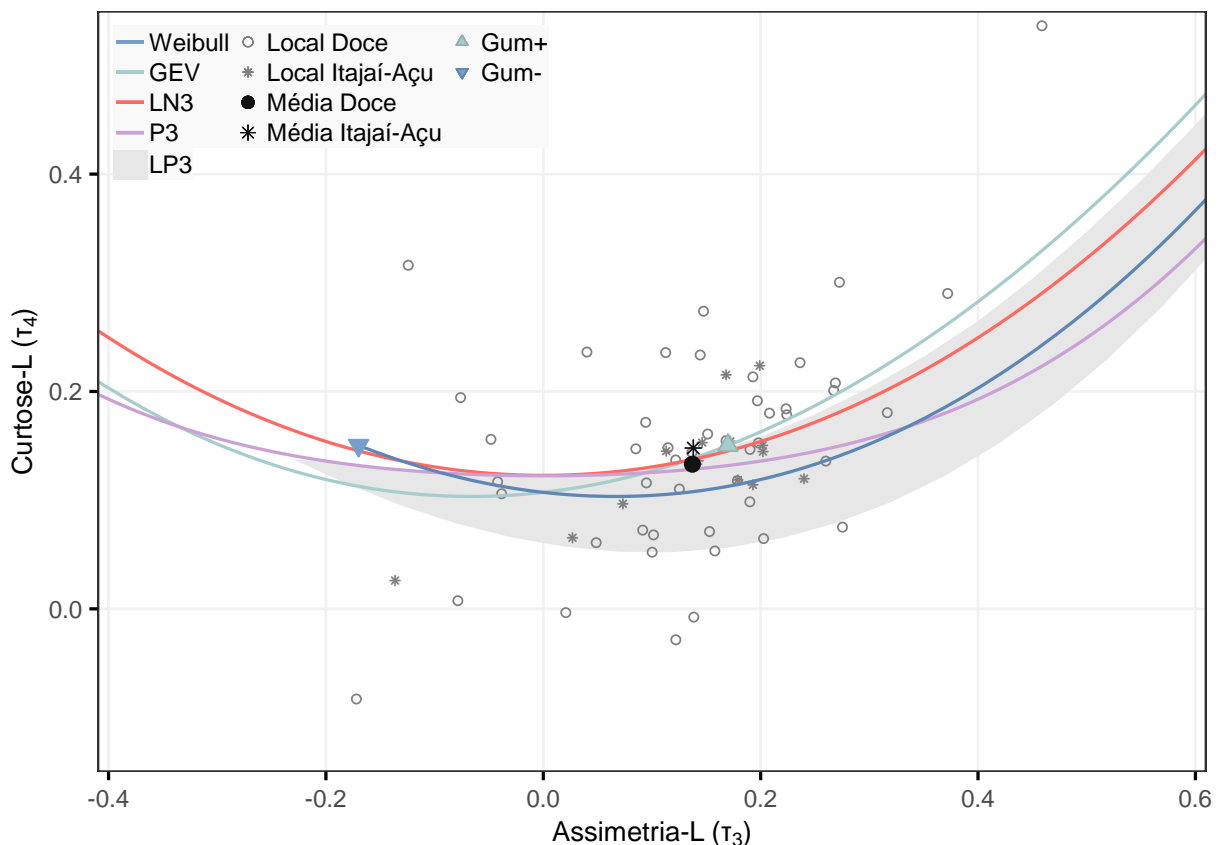
A Figura 5.2 e Figura 5.4, além de mostrar a disponibilidade temporal dos dados, expõe uma característica das séries em estudo. A manifestação de anos secos (em vermelho) ou anos úmidos (azul) tende a ocorrer em conjunto nos postos das bacias, evidenciando que a magnitude das vazões mínimas está relacionada a um processo espacial. A questão da correlação cruzada é abordada posteriormente na análise dos modelos Bayesianos com maior detalhe, mas, a princípio, sua existência é sugerida pelas Figuras Figura 5.2 e Figura 5.4.

5.2 Distribuição de probabilidade candidatas

A análise das distribuições de probabilidade candidatas é averiguada pelo diagrama dos quocientes de momentos-L dos postos fluviométricos (Figura 5.5). O diagrama mostra que a maioria dos postos tem assimetria positiva (54 dos 62 postos). Como consequência, a distribuição de Gumbel com cauda inferior exponencial (GUM-) se

mostrou uma má candidata para a maioria dos postos. As médias dos quocientes de momentos-L das bacias do rio Doce e Itajaí-Açu se aproximaram da distribuição Gumbel com cauda superior exponencial (GUM+), como também das distribuições GEV, LN3 (LN2) e P3 (Ga) e LP3. Os quocientes de momentos-L individuais de cada posto fluviométrico se dispersaram no domínio de valores da Figura 5.5. A dispersão dos pontos pode ser atribuída à variabilidade amostral e à heterogeneidade.

Figura 5.5 - Diagrama dos quocientes de momentos-L dos postos fluviométricos de calibração. O polígono que representa a distribuição LP3 abrange os possíveis pares ordenados da distribuição com o módulo da assimetria inferior a 1,414 (GRIFFIS; STEDINGER, 2007).

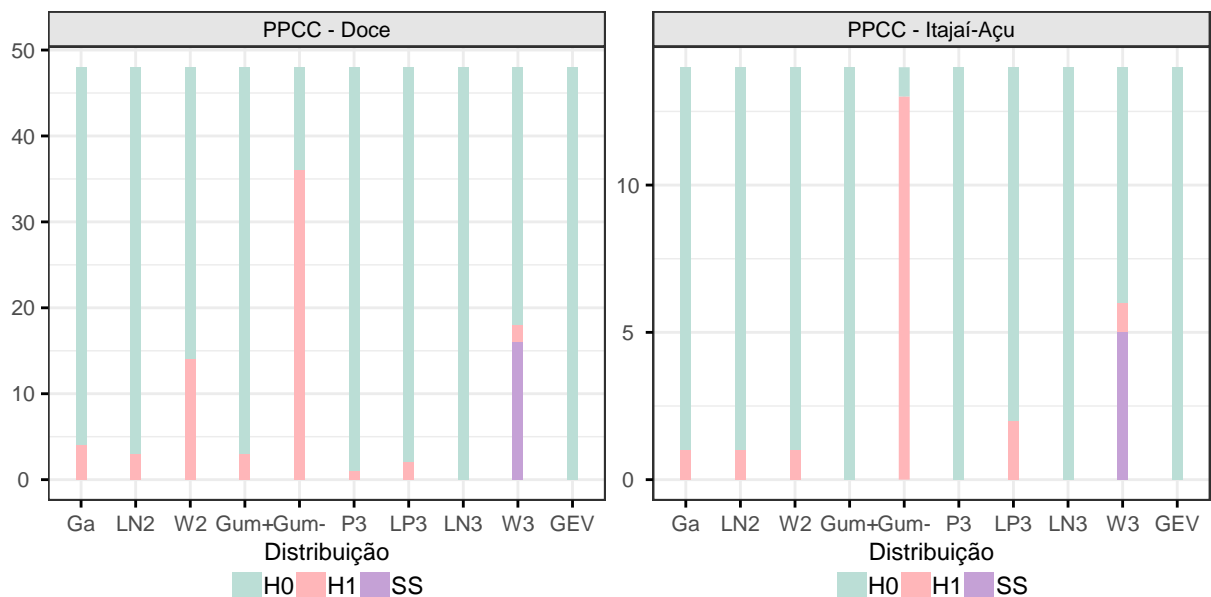


Em seguida, realizou-se o cômputo dos parâmetros das distribuições candidatas pelo método dos momentos-L. O estimador utilizado falhou em estimar o parâmetro que corresponde ao limite inferior da distribuição W3 em vários postos, cujas estimativas desses parâmetros foram superiores às observações. Nesses casos, o posto não foi incluído nas medidas de aderência e, no teste PPCC, ele é marcado como SS – Sem Solução.

O resultado do teste PPCC é apresentado na Figura 5.6. Para o nível de 5% de significância, a distribuição Gum- é atribuída à hipótese alternativa (H1) na maioria dos postos das duas bacias em estudo. Dentre as demais distribuições de dois

parâmetros, a distribuição GUM+ teve menor número de rejeição da hipótese nula do que as demais, e foi seguida pelas distribuições Ga e LN2. Dentre as distribuições de 3 parâmetros, as distribuições LN3, GEV e P3 tiveram os melhores desempenhos. As distribuições LP3 e W3 foram mais frequentemente assinaladas como H1 ou SS. De maneira geral, as distribuições de 3P com melhor desempenho tiveram um número similar de não rejeição da hipótese nula do que as distribuições de 2P com melhor desempenho.

Figura 5.6 - Resultado do teste de aderência PPCC com 5% de significância. H0 e H1 indicam a hipótese nula e a hipótese alternativa do teste. SS é uma abreviação para “Sem Solução”, utilizada para quando o método dos momentos-L falhou em gerar estimativas dos parâmetros da distribuição



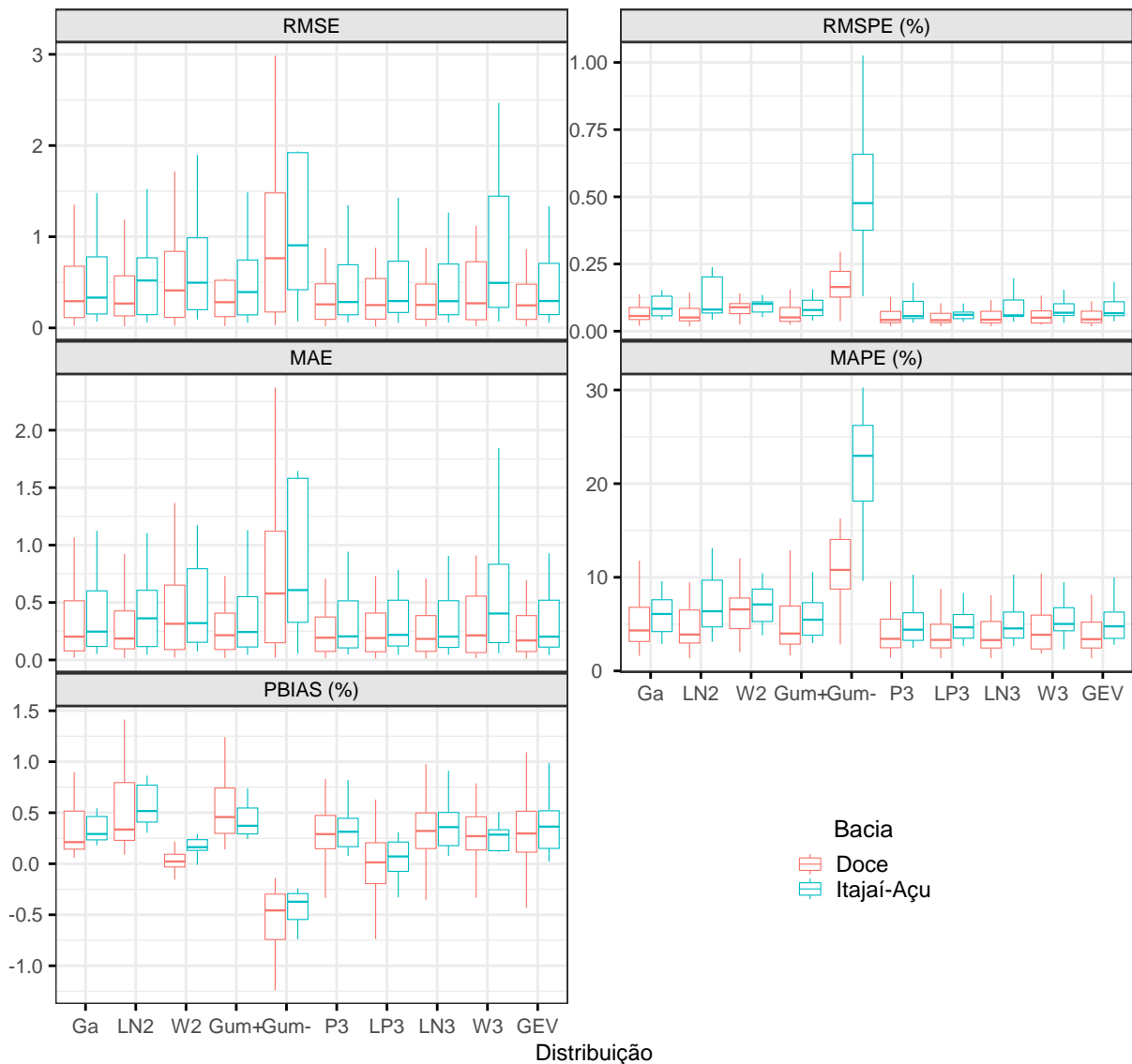
Por fim, as métricas de aderência das distribuições candidatas aos postos fluviométricos é apresentada na Figura 5.7. As distribuições com três parâmetros, em virtude da maior flexibilidade em decorrência do parâmetro adicional, obtiveram melhores métricas em relação às candidatas de 2 parâmetros, embora o desempenho não seja consideravelmente melhor. A escolha de uma das distribuições de dois parâmetros parece ser mais vantajosa pelos seguintes motivos:

- As métricas de ajuste, assim como o teste de aderência, indicam que a qualidade do ajuste das distribuições de 3 parâmetros são similares às de 2 parâmetros;
- Uma distribuição de 3 parâmetros introduziria maior complexidade ao modelo de regionalização;

- Embora um terceiro parâmetro introduza flexibilidade no ajuste para descrever quantis extremos (e.g. TR \geq 100 anos), esses não são o alvo desse estudo. O quantil de interesse está associado a probabilidade de 10% de não superação, relacionado ao critério de outorga da $Q_{7,10}$.

A atenção é então voltada para as distribuições de 2 parâmetros. Dentre elas, a distribuição GUM- é imediatamente descartada. As métricas de ajuste, assim como o teste PPCC e o diagrama de momentos-L evidenciam que essa é a pior candidata para modelar as séries abrangidas neste estudo. A distribuição W2, embora tenha o menor viés das distribuições candidatas, apresentou ajuste inferior às demais candidatas na bacia do rio Doce – maior número de rejeições da H_0 do teste PPCC e piores valores de RMSE, MAE, RMSPE e MAPE. Por fim, as distribuições Ga, LN2 e Gum+ tiveram bons resultados nas métricas de aderência, embora as distribuições Gum+ e Ga tenham se sobressaído ligeiramente em relação à distribuição LN2.

Figura 5.7 - Medidas de aderência das distribuições candidatas aos postos fluviométricos



Os resultados acima sugerem que as distribuições Ga, LN2 e Gum+ são boas candidatas para modelar as vazões mínimas em ambas bacias em estudo.

Em seguida, a capacidade preditiva dos parâmetros dessas distribuições é investigada (Tabela 5.1 e Tabela 5.2). Foram avaliados 5 preditores, a área de drenagem (A), precipitação (P), declividade do talvegue principal (S), densidade de drenagem (DD) e a produção do aquífero (QSUB). Os parâmetros da distribuição Gum+ são explicados, em grande parte ($r > 0,9$), pela área de drenagem. A escalabilidade dos parâmetros dessa distribuição e a área de drenagem também foi relatada por Lima e Lall (2010). Essa mesma relação de escalabilidade com a área de drenagem é observada nos parâmetros de posição da distribuição LN2 e no parâmetro de escala da distribuição Ga. Os demais parâmetros das distribuições Ga e LN2, no

entanto, não possuem preditores que expliquem substancialmente a variação espacial desses parâmetros como a área de drenagem na distribuição Gumbel. Em termos práticos, a inclusão de covariáveis reduz o viés das estimativas em comparação a um modelo sem regressão e baseado apenas no valor médio observado (BEARD, 1962).

Tabela 5.1 - Correlação das covariáveis candidatas com os parâmetros das distribuições Gumbel (Gum+), Log-Normal 2P (LN2) e Gama(Ga) na bacia do rio Doce

Covariável	Gum+		LN2		Ga	
	log(μ)	log(σ)	μ	log(σ)	log(μ)	log(σ)
log(A)	0,932	0,973	0,934	-0,176	0,946	0,176
log(P)	0,091	-0,064	0,086	-0,564	-0,216	0,564
log(S)	-0,201	-0,143	-0,198	0,278	-0,074	-0,275
log(DD)	0,380	0,349	0,379	-0,238	0,293	0,237
log(QSUB)	0,208	0,209	0,210	-0,064	0,195	0,065

Nota: Correlações significativamente diferentes de 0 ($\alpha = 5\%$) estão em negrito

Tabela 5.2 - Correlação das covariáveis candidatas com os parâmetros das distribuições Gumbel (Gum+), Log-Normal 2P (LN2) e Gama(Ga) na bacia do rio Itajaí-Açu

Covariável	Gum+		LN2		Ga	
	log(μ)	log(σ)	μ	log(σ)	log(μ)	log(σ)
log(A)	0,957	0,967	0,958	0,285	0,954	-0,285
log(P)	-0,202	-0,115	-0,197	0,457	-0,042	-0,457
log(S)	-0,133	-0,150	-0,134	-0,127	-0,160	0,127
log(DD)	0,453	0,489	0,455	0,308	0,506	-0,308
log(QSUB)	0,054	0,082	0,055	0,176	0,103	-0,176

Nota: Correlações significativamente diferentes de 0 ($\alpha = 5\%$) estão em negrito

A distribuição escolhida para representar as vazões mínimas na duas áreas de estudo foi a distribuição Gumbel com causa superior exponencial, devido aos bons resultados nas métricas de aderência (teste PPCC e medidas de ajuste), parcimônia do modelo (2 parâmetros) e a superior capacidade preditiva do modelo. Essa distribuição não é limitada inferiormente, portanto, a distribuição de quantis pode alocar massa em valores de vazão negativa, os quais não possuem significado físico. Reconhece-se que uma forma mais adequada para representar a probabilidade das vazões mínimas é necessária, entretanto, acredita-se que para o quantil de interesse (TR de 10 anos), a estimação de vazões negativas seja mínima ou inexistente. Alternativamente, o ajuste de distribuições de probabilidade no inverso das observações de vazões mínimas impediria a atribuição de probabilidade à vazões negativas (BHATTI; KROLL; VOGEL, 2019). Essa abordagem foi avaliada por Bhatti *et al.* (2019), cujos resultados

sugerem que o ajuste de distribuições de probabilidade aos dados transformados e não transformados é similar, mas os autores indicam que mais investigações são necessárias. A transformação dos dados não foi contemplada neste estudo pode ter sido identificada tardiamente na revisão bibliográfica. Entretanto, trata-se de uma opção atraente para fazer uso da relação de escalabilidade dos parâmetros da distribuição Gumbel e estabelecer um limite, de forma a evitar a estimação de quantis negativos.

5.3 Modelo estacionário

5.3.1 Escolha da covariável

A primeira etapa consistiu em filtrar as covariáveis com correlação significativa com os parâmetros da distribuição de probabilidade adotada. Nas Tabela 5.1 e Tabela 5.2 foram apresentados os valores de correlação entre os preditores candidatos e os parâmetros da distribuição Gum⁺. Na bacia do rio Doce, a área de drenagem (A) e a densidade de drenagem (DD) apresentaram correlação significativa. Na bacia do rio Itajaí-Açu, a área de drenagem é novamente discriminada com uma correlação forte e significativa. Outras covariáveis, no entanto, não tiveram correlação significativa. A pequena quantidade de postos nessa bacia torna a tomada de decisão por meio de testes de significância mais difícil, uma vez que a hipótese nula (H_0 : a correlação é diferente de zero) é mais facilmente rejeitada. Optou-se por investigar a densidade de drenagem, além da área de drenagem, em função da sua correlação com os parâmetros da distribuição Gumbel na outra bacia analisada.

Neste momento, algumas ressalvas devem ser feitas em relação a qualidade ou representatividade das covariáveis físicas selecionadas. A densidade de drenagem e principalmente a hidrogeologia foram obtidas de bancos de dados com escalas que podem não refletir de maneira acurada a realidade de pequenas bacias (no caso da hidrogeologia, cuja escala é de 1:1.800.000, a preocupação é maior). Por esse motivo, seria injusto afirmar categoricamente que essas covariáveis não influenciam no processo. Talvez sob uma resolução melhor, correlações maiores poderiam se manifestar entre os dados analisados, no entanto, quantidade e qualidade dos dados é um obstáculo comum em estudos hidrológicos.

No final, 4 modelos foram avaliados partindo de um modelo mais simples apenas com a área de drenagem e considerando a introdução da densidade de drenagem nos parâmetros da distribuição Gum+. Os resultados são apresentados na Tabela 5.3. Os valores do DIC variaram minimamente entre os modelos investigados, tornando-os indistinguíveis em termos de desempenho. O desempenho similar dos modelos é atribuído à alta capacidade preditiva da área de drenagem, que conforme demonstrado na Tabela 5.1 e Tabela 5.2, sozinha explica a maior parte da variabilidades dos parâmetros da distribuição Gumbel. Uma vez o desempenho dos modelos foi semelhante, adota-se o modelo mais simples composto exclusivamente pela área.

Tabela 5.3 - Resultados do DIC para a escolha da covariável

N	Modelo	Doce			Itajaí-Açu		
		\bar{D}	pD	DIC	\bar{D}	pD	DIC
1	$\log(\mu) = \alpha_0 + \alpha_1 \log(A)$ $\log(\sigma) = \beta_0 + \beta_1 \log(A)$	6787,8	82,7	6870,5	4168,0	25,9	4193,9
2	$\log(\mu) = \alpha_0 + \alpha_1 \log(A) + \alpha_2 \log(DD)$ $\log(\sigma) = \beta_0 + \beta_1 \log(A)$	6788,2	83,1	6871,3	4167,8	26,0	4193,8
3	$\log(\mu) = \alpha_0 + \alpha_1 \log(A)$ $\log(\sigma) = \beta_0 + \beta_1 \log(A) + \beta_2 \log(DD)$	6787,8	82,3	6870,2	4168,0	26,2	4194,1
4	$\log(\mu) = \alpha_0 + \alpha_1 \log(A) + \alpha_2 \log(DD)$ $\log(\sigma) = \beta_0 + \beta_1 \log(A) + \beta_2 \log(DD)$	6787,7	82,8	6870,5	4167,7	26,0	4193,7

Nota: o menor DIC está em negrito

5.3.2 Ajuste dos modelos

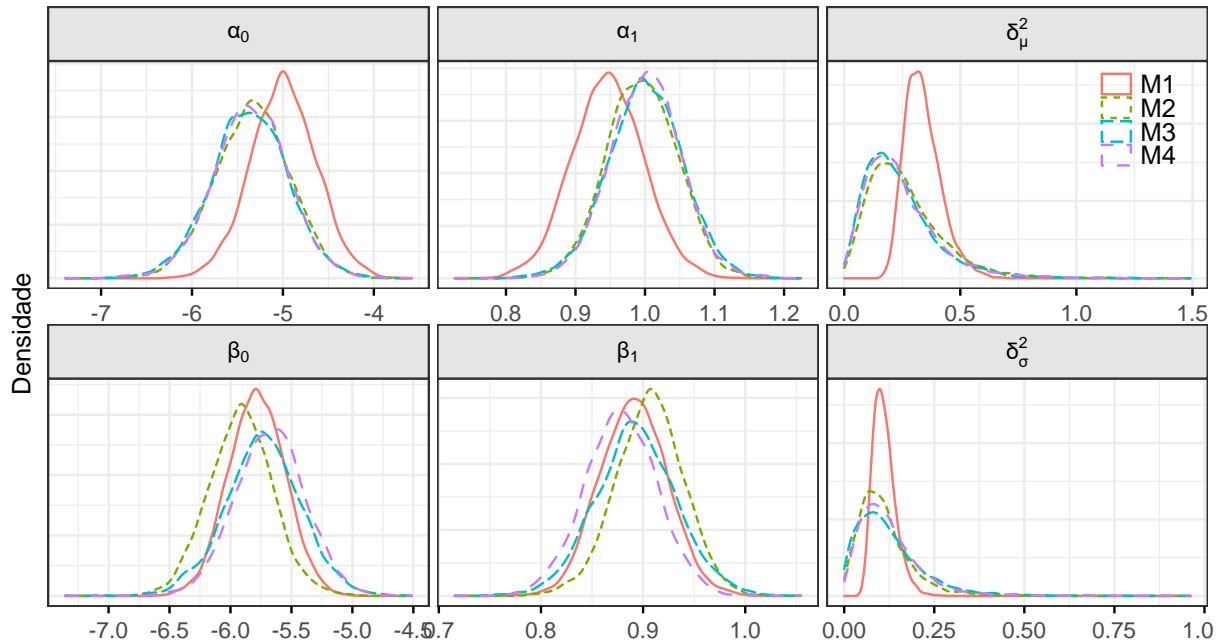
Os resultados apresentados a seguir são organizados para levantar e discutir os aspectos e premissas dos modelos estacionários elaborados, M1 a M4, e por fim, avaliar a capacidade preditiva dos modelos com auxílio dos postos de validação.

Inicialmente, os parâmetros do modelo regressivo de μ e σ dos MHBs são analisados, começando em maior detalhe pela bacia do rio Doce (Figura 5.8). O intercepto (α_0) e a declividade (α_1) da regressão de μ apresentaram diferenças entre os modelos. Ao levar em consideração a dependência espacial dos resíduos da regressão, houve um deslocamento horizontal dessas duas quantidades do modelo M1 para os demais modelos e um pequeno aumento da incerteza no parâmetro α_0 , possivelmente em resposta à introdução dos resíduos no modelo regressivo. As variâncias (δ_{μ}^2) da

regressão de μ também diferiram entre os modelos. No modelo M1, a incerteza desse parâmetro foi inferior aos demais modelos, porém, a maior parte da massa da distribuição está alocada em maiores valores de δ_μ^2 . Com o relaxamento da hipótese de independência dos resíduos, os intervalos da distribuição *a posteriori* aumentaram, permitindo a inferência de menores valores de δ_μ^2 nos modelos M2 a M4. Inicialmente, a expectativa era de que a correlação existente entre as observações inflasse a quantidade de informação disponível e resultasse em parâmetros com menor variância, todavia, a diferença nas estimativas dos parâmetros citados entre os modelos M2, M3 e M4 foi mínima.

Na regressão de σ , por outro lado, diferenças mais expressivas foram observadas entre os modelos em todos os parâmetros. Um deslocamento horizontal pode ser observado entre os modelos M1 e M2 no intercepto (β_0) e na declividade (β_1) da regressão de σ . O relaxamento da hipótese de independência das observações influenciou mais ativamente a distribuição *a posteriori* dos parâmetros da regressão. As variâncias de β_0 e β_1 foram maiores nos modelos M3 e M4 em relação aos demais modelos. Além disso, houve deslocamento da distribuição *a posteriori* desses parâmetros em relação ao modelo M2, sobre o qual os modelos M3 e M4 foram construídos. Por fim, parece que novamente a dependência espacial dos resíduos alterou a incerteza e a média da variância da regressão (δ_σ^2), colocando mais massa *a posteriori* em valores mais baixos de δ_σ^2 e permitindo que valores maiores δ_σ^2 fossem inferidos. A variância de δ_σ^2 também aumentou do modelo M2 para os modelos M3 e M4.

Figura 5.8 - Distribuição a posteriori dos parâmetros das regressões dos modelos M1-M4 na bacia do rio Doce



Os parâmetros da regressão na bacia do rio Itajaí-Açu são apresentados na Figura 5.9. Nessa bacia, as estimativas dos modelos M2-M4 foram diferentes do modelo M1 em todos os parâmetros, com exceção de α_0 e α_1 . Dois pontos notáveis diferenciam a resposta dos parâmetros dos modelos nessa bacia em relação à bacia do rio Doce: (i) os modelos M3 e M4 exerceram mais influência na regressão de μ , enquanto na bacia do rio Doce isso ocorreu na regressão de σ ; e (ii) as médias de δ_μ^2 e δ_σ^2 aumentaram do modelo M1 para os modelos M2-M4, enquanto na bacia do rio Doce essas quantidades diminuíram.

Em seguida, a regressão do modelo M1 é plotada (Figura 5.10) para fins de visualização da escalabilidade com a área de drenagem, das estimativas do modelo e da incerteza preditiva. A estimativa pontual da regressão de μ foi similar entre as bacias, apesar das grandes diferenças entre os regimes pluviométricos, tanto em termo de sazonalidade quanto de totais acumulados, e suas respectivas hidrogeologias. O intervalo preditivo do modelo regressivo de μ é maior na bacia do rio Doce, provavelmente em reflexo à maior heterogeneidade nas características físicas dessa bacia.

