

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

Ana Clara de Sousa Matos

**ESTIMAÇÃO BAYESIANA DOS PARÂMETROS DE MODELOS HIDROLÓGICOS A
PARTIR DO EMPREGO DE ASSINATURAS HIDROLÓGICAS**

Belo Horizonte

2021

Ana Clara de Sousa Matos

**ESTIMAÇÃO BAYESIANA DOS PARÂMETROS DE MODELOS HIDROLÓGICOS A
PARTIR DO EMPREGO DE ASSINATURAS HIDROLÓGICAS**

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

Área de concentração: Recursos Hídricos

Linha de pesquisa: Modelagem de Processos Hidrológicos

Orientador: Francisco Eustáquio Oliveira e Silva

Belo Horizonte

2021

M433e

Matos, Ana Clara de Sousa.

Estimação bayesiana dos parâmetros de modelos hidrológicos a partir do emprego de assinaturas hidrológicas [recurso eletrônico] / Ana Clara de Sousa Matos. – 2021.

1 recurso online (xi, 117 f.: il., color.) : pdf.

Orientador: Francisco Eustáquio Oliveira e Silva.

Dissertação (mestrado) - Universidade Federal de Minas Gerais, Escola de Engenharia.

Apêndices: f. 95-117.

Inclui bibliografia.

Exigências do sistema: Adobe Acrobat Reader.

1. Engenharia sanitária - Teses. 2. Recursos hídricos - Desenvolvimento - Teses. 3. Hidrologia - Modelos - Teses. 4. Teoria bayesiana de decisão estatística - Teses. I. Silva, Francisco Eustáquio Oliveira e. 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

["Estimação Bayesiana dos parâmetros de um modelo hidrológico a partir do emprego de assinaturas hidrológicas"]

[Ana Clara de Sousa Matos]

Dissertação de Mestrado] defendida e aprovada, no dia [31 de agosto 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. Pedro Luiz Borges Chaffe] - Membro Externo]

[UFSC]

[Prof. Veber Afonso Figueiredo Costa] - Membro Interno]

[UFMG]

[Prof. Francisco Eustáquio Oliveira Silva] - Orientador]

[UFMG]

APROVADA PELO COLEGIADO DO PPG SMARH

Sonaly Cristina Rezende Borges de Lima - Coordenadora

Belo Horizonte, 31 de agosto de 2021.

Documento assinado eletronicamente por **Francisco Eustaquio Oliveira e Silva, Professor do**



Magistério Superior, em 01/09/2021, às 10:26, 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, Servidor(a)**, em 01/09/2021, às 10:55, 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 **Pedro Luiz Borges Chaffe, Usuário Externo**, em 01/09/2021, às 11:25, 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 **Sonaly Cristina Rezende Borges de Lima, Coordenador(a) de curso de pós-graduação**, em 25/10/2021, às 09:18, 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 **0935845** e o código CRC **9EE74AE5**.

AGRADECIMENTOS

Agradeço aos meus pais, pelo apoio e amor incondicionais e pelos inúmeros sacrifícios feitos em benefício da minha educação.

Aos meus avós, pelo abraço sempre apertado e as mensagens sempre carinhosas. À minha avó, especialmente, por me mostrar que poder estudar e se dedicar aos próprios objetivos pessoais e profissionais ainda é um privilégio de poucas mulheres.

Ao Carlos, por compreender as minhas numerosas ausências e por ser a paz de que a minha natureza agitada precisa.

Ao meu irmão, por dividir a saudade de casa comigo.

Ao meu orientador, pela constante dedicação, envolvimento e apoio nesse período. Mais do que isso, obrigada por se fazer presente mesmo diante dos enormes desafios que a vida impôs.

Aos demais professores do PPG-SMARH, pelas valiosas lições no decorrer das disciplinas.

À equipe da secretaria do Programa e aos demais funcionários do EHR, pelo suporte sempre necessário.

Aos amigos de turma do PPG-SMARH, pelos momentos de alegria em meio às disciplinas.

Aos amigos da Walm e da Pimenta, pela companhia e pelas muitas caronas durante o período de disciplinas.

Aos professores avaliadores, pela leitura atenciosa e pelos comentários enriquecedores.

RESUMO

A modelagem de processos hidrológicos, apesar dos avanços tecnológicos observados nas últimas décadas, ainda conta com diversos desafios. Dentre eles, destacam-se a análise das incertezas relacionadas à modelagem e à predição em bacias não monitoradas. Este trabalho avaliou a aplicabilidade de assinaturas hidrológicas e de métodos de Aproximação Computacional Bayesiana para fins de estimação de parâmetros de modelos hidrológicos. O objetivo final era verificar a viabilidade da consideração de assinaturas regionalizadas para predição em bacias não monitoradas. Para tanto, foram conduzidas simulações em escalas de tempo horária e mensal, empregando os modelos GR4H (*Génie Rural à 4 paramètres Horaires*) e GR2M (*Génie Rural à 2 paramètres Mensuel*), respectivamente. Buscou-se aproximar as assinaturas calculadas a partir das séries simuladas às estimadas por meio dos registros pluvio-fluviométricos das estações analisadas. Foram consideradas, ainda, combinações de diferentes assinaturas. Os resultados foram avaliados em relação ao desempenho e às distribuições *a posteriori* dos parâmetros do modelo que considerava uma abordagem Bayesiana “clássica”, i.e., com emprego de amostrador MCMC (*Differential Evolution Adaptive Metropolis* – DREAM) e especificação de uma função de verossimilhança (função de verossimilhança generalizada), no domínio do tempo. Como principais resultados, foram encontrados índices de desempenho de modelagem, para as simulações do modelo mensal, muito próximos dos calculados para a solução paradigma, no domínio do tempo. Para o modelo horário, no entanto, foi observada uma redução desses índices em relação à solução paradigma. A combinação de assinaturas, nessa escala de simulação, mostrou-se bastante vantajosa, melhorando a capacidade de representação dos picos e da recessão dos hidrogramas. Além disso, as distribuições *a posteriori* dos parâmetros dos modelos sugerem, para ambas as escalas de simulação, problemas de equifinalidade e identificabilidade. Por fim, foi conduzida uma avaliação sumária do potencial de regionalização das assinaturas em escala horária, que demonstrou uma piora da representação da variabilidade dos hidrogramas, mas índices de desempenho semelhantes aos observados para as assinaturas locais. Como conclusão, observou-se grande potencial para a utilização da metodologia proposta, desde que haja uma criteriosa seleção das assinaturas e poder computacional compatível.

Palavras-chave: modelagem hidrológica; assinaturas hidrológicas; aproximação computacional Bayesiana.

ABSTRACT

Despite the technological advances in the last decades, modelling hydrological processes still faces several challenges. Among them, we highlight uncertainty analysis and prediction in ungauged basins. This work evaluates the applicability of streamflow signatures and ABC methods to estimating parameters of hydrological models. The final purpose was to verify the viability of hydrological signatures to allow prediction in ungauged basins. To do so, simulations were realized on hourly and monthly timescales, using GR4H (*Génie Rural à 4 paramètres Horaires*) and GR2M (*Génie Rural à 2 paramètres Mensuel*) models, respectively, intending to approximate the signatures calculated from the simulated series to those estimated for the monitoring series of selected gauges. The combination of different signatures was also evaluated. Model performance and *posterior* distributions of model parameters were evaluated concerning the results of a “classical” Bayesian approach, using a MCMC sampler (Differential Evolution Adaptive Metropolis – DREAM) and specification of the likelihood function (generalized likelihood function), in the time domain. The main result, in the monthly timescale, was that similar performance measures were found both in the signature and in the time domain. In the hourly model, there was a reduction in the values of these performance indices. The combination of different signatures, in this timescale, proved to be advantageous, improving the ability of the model to reproduce the peaks and the recession of the hydrograph. Furthermore, the *posterior* distributions of model parameters suggest the problems of equifinality and identifiability in both timescales. Finally, a simplified assessment of the potential for regionalization of signatures was conducted on an hourly scale, which showed worse representations of the variability of the hydrograph, but performance measures close to those calculated to local signatures. As a conclusion, it was observed great potential for using the proposed methodology, as long as there are a careful selection of signatures and compatible computational power.

Keywords: hydrological modeling; hydrological signatures; approximate Bayesian computation.

LISTA DE FIGURAS

Figura 3.1 – Síntese do procedimento de calibração de um modelo hidrológico.	19
Figura 4.1 – Fluxograma geral do trabalho.	41
Figura 4.2 – Localização da bacia em estudo.	43
Figura 4.3 – Mapa de declividades percentuais encontradas na bacia.	44
Figura 4.4 – Mapa de solos encontrados na bacia.	44
Figura 4.5 – Série ponderada de precipitações horárias registradas nas estações localizadas na bacia do ribeirão Serra Azul (Alto da Boa Vista, Fazenda Laranjeiras – Jusante, Jardim e Serra Azul).	45
Figura 4.6 – Série histórica de vazões horárias registrada na estação Jardim.	46
Figura 4.7 – Localização da bacia em estudo e das estações analisadas.	48
Figura 4.8 – Croqui esquemático da estrutura do modelo GR4H.	52
Figura 4.9 – Croqui esquemático da estrutura do modelo GR2M.	58
Figura 5.1 – Precipitações médias mensais para as estações hidrométricas adotadas no estudo.	67
Figura 5.2 – Vazões médias mensais para as estações hidrométricas adotadas no estudo.	67
Figura 5.3 – Precipitações e vazões médias mensais para as estações hidrométricas adotadas no estudo.	68
Figura 5.4 – Série temporal de vazões observadas e simuladas no domínio do tempo, para o período de validação.	75
Figura 5.5 – Série temporal de vazões observadas e simuladas no domínio das assinaturas, para o período de validação - P_{FDC}	75
Figura 5.6 – Série temporal de vazões observadas e simuladas no domínio das assinaturas, para o período de validação - S_{FDC}	76
Figura 5.7 – Série temporal de vazões observadas e simuladas no domínio das assinaturas, para o período de validação - F_I	76
Figura 5.8 – Vazões simuladas <i>versus</i> vazões observadas (período de validação)	77
Figura 5.9 – Série temporal de vazões observadas e simuladas no domínio das assinaturas, para o período de validação - $P_{FDC} + F_I$	79
Figura 5.10 – Série temporal de vazões observadas e simuladas no domínio das assinaturas, para o período de validação - $c_a + F_I$	79
Figura 5.11 – Correlogramas total e parcial dos resíduos no domínio do tempo, empregando o algoritmo DREAM e a função de verossimilhança generalizada.	80
Figura 5.12 – Comportamento dos resíduos no domínio do tempo, empregando o algoritmo DREAM e a função de verossimilhança generalizada	81
Figura 5.13 – Gráficos de dispersão dos parâmetros do modelo.	83
Figura 5.14 – Distribuições marginais <i>a posteriori</i> dos parâmetros do modelo GR4H.	86
Figura 5.15 – Distribuições marginais <i>a posteriori</i> dos parâmetros do modelo GR4H para as combinações propostas para as assinaturas.	87
Figura 5.16 – Distribuições <i>a posteriori</i> dos parâmetros do modelo GR4H para as combinações propostas para as assinaturas.	88
Figura 5.17 – Série temporal de vazões observadas e simuladas no domínio do tempo.	90
Figura 5.18 – Série temporal de vazões observadas e simuladas no domínio das assinaturas - P_{FDC}	91
Figura 5.19 – Série temporal de vazões observadas e simuladas no domínio das assinaturas - c_a	91
Figura 5.20 – Correlogramas total e parcial dos resíduos no domínio do tempo, empregando o algoritmo DREAM e a função de verossimilhança generalizada.	92
Figura 5.21 – Comportamento dos resíduos no domínio do tempo.	93
Figura 5.22 – Gráficos de dispersão dos parâmetros do modelo.	95

Figura 5.23 – Distribuições marginais dos parâmetros do modelo GR2M.....	96
Figura 5.24 – Série temporal de vazões observadas e simuladas no domínio das assinaturas, para o período de validação - q_{mean}	99
Figura 5.25 – Série temporal de vazões observadas e simuladas no domínio das assinaturas, para o período de validação - P_{FDC}	99
Figura 5.26 – Série temporal de vazões observadas e simuladas no domínio das assinaturas, para o período de validação - S_{FDC}	100
Figura 5.27 – Série temporal de vazões observadas e simuladas no domínio das assinaturas, para o período de validação - c_a	100

LISTA DE TABELAS

Tabela 3.1 – Exemplos de aplicações de métodos ABC em estudos de modelagem hidrológica	30
Tabela 3.2 – Exemplos de assinaturas comumente adotadas (Fonte: adaptado de Addor <i>et al.</i> , 2018)	34
Tabela 4.1 – Estações fluviométricas com registros diários.....	46
Tabela 4.2 – Estações pluviométricas com registros diários.....	47
Tabela 4.3 – Estatísticas amostrais para as séries de vazão observadas nas estações selecionadas	47
Tabela 4.4 – Estatísticas amostrais para as séries anuais de precipitação observadas nas estações selecionadas	48
Tabela 4.5 – Assinaturas hidrológicas adotadas.....	50
Tabela 4.6 – Parâmetros do modelo GR4H (Fonte: adaptado de PERRIN <i>et al.</i> , 2003).....	51
Tabela 4.7 – Parâmetros do modelo GR2M (Fonte: adaptado de MOUELHI <i>et al.</i> , 2006)	57
Tabela 4.8 – Parâmetros da função de verossimilhança generalizada a serem estimados ...	60
Tabela 4.9 – Variáveis consideradas para utilização do método DREAM (Fonte: adaptado de Silva, 2015).....	61
Tabela 4.10 – Intervalos de variação dos parâmetros do modelo GR4H admitidos <i>a priori</i> ...63	
Tabela 4.11 – Intervalos de variação dos parâmetros do modelo GR2M admitidos <i>a priori</i> ...64	
Tabela 4.12 – Intervalos de variação admitidos <i>a priori</i> para os parâmetros da função de verossimilhança generalizada – Modelagem horária	64
Tabela 4.13 – Intervalos de variação admitidos <i>a priori</i> para os parâmetros da função de verossimilhança generalizada – Modelagem mensal.....	64
Tabela 4.14 – Valores dos parâmetros empregados na utilização do método DREAM.....	65
Tabela 4.15 – Valores dos parâmetros empregados na utilização do algoritmo SABC	66
Tabela 4.16 – Bibliotecas R utilizadas neste trabalho.....	66
Tabela 5.1 – Assinaturas hidrológicas calculadas para as estações selecionadas – Séries horárias e diárias	71
Tabela 5.2 – Assinaturas hidrológicas calculadas para as estações selecionadas – Séries mensais.....	71
Tabela 5.3 – Índices de desempenho da modelagem em escala horária - Período de calibração.....	72
Tabela 5.4 – Índices de desempenho da modelagem em escala horária - Período de validação.....	72
Tabela 5.5 – Percentual das observações contidas no intervalo de credibilidade de 95%....	72
Tabela 5.6 – Índices de desempenho - Combinações de assinaturas – Período de calibração	78
Tabela 5.7 – Índices de desempenho - Combinações de assinaturas – Período de validação	78
Tabela 5.8 – Percentual das observações contidas no intervalo de credibilidade de 95%....	78
Tabela 5.9 – Índices de desempenho da modelagem em escala mensal - Período de calibração.....	89
Tabela 5.10 – Índices de desempenho da modelagem em escala mensal - Período de validação.....	89
Tabela 5.11 – Percentual das observações contidas no intervalo de credibilidade de 95%	89
Tabela 5.12 – Coeficiente de correlação de Pearson entre os parâmetros θ_1 e θ_2 do modelo GR2M.....	93
Tabela 5.13 – Índices de desempenho da modelagem a partir das assinaturas regionais - Período de calibração.....	97

Tabela 5.14 – Índices de desempenho da modelagem a partir das assinaturas regionais - Período de validação	97
Tabela 5.15 – Percentual das observações contidas no intervalo de credibilidade de 95%	97

LISTA DE ABREVIATURAS E SIGLAS

ABC	<i>Approximate Bayesian Computation</i>
BEAR	<i>Bayesian Error Analysis with Reordering</i>
BFI	<i>Baseflow index</i>
DREAM	<i>Differential Evolution Adaptive Metropolis</i>
GL	<i>Generalized likelihood function</i>
GLUE	<i>Generalized Likelihood Uncertainty Estimation</i>
GR2M	<i>Génie Rural à 2 paramètres Mensuel</i>
GR4H	<i>Génie Rural à 4 paramètres Horaires</i>
IGAM	Instituto Mineiro de Gestão das Águas
KGE	Coeficiente de Kling-Gupta Efficiency
MCMC	Monte Carlo via cadeias de Markov
NSE	<i>Nash-Sutcliffe efficiency</i>
RMBH	Região Metropolitana de Belo Horizonte
SA	<i>Simulated Annealing</i>
SABC	<i>Simulated Annealing – Approximate Bayesian Computation</i>
SEP	<i>Skew Exponential Power</i>
TOSSH	<i>Toolbox for Streamflow Signatures in Hydrology</i>

LISTA DE SÍMBOLOS

θ	Parâmetros do modelo
Θ	Espaço paramétrico
\tilde{y}	Observações de uma variável aleatória
\mathcal{Y}	Espaço amostral das observações
\mathcal{M}	Modelo
y	Pseudo-observações de uma variável aleatória
ε	Nível de tolerância aceito para aproximação em métodos ABC
η	Estatística de resumo de interesse
ρ	Função da distância para o algoritmo SABC
u	Realização de uma distribuição uniforme
\tilde{P}	Observações de precipitação
\tilde{Q}	Observações de vazão
Z	Vetor de vazões médias considerado para o cálculo da verossimilhança
δ	Vetor de resíduos aleatórios com média zero – GL
t	Intervalo de tempo
i	Dados de entrada – GL
μ_M	Viés na vazão modelada – GL
θ_e	Parâmetros de caracterização dos resíduos
a_t	Erros aleatórios independentes – GL
σ_t	Desvio padrão no tempo t – GL
ξ	Assimetria – GL
β	Curtose – GL
σ_0	Coefficiente linear para consideração da heterocedasticidade – GL
σ_1	Coefficiente angular para consideração da heterocedasticidade – GL
ϕ_i	Coefficiente de autocorrelação – GL
Q	Série de vazões simuladas
$g(q^{(j)})$	Assinaturas calculadas a partir da série modelada
\tilde{g}	Assinaturas calculadas a partir dos dados
Ψ	Função auxiliar para consideração de assinaturas potencialmente vetoriais
r	Coefficiente de correlação de Pearson
α	Razão entre desvios padrões dos valores simulados e observados – KGE
γ	Razão entre a média dos valores simulados e a média dos valores observados – KGE

ε	Erro aleatório a ser incorporado para consideração de um modelo probabilístico
$h(\theta_h, x)$	Série simulada a partir do modelo determinístico
$Q(\theta, x)$	Série simulada a partir do modelo probabilístico
W_t	Inovações de um modelo AR
c_a	Coefficiente de escoamento anual
F_I	Taxa de mudança
q_{mean}	Vazão média mensal
S_{FDC}	Declividade da curva de permanência
P_{FDC}	Percentis da curva de permanência – Q_5 , Q_{10} , Q_{90} e Q_{95}

SUMÁRIO

1	INTRODUÇÃO.....	15
2	OBJETIVOS.....	17
2.1	OBJETIVO GERAL	17
2.2	OBJETIVOS ESPECÍFICOS.....	17
3	REVISÃO DA LITERATURA	18
3.1	CONSIDERAÇÕES SOBRE MODELOS DE SIMULAÇÃO HIDROLÓGICA	18
3.1.1	<i>Classificação de modelos chuva-vazão</i>	<i>19</i>
3.1.2	<i>Incertezas na modelagem chuva-vazão</i>	<i>21</i>
3.2	APPROXIMATE BAYESIAN COMPUTATION (ABC).....	24
3.2.1	<i>Emprego de métodos ABC em estudos de modelagem hidrológica.....</i>	<i>28</i>
3.3	ASSINATURAS HIDROLÓGICAS	32
3.3.1	<i>Seleção de assinaturas e suas aplicações</i>	<i>34</i>
3.3.2	<i>Emprego de assinaturas hidrológicas para fins de modelagem hidrológica.....</i>	<i>36</i>
4	MATERIAL E MÉTODOS	40
4.1	ÁREA DE ESTUDO E INFORMAÇÕES HIDROLÓGICAS	41
4.1.1	<i>Caracterização da bacia do ribeirão Serra Azul.....</i>	<i>42</i>
4.1.2	<i>Estações pluviométricas, fluviométricas e evaporimétricas adotadas.....</i>	<i>45</i>
4.2	CÁLCULO DAS ASSINATURAS HIDROLÓGICAS.....	48
4.3	MODELOS HIDROLÓGICOS E DADOS DE ENTRADA DO MODELO.....	51
4.3.1	<i>Génie Rural à 4 paramètres Horaires – GR4H.....</i>	<i>51</i>
4.3.2	<i>Génie Rural à 2 paramètres Mensuel – GR2M</i>	<i>57</i>
4.4	ESTIMATIVA DOS PARÂMETROS DO MODELO PARA A ESTAÇÃO JARDIM NO DOMÍNIO DO TEMPO.....	58
4.5	ESTIMATIVA DOS PARÂMETROS DO MODELO CONSIDERANDO ASSINATURAS HIDROLÓGICAS	61
4.6	DADOS DE ENTRADA E PARÂMETROS DAS SIMULAÇÕES.....	63
5	RESULTADOS E DISCUSSÃO.....	67
5.1	CÁLCULO DAS ASSINATURAS HIDROLÓGICAS.....	67
5.2	DOMÍNIO TEMPO VERSUS DOMÍNIO DAS ASSINATURAS: ESCALA HORÁRIA	72
5.2.1	<i>Avaliação do desempenho da modelagem</i>	<i>72</i>
5.2.2	<i>Avaliação das distribuições a posteriori dos parâmetros do modelo.....</i>	<i>82</i>
5.3	DOMÍNIO TEMPO VERSUS DOMÍNIO ASSINATURAS: ESCALA MENSAL.....	89
5.3.1	<i>Avaliação do desempenho da modelagem</i>	<i>89</i>
5.3.2	<i>Avaliação das distribuições a posteriori dos parâmetros dos modelos.....</i>	<i>93</i>
5.4	DESEMPENHO DA MODELAGEM NO DOMÍNIO DAS ASSINATURAS INFERIDAS A PARTIR DE REGISTROS FLUVIOMÉTRICOS REGIONAIS	97
6	CONCLUSÕES E RECOMENDAÇÕES	101
	REFERÊNCIAS	104
	APÊNDICE I – DOMÍNIO TEMPO VERSUS DOMÍNIO DAS ASSINATURAS: ESCALA HORÁRIA	112
	APÊNDICE II – DOMÍNIO TEMPO VERSUS DOMÍNIO DAS ASSINATURAS: ESCALA MENSAL.....	123
	APÊNDICE III – ASSINATURAS INFERIDAS A PARTIR DE REGISTROS REGIONAIS	129

1 INTRODUÇÃO

Os modelos hidrológicos constituem um ferramental que permite a análise quantitativa, a extrapolação e a predição dos eventos (BEVEN, 2012), permitindo uma estimativa das variáveis envolvidas, em cenários ainda não observados. Apesar dos avanços tecnológicos observados nas últimas décadas, como o aumento da capacidade computacional, que impactou diretamente os estudos de modelagem de processos hidrológicos, vários ainda são os desafios a serem enfrentados. Dentre eles, destaca-se a consideração das incertezas associadas à modelagem, como às relacionadas aos dados de entrada, aos parâmetros e à própria estrutura do modelo.

Em geral, a avaliação das incertezas da modelagem, especialmente a relacionada à estimação dos parâmetros, é incorporada por meio da inferência Bayesiana, que considera que os parâmetros dos modelos hidrológicos são variáveis aleatórias, cuja distribuição *a posteriori* é estimada a partir da incerteza *a priori* e do comportamento probabilístico dos resíduos.

Nesse cenário, pode-se proceder a análise sob um ponto de vista “informal”, que prioriza a semelhança entre saídas do modelo e observações, em lugar da especificação de um modelo probabilístico formal. Um exemplo desta abordagem é o método GLUE (*Generalized Likelihood Uncertainty Estimation* – BEVEN e BINLEY, 1992). Outra possibilidade é o emprego de distribuições teóricas na especificação do modelo de resíduos, em uma abordagem dita “formal” para caracterização do comportamento probabilístico dos resíduos. Todavia, a função de verossimilhança pode ser de difícil tratamento. Nesses casos, os métodos de aproximação computacional Bayesiana (“ABC” – *Approximate Bayesian Computation*) são uma alternativa para avaliação da distribuição *a posteriori* do ponto de vista formal.

Segundo Kavetski *et al.* (2018), o conjunto de métodos ABC preconiza a avaliação das saídas do modelo, a partir de estatísticas de resumo, permitindo, desse modo, estimar a incerteza paramétrica. As estatísticas de resumo utilizadas para a aproximação, por sua vez, podem ser desde simples medidas de tendência central até conjuntos de diferentes assinaturas hidrológicas.

As assinaturas hidrológicas buscam sintetizar informações acerca do comportamento da bacia, como a identificação de processos dominantes e a representação de diferentes aspectos do hidrograma. Possuem as mais diversas aplicações, dentre as quais destacam-se estudos como os de Sahraei *et al.* (2020) e Fenicia *et al.* (2018), por exemplo, que fazem uso de assinaturas hidrológicas para calibração dos parâmetros de modelos. Dentre as vantagens do uso de assinaturas para a estimação de parâmetros, destaca-se a possibilidade de predição em bacias não monitoradas, por meio da regionalização das assinaturas (MCMILLAN, 2021).

Além da metodologia empregada para estimação dos parâmetros, outro ponto a ser levado em consideração em estudos de modelagem hidrológica diz respeito à escala de tempo adotada para simulação, uma vez que modelos com menor escala temporal normalmente estão associados a uma estrutura mais complexa e, possivelmente, a uma maior incerteza paramétrica (MASSMANN, 2020). Especialmente no contexto brasileiro, em que a rede de monitoramento hidrométrico é, por vezes, escassa ou falha, a seleção de um intervalo adequado de simulação pode representar a viabilidade ou não da modelagem, dadas as limitações relacionadas à aquisição de séries temporais com maior discretização.

O presente trabalho insere-se nesse cenário e avalia o emprego de métodos ABC e de assinaturas hidrológicas para fins de estimação de parâmetros de modelos hidrológicos, com o objetivo final de lançar luz sobre a predição em bacias não monitoradas, por meio da consideração de assinaturas regionalizadas. Em um primeiro momento, é conduzida uma abordagem “clássica” para estimação dos parâmetros do modelo, sob o paradigma Bayesiano. Para tanto, é empregada a função de verossimilhança generalizada e o algoritmo DREAM (*Differential Evolution Adaptive Metropolis*) de simulação MCMC para estimação dos parâmetros do modelo GR4H (*Génie Rural à 4 paramètres Horaires*). O objetivo é o estabelecimento de um *benchmark* para avaliação das etapas posteriores.

Na sequência, utilizando dados horários para estações hidrométricas localizadas na bacia do ribeirão Serra Azul, os parâmetros do modelo GR4H são estimados usando o algoritmo ABC denominado SABC (acrônimo para *Simulated Annealing* e *Approximate Bayesian Computation*). As distribuições *a posteriori* dos parâmetros são avaliadas, além da incerteza paramétrica associada e do desempenho da modelagem, em relação à solução paradigma. Em seguida, com o propósito de se identificar a influência da escala temporal sob os resultados, as mesmas avaliações são conduzidas em escala mensal, adotando o modelo GR2M (*Génie Rural à 2 paramètres Mensuel*). Por fim, tendo em vista a possibilidade de regionalização das assinaturas, permitindo a predição em bacias não monitoradas, é realizada uma avaliação sumária do desempenho da modelagem quando da estimação dos parâmetros sob a consideração de assinaturas “regionais”, i.e., estimadas a partir de registros de estações vizinhas à bacia avaliada.

O presente documento está estruturado em 5 capítulos complementares a este, de natureza introdutória. No segundo capítulo, são apresentados os objetivos deste estudo, sendo a literatura referente aos principais temas abordados apresentada no capítulo 3. A metodologia empregada é descrita no capítulo 4. Os resultados alcançados são demonstrados no capítulo 5. Por fim, no capítulo 6, são apresentadas as conclusões e as recomendações deste trabalho.

2 OBJETIVOS

2.1 *Objetivo geral*

O objetivo geral deste trabalho é avaliar a aplicabilidade de assinaturas hidrológicas e de métodos de Aproximação Computacional Bayesiana para fins de estimação de parâmetros de modelos hidrológicos, como forma de viabilizar a predição em bacias não monitoradas a partir do emprego de assinaturas regionalizadas.

2.2 *Objetivos específicos*

Os objetivos específicos são:

- Caracterizar a qualidade do ajuste e a incerteza paramétrica estimada a partir das séries de vazões modeladas nos domínios do tempo e das assinaturas;
- Analisar o impacto da combinação de diferentes assinaturas hidrológicas sobre os resultados da modelagem;
- Caracterizar a influência das assinaturas hidrológicas quando o fenômeno é modelado em diferentes escalas temporais; e
- Avaliar o impacto da consideração de assinaturas hidrológicas estimadas a partir de informações regionais sobre os resultados da modelagem.

3 REVISÃO DA LITERATURA

3.1 *Considerações sobre modelos de simulação hidrológica*

Um sistema hidrológico compreende as dinâmicas e os processos relacionados à água e a suas interconexões em determinada escala, que pode variar desde avaliações a nível molecular até a consideração de todo o ciclo hidrológico (JAYAWARDENA, 2013). Os mecanismos de produção das saídas do sistema hidrológico dependem de diversos aspectos, como das entradas e da física que governa o comportamento do sistema.

Essa complexa dinâmica é de difícil representação, fazendo necessária a simplificação de determinados aspectos, em busca do estabelecimento de relações entre entradas e as respectivas saídas do sistema. Neste contexto, o ciclo hidrológico, em um sentido amplo, pode ser compreendido como um sistema fechado composto por sistemas menores e abertos, a nível de bacia hidrográfica, que apresentam processos de transferência entre si (JAYAWARDENA, 2013). Dentre esses diversos microsistemas e suas interações, destaca-se o interesse, por parte dos hidrólogos, quanto ao estabelecimento de relações entre a precipitação registrada em uma determinada bacia e as correspondentes vazões observadas no leito fluvial.

De acordo com Beven (2012), muitas vezes, modelos hidrológicos são aplicados com objetivo puramente técnico, a fim de melhor compreender os sistemas hidrológicos, contrapondo variáveis simuladas e observadas. Contudo, o autor destaca a importância da predição para a tomada de decisão relacionada a processos hidrológicos, como na gestão dos recursos hídricos, no controle de cheias e na mitigação da contaminação.

Outras possíveis aplicações dos modelos hidrológicos são estudos para a restauração de rios, para o manejo das águas subterrâneas, para irrigação de culturas e para projetos de práticas de conservação dos solos. Destaca-se também o emprego de modelos hidrológicos em bacias não-monitoradas, por meio da regionalização dos parâmetros do modelo ou de suas variáveis.

Os modelos hidrológicos são descritos por parâmetros, que, por meio das equações e premissas admitidas, buscam sintetizar o comportamento hidrológico da bacia. De acordo com Singh (1995), os parâmetros da maioria dos modelos podem ser agrupados em dois conjuntos: os físicos, que podem ser avaliados a partir de características fisiográficas da bacia, ou os de processos, relacionados a equacionamentos matemáticos que buscam descrever de forma simplificada os processos físicos presentes na bacia. Esses últimos guardam relação com as propriedades físicas da bacia, mas não podem ser facilmente mensurados ou quantificados a partir de tais propriedades.

O processo de estimativa dos parâmetros do modelo é denominado calibração, e pode ser conduzido de forma “manual”, por meio de um processo de tentativa e erro, ou automatizada. Nesse sentido, o problema da calibração, no contexto de modelagem hidrológica, está relacionado à busca de um conjunto de parâmetros que melhor aproxime as saídas do modelo ao comportamento observado no sistema, conforme apresentado na Figura 3.1.

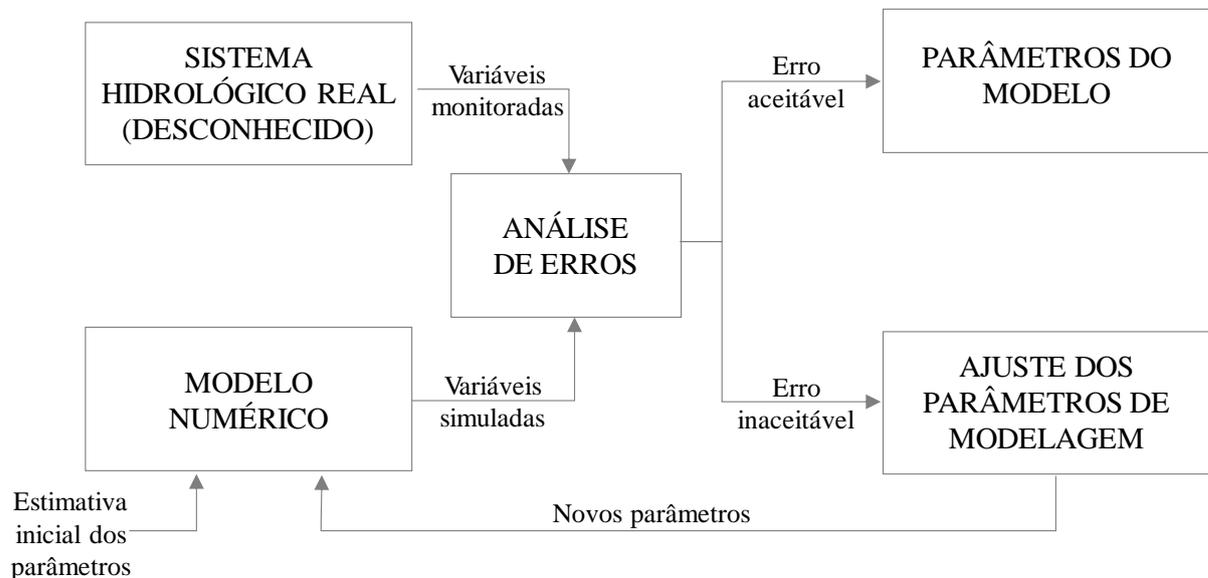


Figura 3.1 – Síntese do procedimento de calibração de um modelo hidrológico.