Figura 5.9 - Distribuição a posteriori dos parâmetros das regressões dos modelos M1-M4 na bacia do rio Itajaí-Açu

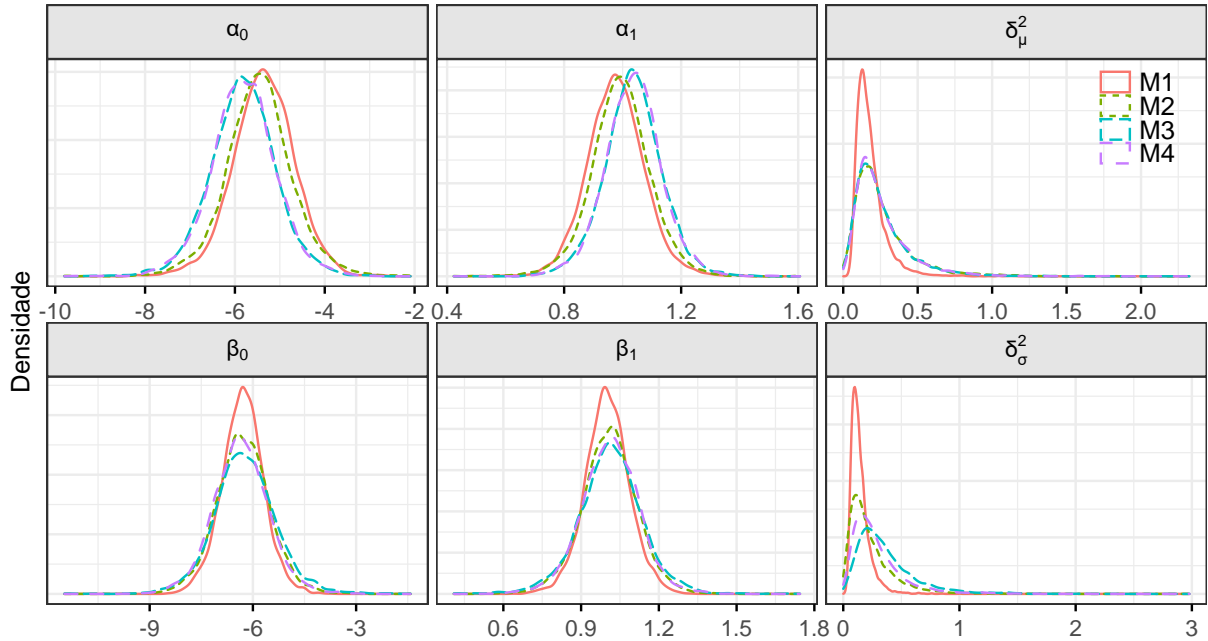
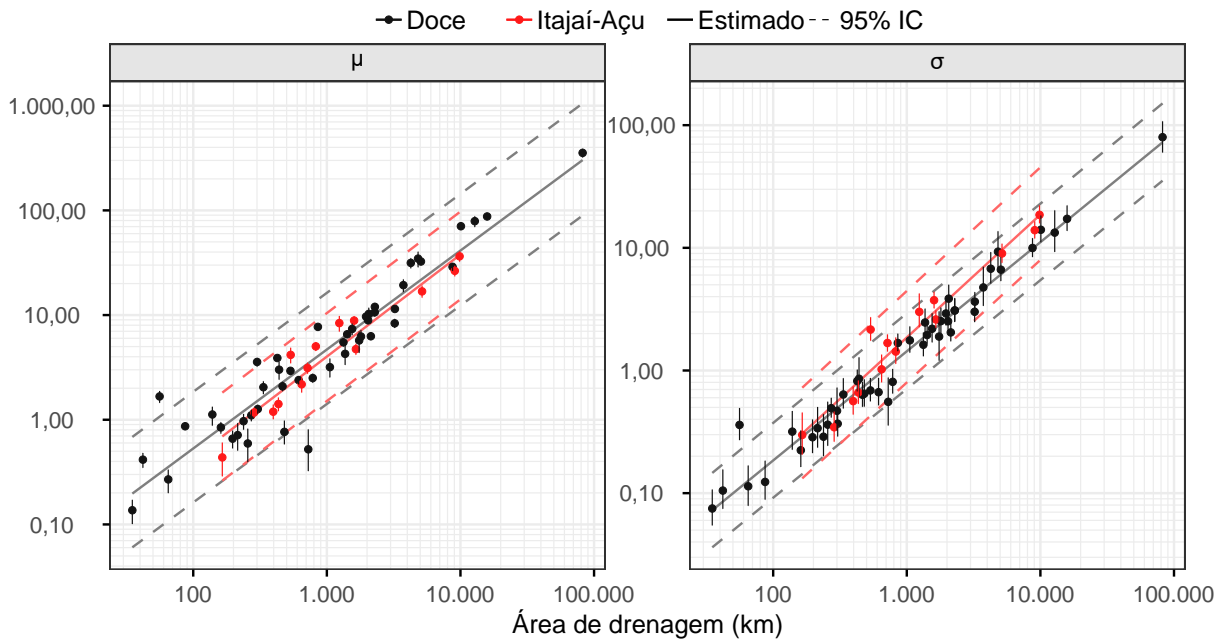


Figura 5.10 - Regressão dos parâmetros de posição e escala da distribuição Gum+ pelo modelo M1. Linha tracejada representa o intervalo de 95% de credibilidade preditivo. As estimativas pontuais e o intervalo de 95% de credibilidade desses parâmetros são representadas, respectivamente, por pontos e linhas verticais.



Quanto à regressão de σ , que controla o achatamento da função de distribuição de probabilidade, a evolução da escalabilidade do parâmetro com a área de drenagem ocorre de maneira diferente entre as bacias. A declividade da reta na bacia do rio Itajaí-Açu é maior, o que sugere que bacias maiores tenham maior variabilidade anual de vazões mínimas (menos massa no centro da distribuição e mais nas caudas) do que na bacia do rio Doce.

De maneira geral, a quantificação da incerteza preditiva dos parâmetros pelas regressões do modelo M1 parecem contemplar adequadamente a incerteza dessas quantidades. O intervalo de credibilidade (IC) da bacia do rio Itajaí-Açu abrangeu todas as estimativas pontuais dos parâmetros da distribuição Gumbel, enquanto na bacia do rio Doce, ao todo, 2 postos fluviométricos ficaram fora do IC de um total de 48 postos, a saber 56385000-São Miguel Do Anta e 56846890-Fazenda Aconchego. O desvio do comportamento regional do último posto pode ser atribuído, em parte, a incerteza devido à curta amostra. Já no primeiro posto citado, o qual possui uma extensa série de mínimos, esse fato deve estar relacionado a um aspecto físico da bacia (não identificado sob a análise das covariáveis investigadas).

A investigação e a disponibilidade de mais covariáveis climáticas poderia auxiliar a descrever parte da variação dos parâmetros μ e σ não capturados pela área de drenagem, e conseqüentemente, melhorar a acurácia e precisão do modelo.

A predição dos parâmetros dos demais modelos não é plotada pois a incerteza preditiva está condicionada ao campo de resíduos das regressões e ao local da predição.

O modelo espacial ajustado aos resíduos dos parâmetros dos modelos é ilustrado na **Erro! Fonte de referência não encontrada.** Conforme citado na metodologia, o processo Gaussiano requer distribuições *a priori* mais informativas devido a problemas de identifiabilidade com o *range* do variograma. Assim, a correlação espacial *a posteriori* foi bastante similar entre os modelos das diferentes bacias, devido à similaridade com a distribuição *a posteriori* do *range* com a distribuição *a priori*.

Os resíduos da regressão dos parâmetros μ e σ do modelo M1, sob a hipótese de independência dos resíduos, foram analisados quanto à existência de autocorrelação espacial por meio da geoestatística. O variograma experimental das medianas dos resíduos é apresentado na Figura 5.11. Na bacia do rio Doce, nota-se em ambos os variogramas experimentais dos parâmetros da distribuição Gumbel um crescimento da variância espacial até aproximadamente uma distância de 50 km, a partir da qual a variância se estabiliza. Esse resultado sugere que os resíduos das regressões de ambos os parâmetros são espacialmente autocorrelacionados. Na bacia do rio Itajaí-

Açu, por outro lado, devido a quantidade limitada de dados, apenas 14, a interpretação e a confiabilidade do variograma experimental são dificultadas. Esses resultados indicam que os resíduos manifestam parte da variabilidade espacial não capturada pelo modelo regressivo, conforme reportado por Kjeldsen e Jones (2007, 2009), o que sugere que a linha de trabalho seguida pelo modelo M2 é coerente.

Figura 5.11 - Dependência espacial dos parâmetros da distribuição Gum+ pelos modelos M2 e M4

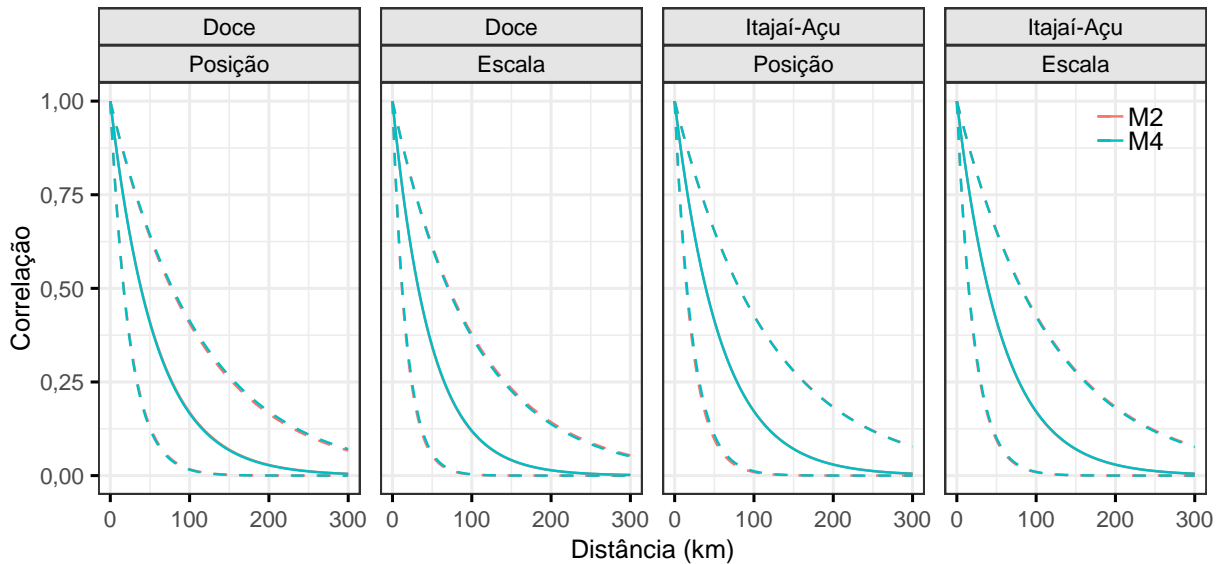
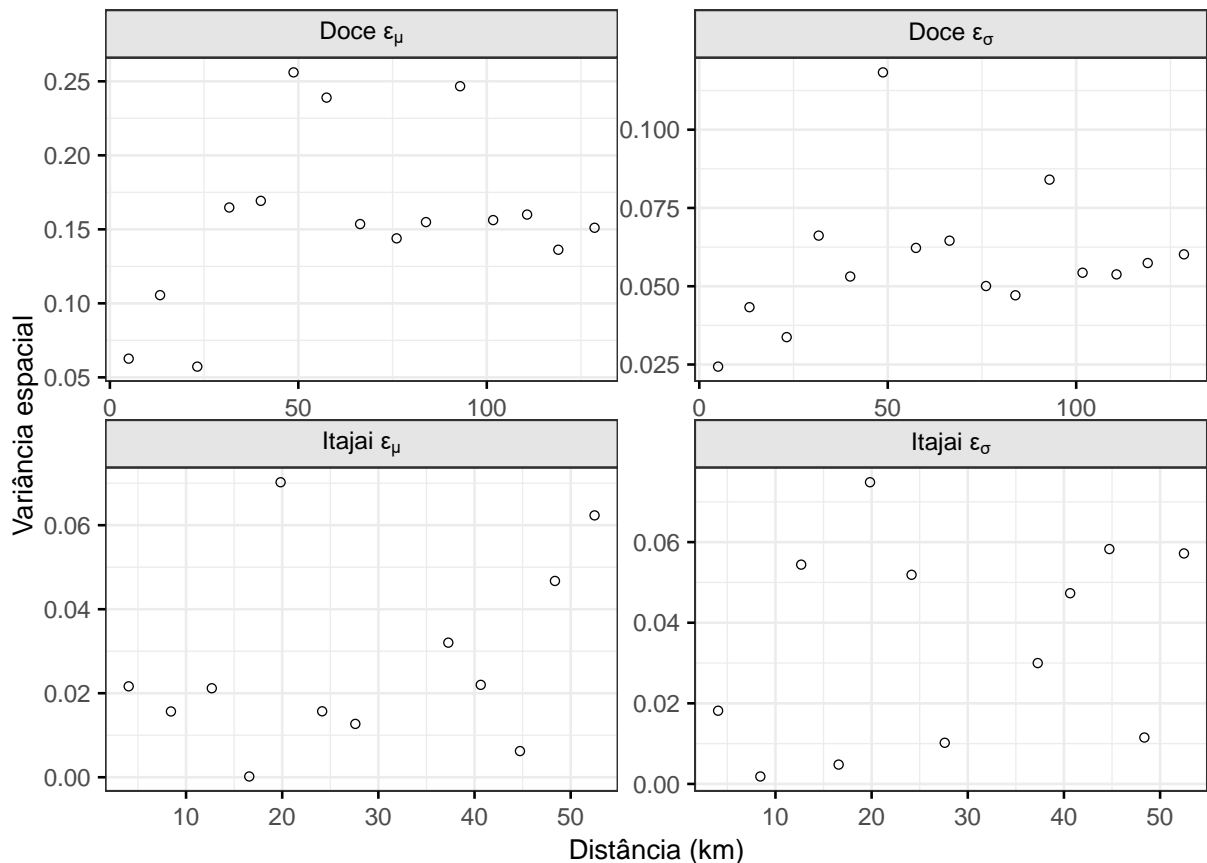


Figura 5.12 - Semivariograma empírico das medianas dos resíduos das regressões do logaritmo dos parâmetros de posição e escala do modelo M1



Com o intuito de avaliar a influência dos resíduos das regressões na predição do modelo e seus respectivos domínios espaciais, estimou-se a distribuição preditiva dos resíduos das regressões dos modelos em 2826 pontos, espaçados entre si em 3'. Esse procedimento foi realizado com o modelo M2, e para fins de investigar a influência da correlação dos dados nessa variável, foi também realizado com o modelo M4.

Os resultados, em termos de mediana dos resíduos dos modelos M2 e M4, são exibidos, respectivamente, nas Figuras Figura 5.14 e Figura 5.13 para a bacia do rio Doce. Na Figura 5.14, fica evidente a aparição de um padrão espacial nos resíduos dos dois parâmetros da distribuição Gumbel, os quais podem ser úteis para transferir informações para locais não monitorados e potencialmente reduzir a incerteza preditiva. Na Figura 5.13, um padrão quase idêntico é concebido pelo modelo M4, no entanto, percebe-se que os resíduos em algumas regiões (mais perceptível em regiões com efeito *bull's eye*) são levemente maiores em valores absolutos.

Figura 5.13 - Resíduos do parâmetro (a) μ e (b) σ do modelo M4 na bacia do rio Doce

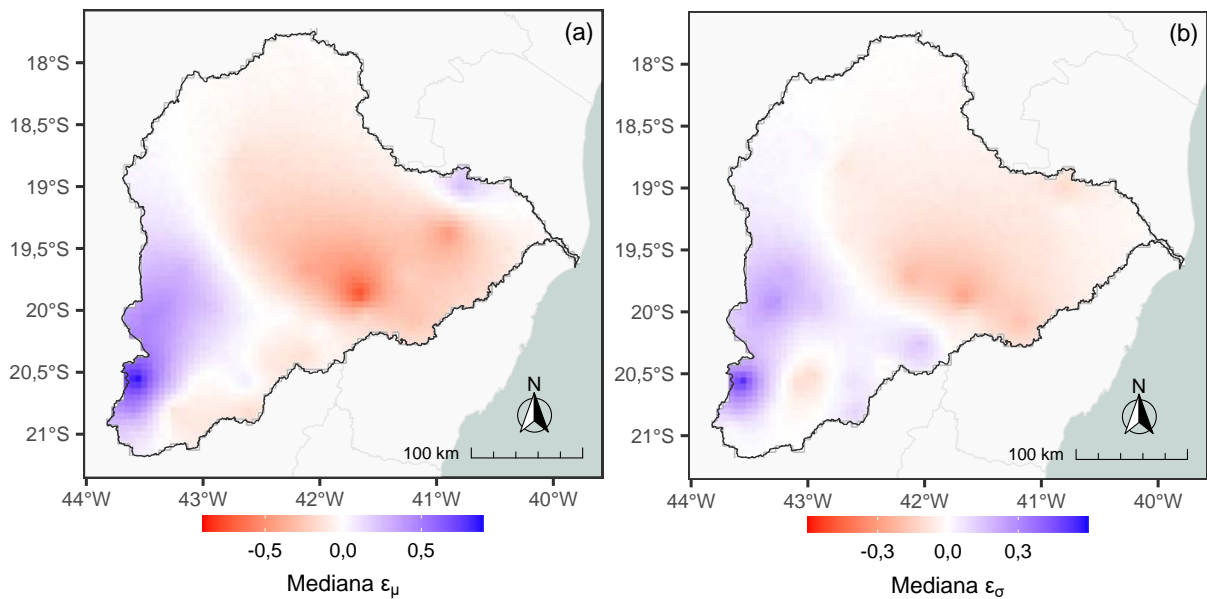
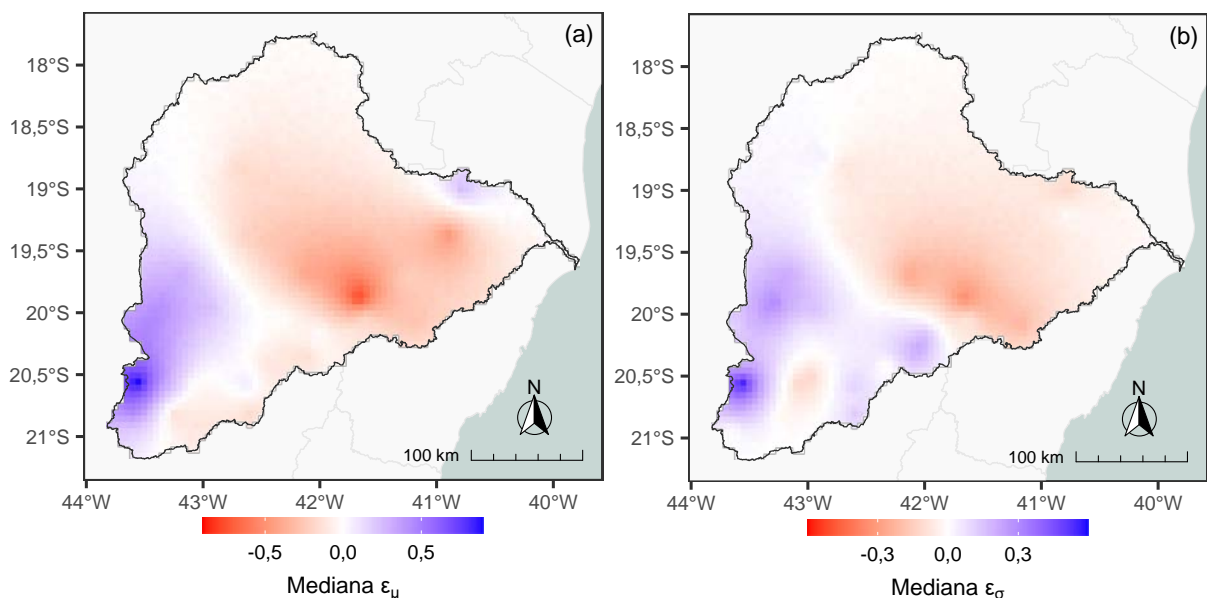


Figura 5.14 - Resíduos do parâmetro (a) μ e (b) σ do modelo M2 na bacia do rio Doce



Nas Figura 5.16 e Figura 5.15 são mostradas as medianas dos resíduos dos modelos M2 e M4 na bacia do rio Itajaí-Açu. Nota-se também a manifestação de um campo espacial de alguma forma similar ao campo dos resíduos dos parâmetro de posição e escala da GEV de Sampaio e Costa (2021), isso é, resíduos positivos à montante e negativos à jusante da bacia. Os resíduos do modelo M4 nessa bacia tem comportamento similar aos da bacia do rio Doce, na medida que repetem o padrão espacial do modelo M2, mas com pequenas mudanças na magnitude da mediana dos resíduos em algumas regiões.

Figura 5.15 - Resíduos do parâmetro (a) μ e (b) σ do modelo M4 na bacia do rio Itajaí-Açu

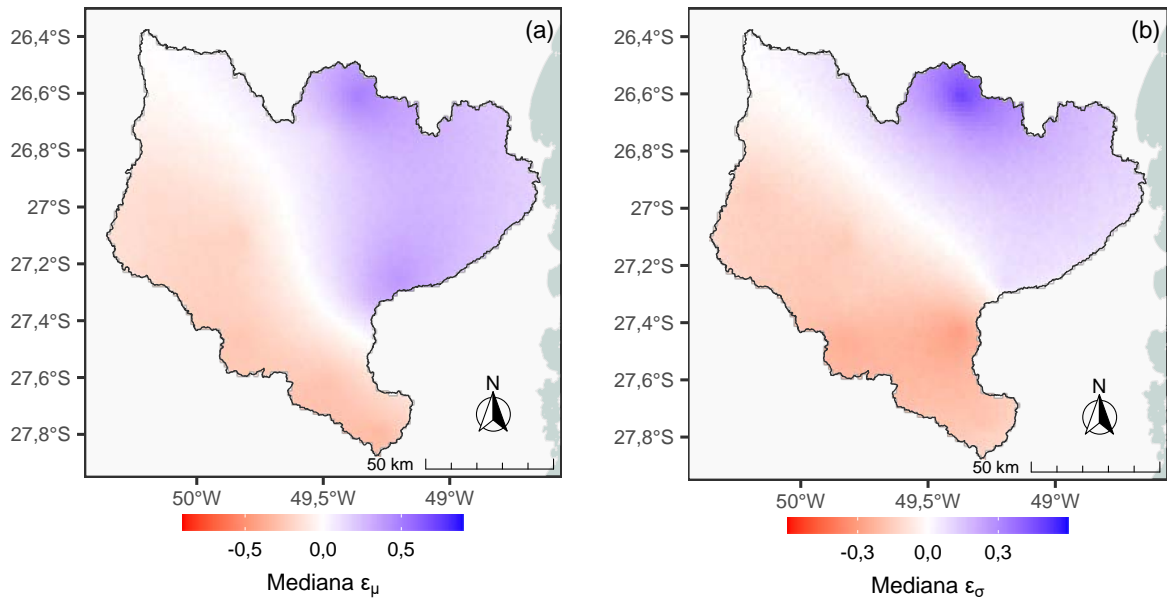
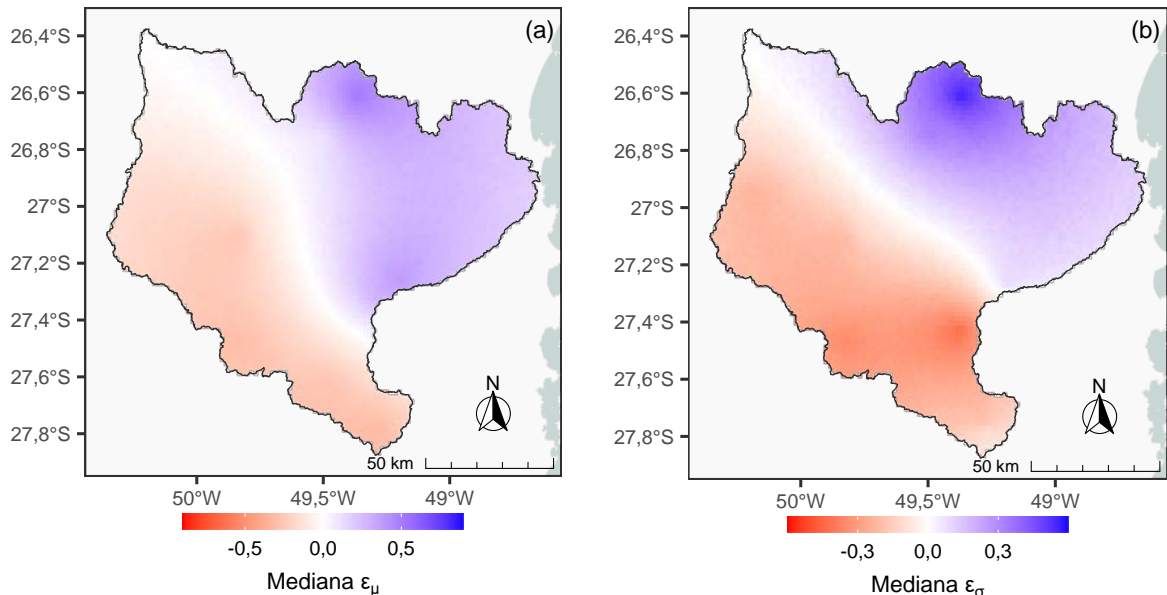


Figura 5.16 - Resíduos do parâmetro (a) μ e (b) σ do modelo M2 na bacia do rio Itajaí-Açu



Esses resultados sugerem que mudanças na distribuição *a posteriori* (e.g. média e variância) dos parâmetros das regressões decorrentes da introdução da dependência conjunta de observações são capturadas pelos resíduos da regressão. Os resíduos, por sua vez, parecem ter maiores valores absolutos, possivelmente para corrigir mudanças nas expectativas das regressões.

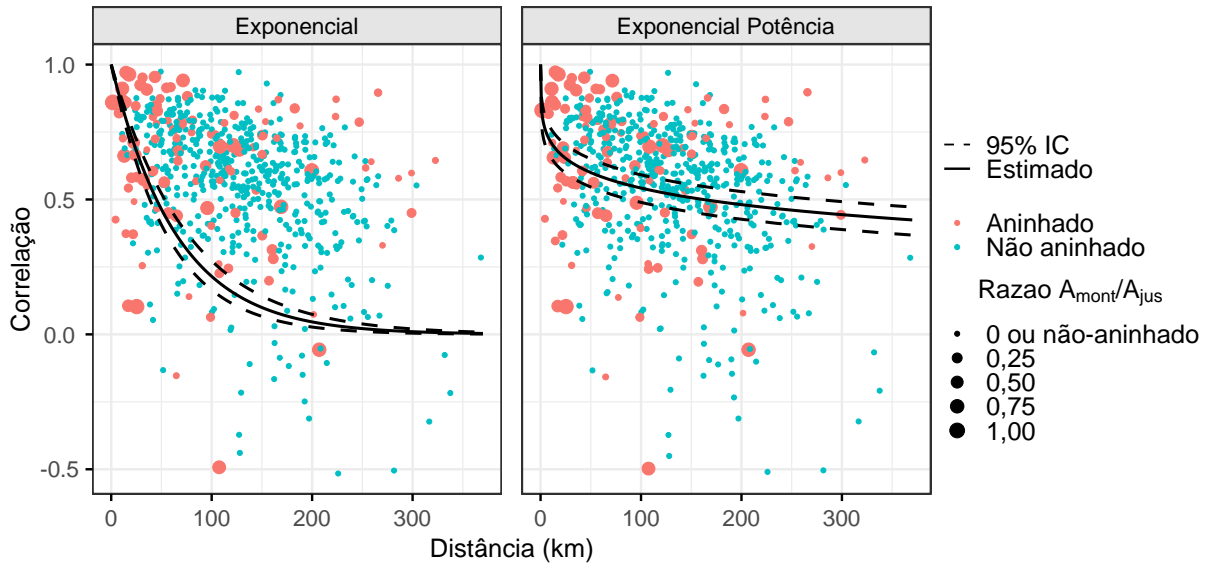
A atenção agora é voltada para a estrutura de dependência dos dados, avaliada nos modelos M3 e M4. As Figura 5.17 e Figura 5.18 trazem a correlação espacial inferida pelos MHBs junto com os pares de correlação inferidos dos dados. A correlação para um par genérico de postos A e B é calculada pela mediana da correlação entre $\Phi^{-1}(F(\mathbf{y}_A|\mu_A, \sigma_A))$ e $\Phi^{-1}(F(\mathbf{y}_B|\mu_B, \sigma_B))$, na qual F é a função acumulada de

probabilidade da distribuição Gumbel e Φ^{-1} é a função inversa da distribuição normal, calculada para cada realização de μ_A, σ_A, μ_B e σ_B . Esse cálculo visa manter a consistência com a metodologia de cálculo da cópula Gaussiana, cuja matriz de dependência (dependograma) é inferida a partir da aplicação da função normal inversa Gaussiana nas probabilidades acumuladas da distribuição marginal do MHB. A estrutura de correlação espacial discutida a seguir é baseada na correlação da probabilidade de não-excedências das observações, mas por simplificação do termo, ela será referida como correlação dos dados/vazão. Além disso, por clareza nos resultados, ressalta-se que a correlação dos dados não deve ser vista como uma correlação empírica ou não-paramétrica, uma vez que ela parte do ajuste dos modelos aos dados e, portanto, está condicionada à suas respectivas hipóteses (e.g. distribuição marginal, distribuição conjunta, dependograma).

Os resultados da correlação espacial dos dados na bacia do rio Doce são apresentados na Figura 5.17. O ajuste das funções exponencial (modelo M3) e exponencial de potência (modelo M4) delineiam duas estruturas distintas de correlação espacial. A função exponencial decai em função da distância até chegar em valores próximos de 0 para distâncias maiores que 250 km. A função exponencial de potência, por outro lado, tem um decaimento muito menos acentuado e define maiores valores de correlação em relação à função exponencial para distâncias maiores que 40 km. O modelo M4 parece descrever melhor a dependência espacial dos dados, que continua elevada para grandes distâncias.

Figura 5.17 - Dependência espacial das observações da bacia do rio Doce. As linhas representam a correlação espacial com um modelo exponencial (M3), painel à esquerda, e um

modelo exponencial potência (M4), painel à direita. Apenas pares de postos fluviométricos com pelo menos 10 observações concomitantes são plotados

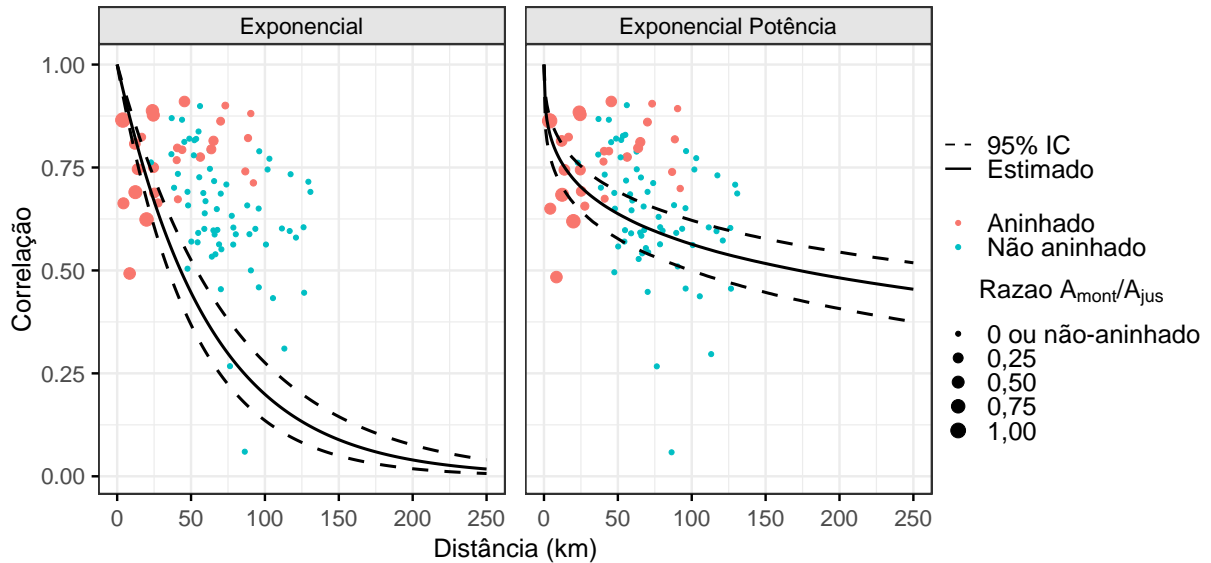


A natureza dos pares de correlação em relação a rede hidrográfica também é assinalada pela Figura 5.17. Valores elevados de correlação são observados para pares aninhados com áreas de drenagem semelhante, o que é bastante intuitivo. Por outro lado, à medida que a razão entre as áreas dos pares aninhados diminui, a correlação da magnitude das vazões nessas bacias varia substancialmente e se torna similar à correlação dos pares de bacias não aninhadas, mesmo para razões de área montante/jusante consideráveis, e.g., 0,5 a 0,75.

A seguir, a correlação das vazões na bacia do rio Itajaí-Açu é apresentada na Figura 5.18. Novamente, a maior flexibilidade da função exponencial potência parece capturar melhor a estrutura de dependência com menor decaimento manifestada pela correlação inferida com os dados. Nessa bacia há uma disparidade mais evidente entre a estrutura de correlação de postos aninhados e não aninhados. Os postos aninhados parecem manter uma correlação mais forte em comparação a postos não aninhados, mesmo para maiores distâncias. Contudo, para distâncias menores que 25 km, é possível notar muitos pares de postos aninhados com áreas semelhantes apresentando valores de correlação não muito altos. Acredita-se que essas diferenças de comportamento estão relacionadas à homogeneidade/heterogeneidade interbacias aninhadas.

Figura 5.18 - Dependência espacial das observações da bacia do rio Itajaí-Açu. As linhas representam a correlação espacial com um modelo exponencial (M3), painel à esquerda, e um

modelo exponencial potência (M4), painel à direita. Apenas para pares com pelo menos 10 observações concomitantes



Uma análise visual dos pares de correlação dos dados entre as duas bacias estudadas não parece sugerir que uma bacia manifeste dependência entre dados mais forte do que a outra. Além disso, o modelo mais flexível de dependência dos dados, exponencial potência, parece descrever estruturas de dependência semelhantes para as duas bacias (i.e. decaimento similar da correlação espacial). Esses fatores sugerem que, embora pareça haver uma dependência mais forte entre bacias aninhadas no rio Itajaí-Açu, a dependência global dos dados nas bacias é similar.

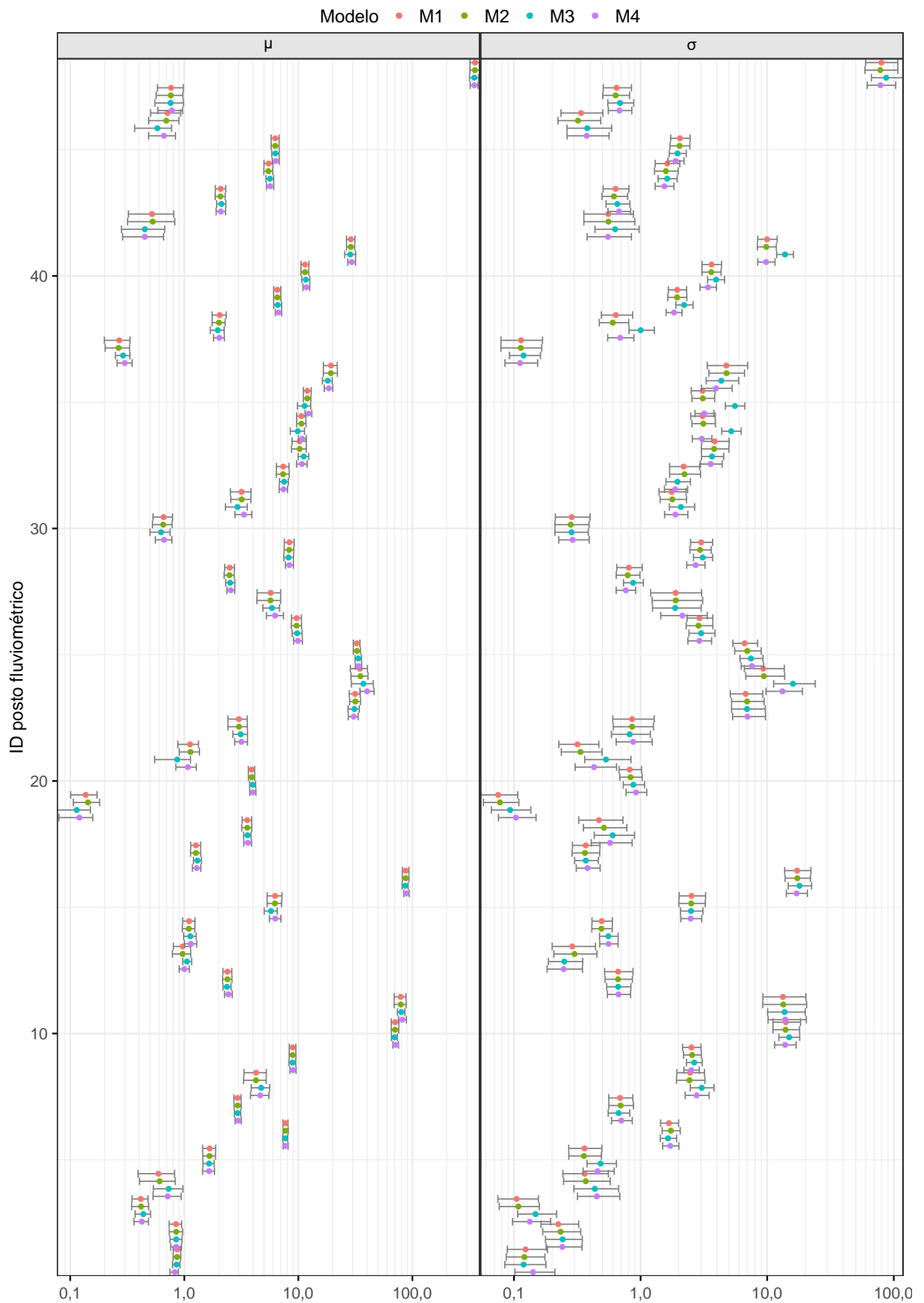
Gudmundsson *et al.* (2011), em um estudo de larga escala na Europa, observaram que a correlação cruzada de vazões mínimas é menor do que de vazões médias/máximas para curtas distâncias (< 400 km) e atribuíram essa diferença à grande influência das propriedades das bacias, como hidrogeologia e topografia. Em outras palavras, a heterogeneidade espacial da bacia teria um papel muito importante nas vazões mínimas. A grande dispersão de pontos de bacias aninhadas relatada na bacia do rio Doce pode estar associada ao domínio do processo pelas características físicas da bacia, como teorizado por Gudmundsson *et al.* (2011). Nessa concepção, mesmo que uma parcela da vazão do posto a jusante tenha origem em comum com a vazão do posto a montante, a variabilidade espacial da característica física de outras áreas de contribuição estaria exercendo grande influência nos dados. Não se descarta também a influência antrópica, como a captação de água para abastecimento humano e irrigação.

A correlação espacial identificada nesses resultados traz implicações práticas na gestão de recursos hídricos. Em um cenário de escassez hídrica, em função da correlação espacial dos dados, é provável que a seca atinja grandes extensões das bacia hidrográficas.

A discussão agora é direcionada para os parâmetros e quantis estimados nos postos de calibração para verificar os ajustes dos parâmetros da distribuição Gumbel e as curvas de quantis.

Na Figura 5.19 foram plotados os parâmetros estimados da distribuição Gumbel na bacia do rio Doce. As estimativas dos modelos M1 e M2 foram praticamente idênticas para ambos os parâmetros da distribuição Gumbel. Algo similar foi constatado por Sampaio e Costa (2021) para a distribuição GEV, família da qual a distribuição Gumbel é um caso particular. Em locais monitorados, portanto, a regionalização pelo M2 não traz benefícios em relação à sua contraparte mais parcimoniosa, o modelo M1.

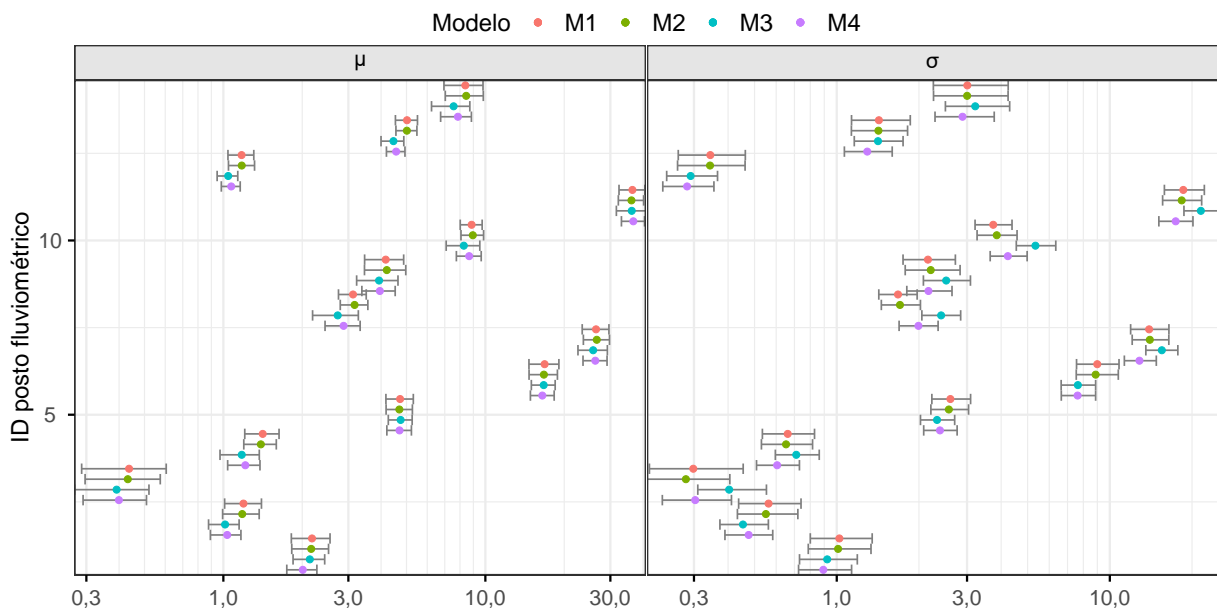
Figura 5.19 - Parâmetros estimados nos postos de calibração da bacia do rio Doce. Pontos e barras de erro representam, respectivamente, a estimativa pontual e seu intervalo de credibilidade de 95%



Por outro lado, as estimativas dos modelos M3 e M4 diferem entre si e dos modelos M1/M2, tanto em termos da estimativa pontual dos parâmetros da distribuição Gumbel, quanto de seus intervalos de credibilidade. As diferenças observadas não são consistentes no domínio do espaço, isso é, o viés das estimativas dos modelos M3/M4 em relação às estimativas do modelo M1/M2 não teve um sinal fixo. Além disso, em alguns postos os ICs dos modelos M3 e M4 foram mais largos em comparação aos modelos mais parcimoniosos, em outros postos seus ICs foram mais amplos. Essa variação nos parâmetros estimados entre os modelos surge tanto em postos com muitos dados (e.g. 29, 41), quanto postos com menos dados (e.g. 5, 21).

O cenário se repete na bacia do rio Itajaí-Açu (Figura 5.20). As estimativas dos modelos M1 e M2 convergem entre si, enquanto que as estimativas dos modelos M3 e M4 diferem das demais em muitos postos.

Figura 5.20 - Parâmetros estimados nos postos de calibração da bacia do rio Itajaí-Açu. Pontos e barras de erro representam, respectivamente, a estimativa pontual e seu intervalo de credibilidade de 95%



As implicações das estimativas heterogêneas entre os modelos é mostrada nas curvas de quantis em três postos de calibração de cada bacia. Nessas curvas de quantis, assim como nas demais deste trabalho, emprega-se posição de plotagem de Cunnane para o cálculo do tempo de retorno empírico dos dados. Na Figura 5.21 e

Figura 5.22, são mostrados os ajustes nos postos de calibração da bacia do rio Doce e Itajaí-Açu, respectivamente. Postos com menos dados são mais sensíveis às especificações do modelo, e.g. IDs 5 e 21 na bacia do rio Doce e ID 3 na bacia do rio Itajaí Açu, enquanto que em postos com mais dados, em função do maior peso da verossimilhança, a diferença entre ajustes de diferentes modelos é reduzida.

Figura 5.21 - Curva de quantis em três postos de calibração da bacia do rio Doce. Linhas contínuas e pontilhadas representam, respectivamente, as estimativas pontuais e o intervalo de predição com 95% de credibilidade dos modelos. No cabeçalho dos painéis está código e o ID (entre parêntese) dos postos

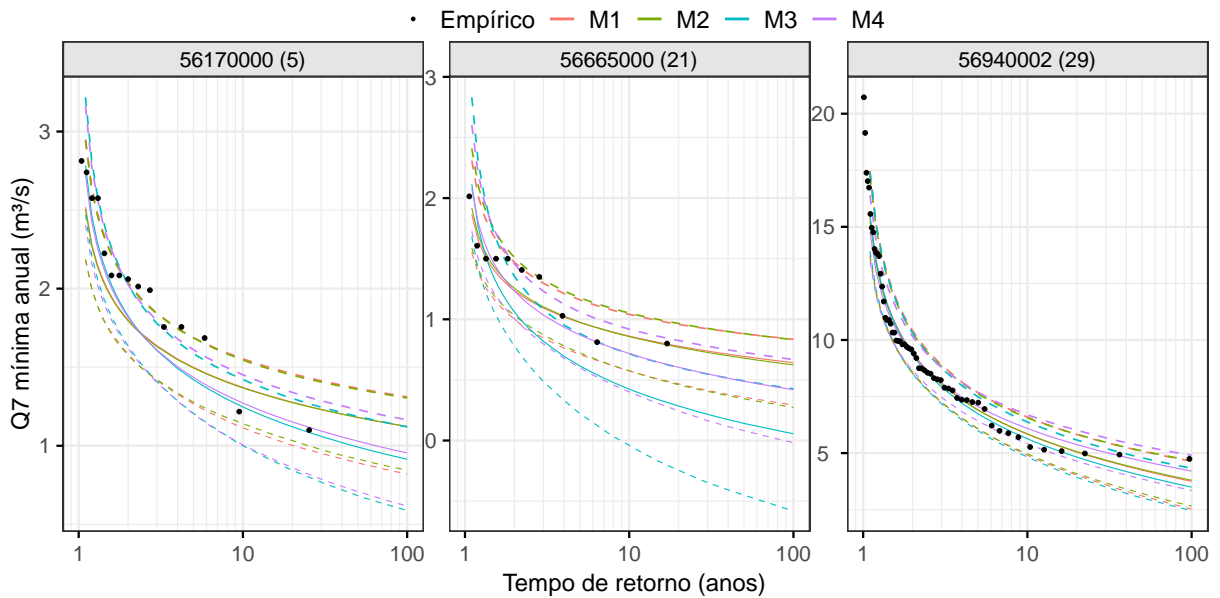
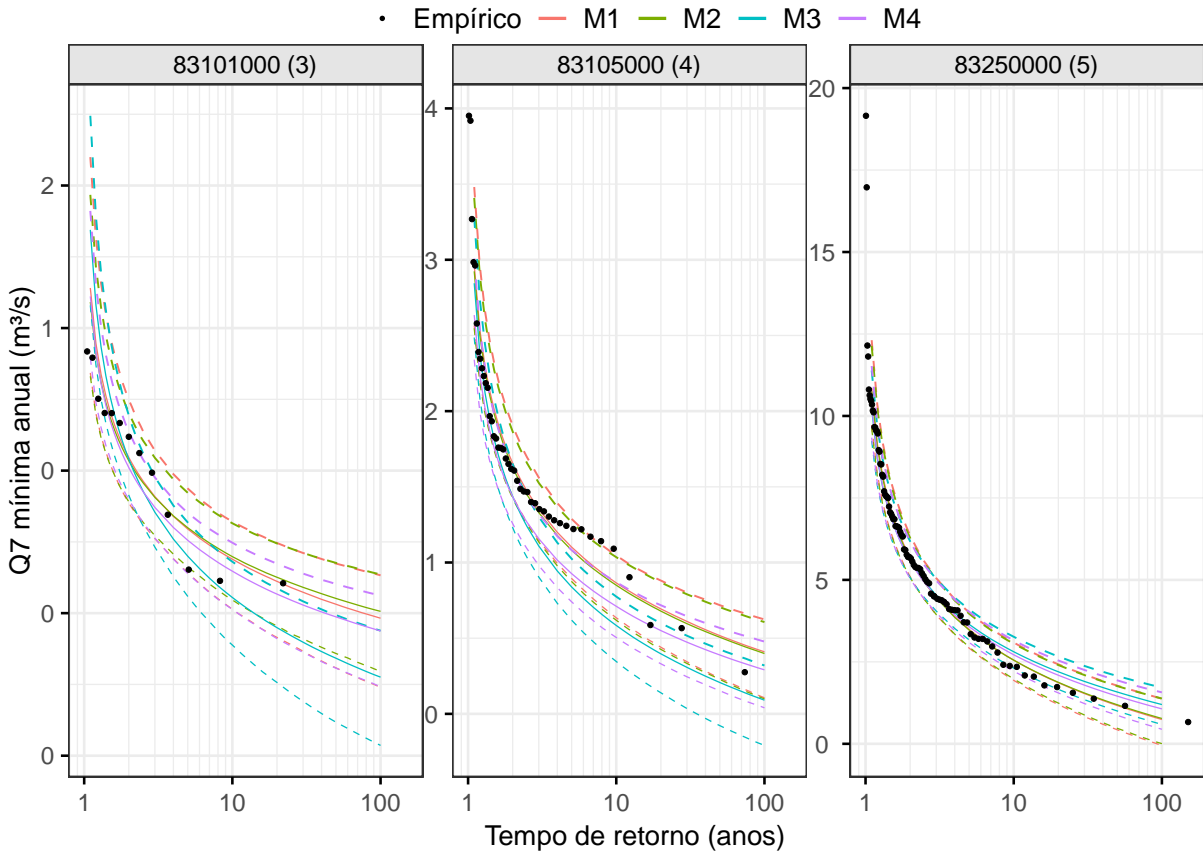


Figura 5.22 - Curva de quantis em três postos de calibração da bacia do rio Itajaí-Açu. Linhas contínuas e pontilhadas representam, respectivamente, as estimativas pontuais e o intervalo de predição com 95% de credibilidade dos modelos. No cabeçalho dos painéis está código e o ID (entre parêntese) dos postos



Para quantificar o desempenho dos modelos nos postos de calibração, recorreu-se novamente ao DIC. Na Tabela 5.4, são apresentados os valores do DIC calculados. Eles permitem confirmar que os modelos M1 e M2 tem desempenho semelhante nos postos de calibração, conforme esperado, em função das estimativas semelhantes dos parâmetros da distribuição Gumbel. Por outro lado, o modelo M3 teve pior desempenho dentre os modelos avaliados nas duas bacias estudadas. Na bacia do rio Doce, o modelo M4 teve ajuste superior aos demais, enquanto o modelo M4 foi inferior aos modelos M1 e M2 bacia do rio Itajaí-Açu.

Tabela 5.4 - Resultados do DIC para os modelos M1 a M4 nos postos de calibração

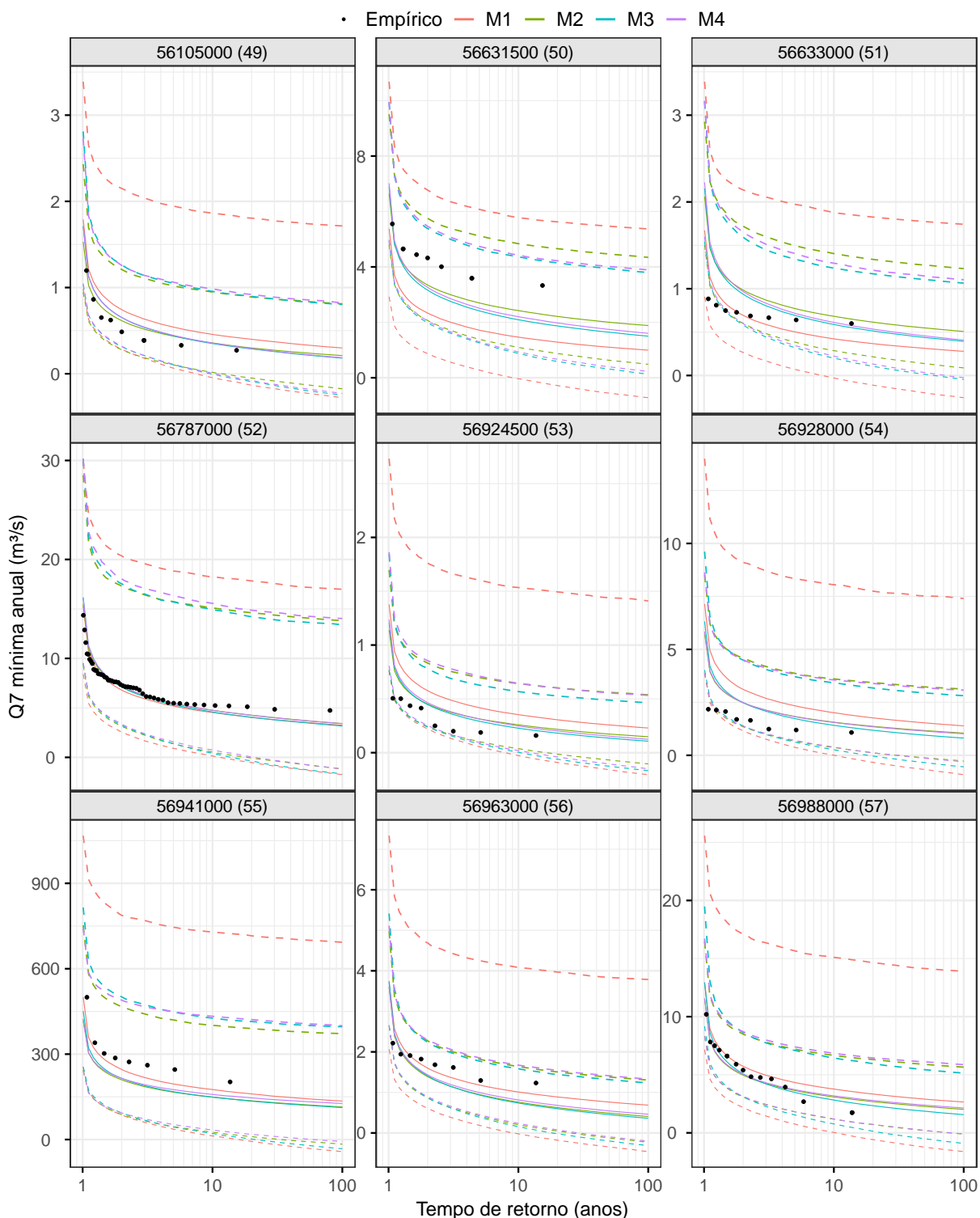
Modelo	Doce			Itajaí-Açu		
	\bar{D}	pD	DIC	\bar{D}	pD	DIC
M1	6787,5	82,7	6870,3	4168,0	25,8	4193,9
M2	6787,9	82,0	6869,9	4166,7	25,5	4192,3
M3	6865,4	55,3	6920,7	4237,1	18,0	4255,0
M4	6790,7	56,6	6847,3	4190,9	19,5	4210,5

Esses resultados sugerem que uma especificação mais simples da estrutura de

dependência entre os dados pode deteriorar as estimativas dos MHBs nos postos de calibração, até mesmo para quantis de tempo de retorno moderado, como a $Q_{7,10}$. Por outro lado, uma estrutura de dependência espacial mais consoante com os dados pode melhorar as estimativas.

A predição dos MHBs estacionários nos postos fluviométricos de validação (não utilizados para calibrar os modelos) é apreciada a seguir. Primeiro, comentários são feitos acerca das estimativas pontuais dos modelos na bacia do rio Doce (Figura 5.23). No posto 52, as estimativas de quantis dos quatro modelos se assemelham fortemente às contrapartes empíricas. Além disso, não houve diferenças notáveis entre os modelos, uma vez que a introdução de informação pelos resíduos dos modelos foi limitada devido à ausência de postos vizinhos próximos. Esse resultado também indica que o impacto da correlação dos dados na predição pontual foi mínimo, apesar de alterar a distribuição *a posteriori* dos parâmetros da regressão de σ . No posto 57, sob a luz de mais informações para a predição pelos resíduos, as estimativas pontuais dos modelos M2 a M4 se ajustaram melhor aos dados empíricos. No posto 57 e nos demais postos com menos de 10 observações, maiores diferenças foram observadas entre os modelos M2 a M4, provavelmente em função da interação dos parâmetros da regressão de σ com os resíduos da regressão, uma vez que esses postos estão em regiões mais densamente monitoradas que o posto 57. A diferença entre os modelos M3 e M4 foi mínima. Em suma, as estimativas do modelo M4 foram marginalmente superiores à do modelo M3. Por último, as predições pontuais dos modelos M2 a M4 tiveram menor viés em todos os postos, com exceção dos postos 55 e 56, nos quais as estimativas do modelo M1 foram mais precisas.

Figura 5.23 - Curva de quantis dos postos de validação na bacia do rio Doce. Linhas contínuas e pontilhadas representam, respectivamente, as estimativas pontuais e o intervalo de predição com 95% de credibilidade. No cabeçalho dos painéis está código e o ID (entre parêntese) dos postos

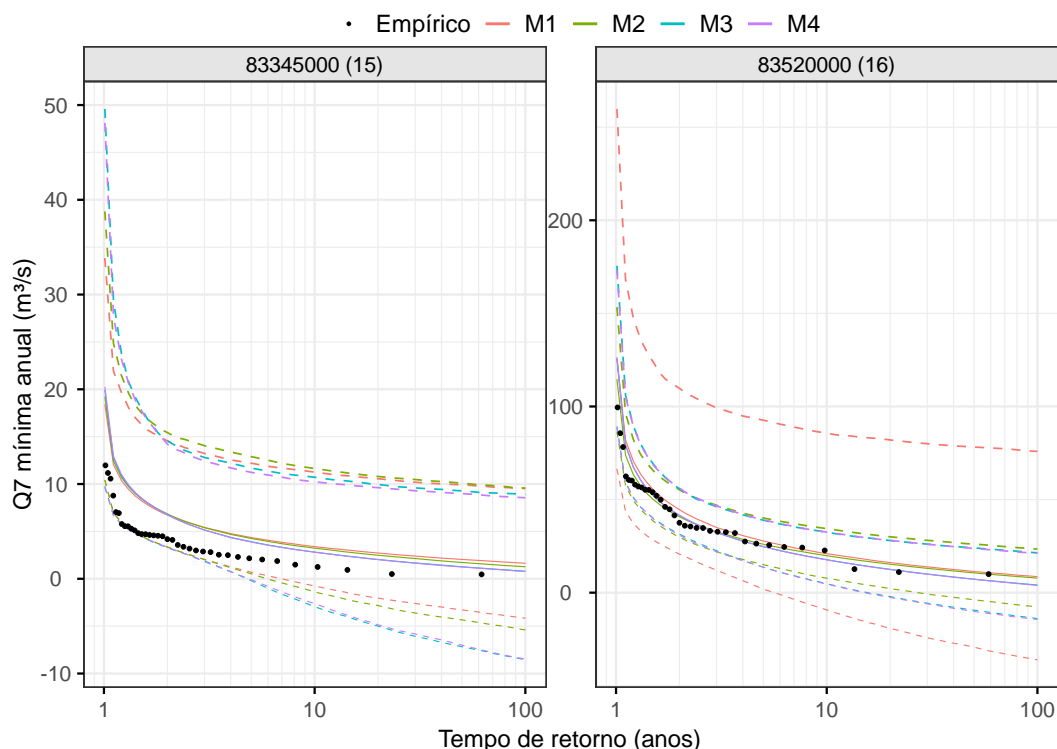


No que concerne a incerteza preditiva, observa-se que os todos os dados empíricos foram encapsulados pelo intervalo de credibilidade (IC) dos modelos, o que indica que a incerteza preditiva não é subestimada. O modelo M1, sem o benefício de utilizar os resíduos da calibração para fazer a predição nos postos de validação, resultou em intervalos de credibilidade mais largos. A grande amplitude do IC do modelo M1

representa a grande heterogeneidade espacial na bacia do rio Doce. O relaxamento da hipótese de independência dos resíduos da regressão diminuiu consideravelmente o intervalo de credibilidade nos modelos M2 a M4 nos postos com mais vizinhos próximos, como os postos 56 e 57, localizados nos rios Manhuaçu e José Pedro, respectivamente. A expectativa era de que, ao se considerar a correlação nos dados, aumentasse o IC preditivo devido à redução na quantidade efetiva de dados. De fato, observou-se aumento da incerteza em alguns dos hiper-parâmetros das regressões desses modelos em relação ao modelo M4. Entretanto, os ICs dos modelos M2 a M4 manifestaram pouca discordância. Isso se deve, provavelmente, à interação dos parâmetros com os resíduos dos modelos, que são diferentes entre si, com o cálculo da variância condicional, $\sigma^2_{k,conditional}$, na Eq. 4.13.

Na bacia do rio Itajaí-Açu, os postos utilizados para validação têm os IDs 15 e 16 e estão, respectivamente, em regiões escarpadas e densas em monitoramento. Em ambos os postos as estimativas pontuais dos modelos foram bastante similares entre si (Figura 5.24). No posto 15, a estimativa pontual dos modelos foi próxima aos quantis empíricos, enquanto no posto 16 essas quantidades foram basicamente coincidentes. Em termos de incerteza preditiva, o IC dos modelos no posto 15 foi largo em função da ausência de postos vizinhos para auxiliar na predição. Os ICs do modelo M3 e M4 foram praticamente coincidentes, porém mais largos do que os modelos M1 e M2, principalmente para quantis associados a maiores tempos de retorno.

Figura 5.24 - Curva de quantis dos postos de validação na bacia do rio Doce. Linhas contínuas e pontilhadas representam, respectivamente, as estimativas pontuais e o intervalo de predição com 95% de credibilidade. No cabeçalho dos painéis está código e o ID (entre parêntese) dos postos



O aumento da complexidade do modelo M1 para o modelo M2, considerando a dependência espacial do processo na inferência, é justificável em regiões nas quais há postos de calibração perto para fazer proveito do resíduo do ajuste nesses locais. Nessa condição, observou-se redução da incerteza preditiva e do viés das estimativas. Por outro lado, na ausência de postos de calibrações próximos, o modelo M1 parece ser mais vantajoso na medida em que o modelo M2 tende a convergir para o modelo M1. O aumento da complexidade do modelo em regiões escurças de monitoramento não é justificável pois não há acompanhamento do aumento de informação (disponibilidade de resíduos do ajuste do modelo para auxiliar a predição).