Fonte: adaptado de FREEZE e CHERRY (1979) *apud* NORTH *et al.* (2014).

A abordagem clássica para calibração consiste na seleção de um conjunto “ótimo” de parâmetros que minimize o somatório dos erros de modelagem, ou seja, as diferenças entre variáveis simuladas e seus valores observados. Dessa forma, os parâmetros do modelo são considerados invariantes no tempo (DENG *et al.*, 2018) e empregados para predição em períodos distintos do usado para calibração.

A qualidade do ajuste pode ser avaliada por meio de comparação visual, como em gráficos de dispersão das variáveis simuladas e observadas ou de distribuição dos resíduos, ou de métricas e índices de desempenho (NORTH *et al.*, 2014), como o erro médio quadrático ou os coeficientes de Nash-Sutcliffe (NASH e SUTCLIFFE, 1970) ou de eficiência de Kling-Gupta (GUPTA *et al.*, 2009).

3.1.1 Classificação de modelos chuva-vazão

Dada a complexidade dos processos hidrológicos, reflexo da interdependência de diversos processos que ocorrem na atmosfera, no solo e nos corpos de água em diferentes escalas temporais e espaciais, não há consenso acerca da identificação e representação dos processos nos modelos hidrológicos (WAGENER *et al.*, 2004).

A descrição qualitativa do ciclo da água é, portanto, complexa e, segundo Beven (2012), dependente das percepções do modelador e influenciadas, sobretudo, por suas experiências prévias acerca dos sistemas a serem modelados. O ajuste de equações e teorias matemáticas às percepções acerca dos fenômenos modelados introduz um caráter quantitativo à análise, caracterizando um modelo conceitual (ou paramétrico). Nesse tipo de modelo, os processos frequentemente são representados por meio de reservatórios fictícios, que podem ou não ser lineares. O modelo pode ser codificado, a fim de permitir análises computacionais, tendo soluções analíticas. Nos casos em que as equações empregadas não podem ser assim resolvidas, esse tipo de modelo pode necessitar de métodos de cálculo numérico para aproximações sucessivas do resultado. A transformação de equações matemáticas em algoritmos computacionais pode ser uma importante fonte de erro na modelagem (BEVEN, 2012).

Além dos conceituais, os modelos hidrológicos podem ser classificados, com relação à estrutura de representação dos fenômenos, em físicos ou empíricos (caixa-preta). Modelos físicos se baseiam nas equações de conservação de massa, *momentum* e energia, empregando métodos numéricos como o de diferenças finitas (JAISWAL *et al.*, 2020). Os modelos empíricos, por sua vez, são caracterizados pela dedução da estrutura do modelo e dos valores dos parâmetros a partir de séries temporais disponíveis, não guardando relação com uma descrição física do fenômeno de transformação da chuva em vazão.

Os modelos hidrológicos podem ser classificados também com relação a diversos outros aspectos para além da representação dos fenômenos, como de acordo com as escalas temporal e espacial de simulação, bem como da natureza dos processos envolvidos. Quanto à escala de tempo, os modelos podem ser contínuos, i.e., que buscam representar o balanço hídrico na bacia “continuamente”, inclusive, quando da ausência de precipitação, ou por eventos, quando o objetivo é a simulação da resposta da bacia a um “evento” pluviométrico específico.

Com relação à escala espacial, os modelos podem ser concentrados ou distribuídos. No primeiro caso, a bacia é tratada como uma unidade uniforme, em que as variáveis de entrada e as características da bacia são representadas por valores ponderados para a área de contribuição. Os modelos distribuídos, por sua vez, admitem a divisão da bacia em sub-regiões, sendo os fenômenos modelados de forma independente em cada uma delas. Encontra-se na literatura também a classificação em modelos semi-distribuídos (BEVEN, 2012). Segundo Jaiswal *et al.* (2020), modelos conceituais são geralmente concentrados, enquanto os físicos são essencialmente distribuídos.

Quanto à natureza dos processos, os modelos podem ser classificados como: determinísticos, quando os mesmos *inputs* geram séries modeladas idênticas; estocásticos, quando é considerada a aleatoriedade

decorrente das incertezas envolvidas na modelagem; ou híbridos, que combinam componentes determinísticas e estocásticas.

3.1.2 Incertezas na modelagem chuva-vazão

A calibração dos parâmetros de modelos hidrológicos e sua utilização estão sujeitas à incerteza, uma vez que são consideradas simplificações, premissas e equacionamentos que não são capazes de refletir completamente os processos hidrológicos envolvidos. Além disso, os dados de monitoramento utilizados para calibração também estão sujeitos a erros, como os de medição ou os relacionados à variabilidade natural, as chamadas incertezas aleatórias.

As incertezas aleatórias (ou estocásticas) são decorrentes da aleatoriedade dos fenômenos naturais, sendo independentes da metodologia adotada. Portanto, sua grandeza não pode ser alterada. Por sua vez, as incertezas epistêmicas são resultado do desconhecimento (de parte) dos processos envolvidos e, dessa forma, podem ser reduzidas por meio de experimentos e observações. Beven (2012), porém, comenta que a redução das incertezas epistêmicas pode acarretar uma maior complexidade dos modelos, sem necessariamente melhorar a predição de forma significativa.

Várias são as fontes de incerteza em um estudo de modelagem, fruto da heterogeneidade dos processos e das variáveis hidrológicas, geológicas e climáticas no tempo e no espaço. Dentre elas, destacam-se às associadas às condições iniciais e de contorno, aos dados de entrada do modelo, à sua estrutura e à estimativa de seus parâmetros.

Incerteza com relação aos dados de entrada do modelo

No que concerne aos dados de entrada do modelo, as incertezas estão relacionadas à variabilidade natural e aos erros de medição das variáveis. McMillan *et al.* (2018) mencionam os erros decorrentes de medições pontuais de variáveis hidrológicas, bem como os erros decorrentes da derivação de informação, como a estimativa de vazões a partir da curva-chave. Ainda, os autores citam as incertezas decorrentes da interpolação das variáveis, como a consideração de valores ponderados para a precipitação, e de efeitos de escala relacionados às variáveis. Estima-se que a incerteza associada aos dados de entrada varie de 10% a 40% (MCMILLAN *et al.*, 2018).

Além disso, de acordo com Tucci (1998), os volumes de precipitação observados tendem a apresentar maior variabilidade temporal e espacial do que a evaporação, o que resulta de fatores como topografia, umidade atmosférica e latitude.

Incerteza com relação à estrutura do modelo

Os modelos fazem uso de equações para representação dos complexos fenômenos simulados, sendo necessário o estabelecimento de premissas e de simplificações que tornem possível a estimação das variáveis de interesse. Além disso, há certo grau de dificuldade para representação dos efeitos de mudança temporal de variáveis, como do uso do solo, das características climáticas e da poluição.

De acordo com Moges *et al.* (2021), o desempenho do modelo depende significativamente de sua estrutura, o que pode acarretar uma predominância dessas incertezas em relação às demais. Os estudos de Troin *et al.* (2018) e de Højberg e Refsgaard (2005) corroboram essa observação, tendo demonstrado a dominância da incerteza estrutural especialmente quando o modelo é utilizado para períodos distintos daquele usado para a calibração. Rojas *et al.* (2008) estimam que a incerteza estrutural pode corresponder a até cerca de 30% da incerteza preditiva.

Incerteza com relação à estimação dos parâmetros

De acordo com Sorooshian e Gupta (1995), os principais problemas quanto à estimação de parâmetros de modelos estão relacionados a variações no processo de convergência, como a possibilidade de diversos mínimos locais ou de regiões de atração, à escolha da função-objetivo adotada ou aos diferentes graus de sensibilidade dos parâmetros, que podem estar correlacionados ou, ainda, se compensarem.

Em casos de modelos concentrados, outra fonte de incerteza diz respeito à representação uniforme das características da bacia, o que pode impactar a estimação dos parâmetros “ótimos”. Em bacias não monitoradas, podem ser empregados parâmetros regionalizados. Contudo, nessas circunstâncias, deve-se, idealmente, considerar as incertezas associadas às técnicas de regionalização, além de todas as envolvidas nas demais etapas de modelagem.

Moges *et al.* (2021) ressaltam que, mesmo que o modelo fosse uma representação exata do sistema hidrológico, poderiam existir incertezas em relação aos parâmetros, decorrentes daquelas associadas aos dados utilizados para calibração.

Este trabalho se propõe a quantificar, predominantemente, as incertezas paramétricas relacionadas à modelagem hidrológica, por meio de inferência Bayesiana.

Avaliação das incertezas

Por meio da avaliação das incertezas, é possível identificar limitações e propor melhorias com relação aos modelos (BULYGINA e GUPTA, 2009; FENICIA *et al.*, 2018), orientar a aquisição de novos dados (LU *et al.*, 2018) e quantificar a incerteza preditiva. A avaliação das incertezas requer a consideração de métodos para além da estrita utilização dos modelos hidrológicos (MOGES *et al.*, 2021), a fim de permitir uma análise quantitativa das diversas fontes de incerteza. Tal avaliação fornece uma quantificação da confiança de determinada variável hidrológica, seja ela observada ou inferida a partir de determinado modelo, e pode ser expressa pela estimativa do erro dessa variável ou pelo padrão de variabilidade de seu valor em um intervalo pré-estabelecido (MONTANARI, 2011).

Há diversos métodos para avaliação das incertezas, que podem ser categorizados em probabilísticos e não probabilísticos (MONTANARI, 2011). No primeiro grupo, são encontrados métodos embasados pela teoria da probabilidade, que fazem uso de distribuições teóricas na especificação de modelo de resíduos, enquanto, nos métodos não probabilísticos, algumas premissas teóricas podem ser relevadas. Um resumo dos principais métodos para avaliação das incertezas pode ser encontrado em Moges *et al.* (2021).

Tanto os métodos probabilísticos quanto os não probabilísticos são empregadas para fins de inferência Bayesiana. Sob esse paradigma, os parâmetros do modelo são interpretados como variáveis aleatórias e, admitidas distribuições *a priori* e uma função de verossimilhança adequada, é possível avaliar a distribuição conjunta *a posteriori* para esses parâmetros.

A avaliação das incertezas sob um paradigma Bayesiano é viabilizada, em geral, por técnicas de simulação e amostragem baseadas em métodos de Monte Carlo via cadeias de Markov (MCMC), os quais aproximam a distribuição *a posteriori* conjunta das variáveis aleatórias do problema, como os parâmetros θ , a fim de obter múltiplas realizações das saídas do modelo.

Dentre os diversos métodos de avaliação da densidade de probabilidade *a posteriori* dos parâmetros segundo simulações de Monte Carlo, o método *Differential Evolution Adaptive Metropolis* – DREAM (VRUGT *et al.*, 2008) é amplamente aplicado em estudos de modelagem hidrológica (DIAO *et al.*, 2021; HOPP *et al.*, 2020; SHENG *et al.*, 2020). O algoritmo se baseia em simulações de Monte Carlo via cadeias de Markov (MCMC), permitindo a simulação de múltiplas cadeias concomitantes e o ajuste da escala e da forma dos modelos distributivos propostos no decorrer das simulações (VRUGT *et al.*, 2009).

Apesar da larga aplicação, a inferência Bayesiana a partir de métodos MCMC preconiza a definição da função de verossimilhança, que muitas vezes é desconhecida ou intratável. Nesses casos, os resultados

da análise devem ser tomados com cautela, especialmente quando os resíduos do modelo são correlacionados, não-estacionários e não-normais (BENNETT, 2019).

Tendo em vista a dificuldade de se justificar as premissas assumidas com relação à natureza dos erros, Beven e Binley (1992) propuseram o método GLUE (*Generalized Likelihood Uncertainty Estimation*). O método rejeita a teoria da probabilidade em favor da estimativa de um conjunto de representações (entradas, estrutura, conjuntos de parâmetros e erros do modelo) cujas saídas do modelo sejam consistentes com as observações (SADEGH e VRUGT, 2013). O método evita a especificação de um modelo probabilístico formal, em busca de regiões do espaço paramétrico cujas previsões se assemelhem ao monitoramento dentro de um limite de aceitação, podendo encontrar múltiplas regiões que conduzam a séries modeladas semelhantes (NOTT *et al.*, 2012), situação relatada na literatura como “equifinalidade” (BEVEN, 1993).

Apesar da capacidade de estimar a incerteza paramétrica (RAGAB *et al.*, 2020; YAN *et al.*, 2020), o método GLUE é alvo de críticas por parte da comunidade científica, que advoga que a quantificação da verossimilhança de maneira subjetiva impede a verificação *a posteriori* das premissas admitidas *a priori*, resultando em distribuições preditivas e dos parâmetros que são estatisticamente incoerentes e/ou questionáveis (VRUGT *et al.*, 2009).

Neste contexto, os métodos de aproximação computacional Bayesiana (“ABC” – *Approximate Bayesian Computation*) surgem como alternativa para avaliação das incertezas sob um ponto de vista Bayesiano formal. De acordo com Kavetski *et al.* (2018), o uso de métodos ABC mantém a necessidade de especificação de um modelo probabilístico formal, como nos métodos MCMC, porém a faz por meio da amostragem de realizações das saídas do modelo, ao invés da avaliação da função de verossimilhança.

O conjunto de métodos ABC elimina a necessidade de definição de uma função de verossimilhança explícita, avaliando o desempenho das simulações por meio de estatísticas sumárias. Caso seja selecionado um conjunto suficiente de estatísticas sumárias, é possível estimar empiricamente a distribuição *a posteriori* das variáveis (BEAUMONT, 2019).

Por constituírem a principal base teórica e ferramental para estimação dos parâmetros e avaliação dos resultados dos modelos neste trabalho, os métodos ABC são discutidos em detalhes no item subsequente.

3.2 Approximate Bayesian Computation (ABC)

Supondo um modelo Bayesiano em que $\pi(\theta)$ denota a densidade *a priori* do conjunto de parâmetros $\theta \in \Theta$ e $f(\tilde{y}|\theta)$ denota a função de verossimilhança das observações $\tilde{y} \in \mathcal{Y}$, inferências podem ser

realizadas a partir da densidade *a posteriori* resultante $\pi(\tilde{\mathbf{y}}|\boldsymbol{\theta})$. Destaca-se que $\tilde{\mathbf{y}}$ pode ser composto por séries temporais de vazão, precipitação, assinaturas hidrológicas ou outras variáveis quantificáveis. De acordo com Sisson *et al.* (2019), a distribuição *a posteriori*, representada pelo termo a esquerda da igualdade indicada na equação (3.1), contém as informações necessárias para análise do modelo \mathcal{M} , incluindo sua verificação e validação, além de permitir predições e auxiliar a tomada de decisão.

$$\pi(\boldsymbol{\theta}|\tilde{\mathbf{y}}) = \frac{f(\tilde{\mathbf{y}}|\boldsymbol{\theta})\pi(\boldsymbol{\theta})}{\int_{\boldsymbol{\theta}} f(\tilde{\mathbf{y}}|\boldsymbol{\theta})\pi(\boldsymbol{\theta}) d\boldsymbol{\theta}} \quad (3.1)$$

Em geral, a complexidade do modelo ou da distribuição *a priori* pode acarretar a não existência de forma analítica fechada para a distribuição *a posteriori*, sendo necessário o emprego de métodos numéricos. Nesses casos, é frequente a adoção de abordagens baseadas na integração de Monte Carlo, convencional ou via cadeias de Markov, para aproximação da *posteriori*.

Todavia, em muitos casos a função de verossimilhança também é difícil de ser calculada, o que compromete a amostragem a partir da *posteriori*. Isso pode ocorrer devido a questões matemáticas, como a impossibilidade de se formular explicitamente a função, ou ferramentais, devido ao elevado custo computacional para avaliação da verossimilhança na ausência de um conjunto reduzido de estatísticas suficientes.

De acordo com Sisson *et al.* (2019), nessas circunstâncias, uma opção é a substituição do modelo \mathcal{M} adotado por um que tenha implementação mais simplificada. A principal desvantagem dessa abordagem está relacionada à adequação do modelo para representação do fenômeno de interesse, a qual pode ser prejudicada. Outra possibilidade, cuja principal vantagem é a manutenção do modelo selecionado, às custas de um certo erro de aproximação pré-definido, é a adoção de métodos baseados em aproximação computacional conhecidos como “*likelihood-free*”. Entre esses, destacam-se os métodos de aproximação computacional Bayesiana (*Approximate Bayesian computation* – ABC).

É importante ressaltar que, apesar de comumente utilizadas como sinônimos e empregadas para descrever métodos computacionais Bayesianos viáveis quando a verossimilhança é intratável ou impossível de ser calculada, as terminologias “*likelihood-free*” e ABC não são equivalentes. Sisson *et al.* (2019) apontam, inclusive, a incoerência do termo “*likelihood-free*”, uma vez que os métodos assim denominados continuam envolvendo a função de verossimilhança na análise, contudo, isso apenas ocorre de forma indireta.

Nesse contexto, tais métodos podem ser entendidos como aqueles que não promovem a avaliação direta da verossimilhança, enquanto os métodos ABC, termo cunhado por Beaumont *et al.* (2002), constituem um caso específico dos métodos “*likelihood-free*” em que ocorre uma aproximação da distribuição *a posteriori* a partir da comparação entre saídas do modelo e dados de monitoramento (ou estatísticas correspondentes).

De acordo com Del Moral *et al.* (2012), os métodos de aproximação computacional Bayesiana (ABC) requerem apenas a possibilidade de se amostrar pseudo-observações $\mathbf{y} \in \mathcal{Y}$ a partir de $f(\cdot | \boldsymbol{\theta})$. Em sua forma mais comum, as inferências, usando métodos ABC, baseiam-se na densidade *a posteriori* apresentada na equação (3.2):

$$\pi_{\varepsilon}(\boldsymbol{\theta}, \mathbf{y} | \tilde{\mathbf{y}}) = \frac{\pi(\boldsymbol{\theta})f(\mathbf{y} | \boldsymbol{\theta})I_{A_{\varepsilon, \tilde{\mathbf{y}}}}(\mathbf{y})}{\int_{A_{\varepsilon, \tilde{\mathbf{y}}} \times \boldsymbol{\theta}} \pi(\boldsymbol{\theta})f(\mathbf{y} | \boldsymbol{\theta}) d\mathbf{y} d\boldsymbol{\theta}} \quad (3.2)$$

em que: $\varepsilon > 0$ representa o nível de tolerância aceito; $I_B(\cdot)$ é a função indicadora para um dado conjunto B; \mathbf{y} corresponde a pseudo-observações; e $A_{\varepsilon, \tilde{\mathbf{y}}}$ corresponde ao conjunto de pseudo-observações próximas às observações reais, em algum sentido considerado.

O conjunto das pseudo-observações $A_{\varepsilon, \tilde{\mathbf{y}}}$ é definido como na equação (3.3):

$$A_{\varepsilon, \tilde{\mathbf{y}}} = \{\rho(\boldsymbol{\eta}(\mathbf{y}), \boldsymbol{\eta}(\tilde{\mathbf{y}})) < \varepsilon\} \quad (3.3)$$

em que: $\boldsymbol{\eta}$ representa estatísticas de resumo de interesse; e ρ representa uma função da distância.

Segundo os autores, algumas poucas variações da *posteriori* apresentada nas equações acima foram propostas, sendo essa formulação a mais recorrente. Além disso, de acordo com Marin *et al.* (2012), a definição do valor da tolerância ε é, primordialmente, um problema de custo computacional: baixos valores de tolerância, em geral, acarretam maior custo computacional, sendo prática a adoção de um pequeno percentil para esta variável.

Ainda segundo Marin *et al.* (2012), a primeira menção a métodos ABC ocorreu em Rubin (1984), em que se propõe uma solução quase automatizada para modelos intratáveis, mas a partir dos quais podem ser realizadas simulações. O primeiro algoritmo ABC, proposto conceitualmente por Rubin (1984) e aplicado por Tavaré *et al.* (1997) constitui um caso especial do método de simples rejeição, em que o conjunto de parâmetros θ é gerado a partir da distribuição *a priori* $\pi(\theta)$ e a aceitação está condicionada ao fato de a simulação correspondente ser praticamente idêntica às observações $\tilde{\mathbf{y}}$. Por isso, o algoritmo, apresentado esquematicamente abaixo, era limitado a um pequeno conjunto de aplicações.

Pseudo-Algoritmo 1

Tavaré *et al.* (1997) – Adaptado de Marin *et al.* (2012)

- 1) Gerar θ' a partir da distribuição *a priori* $\pi(\theta)$
 - 2) Gerar \mathbf{y} a partir do modelo \mathcal{M} e do conjunto de parâmetros θ'
 - 3) Repetir etapas 1 e 2 até que $\mathbf{y} = \tilde{\mathbf{y}}$
 - 4) $\theta = \theta'$
-

Em seguida, o algoritmo foi estendido por Pritchard *et al.* (1999) para o caso de espaços amostrais contínuos, gerando o primeiro algoritmo ABC genuíno (MARIN *et al.*, 2012). Essa ampliação do algoritmo, apresentado abaixo, baseia-se na ideia de que, adotados um conjunto representativo (suficiente) de estatísticas η e uma pequena tolerância ε , pode-se obter uma boa aproximação da distribuição *a posteriori*, ou seja, $\pi_\varepsilon(\theta|\tilde{\mathbf{y}}) = \int \pi_\varepsilon(\theta, \mathbf{y}|\tilde{\mathbf{y}}) d\mathbf{y} \approx \pi(\theta|\tilde{\mathbf{y}})$.

Pseudo-Algoritmo 2

Pritchard *et al.* (1999) – Adaptado de Marin *et al.* (2012)

- 1) Gerar θ' a partir da distribuição *a priori* $\pi(\theta)$
 - 2) Gerar \mathbf{y} a partir do modelo \mathcal{M} e do conjunto de parâmetros θ'
 - 3) Repetir etapas 1 e 2 até que $\rho(\eta(\mathbf{y}), \eta(\tilde{\mathbf{y}})) < \varepsilon$
 - 4) $\theta = \theta'$
-

Na prática, a adoção de uma distribuição *a priori* não-informativa $\pi(\cdot)$ pode conduzir a amostras localizadas em regiões de baixa probabilidade *a posteriori*, o que, associado aos algoritmos de simples rejeição apresentados anteriormente, pode apresentar-se ineficiente para aproximação da *posteriori*. Tendo isso em vista, Marjoram *et al.* (2003) propuseram um algoritmo MCMC-ABC, apresentado esquematicamente no pseudo-algoritmo 3.

Pseudo-Algoritmo 3Marjoram *et al.* (2003) – Adaptado de Marin *et al.* (2012)

-
- 1) Usar algoritmo 2 para gerar realização $(\boldsymbol{\theta}^{(0)}, \mathbf{y}^{(0)})$ a partir da distribuição alvo $\pi_\varepsilon(\boldsymbol{\theta}, \mathbf{y}|\tilde{\mathbf{y}})$
 - a. Gerar $\boldsymbol{\theta}'$ a partir do kernel $q(\cdot|\boldsymbol{\theta}^{(t-1)})$
 - b. Gerar \mathbf{y}' a partir do modelo \mathcal{M} e do conjunto de parâmetros $\boldsymbol{\theta}'$
 - c. Gerar u a partir de uma distribuição uniforme $\mathcal{U}_{[0,1]}$

- i. Se $u \leq \frac{\pi(\boldsymbol{\theta}')q(\boldsymbol{\theta}^{(t-1)}|\boldsymbol{\theta}')}{\pi(\boldsymbol{\theta}^{(t-1)})q(\boldsymbol{\theta}'|\boldsymbol{\theta}^{(t-1)})}$ e $\rho(\boldsymbol{\eta}(\mathbf{y}'), \boldsymbol{\eta}(\tilde{\mathbf{y}})) \leq \varepsilon$,

$$(\boldsymbol{\theta}^{(t)}, \mathbf{y}^{(t)}) = (\boldsymbol{\theta}', \mathbf{y}')$$

- ii. Se $u > \frac{\pi(\boldsymbol{\theta}')q(\boldsymbol{\theta}^{(t-1)}|\boldsymbol{\theta}')}{\pi(\boldsymbol{\theta}^{(t-1)})q(\boldsymbol{\theta}'|\boldsymbol{\theta}^{(t-1)})}$ ou $\rho(\boldsymbol{\eta}(\mathbf{y}'), \boldsymbol{\eta}(\tilde{\mathbf{y}})) > \varepsilon$,

$$(\boldsymbol{\theta}^{(t)}, \mathbf{y}^{(t)}) = (\boldsymbol{\theta}^{(t-1)}, \mathbf{y}^{(t-1)})$$

Além desses, são encontrados diversos outros algoritmos de aproximação computacional Bayesiana, muitos dos quais envolvem a adoção de estatísticas sumárias para a aproximação. Cita-se, como referência de outros algoritmos além dos aqui citados, o trabalho de Fan e Sisson (2018) e, como exemplo, os trabalhos de Albert (2015) e Stoica *et al.* (2021). Apesar dos diversos algoritmos disponíveis, Sisson *et al.* (2019) destacam que a definição das estatísticas de resumo, quando utilizadas, configuram um dos maiores desafios para a implementação de métodos ABC, uma vez que pode não ser viável a incorporação de novas estatísticas, devido ao aumento do custo computacional, por exemplo.

Neste trabalho, os métodos ABC são empregados para estimação dos parâmetros de um modelo hidrológico conceitual sob um ponto de vista formal. No item a seguir, são apresentadas as principais aplicações desses métodos, bem como uma síntese dos trabalhos que empregaram métodos ABC em estudos de modelagem hidrológica.

3.2.1 Emprego de métodos ABC em estudos de modelagem hidrológica

Originados em estudos de genética populacional, os métodos ABC têm hoje as mais diversas aplicações, sendo alguns exemplos: arqueologia (WILKINSON e TAVARÉ, 2009), história evolucionária de mosquitos (BENNETT *et al.*, 2016) e dos caracóis (CHUECA *et al.*, 2021), sistemas parasitários (BAUDET *et al.*, 2015), doenças infecciosas (AANDAHL *et al.*, 2012; DUTTA *et al.*, 2018) e cardiovasculares (DUTTA *et al.*, 2020), biomedicina (DUTTA *et al.*, 2018; WEYLAND *et al.*, 2020), ecologia (BEAUMONT, 2010), distribuição de quantis (DROVANDI e PETTITT, 2011), psicologia

(TURNER e VAN ZANDT, 2012), modelos estocásticos não lineares (ALBERT e ULZEGA, 2015), econometria (BIVAND *et al.*, 2014) e engenharia de comunicação (PETERS *et al.*, 2010).

Com relação à mecânica dos fluidos e aos recursos hídricos, são encontrados trabalhos relacionados à dinâmica dos fluidos (CHRISTOPHER *et al.*, 2021; DORONINA *et al.*, 2018), à modelagem hidrogeológica (CUI *et al.*, 2018), a dados de precipitação obtidos por meio de imagens de satélites (MA *et al.*, 2019), à estimativa de propriedades geoestatísticas, como a condutividade hidráulica (VISENTINI *et al.*, 2020), entre outras.

Segundo Sadegh e Vrugt (2014), a primeira referência ao emprego de métodos ABC em hidrologia é encontrada em Nott *et al.* (2012), que estudaram as semelhanças e diferenças entre métodos ABC e o já citado GLUE. Posteriormente, Sadegh e Vrugt (2013) utilizaram métodos ABC e assinaturas hidrológicas para diagnóstico de modelagem, a fim de evidenciar problemas no modelo. Outras aplicações são sumariamente apresentadas na Tabela 3.1.

Mais detalhadamente, os trabalhos de Kavetski *et al.* (2018) e de Fenicia *et al.* (2018) focaram na inferência Bayesiana no domínio das assinaturas, por meio do emprego de métodos ABC. O primeiro aborda aplicações prévias de métodos ABC e aspectos teóricos, como a viabilidade do emprego da metodologia proposta, a seleção da tolerância ε e a possibilidade de erros de aproximação. Adicionalmente, são discutidos os impactos da adoção de um modelo determinístico, desconsiderando as incertezas associadas à modelagem, e sugere-se a incorporação de um erro aleatório $\mathcal{E}(\theta_\varepsilon)$ à série temporal $\mathbf{h}(\theta_h, \mathbf{x})$ simulada pelo modelo determinístico. Desse modo, a estimação dos parâmetros de modelagem é feita a partir do modelo probabilístico $\mathbf{Q}(\theta, \mathbf{x})$, em que θ corresponde aos parâmetros considerados e \mathbf{x} aos dados e configurações empregados. Os autores afirmam que a consideração do modelo probabilístico representa genuinamente um modelo estocástico, tendo encontrado diferenças significativas para a incerteza estimada a partir de $\mathbf{h}(\theta_h, \mathbf{x})$ e de $\mathbf{Q}(\theta, \mathbf{x})$.

Tabela 3.1 – Exemplos de aplicações de métodos ABC em estudos de modelagem hidrológica

Referência	Principais pontos abordados e resultados encontrados
Sadegh e Vrugt (2014)	Buscando aumentar a eficiência dos algoritmos ABC implementados até então, os autores propuseram uma metodologia que combina métodos ABC e o algoritmo DREAM. Os testes conduzidos em 3 cenários distintos demonstraram uma redução do custo computacional em relação ao método de simples rejeição ABC. Além disso, o denominado DREAM _{ABC} mostrou-se mais receptivo à implementação de modelos com múltiplos processadores.
Chowdhury <i>et al.</i> (2016)	Os autores desenvolveram um <i>framework</i> em linguagem R para calibração automática de um modelo hidrológico de simulação de eventos, utilizando métodos ABC, e compararam os resultados para 3 pequenas bacias urbanas da Austrália com os obtidos utilizando o <i>software</i> comercial de modelagem MIKE URBAN.
Davison <i>et al.</i> (2017)	Tentativa de avaliação conjunta de previsão e incertezas associadas, utilizando uma abordagem denominada <i>Parameter-State Ensemble Data Assimilation (P-SEDA) Filter</i> . No entanto, devido à estrutura conceitual adotada e apesar dos bons resultados, a abordagem limita-se a simulações de curtos períodos de tempo.
Romero-Cuellar <i>et al.</i> (2019)	Os autores propuseram uma metodologia para quantificação da incerteza preditiva condicional no pós-processamento de modelos hidrológicos, utilizando algoritmo MCMC para <i>benchmark</i> . As duas abordagens (ABC e MCMC) foram testadas em dois cenários distintos, sendo o primeiro baseado em um modelo concentrado empregado em uma bacia tropical colombiana e o segundo, em um modelo distribuído utilizado em uma bacia espanhola. Como resultado, observou-se que as previsões realizadas a partir do pós-processamento ABC eram praticamente idênticas às do pós-processamento MCMC, com a vantagem de maior liberdade do modelo.
Bennett (2019)	Para tentar reduzir o custo computacional geralmente associado a métodos ABC, propôs uma combinação de <i>machine learning</i> e do método de simples rejeição de Monte Carlo para estimação de parâmetros de modelagem e incertezas associadas.

Tabela 3.1 – Exemplos de aplicações de métodos ABC em estudos de modelagem hidrológica (Continuação)

Referência	Principais pontos abordados e resultados encontrados
Nourali (2021)	Foram consideradas diferentes versões do algoritmo DREAM para determinação da incerteza paramétrica de um modelo chuva-vazão de simulação de eventos. Os algoritmos empregados foram o DREAM _{ZS} , baseado em métodos MCMC, e o DREAM _{ABC} , que constitui uma adaptação do algoritmo para incorporar métodos ABC. O autor destaca, como principal conclusão, a superioridade do DREAM _{ABC} em relação ao DREAM _{ZS} , em caso de espaços paramétricos multidimensionais.
Reichert <i>et al.</i> (2021)	Aborda a adoção de parâmetros estocásticos e dependentes no tempo para transformação de modelos determinísticos em estocásticos e, dentre as possibilidades de implementação, destaca os métodos de aproximação computacional Bayesiana. Como vantagem, os autores destacam a possibilidade de identificação e redução de deficiências no modelo.
Wu <i>et al.</i> (2021)	Apresenta o método <i>Bayesian Error Analysis with Reordering</i> (BEAR), que se baseia em uma aplicação de abordagem ABC e tem como objetivo a estimativa de erros de determinada origem (precipitação, por exemplo), ao invés da incerteza total da modelagem, decorrente da consideração de todas as fontes de incerteza. O algoritmo, avaliado em estudos de caso com dados de monitoramento e sintéticos, mostrou-se uma potencial ferramenta para avaliação de modelo ambientais cujas respostas apresentam retardo ou têm correlação entre si.

Fenicia *et al.* (2018), por sua vez, baseiam-se na discussão de Kavetski *et al.* (2018) para apresentar uma análise empírica das propriedades fundamentais relacionadas à análise desenvolvida, como extensão da série, número de assinaturas e custo computacional. Os resultados da modelagem são comparados à abordagem “clássica”, no domínio do tempo. Foram investigados aspectos como as incertezas associadas às séries temporais de vazão e à adoção de conjuntos particulares de assinaturas, por meio de experimentos que empregaram séries sintéticas e monitoradas. Como resultados principais, destacam-se a consistência entre os resultados nos domínios do tempo e das assinaturas, a possibilidade de se estimar a incerteza preditiva mesmo quando a estimação dos parâmetros é realizada por meio de assinaturas e a robustez potencial da inferência no domínio das assinaturas, quando há inconsistências no modelo adotado ou nos dados de entrada, como a ocorrência de erros nas séries temporais (como antecipação ou retardo de picos).

Em resumo, os trabalhos de Kavetski *et al.* (2018) e de Fenicia *et al.* (2018) demonstram a capacidade do ABC de promover estimativas satisfatórias para a incerteza preditiva, mesmo quando os parâmetros são estimados a partir de assinaturas hidrológicas. Dessa forma, o emprego desses métodos permitiria a

predição de vazões em bacias com pouco ou mesmo nenhum monitoramento, desde que as assinaturas possam ser calculadas por meio de regionalização.

Neste trabalho, as métricas de resumo para aproximação da distribuição *posteriori* dos parâmetros do modelo, por meio de métodos ABC, baseiam-se no cálculo de assinaturas hidrológicas. Portanto, o tópico subsequente é dedicado à sua caracterização e de suas principais aplicações.