A adição, ao modelo M2, de uma componente para modelar a dependência conjunta dos dados pelos modelos M3 e M4 na bacia do rio Doce não pareceu ser uma linha de trabalho vantajosa. De fato, as predições dos modelos M3 e M4 foram bastante similares às do modelo M2, apenas com um pequeno deslocamento vertical das estimativas e do IC. Na bacia do rio Itajaí-Açu, por outro lado, os modelos M3 e M4 trouxeram resultados mais expressivos. Ambos os modelos inflaram o IC dos quantis no posto localizado em região escurça em monitoramento, indicando que ignorar a

dependência das observações pode implicar em subestimação da incerteza preditiva. Esse efeito é mais pronunciado em quantis extremos, o que sugere que essa abordagem possa ser interessante em estudos de quantis extremos, nos quais a quantificação apurada da incerteza das estimativas é vital para a tomada de decisão. Por outro lado, a alteração no IC é menos pronunciada para os menores quantis, que são de interesse para outorgas.

Entre os modelos M3 e M4, a estrutura do modelo M4 representou melhor a dependência ajustada aos dados, entretanto, os dois modelos reproduzem estimativas aproximadas. Ao se utilizar distâncias Euclidianas, parece que um modelo mais simples, como o M3, seria adequado para capturar a dependência dos dados. Estudos mais aprofundados são necessários para investigar se uma melhor descrição do meio físico, como distância em termos de rede de drenagem ou mesmo covariâncias que descrevam a hierarquia da rede de drenagem (MÜLLER; THOMPSON, 2015), reproduzam resultados diferentes. A princípio, similar a Renard (2011), chega-se à conclusão de que a estrutura de dependência espacial no nível do processo tem mais influência na predição do que aquela no nível dos dados. No entanto, ressalta-se que em regiões pouco monitoradas a dependência dos dados pode desempenhar um papel mais importante (vide posto 15 da Figura 5.24).

Por fim, optou-se pela escolha do modelo M2 para compor o modelo estacionário de referência a ser comparado com o modelo não estacionário. Essa escolha baseia-se na capacidade preditiva dos modelos estudados, a qual não melhorou a partir do aumento da complexidade dos modelos desenvolvidos a partir do modelo M2. Além disso, a introdução da não estacionariedade acarreta em um considerável maior esforço computacional. Logo, um modelo estacionário mais simples irá agilizar ou mesmo viabilizar a continuidade do estudo.

5.4 Teleconexão

A análise de componentes principais nas duas bacias hidrográficas foi utilizada para reduzir as dimensões das múltiplas séries de vazões mínimas em componentes principais (CPs), e dessa forma, simplificar a análise. As componentes principais das bacias são apresentadas na Tabela 5.5. A primeira componente das duas bacias explica uma grande porcentagem da variância dos dados – 78,7% na bacia do rio Doce e 74,9% na bacia do rio Itajaí-Açu. Devido à expressividade da variância

explicada pela primeira componente em relação às suas contrapartes, a correlação da temperatura da superfície do mar (TSM) será analisada apenas em relação a essa componente.

Tabela 5.5 - Componentes principais das séries de vazões mínimas anuais de sete dias de duração da bacia do rio Doce e Itajaí-Açu

CP – Bacia	Variância explicada (%)	Variância explicada acumulada (%)
CP 1 - Doce	78,7	78,7
CP 2 - Doce	7,2	85,9
CP 3 - Doce	4,6	90,5
CP 1 - Itajaí-Açu	74,9	74,9
CP 2 - Itajaí-Açu	13,5	88,4
CP 3 - Itajaí-Açu	6,2	94,6

A correlação entre as séries mensais de TSM e a CP 1 da bacia do rio Doce ficou restrita ao intervalo [-0,55, 0,63] (Figura 5.25). Nos meses de janeiro a julho, uma extensa mancha com correlação negativa entre a TSM e a CP 1 é observada entre a costa brasileira e a costa africana. A mancha começa em janeiro com um formato e posição similar ao mecanismo descrito por Rodrigues *et al.* (2019) e maiores valores absolutos de correlação em comparação aos demais meses. Nos meses seguintes, a mancha se desloca para o norte e os valores da correlação diminuem em intensidade.

A CP 1 nessa bacia é negativamente correlacionada às séries de vazões mínima e, conseqüentemente, o sinal da correlação entre a TSM e as séries de vazões mínimas é o inverso do sinal entre a CP 1 e a TSM. Isso indica que no mês de janeiro, a TSM na região delimitada pela mancha azul adjacente à costa brasileira é negativamente correlacionada às séries de vazões mínimas. Trata-se, possivelmente, do mecanismo descrito por Rodrigues *et al.* (2019), pois, assim como a correlação assinalada, ele ocorre no verão e associa a altas temperaturas da TSM com bloqueios atmosféricos que inibem as precipitações no sudeste do Brasil. Em função dos maiores valores absolutos de correlação em janeiro e à similaridade ao mecanismo gerador de secas descrito por Rodrigues *et al.* (2019), escolheu-se a TSM desse mês para compor o índice climático customizado para a bacia do rio Doce.

Devido à gênese do processo de geração de vazões mínimas na bacia do rio Doce, era esperado que o sinal mais forte ocorresse durante o período chuvoso (out-mar),

por causa da forte sazonalidade no regime pluviométrico. Nessa bacia, as vazões mínimas ocorrem com maior frequência na transição do período de estiagem (set-out) e são dependentes da precipitação durante o período chuvoso, que restabelecem as reservas de águas subterrâneas que irão alimentar o escoamento de base no período seco.

A correlação entre as séries mensais de TSM e a CP 1 da bacia do rio Itajaí-Açu ficou restrita ao intervalo $[-0,59, 0,53]$ (Figura 5.25). Na região delimitada pelos índices climáticos relacionados ao ENSO, a correlação da TSM e a CP 1 é mais forte nos meses de janeiro a maio, com maior intensidade no mês de abril. Esse período engloba a maior parte da estação chuvosa principal na bacia – dezembro a março, assim como está contido no período identificado por Penalba e Rivera (2016), no qual o El Niño teria um sinal mais forte com os totais pluviométricos da região na qual a bacia está inserida.

A CP 1 das vazões mínimas do rio Itajaí-Açu também é negativamente correlacionada com as vazões mínimas. Portanto, o sinal observado na região dos índices climáticos do ENSO indica uma relação positiva entre a TSM e as vazões mínimas nessa bacia. Esse resultado vai de encontro com o sinal identificado por Penalba e Rivera (2016), no qual menores TSM estão associadas à estiagem.

Nessa bacia, o mês de abril foi escolhido para ser utilizado no índice climático customizado, em função da maior intensidade do sinal de teleconexão.

Nas Figuras Figura 5.25 e Figura 5.26, o foco da análise é direcionado para os polígonos de interesse em função do suporte da literatura. Apesar disso, é notável que áreas externas aos polígonos delimitados exibiram correlação significativa com a variável em estudo, que podem representar outras associações das vazões mínimas com padrões climáticos de larga escala. O foco deste estudo, no entanto, não está em capturar sinais e investigar relações de causalidade com as vazões mínimas, portanto, a restrição da área de estudo é importante para manter um embasamento no entendimento atual dos mecanismos que afetam os processos hidrológicos.

Por fim, os valores de correlação encontrados foram moderados ($\rho > 0,5$), o que sugere que a inclusão da TSM como covariável temporal tem potencial pra melhorar as estimativas e predições do modelo não estacionário.

Figura 5.25 - Correlação de Spearman (ρ) entre a primeira componente principal das séries de vazões mínimas da bacia do rio Doce e a malha TSM no Oceano Atlântico sul. A bacia do rio Doce está simbolizada por um ponto preto. Apenas células da malha nas quais ρ foi significativamente diferente de 0 (valor- $p \leq 0,05$) são exibidas. Datum horizontal WGS 84

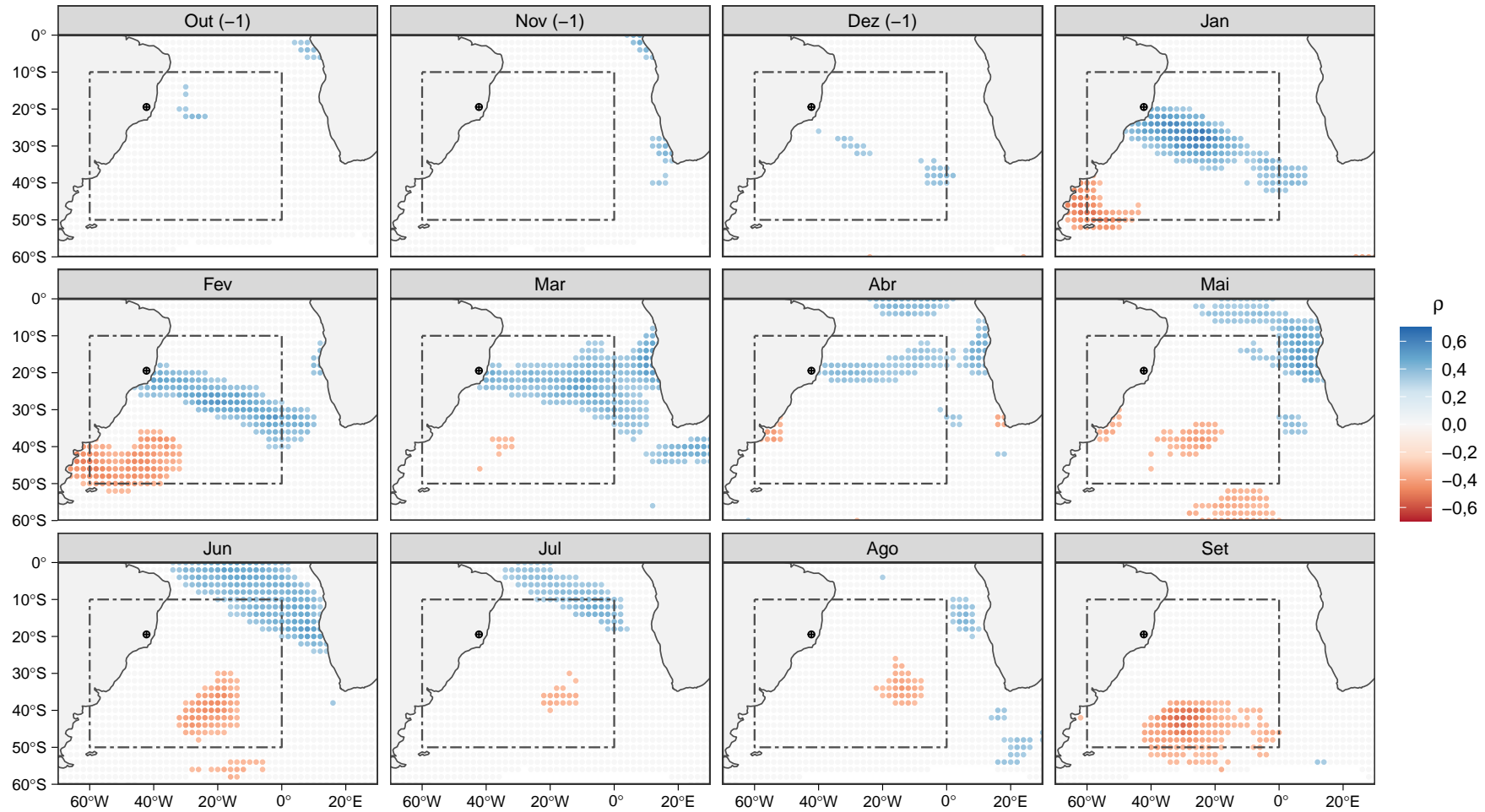
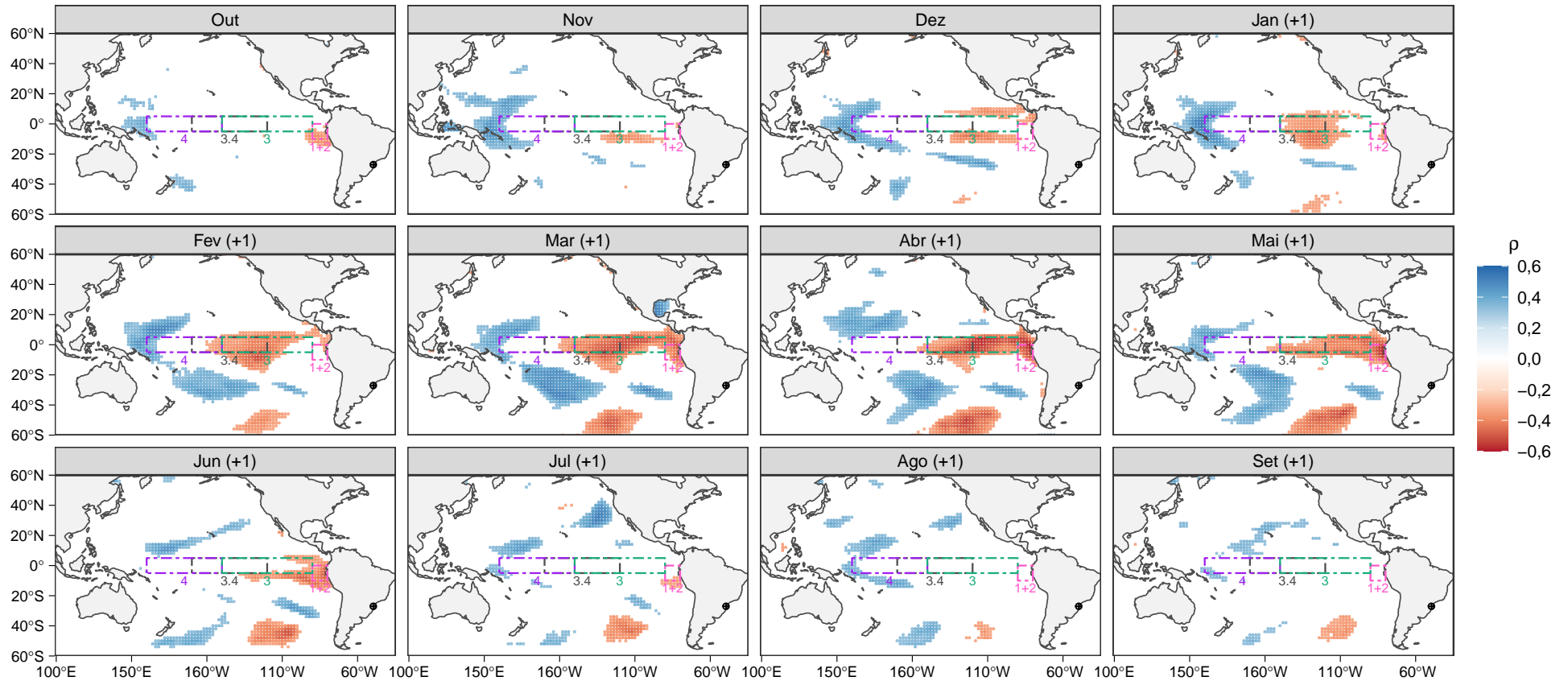


Figura 5.26 - Correlação de Spearman (ρ) entre a primeira componente principal das séries de vazões mínimas da bacia do rio Itajaí-Açu e a malha de TSM no oceano Pacífico. A bacia do rio Itajaí-Açu está simbolizada por um ponto preto. Apenas células da malha nas quais ρ foi significativamente diferente de 0 (valor-p $\leq 0,05$) são exibidas. As regiões relacionada a quatro índices do ENSO - Niño 1+2, 3, 4 e 3.4 - também é apresentada. Datum horizontal WGS 84



5.5 Modelo não estacionário

Três modelos não estacionários foram elaborados com o parâmetro de posição (μ) e escala (σ) da distribuição Gumbel variando temporalmente. A configuração desses modelos é sumarizada na Tabela 5.6.

Tabela 5.6 - Modelos não estacionários M5, M6 e M7

Modelo	Regressão dos parâmetros da distribuição Gumbel
M5	$\log(\mu[\tilde{S}_i]) = MVN(\alpha_0 + \alpha_1 \log(A[\tilde{S}_i]), \Sigma_\mu) + \gamma_\mu[\tilde{S}_i]\eta_n$ $\log(\sigma[\tilde{S}_i]) = MVN(\beta_0 + \beta_1 \log(A[\tilde{S}_i]), \Sigma_\mu)$
M6	$\log(\mu[\tilde{S}_i]) = MVN(\alpha_0 + \alpha_1 \log(A[\tilde{S}_i]), \Sigma_\mu)$ $\log(\sigma[\tilde{S}_i]) = MVN(\beta_0 + \beta_1 \log(A[\tilde{S}_i]), \Sigma_\mu) + \gamma_\sigma[\tilde{S}_i]\eta_n$
M7	$\log(\mu[\tilde{S}_i]) = MVN(\alpha_0 + \alpha_1 \log(A[\tilde{S}_i]), \Sigma_\mu) + \gamma_\mu[\tilde{S}_i]\eta_n$ $\log(\sigma[\tilde{S}_i]) = MVN(\beta_0 + \beta_1 \log(A[\tilde{S}_i]), \Sigma_\mu) + \gamma_\sigma[\tilde{S}_i]\eta_n$

A apresentação dos resultados foi segmentada por área de estudo nas seções abaixo. Primeiro, uma comparação entre os modelos não estacionários é realizada por meio do DIC e, em seguida, o melhor modelo dentre os candidatos é detalhado para mostrar suas características e comparar suas estimativas com o modelo estacionário de referência.

5.5.1 Bacia do rio Doce

O DIC calculado para os três modelos na bacia do rio Doce é exibido na Tabela 5.7, junto com o resultado já apresentado do modelo estacionário de referência, M2. Os resultados indicam que a adição de uma componente não estacionária no parâmetro de posição melhorou o desempenho do modelo. Já a adição dessa componente no parâmetro de escala manifestou menor ganho de desempenho. Esses resultados indicam, portanto, que os modelos não estacionários desenvolvidos tiveram melhor desempenho do que o modelo estacionário M2. De maneira geral, o modelo M7, que considera que ambos os parâmetros da distribuição Gumbel variam no tempo, apresentou o melhor desempenho dentre os modelos avaliados de acordo com o DIC. O modelo M7 é, portanto, selecionado dentre os modelos não estacionários candidatos.

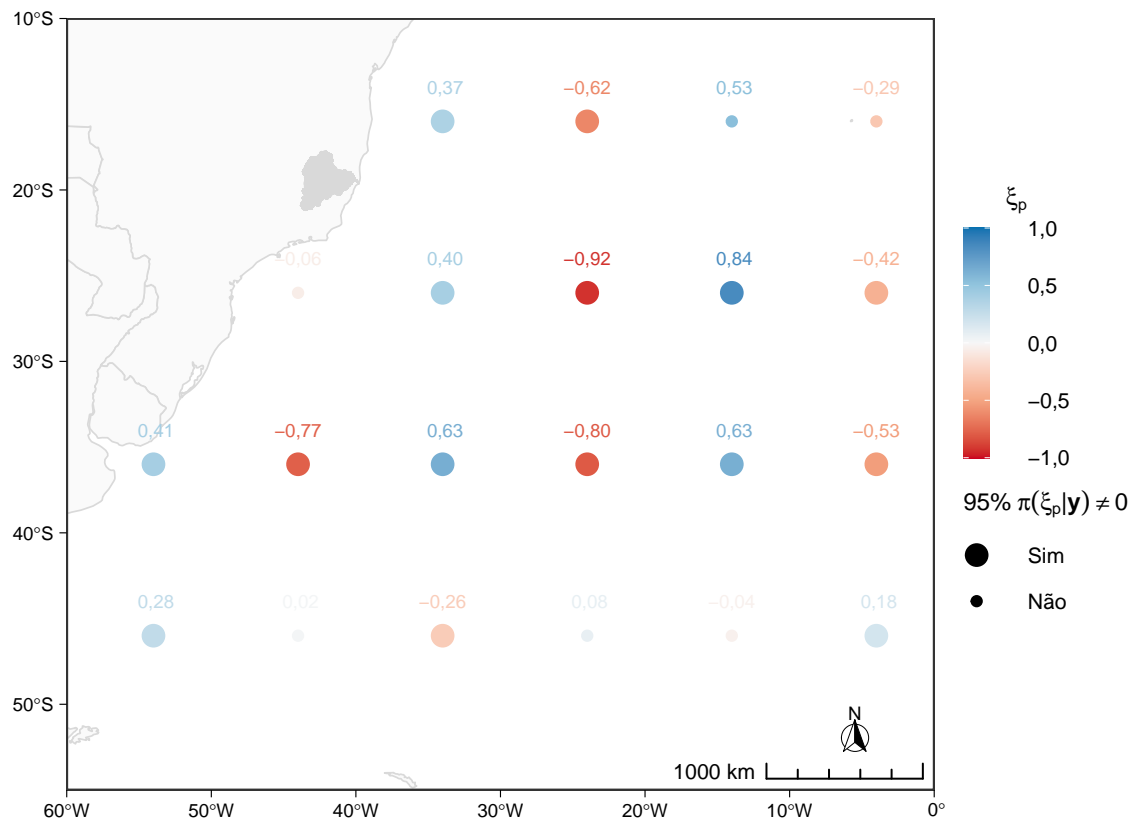
Tabela 5.7 - DIC no modelo estacionário de referência, M2, e nos modelos não estacionários, M5 a M7, na bacia do rio Doce

Modelo	Doce		
	\bar{D}	pD	DIC
M2	6787,9	82,0	6869,9
M5	6316,6	107,8	6424,4
M6	6695,1	99,2	6794,3
M7	6262,1	110,7	6372,7

Nota: o menor DIC está em negrito

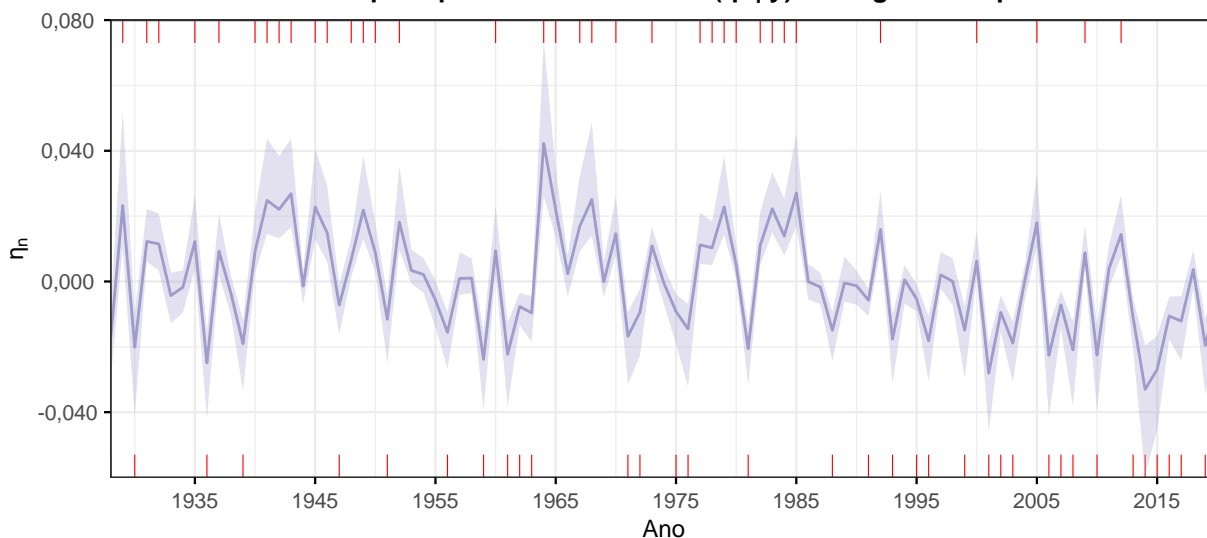
Os pesos (ξ_p) atribuídos aos componentes do índice climático inferido na bacia do rio Doce podem ser visualizados na Figura 5.27. Nota-se que os pesos assumem tanto valores negativos quanto positivos em toda a área de estudo, mas que os dados de TSM próximos à latitude de 30°S colaboram mais intensamente com a definição do índice climático customizado (i.e., possuem pesos maiores). Além disso, observa-se que as estimativas de pesos em uma mesma longitude tenderam a compartilhar o mesmo sinal (positivo ou negativo) e que houve uma intercalação nos sinais dos pesos de longitudes diferentes, isso é, pesos em uma mesma longitude tenderam a ter sinais opostos aos pesos de longitudes adjacentes. Acredita-se que os padrões espaciais observados estejam relacionados à correlação espacial da TSM.

Figura 5.27 - Mediana dos pesos (ξ_p) atribuídos à malha de TSM em escala reduzida ($10^\circ \times 10^\circ$) no modelo M7 aplicado à bacia do rio Doce. Datum horizontal WGS 84



Em seguida, a variação anual do índice climático customizado (η_n) é apresentada na Figura 5.28. O índice varia temporalmente com sinais positivos e negativos, indicando um comportamento complexo e sem aparente persistência temporal. Esse comportamento seria dificilmente captável por meio da abordagem tradicional sob o uso de regressões polinomiais utilizando o tempo como variável independente. Nessa abordagem, a componente não estacionária exibe forte dependência temporal (usualmente descrita por meio de regressões de primeiro grau, na qual a dependência temporal se propaga por meio de tendências monotônicas), que pode não descrever a alta variabilidade interanual modelada.

Figura 5.28 - Índice climático customizado (η_n) para a bacia do rio Doce. Traços vermelhos indicam anos nos quais pelo menos 95% de $\pi(\eta_n | y)$ foi negativa ou positiva



À primeira vista, a variação da estimativa pontual de η_n parece pequena. No entanto, deve-se ater ao fato de que esse índice é multiplicado por um coeficiente espacial e que o índice está relacionado ao logaritmo de μ e σ . Logo, a variabilidade de η_n pode impor mudanças razoáveis nas estimativas dos parâmetros da distribuição Gumbel e, conseqüentemente, nos quantis estimados conforme discutido adiante.

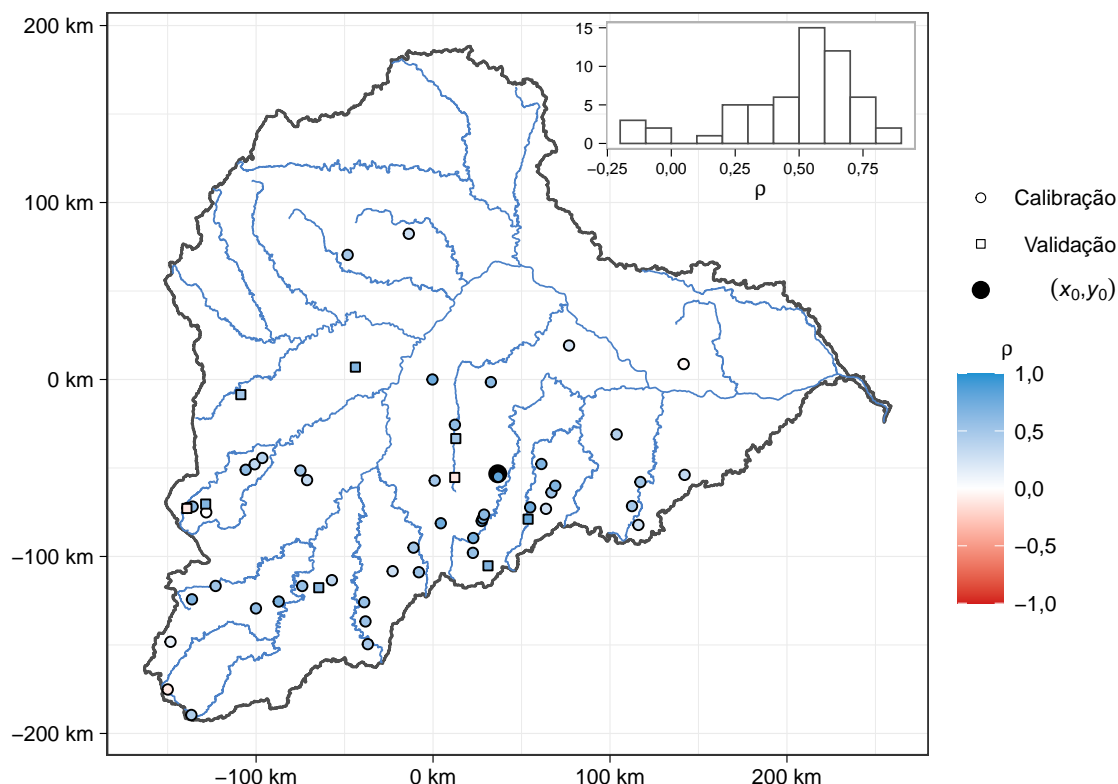
Nota-se que 2014 e 2015, início do longo período de seca na bacia do rio Doce, apresentaram, respectivamente, o 1° e 3° menor valor de η_n em toda a série do índice η_n , e que a estimativa desse parâmetro teve sinal negativo praticamente em todo o período de 2013 a 2020. O índice η_n , cuja inferência é exclusivamente baseada no ajuste aos dados antecedentes a 2014, parece refletir de maneira coerente o período de seca citado. Esse resultado, portanto, reforça a conjectura de que o índice climático proposto explica parte da variabilidade temporal das vazões mínimas na bacia do rio Doce.

A existência e a causa de períodos ricos em secas/enchentes é um dos temas na hidrologia que carecem de informações (BLÖSCHL *et al.*, 2019), e que usualmente é investigado com dados extensos temporalmente ou espacialmente (BLÖSCHL *et al.*, 2020; LUN, 2020). Apesar de não ser o foco deste estudo, o índice η_n transpassa uma percepção sobre o agrupamento temporal de períodos propensos a secas (maior probabilidade atribuída a quantis da cauda inferior da distribuição Gumbel, η_n negativo), como de períodos menos propensos a esse fenômeno (η_n positivo) na bacia

de estudo. No período de 1986 a 2020, por exemplo, houve 19 (5) anos em que pelo menos 95% da massa da distribuição a posteriori do índice η_n esteve em valores negativos (positivos). Esses 35 anos, portanto foram preponderantemente propensos à ocorrência de secas. Os anos 1985 e 1986 parecem dividir a série em um período menos propenso à secas e outro mais propenso. Esse resultado sugere que um agrupamento nas realizações do processo poderia ser caracterizado pela variabilidade climática, representada pelo índice η_n .

O posicionamento da coordenada (\hat{x}_0, \hat{y}_0) na bacia do rio Doce é exibido na Figura 5.29, junto com as correlações de Spearman (ρ) entre as observações dos postos fluviométricos e o índice η_n . O centro da teleconexão com o índice customizado está localizado à sudeste do centroide da bacia do rio Doce e próximo às bacias dos rios Caratinga, Manhuaçu e José Pedro, região na qual se observa correlações mais fortes entre as observações de Q_7 e η_n . Também é possível aferir pela Figura 5.29 que há indícios de heterogeneidade na resposta ao índice climático. A maioria dos postos manifestam correlação moderada a forte com o índice (e.g. $|\rho| > 0,5$), no entanto, uma parcela razoável de postos possuem correlação fraca ou desprezível.

Figura 5.29 - Localização dos centroides dos postos de validação e calibração e a mediana do par de coordenadas inferidas (\hat{x}_0, \hat{y}_0) do sinal do índice climático customizado. A cor dos postos indica a correlação entre as observações e o índice climático customizado. Eixo das coordenadas com origem no centroide da bacia do rio Doce.



A dependência espacial do índice climático nos parâmetros da distribuição Gumbel é retratada a seguir. As distribuições *a posteriori* e *a priori* dos parâmetros da função de decaimento do coeficiente do índice η_n na regressão do parâmetro μ são apresentada na Figura 5.30b e Figura 5.30c. Nota-se que esses parâmetros foram identificados e que as distribuições *a priori* assinaladas a eles são pouco informativas. O coeficiente espacial do índice climático resultante, γ_μ , é exibido em termos de sua mediana na Figura 5.30a. Nessa figura, $\gamma_\mu \in [4,9, 13,9]$ e seu valor decai a partir do par de coordenadas (\hat{x}_0, \hat{y}_0) exibidas na Figura 5.29. A grande amplitude do intervalo de γ_μ é reflexo da heterogeneidade na resposta ao índice climático discutida anteriormente. Locais próximos às coordenadas (\hat{x}_0, \hat{y}_0) possuem um peso cerca de três vezes maior atribuído ao índice η_n do que locais próximos à foz do rio Doce e à vertente noroeste da bacia, por exemplo. Em respeito à dependência espacial do índice climático relativa ao parâmetro σ , algumas diferenças são observadas: (i) o decaimento é menor, ou seja, a variação espacial de γ_σ é menor e o sinal do índice η_n é mais homogêneo na

bacia (ii) o sinal de teleconexão tem um peso maior no parâmetro σ do que no parâmetro μ .

Figura 5.30 - Dependência espacial da influência do índice climático customizado no parâmetro μ da bacia do rio Doce. Mapa com a mediana de γ_μ (a) e as distribuições a priori e a posteriori do decaimento (b) e do sill (c) da função de decaimento de γ_μ .

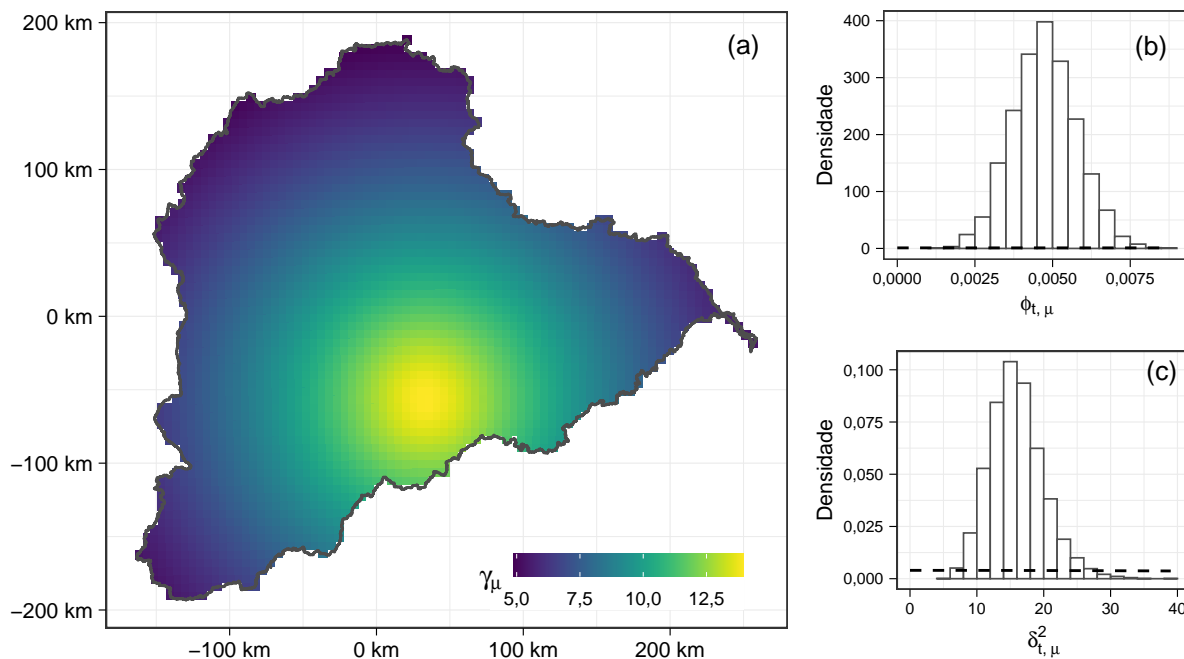
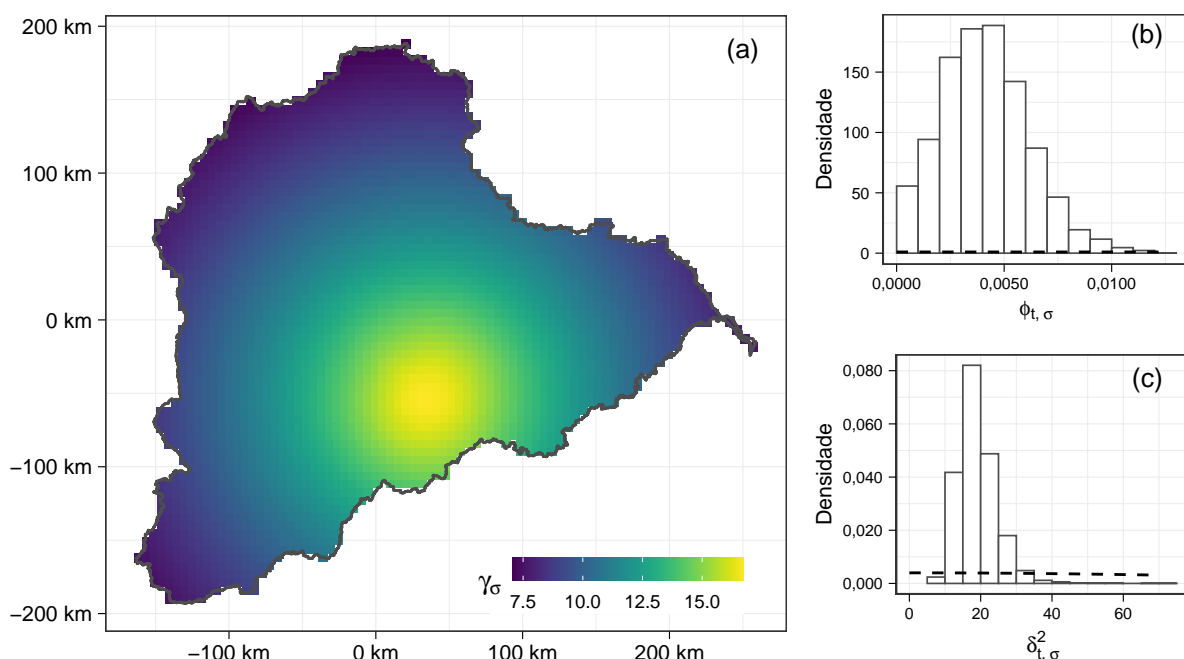


Figura 5.31 - Dependência espacial da influência do índice climático customizado no parâmetro σ da bacia do rio Doce. Mapa com a mediana de γ_σ (a) e as distribuições a priori e a posteriori do decaimento (b) e do sill (c) da função de decaimento de γ_σ .

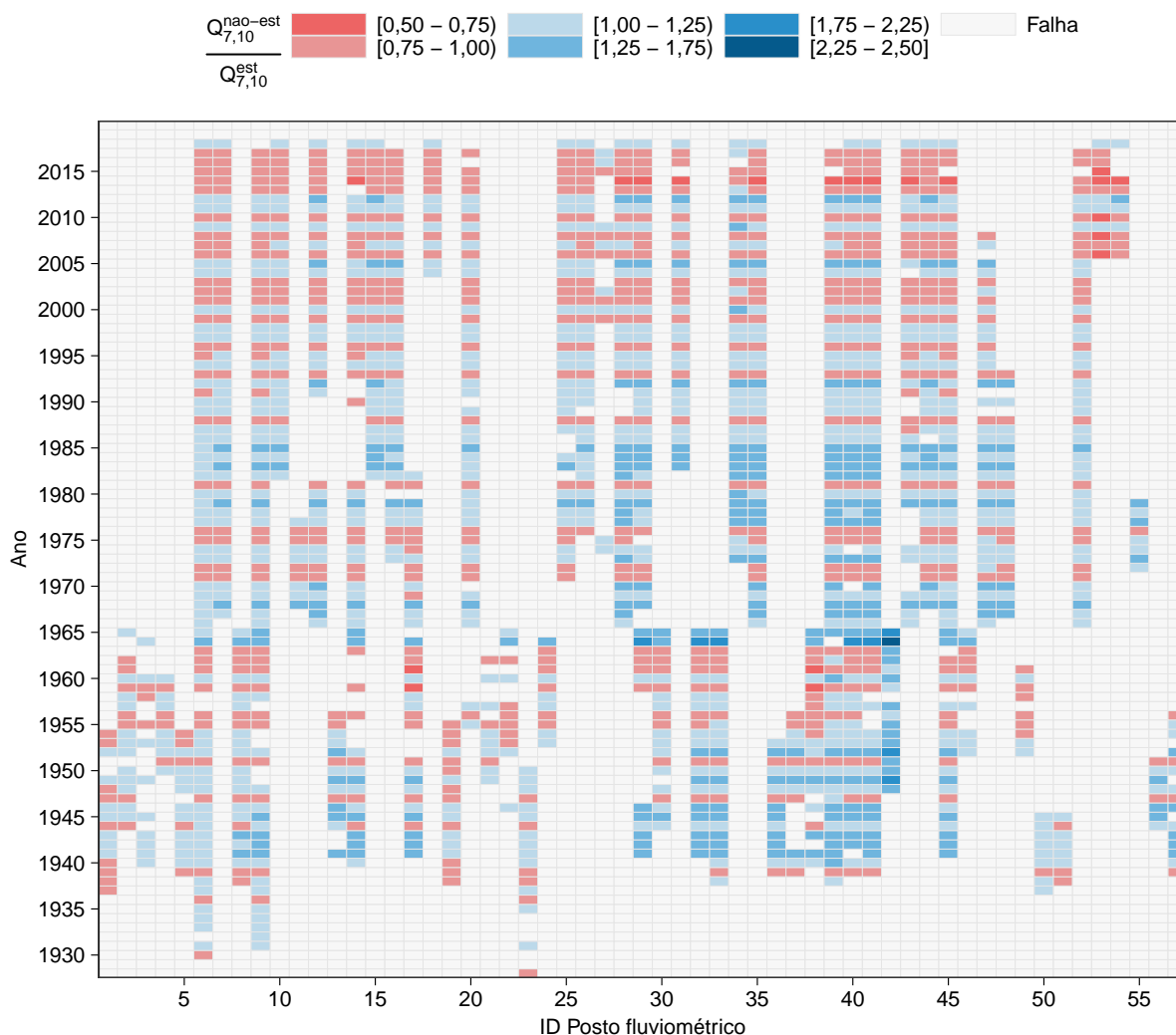


Por fim, a $Q_{7,10}$ dos modelos M2 e M7 são comparadas para averiguar a influência da componente não estacionária nas estimativas desse quantil. A Figura 5.32 exibe a razão entre as estimativas da $Q_{7,10}$ entre os modelos M7 e M2, a qual segue as paletas de cor vermelha e azul em anos que a $Q_{7,10}$ do modelo estacionária foi, respectivamente, superior e inferior à estimativa do modelo não estacionário. Primeiro, nota-se que a razão da $Q_{7,10}$ foi inferior a um durante 2013 a 2017, indicando que o modelo M7, alimentado pelas informações climáticas que compuseram o índice η_n , associou probabilidades maiores de não excedência às vazões mínimas em relação ao modelo estacionário. Segundo, as razões de $Q_{7,10}$ encontradas variaram de 0,63 a 2,33, indicando que a componente climática introduzida no modelo M7, de fato, influencia substancialmente a estimação de quantis, mesmo para TRs não elevados, como o de 10 anos.

O modelo M7 teve melhor desempenho em relação ao modelo estacionário M2 a partir do DIC, indicando que ele potencialmente deve fornecer uma estimativa mais precisa da probabilidade de não excedência das vazões mínimas. Ou seja, considerar que o processo não é estacionário e que parte da variabilidade anual das vazões mínimas está condicionada à covariável temporal utilizada parece refletir a variável hidrológica de maneira mais consistente do que a abordagem estacionária.

A componente não estacionária do processo altera substancialmente os quantis de vazões mínimas. Esse fato traz implicações importantes para a gestão de recursos hídricos. A hipótese de que o processo é estacionário e que os fatores que controlam o processo, como o clima, não variam temporalmente pode não ser realística. Na verdade, os resultados indicam que a variabilidade climática controla uma parte substancial do processo, como a magnitude das vazões e potencialmente o agrupamento de períodos ricos em secas. Segundo, em função dos valores observados de TSM, valores extremos de vazões mínimas possuem maior probabilidade de ocorrência do que na abordagem estacionária (e.g., 2013 a 2017). Uma vez que a covariável temporal utilizada é observada em janeiro, meses antes da ocorrência das vazões mínimas nessa bacia (em geral em setembro e outubro), o índice climático desenvolvido tem potencial para ser utilizado por gestores de recursos hídricos em um cenário de previsão.

Figura 5.32 - Quociente entre as estimativas pontuais da $Q_{7,10}$ dos modelos não estacionário (M7) e estacionário (M2) nos anos com observação. Estimativas dos postos 49 a 57 são obtidas a partir das predições dos modelos. Os anos de 2014 a 2019 não foram utilizados na inferência.



5.5.2 Bacia do rio Itajaí-Açu

O DIC calculado para os três modelos na bacia do rio Itajaí-Açu é exibido na Tabela 5.8, junto com o resultado já apresentado do modelo estacionário de referência, M2. De maneira similar aos resultados encontrados na bacia do rio Doce, os resultados nessa bacia indicam que a adição de uma componente não estacionária no parâmetro de posição melhorou muito o desempenho do modelo. Já a adição dessa componente no parâmetro de escala manifestou menor ganho de desempenho. Os modelos não estacionários desenvolvidos tiveram melhor desempenho do que o modelo estacionário M2 de acordo com o DIC. O DIC do modelo M5 foi o menor obtido, entretanto, sua pequena diferença com o DIC do modelo M7 os torna indistinguíveis

em termos de desempenho. Opta-se por utilizar o modelo M5 por ser mais parcimonioso.

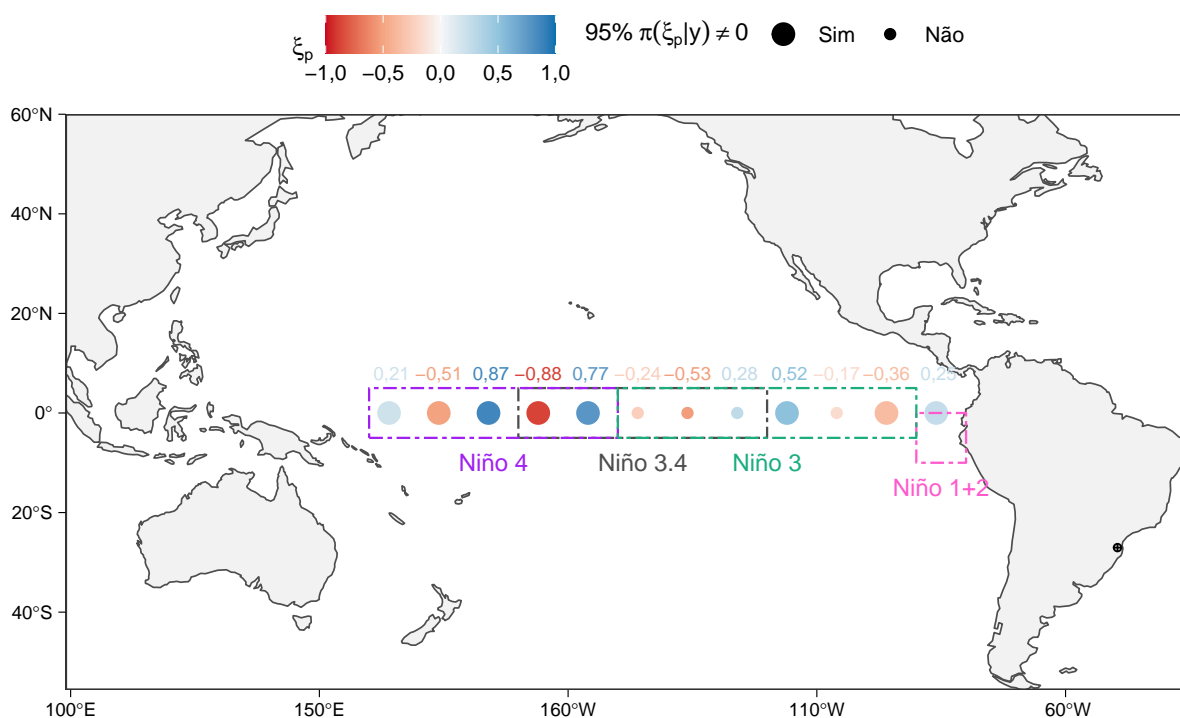
Tabela 5.8 - DIC no modelo estacionário de referência, M2, e nos modelos não estacionários, M5 a M7, na bacia do rio Itajaí-Açu

Modelo	Itajaí-Açu		
	\bar{D}	pD	DIC
M2	4166,7	25,5	4192,3
M5	3973,5	37,6	4011,1
M6	4132,8	33,0	4165,8
M7	3972,8	38,5	4011,2

Nota: o menor DIC está em negrito

Os pesos (ξ_p) atribuídos aos componentes do índice climático inferido na bacia do rio Itajaí-Açu podem ser visualizados na Figura 5.33. Pesos maiores foram atribuídos ao oeste da área investigada, associada à região do índice climático Niño 4. Embora haja pesos com sinais opostos, o sinal negativo é predominante na região desse índice. Na região leste dos índices climáticos do ENSO (e.g. Niño 3 e Niño 1+2), na qual Penalba e Rivera (2016) identificaram que as déficits de precipitação ocorriam simultaneamente com anomalias negativas de TSM, não se reconheceu um padrão nos pesos ξ_p que remetesse aos resultados dos autores citados. O mecanismo de pesos atribuídos às diferentes áreas de atuação do ENSO pode ter capturado uma relação mais complexa entre a TSM dessa região e o regime de estiagem. Além disso, anomalias em dados de TSM, utilizados no estudo de Penalba e Rivera (2016), dependem do período utilizado como referência para estabelecer o que é uma anomalia e, portanto, a comparação não é direta.

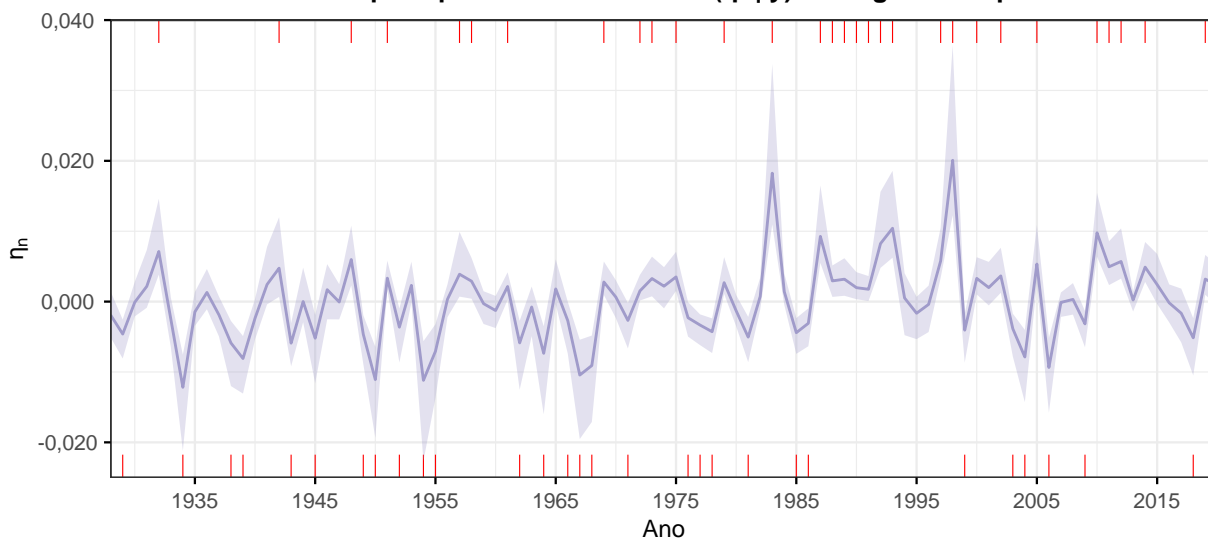
Figura 5.33 - Mediana dos pesos (ξ_p) atribuídos à malha de TSM em escala reduzida ($10^\circ \times 10^\circ$) no modelo M5 aplicado à bacia do rio Itajaí-Açu. Datum horizontal WGS 84



A descrição temporal do índice climático customizado (η_n) inferido na bacia do rio Itajaí-Açu é apresentada na Figura 5.34. As estimativas do índice η_n foram menores, em valores absolutos, do que as estimativas na bacia do rio Doce, o que indica que a variabilidade climática nessa bacia influencia menos o processo. A série temporal do índice η_n também apresentou alta variabilidade interanual. Os picos de 1983 e 1998 coincidiram com anos com El Niño muito forte de acordo com o índice ONI¹⁷ (medido na região do Niño 3.4), sugerindo que as vazões mínimas são maiores em anos em que a TSM do oceano Pacífico Tropical são mais elevadas. Quanto aos vales do índice, relacionados a ocorrência de vazões menores, não se identificou um padrão com o índice ONI.

¹⁷ Disponível em : <https://ggweather.com/enso/oni.htm>

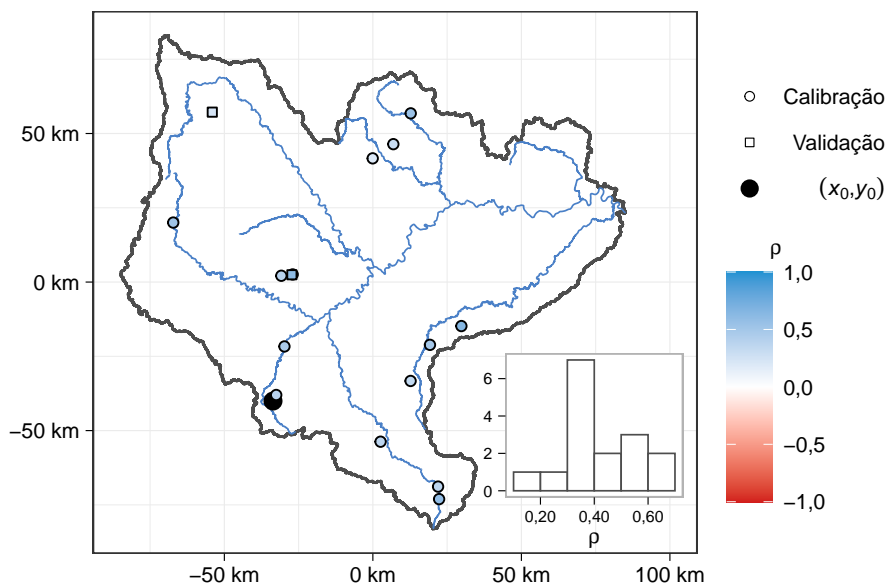
Figura 5.34 - Índice climático customizado (η_n) para a bacia do rio Itajaí-Açu. Traços vermelhos indicam anos nos quais pelo menos 95% de $\pi(\eta_n | y)$ foi negativa ou positiva



A manifestação de períodos mais ou propensos à secas também é observada na bacia do rio Itajaí-Açu. Entre 1987 e 2020, por exemplo, o índice manteve valores positivos (negativos), para 95% de significância, em 17 (5) anos.

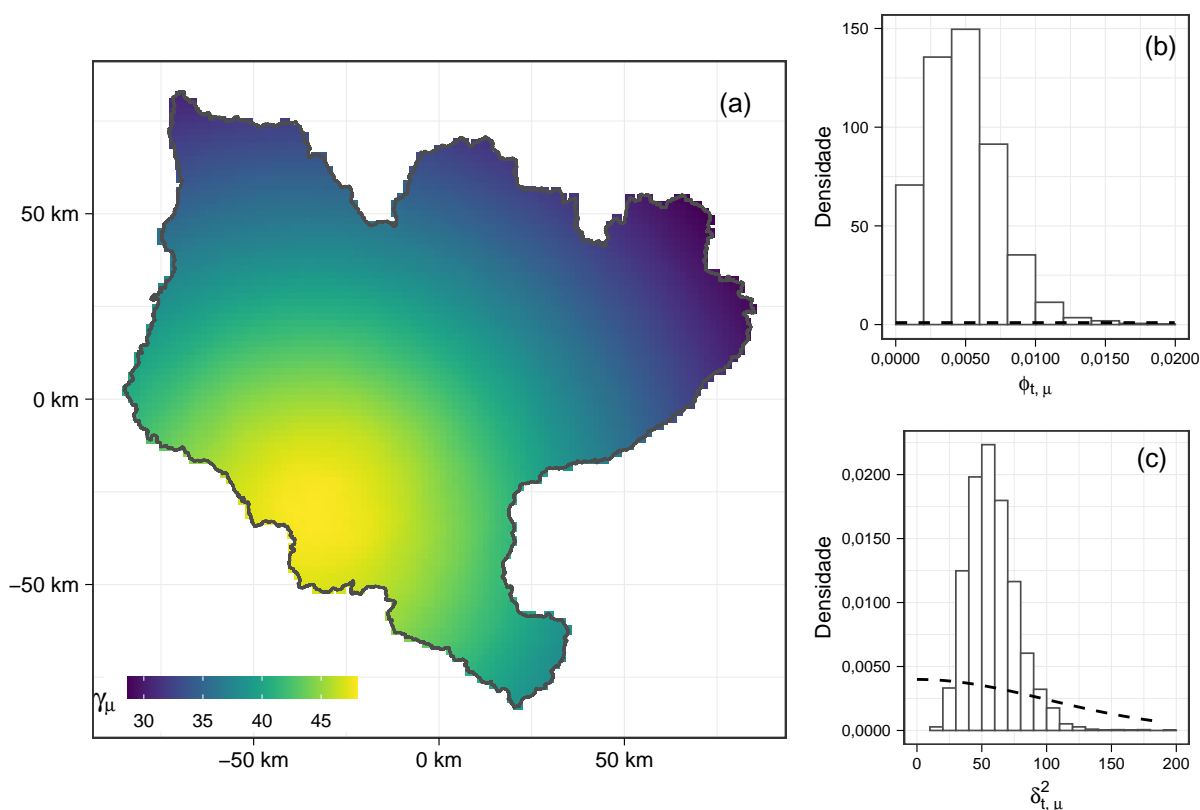
O posicionamento da coordenada (\hat{x}_0, \hat{y}_0) na bacia do rio Itajaí-Açu é exibido na Figura 5.35, Figura 5.29 junto com as correlações de Spearman (ρ) entre as observações dos postos fluviométricos e o índice η_n . O centro da teleconexão com o índice customizado está localizado no extremo oeste da bacia do rio Itajaí-Açu, próximo à bacia do rio Trombudo. Nessa bacia, as correlações entre as observações de Q_7 e o índice η_n foram menores do que na bacia do rio Doce, com a maioria dos postos apresentando correlação fraca ou desprezível e apenas 5 postos com correlação acima de 0,5. Há indícios, portanto, de que o índice climático inferido explique uma menor variabilidade do processo em comparação à bacia do rio Doce.

Figura 5.35 - Localização da mediana do par de coordenadas inferidas (\hat{x}_0, \hat{y}_0) do sinal do índice climático customizado. A cor dos postos indica a correlação entre as observações e o índice climático customizado. Eixo das coordenadas com origem no centroide da bacia



A descrição espacial do índice climático customizado é resumida pela Figura 5.36. Os resultados indicam que o índice se manifesta mais fortemente no extremo oeste da bacia e que sua variação espacial não é desprezível, alterando em até cerca em 40% do valor máximo estimado. A bacia do rio Itajaí-Açu é de grande porte, mas ainda assim, é cerca de 5,5 vezes menor do que a bacia do rio Doce. Provavelmente em função de seu menor tamanho, a variabilidade espacial do índice η_n no parâmetro de posição da distribuição Gumbel foi menor em comparação com a bacia do rio Doce. Intuitivamente, áreas maiores são mais suscetíveis à heterogeneidade física e climática.

Figura 5.36 - Dependência espacial da influência do índice climático customizado na bacia do rio Itajaí-Açu. Mapa com a mediana de γ_μ (a) e as distribuições a priori e a posteriori do decaimento (b) e do sill (c) da função de decaimento de γ_μ .



Por fim, a comparação entre a $Q_{7,10}$ pontual estimada pelos modelos M2 e M7 é realizada na Figura 5.37. A razão da $Q_{7,10}$ na área de estudo variou de 0,32 a 4,00, indicando que o índice climático customizado altera em grande extensão a estimação de quantis no modelo não estacionário. Nos postos de calibração (ID de 1 a 14), a razão da $Q_{7,10}$ variou anualmente oscilando entre anos nos quais o modelo estacionário estimou vazões maiores para o tempo de retorno de 10 anos e anos em que suas estimativas foram menores do que o modelo não estacionário. Nos postos de validação, por outro lado, nota-se que a razão foi superior a um em todos os anos com observações, variando aproximadamente de 1,0 a 1,2. Isso se deve à alteração de parâmetros relacionados à componente estacionária, cujas estimativas foram diferentes em relação ao modelo estacionário. Essas diferenças são ilustradas na Figura 5.38, na qual observa-se que os hiper-parâmetros relacionados ao parâmetro μ (no qual está inserido a componente não estacionária) variaram mais entre os modelos. Esse resultado implica que um viés nas predições pode surgir ao se ignorar a variabilidade climática.

Figura 5.37 - Quociente entre as estimativas pontuais da Q_{7,10} dos modelos não estacionário (M5) e estacionário (M2) nos anos com observação. Estimativas dos postos 15 e 16 são obtidas a partir das predições dos modelos. O ano indicado no gráfico corresponde ao ano das observações de out-dez.