3.3 Assinaturas hidrológicas

A abordagem mais comum para estimativa de parâmetros de modelos hidrológicos busca igualar séries temporais monitoradas e simuladas, por meio de funções-objetivo e da avaliação de métricas de desempenho. Essa é referenciada como abordagem no domínio do tempo, sendo a tradicionalmente empregada. Outra abordagem possível é por meio da comparação de assinaturas hidrológicas calculadas a partir das variáveis observadas e simuladas (FENICIA *et al.*, 2018), aqui referenciada como abordagem no domínio das assinaturas.

Apesar de amplamente adotadas em estudos sobre processos hidrológicos há décadas, a formalização do conceito de assinaturas hidrológicas foi feita por GUPTA *et al.* (2008), que as descreveram como a mínima representação relevante das informações hidrológicas contidas em um conjunto de dados. Podem ser entendidas como características derivadas de dados de monitoramento ou de séries modeladas de dados hidrológicos, como chuva, vazão ou umidade do solo. Podem englobar desde estatísticas simples, como média ou quantis de uma série temporal, até métricas mais complexas, como as que descrevem a recessão e são relacionadas ao armazenamento na bacia. Ademais, outras nomenclaturas encontradas na literatura são “métricas hidrológicas”, “índices hidrológicos” ou “assinaturas diagnósticas”.

De acordo com McMillan (2021), as assinaturas hidrológicas permitem inferências a respeito das funções hidrológicas da bacia, uma vez que podem caracterizar os fenômenos por meio de variáveis mensuráveis. Dessa forma, é possível extrair informações relevantes sobre o comportamento hidrológico de determinada região, como a identificação de processos dominantes, além de permitir a determinação da variabilidade da resposta da transformação chuva-vazão da bacia (MCMILLAN *et al.*, 2017).

Em uma primeira tentativa de categorização das assinaturas, Richter *et al.* (1996) as dividiram em 5 classes, cada uma correspondendo a uma característica do regime de escoamento: magnitude, distribuição temporal, frequência, duração e taxa de variação. Como exemplo de cada uma dessas classes, tem-se, respectivamente: vazão média mensal, período de ocorrência da máxima anual,

frequência de ocorrência de eventos de cheia, duração total do período em que a vazão observada é superior a um limite definido e taxa de variação da vazão.

Mais recentemente, McMillan (2020) propôs uma nova divisão, baseada na separação entre assinaturas baseadas em estatísticas, como a vazão média, e as baseadas nos processos que ocorrem na bacia, como o formato da curva de recessão. Para a referida autora, o primeiro grupo pode ser calculado para períodos sazonais ou de longas escalas temporais, requerem somente a distribuição das vazões e tipicamente estão relacionadas ao armazenamento na bacia, enquanto as assinaturas baseadas nos processos descrevem padrões ou formatos das séries temporais, estando normalmente associadas ao particionamento do escoamento entre diferentes caminhos (superficial e de base, por exemplo).

Dentro desses dois grandes grupos, as assinaturas podem ser divididas em 5 categorias: análise visual das séries temporais, quantificação da dinâmica de eventos, quantificação da dinâmica sazonal, estatísticas sazonais e minimodelo (quantidades derivadas de modelos simplificados). Exemplos de cada uma dessas categorias são, respectivamente: distribuição de vazões multimodal, formato da recessão, *baseflow index* (BFI), coeficiente de escoamento superficial e relações de regressão.

McMillan (2020) destaca que, inicialmente, as assinaturas eram, em geral, adotadas individualmente, tendo a combinação de diferentes assinaturas ganhado destaque em estudos recentes. A Tabela 3.2 apresenta alguns exemplos de assinaturas comumente utilizadas.

Contudo, enquanto a diversidade de assinaturas conhecidas até então permite uma caracterização de uma grande variedade de comportamentos hidrológicos, também dificulta a adoção de assinaturas adequadas e a comparação entre diferentes estudos (MCMILLAN *et al.*, 2017). Por isso, a seleção das assinaturas a serem consideradas em determinado estudo deve ser feita de forma criteriosa, tendo em vista os processos que se deseja representar.

Tabela 3.2 – Exemplos de assinaturas comumente adotadas (Fonte: adaptado de Addor *et al.*, 2018)

Assinatura hidrológica	Referências
<i>Vazão média</i>	Westerberg <i>et al.</i> (2016)
<i>Coefficiente de escoamento superficial</i>	Shafii e Tolson (2015)
<i>Elasticidade da vazão em relação a mudanças de precipitação</i>	Addor <i>et al.</i> (2018)
<i>Declividade da curva de permanência</i>	Nijzink <i>et al.</i> (2018)
<i>Baseflow index</i>	Fencia <i>et al.</i> (2018)
<i>Q₅ (Vazão com 5% de permanência)</i>	Westerberg <i>et al.</i> (2016)
<i>Q₉₅ (Vazão com 95% de permanência)</i>	Westerberg <i>et al.</i> (2016)
<i>Frequência de dias com vazões acima de um limite pré-definido</i>	Sadegh <i>et al.</i> (2015a)
<i>Duração média de eventos de cheia</i>	Addor <i>et al.</i> (2018)
<i>Frequência de dias com vazões abaixo de um limite pré-definido</i>	Sadegh <i>et al.</i> (2015a)
<i>Duração média de eventos de estiagem</i>	Addor <i>et al.</i> (2018)
<i>Frequência de dias com vazão nula (Q = 0 mm/dia)</i>	Sadegh <i>et al.</i> (2015a)

3.3.1 Seleção das assinaturas e suas aplicações

Para seleção das assinaturas a serem consideradas em um estudo, McMillan *et al.* (2017) propõem cinco diretrizes a serem satisfeitas:

- **Identificabilidade:** a incerteza com relação ao valor da assinatura deve ser pequena, quando comparada à faixa de valores possíveis para essa mesma assinatura, considerando diversas bacias. Dessa forma, diferenças relacionadas às características das bacias não serão encobertas pela incerteza associada;
- **Robustez:** o valor da assinatura deve ser independente das particularidades de medição de variáveis de campo, como as associadas a instrumentos de medição, extensão da série, períodos de registro, etc. Em outras palavras, a assinatura selecionada deve caracterizar a bacia, não somente uma medida de determinado retrato espaço-temporal da mesma;
- **Consistência:** assinaturas devem poder ser comparadas entre diferentes bacias, bem como ser insensíveis a fatores irrelevantes. Por exemplo, algumas assinaturas podem ser ponderadas em função da área de drenagem da bacia;

- Representatividade: a assinatura deve representar o comportamento médio de toda a bacia. Em se tratando de assinaturas baseadas em vazões dos rios, por exemplo, essas devem representar a dinâmica também dos tributários a montante do exutório considerado;
- Diferenciação: o conhecimento acerca dos valores das assinaturas deve contribuir para o conhecimento acerca dos processos hidrológicos. Assim, bacias com processos dominantes parecidos devem apresentar valores de assinaturas semelhantes.

Quanto às aplicações, Addor *et al.* (2018) apontam que as assinaturas hidrológicas são particularmente úteis para caracterizar e comparar a dinâmica de um extenso conjunto de bacias em que há a predominância de registros de vazão e escassez de dados como evapotranspiração e nível do lençol freático. Em outras palavras, as assinaturas podem ser uma importante fonte de informação indireta acerca dos processos hidrológicos da bacia, quando tais processos não podem ser isolados devido à ausência de dados de monitoramento.

Wlostowski *et al.* (2021), por exemplo, analisaram assinaturas para 15 bacias americanas, com o objetivo de avaliar o particionamento do escoamento. Como resultado, os autores observaram que bacias com baixos valores de *baseflow index* (BFI) e elevado escoamento superficial tendem a ter a chuva como origem do escoamento, além de solo composto em sua maior parte por argilas, enquanto elevados valores de BFI e reduzido escoamento superficial estão, em geral, associados à neve como origem do escoamento.

Outros campos de estudos ambientais fazem uso das assinaturas de forma semelhante, buscando extrair informações acerca de determinada região. Hynek *et al.* (2010), por exemplo, avaliaram assinaturas baseadas na densidade de drenagem para estudar um possível ciclo hidrológico primitivo em Marte. Trabalhos como os de Panigrahi e Das (2018) e Park e Latrubesse (2015) buscam analisar propriedades da superfície terrestre, como padrões de sedimentação de corpos hídricos e umidade do solo, por meio das chamadas assinaturas espectrais (reflectância relativa de uma superfície para diferentes comprimentos de onda). Outra aplicação possível é encontrada em Sprenger *et al.* (2019), que aborda a utilização de proporções de isótopos em uma amostra de água para identificação de fontes de água em função do tempo e do local. Essas proporções são também chamadas de “assinaturas”.

Adicionalmente às aplicações em outras áreas de estudo, destaca-se que as assinaturas hidrológicas podem ser calculadas também a partir de outras fontes de dados hidrológicos para além da vazão. Por exemplo, Horner *et al.* (2020) utilizaram uma combinação de assinaturas calculadas a partir de dados de vazão e temperatura para avaliação do regime nival em bacias dos Estados Unidos. Por sua vez, o

trabalho de Heudorfer *et al.* (2019) utilizou assinaturas para a caracterização da dinâmica hidrogeológica.

Em resumo, McMillan (2021) apresenta três áreas principais de pesquisas: ecoidrologia, processos hidrológicos na bacia e modelagem. No primeiro tipo, são encontrados trabalhos que objetivam a avaliação do regime de vazões e seu impacto no habitat ripário. Dessa forma, a fim de classificar a aptidão de determinada região, pesquisadores buscam assinaturas capazes de explicar a abundância de determinadas espécies (JOWETT e DUNCAN, 1990), por exemplo. Outra possibilidade são os estudos que pretendem detectar mudanças no regime hidrológico, o que pode levar ao aparecimento de espécies invasoras e à redução de espécies nativas (BUNN e ARTHINGTON, 2002). Adicionalmente, menciona-se o trabalho de Fan *et al.* (2017), que relacionaram a profundidade média das raízes de diversas espécies e as características de escoamento, em escala global.

Com relação à identificação e à caracterização dos processos hidrológicos, o objetivo principal é a avaliação das propriedades geomorfológicas da bacia e seu impacto no regime fluvial. A consideração dessas características permite a extrapolação para casos de bacias não monitoradas, por meio do estabelecimento de relações entre assinaturas e processos.

O grande desafio, nesse contexto, é a definição de assinaturas capazes de resumir ou representar os processos envolvidos. Eng *et al.* (2017), por exemplo, conduziram um estudo com 612 assinaturas, das quais apenas 40% podiam ser estimadas adequadamente a partir de atributos de bacias americanas. Uma possível justificativa dada por Addor *et al.* (2018) é que as propriedades climáticas, como aridez, são mais eficientes para predição do que propriedades geomorfológicas da bacia, como tipo de solo e declividade do terreno.

Além disso, é possível comparar a similaridade entre bacias distintas, como em Jehn *et al.* (2020) e em Praskievicz e Luo (2019). Dessa forma, é possível prever impactos ambientais decorrentes de mudanças climáticas ou de diferentes graus de antropização, bem como propor redes e modelos de monitoramento.

Por fim, em estudos de modelagem hidrológica, as assinaturas hidrológicas podem ser utilizadas em todas as etapas do processo, desde a seleção do modelo até a calibração e validação dos resultados. Devido à relevância deste tema para este trabalho, esta abordagem é detalhada no subitem a seguir.

3.3.2 Emprego de assinaturas hidrológicas para fins de modelagem hidrológica

Segundo Shafii e Tolson (2015), estudos de modelagem que adotam assinaturas hidrológicas podem ser categorizados em duas classes distintas: seleção ou calibração de parâmetros de modelos. No primeiro caso, as assinaturas auxiliam a seleção do modelo a ser adotado para a bacia, dentre uma série de candidatos. Nesse caso, os parâmetros do modelo são calibrados, independentemente das assinaturas

escolhidas. Por outro lado, na calibração de parâmetros de modelos baseada em assinaturas, estas são usadas para a calibração de um modelo particular, já definido, estando os parâmetros calibrados, conseqüentemente, associados às assinaturas consideradas.

Como exemplo do uso de assinaturas hidrológicas para seleção de modelos, tem-se o trabalho de David (2020), que avaliou 508 bacias brasileiras com área de contribuição inferior a 10.000 km², sob 4 modelos desenvolvidos a partir do *framework* SUPERFLEX (FENICIA *et al.*, 2011). As assinaturas foram empregadas, conjuntamente com características físicas das bacias, para verificação da existência de relação entre estrutura do modelo e características hidrológicas. Aplicação semelhante é encontrada em Martinez e Gupta (2011).

Em seu trabalho, Sahraei *et al.* (2020) propõem uma abordagem para calibração de modelos hidrológicos a partir de assinaturas hidrológicas, baseada em uma envoltória resultante da ponderação de segmentos da curva de permanência. A metodologia foi testada em 4 diferentes modelos, de diferentes graus de complexidade e estrutura, sendo alcançado, para a metodologia proposta, NSE igual a 0,76, contra 0,44 do modelo que apresentou melhor desempenho isoladamente. Outros exemplos podem ser encontrados em Bulygina *et al.* (2012), em Fenicia *et al.* (2018) e em Kavetski *et al.* (2018).

Além da seleção e da calibração de parâmetros de modelos, é possível realizar predições em bacias não monitoradas por meio da adoção de assinaturas hidrológicas regionalizadas, ao invés da regionalização dos parâmetros do modelo. Isso é possível uma vez que os atributos da bacia, em geral, estão mais relacionados às assinaturas do que aos parâmetros do modelo, e porque a regionalização das assinaturas independe da escolha do modelo de predição ou do modelo de erros (MCMILLAN, 2021).

Nesse sentido, é válido mencionar os avanços nos métodos de regionalização, por meio da incorporação de redes neurais (BECK *et al.*, 2015) e *random forests* (PRIETO *et al.*, 2019). Contudo, Addor *et al.* (2018) ressalta que a escolha das assinaturas deve ser feita criteriosamente, buscando aquelas que têm variação espacial mais suave e são mais eficientemente estimadas por meio de atributos da bacia.

É importante mencionar também trabalhos como o de Todorović *et al.* (2019), que fazem uso de assinaturas, entre outras métricas, para avaliação da eficiência de um novo modelo proposto. Vrugt e Sadegh (2013) já haviam proposto uma abordagem baseada em métodos ABC para avaliação de modelos hidrológicos por meio de assinaturas. Por fim, destaca-se o trabalho de Mendoza *et al.* (2016), que utiliza assinaturas para avaliação da resposta de 4 modelos de previsão, diante de diferentes premissas de modelagem, para um cenário climático futuro.

Especificamente com relação aos trabalhos que visam à calibração, as principais motivações para o uso de assinaturas no processo de estimativa de parâmetros de modelos hidrológicos, segundo Fenicia *et al.* (2018), são:

- A possibilidade de calibração e predição em regiões com escasso monitoramento, por meio da consideração de caráter aproximadamente constante das assinaturas ao longo do tempo. Desse modo, séries não-concomitantes de precipitação e vazão podem ser utilizadas (SADEGH *et al.*, 2015). Ainda, algumas assinaturas, como as baseadas na curva de permanência, podem ser estimadas por meio de regionalização, como em Boscarello *et al.* (2016), e posteriormente utilizadas para estimativa de parâmetros (YADAV *et al.*, 2007), como em Pinheiro e Naghettini (2010);
- Enfatizar ou desconsiderar o ajuste de componentes hidrológicas particulares, permitindo estimar parâmetros para modelos capazes de simular processos específicos com maior acurácia, como em Jepsen *et al.* (2016). Além disso, promove calibrações mais robustas contra deficiências em funções-objetivo em séries temporais de vazões, por meio do emprego de assinaturas menos sensíveis a comportamentos hidrológicos reproduzidos de forma insatisfatória pelos hidrogramas. Como exemplo, Liu *et al.* (2011) concluem que assinaturas baseadas na curva de permanência, apesar de insensíveis a erros na variação horária e nos picos dos hidrogramas, podem, em teoria, produzir estimativas mais precisas dos parâmetros, quando comparadas a modelos que empregam as séries temporais de vazão.

No entanto, a utilização de assinaturas para estimação de parâmetros de modelos hidrológicos também apresenta desvantagens. A principal delas é a possível perda de informação, que pode ter impacto negativo na estimação de parâmetros e na predição. Fenicia *et al.* (2018) destacam que isso pode ser contornado por meio da adoção de um conjunto suficiente de assinaturas. Contudo, os autores ressaltam que não há diretrizes para definição de um conjunto “estritamente suficiente”, sendo adotado, na prática, um conjunto “informativo” de assinaturas. Martinez e Gupta (2011) mencionam também que o uso de múltiplas assinaturas pode promover uma melhor representação de diversos processos físicos da bacia.

Outro ponto negativo é o fato de que, apesar da adoção das assinaturas para estimativa de parâmetros, o objetivo da modelagem, em geral, é estimar outras variáveis para além das assinaturas, como vazão, além da incerteza preditiva associada (FENICIA *et al.*, 2018). Apesar de o processo ser indireto, estudos como o de Westerberg *et al.* (2011) mostram que bons resultados podem ser encontrados para a predição de vazões em modelos cujos parâmetros foram estimados por meio de assinaturas. Em concordância, Beck *et al.* (2015) sugerem que as assinaturas podem ser regionalizadas com base em dados climáticos e fisiográficos da região de interesse.

Para a calibração de parâmetros de modelos a partir de assinaturas hidrológicas, várias são as abordagens que podem ser adotadas (SHAFII e TOLSON, 2015), sendo a maioria baseada em níveis de tolerância aceitáveis e na estimativa de incertezas na modelagem. Neste trabalho, será realizada a estimação dos parâmetros de dois modelos hidrológicos de escalas temporais distintas, a partir da consideração de assinaturas hidrológicas e um algoritmo ABC. A metodologia empregada é apresentada na sequência.

4 MATERIAL E MÉTODOS

Este trabalho está estruturado em 3 avaliações distintas, as quais buscam analisar diferentes aspectos concernentes à modelagem hidrológica e à proposta metodológica:

- Avaliação I: com o objetivo de avaliar o desempenho da modelagem e a incerteza paramétrica, desenvolveram-se simulações nos domínios do tempo, para estabelecimento de uma solução paradigma, e das assinaturas, via métodos de aproximação computacional Bayesiana. As assinaturas foram calculadas a partir dos registros horários das estações localizadas na área de contribuição da bacia do ribeirão Serra Azul. A modelagem foi desenvolvida, também em escala horária, por meio do modelo conceitual GR4H, empregando o método DREAM e a função de verossimilhança generalizada (GL), para o domínio do tempo, e o algoritmo SABC, para o domínio das assinaturas.
 - Busca-se também avaliar o impacto, nos resultados da modelagem, da consideração de combinações de assinaturas distintas;
- Avaliação II: tendo em vista a escassez de dados horários no contexto brasileiro e anteendo-se a necessidade de se empregar assinaturas hidrológicas produzidas a partir de registros hidrométricos díspares, em termos de escala temporal, daqueles adequados à representação dos processos hidrológicos em estudos de abrangência regional, buscou-se caracterizar a influência da escala temporal sobre os resultados obtidos. Simulações em escala mensal foram realizadas por meio do modelo GR2M (uma variação do modelo GR4H, como será detalhado adiante). Nesse caso, ambos os domínios foram simulados a partir de médias mensais dos registros horários na bacia em estudo. Apesar da disponibilidade de dados diários para as estações selecionadas, a consideração dos registros médios mensais reduz a influência de eventos isolados, como chuvas intensas e concentradas, os quais teriam as magnitudes e tempos associados às vazões de pico representados inadequadamente na bacia em estudo;
- Avaliação III: por fim, um ensaio preliminar de estimativa em condições de ausência de monitoramento fluviométrico local foi realizado por meio de simulações que se valeram de assinaturas estimadas a partir de registros hidrométricos diários de estações vizinhas à bacia analisada. A modelagem foi realizada em escala horária, usando o modelo GR4H e o algoritmo SABC.

A Figura 4.1 apresenta o fluxograma das etapas desenvolvidas.

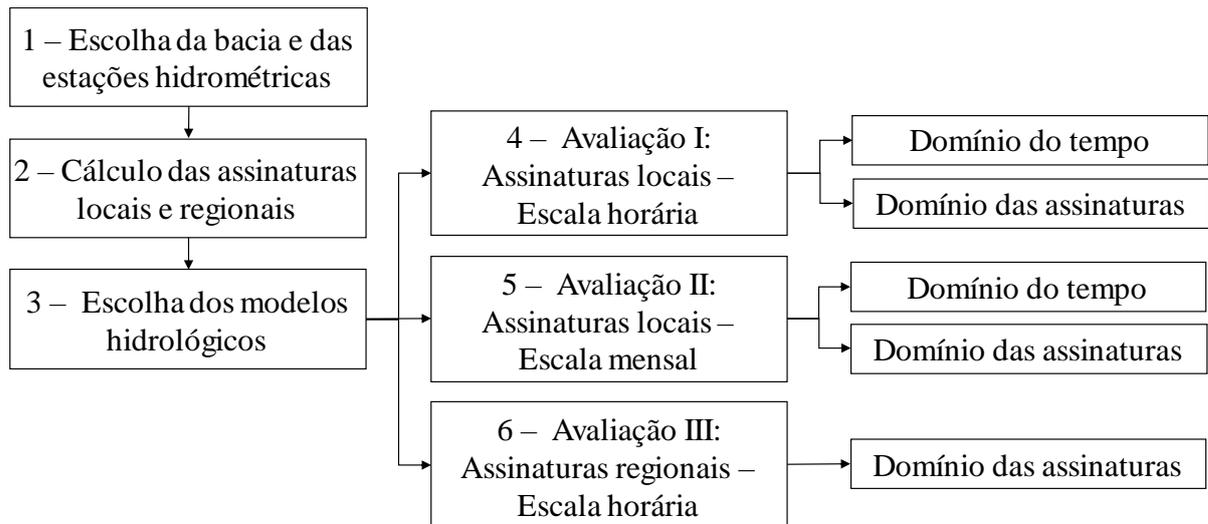


Figura 4.1 – Fluxograma geral do trabalho.

A bacia, as estações e as assinaturas selecionadas para este estudo são apresentadas na sequência, bem como as descrições dos modelos e do processo de estimação dos parâmetros nos domínios do tempo e das assinaturas.

4.1 Área de estudo e informações hidrológicas

A sub-bacia do ribeirão Serra Azul, no estado de Minas Gerais, foi escolhida para o desenvolvimento do estudo, devido à existência de séries históricas horárias e à importância socioeconômica da bacia para a Região Metropolitana de Belo Horizonte (RMBH). Sua área de contribuição está inserida na bacia representativa do ribeirão Juatuba, contando com maior número de estações para monitoramento de variáveis hidrológicas e com observações em intervalo de tempo adequado à proposta deste trabalho. Além disso, o reservatório conformado pelo barramento do ribeirão Serra Azul, em conjunto com os reservatórios Vargem das Flores e Manso, integra o Sistema Paraopeba de abastecimento. Tal sistema é responsável pelo abastecimento de cerca de 53% da população residente na RMBH (MINAS GERAIS, 2021). Isoladamente, estima-se que o reservatório seja capaz de abastecer aproximadamente 20% da população da RMBH (SANTOS e SILVA, 2015).

Embora fundamental para a manutenção das atividades socioeconômicas na RMBH, o reservatório Serra Azul foi intensamente afetado pela chamada “crise hídrica” que afetou os estados do sudeste brasileiro em meados da década passada (COELHO *et al.*, 2016; NOBRE *et al.*, 2016). Atingindo 5,2% de seu volume útil em novembro de 2014, o reservatório foi o mais impactado do sistema, tendo sido declarada situação crítica de escassez hídrica, devido ao risco iminente de desabastecimento de parte da população da RMBH (SANTOS e SILVA, 2015). Imposta pela Portaria IGAM nº 014, de 08 de abril de 2015, a

situação crítica de escassez condicionava as outorgas existentes a reduções das vazões de captação, de acordo com a finalidade de uso.

Dentre as inúmeras dificuldades enfrentadas pelo órgão gestor no enfrentamento da crise, Melo *et al.* (2020) destacam a vulnerabilidade do sistema, que não previa um plano de contingência para estabelecimento de ações emergenciais. Os autores destacam que, para além da crise, a gestão das águas em áreas urbanas é um mecanismo complexo, que deve compatibilizar quantidade e qualidade da água com seus usos, como o abastecimento. Nesse contexto, a modelagem dos processos hidrológicos da bacia constitui uma importante ferramenta, podendo auxiliar na gestão dos recursos hídricos e viabilizar previsões para elaboração dos planos de contingência para cenários futuros.

4.1.1 Caracterização da bacia do ribeirão Serra Azul

A sub-bacia do ribeirão Serra Azul integra as bacias dos rios Paraopeba/São Francisco, conforme apresentado na Figura 4.2. A bacia está compreendida entre os paralelos de $-20^{\circ}15'$ e $-20^{\circ}00'$ e os meridianos de $44^{\circ}50'W$ e $44^{\circ}35'W$ e seu reservatório tem uma altitude de cerca de 760,0 m. Distante cerca de 55 km da capital do estado, está localizada entre os municípios de Mateus Leme, Juatuba, Igarapé e Itaúna, estando inserida em uma Área de Proteção Especial.

O ribeirão Serra Azul nasce sob o nome de ribeirão dos Freitas, na Serra Azul, no município de Itaúna, a aproximadamente 1.160 m de altitude. Recebe a denominação da serra após a confluência com o córrego da Matinha ou Mato Frio, pela margem esquerda. Após a confluência com o ribeirão Mateus Leme, passa a ser denominado ribeirão Juatuba, que desagua no rio Paraopeba por volta da altitude 700 m, após mais de 44 km de talvegue (FERNANDES, 2012).

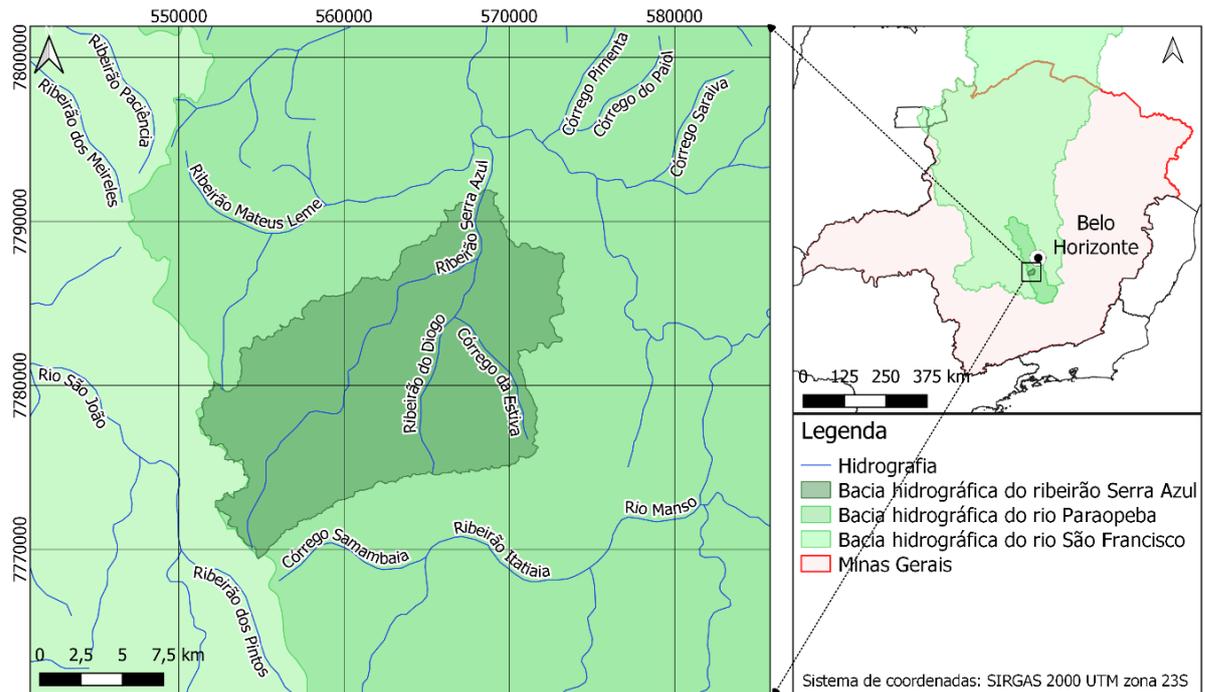


Figura 4.2 – Localização da bacia em estudo.

A bacia de contribuição do reservatório tem uma área de drenagem total de aproximadamente 260 km², tendo como vegetação característica variações típicas do cerrado, como matas de galeria, cerradão, campos sujos e limpos e mata estacional semidecidual. O uso do solo é predominantemente rural, com áreas de silvicultura, pastagens e atividades agrícolas (IBGE, 2021). Também são encontradas áreas degradadas pela atividade minerária e pelas explorações por dragagem de areia e cascalho, destinados a obras civis (FERNANDES, 2012). As Figuras 4.3 e 4.4 apresentam a caracterização das declividades e dos solos encontrados na bacia.

A região em que se localiza a bacia é caracterizada por duas estações definidas, sendo uma quente e chuvosa, entre os meses de novembro e março, e outra seca e fria, entre abril e outubro. A temperatura média do ar varia entre 22 °C e 15 °C. A umidade relativa média do ar é de 70%, a precipitação média anual é de 1476 mm e estima-se que 70% desse valor corresponda à evapotranspiração (NEVES e RODRIGUES, 2007).

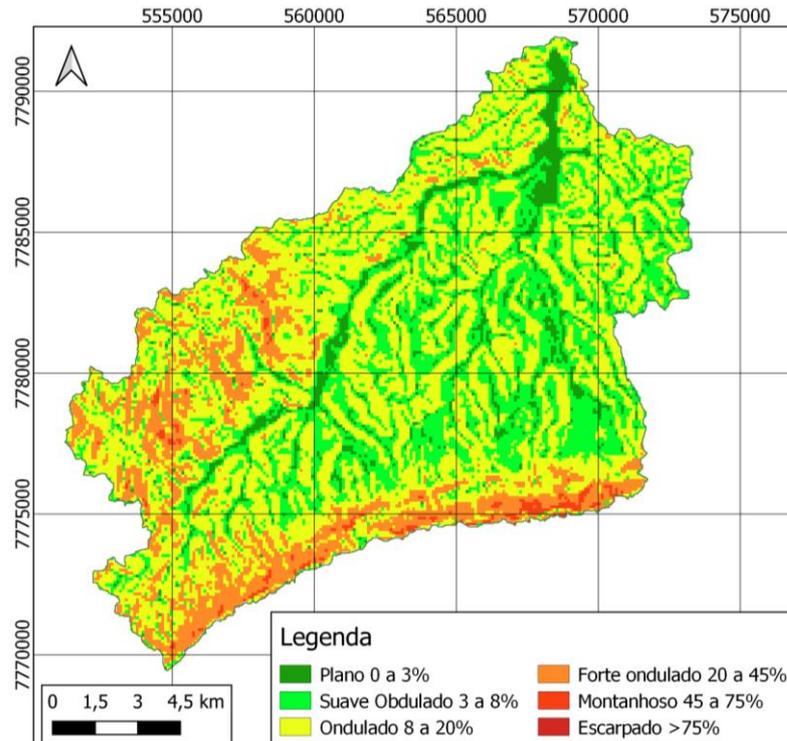


Figura 4.3 – Mapa de declividades percentuais encontradas na bacia.
 Fonte: adaptado de CPRM, 2010.

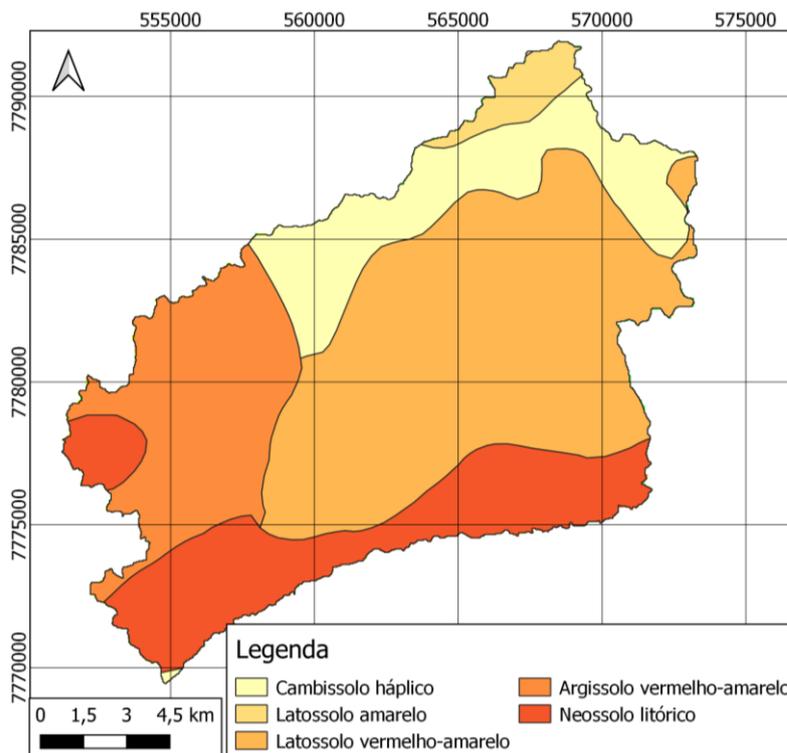


Figura 4.4 – Mapa de solos encontrados na bacia.
 Fonte: adaptado de UFV, 2010.

4.1.2 Estações pluviométricas, fluviométricas e evaporimétricas adotadas

Além da relevância socioeconômica, a bacia possui um histórico de monitoramento hidrométrico adequado à proposta, permitindo a adoção de modelos horários. Para o período compreendido entre 08/01/1997 e 28/05/2008, há registros horários de vazão para a estação Jardim (Cód. 40811100), cuja área de drenagem é de aproximadamente 113 km². Para o mesmo período, são encontradas séries de precipitação para as estações Alto da Boa Vista (Cód. 2044021), Fazenda Laranjeiras – Jusante (Cód. 2044041), Jardim (Cód. 2044052) e Serra Azul (Cód. 2044054). Portanto, essas foram as estações selecionadas para cálculo das assinaturas hidrológicas da bacia do ribeirão Serra Azul.

Tendo em vista a extensão da série de vazões horárias disponível para a estação Jardim, o período entre Jan-1997 e Nov-1997 foi utilizado para aquecimento do modelo. Os anos hidrológicos de 1997/1998 a 2007/2008 foram adotados para as simulações, sendo o período entre Dez-1997 e Fev-2003 selecionado para a calibração e o restante para a validação. Destaca-se que o último ano hidrológico, excepcionalmente, foi considerado apenas até o dia 28 (Mai-2008), devido à ausência de dados para os dias restantes do ano.