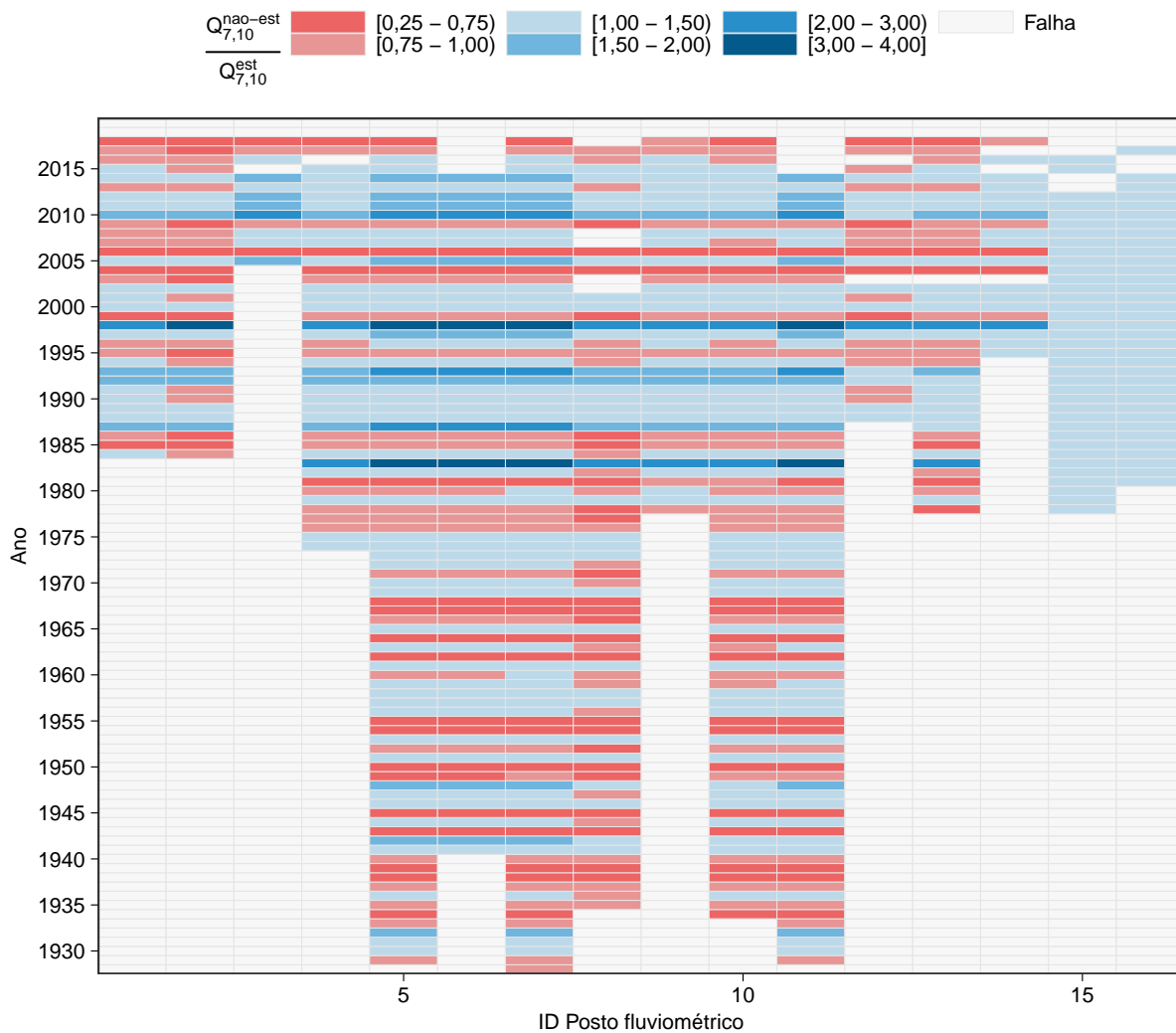
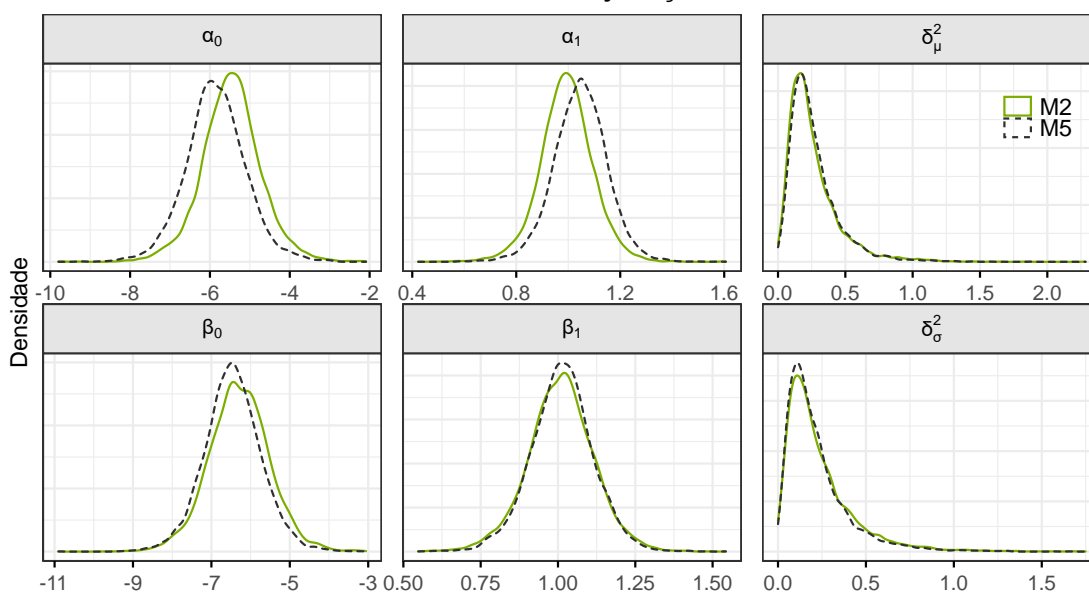


Figura 5.38 - Distribuição a posteriori dos parâmetros das regressões dos modelos M2 e M5 na bacia do rio Itajaí-Açu



6 CONCLUSÕES

Esta dissertação teve como objetivo investigar aspectos espaciais em modelos regionais de análise de frequência de vazão mínima e avaliar o desempenho de um modelo não estacionário. A linha de trabalho desenvolvida foi baseada no paradigma Bayesiano, devido a sua facilidade em incorporar estruturas flexíveis para modelar aspectos de dependência espacial e temporal, além de quantificar naturalmente a incerteza da inferência.

Em retrospecto aos objetivos específicos, foram identificadas covariáveis espaciais e temporais que auxiliem a explicar o processo estocástico em análise; avaliou-se os efeitos da incorporação da dependência espacial na estimação dos parâmetros e na incerteza preditiva; avaliou-se a influência da interdependência espacial entre as observações de vazão mínima na inferência estatística; e avaliou-se o desempenho de um modelo não estacionário que considere a variabilidade climática. Um breve resumo do trabalho é descrito a seguir, seguido por considerações finais e conclusões obtidas.

Neste estudo, duas bacias hidrográficas – bacia do rio Doce e Itajaí-Açu – foram utilizadas para o ajuste dos Modelos Hierárquicos Bayesianos (MHBs) elaborados. A escolha dessas bacias baseou-se na disponibilidade prévia de dados e na existência de relatos na literatura de teleconexões com variações climáticas de larga escala.

Na concepção dos MHB de regionalização, primeiramente foi necessário definir a distribuição de probabilidade dos dados. Essa etapa do trabalho foi realizada com auxílio de métricas de ajuste, teste de aderência e uma inspeção visual no diagrama de quocientes de momentos-L. Em seguida, por meio de modelos aninhados, avaliou-se a construção de modelos que considerassem (i) a dependência espacial manifestada pelos resíduos da regressão dos parâmetros da distribuição de probabilidade, e (ii) a dependência espacial entre observações concomitantes sob diferentes estruturas de dependência dos dados.

Finalmente, a teleconexão entre a TSM as vazões mínimas foi investigada para a seleção do mês com sinal mais forte. Posteriormente, elaborou-se modelos não estacionários que inferiram um índice climático customizado a partir das TSM, o qual foi introduzido nos parâmetros de probabilidade da distribuição Gumbel. O índice foi permitido variar espacialmente, flexibilizando o modelo para a existência de heterogeneidade na resposta ao índice climático. O desempenho dos modelos não estacionários foi avaliado por meio do DIC.

Em princípio, os MHBs estacionários desenvolvidos neste estudo podem ser empregados em qualquer bacia. Por outro lado, os modelos não estacionários desenvolvidos não podem ser estendidos para outras bacias sem devidas alterações. Ao longo deste estudo, defendeu-se a ideia de que caso se introduza uma componente não estacionária na regionalização, ela deve ser devidamente embasada, caso contrário, o modelo gerado é especulativo ou pode não representar adequadamente a física por trás do mecanismo gerador das vazões mínimas. Por esse motivo, a covariável temporal utilizada neste trabalho (temperatura da superfície do mar) foi cuidadosamente selecionada para refletir sinais identificados na literatura das bacias do rio Doce e Itajaí-Açu.

Nesse trabalho, a dependência serial das vazões mínimas (memória do processo) foi constatada em postos fluviométricos das duas bacias hidrográficas estudadas. Na bacia do rio Doce, a porcentagem de postos com dependência serial foi maior do que na bacia do rio Itajaí-Açu, potencialmente em função de um evento de seca extrema plurianual que se instalou na bacia nos últimos anos. A estratégia adotada foi de descartar essas estações para facilitar o alcance dos objetivos visados, e no caso da bacia do rio Doce, descartar também o período de seca extrema citado. Essa estratégia não pode ser generalizada e foi adotada exclusivamente em função dos objetivos listados.

Em face à metodologia empregada neste estudo, as principais conclusões obtidas nesse estudo foram:

- Na identificação das distribuições candidatas para as vazões mínimas, identificou-se que a assimetria dessa variável é majoritariamente positiva dentro do conjunto de dados avaliados. Naturalmente, distribuições com assimetria positiva se ajustaram melhor aos dados. As distribuições Gumbel com assimetria positiva, Gama e Lognormal tiveram melhores ajustes dentre as distribuições de dois parâmetros;
- A relação de escalabilidade com a área de drenagem com os parâmetros da distribuição Gumbel foi utilizada para fazer previsões em locais não monitorados, compondo a base da variação espacial dos modelos estacionários desenvolvidos. Essa relação foi identificada para o parâmetro de posição da distribuição LN2, assim como o parâmetro de escala da distribuição Gama.
- Na concepção dos modelos hierárquicos estacionários, identificou-se que em regiões pouco monitoradas o modelo M1, sob a hipótese de independência dos resíduos da regressão, é preferível em função da subutilização dos resíduos dos postos de calibração. Nessas regiões, o aumento da complexidade do modelo não é acompanhado por um aumento na quantidade de dados (neste caso, os resíduos). Contudo, em regiões mais densamente monitoradas, o relaxamento dessa hipótese (modelo M2) conjunto ao estabelecimento de uma estrutura espacial de dependência dos resíduos trouxe benefícios para a acurácia e precisão das previsões.
- A hipótese de dependência dos dados foi relaxada a partir do ajuste da distribuição conjunta de observações concomitantes por meio de cópulas elípticas Gaussianas (modelos M3 e M4). Essa abordagem, no entanto, apresentou pouca diferença em comparação ao modelo M2 em termos de previsão. Esse resultado sugere que ao menos nas bacias abrangidas por este estudo, ignorar a dependência dos dados, assim como no *index flood* e em outras abordagens de regionalização de vazões mínimas, não traz malefícios para a previsão de quantis de vazões mínimas. Por outro lado, o ajuste nos postos de calibração sofreu mais influência da estrutura de dependência dos dados. Na bacia do rio Doce, o modelo M4 teve melhor desempenho entre os 4 modelos avaliados, enquanto o modelo M3 teve o pior. Na bacia do rio Itajaí-Açu, o modelo M3 também teve o pior desempenho, enquanto que o modelo M2 teve o melhor desempenho. Estudos em mais bacias seriam necessários para chegar a uma conclusão assertiva se os modelos M3 e M4 tendem a gerar estimativas

melhores nos postos de calibração. Entretanto, o objetivo principal da AFR de vazões mínimas reside na predição em locais não monitorados.

- Com relação às funções de dependência espacial dos dados empregadas nos modelos M3 e M4, modelos exponencial simples e exponencial potência, elas são mais influentes nas estimativas dos postos de calibração. No entanto, a predição por meio de ambas é pouco diferente. O modelo M4 representou mais adequadamente a estrutura de dependência dos dados, a qual se manifesta intensamente para curtas distâncias e possui um decaimento lento, indicando que a extensão espacial da dependência dos dados é grande.
- Sinais de teleconexão entre vazões mínimas e TSM foram identificados nas duas bacias de estudo. O sinal mais forte identificado na bacia do rio Doce ocorreu em janeiro, próximo à costa brasileira, possivelmente o mesmo sinal identificado por Rodrigues *et al.* (2019). Na bacia do rio Itajaí, sinais na área dos índices climáticos relacionados ao ENSO ocorreram mais fortemente nos meses de março a maio, com maiores correlações observadas em abril.
- O DIC foi calculado para três modelos não estacionários (i) parâmetro de posição variando temporalmente, M5, (ii) parâmetro de escala variando temporalmente, M6, (iii) ambos os parâmetros variando temporalmente e comparado com o modelo estacionário de referência, M2. Os resultados indicaram que a introdução do índice climático customizado na inferência melhorou o desempenho dos modelos. Na bacia do rio Doce o melhor modelo foi o M7 e na bacia do rio Itajaí-Açu foi o M5.
- O índice climático customizado inferido em ambas bacias de estudo manifestou características não estacionárias, como tendências monotônicas e pontos de mudança. Essas características não foram avaliadas nas séries de vazões mínimas, mas caso sejam sugeridas por testes não paramétricos, elas podem estar relacionadas à variabilidade climática modelada pelo índice customizado inferido neste estudo. Os índices também exibiram agrupamentos temporais no processo, nos quais períodos extensos foram mais/menos propensos à secas. Embora esses períodos tenham sido identificados nas séries temporais do índice climático customizado, a caracterização e o entendimento físico por trás da geração desses períodos carecem de mais investigações.

- Os maiores picos no índice climático na bacia do rio Itajaí-Açu ocorrem simultaneamente com eventos fortes de El Niño, o que sugere que o El Niño possa estar relacionado à ocorrência de maiores vazões nessa bacia e que secas em anos como esses são menos prováveis.
- Na bacia do rio Doce, a coordenada na qual o índice climático incide maior influência ficou localizada próximas às bacias dos rio Manhuaçu, José Pedro e Caratinga, na porção sudeste da bacia do rio Doce. Na bacia do rio Itajaí-Açu, essa coordenada ficou localizada no extremo oeste da bacia. Em ambas as bacias, o índice variou espacialmente, exibindo heterogeneidade espacial nas respostas, embora em menor grau na bacia do rio Itajaí-Açu, possivelmente em função de seu tamanho.
- Nas duas bacias de estudo, uma extensa parte da variabilidade do processo é explicada pelo índice climático e, portanto, a $Q_{7,10}$ variou substancialmente em função dos valores anuais do índice em comparação com a estimativa estacionária. A utilização do modelo não estacionário, em função do melhor desempenho indicado pelo DIC, deve representar mais adequadamente o risco de falha e auxiliar na tomada de decisão por parte dos gestores de recursos hídricos.

7 RECOMENDAÇÕES

As conclusões apresentadas permitem afirmar que, de maneira geral, os objetivos propostos nessa dissertação foram atingidos. Entretanto, alguns aspectos deste trabalho poderiam ser melhor explorados em futuras pesquisas. Algumas recomendações nesse sentido:

Neste estudo postos fluviométricos cujas séries manifestam autocorrelação serial foram desconsiderados. A autocorrelação, no entanto, pode ser modelada por algumas abordagens estatísticas, como modelos autorregressivos de média móvel (SALAS *et al.*, 2019). A regionalização de vazões mínimas de bacias sob essa condição carece de estudos e essa pode ser uma via promissora para investigar de aspectos espaciais da autocorrelação serial.

Parte da distribuição *a posteriori* da $Q_{7,10}$ abrangeu quantis de vazões negativas, sem significado físico. Não foi investigado neste estudo o ajuste de distribuições de probabilidade a partir dos dados transformado por seu valor inverso (BHATTI; KROLL; VOGEL, 2019). Essa abordagem tem potencial para contornar problemas práticos de estimação de vazões negativas em distribuições não limitadas inferiormente. Estudos que investiguem o qualidade dos ajustes por essa abordagem podem auxiliar a clarear seu potencial.

A modelagem por meio do processo Gaussiano para capturar a dependência espacial manifestada pelos resíduos da regressão dos parâmetros dos MHBs mostrou ser uma linha de trabalho promissora para melhorar a predição de modelos regionais de análise de frequência. Entretanto, por problemas de identificabilidade, distribuições *a priori* informativas devem ser assinalada para os parâmetros do variograma. Estudos regionais com modelos similares ao modelo M1 podem ajudar a melhorar o entendimento acerca da variação espacial dos resíduos para a variável de interesse, seja vazão mínima, como desenvolvido neste estudo, vazão máxima sob os moldes de Lima *et al.* (2016) e Sampaio e Costa (2021), ou precipitação.

Neste estudo, bacias aninhadas foram utilizadas as *is* nas análises e inferências para verificar o impacto das hipóteses de dependência espacial dos dados e dos resíduos

dos modelos. Devido a estrutura hierárquica da rede de drenagem e do modelo espacial utilizado, o uso dessas bacias *as is* pode resultar em inferências de quantis menores em postos com maior área de drenagem, característica não desejável por ser fisicamente inconsistente, além de aumentar o efeito da dependência entre observações concomitantes. Uma abordagem alternativa para reduzir os problemas citados consiste em trabalhar com bacias/vazões incrementais para diminuir a dependência espacial de bacias aninhadas.

Apenas uma covariável temporal foi avaliada neste estudo. No entanto, é provável que outros agentes relacionados à climatologia de larga escala influenciem o comportamento regional das vazões mínimas. Outros estudos poderiam investigar a utilização de outras componentes temporais na AFR, como indicadores climáticos de bloqueio atmosférico.

FINANCEAMENTO

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Código de Financiamento 001.

REFERÊNCIAS

AGÊNCIA NACIONAL DE ÁGUAS E SANEAMENTO BÁSICO (BRASIL). *Manual de Usos Consuntivos da Água no Brasil*. . Brasília: [s.n.], 2019.

AHN, K.-H.; PALMER, R.; STEINSCHNEIDER, S. A hierarchical Bayesian model for regionalized seasonal forecasts: Application to low flows in the northeastern United States. *Water Resources Research*, v. 53, n. 1, p. 503–521, jan. 2017. Disponível em: <<http://doi.wiley.com/10.1002/2016WR019605>>.

ALVARES, C. A. *et al.* Köppen's climate classification map for Brazil. *Meteorologische Zeitschrift*, v. 22, n. 6, p. 711–728, 2013.

BADJANA, H. M. *et al.* Bayesian trend analysis in annual rainfall total, duration and maximum in the Kara River basin (West Africa). *Journal of Hydrology: Regional Studies*, v. 13, n. August, p. 255–273, 2017. Disponível em: <<http://dx.doi.org/10.1016/j.ejrh.2017.08.009>>.

BANERJEE, S.; CARLIN, B. P.; GELFAND, A. E. *Hierarchical Modeling and Analysis for Spatial Data*. 2. ed. New York: Chapman and Hall/CRC, 2014. Disponível em: <<https://www.taylorfrancis.com/books/9781439819180>>.

BARREIRO, M. Influence of ENSO and the South Atlantic Ocean on climate predictability over Southeastern South America. *Climate Dynamics*, v. 35, n. 7–8, p. 1493–1508, 22 dez. 2010. Disponível em: <<http://link.springer.com/10.1007/s00382-009-0666-9>>.

BARTIKO, D.; CHAFFE, P. L. B.; BONUMÁ, N. B. Não-estacionariedade em séries de vazões máximas diárias do Sul do Brasil. *Revista Brasileira de Recursos Hídricos*, v. 22, 2017.

BEARD, L. R. *Statistical methods in hydrology*. . Sacramento, CA: [s.n.], 1962.

BHATTI, S. J.; KROLL, C. N.; VOGEL, R. M. Revisiting the Probability Distribution of Low Streamflow Series in the United States. *Journal of Hydrologic Engineering*, v. 24, n. 10, p. 04019043, 2019. Disponível em: <<http://ascelibrary.org/doi/10.1061/%28ASCE%291084-0699%282002%297%3A2%28137%29>>.

BLÖSCHL, G. *et al.* Current European flood-rich period exceptional compared with past 500 years. *Nature*, v. 583, n. 7817, p. 560–566, 2020. Disponível em: <<https://doi.org/10.1038/s41586-020-2478-3>>.

BLÖSCHL, G. *et al.* Twenty-three unsolved problems in hydrology (UPH) – a community perspective. *Hydrological Sciences Journal*, v. 64, n. 10, p. 1141–1158, 2019.

BRACKEN, C. *et al.* A Bayesian Hierarchical Approach to Multivariate Nonstationary Hydrologic Frequency Analysis. *Water Resources Research*, v. 54, n. 1, p. 243–255, 19 jan. 2018. Disponível em: <<https://onlinelibrary.wiley.com/doi/abs/10.1002/2017WR020403>>.

BRACKEN, C. *et al.* Spatial Bayesian hierarchical modeling of precipitation extremes over a large domain. *Water Resources Research*, v. 52, n. 8, p. 6643–6655, ago. 2016. Disponível em: <<http://doi.wiley.com/10.1002/2016WR018768>>.

BRUNNER, M. I.; FURRER, R.; FAVRE, A. C. Modeling the spatial dependence of floods using the Fisher copula. *Hydrology and Earth System Sciences*, v. 23, n. 1, p. 107–124, 2019.

BURN, D. H. DELINEATION OF GROUPS FOR REGIONAL FLOOD FREQUENCY ANALYSIS The problem of assigning a flood risk to a particular flow value is one which has received considerable attention in the literature . The estimation of flood risk through the evaluation of a floo. v. 104, p. 345–361, 1988.

BURN, D. H. Evaluation of regional flood frequency analysis with a region of influence approach. *Water Resources Research*, v. 26, n. 10, p. 2257–2265, 1990.

CAI, W. *et al.* Climate impacts of the El Niño–Southern Oscillation on South America. *Nature Reviews Earth & Environment*, v. 1, n. 4, p. 215–231, 10 abr. 2020. Disponível em: <<http://www.nature.com/articles/s43017-020-0040-3>>.

CAPOZZOLI, C. R.; DE OLIVEIRA CARDOSO, A.; FERRAZ, S. E. T. Padrões de variabilidade de vazão de rios nas principais bacias Brasileiras e associação com índices climáticos. *Revista Brasileira de Meteorologia*, v. 32, n. 2, p. 243–254, 2017.

CASTIGLIONI, S. *et al.* 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, 2011.

CASTIGLIONI, S.; CASTELLARIN, A.; MONTANARI, A. Prediction of low-flow indices in ungauged basins through physiographical space-based interpolation. *Journal of Hydrology*, v. 378, n. 3–4, p. 272–280, 2009. Disponível em: <<http://dx.doi.org/10.1016/j.jhydrol.2009.09.032>>.

CECAV. *Áreas de Ocorrência de Cavernas do Brasil*. Disponível em: <<https://www.icmbio.gov.br/cecav/projetos-e-atividades/provincias-espeleologicas.html>>. Acesso em: 21 abr. 2020.

CECAV. *Cavidades Naturais Subterrâneas Brasileiras*. Disponível em: <<https://www.icmbio.gov.br/cecav/canie.html>>. Acesso em: 21 abr. 2020.

CHAVES, H. *et al.* Regionalização de Vazões Mínimas em Bacias Através de

Interpolação em Sistemas de Informação Geográfica. *Revista Brasileira de Recursos Hídricos*, v. 7, n. 3, p. 43–51, 2002.

CHEN, Y. D. *et al.* Regional analysis of low flow using L-moments for Dongjiang basin, South China. *Hydrological Sciences Journal*, v. 51, n. 6, p. 1051–1064, 2006.

CHOKMANI, K.; OUARDA, T. B. M. J. Physiographical space-based kriging for regional flood frequency estimation at ungauged sites. *Water Resources Research*, v. 40, n. 12, p. 1–13, 2004.

COELHO, C. A. S. *et al.* The 2014 southeast Brazil austral summer drought: regional scale mechanisms and teleconnections. *Climate Dynamics*, v. 46, n. 11–12, p. 3737–3752, 2016.

COMITÊ DO ITAJAÍ. *Caderno Síntese: Plano de recursos hídricos da bacia do Itajaí*. Blumenau: [s.n.], 2010.

COOLEY, D. *et al.* A survey of spatial extremes: Measuring spatial dependence and modeling spatial effects. *Revstat Statistical Journal*, v. 10, n. 1, p. 135–165, 2012.

COOLEY, D. *et al.* Bayesian spatial modeling of extreme precipitation return levels. *Journal of the American Statistical Association*, v. 102, n. 479, p. 824–840, 2007.

COSTA, V. Correlation and Regression. In: NAGHETTINI, M. (Org.). *Fundamentals of Statistical Hydrology*. Cham: Springer International Publishing, 2017. p. 391–440. Disponível em: <http://link.springer.com/10.1007/978-3-319-43561-9_9>.

COSTA, V. *et al.* Regional Modeling of Long-Term and Annual Flow Duration Curves: Reliability for Information Transfer with Evolutionary Polynomial Regression. *Journal of Hydrologic Engineering*, v. 26, n. 2, p. 04020067, 2021.

COSTA, V.; FERNANDES, W.; STARICK, Â. Identifying Regional Models for Flow Duration Curves with Evolutionary Polynomial Regression: Application for Intermittent Streams. *Journal of Hydrologic Engineering*, v. 25, n. 1, p. 1–12, 2020.

COSTA, V.; SAMPAIO, J. A Bayesian approach for estimating the distribution of annual maximum floods with a mixture model. *Journal of Hydrologic Engineering*, [S.d.].

CPRM. *Atlas Pluviométrico do Brasil*. Brasília: [s.n.], 2011. Disponível em: <<http://www.cprm.gov.br/publique/Hidrologia/Mapas-e-Publicacoes/Atlas-Pluviometrico-do-Brasil-1351.html>>. Acesso em: 20 abr. 2020.

CPRM. *Regionalização de vazões. Sub-bacias 40 e 41: Relatório Final*. Belo Horizonte: CPRM/ANEEL. , 2001

CUNHA, A. P. M. A. *et al.* Extreme drought events over Brazil from 2011 to 2019. *Atmosphere*, v. 10, n. 11, 2019.

CUPOLILLO, F. *Diagnóstico hidroclimatológico da bacia do rio Doce*. 2008. 153 f. Universidade Federal de Minas Gerais, 2008.

DALRYMPLE, T. Flood-Frequency Analyses. Manual of Hydrology Part 3. Flood-flow techniques. *Usgpo*, v. 1543- A, p. 80, 1960. Disponível em: <<http://pubs.usgs.gov/wsp/1543a/report.pdf>>.

DAVISON, A. C.; PADOAN, S. A.; RIBATET, M. Statistical modeling of spatial extremes. *Statistical Science*, v. 27, n. 2, p. 161–186, 2012.

DE JESUS, E. T. *et al.* Meteorological and hydrological drought from 1987 to 2017 in doce river basin, Southeastern Brazil. *Revista Brasileira de Recursos Hídricos*, v. 25, p. 1–10, 2020.

DEVINENI, N. *et al.* A tree-ring-based reconstruction of delaware river basin streamflow using hierarchical Bayesian regression. *Journal of Climate*, v. 26, n. 12, p. 4357–4374, 2013.

DO VALE MOREIRA, J. G.; NAGHETTINI, M.; ELEUTÉRIO, J. C. Frequência e risco sob não-estacionariedade em registros pluviométricos da bacia do alto rio Tarauacá, Acre. *Revista Brasileira de Recursos Hídricos*, v. 21, n. 1, p. 232–241, 2016.

DODANGEH, E. *et al.* Flood Frequency Analysis of Interconnected Rivers by Copulas. 2020.

DUDLEY, R. W. *et al.* Low streamflow trends at human-impacted and reference basins in the United States. *Journal of Hydrology*, v. 580, n. October 2019, p. 124254, 2020. Disponível em: <<https://doi.org/10.1016/j.jhydrol.2019.124254>>.

DYRRDAL, A. V. *et al.* Bayesian hierarchical modeling of extreme hourly precipitation in Norway. *Environmetrics*, v. 26, n. 2, p. 89–106, mar. 2015. Disponível em: <<http://doi.wiley.com/10.1002/env.2301>>.

ELESBON, A. A. A. *et al.* Multivariate statistical analysis to support the minimum streamflow regionalization. *Engenharia Agrícola*, v. 35, n. 5, p. 838–851, 2015.

ELETROBRÁS – CENTRAIS HIDRELÉTRICAS BRASILEIRAS S.A. *Metodologia para Regionalização de Vazões*. v. 1 ed. Rio de Janeiro: [s.n.], 1985.

ENGELAND, K.; HISDAL, H. A comparison of low flow estimates in ungauged catchments using regional regression and the HBV-model. *Water Resources Management*, v. 23, n. 12, p. 2567–2586, 2009.

FARSADNIA, F. *et al.* Identification of homogeneous regions for regionalization of watersheds by two-level self-organizing feature maps. *Journal of Hydrology*, v. 509, p. 387–397, 2014. Disponível em: <<http://dx.doi.org/10.1016/j.jhydrol.2013.11.050>>.

FELIPPE, M. F. *et al.* Conexões geo-históricas e contemporâneas entre ocupação territorial, degradação ambiental e rarefação hídrica na Bacia do Rio Doce. *Geografias*, n. Edição Especial Vale do Rio Doce, p. 203–222, 2016.

FERNANDES, L. G.; RODRIGUES, R. R. Changes in the patterns of extreme rainfall events in Southern Brazil. *International Journal of Climatology*, v. 38, n. 3, p. 1337–1352, 2018.

GARCÍA, J. A. *et al.* A Bayesian hierarchical spatio-temporal model for extreme rainfall in Extremadura (Spain). *Hydrological Sciences Journal*, v. 63, n. 6, p. 878–894, 2018. Disponível em: <<https://doi.org/10.1080/02626667.2018.1457219>>.

GELMAN, A. *et al.* *Bayesian Data Analysis*. 3. ed. New York: Chapman and Hall/CRC, 2013. Disponível em: <<https://www.taylorfrancis.com/books/9781439898208>>.

GELMAN, A. Prior distributions for variance parameters in hierarchical models (comment on article by Browne and Draper). *Bayesian Analysis*, v. 1, n. 3, p. 515–534, set. 2006. Disponível em: <<http://projecteuclid.org/euclid.ba/1340371048>>.

GOMES, E. P.; BLANCO, C. J. C.; PESSOA, F. C. L. Regionalization of precipitation with determination of homogeneous regions via fuzzy c-means. *Rbrh*, v. 23, n. 0, 2018.

GRANEMANN, A. R. B.; MINE, M. R. M. Distribuição de probabilidades para vazões mínimas: uma revisão bibliográfica. *Simpósio Brasileiro de Recursos Hídricos*, v. 21, p. 1–8, 2015.

GRIFFIS, V. W.; STEDINGER, J. R. Log-Pearson Type 3 Distribution and Its Application in Flood Frequency Analysis. I: Distribution Characteristics. *Journal of Hydrologic Engineering*, v. 12, n. 5, p. 482–491, 2007.

GRIMM, A. M.; TEDESCHI, R. G. ENSO and Extreme Rainfall Events in South America. *Journal of Climate*, v. 22, n. 7, p. 1589–1609, abr. 2009. Disponível em: <<http://journals.ametsoc.org/doi/10.1175/2008JCLI2429.1>>.

GUDMUNDSSON, L.; TALLAKSEN, L. M.; STAHL, K. Spatial cross-correlation patterns of European low, mean and high flows. *Hydrological Processes*, v. 25, n. 7, p. 1034–1045, 2011.

GUSE, B. *et al.* Effects of intersite dependence of nested catchment structures on probabilistic regional envelope curves. *Hydrology and Earth System Sciences*, v. 13,

n. 9, p. 1699–1712, 2009.

HAINING, R.; LI, G. *Modelling Spatial and Spatial-Temporal Data*. Boca Raton, FL: Chapman and Hall/CRC, 2020. Disponível em: <<https://www.taylorfrancis.com/books/9781482237436>>.

HEO, J. H. *et al.* Regression equations of probability plot correlation coefficient test statistics from several probability distributions. *Journal of Hydrology*, v. 355, n. 1–4, p. 1–15, 2008.

HOFFMAN, M. D.; GELMAN, A. The no-U-turn sampler: Adaptively setting path lengths in Hamiltonian Monte Carlo. *Journal of Machine Learning Research*, v. 15, p. 1593–1623, 2014.

HORA, A. M. *et al.* Da exploração econômica da bacia hidrográfica do Rio Doce ao atual processo de degradação de seus recursos naturais. In: GUEDES, G. R. O. R. (Org.). *Território, mobilidade populacional e ambiente*. Governador Valadares: Univale/CBH-Doce, 2012. p. 201–234.

HOSKING, J. R. M. *L-Moments*. [S.l.: s.n.]. Disponível em: <<https://cran.r-project.org/package=lmom>>. , 2019

HOSKING, J. R. M.; WALLIS, J. R. *Regional frequency analysis: an approach based on L-moments*. Cambridge: Cambridge Cambridge University Press, 1997.

HUANG, B. *et al.* Extended Reconstructed Sea Surface Temperature, Version 5 (ERSSTv5): Upgrades, Validations, and Intercomparisons. *Journal of Climate*, v. 30, n. 20, p. 8179–8205, 15 out. 2017. Disponível em: <[10.1175/JCLI-D-16-0836.1](https://doi.org/10.1175/JCLI-D-16-0836.1)>.

IBGE. *Mapa Hidrogeológico Região Sudeste*. Disponível em: <<https://www.ibge.gov.br/geociencias/informacoes-ambientais/geologia/15824-hidrogeologia.html?=&t=downloads>>. Acesso em: 21 abr. 2020a.

IBGE. *Mapa Hidrogeológico Região Sul*. Disponível em: <<https://www.ibge.gov.br/geociencias/informacoes-ambientais/geologia/15824-hidrogeologia.html?=&t=downloads>>. Acesso em: 21 abr. 2020b.

IGAM. *PLANO INTEGRADO DE RECURSOS HÍDRICOS DA BACIA HIDROGRÁFICA DO RIO DOCE E PLANOS DE AÇÕES PARA AS UNIDADES DE PLANEJAMENTO E GESTÃO DE RECURSOS HÍDRICOS NO ÂMBITO DA BACIA DO RIO DOCE - Volume 1*. . Belo Horizonte: [s.n.], 2010.

KIM, S. U.; LEE, K. S. Regional low flow frequency analysis using Bayesian regression and prediction at ungauged catchment in Korea. *KSCE Journal of Civil Engineering*, v. 14, n. 1, p. 87–98, 2010.

KJELDSEN, T. R.; JONES, D. A. An exploratory analysis of error components in hydrological regression modeling. *Water Resources Research*, v. 45, n. 2, p. 1–13, 2009.

KJELDSEN, T. R.; JONES, D. A. Estimation of an index flood using data transfer in the UK. *Hydrological Sciences Journal*, v. 52, n. 1, p. 86–98, 2007.

KOUTSOYIANNIS, D. Nonstationarity versus scaling in hydrology. *Journal of Hydrology*, v. 324, n. 1–4, p. 239–254, 2006.

KOUTSOYIANNIS, D.; MONTANARI, A. Negligent killing of scientific concepts: the stationarity case. *Hydrological Sciences Journal*, v. 60, n. 7–8, p. 1174–1183, 2015. Disponível em: <<http://dx.doi.org/10.1080/02626667.2014.959959>>.

KUCZERA, G. Combining site-specific and regional information: An empirical Bayes Approach. *Water Resources Research*, v. 18, n. 2, p. 306–314, 1982.

KUCZERA, G. Comprehensive at-site flood frequency analysis using Monte Carlo Bayesian inference. *SpringerReference*, v. 35, n. 5, p. 1551–1557, 1999.

LAAHA, G. *et al.* Prediction of low flows in ungauged basins. In: BLOSCHL, G. *et al.* (Org.). *Runoff Prediction in Ungauged Basins*. Cambridge: Cambridge University Press, 2013. p. 163–188. Disponível em: <https://www.cambridge.org/core/product/identifier/9781139235761%23c02818-8-1/type/book_part>.

LEME, E. J.; CHAUDHRY, F. Vazão Mínima do Vale Médio do Rio Jaguari Mirim. *Revista Brasileira de Recursos Hídricos*, v. 10, n. 4, p. 127–136, 2005.

LEMOINE, N. P. Moving beyond noninformative priors: why and how to choose weakly informative priors in Bayesian analyses. *Oikos*, v. 128, n. 7, p. 912–928, 2019.

LIANG, Z.; CHANG, W.; LI, B. Bayesian flood frequency analysis in the light of model and parameter uncertainties. *Stochastic Environmental Research and Risk Assessment*, v. 26, n. 5, p. 721–730, 2012.

LIMA, C. H. R. *et al.* A hierarchical Bayesian GEV model for improving local and regional flood quantile estimates. *Journal of Hydrology*, v. 541, p. 816–823, 2016. Disponível em: <<http://dx.doi.org/10.1016/j.jhydrol.2016.07.042>>.

LIMA, C. H. R.; LALL, U. Spatial scaling in a changing climate: A hierarchical bayesian model for non-stationary multi-site annual maximum and monthly streamflow. *Journal of Hydrology*, v. 383, n. 3–4, p. 307–318, 2010. Disponível em: <<http://dx.doi.org/10.1016/j.jhydrol.2009.12.045>>.

LINS, H. F.; COHN, T. A. Stationarity: Wanted dead or alive? *Journal of the American Water Resources Association*, v. 47, n. 3, p. 475–480, 2011.

LISBOA, L. *et al.* System for analysis of water abstraction and effluent dilution in piracicaba river basin, Minas Gerais State, Brazil. *Engenharia Sanitaria e Ambiental*, v. 24, n. 5, p. 929–937, 2019.

LUIZ, S. F. *Regionalização hidrológica de vazões mínimas por meio do Método dos Mínimos Quadrados Generalizados aplicada à bacia do Alto São Francisco*. 2013. 134 f. Universidade Federal de Minas Gerais, 2013.

LUIZ, S.; FERNANDES, W.; JUNIOR, D. Regionalização Hidrológica de Vazões Mínimas por meio dos Métodos OLS e WLS Aplicada à Bacia do Alto São Francisco. *Revista Brasileira de Recursos Hídricos*, v. 18, n. 4, p. 231–241, 2013.

LUN, D. Detecting Flood - Rich and Flood - Poor Periods in Annual Peak Discharges Across Europe Water Resources Research. 2020.

MADSEN, H.; ROSBJERG, D. Generalized least squares and empirical Bayes estimation in regional partial duration series index-flood modeling. *Water Resources Research*, v. 33, n. 4, p. 771–781, 1997.

MARTINS, E. S.; STEDINGER, J. R. Generalized maximum-likelihood generalized extreme-value quantile estimators for hydrologic data. *Water Resources Research*, v. 36, n. 3, p. 737–744, mar. 2000. Disponível em: <<http://doi.wiley.com/10.1029/1999WR900330>>.

MATALAS, N. C. Comment on the announced death of stationarity. *Journal of Water Resources Planning and Management*, v. 138, n. 4, p. 311–312, 2012.

MATOS, T. S. *et al.* Regionalization of maximum, minimum and mean streamflows for the Juruena River basin, Brazil. *Ambiente e Agua - An Interdisciplinary Journal of Applied Science*, v. 15, n. 3, p. 1, 15 maio 2020. Disponível em: <https://www.scielo.br/scielo.php?script=sci_arttext&pid=S1980-993X2020000300300&lng=en&nrm=iso&tlng=en>.

MEHMOOD, A. *et al.* Non-stationary bayesian modeling of annual maximum floods in a changing environment and implications for flood management in the Kabul River Basin, Pakistan. *Water (Switzerland)*, v. 11, n. 6, 2019.

MELLO, F. W. P. J.; PESSOA, F. C. L.; SANTANA, L. R. Regionalization of minimum flows in the hydrographic basin of the Rio Mearim in the state of Maranhão. *Research, Society and Development*, v. 9, n. 6, p. 1–23, 2020.

MELO, L. S. *Distribuições e tendências de vazões na bacia hidrográfica do rio das velhas*. 2016. Universidade Federal de São João Del Rei, 2016.

MERZ, B *et al.* HESS Opinions “ More efforts and scientific rigour are needed to attribute trends in flood time series ”. p. 1379–1387, 2012.

MERZ, BRUNO; THIEKEN, A. H. Separating natural and epistemic uncertainty in flood frequency analysis. *Journal of Hydrology*, v. 309, n. 1–4, p. 114–132, 2005.

MILLY, P. C. D. *et al.* Climate change: Stationarity is dead: Whither water management? *Science*, v. 319, n. 5863, p. 573–574, 2008.

MISHRA, A. K.; SINGH, V. P.; KONAPALA, G. Low Flow and Drought Analysis. In: SINGH, V. P. (Org.). . *Handbook of Applied Hydrology, Second Edition*. 2. ed. [S.l.]: McGraw-Hill Educational, 2016. .

MMA. *Bases do PNRH*. . [S.l: s.n.], 2005.

MMA. *Caderno da Região Hidrográfica Atlântico Sudeste*. . Brasília: [s.n.], 2006.

MODARRES, R. Regional frequency distribution type of low flow in North of Iran by L-moments. *Water Resources Management*, v. 22, n. 7, p. 823–841, 2008.

MONTANARI, A.; KOUTSOYIANNIS, D. Modeling and mitigating natural hazards: Stationarity is immortal! *Water Resources Research*, v. 50, n. 12, p. 9748–9756, dez. 2014. Disponível em: <<http://doi.wiley.com/10.1002/2014WR016092>>.

MÜLLER, M. F.; THOMPSON, S. E. TopREML: A topological restricted maximum likelihood approach to regionalize trended runoff signatures in stream networks. *Hydrology and Earth System Sciences*, v. 19, n. 6, p. 2925–2942, 2015.

NAGHETTINI, M. Parameter and Quantile Estimation. In: NAGHETTINI, M. (Org.). . *Fundamentals of Statistical Hydrology*. Cham: Springer International Publishing, 2017a. p. 203–250. Disponível em: <http://link.springer.com/10.1007/978-3-319-43561-9_6>.

NAGHETTINI, M. Statistical Hypothesis Testing. In: NAGHETTINI, M. (Org.). . *Fundamentals of Statistical Hydrology*. Cham: Springer International Publishing, 2017b. p. 251–309. Disponível em: <http://link.springer.com/10.1007/978-3-319-43561-9_7>.

NAGHETTINI, M.; PINTO, E. J. DE A. Regional Frequency Analysis of Hydrologic Variables. In: NAGHETTINI, M. (Org.). . *Fundamentals of Statistical Hydrology*. Cham: Springer International Publishing, 2017. p. 441–495. Disponível em: <http://link.springer.com/10.1007/978-3-319-43561-9_8>.

NAJAFI, M. R. *et al.* Analysis of runoff extremes using spatial hierarchical Bayesian modeling. *Water Resources Research*, v. 49, n. 10, p. 6656–6670, out. 2013.

Disponível em: <<http://doi.wiley.com/10.1002/wrcr.20381>>.

NASA JPL. *NASA Shuttle Radar Topography Mission Global 1 arc second*.

Disponível em: <<https://doi.org/10.5067/MEaSURES/SRTM/SRTMGL1.003>>. Acesso em: 22 abr. 2020.

OPITZ, T. *et al.* INLA goes extreme: Bayesian tail regression for the estimation of high spatio-temporal quantiles. *Extremes*, v. 21, n. 3, p. 441–462, 25 set. 2018.

Disponível em: <<http://link.springer.com/10.1007/s10687-018-0324-x>>.

OUARDA, T. B. M. J. *et al.* Regional flood frequency estimation with canonical correlation analysis. *Journal of Hydrology*, v. 254, n. 1–4, p. 157–173, 2001.

OUARDA, T. B. M. J.; CHARRON, C. Changes in the distribution of hydro-climatic extremes in a non-stationary framework. *Scientific Reports*, v. 9, n. 1, p. 1–8, 2019.

Disponível em: <<http://dx.doi.org/10.1038/s41598-019-44603-7>>.

OUARDA, T. B. M. J.; CHARRON, C.; ST-HILAIRE, A. Statistical Models and the Estimation of Low Flows. *Canadian Water Resources Journal*, v. 33, n. 2, p. 195–206, 2008.

PENALBA, O. C.; RIVERA, J. A. Precipitation response to El Niño/La Niña events in Southern South America - Emphasis in regional drought occurrences. *Advances in Geosciences*, v. 42, n. 2005, p. 1–14, 2016.

PIERCE, D. *ncdf4: Interface to Unidata netCDF (Version 4 or Earlier) Format Data Files*. [S.l.: s.n.]. Disponível em: <<https://cran.r-project.org/package=ncdf4>>. , 2019

PILON, P. J. The Weibull distribution applied to regional low flow frequency analysis. In: BERAN, M. A. *et al.* (Org.). *Regionalization in Hydrology, Proceedings of a symposium held at Ljubljana, April 1990*. Ljubljana: [s.n.], 1990. p. 227–237.

PROJETO MAPBIOMAS. *Mapbiomas versão 4.1*. Disponível em: <<https://mapbiomas.org/download>>. Acesso em: 8 maio 2020.

RAO, A. R.; HAMED, K. H. *Flood Frequency Analysis*. Boca Raton: CRC Press, 2000.

RAO, A. R.; SRINIVAS, V. V. Regionalization of watersheds by fuzzy cluster analysis. *Journal of Hydrology*, v. 318, n. 1–4, p. 57–79, 2006.

REBOITA, M. S. *et al.* Regimes de precipitação na América do Sul: uma revisão bibliográfica. *Revista Brasileira de Meteorologia*, v. 25, n. 2, p. 185–204, 2010.

REICH, B. J.; GHOSH, S. K. *Bayesian Statistical Methods*. Boca Raton, FL:

Chapman and Hall/CRC, 2019. Disponível em:
<<https://www.taylorfrancis.com/books/9780429510915>>.

RENARD, B. A Bayesian hierarchical approach to regional frequency analysis. *Water Resources Research*, v. 47, n. 11, p. 1–21, nov. 2011. Disponível em:
<<http://doi.wiley.com/10.1029/2010WR010089>>.

RENARD, B.; LALL, U. Regional frequency analysis conditioned on large-scale atmospheric or oceanic fields. *Water Resources Research*, v. 50, n. 12, p. 9536–9554, dez. 2014. Disponível em: <<http://doi.wiley.com/10.1002/2014WR016277>>.

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

REQUENA, A. I.; OUARDA, T. B. M. J.; CHEBANA, F. Low-flow frequency analysis at ungauged sites based on regionally estimated streamflows. *Journal of Hydrology*, v. 563, n. June, p. 523–532, 2018.

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

RIES III, K. G.; FRIESZ, P. J. *Methods for estimating low-flow statistics for Massachusetts streams*. [S.l.: s.n.], 2000. Disponível em:
<<https://pubs.er.usgs.gov/publication/wri004135>>.

RODRIGUES, R. R. *et al.* Common cause for severe droughts in South America and marine heatwaves in the South Atlantic. *Nature Geoscience*, v. 12, n. 8, p. 620–626, 2019.

SALAS, J. D. *et al.* Low Flows and Droughts. *Statistical Analysis of Hydrologic Variables*. Reston, VA: American Society of Civil Engineers, 2019. p. 269–332. Disponível em: <<http://ascelibrary.org/doi/10.1061/9780784415177.ch08>>.

SAMPAIO, J.; COSTA, V. Bayesian regional flood frequency analysis with GEV hierarchical models under spatial dependency structures. *Hydrological Sciences Journal*, v. 0, n. ja, p. null, 2021. Disponível em:
<<https://doi.org/10.1080/02626667.2021.1873997>>.

SEAMAN, J. W.; SEAMAN, J. W.; STAMEY, J. D. Hidden dangers of specifying noninformative priors. *American Statistician*, v. 66, n. 2, p. 77–84, 2012.

SERINALDI, F.; KILSBY, C. G.; LOMBARDO, F. Untenable nonstationarity: An assessment of the fitness for purpose of trend tests in hydrology. *Advances in Water Resources*, v. 111, n. June 2017, p. 132–155, 2018.

- SHI, P. *et al.* Regional frequency analysis of low flow based on L moments: Case study in karst area, Southwest China. *Journal of Hydrologic Engineering*, v. 15, n. 5, p. 370–377, 2010.
- SHIAU, J.-T.; WU, F.-C. Regionalization of natural flow regime: application to environmental flow optimization at ungauged sites. *River Research and Applications*, v. 25, n. 9, p. 1071–1089, nov. 2009. Disponível em: <<http://doi.wiley.com/10.1002/rra.1207>>.
- SILVA, A. T.; NAGHETTINI, M.; PORTELA, M. M. On some aspects of peaks-over-threshold modeling of floods under nonstationarity using climate covariates. *Stochastic Environmental Research and Risk Assessment*, v. 30, n. 1, p. 207–224, 19 jan. 2016. Disponível em: <<http://dx.doi.org/10.1007/s00477-015-1072-y>>.
- SIMPSON, D. *et al.* Penalising model component complexity: A principled, practical approach to constructing priors. *Statistical Science*, v. 32, n. 1, p. 1–28, 2017.
- SKØIEN, J. O.; MERZ, R.; BLÖSCHL, G. Top-kriging - geostatistics on stream networks. *Hydrology and Earth System Sciences*, v. 10, n. 2, p. 277–287, 26 abr. 2006. Disponível em: <<http://www.hydrol-earth-syst-sci.net/10/277/2006/>>.
- SMAKHTIN, V. U. Low flow hydrology: A review. *Journal of Hydrology*, v. 240, n. 3–4, p. 147–186, 2001.
- SOUZA, R. M. *Estimativa de Vazões Mínimas na Bacia do Altíssimo Rio Negro, Região Sul-Brasileira, com Aplicação do Modelo SWAT*. 2011. 122 f. Universidade Federal do Paraná, 2011.
- SPIEGELHALTER, D. J. *et al.* Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, v. 64, n. 4, p. 583–639, out. 2002. Disponível em: <<http://doi.wiley.com/10.1111/1467-9868.00353>>.
- SPIEGELHALTER, D. J. *et al.* The deviance information criterion: 12 years on. *Journal of the Royal Statistical Society. Series B: Statistical Methodology*, v. 76, n. 3, p. 485–493, 2014.
- STAN DEVELOPMENT TEAM. *RStan: the R interface to Stan*. [S.l.: s.n.]. Disponível em: <<http://mc-stan.org/>>. , 2020
- STEDINGER, J. R.; GRIFFIS, V. W. Getting from here to where? Flood frequency analysis and climate. *Journal of the American Water Resources Association*, v. 47, n. 3, p. 506–513, 2011.
- STEDINGER, J. R.; TASKER, G. D. Regional Hydrologic Analysis - 1. Ordinary, Weighted, and Generalized Least Squares Compared. *Water Resources Research*,

v. 21, n. 9, p. 1421–1432, set. 1985. Disponível em:
<<http://doi.wiley.com/10.1029/WR021i009p01421>>.

STEDINGER, J. R.; VOGEL, R. M.; FOUFOULA-GEORGIU, E. Frequency analysis of extreme events. In: MAIDMENT, D. R. (Org.). *Handbook of Hydrology*. New York: McGraw-Hill, 1993. .

STEINSCHNEIDER, S.; LALL, U. A hierarchical Bayesian regional model for nonstationary precipitation extremes in Northern California conditioned on tropical moisture exports. *Water Resources Research*, v. 51, n. 3, p. 1472–1492, 2015. Disponível em:
<<https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2014WR016664>>.

SUN, X. *et al.* A general regional frequency analysis framework for quantifying local-scale climate effects: A case study of ENSO effects on Southeast Queensland rainfall. *Journal of Hydrology*, v. 512, p. 53–68, 2014. Disponível em:
<<http://dx.doi.org/10.1016/j.jhydrol.2014.02.025>>.

SUN, X. *et al.* Hierarchical Bayesian clustering for nonstationary flood frequency analysis: Application to trends of annual maximum flow in Germany. *Water Resources Research*, v. 51, n. 8, p. 6586–6601, 2015. Disponível em:
<<https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2015WR017117>>.

TASKER, G. D. Hydrologic regression with weighted least squares. *Water Resources Research*, v. 16, n. 6, p. 1107–1113, 1980.

TASKER, G. D.; STEDINGER, J. R. An operational GLS model for hydrologic regression. *Journal of Hydrology*, v. 111, n. 1–4, p. 361–375, jan. 1989. Disponível em: <<https://linkinghub.elsevier.com/retrieve/pii/0022169489902680>>.

TASKER, G. D.; STEDINGER, J. R. Regional skew with weighted ls regression. *Journal of Water Resources Planning and Management*, v. 112, n. 2, p. 225–237, 1986.

TEDESCHI, R. G.; GRIMM, A. M.; CAVALCANTI, I. F. A. Influence of Central and East ENSO on extreme events of precipitation in South America during austral spring and summer. *International Journal of Climatology*, v. 35, n. 8, p. 2045–2064, 2015.

TEDESCHI, R. G.; GRIMM, A. M.; CAVALCANTI, I. F. A. Influence of Central and East ENSO on precipitation and its extreme events in South America during austral autumn and winter. *International Journal of Climatology*, v. 36, n. 15, p. 4797–4814, 2016.

THORARINSDOTTIR, T. L. *et al.* Bayesian Regional Flood Frequency Analysis for Large Catchments. *Water Resources Research*, v. 54, n. 9, p. 6929–6947, set. 2018. Disponível em: <<http://doi.wiley.com/10.1029/2017WR022460>>.

TSAKIRIS, G.; NALBANTIS, I.; CAVADIAS, G. Regionalization of low flows based on Canonical Correlation Analysis. *Advances in Water Resources*, v. 34, n. 7, p. 865–872, 2011. Disponível em: <<http://dx.doi.org/10.1016/j.advwatres.2011.04.007>>.

VEHTARI, A. *et al.* loo: Efficient leave-one-out cross-validation and WAIC for Bayesian models. . [S.l: s.n.]. Disponível em: <<https://mc-stan.org/loo/>>. , 2020

VEHTARI, A. *et al.* Rank-Normalization, Folding, and Localization: An Improved \widehat{R} for Assessing Convergence of MCMC. *Bayesian Analysis*, p. 1–26, jul. 2019. Disponível em: <<http://dx.doi.org/10.1214/20-BA1221>>.

VEHTARI, A.; GELMAN, A.; GABRY, J. Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. *Statistics and Computing*, v. 27, n. 5, p. 1413–1432, 2017.

VIBRANS, A. C. *A Cobertura Florestal da Bacia do Rio Itajaí Elementos para uma Análise Histórica*. 2003. 240 f. Universidade Federal de Santa Catarina, 2003.

VOGEL, R. M.; KROLL, C. N. Generalized Low-Flow Frequency Relationships for Ungaged Sites in Massachusetts. *JAWRA Journal of the American Water Resources Association*, v. 26, n. 2, p. 241–253, 1990.

WANG, H.; XUAN, Y. Spatial Dependency in Nonstationary GEV Modelling of Extreme Precipitation over Great Britain. *Hydrology and Earth System Sciences Discussions*, n. February, p. 1–24, 2020.

WARD, P. J. *et al.* Annual flood sensitivities to el nintild;O-Southern Oscillation at the global scale. *Hydrology and Earth System Sciences*, v. 18, n. 1, p. 47–66, 2014.

WATANABE, S. Asymptotic equivalence of Bayes cross validation and widely applicable information criterion in singular learning theory. *Journal of Machine Learning Research*, v. 11, p. 3571–3594, 2010.

WU, YENAN *et al.* Local and regional flood frequency analysis based on hierarchical Bayesian model: application to annual maximum streamflow for the Huaihe River basin. *Hydrology and Earth System Sciences Discussions*, n. January, p. 1–21, 2018.

WU, YUN BIAO; XUE, L. QING; LIU, Y. HONG. Local and regional flood frequency analysis based on hierarchical Bayesian model in Dongting Lake Basin, China. *Water Science and Engineering*, v. 12, n. 4, p. 253–262, 2019. Disponível em: <<https://doi.org/10.1016/j.wse.2019.12.001>>.

YAN, H.; MORADKHANI, H. A regional Bayesian hierarchical model for flood frequency analysis. *Stochastic Environmental Research and Risk Assessment*, v. 29, n. 3, p. 1019–1036, 18 mar. 2014. Disponível em: <<http://link.springer.com/10.1007/s00477-014-0975-3>>.

YAN, H.; MORADKHANI, H. Toward more robust extreme flood prediction by Bayesian hierarchical and multimodeling. *Natural Hazards*, v. 81, n. 1, p. 203–225, 2016.

ZENG, H. *et al.* Nonstationary extreme flood/rainfall frequency analysis informed by large-scale oceanic fields for Xidayang Reservoir in North China. *International Journal of Climatology*, v. 37, n. 10, p. 3810–3820, 2017. Disponível em: <<https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/joc.4955>>.

APÊNDICE A – Caracterização das bacias hidrográficas

Figura A.0.1 - Caracterização da bacia do rio Doce

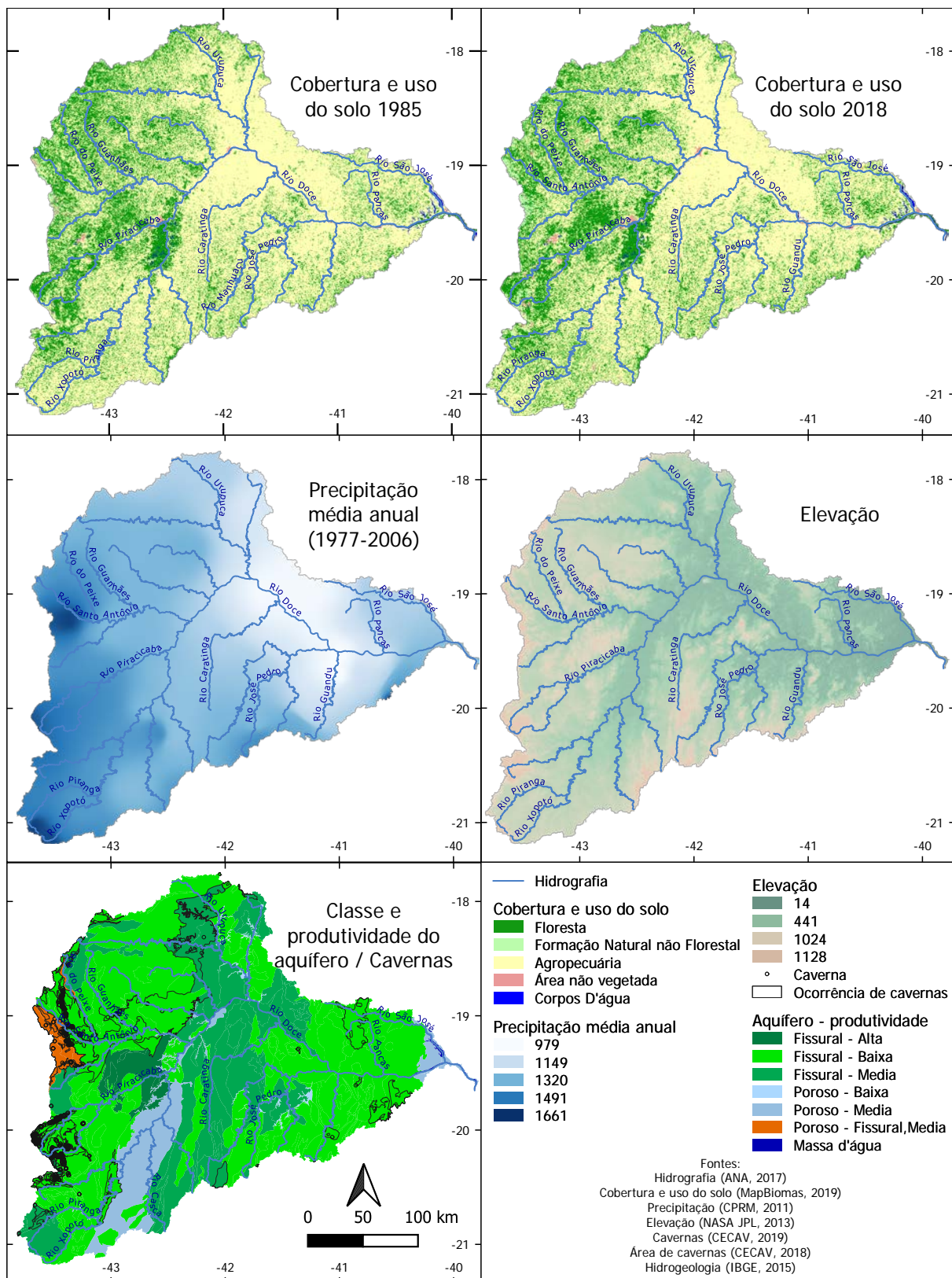
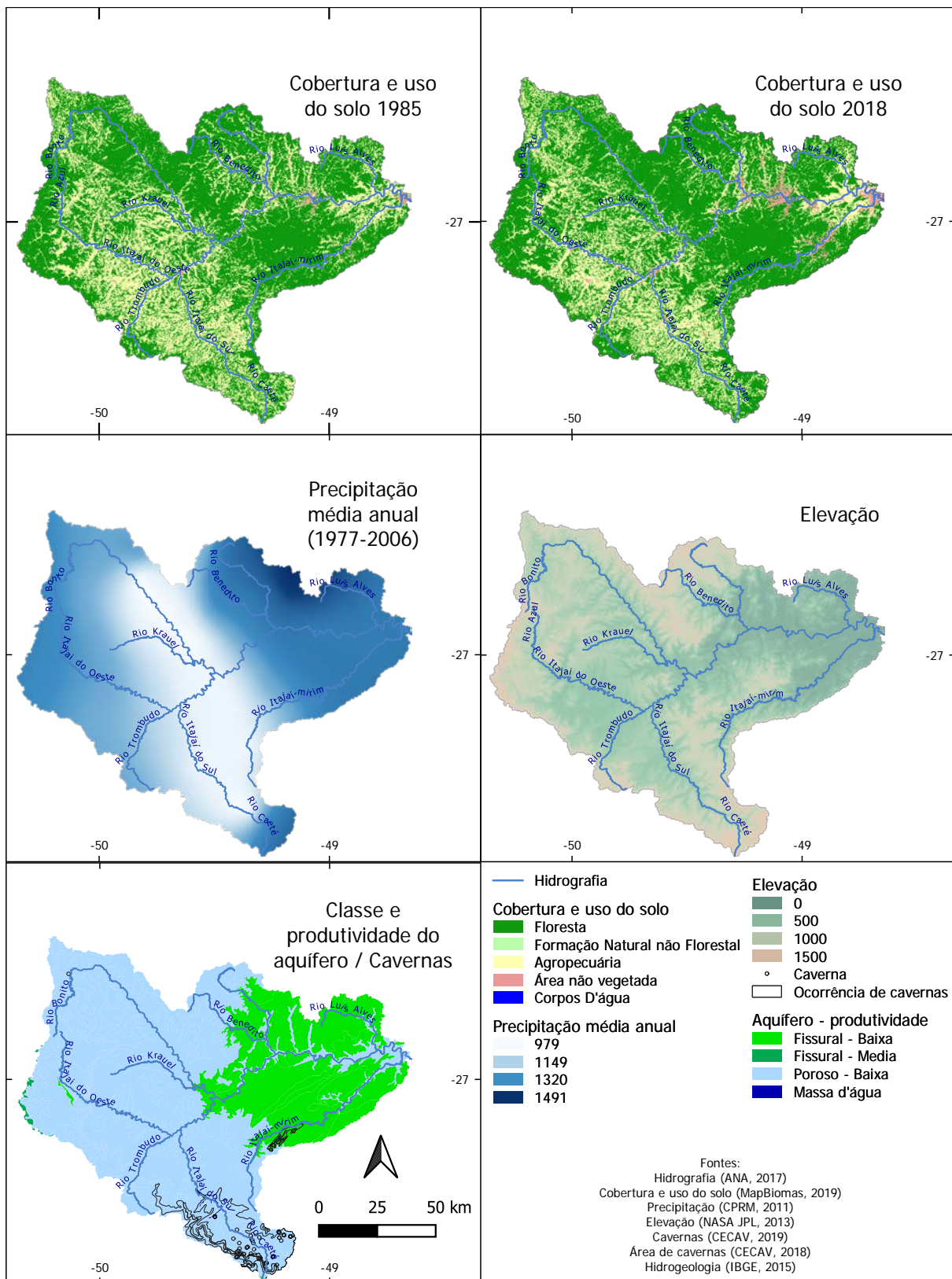