As Figuras 4.5 e 4.6 apresentam as séries históricas disponíveis para as estações localizadas na área em estudo, bem como os períodos considerados para aquecimento do modelo, calibração e validação.

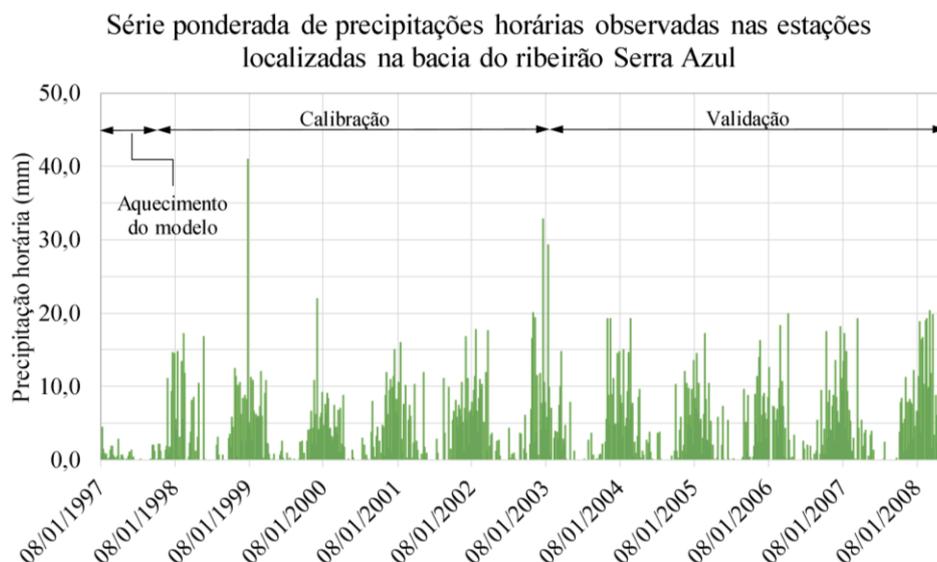


Figura 4.5 – Série ponderada de precipitações horárias registradas nas estações localizadas na bacia do ribeirão Serra Azul (Alto da Boa Vista, Fazenda Laranjeiras – Jusante, Jardim e Serra Azul).

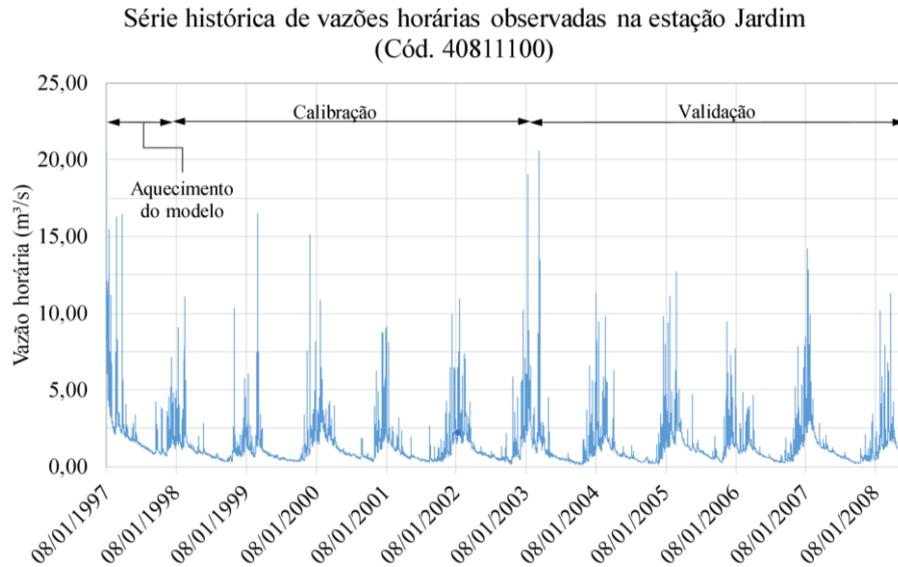


Figura 4.6 – Série histórica de vazões horárias registrada na estação Jardim.

Além dessas, foram selecionadas 5 estações fluviométricas e 5 estações pluviométricas, com monitoramento diário, no entorno da bacia, determinando uma região homogênea agrupada por conveniência geográfica. Foi considerado um raio máximo de 100 km de distância entre as estações escolhidas e a estação Jardim (Cód. 40811100). Essas estações estão apresentadas nas Tabelas 4.1 e 4.2, respectivamente. Elas foram selecionadas, também, por constituírem, de fato, uma única estação de monitoramento hidrométrico, à exceção da estação fluviométrica Alberto Flores (Cód. 40740000), cuja estação pluviométrica homônima não dispunha de dados. Foi adotada, então, a estação pluviométrica mais próxima com dados disponíveis (estação Melo Franco – Cód. 2044008).

Tabela 4.1 – Estações fluviométricas com registros diários

Código	Estação	Área de drenagem (km²)	Período de dados	Distância em relação à estação Jardim (km)
40150000	Carmo do Cajuru	2500	1936-2020	39
40330000	Velho da Taipa	7350	1938-2020	68
40740000	Alberto Flores	4120	1963-2019	31
40850000	Ponte da Taquara	8750	1967-2020	75
41250000	Vespasiano	709	1938-2020	65

Tabela 4.2 – Estações pluviométricas com registros diários

Código	Estação	Período de dados	Distância em relação à estação Jardim (km)
2044003	Carmo do Cajuru	1941-2020	39
1944021	Velho da Taipa	1959-2019	68
2044008	Melo Franco	1941-2019	34
1944031	Ponte da Taquara	1962-2020	75
1943009	Vespasiano	1941-2020	65

As Tabelas 4.3 e 4.4 apresentam, respectivamente, estatísticas amostrais para as séries das estações pluviométricas e pluviométricas consideradas neste trabalho.

Tabela 4.3 – Estatísticas amostrais para as séries de vazão observadas nas estações selecionadas

Estatística	Unidade	Jardim	Velho da Taipa	Ponte da Taquara	Carmo do Cajuru	Melo Franco	Vespasiano
Média	m ³ /s.km ²	0,012	0,013	0,013	0,014	0,017	0,014
Mediana	m ³ /s.km ²	0,009	0,009	0,009	0,010	0,012	0,010
Moda	m ³ /s.km ²	0,008	0,007	0,008	0,008	0,007	0,001
Máximo	m ³ /s.km ²	0,182	0,215	0,344	0,370	0,740	0,417
Mínimo	m ³ /s.km ²	0,001	0,000	0,001	0,001	0,001	0,000
Desvio-padrão	m ³ /s.km ²	0,011	0,013	0,014	0,013	0,018	0,015
Coeficiente de variação	-	0,958	0,983	1,044	0,927	1,097	1,042
Assimetria	-	4,182	3,819	4,962	4,725	6,543	6,367

Os dados de temperatura média do ar foram extraídos dos registros da estação climatológica INMET Florestal (Cód. 83581), de onde também foram sintetizados os dados para cálculo da evapotranspiração potencial, segundo metodologia FAO Penman-Monteith. Ressalta-se que, para o período simulado, a estação apresentava apenas 3 registros diários, que foram interpolados linearmente para estimativa das informações nos demais intervalos horários.

A Figura 4.7 apresenta a localização de todas as estações consideradas no presente trabalho.

Tabela 4.4 – Estatísticas amostrais para as séries anuais de precipitação observadas nas estações selecionadas

Estadística	Unidade	Jardim*	Velho da Taipa	Ponte da Taquara	Carmo do Cajuru	Melo Franco	Vespasiano
Média	mm	1353,3	1366,6	1290,9	1409,6	1377,5	1318,6
Mediana	mm	1307,0	1361,2	1265,5	1399,5	1360,8	1306,4
Máximo	mm	1592,9	1995,8	1863,1	2122,7	2082,9	1896,5
Mínimo	mm	1196,2	733,5	955,9	929,7	679,3	836,9
Desvio-padrão	mm	154,4	242,9	217,5	240,6	266,5	229,2
Coefficiente de variação	-	0,11	0,18	0,17	0,17	0,19	0,17

* Média ponderada por polígonos de Thiessen para as estações localizadas na bacia do ribeirão Serra Azul (Alto da Boa Vista, Fazenda Laranjeiras – Jusante, Jardim e Serra Azul)

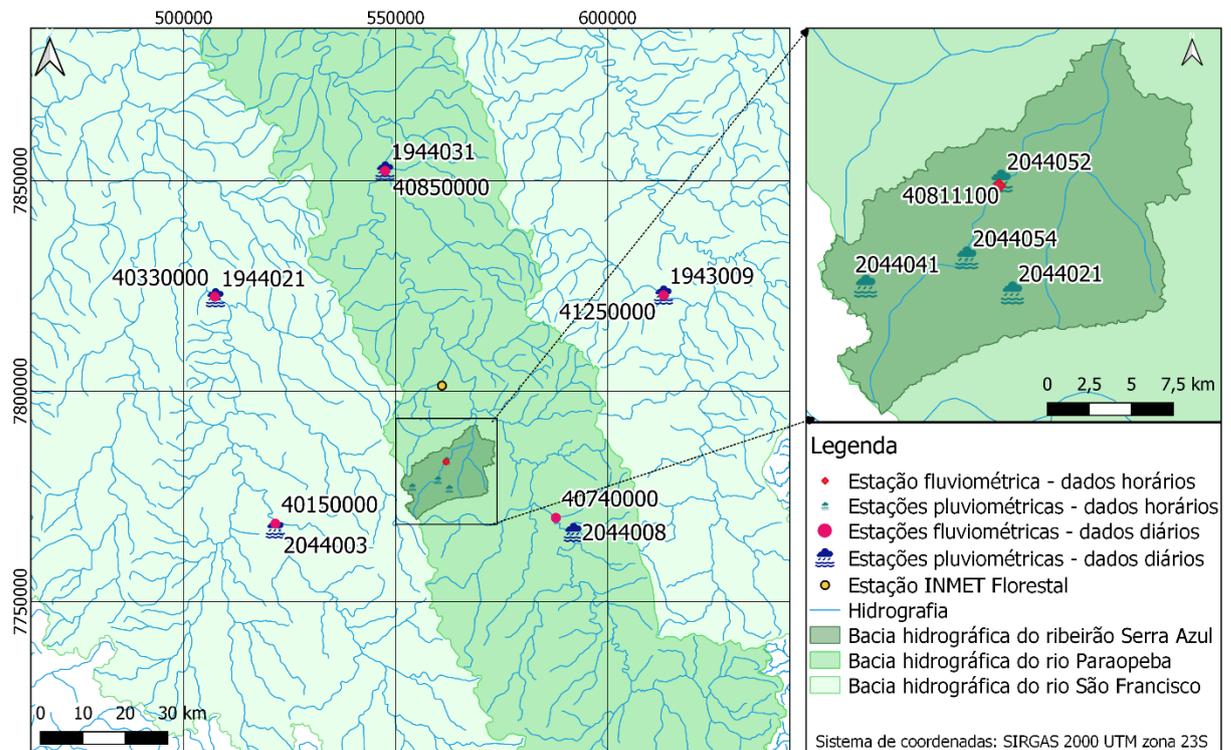


Figura 4.7 – Localização da bacia em estudo e das estações analisadas.

A seguir, são apresentadas as assinaturas calculadas a partir das séries monitoradas nessas estações.

4.2 Cálculo das assinaturas hidrológicas

Dadas as inúmeras aplicações possíveis, encontra-se na literatura uma gama de assinaturas hidrológicas empregadas com os mais diversos objetivos. Neste trabalho, foram consideradas 5 assinaturas hidrológicas para caracterização da resposta da bacia: vazão média mensal, percentis da curva de

permanência, declividade da curva de permanência, coeficiente de escoamento anual e a taxa de mudança.

Nas equações apresentadas a seguir, considera-se t o instante de tempo (hora, no caso das assinaturas locais, e dia, no caso das assinaturas regionais), $\mathbf{P} = (p_1, p_2, \dots, p_N)$ o conjunto das N pseudo-observações de precipitação e $\mathbf{Q} = (q_1, q_2, \dots, q_N)$ as de vazão. O ano hidrológico é considerado de outubro a setembro.

Buscando-se avaliar a dinâmica sazonal da bacia, foi adotada a vazão média mensal. Segundo Zhang *et al.* (2020), a consideração da vazão média mensal é comum em estudos de gestão de recursos hídricos, especialmente para operação de reservatórios. Os respectivos autores destacam também a crescente aplicação desta assinatura em estudos que buscam avaliar a vazão ambiental, relacionada à manutenção dos ecossistemas aquáticos.

A curva de permanência das vazões permite uma análise gráfica e estatística da variabilidade e distribuição empírica das vazões, e seu formato é determinado pelo padrão da precipitação, pelo uso do solo e pelas características fisiográficas da bacia (CHILES, 2019). Neste estudo, buscando avaliar o comportamento nas extremidades da curva de permanência, foram consideradas as vazões iguais ou excedidas em 5% (Q_5), 10% (Q_{10}), 90% (Q_{90}) e 95% (Q_{95}) do tempo. Foi adotada a posição de plotagem de Weibull e uma única curva de permanência para todos os registros da série de monitoramento.

Ainda com relação à curva de permanência, foi adotado também o índice relacionado a sua declividade entre os percentis 33% (Q_{33}) e 66% (Q_{66}). O objetivo do emprego dessa assinatura é a avaliação do armazenamento na bacia, bem como de sua redistribuição vertical (MCMILLAN, 2020). O valor dessa assinatura pode ser estimado a partir da equação (4.1) (SAWICZ *et al.*, 2011):

$$S_{FDC} = \frac{\ln(Q_{33}) - \ln(Q_{66})}{0,66 - 0,33} \quad (4.1)$$

A fim de avaliar a dinâmica entre superfície e subsolo, além das perdas para o lençol freático (MCMILLAN, 2020), foi adotado o coeficiente de escoamento superficial anual. A equação (4.2) apresenta o cálculo da assinatura em questão a partir dos dados do monitoramento diário (VRUGT e SADEGH, 2013). Para o cálculo das assinaturas locais, a partir da série de registros horários, o limite superior do somatório foi adequado. Destaca-se que foi considerada a média dos coeficientes calculados para cada ano da série histórica.

$$c_a = \frac{\sum_{t=1}^{365} q_t}{\sum_{t=1}^{365} p_t} \quad (4.2)$$

A taxa de mudança, tradução de *flashiness index* encontrada em Rocha (2010) e em Chiles (2019), é calculada por meio da equação (4.3) e busca caracterizar a rapidez da alteração da vazão do curso d'água. Segundo Baker *et al.* (2004), determinadas bacias têm dinâmicas pouco estáveis e resposta rápida, enquanto outras podem apresentar respostas mais lentas. O valor do índice aumenta à medida que se intensifica a instabilidade da resposta da bacia.

$$F_I = \frac{\sum_{t=2}^n |q_t - q_{t-1}|}{\sum_{t=2}^n q_t} \quad (4.3)$$

Destaca-se que os valores finais admitidos para as assinaturas regionais foram estimados a partir da mediana das assinaturas calculadas para cada estação selecionada (Tabela 4.1), a partir da série de vazões Q , dividida pela sua área de contribuição. A Tabela 4.6 apresenta a simbologia escolhida para futuras referências a cada uma das assinaturas apresentadas.

Tabela 4.5 – Assinaturas hidrológicas adotadas

Símbolo	Descrição
q_{mean}	Vazão média mensal
P_{FDC}	Percentis da curva de permanência – Q_5 , Q_{10} , Q_{90} e Q_{95}
S_{FDC}	Declividade da curva de permanência
c_a	Coefficiente de escoamento anual
F_I	Taxa de mudança

4.3 Modelos hidrológicos e dados de entrada do modelo

A seleção de um modelo hidrológico deve considerar as particularidades do problema que se deseja simular, ponderando vantagens e desvantagens dos modelos propostos. Jayawardena (2013) destaca que alguns modelos podem levar a melhores índices de desempenho, porém a um aumento significativo do custo computacional. Além disso, estudos como o de Jaiswal *et al.* (2020) demonstram a capacidade de modelos conceituais obterem boas estimativas da resposta da bacia, mesmo com as limitações decorrentes das simplificações e premissas assumidas.

Nesta pesquisa, foram adotados os modelos conceituais GR4H e GR2M, descritos nos itens subsequentes. Dado o elevado número de simulações necessárias neste trabalho, a escolha se deve ao baixo custo computacional e à parametrização parcimoniosa, permitindo simulações rápidas para um elevado número de conjuntos de parâmetros. Além disso, destaca-se que, apesar de serem encontrados registros diários para as estações escolhidas, optou-se pela adoção do modelo mensal devido às características fisiográficas da bacia em estudo. Com pequena área de drenagem e declividades por vezes acentuadas, o tempo de concentração da bacia é inferior ao intervalo diário, o que poderia acarretar uma sub-representação dos picos dos hidrogramas. Este problema é minimizado com a adoção do modelo mensal, em que foram consideradas as médias das observações de cada mês do período analisado.

4.3.1 Génie Rural à 4 paramètres Horaires – GR4H

Para as simulações com escala temporal horária, o modelo computacional adotado para estimativa das vazões é o GR4H (*Génie Rural à 4 paramètres Horaires*), variante horária do modelo chuva-vazão concentrado GR4J, desenvolvido por Perrin *et al.* (2003).

Trata-se de um modelo chuva-vazão concentrado conceitual de 4 parâmetros, os quais estão apresentados na Tabela 4.6.

Tabela 4.6 – Parâmetros do modelo GR4H (Fonte: adaptado de PERRIN *et al.*, 2003)

Parâmetro	Descrição
θ_1	Capacidade do reservatório de produção (mm)
θ_2	Coefficiente de troca de águas subterrâneas (mm/hora)
θ_3	Capacidade máxima do reservatório de distribuição (mm)
θ_4	Tempo de base do hidrograma unitário UH1 (horas)

Para o cálculo da saída em cada intervalo de tempo, o modelo GR4H utiliza como dados de entrada séries temporais de precipitação e evapotranspiração potencial, aqui referenciadas como $\mathbf{P} = (p_1, p_2, \dots, p_n)$ e $\mathbf{E} = (e_1, e_2, \dots, e_n)$, respectivamente. Todas as variáveis, sejam de entrada, saída ou mesmo as internas, são expressas em mm. Na Figura 4.8, é apresentado um diagrama esquemático do modelo para um dado passo de tempo.

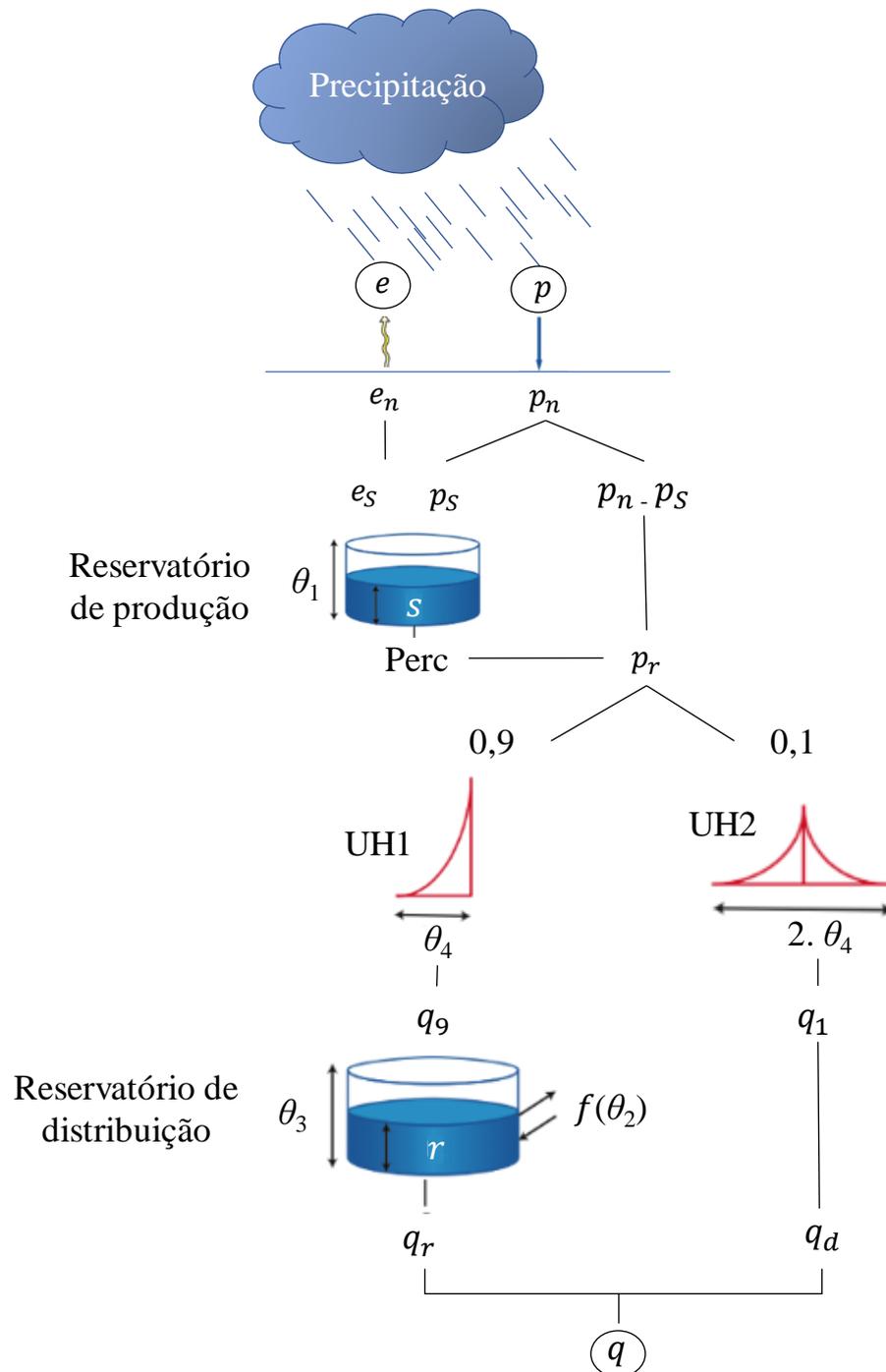


Figura 4.8 – Croqui esquemático da estrutura do modelo GR4H.

Fonte: adaptado de Perrin *et al.*, 2003.

Inicialmente, tendo-se p e e , calcula-se uma aproximação da precipitação efetiva ($net\ rainfall - p_n$) ou da evapotranspiração real ($net\ evapotranspiration - e_n$), por meio das equações (4.4) e (4.5):

$$\text{Se } p \geq e: \quad \begin{cases} p_n = p - e \\ e_n = 0 \end{cases} \quad (4.4)$$

$$\text{Se } p < e: \quad \begin{cases} p = 0 \\ e_n = e - p \end{cases} \quad (4.5)$$

Quando do primeiro caso, em que $p_n > 0$, uma parcela de p_n , simbolizada por p_s e expressa pela equação (4.6), contribui para a umidade do solo e redução do *runoff*.

$$p_s = \frac{\theta_1 \cdot \left(1 - \left(\frac{s}{\theta_1}\right)^2\right) \cdot \tanh\left(\frac{p_n}{\theta_1}\right)}{1 + \frac{s}{\theta_1} \cdot \tanh\left(\frac{p_n}{\theta_1}\right)} \quad (4.6)$$

Sendo θ_1 o primeiro parâmetro apresentado na Tabela 4.6 e s o nível do chamado reservatório de produção.

Todavia, se $e_n > 0$, considera-se uma redução do nível do reservatório decorrente da evaporação, conforme apresentado na equação (4.7):

$$e_s = \frac{s \cdot \left(2 - \frac{s}{\theta_1}\right) \cdot \tanh\left(\frac{e_n}{\theta_1}\right)}{1 + \left(1 - \frac{s}{\theta_1}\right) \cdot \tanh\left(\frac{e_n}{\theta_1}\right)} \quad (4.7)$$

O nível do reservatório de produção, portanto, é atualizado de acordo com a equação (4.8):

$$s = s - e_s + p_s \quad (4.8)$$

A percolação é calculada também com base no nível do reservatório de produção, conforme equação (4.9):

$$Perc = s \cdot \left\{1 - \left[1 + \left(\frac{4}{9} \cdot \frac{s}{\theta_1}\right)\right]\right\} \quad (4.9)$$

E, novamente, atualiza-se o nível do reservatório:

$$s = s - Perc \quad (4.10)$$

Dessa forma, a parcela que será propagada na fase de canal do modelo é representada por p_r e pode ser estimada pela equação (4.11):

$$p_r = Perc + (p_n - p_s) \quad (4.11)$$

A propagação da vazão escoada é interpretada pelo modelo como o processo resultante de duas componentes, sendo 90% do fluxo propagado com base em um hidrograma unitário UH1, com tempo de base igual a θ_4 . Na sequência, essa parcela é direcionada para um novo reservatório, cuja propagação é não-linear. Os 10% restantes da vazão total são propagados com base em outro hidrograma unitário, denominado UH2, cujo tempo de base é igual a $2 \cdot \theta_4$.

As ordenadas dos hidrogramas unitários UH1 e UH2, SH1 e SH2, respectivamente, são calculadas com base em curvas S, conforme apresentadas nas equações (4.12) e (4.13):

$$SH1(t) = \begin{cases} 0, & t \leq 0 \\ \left(\frac{t}{\theta_4}\right)^{5/2}, & 0 < t < \theta_4 \\ 1, & t \geq \theta_4 \end{cases} \quad (4.12)$$

$$SH2(t) = \begin{cases} 0, & t \leq 0 \\ \frac{1}{2} \left(\frac{t}{\theta_4}\right)^{5/2}, & 0 < t < \theta_4 \\ 1 - \frac{1}{2} \left(2 - \frac{t}{\theta_4}\right)^{5/2}, & \theta_4 < t < 2 \cdot \theta_4 \\ 1, & t \geq 2 \cdot \theta_4 \end{cases} \quad (4.13)$$

As ordenadas de UH1 e UH2 são então calculadas por meio das equações (4.14) e (4.15):

$$UH1(j) = SH1(j) - SH1(j - 1) \quad (4.14)$$

$$UH2(j) = SH2(j) - SH2(j - 1) \quad (4.15)$$

Para cada passo de tempo i , as vazões referentes aos 90% e 10% descritos anteriormente são calculadas por meio das equações (4.16) e (4.17):

$$q9(i) = 0,9 \cdot \sum_{k=1}^l UH1(k) \cdot p_r(i - k + 1) \quad (4.16)$$

$$q1(i) = 0,1 \cdot \sum_{k=1}^m UH2(k) \cdot p_r(i - k + 1) \quad (4.17)$$

Sendo $l = \text{inteiro}(\theta_4) + 1$ e $m = \text{inteiro}(2 \cdot \theta_4) + 1$

As trocas relacionadas à dinâmica hidrogeológica são modeladas a partir da consideração de um novo reservatório não-linear, cujo nível é referenciado como r e que tem capacidade de armazenamento denotada pelo parâmetro θ_3 . O parâmetro θ_2 guarda relação com a taxa de transferência entre solo e subsolo, podendo ser positivo, quando o aquífero tende a receber a contribuição das camadas superiores do solo, ou negativo, quando o aquífero tende a contribuir diretamente para o escoamento no canal. Para θ_2 igual a zero, considera-se que não há trocas significativas. Com base nessas variáveis, calcula-se o termo f :

$$f = \theta_2 \cdot \left(\frac{r}{\theta_3}\right)^{7/2} \quad (4.18)$$

O nível do reservatório r é atualizado a cada passo de tempo segundo a equação (4.19):

$$r = \max(0, r + q9 + f) \quad (4.19)$$

A saída q_r do reservatório é estimada pela equação (4.20):

$$q_r = r \cdot \left\{ 1 - \left[1 + \left(\frac{r}{\theta_3}\right)^4 \right]^{-1/4} \right\} \quad (4.20)$$

Dessa forma, o novo nível do reservatório é dado pela equação (4.21):

$$r = r - q_r \quad (4.21)$$

A parcela $q1$, propagada segundo hidrograma UH2, também sofre influência das trocas com o aquífero, sendo sua vazão no exutório calculada pela equação (4.22):

$$q_d = \max(0, q_1 + f) \quad (4.22)$$

A vazão total calculada no exutório, portanto, é dada pela equação (4.23):

$$q = q_r + q_d \quad (4.23)$$

4.3.2 *Génie Rural à 2 paramètres Mensuel – GR2M*

Foi empregado também o modelo GR2M (*Génie Rural à 2 paramètres Mensuel – MOUELHI et al., 2006*) para as simulações com dados de vazões médias mensais. Essa variante do já citado modelo GR4J, apresenta escala mensal e apenas dois parâmetros a serem estimados.

A variável mensal apresenta algumas simplificações e adaptações em relação às versões horária e diária de 4 parâmetros, sendo a principal delas a consideração de que todo o escoamento é condicionado à dinâmica do reservatório R, ou seja, $Q1 = 0$ (MOUELHI *et al.*, 2006). A Tabela 4.7 apresenta os parâmetros livres considerados no modelo GR2M, enquanto a Figura 4.9 apresenta o diagrama esquemático do modelo, bem como as equações que regem o cálculo da vazão.

Tabela 4.7 – Parâmetros do modelo GR2M (Fonte: adaptado de MOUELHI *et al.*, 2006)

Parâmetro	Descrição
θ_1	Capacidade do reservatório de produção (mm)
θ_2	Coefficiente de troca com bacias vizinhas (mm/mês)

Okkan e Fistikoglu (2014) destacam que o parâmetro θ_1 controla a resposta da bacia aos eventos chuvosos e, em certo grau, a variabilidade da vazão modelada. Valores elevados de θ_1 tendem a gerar um significativo armazenamento na bacia, tornando o *runoff* menos dependente da chuva instantânea, porém mais dependentes dos eventos antecedentes. Do contrário, para valores θ_1 de menor magnitude, o armazenamento é reduzido e o escoamento direto tem seu valor aumentado.

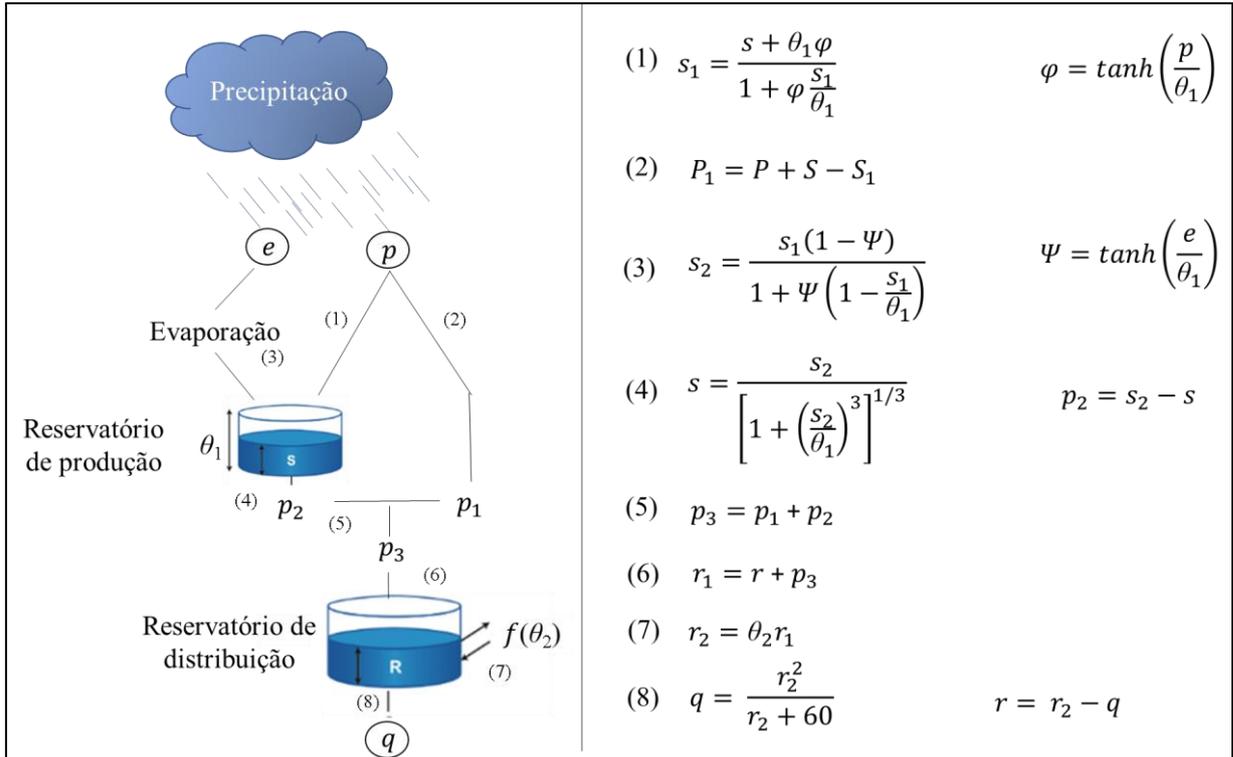


Figura 4.9 – Croqui esquemático da estrutura do modelo GR2M.

Fonte: adaptado de MOUELHI *et al.*, 2006.

4.4 Estimativa dos parâmetros do modelo para a estação Jardim no domínio do tempo

Para estabelecimento da solução paradigmática, a fim de permitir a avaliação dos resultados obtidos no domínio das assinaturas, foi empregada a abordagem comumente utilizada para estimativa dos parâmetros de um modelo, segundo uma abordagem Bayesiana (DIAO *et al.*, 2021; HOPP *et al.*, 2020; NOURALI, 2021; SHENG *et al.*, 2020). Nesse sentido, buscando permitir maior flexibilidade a representação adequada dos resíduos da modelagem, foi adotada a função de verossimilhança generalizada (SCHOUPS e VRUGT, 2010).

A função de verossimilhança generalizada permite caracterizar resíduos heterocedásticos, autocorrelacionados e modelados segundo a função densidade de probabilidade exponencial assimétrica. Seu desenvolvimento é baseado em um modelo de regressão não linear aditivo, descrito pela equação (4.24):

$$\tilde{Q} = Z + \delta \quad (4.24)$$

Em que \tilde{Q} corresponde às N observações da vazão, Z é um vetor de vazões médias e δ , um vetor de resíduos aleatórios com média zero.

As vazões médias em cada intervalo de tempo, Z_t , são calculadas a partir das vazões modeladas, q_t , por meio do modelo hidrológico \mathcal{M} , conforme apresenta a equação (4.25):

$$Z_t = q_t(\mathbf{x}|\boldsymbol{\theta}) \cdot \mu_t \quad (4.25)$$

Em que q_t é função das entradas \mathbf{x} e dos parâmetros $\boldsymbol{\theta}$ do modelo e μ_t corresponde a um fator multiplicativo que busca caracterizar o viés introduzido às saídas do modelo devido aos erros nas observações e na estrutura conceitual admitida. Sendo μ_M um parâmetro que representa o viés estimado a partir dos dados de entrada, o valor de μ_t é calculado por meio da equação (4.26):

$$\mu_t = \exp(\mu_M \cdot q_t) \quad (4.26)$$

Para consideração da autocorrelação e da dependência, os resíduos δ – equação (4.24) – são caracterizados pelo conjunto de parâmetros $\boldsymbol{\theta}_\delta$ e por uma função densidade de probabilidade, modelada segundo a equação (4.27):

$$\begin{aligned} \Phi_p(B)\delta_t &= \sigma_t \cdot a_t \\ a_t &\sim SEP(0, 1, \xi, \beta) \end{aligned} \quad (4.27)$$

Na equação (4.27), $\Phi_p(B) = 1 - \sum_{j=1}^p \phi_j B^j$ é um polinômio autorregressivo com p parâmetros ϕ_j , B é o operador de defasagem ($B^j \delta_t = \delta_{t-j}$), σ_t é o desvio padrão no tempo t , a_t expressa erros aleatórios independentes e igualmente distribuídos, com média igual a zero e desvio padrão unitário.

A equação (4.28) apresenta o modelo proposto para a função densidade de probabilidade exponencial assimétrica (SEP – *Skew Exponential Power*), com parâmetros ξ (assimetria) e β (curtose), que descreve a_t :

$$p(a_t | \xi, \beta) = \frac{2\sigma_\xi}{\xi + \xi^{-1}} \cdot \omega_\beta \cdot \exp\left\{-c_\beta |a_{\xi,t}|^{2/(1+\beta)}\right\} \quad (4.28)$$

Sendo $a_{\xi,t} = \xi^{-\text{sign}(\mu_{\xi} + \sigma_{\xi} \cdot a_t)} \cdot (\mu_{\xi} + \sigma_{\xi} \cdot a_t)$ e os valores de μ_{ξ} , σ_{ξ} , c_{β} e ω_{β} calculados a partir da assimetria ξ e da curtose β . Para $\xi = 1$, a densidade calculada é simétrica, sendo assimétrica positiva ou negativa para valores maiores ou menores do que a unidade, respectivamente. Além disso, caso seja observada simetria, o modelo probabilístico assume a forma da distribuição uniforme quando $\beta = -1$, Gaussiana quando $\beta = 0$ e de Laplace para $\beta = 1$.

A heterocedasticidade é explicitamente considerada por meio da equação (4.29), que admite variação linear para o desvio padrão, em função da vazão Z_t :

$$\sigma_t = \sigma_0 + \sigma_1 \cdot Z_t \quad (4.29)$$

Os valores dos coeficientes linear σ_0 e angular σ_1 são estimados a partir dos registros de monitoramento. Tal formulação busca representar as incertezas associadas aos ramos superiores da curva-chave (SCHOUUPS e VRUGT, 2010).

O logaritmo da função de verossimilhança derivada do modelo de resíduos descrito é apresentado na equação (4.30):

$$L(\theta|\tilde{Q}) = N \cdot \log \frac{2\sigma_{\xi}\omega_{\beta}}{\xi + \xi^{-1}} - \sum_{t=1}^n \log \sigma_1 - c_{\beta} \sum_{t=1}^n |a_{\xi,t}|^{2/(1+\beta)} \quad (4.30)$$

Neste trabalho, foi empregada a implementação da função de verossimilhança generalizada realizada por Silva (2015) em linguagem R, a partir da versão original desenvolvida em MATLAB por Schoups e Vrugt (2010). Cabe destacar que os parâmetros da função de verossimilhança, resumidos na Tabela 4.8, devem ser estimados concomitantemente aos parâmetros dos modelos, já sintetizados nas Tabelas 4.6 e 4.7.

Tabela 4.8 – Parâmetros da função de verossimilhança generalizada a serem estimados

Parâmetro	Descrição
σ_0	Heterocedasticidade: intercepto
σ_1	Heterocedasticidade: inclinação
β	Curtose
ξ	Assimetria
ϕ_j	Coefficiente de autocorreção
μ_M	Viés na vazão modelada

Para estimar a densidade de probabilidade *a posteriori* dos parâmetros do modelo e da função de verossimilhança, foi adotado o algoritmo DREAM (*Differential Evolution Adaptive Metropolis*). Esse

algoritmo de simulação de Monte Carlo via cadeias de Markov (MCMC) admite distribuições *a priori* uniformes para os parâmetros do modelo e permite a simulação de múltiplas cadeias simultaneamente. Além disso, ao longo da simulação, a escala e a forma dos modelos distributivos são constantemente atualizadas, o que resulta em maior eficiência na simulação de distribuições-alvo complexas, não lineares ou multimodais (VRUGT *et al.*, 2009).

A descrição detalhada do algoritmo DREAM foge do escopo deste trabalho, que busca avaliar de forma mais detalhada a modelagem no domínio das assinaturas por meio métodos de aproximação computacional Bayesiana e faz uso desta aplicação MCMC para definição de métricas e índices de desempenho de referência. Para mais informações acerca do DREAM, são sugeridos os trabalhos de Vrugt *et al.* (2008, 2009).

A Tabela 4.9 apresenta as principais variáveis consideradas para utilização do algoritmo DREAM.

Tabela 4.9 – Variáveis consideradas para utilização do método DREAM (Fonte: adaptado de Silva, 2015)

Variável	Descrição
<i>pars</i>	Parâmetro do modelo e suas distribuições <i>a priori</i>
<i>nseq</i>	Número de cadeias avaliadas
<i>ndraw</i>	Número máximo de iterações
<i>burn-in</i>	Número de iterações desconsideradas após o início da simulação
<i>thin.t</i>	Intervalos entre avaliações da função de verossimilhança utilizados para composição da amostra
<i>Rthres</i>	Estatística de Gelman e Rubin (R^j) para diagnóstico da convergência

Ressalta-se diversos são os exemplos de aplicações semelhantes do DREAM a estudos de modelagem hidrológica e hidrodinâmica (ESTÁCIO, 2020; HOPP *et al.*, 2020; LIU *et al.*, 2017; PINHEIRO *et al.*, 2019; RAMPINELLI, 2016).

4.5 Estimativa dos parâmetros do modelo considerando assinaturas hidrológicas

Os parâmetros dos modelos foram estimados também no domínio das assinaturas, procurando aproximar as assinaturas calculadas para as séries temporais, descritas no item 4.2, às calculadas a partir das séries simuladas.

Foi empregado o algoritmo SABC, descrito por Albert *et al.* (2014), que agrega conhecimentos dos métodos de recozimento simulado (SA – *Simulated Annealing*) e de aproximação computacional Bayesiana (ABC) para permitir a avaliação das distribuições *a posteriori* dos parâmetros. O algoritmo

permite a adaptação sucessiva do nível de tolerância ε e sua principal vantagem está relacionada ao fato de não ser baseado em esquemas de amostragem por importância, não sofrendo, portanto, de prejuízos para o tamanho da amostra efetiva (ALBERT, 2015). Para mais detalhes com relação ao SABC, recomenda-se a literatura de referência (ALBERT *et al.*, 2014; ALBERT, 2015).

Pra cada conjunto de parâmetros θ , o modelo hidrológico calcula a série de vazão simulada Q , a partir da qual são computados os valores das assinaturas g . Em seguida, os valores das assinaturas calculadas para as séries observada e simulada são comparados por meio da métrica de distância ρ , que deve ser menor ou igual a um limite aceitável e sucessivamente reduzido ε . As etapas de cálculo adotadas são apresentadas esquematicamente a seguir:

Pseudo-Algoritmo para avaliação das assinaturas (Adaptado de Kavetski *et al.*, 2018)

- 1) Gerar $\theta^{(j)}$ a partir da distribuição *a priori* $\pi(\theta)$
 - 2) Gerar a série simulada $q^{(j)} \leftarrow Q(\theta^{(j)}, x)$ a partir dos parâmetros $\theta^{(j)}$ e o do modelo cujas propriedades são denotadas por x
 - 3) Calcular $g(q^{(j)})$
 - 4) Aceitar $\theta^{(j)}$ se $\rho(\tilde{g}, g(q^{(j)})) < \varepsilon$
 - 5) Repetir os passos de 1 a 5 para $j = 1, 2, \dots, N$ conjuntos de parâmetros desejados
-

Sendo $\tilde{g} = \{\tilde{g}_1, \tilde{g}_2, \dots, \tilde{g}_N\}$ o vetor composto pela série de N assinaturas calculadas a partir dos dados de monitoramento e $g(q^{(j)}) = \{g_1, g_2, \dots, g_N\}$ o vetor estimado a partir da série simulada, a distância ρ considerada para aproximação da distribuição *posteriori* é descrita pela equação (4.31):

$$\rho(\tilde{g}, g) = \max[\Psi(\tilde{g}_1, g_1), \Psi(\tilde{g}_2, g_2), \dots, \Psi(\tilde{g}_N, g_N)] \quad (4.31)$$

em que Ψ denota uma função auxiliar para consideração de assinaturas potencialmente vetoriais, dada pela equação (4.32):

$$\Psi(\tilde{g}_k, g_k) = \frac{1}{n_k} \sum_{l=1}^{n_k} \left| \frac{\tilde{g}_{k,l} - g_{k,l}}{\tilde{g}_{k,l}} \right| \quad (4.32)$$

A adimensionalização da métrica de distância ρ tem como objetivo a ponderação de assinaturas (e seus elementos) de maior valor absoluto. Além disso, uma vez que serão consideradas, no presente trabalho,

assinaturas vetoriais e escalares, a consideração do erro médio busca permitir a comparação das diferentes saídas calculadas.

As etapas descritas anteriormente foram conduzidas para o modelo horário GR4H e, posteriormente, para o modelo mensal GR2M. Em ambos os casos, buscou-se aproximar as assinaturas calculadas para cada conjunto de parâmetro amostrado às obtidas a partir dos registros horários das estações hidrométricas localizadas na área de contribuição da bacia do ribeirão Serra Azul. Para facilitar a apresentação dos resultados, tais assinaturas serão referenciadas neste trabalho como *assinaturas locais*.

Adicionalmente, novas simulações em escala horária foram conduzidas, sendo a comparação entre saídas do modelo e variáveis monitoradas feita com base nas assinaturas calculadas a partir dos registros das estações selecionadas no entorno da bacia do ribeirão Serra Azul (Figura 4.7), doravante referenciadas como *assinaturas regionais*.

4.6 Dados de entrada e parâmetros das simulações

Para determinação das distribuições a priori a serem consideradas para os parâmetros dos modelos, foram realizadas simulações preliminares, variando-se gradativamente os intervalos admitidos e avaliando a resposta simulada da bacia. Os intervalos avaliados inicialmente foram aqueles apresentados por Perrin et al. (2003). Os conjuntos de parâmetros que continham vazões simuladas suficientemente próximas das observadas, dada uma tolerância admitida ($NSE > 0,5$, $KGE > 0,5$ e razão entre volumes simulado e observado compreendida entre 0,95 e 1,05), eram selecionados. Com base nesses resultados, os intervalos de variação dos parâmetros eram ampliados ou reduzidos e o número de simulações realizadas era aumentado. Quando foi observada relativa estabilidade entre duas rodadas de simulações, os limites da distribuição a priori de cada parâmetro foram, então, definidos. Os intervalos resultantes dessa avaliação são apresentados na Tabela 4.10, para o modelo GR4H, e na Tabela 4.11, para o GR2M. Destaca-se que foram adotadas distribuições uniformes para todos os parâmetros.

Tabela 4.10 – Intervalos de variação dos parâmetros do modelo GR4H admitidos *a priori*

Parâmetro	Limite inferior	Limite superior
θ_1	500	5000
θ_2	-10	10
θ_3	1	500
θ_4	0,5	20

Tabela 4.11 – Intervalos de variação dos parâmetros do modelo GR2M admitidos *a priori*

Parâmetro	Limite inferior	Limite superior
θ_1	500	6000
θ_2	0,01	1

Para as simulações do domínio do tempo, é necessária a definição dos parâmetros da função de verossimilhança generalizada, bem como do método DREAM. Com relação à primeira, devido ao elevado custo computacional para a convergência, os valores dos coeficientes ξ (assimetria), β (curtose) e ϕ_i (autocorrelação) foram fixados um a um, após a realização de simulações iniciais. Nessas simulações, a complexidade do modelo, representada pelo número de parâmetros considerados na análise, foi aumentada gradativamente, em busca de tendências de convergência de determinado parâmetro em torno de um pequeno intervalo de incerteza. Os intervalos de variação considerados para os demais parâmetros foram definidos a partir dos dados e dos limites numéricos de validade das formulações. As Tabelas 4.12 e 4.13 apresentam os valores e os limites da distribuição uniforme definidos para as simulações em escala horária e mensal, respectivamente.

Tabela 4.12 – Intervalos de variação admitidos *a priori* para os parâmetros da função de verossimilhança generalizada – Modelagem horária

Parâmetro	Limite inferior	Limite superior	Valor fixo
σ_0	-0,05 mm/h	0,05 mm/h	-
σ_1	0	1	-
ξ	-	-	1,2
β	-	-	0,4
ϕ_1	-	-	0,6
ϕ_2	-	-	0,1
μ_M	-	-	0 mm/h

Tabela 4.13 – Intervalos de variação admitidos *a priori* para os parâmetros da função de verossimilhança generalizada – Modelagem mensal

Parâmetro	Limite inferior	Limite superior	Valor fixo
σ_0	-3 mm/mês	3 mm/mês	-
σ_1	0	1	-
ξ	-	-	1,0
β	-	-	0,2
ϕ_1	-	-	0,55
ϕ_2	-	-	0
μ_M	-	-	0 mm/mês

Por sua vez, os parâmetros considerados na utilização do algoritmo DREAM são apresentados na Tabela 4.14. O número máximo de simulações (*ndraw*) foi definido sabendo-se que a paralisação das simulações ocorre quando há a convergência das cadeias. O *burn-in* considerado foi o sugerido pelos autores do algoritmo, enquanto o valor da estatística de Rubin e Gelman (*Rthres*), empregado para verificação da convergência, é menor do que o sugerido pelos autores (1,2), observado que nem sempre a convergência das cadeias estava caracterizada quando da adoção do valor recomendado.

Após a convergência das cadeias, foram produzidas 30.000 (20.000) novas simulações do modelo horário (mensal), para a seleção de 1.000 conjuntos de parâmetros.

Tabela 4.14 – Valores dos parâmetros empregados na utilização do método DREAM

Parâmetro	Horário	Mensal
<i>nseq</i>	6	4
<i>ndraw</i>	300.000	300.000
<i>burn-in</i>	30%	30%
<i>thin.t</i>	15	10
<i>Rthres</i>	1,01	1,01

Para as simulações no domínio das assinaturas, o número de conjuntos de parâmetros (*n.sample*) foi adotado igual ao empregado por Fenicia *et al.* (2018). O número de iterações (*iter.max*) foi definido após testes iniciais de convergência, a partir da avaliação da variabilidade dos conjuntos ao longo das simulações. Foram traçados os diagramas do tipo *boxplot* para os conjuntos de parâmetros amostrados para determinadas iterações ao longo de toda a simulação e, para todas as assinaturas consideradas, foi observada uma estabilização da variabilidade para um número de realizações do modelo inferior a 1.000.000.

Por fim, o valor inicial da tolerância ε foi selecionado também com base em simulações-teste, avaliando a convergência e as distribuições *a posteriori* dos parâmetros. Foram adotados os menores valores de ε que permitiram a amostragem dos 5.000 conjuntos de parâmetros desejados, uma vez que valores inferiores aos adotados não conduziram a nenhuma solução. É importante destacar que o valor da tolerância é reduzido ao longo das simulações e a adoção de valores iniciais muito pequenos pode ocasionar lenta convergência, enquanto valores iniciais elevados podem produzir distribuições *a posteriori* muito semelhantes à distribuição admitida *a priori* (KAVETSKI *et al.*, 2018). A Tabela 4.15 apresenta os valores adotados.

Todas as etapas deste trabalho foram implementadas em linguagem R, com o auxílio das bibliotecas listadas na Tabela 4.16. Pela disponibilização destas ferramentas, agradecemos à comunidade científica.

Tabela 4.15 – Valores dos parâmetros empregados na utilização do algoritmo SABC

Parâmetro	GR4H	GR2M
<i>n.sample</i>	5000	5000
<i>iter.max</i>	1.000.000	1.000.000
ε	0,10	0,15

Tabela 4.16 – Bibliotecas R utilizadas neste trabalho

Biblioteca	Biblioteca
<i>dplyr</i>	<i>tictoc</i>
<i>tidyr</i>	<i>airGR</i>
<i>tibble</i>	<i>EasyABC</i>
<i>readxl</i>	<i>magrittr</i>
<i>ggplot2</i>	<i>dream</i>
<i>hydroGOF</i>	<i>lattice</i>
<i>lubridate</i>	<i>signal</i>
<i>matrixStats</i>	<i>purrr</i>

Para a avaliação do desempenho da modelagem, foi considerada a série composta pela mediana, para cada intervalo de tempo, das vazões simuladas por meio dos conjuntos de parâmetros selecionados e do modelo determinístico. O desempenho da modelagem foi avaliado com base na análise do índice KGE (GUPTA *et al.*, 2009) e de suas componentes, bem como da análise gráfica das séries simuladas. O KGE é calculado a partir de 3 componentes: o coeficiente de correlação de Pearson (r), a razão entre a média dos valores simulados (no caso, as vazões) e a média dos valores observados (γ) e a razão entre desvios padrões dos valores simulados e observados (α). Idealmente, as três componentes, bem como o próprio KGE, são iguais à unidade.

Adicionalmente, foi avaliada também a raiz do erro médio quadrático (RMSE – *Root Mean Square Error*), uma vez que essa métrica se mostra sensível a *outliers*. Possui a mesma unidade da variável simulada e pode ser interpretada como uma medida do desvio médio entre variáveis observadas e simuladas. Idealmente, seu valor é igual a zero.

Por fim, ressalta-se que, como mencionado no item 3.2.1, Kavetski *et al.* (2018) propõem a consideração de um erro aleatório à série simulada pelo modelo determinístico. No presente trabalho, foram realizadas simulações a fim de avaliar a incerteza quando da consideração ou não de um modelo de resíduos. Foram avaliados os modelos de erros propostos por Kavetski *et al.* (2018), baseado em um processo AR(1), $\varepsilon_t = \phi\varepsilon_{t-1} + W_t$, cujas inovações W_t são realizações de uma distribuição Gaussiana truncada de média zero e desvio-padrão σ_W , e por Schoups e Vrugt (2010), a partir de amostras independentes geradas a partir da já abordada distribuição SEP.

5 RESULTADOS E DISCUSSÃO

5.1 Cálculo das assinaturas hidrológicas

Inicialmente, a fim de possibilitar as avaliações propostas, foram calculadas as assinaturas selecionadas, em escalas local e regional. Para as assinaturas locais, a série de precipitação foi calculada a partir da média ponderada dos registros horários das estações localizadas na bacia, segundo polígonos de Thiessen. As vazões da estação Jardim foram divididas pela sua área de drenagem (113 km²), sendo as assinaturas calculadas a partir da série em mm/h.

Já para o cálculo das assinaturas regionais, a primeira etapa foi avaliar as distribuições da chuva e das vazões para as estações selecionadas, comparando-as com os registros das estações localizadas na área de contribuição direta da bacia. As Figuras 5.1 a 5.3 apresentam os gráficos de precipitação e vazão médias mensais para as estações.

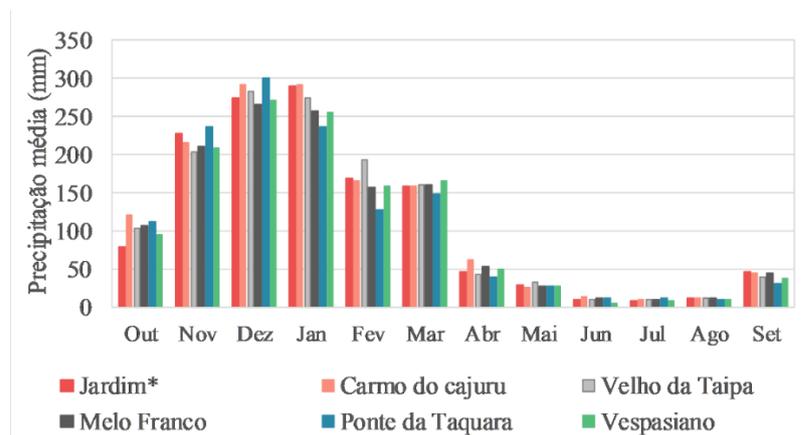


Figura 5.1 – Precipitações médias mensais para as estações hidrométricas adotadas no estudo.
* Média ponderada por polígonos de Thiessen para as estações localizadas na bacia do ribeirão Serra Azul (Alto da Boa Vista, Fazenda Laranjeiras – Jusante, Jardim e Serra Azul)

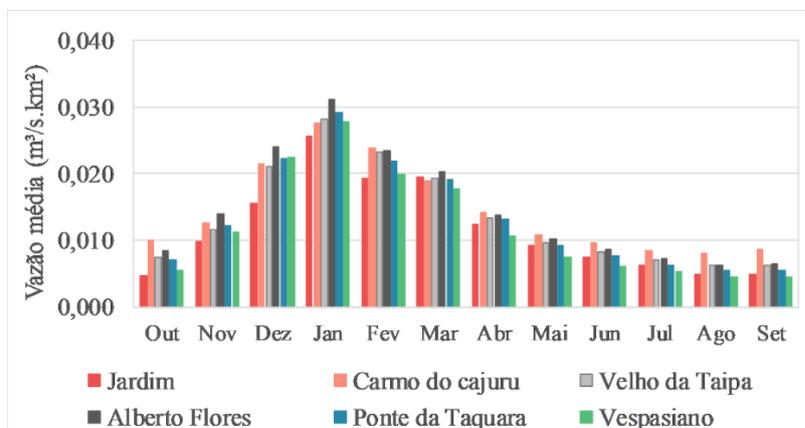


Figura 5.2 – Vazões médias mensais para as estações hidrométricas adotadas no estudo.

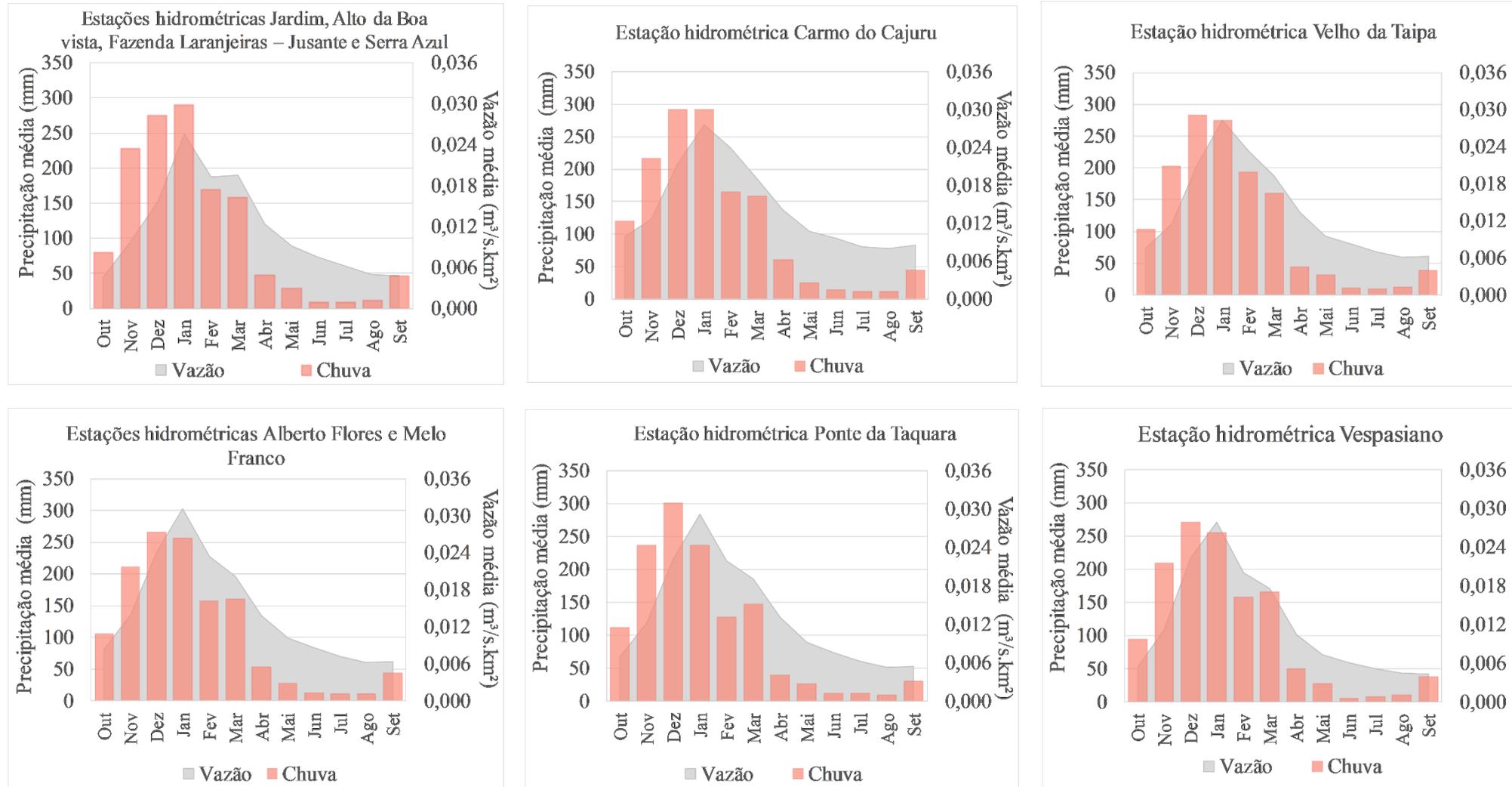


Figura 5.3 – Precipitações e vazões médias mensais para as estações hidrométricas adotadas no estudo.

Observa-se, em geral, consistência entre os dados de precipitação de todas as estações. Em relação ao monitoramento fluviométrico, no entanto, a estação Jardim apresenta, em geral, vazões menores do que as registradas nas demais estações analisadas. Essas diferenças podem estar relacionadas a características físicas da bacia, como sua cobertura vegetal e sua dinâmica hidrogeológica, mas é importante destacar também as disparidades com relação às extensões das séries: para a estação Jardim, há aproximadamente 10 anos de dados horários, enquanto as demais estações contam com séries diárias que variam entre cerca de 50 a 80 anos.

A própria escala do monitoramento pode contribuir para essas diferenças, uma vez que esse padrão não foi observado em relação à chuva, cujo monitoramento, via de regra, é feito com base no volume acumulado em um determinado período de tempo, independentemente da sua distribuição ao longo desse intervalo. Ponderadas as limitações da consideração de um caráter regional para as estações selecionadas, as assinaturas foram calculadas considerando a mediana das séries de vazões específicas de cada estação selecionada no entorno da bacia.

Por fim, as assinaturas para as simulações mensais foram calculadas com base nas médias mensais dos registros pluviométrico e fluviométrico das estações selecionadas.

Os valores calculados para cada uma das assinaturas, em ambas as escalas espaciais de monitoramento, são apresentados nas Tabelas 5.1 e 5.2. Também são apresentadas as diferenças percentuais das assinaturas regionais, em relação às locais.

Como esperado, devido às menores magnitudes das vazões registradas na estação Jardim, considerando o período de dados horários disponíveis, as estações no entorno da bacia conduziram, em geral, a valores maiores para as assinaturas do que aqueles estimados a partir das estações localizadas na bacia.

Ressalva se faz em relação ao valor da taxa de mudança F_t , que apresentou variações expressivas de acordo com a escala considerada. Tais diferenças estão, provavelmente, associadas à escala temporal de monitoramento das variáveis: dado que as vazões da estação Jardim foram observadas a cada intervalo horário, as variações do hidrograma entre dois intervalos de tempo tendem a ser mais suaves.

Por outro lado, as estações localizadas no entorno da bacia contam com apenas duas medições diárias, sintetizadas em um único valor médio de vazão por dia. Dessa forma, mesmo com áreas de drenagem significativamente maiores do que as da estação Jardim, e, conseqüentemente, tempos de concentração mais elevados, a representação da variabilidade das vazões é prejudicada. Assim, as diferenças entre as vazões q_t e q_{t-1} se tornam relevantes, acarretando maiores valores de F_t .

Tendo isso em vista, a taxa de mudança F_I não se enquadra no critério de robustez proposto por McMillan *et al.* (2017) para seleção de assinaturas. Contudo, devido aos bons resultados encontrados por Fenicia *et al.* (2018) para a combinação dessa assinatura com a curva de permanência em escala diária a partir de dados locais, optou-se por manter F_I nas avaliações a partir das assinaturas locais nas simulações horárias. Assim, nas simulações horárias a partir de assinaturas regionais, bem como nas mensais, essa assinatura não foi considerada.

Tabela 5.1 – Assinaturas hidrológicas calculadas para as estações selecionadas – Séries horárias e diárias

Escala	q_{mean} (mm/h)												P_{FDC} (mm/h)				S_{FDC}	c_a	F_I
	Jan	Fev	Mar	Abr	Mai	Jun	Jul	Ago	Set	Out	Nov	Dez	Q ₅	Q ₁₀	Q ₉₀	Q ₉₅			
<i>Local</i>	0,090	0,070	0,070	0,044	0,033	0,027	0,022	0,018	0,017	0,017	0,035	0,056	0,108	0,081	0,013	0,011	1,970	0,251	0,021
<i>Regional</i>	0,102	0,084	0,069	0,048	0,034	0,030	0,025	0,022	0,023	0,027	0,044	0,080	0,139	0,097	0,017	0,014	1,815	0,320	0,130
<i>Diferença</i>	+13%	+21%	-3%	+8%	+5%	+10%	+13%	+25%	+31%	+63%	+25%	+43%	+29%	+20%	+37%	+29%	-8%	+28%	528%

Tabela 5.2 – Assinaturas hidrológicas calculadas para as estações selecionadas – Séries mensais

Escala	q_{mean} (mm/mês)												P_{FDC} (mm/mês)				S_{FDC}	c_a	F_I
	Jan	Fev	Mar	Abr	Mai	Jun	Jul	Ago	Set	Out	Nov	Dez	Q ₅	Q ₁₀	Q ₉₀	Q ₉₅			
<i>Local</i>	68,54	47,16	52,40	31,96	0,40	19,48	16,58	13,23	12,41	12,29	25,19	41,63	66,79	60,00	10,36	9,00	1,95	0,25	0,40

5.2 Domínio tempo versus domínio das assinaturas: escala horária

Neste item, são apresentados os resultados das simulações horárias, considerando a série modelada a partir do algoritmo DREAM e da função de verossimilhança generalizada como *benchmark* para comparação dos resultados no domínio das assinaturas.

5.2.1 Avaliação do desempenho da modelagem

As Tabelas 5.3 e 5.4 apresentam as métricas de desempenho para os períodos de calibração e validação, respectivamente. A Tabela 5.5 apresenta os percentuais das observações que estão contidas no intervalo de credibilidade de 95%.

Tabela 5.3 – Índices de desempenho da modelagem em escala horária - Período de calibração

Estatística	Domínio do tempo	Domínio das assinaturas				
		q_{mean}	P_{FDC}	S_{FDC}	c_a	F_I
r	0,89	0,65	0,80	0,81	0,78	0,89
γ	1,10	1,02	0,97	0,68	0,99	0,98
α	1,19	0,77	0,89	0,60	0,76	1,48
KGE	0,76	0,58	0,77	0,46	0,67	0,51
$RMSE (mm/h)$	0,02	0,03	0,02	0,03	0,02	0,03

Tabela 5.4 – Índices de desempenho da modelagem em escala horária - Período de validação

Estatística	Domínio do tempo	Domínio das assinaturas				
		q_{mean}	P_{FDC}	S_{FDC}	c_a	F_I
r	0,88	0,68	0,79	0,81	0,78	0,87
γ	1,22	1,20	1,09	0,82	1,20	1,00
α	1,06	0,74	0,81	0,56	0,72	1,26
KGE	0,74	0,54	0,70	0,49	0,60	0,71
$RMSE (mm/h)$	0,02	0,03	0,02	0,02	0,02	0,02

Tabela 5.5 – Percentual das observações contidas no intervalo de credibilidade de 95%

Período	Domínio do tempo	Domínio das assinaturas				
		q_{mean}	P_{FDC}	S_{FDC}	c_a	F_I
Calibração	2%	32%	23%	100%	87%	100%
Validação	1%	20%	19%	100%	89%	100%

A partir das tabelas apresentadas, é possível verificar uma tendência, no domínio do tempo, de vazões simuladas mais elevadas dos que as observadas ($\gamma > 1$), além da maior variabilidade

($\alpha > 1$). Em geral, a situação contrária ($\gamma < 1$ e $\alpha < 1$) é encontrada para as simulações no domínio das assinaturas, com exceção da taxa de mudança F_I .

Essas variações ficam evidentes quando são analisados os hidrogramas simulados: tanto no domínio do tempo quanto para F_I . Apesar de haver superestimativas pontuais das vazões simuladas, houve uma melhor representação dos picos dos hidrogramas do que para as demais assinaturas. A recessão, em geral, foi bem caracterizada no domínio do tempo e para a maioria das assinaturas, excluindo-se F_I , cujas vazões de recessão simuladas foram inferiores às observadas. Para F_I , portanto, é possível inferir, no cálculo do coeficiente γ , um efeito compensatório com relação ao viés nos diferentes trechos do hidrograma: apesar de superestimar os picos e subestimar a recessão, a vazão média simulada é bastante semelhante àquela observada, o que é traduzido por valores de γ muito próximos do ideal.