Figura A.0.2 - Caracterização da bacia do rio Itajaí-Açu



APÊNDICE B – Descrição dos postos fluviométricos

Tabela B.1 - Postos fluviométricos da bacia do rio Doce

ID	Código	CC	Curso d'Água	Nome	Tipo	Latitude	Longitude	N	A (km ²)	Perímetro (km)	Lat_c (°)	Long_c (°)	P (mm)	S (m/m)	DD (km/km ²)	QSUB (m ³ /h)
1	56005000		Rio Piranga	Fazenda Do Retiro	C	-20,9333	-43,6833	16	87,0	335,4	-19,0055	-40,7895	1.566	14,18	0,7482	25,00
2	56012000		Rio Guarara	Cachoeira Do Guarara	C	-20,75	-43,5667	19	161,0	159,9	-20,9178	-43,6583	1.464	14,50	0,7595	25,00
3	56040000		Rio Xopotó	Desterro Do Melo	C	-21,15	-43,5167	12	42,0	259,9	-20,7983	-43,5886	1.600	17,88	0,6579	6,50
4	56130000		Rio Oratórios	Usina Ana Florência	C	-20,35	-42,85	11	256,0	489,2	-20,7985	-43,2588	1.338	14,24	0,5823	25,00
5	56170000		Ribeirão Da Cachoeira	Vargem Do Tijucal	C	-20,4833	-43,55	15	56,0	38,2	-20,5560	-43,5589	1.478	14,01	0,6776	6,50
6	56240000		Rio Gualaxo Do Sul	Fazenda Paraíso	C	-20,3875	-43,1817	81	857,0	41,7	-20,4388	-43,5679	1.460	15,46	0,6974	6,53
7	56385000		Rio Casca	São Miguel Do Anta	C	-20,6831	-42,6731	47	534,0	86,8	-20,1277	-43,0460	1.339	16,39	0,7057	24,06
8	56400000		Rio Casca	Jequeri	C	-20,45	-42,6667	28	1.370,0	125,6	-20,7937	-42,6060	1.328	16,23	0,7271	22,49
9	56415000		Rio Casca	Rio Casca	C	-20,2267	-42,6503	78	2.036,0	217,3	-20,6778	-42,6181	1.337	16,09	0,7326	23,31
10	56425000		Rio Doce	Fazenda Cachoeira D'antas	C	-19,9944	-42,6744	32	10.080,0	278,8	-20,5797	-42,6245	1.418	14,84	0,7291	13,83
11	56430000		Rio Doce	Ponte Do Peres	C	-19,9	-42,5833	10	12.800,0	689,8	-20,6082	-43,2124	1.402	14,96	0,7328	15,92
12	56460000		Rio Matipó	Matipo	C	-20,2772	-42,3256	39	615,0	767,6	-20,5740	-43,0887	1.302	15,67	0,7015	19,30
13	56470000		Ribeirão Pernambuco	Caputira	C	-20,1667	-42,2667	14	238,0	134,4	-20,4277	-42,3276	1.286	16,38	0,7489	24,95
14	56500000		Rio Santana	Abre Campo	C	-20,3006	-42,4778	55	272,0	216,0	-20,2968	-42,3093	1.300	17,44	0,6364	24,16
15	56510000		Rio Matipó	Instituto Florestal Raul Soares	C	-20,0967	-42,4594	32	1.800,0	94,5	-20,4219	-42,4696	1.291	15,99	0,7471	20,38
16	56539000		Rio Doce	Cachoeira Dos Oculos-Montante	C	-19,7769	-42,4764	40	15.836,0	71,7	-20,0196	-42,2897	1.380	14,93	0,7338	17,31
17	56565000		Ribeirão Sacramento	Bom Jesus Do Galho	C	-19,8167	-42,3167	37	304,0	841,5	-20,4954	-42,9610	1.272	15,82	0,6989	18,53
18	56631900		Cabeceira Do Rio Conceição	ETA (São Bento Mineração)	C	-19,9708	-43,4761	10	300,5	106,4	-20,0947	-43,5810	1.531	17,84	0,7101	6,52
19	56632000		Ribeirão Do Caraca	Colégio Caraca	C	-20,0833	-43,4833	18	35,0	110,5	-20,0922	-43,5787	1.498	18,30	0,6416	6,50
20	56640000		Ribeirão Santa Bárbara	Carrapato (Brumal)	C	-19,9714	-43,4561	49	426,0	50,9	-20,0732	-43,4793	1.521	17,74	0,7059	6,52

ID	Código	CC	Curso d'Água	Nome	Tipo	Latitude	Longitude	N	A (km ²)	Perímetro (km)	Lat_c (°)	Long_c (°)	P (mm)	S (m/m)	DD (km/km ²)	QSUB (m ³ /h)
21	56665000		Rio Da Prata	São Domingos Do Prata	C	-19,8833	-42,95	10	139,0	336,8	-19,9558	-43,3500	1.402	14,99	0,7496	6,50
22	56667000		Rio Da Prata	Fazenda Da Vargem	C	-19,8	-43	11	439,0	73,0	-19,9551	-42,9303	1.411	15,21	0,7299	11,42
23	56681000		Rio Piracicaba	Antônio Dias	C	-19,65	-42,8667	18	4.250,0	119,4	-19,6622	-43,1615	1.447	14,77	0,7169	13,53
24	56690000		Rio Piracicaba	Ana Matos	C	-19,6106	-42,7711	12	4.816,0	421,1	-19,8999	-43,2608	1.437	15,21	0,7169	14,02
25	56696000	56695000 56700000	Rio Piracicaba	Mario De Carvalho	C	-19,5242	-42,6408	40	5.060,0	451,5	-19,8727	-43,2131	1.425	15,67	0,7483	18,06
26	56846000		Rio Corrente Grande	Porto Santa Rita	C	-18,9506	-42,3569	37	1.965,0	195,2	-18,7325	-42,8574	1.298	16,26	0,7136	6,50
27	56846890	56846900 56847000	Rio Suaçui Pequeno	Fazenda Aconchego	C	-18,9317	-42,1028	10	1.745,0	317,0	-18,8081	-42,7047	1.287	15,93	0,6956	11,68
28	56935000		Rio Caratinga	Dom Cavati	C	-19,3847	-42,1111	45	784,0	171,8	-19,7466	-42,1262	1.243	15,73	0,7277	24,11
29	56940002	56940000	Rio Cuité	Barra Do Cuieté Jusante	C	-19,0619	-41,5333	58	3.220,0	224,1	-19,6764	-42,1309	1.172	16,02	0,7547	22,54
30	56943000		Rio João Pinto Grande	Conselheiro Pena	C	-19,1667	-41,4667	21	197,0	1.701,8	-19,3840	-42,6689	1.042	17,67	0,7078	6,50
31	56960005		Rio Manhuaçu	Fazenda Vargem Alegre	C	-20,1792	-41,9611	31	1.054,0	145,4	-20,3350	-42,1055	1.333	17,73	0,7201	18,51
32	56967000		Rio Manhuaçu	Santana De Manhuaçu	C	-20,1125	-41,9222	25	1.547,0	101,0	-20,3955	-41,9529	1.323	17,99	0,6891	18,85
33	56974000		Rio Manhuaçu	Cachoeira Da Neblina	C	-19,8167	-41,7833	28	2.060,0	243,0	-20,2540	-42,0327	1.312	18,05	0,7499	16,32
34	56976000		Rio Manhuaçu	Fazenda Bragança	C	-19,7497	-41,7808	41	2.282,0	292,1	-20,1671	-41,9869	1.310	18,03	0,7025	16,22
35	56978000		Rio Manhuaçu	Santo Antônio Do Manhuaçu	C	-19,6769	-41,8336	48	2.287,0	307,1	-20,1540	-41,9809	1.307	18,02	0,7364	15,99
36	56982000		Rio Manhuaçu	Paraíso De Ipanema	C	-19,5167	-41,45	15	3.740,0	319,8	-20,1348	-41,9741	1.265	18,50	0,6849	16,98
37	56985000		Ribeirão Santana	Chalé	C	-20,0333	-41,75	13	65,0	123,8	-20,2578	-41,7695	1.294	19,03	0,6958	12,79
38	56986000		Ribeirão São Domingos	Fazenda Boa Esperança	C	-19,9947	-41,7081	23	335,0	52,9	-20,0213	-41,6098	1.319	19,18	0,7240	8,12
39	56988500		Rio Jose Pedro	Ipanema	C	-19,7986	-41,7067	75	1.420,0	204,7	-20,1579	-41,7358	1.322	18,36	0,7362	10,67
40	56989400		Rio Jose Pedro	Assarai-Montante	C	-19,5944	-41,4583	72	3.230,0	202,5	-19,9858	-41,4508	1.266	17,45	0,7487	12,63
41	56990000		Rio Manhuaçu	São Sebastião Da Encruzilhada	C	-19,4917	-41,1617	72	8.728,0	351,7	-19,9874	-41,5891	1.224	17,66	0,7383	13,73
42	56990500		Ribeirão Do Capim	Barra Do Capim	C	-19,5	-41,25	17	726,0	622,2	-19,8768	-41,6646	1.071	15,86	0,7249	10,75

ID	Código	CC	Curso d'Água	Nome	Tipo	Latitude	Longitude	N	A (km ²)	Perímetro (km)	Lat_c (°)	Long_c (°)	P (mm)	S (m/m)	DD (km/km ²)	QSUB (m ³ /h)
43	56990990		Rio Guandú	Afonso Cláudio Montante	C	-20,0783	-41,1217	38	466,0	174,5	-19,7249	-41,2581	1.247	19,21	0,6731	6,50
44	56991500		Rio Guandú	Laranja Da Terra	C	-19,9014	-41,0575	45	1.331,0	137,4	-20,1860	-41,1390	1.208	19,30	0,7232	6,53
45	56992000		Rio Guandú	Baixo Guandu	C	-19,5236	-41,0136	71	2.135,0	238,2	-20,0901	-41,1740	1.146	18,68	0,7271	6,52
46	56992500		Rio Mutum Preto	Barra Do Mutum	C	-19,5	-40,9	11	215,0	1.894,4	-19,4523	-42,4382	1.076	18,41	0,7586	6,50
47	56993002	56993000	Rio Santa Joana	Itaguaçu - Jusante	C	-19,7792	-40,8508	40	481,0	115,7	-19,3648	-40,8984	1.168	17,94	0,7083	6,50
48	56998000		Rio Doce	Linhares	C	-19,4075	-40,0639	25	82.100,0	214,0	-19,1897	-40,7868	1.265	15,47	0,7370	15,38
49	56105000		Ribeirão Vau-Açu	Ponte Do Carvalho	V	-20,4333	-42,8833	9	126,0	109,5	-20,8425	-42,9492	1.330	14,94	0,7710	24,49
50	56631500		Rio Conceição	Ponte Itajuru	V	-19,9667	-43,5	9	422,0	77,7	-20,1273	-43,5987	1.532	17,84	0,4928	6,52
51	56633000		Ribeirão Do Caraca	Ponte Do Brumado	V	-20	-43,4667	8	118,0	27,7	-20,1155	-43,4766	1.498	17,78	0,6993	6,50
52	56787000		Rio Do Tanque	Fazenda Barraca	V	-19,3325	-43,07	48	1.319,0	436,4	-19,0377	-43,3526	1.398	15,18	0,7098	16,28
53	56924500		Rio Caratinga	Santa Rita De Minas	V	-19,8811	-42,138	8	95,0	1.664,4	-19,3901	-42,7369	1.279	17,13	0,6831	25,00
54	56928000		Rio Caratinga	Inhapim	V	-19,5489	-42,1239	8	575,0	81,8	-19,8875	-42,1370	1.252	15,47	0,7354	24,42
55	56941000		Rio Doce	Barra Do Cuieté	V	-19,0664	-41,5356	8	59.500,0	434,8	-19,4590	-41,9370	1.300	15,11	0,7326	17,34
56	56963000		Rio Jequitibá	Parada Da Independência	V	-20,2	-41,95	8	282,0	188,8	-20,3295	-42,0347	1.367	18,96	0,7439	11,22
57	56988000		Rio José Pedro	Fazenda Nova Floresta	V	-19,95	-41,7	15	1.100,0	120,4	-20,1050	-41,6393	1.339	18,80	0,7414	8,22
58	56010000		Rio Piranga	Ponte Do São Lourenço	ACF	-20,7833	-43,5667	24	558,0	69,4	-21,0172	-43,6951	1.507	13,37	0,7361	25,00
59	56028000		Rio Piranga	Piranga	ACF	-20,6881	-43,3006	75	1.395,0	73,0	-20,7741	-43,6773	1.473	13,92	0,7359	19,40
60	56050000		Rio Xopotó	Alto Rio Doce	ACF	-21,0333	-43,4	25	222,0	33,6	-21,1484	-43,5680	1.546	15,74	0,7278	9,87
61	56055000		Rio Xopotó	Braz Pires	ACF	-20,8478	-43,2419	75	1.089,0	88,7	-21,1171	-43,4797	1.489	14,79	0,7371	15,38
62	56065000		Rio Turvo	Senador Firmino	ACF	-20,9072	-43,0983	67	291,0	197,9	-21,0023	-43,4495	1.412	12,57	0,7220	14,73
63	56075000		Rio Piranga	Porto Firme	ACF	-20,6703	-43,0917	76	4.251,0	94,4	-21,0099	-43,0728	1.450	14,35	0,7298	14,74
64	56085000		Rio Turvo Sujo	Seriquite	ACF	-20,724	-42,923	72	342,0	378,5	-20,8698	-43,3928	1.297	13,26	0,7369	14,18
65	56090000		Rio Turvo Limpo	Fazenda Varginha	ACF	-20,714	-43	66	324,0	108,5	-20,7758	-42,8179	1.339	13,21	0,7277	21,86
66	56110005	56110000	Rio Piranga	Ponte Nova-Jusante (Pcd)	ACF	-20,3839	-42,9028	73	6.132,0	494,2	-20,7997	-43,2600	1.417	14,34	0,7439	14,84
67	56145000		Ribeirão Do	São Caetano De	ACF	-20,35	-43,3667	34	138,0	83,9	-20,4663	-42,7991	1.489	16,03	0,6745	15,84

ID	Código	CC	Curso d'Água	Nome	Tipo	Latitude	Longitude	N	A (km ²)	Perímetro (km)	Lat_c (°)	Long_c (°)	P (mm)	S (m/m)	DD (km/km ²)	QSUB (m ³ /h)
			Carmo	Mariana												
68	56148000		Ribeirão Do Carmo	Furquim	ACF	-20,3667	-43,2	24	305,0	75,5	-20,3814	-43,4679	1.471	14,72	0,7137	10,68
69	56150000		Ribeirão Água Limpa	Marimbondo	ACF	-20,5167	-43,6	13	56,0	127,8	-20,3555	-43,3880	1.480	13,02	0,6631	6,97
70	56152000		Córrego Do Garcia	Pai Tomás	ACF	-20,5	-43,6	13	13,0	47,0	-20,5599	-43,6179	1.492	14,50	0,6060	6,50
71	56182000		Ribeirão Falcão	Chapada	ACF	-20,4333	-43,5667	25	48,0	37,1	-20,5586	-43,5585	1.499	11,98	0,6790	6,50
72	56335001	56335000	Rio Do Carmo	Acaiaca Jusante	ACF	-20,3617	-43,1394	73	1.371,0	211,0	-20,4914	-43,4297	1.459	15,42	0,6864	7,48
73	56337000		Rio Gualaxo Do Norte	Fazenda Ocidente	ACF	-20,2675	-43,1008	74	531,0	228,6	-20,4468	-43,3983	1.459	15,28	0,7220	11,77
74	56340000		Rio Do Peixe	Dom Silvério	ACF	-20,1608	-42,9461	16	269,0	165,6	-20,2545	-43,3099	1.418	16,91	0,7368	6,50
75	56484998		Rio Matipó	Raul Soares Montante	ACF	-20,0978	-42,4372	76	1.347,0	85,9	-20,1781	-42,2082	1.288	15,61	0,7235	18,79
76	56520000		Ribeirão Vermelho	Vermelho Velho	ACF	-19,9989	-42,3483	57	163,0	235,8	-20,3024	-42,3561	1.268	16,24	0,6894	17,83
77	56570000		Ribeirão Sacramento	Pingo D'água	ACF	-19,71	-42,4461	39	814,0	123,3	-19,9606	-42,2405	1.263	15,98	0,7314	17,12
78	56610000		Rio Piracicaba	Rio Piracicaba	ACF	-19,9306	-43,1728	81	1.163,0	204,5	-19,8422	-42,2586	1.462	13,72	0,7183	8,21
79	56631000		Rio Conceição	Conceição Do Rio Acima	ACF	-20,0761	-43,5825	20	194,0	231,7	-20,0646	-43,2892	1.543	17,81	0,7010	6,54
80	56659998	56659999 56660000	Rio Piracicaba	Nova Era Iv	ACF	-19,7692	-43,0328	64	3.203,0	117,1	-20,0858	-43,5498	1.463	14,21	0,6880	9,69
81	56670000		Rio Do Peixe	Fazenda Olaria	ACF	-19,7333	-43,0333	25	399,0	115,2	-19,9062	-42,9655	1.410	15,51	0,7001	20,24
82	56719998	56720000	Rio Doce	Belo Oriente	ACF	-19,3278	-42,3975	70	24.159,0	498,6	-19,8409	-43,1717	1.376	14,99	0,7306	19,30
83	56750000		Rio Santo Antônio	Conceição Do Mato Dentro	ACF	-19,0164	-43,4475	68	302,0	983,6	-20,2472	-42,9407	1.418	12,95	0,7480	6,60
84	56765000		Rio Do Peixe	Dom Joaquim	ACF	-18,9656	-43,2442	68	972,0	111,2	-18,8826	-43,4987	1.384	12,88	0,7517	12,35
85	56775000		Rio Santo Antônio, Jusante Rio Do Peixe	Ferros	ACF	-19,2281	-43,02	67	4.058,0	231,5	-18,7573	-43,3628	1.443	14,22	0,7410	12,78
86	56800000		Rio Guanhães	Senhora Do Porto	ACF	-18,8947	-43,0825	68	1.521,0	221,1	-19,5170	-43,2844	1.339	14,48	0,7026	12,55
87	56825000		Rio Santo Antônio	Naque Velho	ACF	-19,1883	-42,4225	39	10.210,0	622,9	-19,0701	-43,1927	1.377	16,05	0,7252	16,20
88	56845000		Rio Corrente Grande	Fazenda Corrente	ACF	-18,8947	-42,7119	59	1.064,0	674,0	-19,0791	-43,1454	1.302	13,00	0,7212	6,50

ID	Código	CC	Curso d'Água	Nome	Tipo	Latitude	Longitude	N	A (km ²)	Perímetro (km)	Lat_c (°)	Long_c (°)	P (mm)	S (m/m)	DD (km/km ²)	QSUB (m ³ /h)
89	56850000		Rio Doce	Governador Valadares	ACF	-18,8819	-41,9511	49	39.828,0	309,0	-18,7016	-42,3761	1.359	15,24	0,7424	17,66
90	56851000		Rio Itambacuri	Campanário	ACF	-18,2392	-41,7311	66	758,0	1.386,8	-19,7519	-42,9135	1.109	15,10	0,7428	23,97
91	56860000		Rio Suaçui Grande	São Pedro Do Suaçui	ACF	-18,3631	-42,6025	48	2.610,0	150,1	-18,1004	-41,7667	1.289	15,17	0,7072	10,88
92	56870000		Rio São Félix	Santa Maria Do Suaçuí	ACF	-18,2017	-42,4547	48	672,0	337,1	-18,3543	-42,9462	1.221	14,76	0,6911	7,21
93	56880000		Rio Urupuca	Fazenda Urupuca	ACF	-18,25	-42,0667	21	2.670,0	156,7	-18,0701	-42,5230	1.196	15,17	0,7466	12,88
94	56891900	56892000	Rio Suaçui Grande	Vila Marias Montante	ACF	-18,575	-41,9181	75	10.200,0	317,0	-17,9904	-42,1540	1.239	15,11	0,7023	13,14
95	56915500		Rio Itambacuri	Jampruca	ACF	-18,4594	-41,8075	31	1.390,0	734,5	-18,2475	-42,4551	1.104	14,51	0,7338	19,04
96	56920000		Rio Doce	Tumiritinga	ACF	-18,9717	-41,6419	42	55.425,0	238,5	-18,2048	-41,7725	1.314	15,06	0,7270	16,89
97	56925000		Rio Caratinga	Caratinga	ACF	-19,7833	-42,0667	27	162,0	47,0	-19,9447	-42,1321	1.271	16,85	0,7188	24,96
98	56948005	56948000	Rio Doce	Resplendor-Jusante	ACF	-19,3431	-41,2461	63	61.237,0	69,4	-19,2720	-41,5158	1.292	15,19	0,7331	17,19
99	56960000		Rio Manhuaçu	Manhuaçu	ACF	-20,2694	-42,0669	40	550,0	1.814,2	-19,3766	-42,6309	1.314	17,10	0,6848	24,88
100	56983000		Cabeceira Do Rio Jose Pedro	Dores De Manhumirim	ACF	-20,1072	-41,7283	72	390,0	502,3	-19,9420	-41,8979	1.383	19,86	0,7385	9,00
101	56993551	56993600	Rio Santa Joana	Jusante Córrego Da Piaba	ACF	-19,5592	-40,7328	65	893,0	150,3	-19,9297	-40,8903	1.140	17,30	0,7102	6,50
102	56994000		Rio Santa Maria	Santa Maria	ACF	-19,5833	-40,5833	25	872,0	228,5	-19,8172	-40,8658	1.219	17,66	0,7083	6,50
103	56994500	56994502	Rio Doce	Colatina	ACF	-19,5333	-40,6297	73	76.400,0	181,0	-19,7903	-40,6735	1.269	15,54	0,7343	15,89
104	56995500		Rio Pancas	Ponte Do Pancas	ACF	-19,4228	-40,6864	47	920,0	253,9	-19,2255	-40,7539	1.166	19,03	0,7277	6,50
105	56998400		Rio São José	Barra De São Gabriel	ACF	-19,0581	-40,5175	45	1.022,0	2.136,5	-19,4465	-42,2529	1.152	16,37	0,8992	6,50
106	56995000		Rio Pancas	Cachoeira Do Oito	0	-19,4333	-40,6167	23	1.160,0	1.982,5	-19,4607	-42,3794	1.168	18,37	0,7191	6,50

Nota: **CC** indica código complementar, referindo-se às estações cujas séries históricas foram agregadas ao posto fluviométrico da mesma linha; na coluna **Tipo**, **C** designa postos de Calibração, **V** postos de validação e **ACF** postos com autocorrelação significativa para o *lag* 1; **A** é a área de drenagem; **Lat_c** e **Long_c** trazem as coordenadas do centroide da área de drenagem; **P** é a precipitação anual, **S** é a declividade do curso d'água principal; **DD** é densidade de drenagem e **QSUB** é a produção do aquífero.

Tabela B.2 - Postos fluviométricos da bacia do rio Itajaí-Açu

ID	Código	CC	Curso d'Água	Nome	Tipo	Latitude	Longitude	N	A (km ²)	Perímetro (km)	Lat_c (°)	Long_c (°)	P (mm)	S (m/m)	DD (km/km ²)	QSUB (m ³ /h)
1	83029900		Rio Itajaí Do Oeste	Barragem Taió Montante	C	-27,0397	-50,1242	35	648,0	167,3	-26,9433	-50,1770	1.637	15,34	0,6965	6,83
2	83069900		Rio Trombudo	Vila Nova	C	-27,3119	-49,7964	35	397,0	124,5	-27,4683	-49,8294	1.581	10,43	0,6512	6,50
3	83101000		Rio Caeté	Baixo Caeté	C	-27,705	-49,3344	13	165,0	69,4	-27,7847	-49,2739	1.672	15,17	0,6319	6,50
4	83105000		Rio Itajaí Do Sul	Saltinho	C	-27,6839	-49,3631	44	434,0	112,0	-27,7467	-49,2775	1.647	14,81	0,6545	6,50
5	83250000		Rio Itajaí Do Sul	Ituporanga	C	-27,3986	-49,6058	90	1.650,0	279,5	-27,6105	-49,4737	1.552	12,65	0,6915	6,42
6	83300200	83300002	Rio Itajaí-Açu	Rio Do Sul - Novo	C	-27,2078	-49,6292	74	5.160,0	574,1	-27,3211	-49,7998	1.578	11,92	0,7006	6,60
7	83500000	83500002	Rio Itajaí-Açu	Apiúna - Régua Nova	C	-27,0375	-49,3919	92	9.070,0	723,6	-27,1068	-49,8100	1.564	12,41	0,6911	6,55
8	83660000		Rio Benedito	Benedito Novo	C	-26,7875	-49,3642	79	717,0	183,0	-26,7498	-49,5007	1.618	12,37	0,6733	6,50
9	83675000		Rio Dos Cedros	Arroeira	C	-26,7414	-49,2711	41	536,0	157,3	-26,6131	-49,3721	1.754	12,74	0,7151	6,50
10	83677000	83680000	Rio Benedito	Timbó Novo	C	-26,8297	-49,2719	85	1.600,0	236,7	-26,7068	-49,4310	1.678	12,66	0,6214	6,50
11	83690000		Rio Itajaí-Açu	Indaial	C	-26,8911	-49,235	86	9.850,0	793,7	-27,1029	-49,7715	1.566	12,67	0,6920	6,55
12	83892990		Rio Itajaí-Mirim	Salseiro	C	-27,33	-49,33	29	286,0	105,9	-27,4260	-49,3717	1.504	11,84	0,6684	6,50
13	83892998	83893000	Rio Itajaí-Mirim	Botuvera-Montante	C	-27,1972	-49,0878	40	827,0	207,1	-27,3162	-49,3060	1.539	15,77	0,6452	6,50
14	83900000		Rio Itajaí-Mirim	Brusque (Pcd)	C	-27,1006	-48,9167	21	1.240,0	291,0	-27,2590	-49,1995	1.590	16,31	0,6806	6,47
15	83345000		Rio Itajaí Do Norte Ou Hercílio	Barra Do Prata	V	-26,6983	-49,8311	37	1.430,0	235,0	-26,6084	-50,0429	1.604	12,95	0,7228	6,50
16	83520000		Rio Itajaí-Açu	Warnow	V	-26,9417	-49,2897	35	9.790,0	791,0	-27,1038	-49,7741	1.566	12,69	0,6924	6,55
17	83050000		Rio Itajaí Do Oeste	Taió	ACF	-27,1139	-49,9953	89	1.570,0	245,0	-27,0445	-50,1382	1.626	13,35	0,7040	6,90
18	83440000		Rio Itajaí Do Norte Ou Hercílio	Ibirama	ACF	-27,0539	-49,5167	87	3.330,0	396,7	-26,7717	-49,8763	1.554	13,08	0,6816	6,50
19	83800002		Rio Itajaí-Açu	Blumenau (Pcd)	ACF	-26,9186	-49,0656	71	11.803,0	866,8	-27,0398	-49,6934	1.592	12,74	0,7076	6,54
20	83880000		Rio Luiz Alves	Luiz Alves	ACF	-26,7239	-48,9311	66	104,0	89,2	-26,7181	-48,9784	1.809	14,61	1,0314	6,50

Nota: **CC** indica código complementar, referindo-se às estações cujas séries históricas foram agregadas ao posto fluviométrico da mesma linha; na coluna **Tipo**, **C** designa postos de Calibração, **V** postos de validação e **ACF** postos com autocorrelação significativa para o lag 1; **A** é a área de drenagem; **Lat_c** e **Long_c** trazem as coordenadas do centroide da área de drenagem; **P** é a precipitação anual, **S** é a declividade do curso d'água principal; **DD** é densidade de drenagem e **QSUB** é a produção do aquífero.