Quanto aos coeficientes de correlação, os resultados no domínio do tempo mostraram-se levemente superiores aos do domínio das assinaturas. Com relação aos valores de KGE, é possível observar uma redução expressiva dessa estatística para a maioria das assinaturas, quando comparadas ao domínio do tempo. Esse resultado evidencia a insuficiência das assinaturas selecionadas, que não se mostraram capazes de sintetizar toda a informação contida na série temporal. Exceção é feita para os percentis da curva de permanência P_{FDC} , que apresentaram valores de KGE consideravelmente maiores do que os das demais simulações do domínio das assinaturas e muito próximos dos do domínio do tempo, em decorrência dos valores de r , γ e α mais próximos da unidade.

Ressalta-se que a curva de permanência é reportada pela literatura como uma alternativa à estimativa de parâmetros no domínio do tempo (WESTERBERG *et al.*, 2011). Em contrapartida, Fenicia *et al.* (2018) afirmam que modelos cujos parâmetros foram estimados a partir da curva de permanência apresentam intervalos mais largos para a incerteza preditiva e podem conduzir a previsões menos confiáveis do que aqueles que empregam diretamente a série temporal de vazões. Esses resultados foram também observados neste trabalho, como apresentam as Figuras 5.4, 5.5 e 5.8. Apesar disso, é interessante observar que poucos percentis da curva de permanência são capazes de alcançar bons índices de desempenho.

Com relação ao RMSE, todas as assinaturas apresentaram resultados semelhantes aos do domínio do tempo. Para efeito de comparação, o desvio-padrão da série observada na estação Jardim é de 0,04 mm/h, valor superior aos encontrados para o RMSE para todas as assinaturas.

Apesar da superioridade da predição no domínio do tempo, destaca-se que as simulações nesse domínio, bem como para q_{mean} e para P_{FDC} , mostraram-se “excessivamente confiantes” quanto à incerteza

paramétrica, como mostra a Tabela 5.5. Situação contrária foi encontrada para S_{FDC} e F_I , que apresentaram menores valores de KGE e ampla faixa de incerteza com relação à estimação dos parâmetros, o que demonstra que, isoladamente, essas assinaturas não são capazes de sintetizar as informações da série monitorada e conduzem a resultados que devem ser tomados com cautela.

Especificamente para o domínio do tempo, o pequeno intervalo de variação dos parâmetros pode estar relacionado à fixação dos parâmetros da SEP, especialmente à assimetria ξ . Idealmente, todos os parâmetros deveriam ser amostrados, porém o número de iterações necessárias para a convergência, nesse caso, se mostrou computacionalmente proibitivo. Apesar de não ser a condição ideal para simulação, a fixação do valor de determinados parâmetros é comumente encontrada na literatura. É válido mencionar ainda que intervalos de credibilidade notadamente menores para o método DREAM, em comparação ao ABC, já haviam sido reportados por Sadegh e Vrugt (2013).

A seguir, são apresentados os hidrogramas das séries modeladas *versus* a série observada, para as simulações que melhor exemplificam os aspectos descritos anteriormente. As hachuras representam os intervalos com 95% de incerteza relacionados às estimativas das vazões. Apresenta-se, aqui, apenas o período de validação, a fim de facilitar a visualização da variabilidade dos hidrogramas. Os gráficos para as demais assinaturas, bem como para o período de calibração, são apresentados no Apêndice I. Na Figura 5.8, são encontradas as relações entre as vazões simuladas e aquelas observadas, para todas as simulações realizadas.

Destaca-se que, apesar de não serem apresentados neste item para permitir uma melhor fluidez do texto, a análise visual dos hidrogramas modelados considerando a vazão média mensal q_{mean} e o coeficiente de escoamento superficial anual c_a também foi realizada (ver Apêndice I). Para a primeira assinatura, observou-se uma boa caracterização dos períodos de recessão, porém uma superestimativa das vazões para os períodos chuvosos. Apesar dos valores elevados para as vazões nos meses chuvosos, os picos dos hidrogramas também não foram satisfatoriamente reproduzidos. Para c_a , o comportamento é semelhante ao encontrado para P_{FDC} , apresentando maior incerteza quanto às predições do modelo.

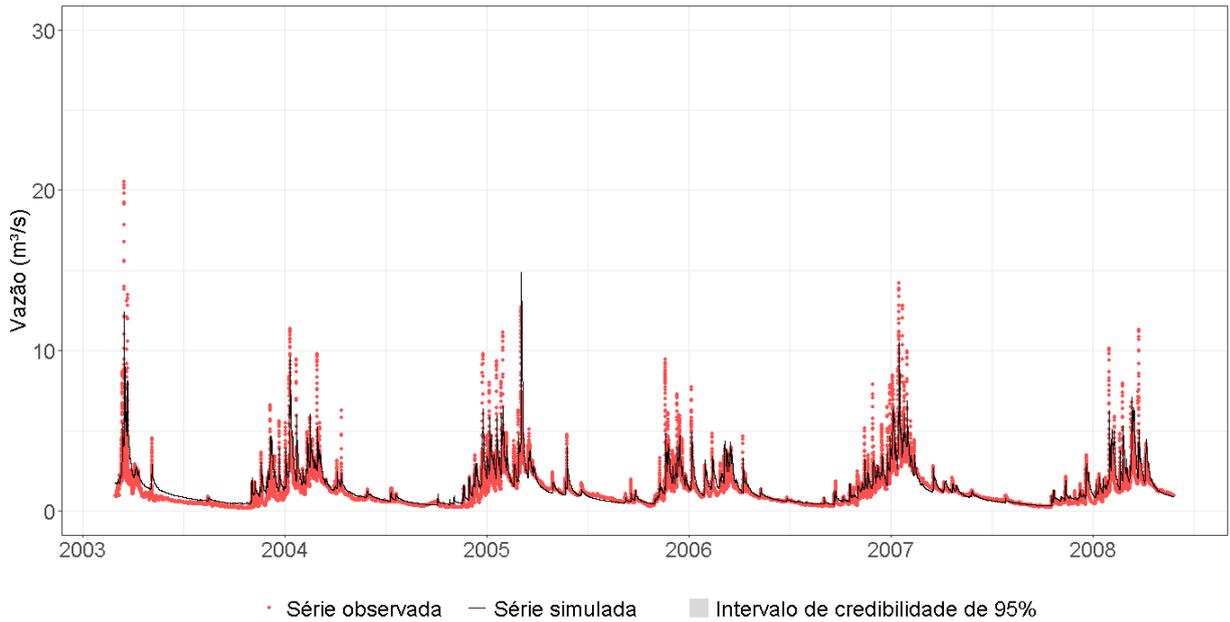


Figura 5.4 – Série temporal de vazões observadas e simuladas no domínio do tempo, para o período de validação.

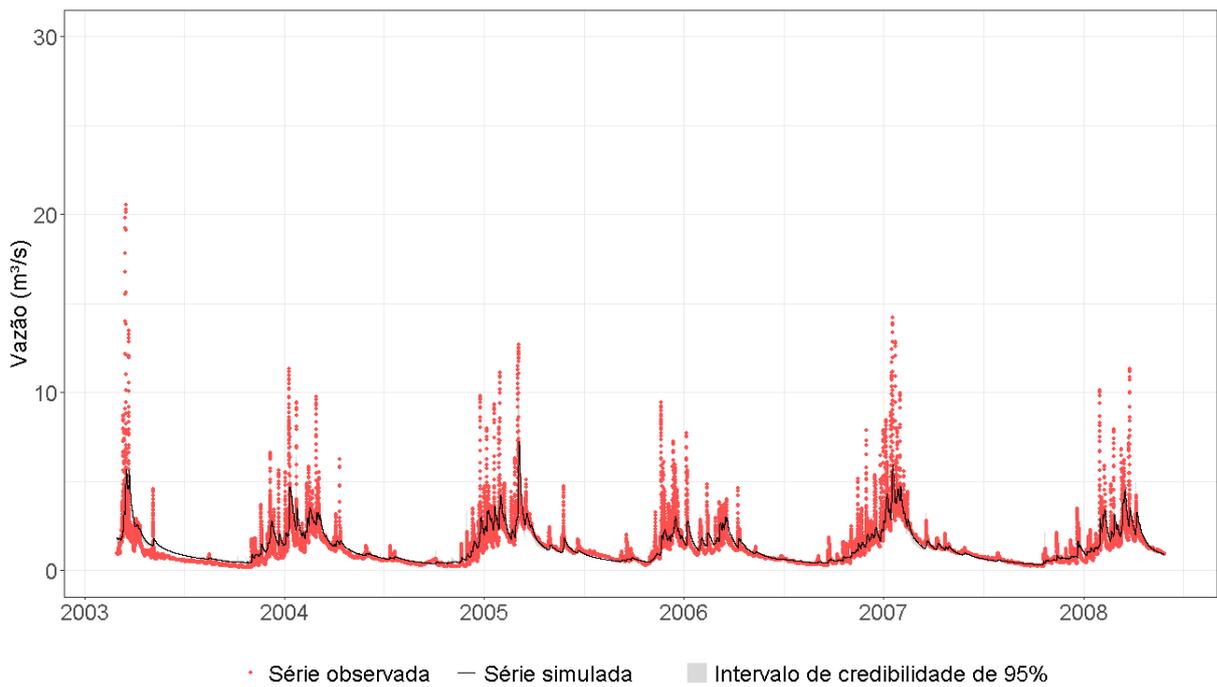


Figura 5.5 – Série temporal de vazões observadas e simuladas no domínio das assinaturas, para o período de validação - P_{FDC}

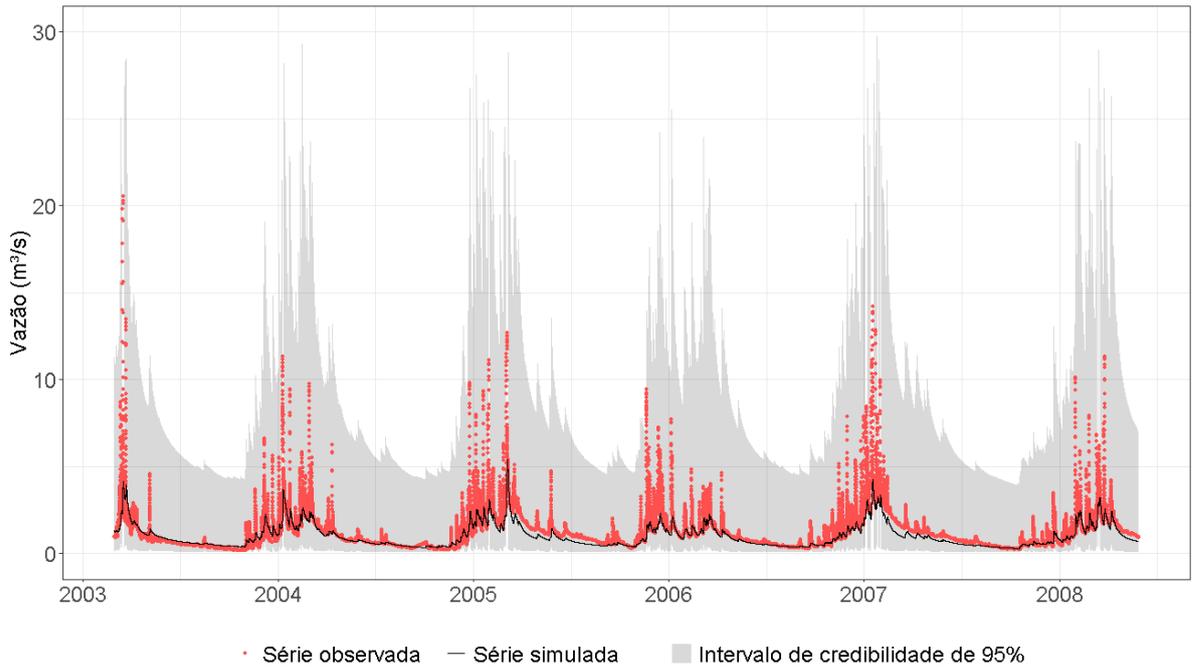


Figura 5.6 – Série temporal de vazões observadas e simuladas no domínio das assinaturas, para o período de validação - S_{FDC}

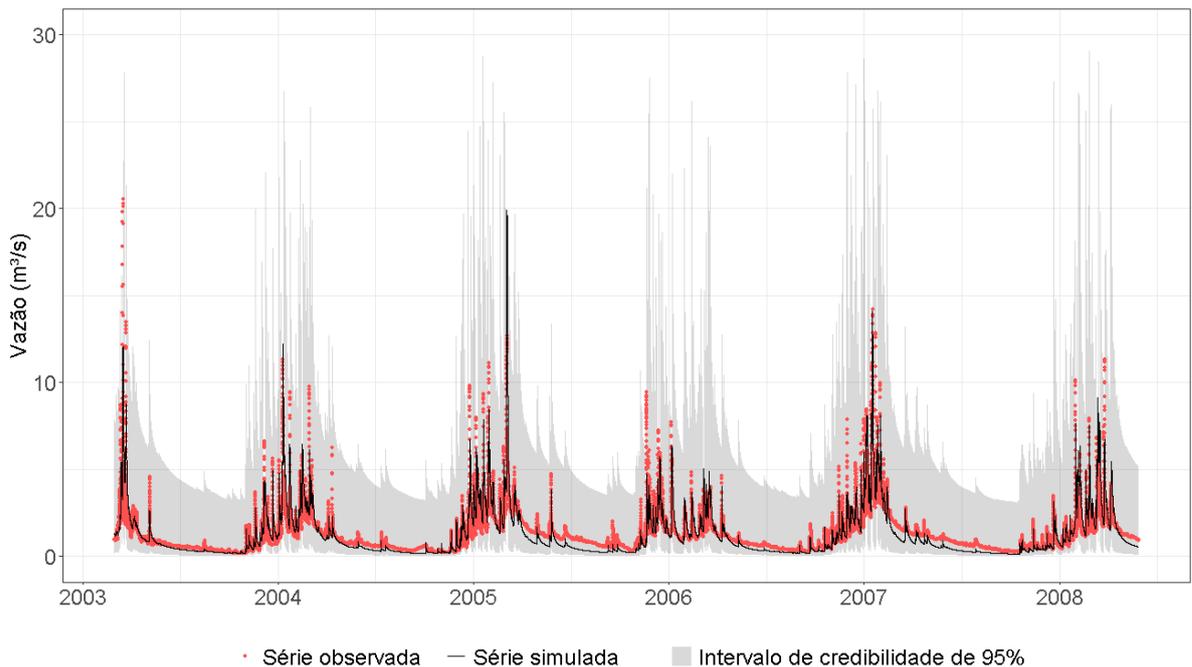


Figura 5.7 – Série temporal de vazões observadas e simuladas no domínio das assinaturas, para o período de validação - F_I

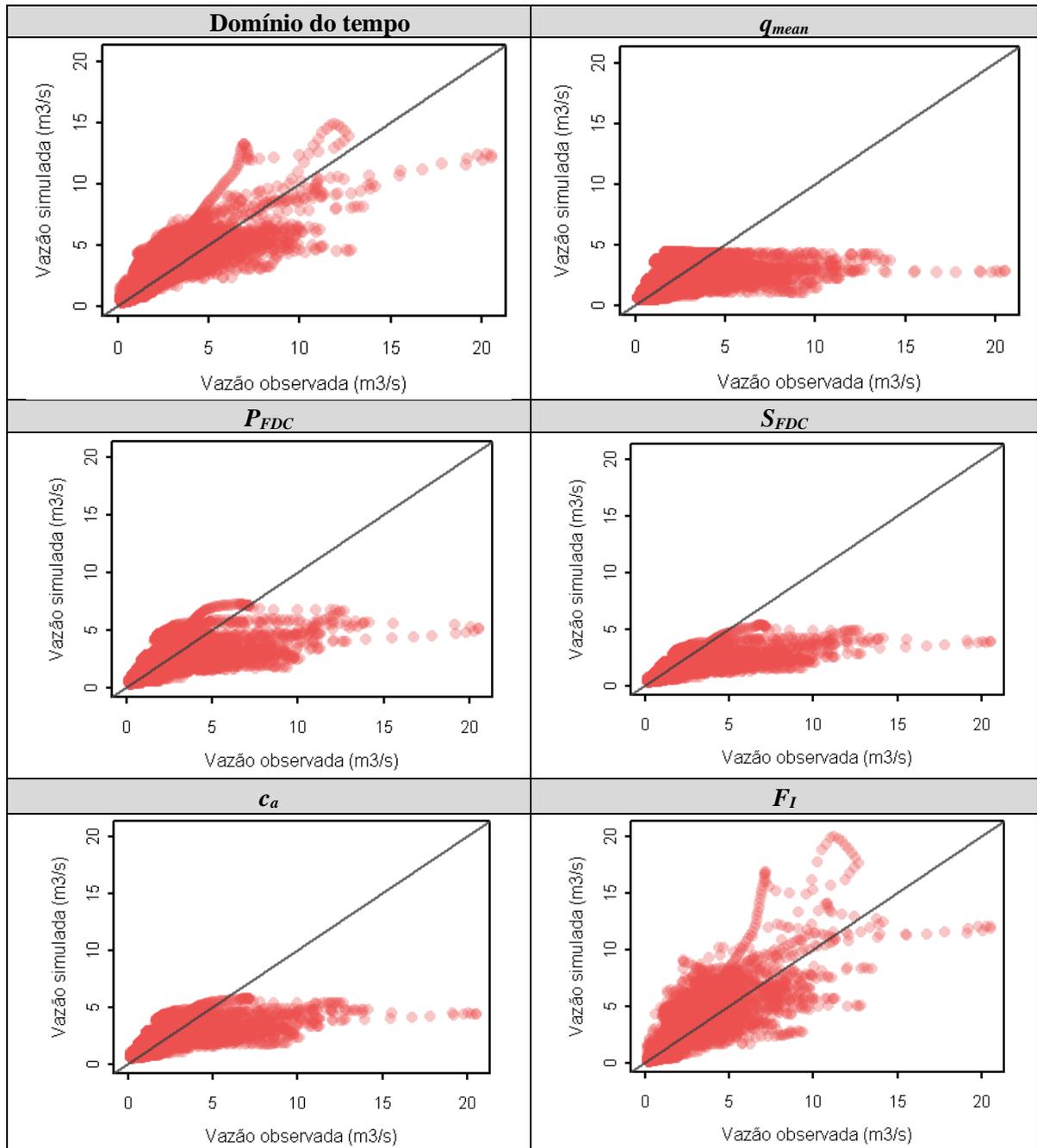


Figura 5.8 – Vazões simuladas versus vazões observadas (período de validação)

Considerando os resultados encontrados para P_{FDC} e c_a , que conseguiram reproduzir satisfatoriamente a resposta média da bacia às custas de subestimativas dos picos dos hidrogramas, e para F_I , cuja simulação conduziu a uma melhor estimativa dos picos, porém a um exagero quanto às vazões da recessão, propôs-se a combinação dessas assinaturas, buscando uma melhor reprodução dos hidrogramas de saída. Dessa forma, foram simuladas as combinações de P_{FDC} e F_I e de c_a e F_I . Os resultados são apresentados nas Tabelas 5.6 e 5.7. A Tabela 5.8 apresenta os percentuais das observações que estão contidas no intervalo de credibilidade de 95%.

Tabela 5.6 – Índices de desempenho - Combinações de assinaturas – Período de calibração

Variável	Domínio do tempo	Domínio das assinaturas				
		P_{FDC}	c_a	F_I	$P_{FDC} + F_I$	$c_a + F_I$
r	0,89	0,80	0,78	0,89	0,83	0,76
γ	1,10	0,97	0,99	0,98	0,96	1,02
α	1,19	0,89	0,76	1,48	0,94	0,87
KGE	0,76	0,77	0,67	0,51	0,82	0,73
$RMSE (mm/h)$	0,02	0,02	0,02	0,03	0,02	0,03

Tabela 5.7 – Índices de desempenho - Combinações de assinaturas – Período de validação

Variável	Domínio do tempo	Domínio das assinaturas				
		P_{FDC}	c_a	F_I	$P_{FDC} + F_I$	$c_a + F_I$
r	0,88	0,79	0,78	0,87	0,82	0,76
γ	1,22	1,09	1,20	1,00	1,10	1,19
α	1,06	0,81	0,72	1,26	0,87	0,82
KGE	0,74	0,70	0,60	0,71	0,75	0,64
$RMSE (mm/h)$	0,02	0,02	0,02	0,02	0,02	0,03

Tabela 5.8 – Percentual das observações contidas no intervalo de credibilidade de 95%

Período	Domínio do tempo	Domínio das assinaturas				
		P_{FDC}	c_a	F_I	$P_{FDC} + F_I$	$c_a + F_I$
Calibração	2%	23%	87%	100%	70%	60%
Validação	2%	19%	89%	100%	73%	67%

Em geral, não foram observadas alterações expressivas quanto aos índices de correlação r , à razão entre as médias das vazões γ nem ao valor do RMSE. Contudo, houve uma melhora considerável quanto à variabilidade das séries simuladas, devido à melhor representação dos picos dos hidrogramas, como apresentam as Figuras 5.9 e 5.10. Além disso, observa-se um aumento da incerteza proveniente da estimação dos parâmetros para $P_{FDC} + F_I$, em relação a P_{FDC} , enquanto para $c_a + F_I$, foi observado efeito contrário.

Ressalta-se que a combinação da curva de permanência e da taxa de mudança também apresentou bons resultados no trabalho de Fenicia *et al.* (2018). Todavia, divergentemente, esses autores observaram uma redução da incerteza quando da combinação das assinaturas.

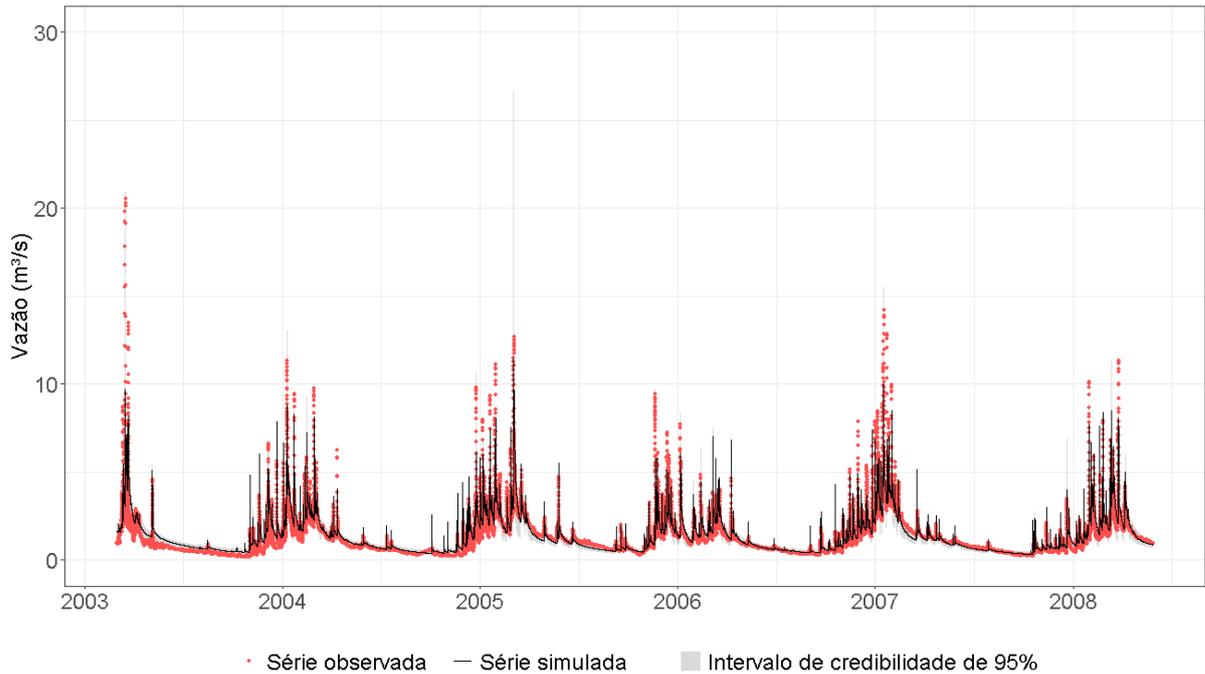


Figura 5.9 – Série temporal de vazões observadas e simuladas no domínio das assinaturas, para o período de validação - $P_{FDC} + F_I$

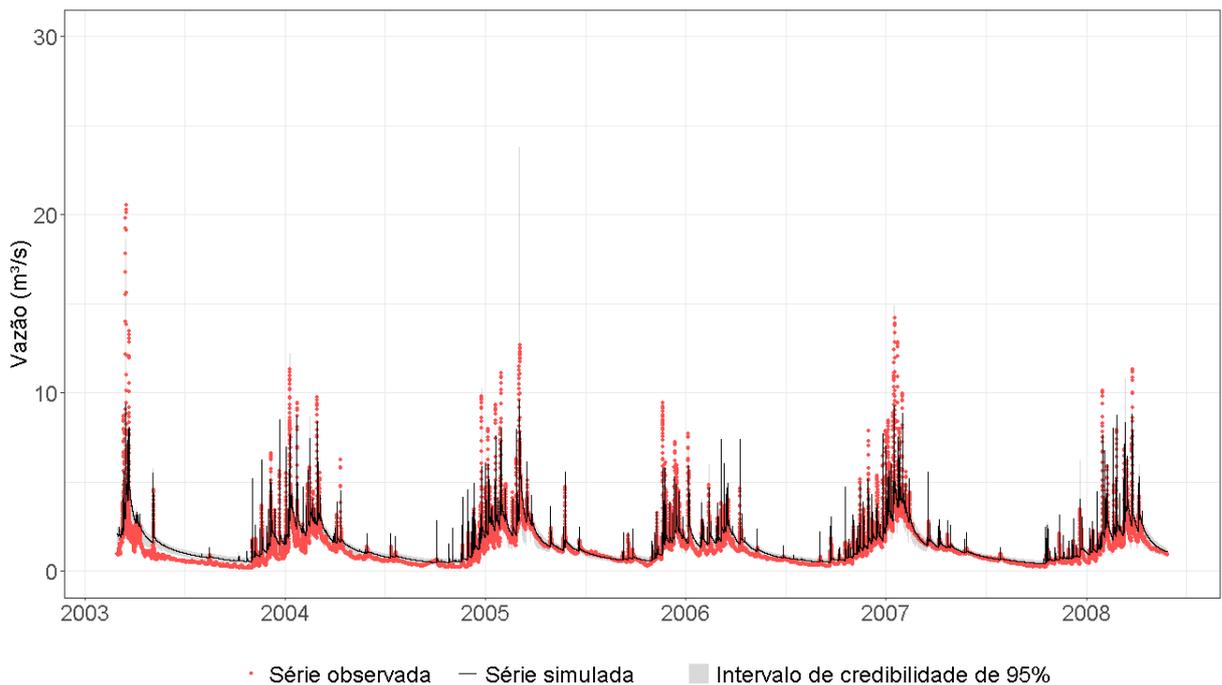


Figura 5.10 – Série temporal de vazões observadas e simuladas no domínio das assinaturas, para o período de validação - $c_a + F_I$

Como comentário geral com relação às modelagens em intervalo horário, seja no domínio do tempo, seja no domínio das assinaturas, destaca-se a grande dificuldade para representação da dinâmica da bacia, especialmente com relação a eventos de curta duração. Essa situação foi abordada por Massmann

(2020), que relatou maior homogeneidade das métricas de desempenho à medida que a escala de tempo da simulação aumenta.

No domínio do tempo, essa limitação fica evidente quando são analisados os resíduos da modelagem, que apresentam autocorrelação significativa mesmo para *lags* elevados (Figura 5.11). Ademais, a função de verossimilhança generalizada originalmente preconiza um modelo AR para tratamento da dependência dos resíduos, enquanto a verificação dos correlogramas total e parcial dos resíduos padronizados da modelagem sugerem uma maior aderência de um modelo do tipo ARMA, uma vez que aparenta haver uma componente de média móvel nos resíduos. Não obstante, os parâmetros considerados para a função de verossimilhança generalizada foram capazes de sintetizar o comportamento dos resíduos. Destaca-se que a consideração do modelo autorregressivo, segundo Schoups e Vrugt (2010), tinha fins práticos de modelagem e os autores reconhecem a possível necessidade de adequação das premissas consideradas.

Em um quadro geral, acredita-se que as limitações observadas em escala horária sejam reduzidas em escalas de tempo maiores. De acordo com Vaze *et al.* (2012) *apud* Jaiswal e Bharti (2020), modelos conceituais são indicados para modelagem em escala diária, enquanto modelos físicos podem ser empregados também em intervalos sub-diários. Ademais, é válido mencionar que não foram encontrados, na literatura, trabalhos que fizessem uso de métodos ABC ou mesmo MCMC sob o paradigma Bayesiano para estimação de parâmetros de um modelo hidrológico horário, nem estudos de modelagem da bacia em estudo, nesta escala de tempo, dificultando a comparação dos resultados.

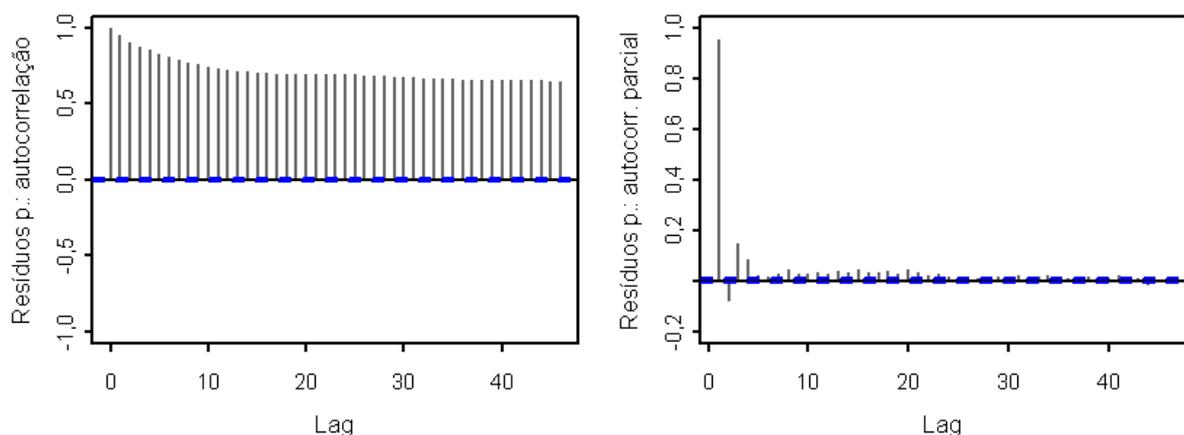


Figura 5.11 – Correlogramas total e parcial dos resíduos no domínio do tempo, empregando o algoritmo DREAM e a função de verossimilhança generalizada

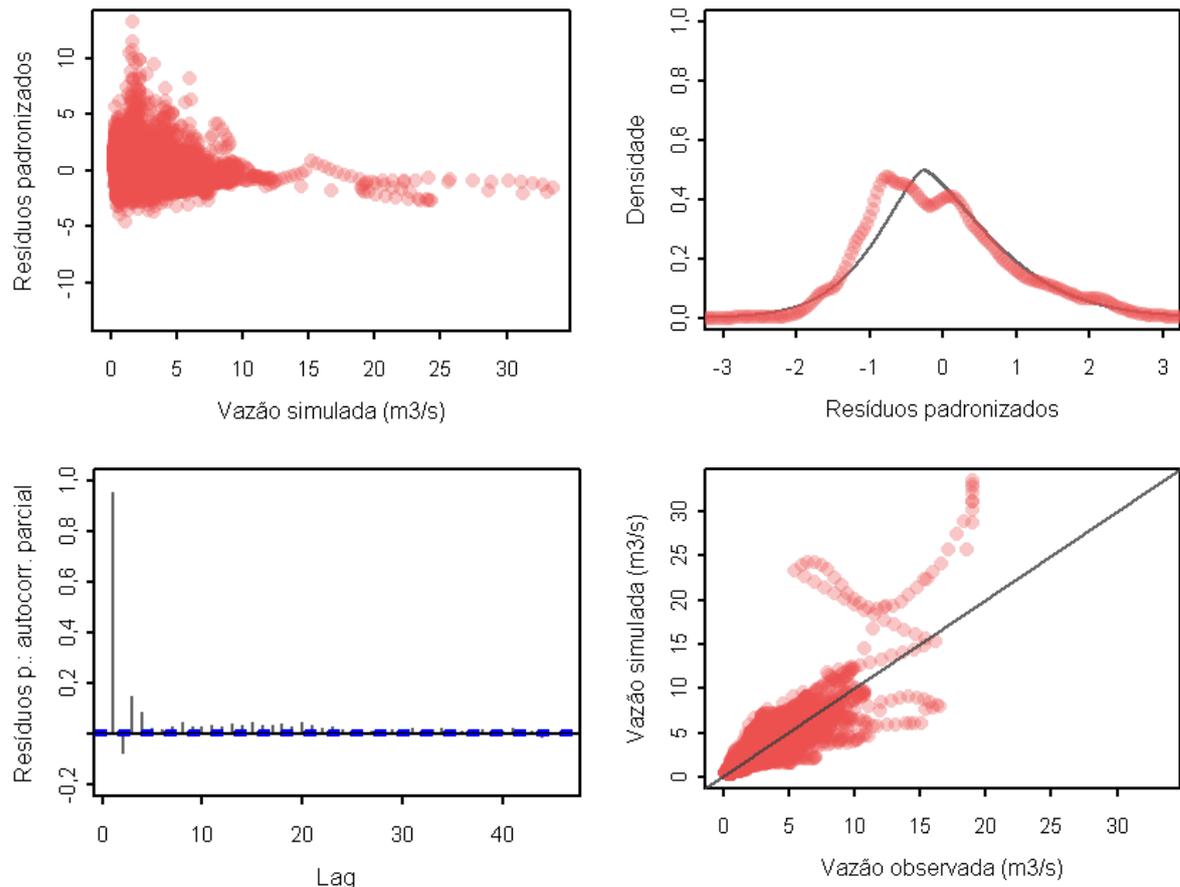


Figura 5.12 – Comportamento dos resíduos no domínio do tempo, empregando o algoritmo DREAM e a função de verossimilhança generalizada

Os resultados aqui apresentados corroboram com a literatura no que se refere à viabilidade do emprego de assinaturas hidrológicas para estimação de parâmetros de modelos hidrológicos (FENICIA *et al.*, 2018; KAVETSKI *et al.*, 2018; SAHRAEI *et al.*, 2020). Todavia, é importante mencionar que a utilização de métodos ABC para esse fim, apesar de eficazes, mostraram-se pouco eficientes, com alto custo computacional. Enquanto foram necessárias cerca de 25.000 iterações para convergência das cadeias no método DREAM, o algoritmo SABC necessitou de mais de 700.000 iterações para aproximação das distribuições *a posteriori* dos parâmetros.

Por fim, quanto à avaliação da incorporação do erro aleatório segundo modelo AR(1) sugerido por Kavetski *et al.* (2018), não foram encontradas evidências que corroborem as conclusões dos autores, tendo sido obtidos resultados praticamente idênticos para $h(\theta_h, x)$ e $Q(\theta, x)$. Por outro lado, para os erros gerados a partir da SEP, não foi possível alcançar a convergência das cadeias, mesmo para um número elevado de iterações. Por isso, todos os resultados apresentados neste trabalho baseiam-se na consideração da série de vazões simuladas a partir do modelo determinístico $h(\theta_h, x)$.

5.2.2 Avaliação das distribuições *a posteriori* dos parâmetros do modelo

Avaliar a contribuição isolada de cada parâmetro de modelagem no hidrograma simulado é, por si só, um grande desafio, devido às correlações entre os mesmos. Neste trabalho, observou-se, ainda, que a correlação varia expressivamente de acordo com a assinatura considerada para estimação dos parâmetros do modelo, conforme apresenta a Figura 5.13. A flutuação é tamanha que dois parâmetros quaisquer (θ_1 e θ_2 , por exemplo), podem estar correlacionados positiva (P_{FDC}) ou negativamente (S_{FDC}), dependendo da simulação. É possível inferir, portanto, que há um problema de identificabilidade quanto à estrutura do modelo (SHIN *et al.*, 2015).

Dessa forma, é possível que haja diferentes dinâmicas de compensação entre os diversos parâmetros do modelo, inviabilizando a verificação de relações diretas entre cada parâmetro e a resposta simulada da bacia. A seguir, são feitos comentários gerais a respeito das distribuições marginais *a posteriori* dos parâmetros do modelo.

O aspecto mais evidente com relação às distribuições *a posteriori* dos parâmetros do modelo é a maior concentração em um pequeno intervalo de variação, para o domínio do tempo, dada a menor incerteza decorrente da estimação dos parâmetros (Figura 5.14). A combinação das assinaturas, embora acarrete uma redução da variação de alguns dos parâmetros, não é capaz de reduzir as discrepâncias (Figura 5.15).

Além disso, as distribuições dos parâmetros θ_1 , θ_2 e θ_3 têm formas bastante variadas entre si, mesmo para assinaturas cujo desempenho é relativamente semelhante, configurando a situação de equifinalidade descrita por Beven (1993). É interessante reparar que, para P_{FDC} , esses parâmetros apresentam distribuição bimodal que engloba os intervalos de variação dos respectivos parâmetros no domínio do tempo.

Quanto ao parâmetro θ_4 , esse parece ser pouco afetado pelos dados para a maioria das assinaturas, à exceção de F_I , o que é mais manifestado quando da combinação desta assinatura com P_{FDC} e com c_a . Os reduzidos valores desse parâmetro, em relação aos estimados a partir das demais assinaturas, podem estar relacionados à melhor representação dos picos dos hidrogramas, devido à redução do tempo de base do UH1.

Dom. Tempo	q_{mean}
-------------------	------------

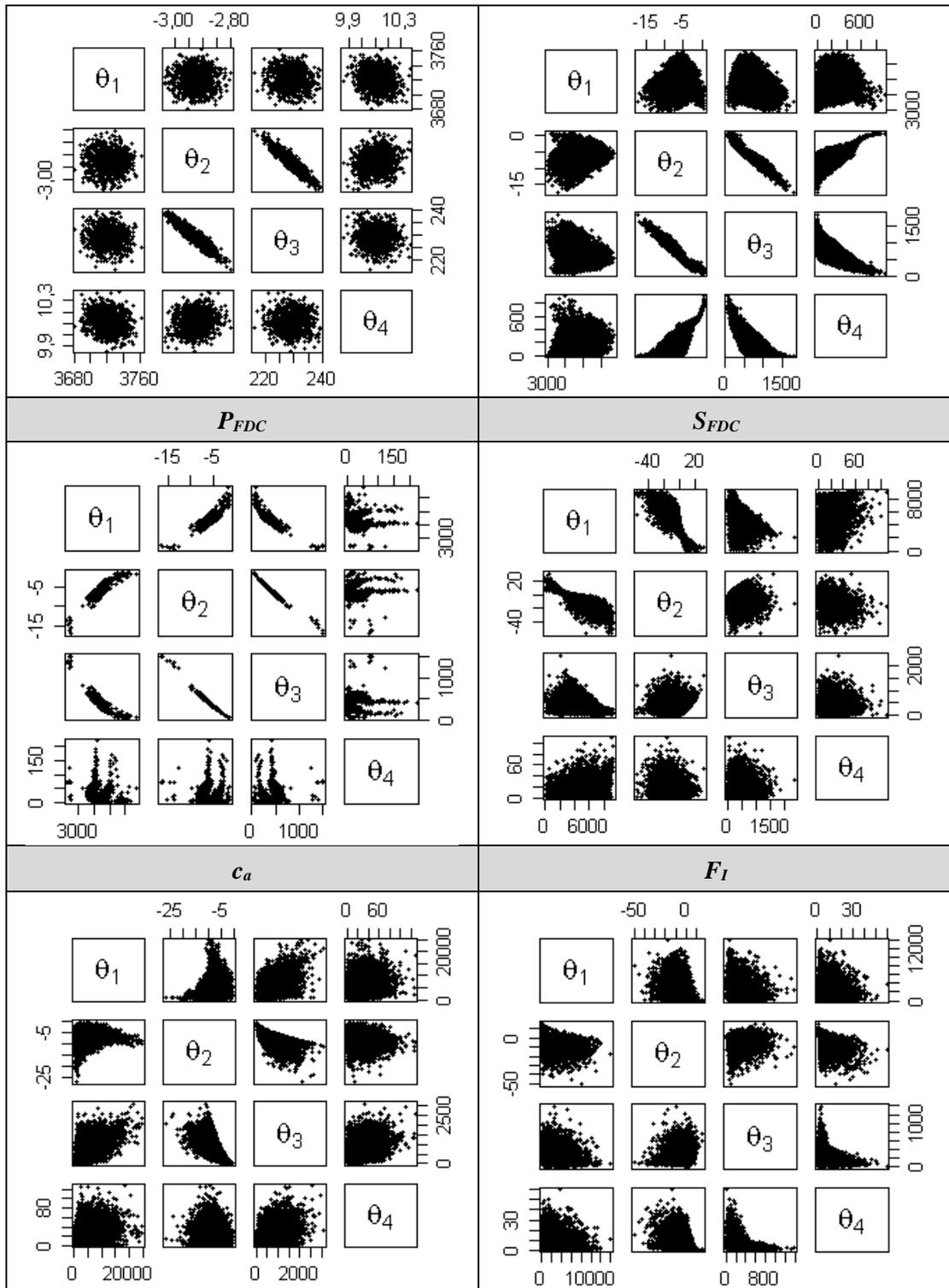


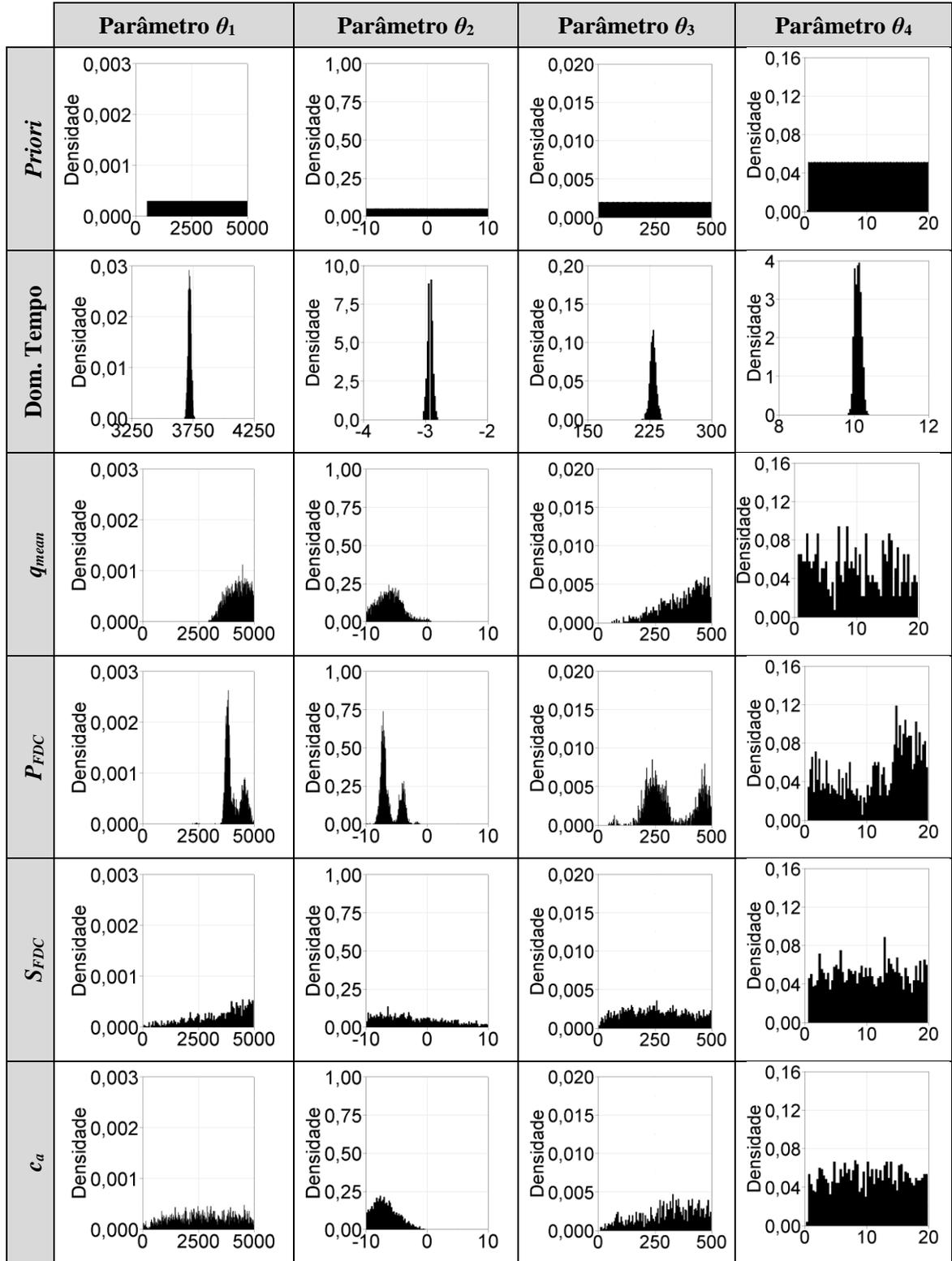
Figura 5.13 – Gráficos de dispersão dos parâmetros do modelo.

É possível verificar ainda que as distribuições dos parâmetros para as assinaturas S_{FDC} e F_I são, em geral, as que apresentam maiores disparidades em relação às demais. Como mencionado no item

anterior, tais assinaturas acarretaram intervalos significativamente mais largos para a incerteza paramétrica.

Por fim, com relação aos parâmetros estimados a partir das combinações de assinaturas, é possível inferir uma conjunção das distribuições de cada assinatura isoladamente: em geral, a distribuição decorrente da combinação de assinaturas apresenta valores intermediários aos das distribuições de cada assinatura isoladamente. Além disso, é notória a similaridade entre as distribuições *a posteriori* de todos os parâmetros do modelo, para as combinações propostas (Figura 5.16).

Os resultados aqui apresentados, dada a grande variabilidade das distribuições *a posteriori* encontradas, indicam a necessidade de novos estudos, avaliando-se diferentes bacias, escalas de tempo e modelos, para que se possa tecer conclusões e recomendações mais assertivas acerca da metodologia proposta e dos resultados encontrados.



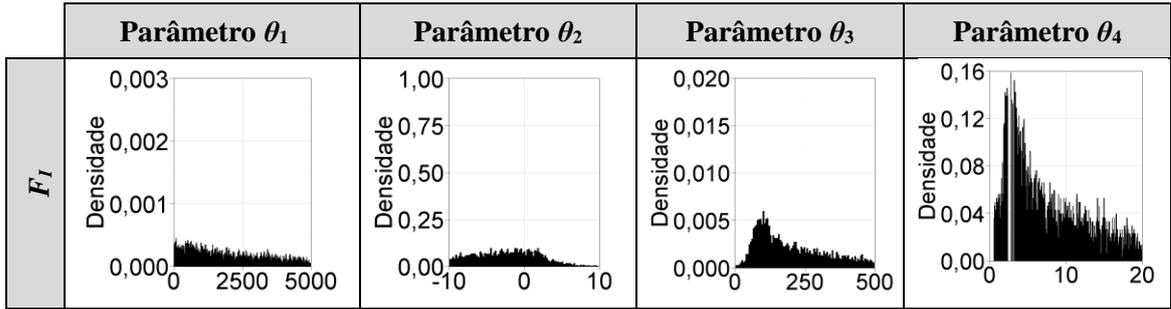
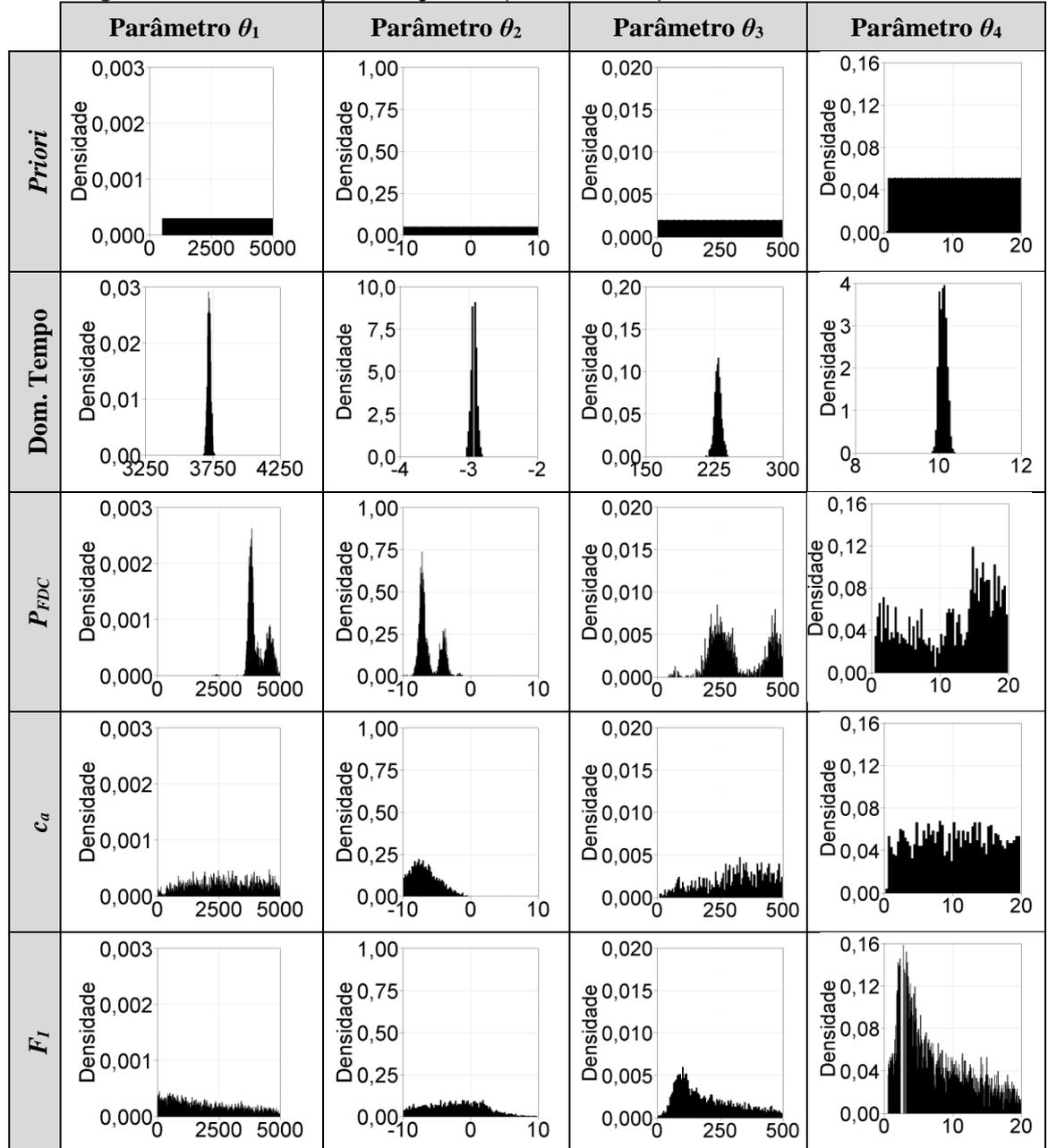


Figura 5.14 – Distribuições marginais *a posteriori* dos parâmetros do modelo GR4H.



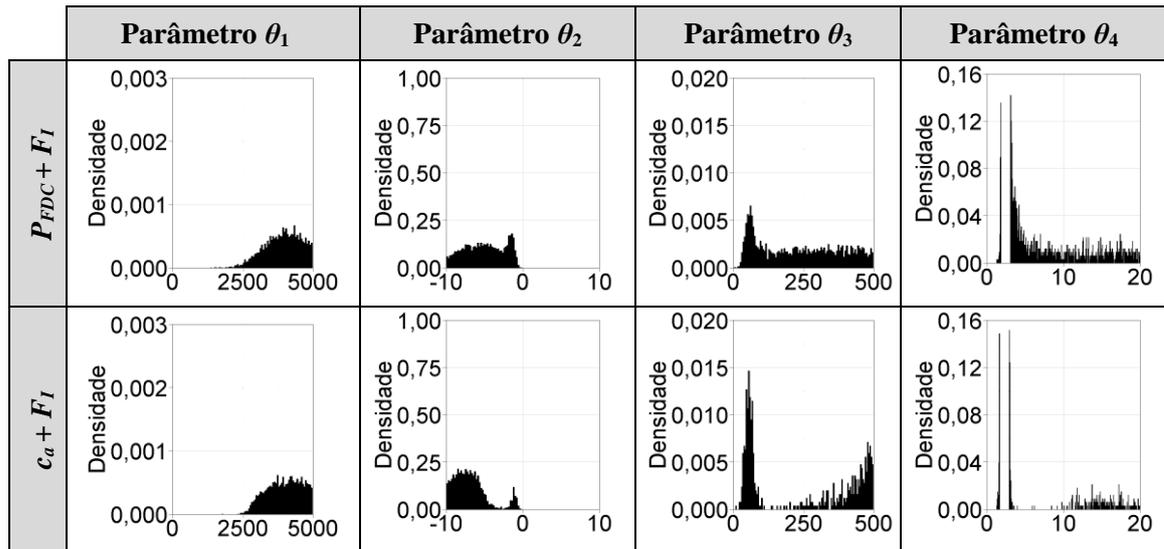


Figura 5.15 – Distribuições marginais *a posteriori* dos parâmetros do modelo GR4H para as combinações propostas para as assinaturas.

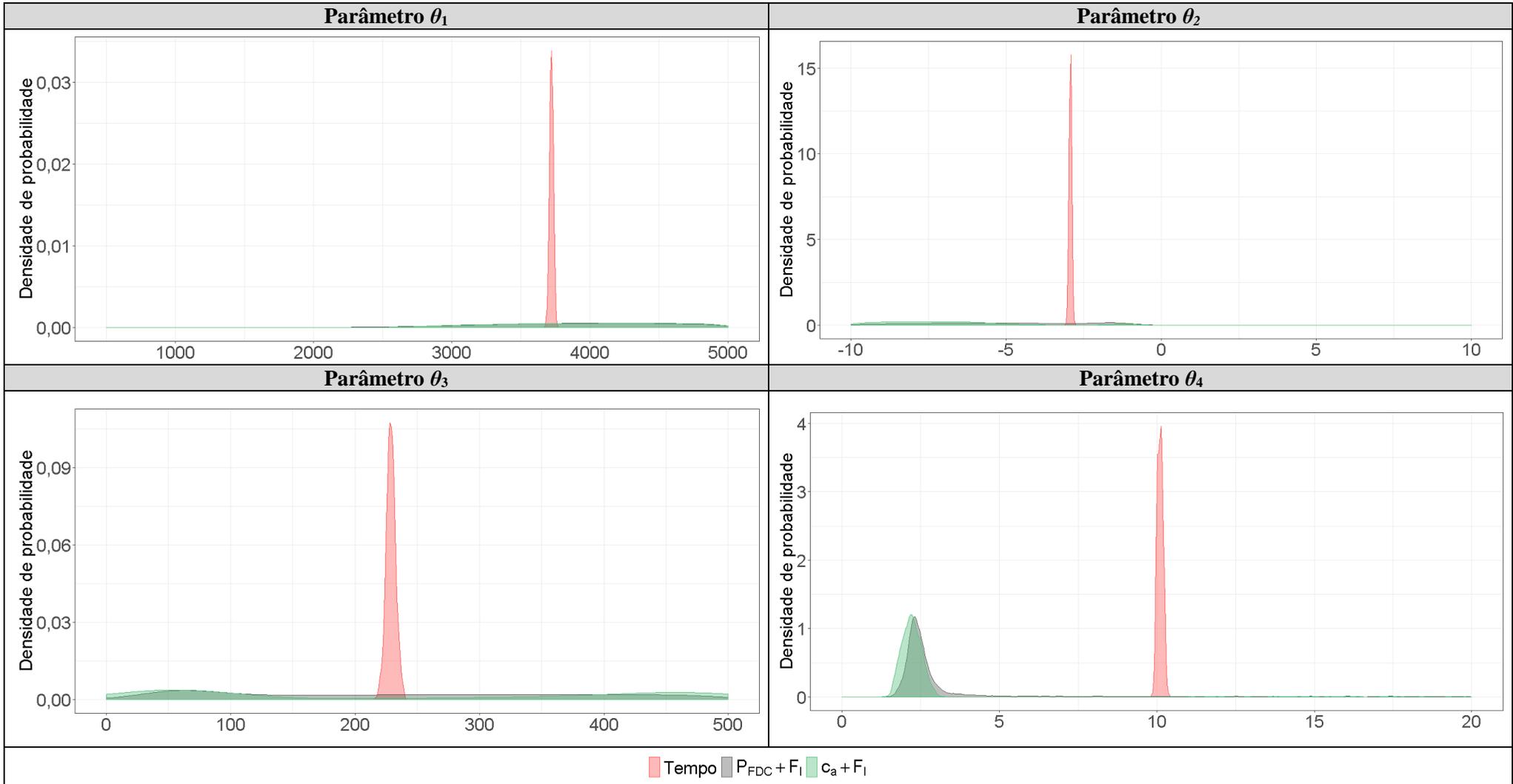


Figura 5.16 – Distribuições *a posteriori* dos parâmetros do modelo GR4H para as combinações propostas para as assinaturas.

5.3 Domínio tempo versus domínio assinaturas: escala mensal

Neste item, são apresentados os resultados das simulações em escala mensal, considerando a série modelada a partir do algoritmo DREAM e da função de verossimilhança generalizada como *benchmark* para comparação dos resultados no domínio das assinaturas.

5.3.1 Avaliação do desempenho da modelagem

As Tabelas 5.9 e 5.10 apresentam as métricas de desempenho para os períodos de calibração e validação, respectivamente. A Tabela 5.11 apresenta os percentuais das observações que estão contidas no intervalo de credibilidade de 95%.

Tabela 5.9 – Índices de desempenho da modelagem em escala mensal - Período de calibração

Estatística	Domínio do tempo	Domínio das assinaturas			
		q_{mean}	$PFDC$	$SFDC$	c_a
r	0,97	0,97	0,97	0,96	0,97
γ	1,01	1,07	0,93	21,38	0,99
α	1,04	1,03	0,95	20,15	1,04
KGE	0,95	0,91	0,91	-26,96	0,95
$RMSE$ (mm/mês)	4,96	5,56	5,27	737,02	4,89

Tabela 5.10 – Índices de desempenho da modelagem em escala mensal - Período de validação

Estatística	Domínio do tempo	Domínio das assinaturas			
		q_{mean}	$PFDC$	$SFDC$	c_a
r	0,96	0,96	0,96	0,93	0,96
γ	1,08	1,15	0,99	21,62	1,05
α	0,94	0,94	0,87	18,42	0,95
KGE	0,90	0,83	0,86	-25,99	0,92
$RMSE$ (mm/mês)	5,60	6,80	5,49	663,54	5,34

Tabela 5.11 – Percentual das observações contidas no intervalo de credibilidade de 95%

Período	Domínio do tempo	Domínio das assinaturas			
		q_{mean}	$PFDC$	$SFDC$	c_a
Calibração	39%	8%	9%	0%	85%
Validação	23%	13%	11%	0%	86%

Nessa escala de simulação, observa-se novamente que as séries modeladas a partir das assinaturas q_{mean} , P_{FDC} e c_a conduziram a melhores índices de desempenho, com boa representação da correlação entre séries simuladas e série observada, além de boa reprodução da resposta da bacia.

Para as demais assinaturas, no entanto, as vazões simuladas mostraram-se significativamente maiores do que as observadas, acarretando valores bem superiores a 1 para γ e para α . É válido mencionar que o valor de S_{FDC} calculado a partir da série simulada para c_a , que apresentou o melhor desempenho em escala mensal dentre as assinaturas propostas, é igual a 2,37. Conforme apresentado na Tabela 5.2, tem-se $S_{FDC} = 1,95$ para os dados da estação Jardim. Dessa forma, o fraco desempenho dessa assinatura não se deve a problemas com relação à aproximação computacional, e sim à capacidade dessa assinatura em prever a resposta da bacia, na escala de tempo considerada para a modelagem. Quanto ao RMSE, os valores encontrados nos domínios do tempo e das assinaturas se mostraram bastante semelhantes e inferiores ao desvio-padrão da série observada (22 mm/mês).

A seguir, nas Figuras 5.17 a 5.20, são apresentados os hidrogramas obtidos para algumas das simulações realizadas, considerando todo o período simulado. Também são apresentados os intervalos de 95% de incerteza relacionada à estimativa dos parâmetros (hachura cinza) e a separação dos períodos de calibração e validação (linha pontilhada). Os demais hidrogramas são apresentados no Apêndice II.

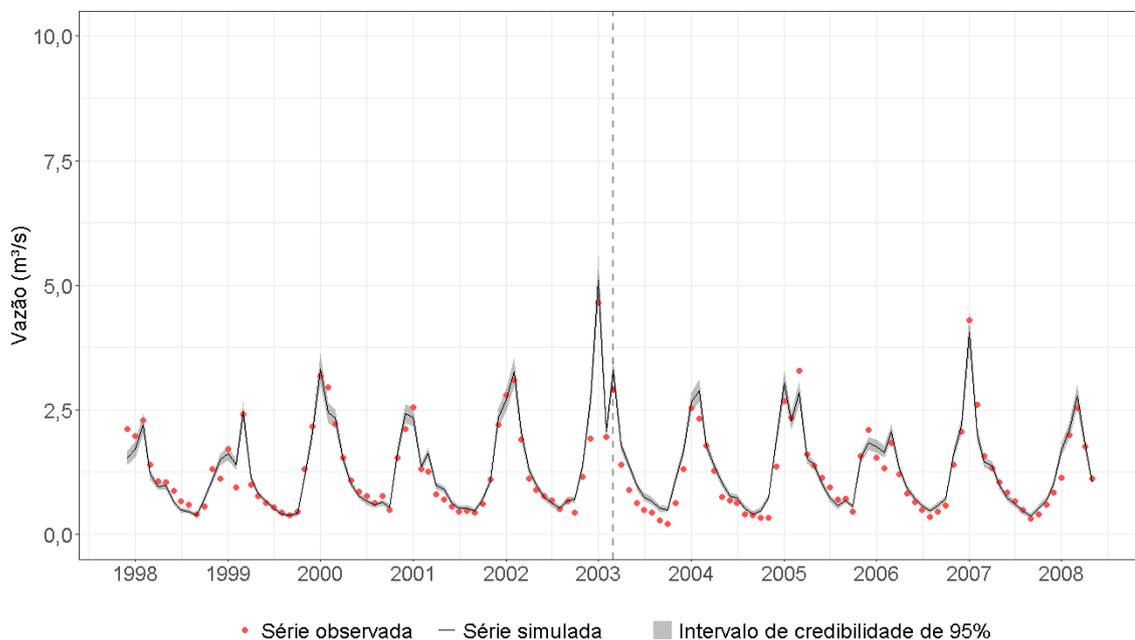


Figura 5.17 – Série temporal de vazões observadas e simuladas no domínio do tempo.

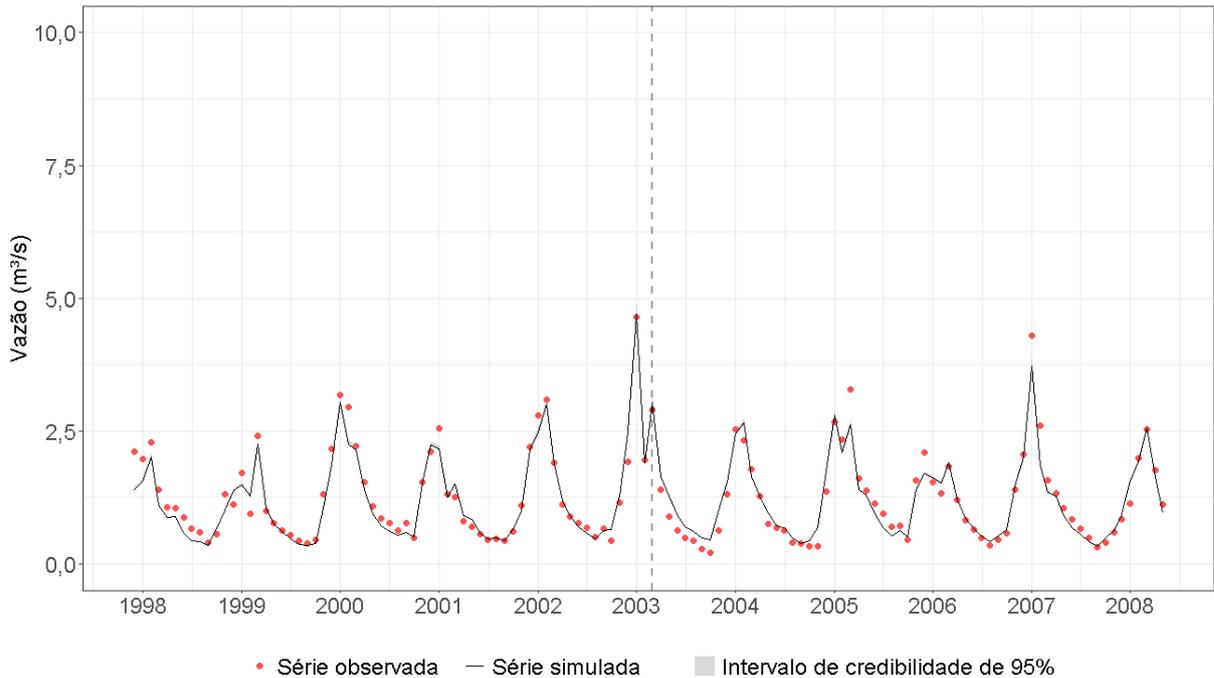


Figura 5.18 – Série temporal de vazões observadas e simuladas no domínio das assinaturas - P_{FDC}

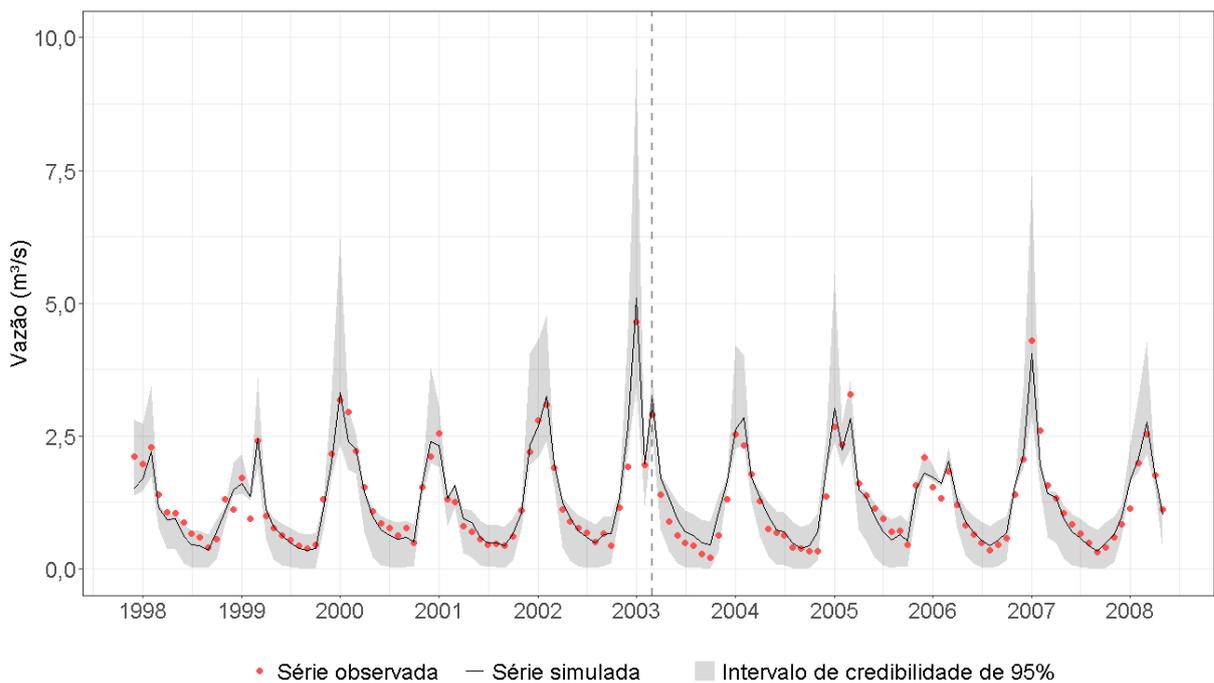


Figura 5.19 – Série temporal de vazões observadas e simuladas no domínio das assinaturas - c_a

Assim como para o modelo horário, as assinaturas P_{FDC} e c_a conduziram a hidrogramas resultantes bastantes semelhantes, sendo o intervalo de credibilidade significativamente mais largo para o coeficiente de escoamento superficial.

Ressalta-se que, conforme esperado e discutido no item 5.2.1, a adoção de uma escala de tempo maior para o modelo mostrou-se mais adequada, tendo sido encontrada maior uniformidade quanto às métricas de desempenho. A consideração de uma única assinatura para estimação dos parâmetros também se mostrou aceitável, alcançando resultados semelhantes aos do domínio do tempo, apesar de uma incerteza relativa maior.

Adicionalmente, a análise da Figura 5.20 permite inferir que a adoção de um modelo AR, para essa escala de tempo, é condizente, não sendo mais encontrados indícios de uma componente de média móvel nos resíduos da simulação. Por fim, a Figura 5.21 apresenta o comportamento dos resíduos no domínio do tempo, para o período de calibração, em que é possível observar que as premissas assumidas com relação à autocorrelação e à homocedasticidade não foram violadas.

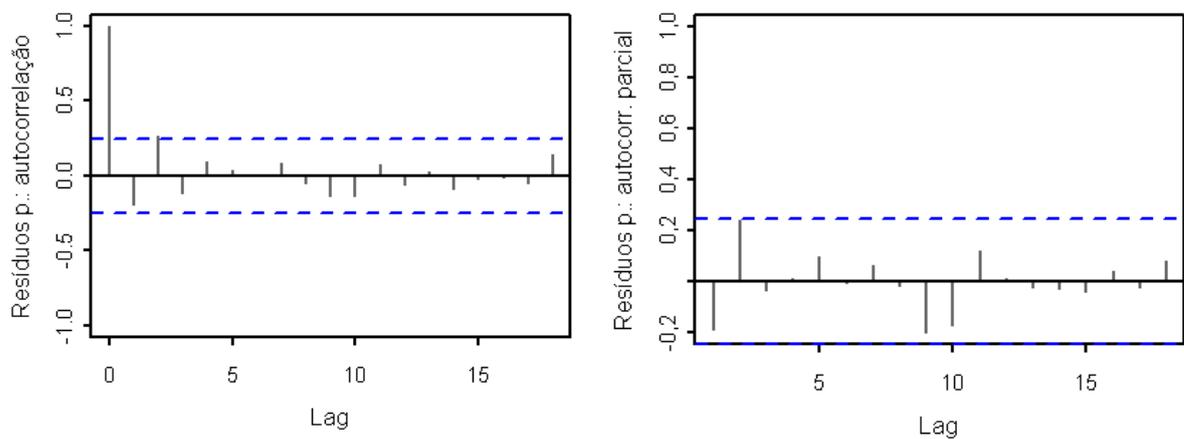


Figura 5.20 – Correlogramas total e parcial dos resíduos no domínio do tempo, empregando o algoritmo DREAM e a função de verossimilhança generalizada.

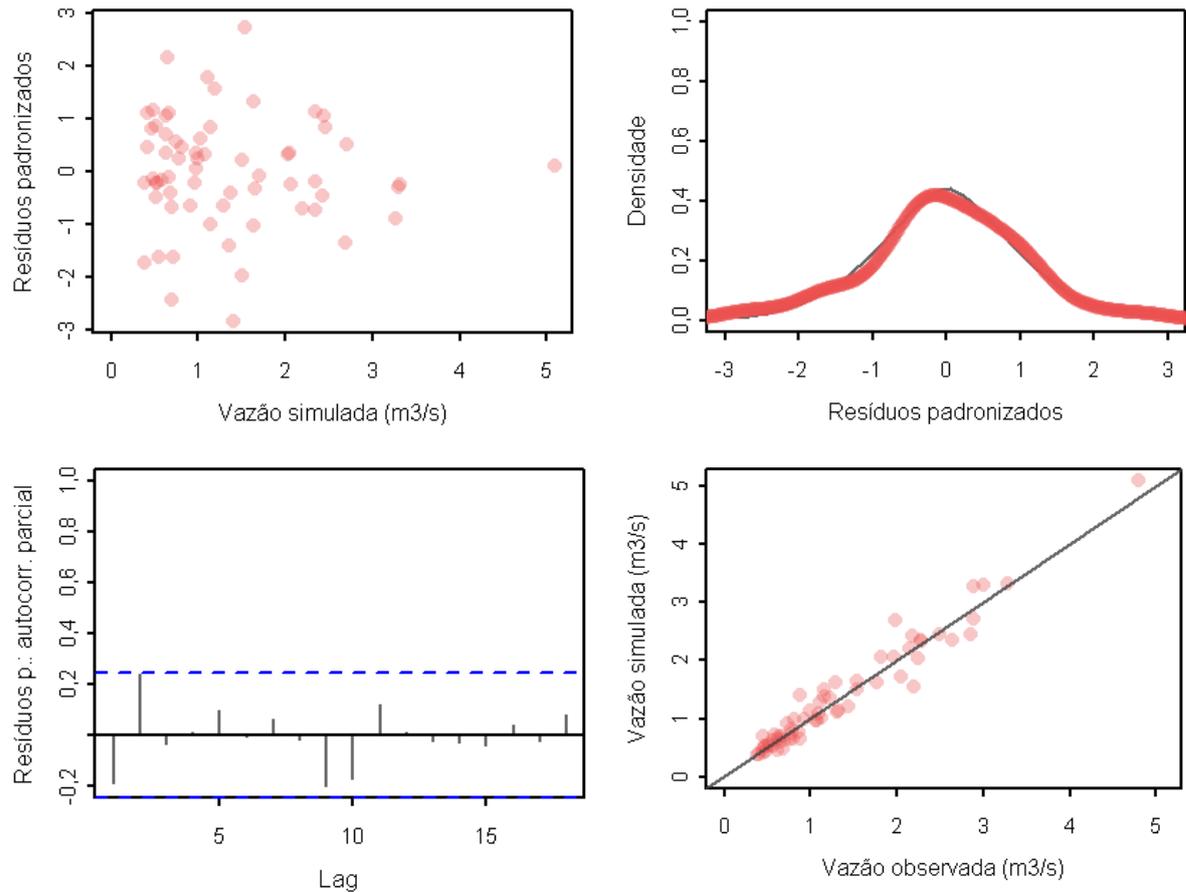


Figura 5.21 – Comportamento dos resíduos no domínio do tempo.

5.3.2 Avaliação das distribuições *a posteriori* dos parâmetros dos modelos

Em comparação com os parâmetros do modelo horário, os parâmetros θ_1 e θ_2 do modelo GR2M apresentaram menor variação do coeficiente de correlação de Pearson, como apresentado na Tabela 5.12. É interessante reparar que, para os parâmetros estimados a partir de c_a , a correlação é quase nula, sendo moderada ou fortemente negativa para as demais assinaturas. A Figura 5.22 apresenta os gráficos de dispersão dos parâmetros do modelo.

Tabela 5.12 – Coeficiente de correlação de Pearson entre os parâmetros θ_1 e θ_2 do modelo GR2M

Estatística	Domínio do tempo	Domínio das assinaturas			
		q_{mean}	P_{FDC}	S_{FDC}	c_a
r	-0,41	-0,99	-0,95	-0,71	0,02

A Figura 5.23 apresenta as distribuições marginais *a posteriori* dos parâmetros do modelo GR2M, em que é possível observar, assim como no modelo horário, grandes diferenças mesmo para assinaturas hidrológicas que apresentaram desempenho semelhante (P_{FDC} e c_a , por exemplo). Isso é mais evidente com relação ao parâmetro θ_1 , relacionado ao reservatório de produção, para o qual a assinatura c_a guarda semelhança com a distribuição *a priori*, enquanto P_{FDC} é fortemente afetado pelos dados.

A distribuição de θ_1 para S_{FDC} , que apresenta assimetria à direita, está contida em intervalos cujos valores são consideravelmente menores do que os estimados para as demais assinaturas, o que, em algum grau, pode estar relacionado às superestimativas das vazões simuladas para essa assinatura.

Com relação ao parâmetro θ_2 , relacionado às trocas de água com as bacias vizinhas, as distribuições para as assinaturas q_{mean} , P_{FDC} e c_a , bem como para o domínio do tempo, são bastante concentradas por volta do valor 0,65 mm/mês. Porém, para S_{FDC} , a distribuição praticamente não foi impactada pelos dados.

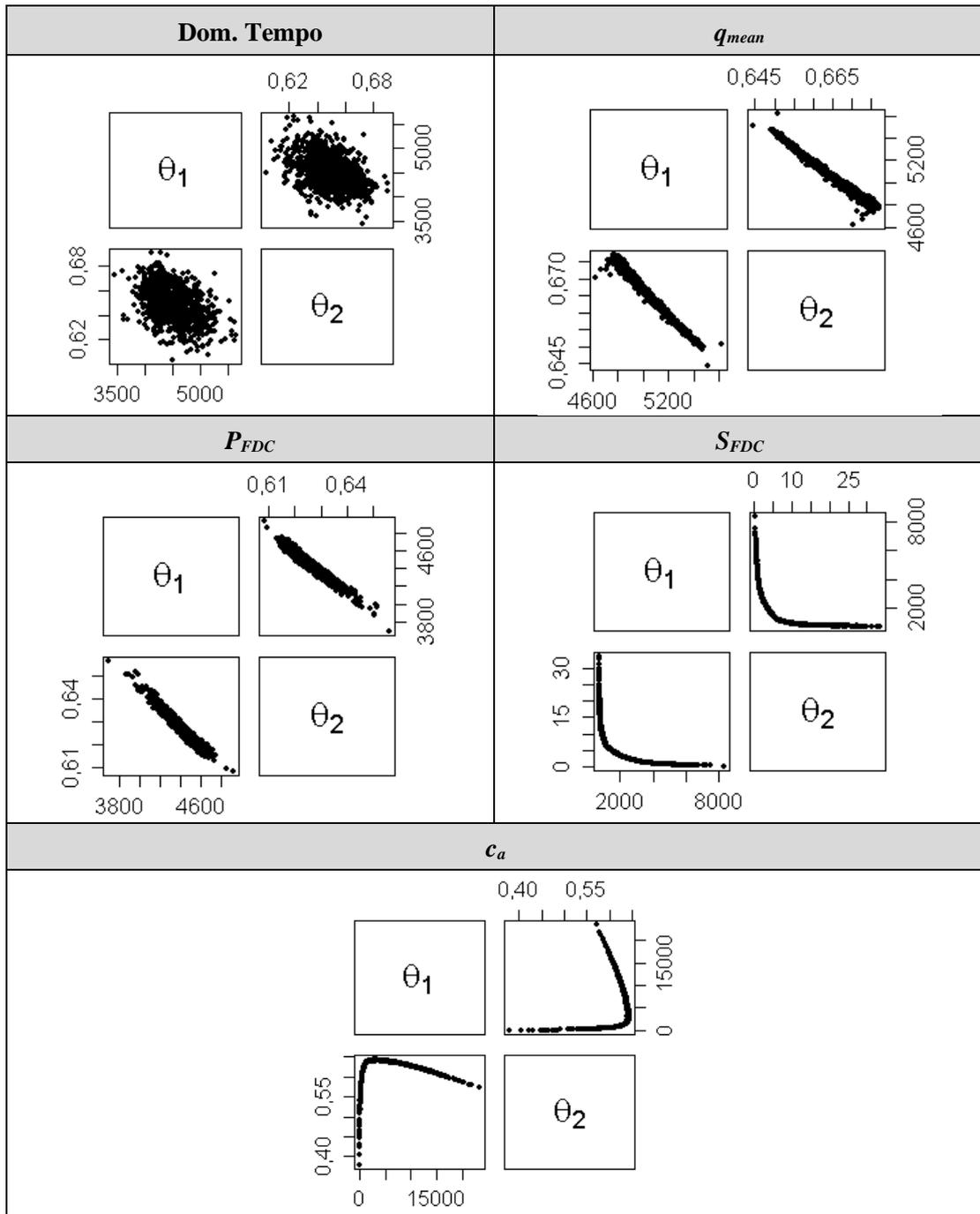


Figura 5.22 – Gráficos de dispersão dos parâmetros do modelo.

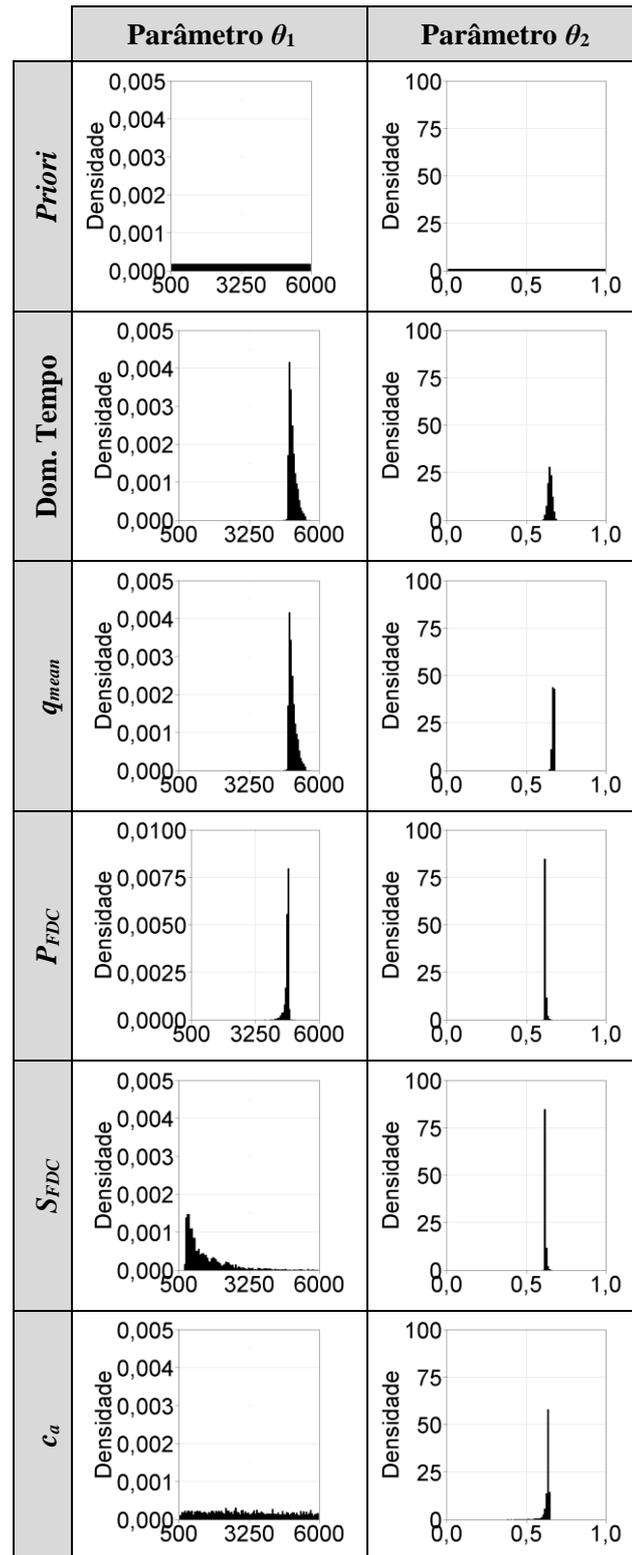


Figura 5.23 – Distribuições marginais dos parâmetros do modelo GR2M

5.4 Desempenho da modelagem no domínio das assinaturas inferidas a partir de registros fluviométricos regionais

Neste item, são apresentados os resultados das simulações horária, considerando a série modelada a partir assinaturas regionais estimadas a partir de registros fluviométricos diários. Destaca-se que, devido à diferença relevante entre os valores estimados para a taxa de mudança F_L , considerando a estação Jardim e a mediana das estações no entorno da bacia, essa assinatura foi desconsiderada nesta avaliação.

As Tabelas 5.13 e 5.14 apresentam as métricas de desempenho para os períodos de calibração e validação, respectivamente. A Tabela 5.15 apresenta os percentuais das observações que estão contidas no intervalo de credibilidade de 95%.

Tabela 5.13 – Índices de desempenho da modelagem a partir das assinaturas regionais - Período de calibração

Estatística	Domínio das assinaturas			
	q_{mean}	P_{FDC}	S_{FDC}	c_a
r	0,68	0,63	0,79	0,81
γ	1,17	0,84	0,75	1,26
α	0,85	0,67	0,61	0,98
KGE	0,61	0,48	0,49	0,67
$RMSE (mm/h)$	0,03	0,03	0,03	0,03

Tabela 5.14 – Índices de desempenho da modelagem a partir das assinaturas regionais - Período de validação

Estatística	Domínio das assinaturas			
	q_{mean}	P_{FDC}	S_{FDC}	c_a
r	0,70	0,66	0,80	0,81
γ	1,42	0,97	0,92	1,48
α	0,81	0,63	0,57	0,93
KGE	0,46	0,50	0,52	0,48
$RMSE (mm/h)$	0,03	0,03	0,02	0,03

Tabela 5.15 – Percentual das observações contidas no intervalo de credibilidade de 95%

Estatística	Domínio das assinaturas			
	q_{mean}	P_{FDC}	S_{FDC}	c_a
Calibração	11%	96%	100%	81%
Validação	14%	96%	100%	83%

Assim como para as assinaturas locais, observou-se uma menor variabilidade entre os valores simulados e a série observada ($\alpha < 1$), menos pronunciada para c_a . Para essa assinatura, inclusive, houve uma melhora expressiva dessa estatística, em relação às simulações a partir das assinaturas locais (de 0,76 para 0,98, para o período de calibração, e de 0,72 para 0,93, para o período de validação). Situação contrária ocorreu para P_{FDC} , que teve reduções de 0,89 e 0,81 para 0,67 e 0,63, respectivamente. Para q_{mean} , houve um sutil aumento do valor de α , e praticamente não foram observadas variações com relação a S_{FDC} .

Quanto à razão entre as médias das vazões simuladas e observadas, foi observado padrão semelhante ao encontrado para as assinaturas locais: em geral, q_{mean} e c_a tendem a superestimar as vazões simuladas, enquanto as assinaturas baseadas na curva de permanência tendem a subestimar. Esse comportamento, no entanto, é mais pronunciado para as séries calculadas a partir das assinaturas regionais, sendo encontrados valores de β mais distantes da unidade. Com relação ao RMSE, observa-se um aumento dos valores dessa métrica, em comparação com as assinaturas locais.

É interessante observar que, com exceção de P_{FDC} , que apresentou uma piora substancial do desempenho da modelagem, e apesar das variações das métricas avaliadas, a resposta simulada da bacia é razoavelmente semelhante à estimada a partir das assinaturas locais. Isso valida a hipótese de regionalização das assinaturas e demonstra o potencial dessa abordagem para predição em bacias não monitoradas.

Adicionalmente, deve-se destacar que a discrepância com relação às extensões das séries utilizadas para cálculo das assinaturas locais e regionais não aparenta ter impacto significativo no desempenho da modelagem. Com efeito, essa observação corrobora com o estudo de Mohammed *et al.* (2021), que, avaliando diferentes extensões de séries de monitoramento, não encontraram diferenças expressivas para desempenho da modelagem no domínio das assinaturas. Contudo, Fenicia *et al.* (2018) relatam uma redução da incerteza paramétrica decorrente da extensão da série monitorada.

Nas Figuras 5.24 a 5.27, são apresentados os hidrogramas obtidos para as simulações realizadas, considerando o período de validação. Também são apresentados os intervalos de 95% de incerteza relacionada à estimativa dos parâmetros. Os demais hidrogramas são apresentados no Apêndice III.

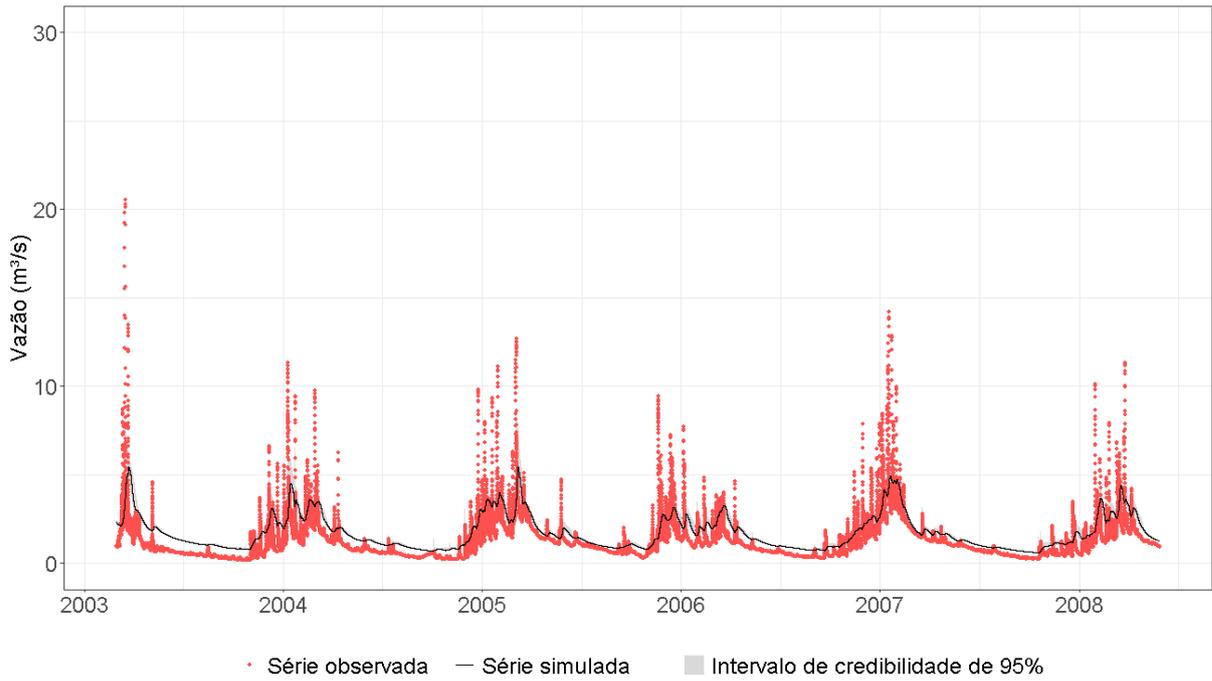


Figura 5.24 – Série temporal de vazões observadas e simuladas no domínio das assinaturas, para o período de validação - q_{mean} .

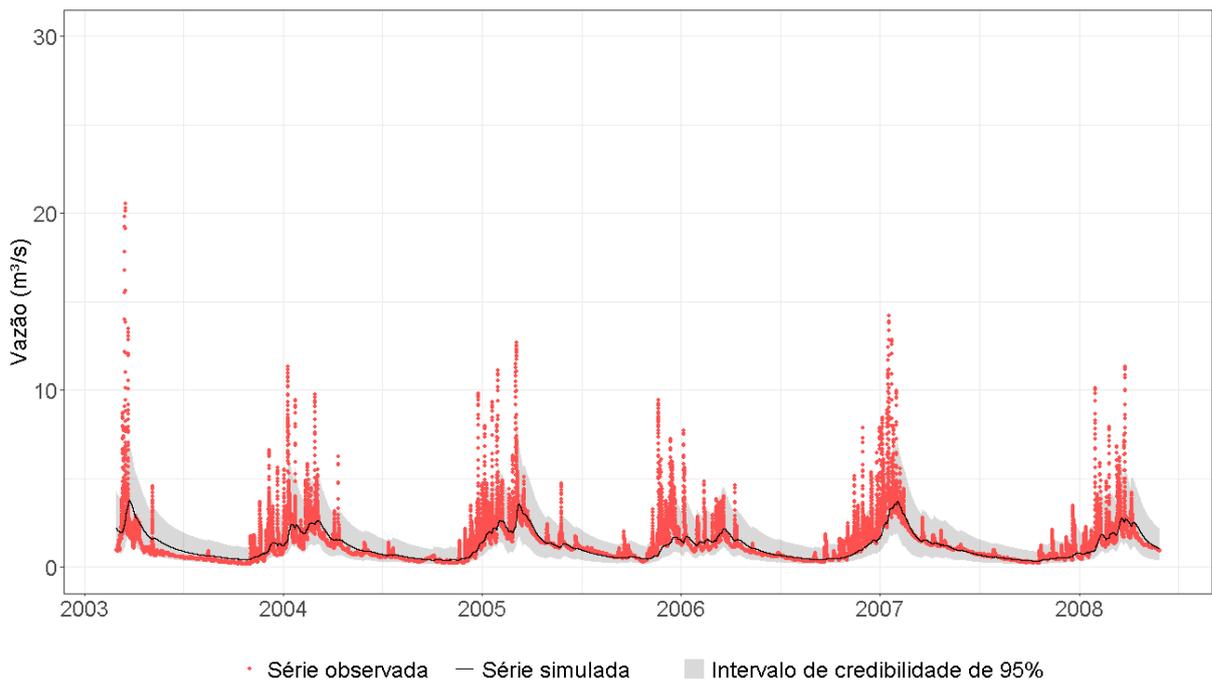


Figura 5.25 – Série temporal de vazões observadas e simuladas no domínio das assinaturas, para o período de validação - P_{FDC} .

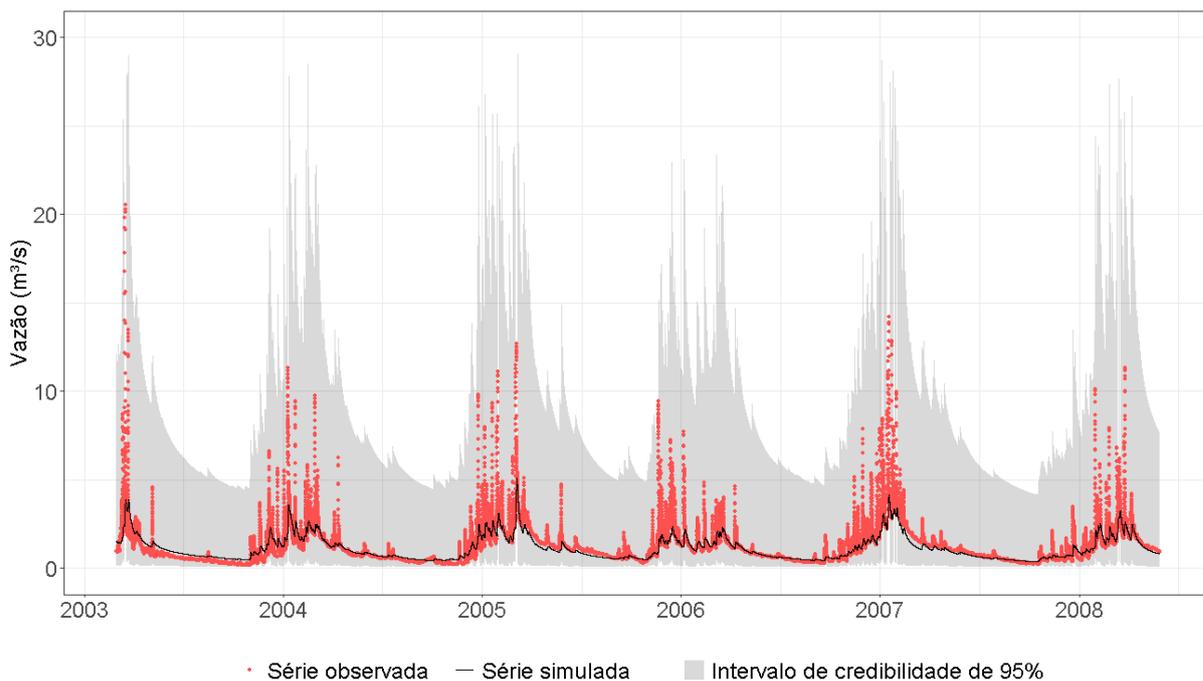


Figura 5.26 – Série temporal de vazões observadas e simuladas no domínio das assinaturas, para o período de validação - S_{FDC} .

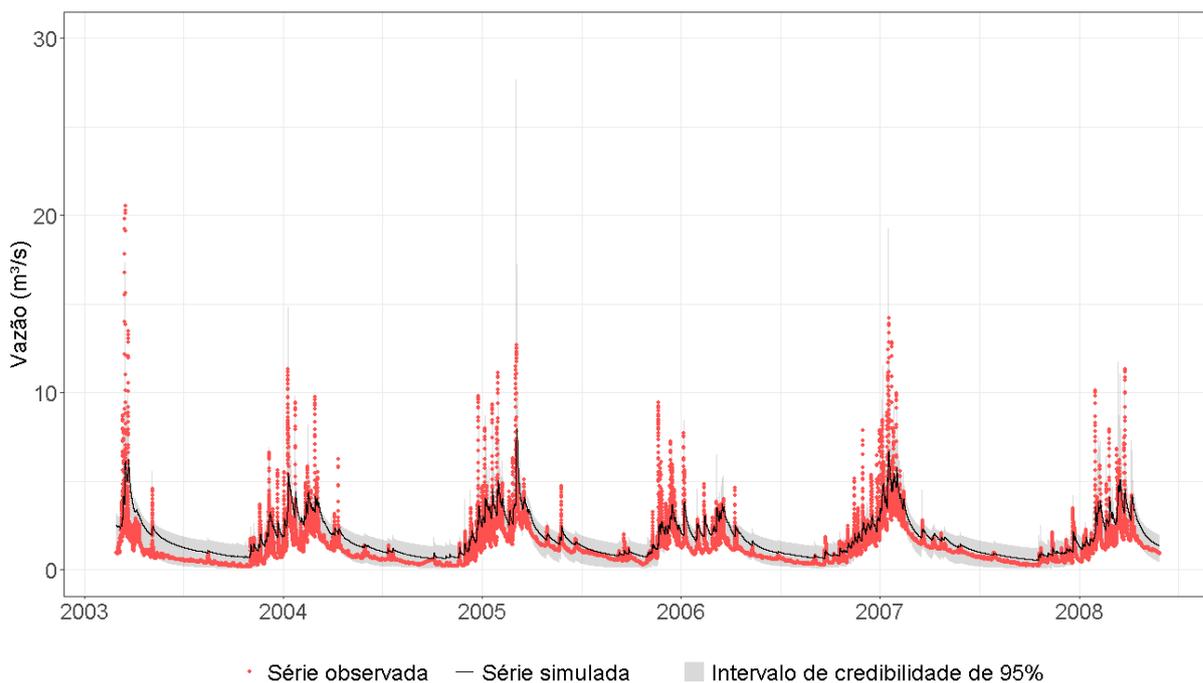


Figura 5.27 – Série temporal de vazões observadas e simuladas no domínio das assinaturas, para o período de validação - c_a .

Mais uma vez, foi observada grande incerteza quanto à estimação dos parâmetros para S_{FDC} , fazendo com que o intervalo de credibilidade contenha as vazões observadas em 100% do tempo. Além disso, houve uma redução injustificada desses intervalos para q_{mean} e para c_a .

6 CONCLUSÕES E RECOMENDAÇÕES

Neste trabalho, um modelo hidrológico conceitual horário foi adotado para avaliação do emprego de assinaturas hidrológicas para estimação de seus parâmetros, por meio de um algoritmo de aproximação computacional Bayesiana. O desempenho da modelagem e a distribuição *a posteriori* dos parâmetros são comparados aos obtidos via técnicas de inferência Bayesiana “clássicas”, com especificação da função de verossimilhança e métodos de amostragem a partir de simulações de Monte Carlo. A mesma avaliação foi realizada considerando um modelo mensal. Adicionalmente, os parâmetros do modelo horário foram estimados a partir de assinaturas calculadas com dados de estações vizinhas à bacia em estudo, e os resultados foram comparados aos da primeira avaliação.

As assinaturas hidrológicas mostraram-se viáveis para a estimação dos parâmetros do modelo, sendo encontrados índices de desempenho semelhantes aos obtidos quando toda a série de monitoramento é utilizada para calibração. Contudo, este resultado depende fortemente da assinatura considerada, sendo fundamental a seleção criteriosa da mesma. Neste trabalho, aspectos relacionados à escala de tempo do modelo (horário ou mensal) e a particularidades da medição das variáveis hidrológicas em campo mostraram-se muito importantes para a seleção das assinaturas, impactando diretamente os resultados das simulações. Nesse sentido, a seleção de uma escala de tempo adequada para modelagem mostrou-se fundamental também para a correta representação dos resíduos da modelagem.

Em geral, considerando as assinaturas locais, os percentis da curva de permanência e o coeficiente de escoamento anual apresentaram os melhores valores para as estatísticas avaliadas. Contudo, há grande variação quanto à incerteza na vazão imposta pela estimação dos parâmetros. Para S_{FDC} e F_L , especialmente, os intervalos de incerteza são bastante amplos e demonstram a incapacidade de estas assinaturas, isoladamente, sintetizarem as informações da série monitorada. Por outro lado, no domínio do tempo, foram estimados parâmetros dentro de uma pequena faixa de variação, o que pode estar relacionado à fixação de valores de certos parâmetros, como do coeficiente de assimetria ξ da SEP. Apesar de potencialmente ampliar o intervalo de incerteza, a consideração da variação deste parâmetro aumenta expressivamente o custo computacional, impedindo a realização do número de simulações propostas neste trabalho.

A combinação de assinaturas demonstrou potencial para melhorar o desempenho da modelagem em escala horária, permitindo que aspectos dos hidrogramas simulados para diferentes assinaturas sejam agregados, alcançando uma melhor representação do hidrograma observado e de sua variabilidade.

Outro resultado está relacionado aos bons índices de desempenho encontrados mesmo para um pequeno número de percentis da curva de permanência, não sendo necessária a consideração completa da curva.

Quanto às distribuições *a posteriori* dos parâmetros dos modelos, foi observado o problema da equifinalidade, em que diferentes conjuntos de parâmetros produzem resultados similares. Essa situação foi verificada tanto em escala horária quanto mensal. Além disso, foi constatado que as correlações entre os parâmetros mudam expressivamente de acordo com a assinatura considerada para estimação dos mesmos, sugerindo um problema de identificabilidade da estrutura dos modelos empregados.

A adoção das assinaturas ditas regionais também se mostrou possível. Com exceção dos percentis da curva de permanência, que apresentaram uma alteração significativa dos resultados, foram encontradas métricas similares às das simulações que consideravam as assinaturas locais, com oscilações de maior magnitude para os coeficientes α e γ do KGE. Destaca-se que, em geral, houve uma redução injustificada dos intervalos de incerteza paramétrica na predição da vazão.

Com relação aos métodos ABC, considera-se que essa abordagem é adequada aos propósitos deste trabalho. Contudo, é importante destacar o elevado custo computacional relacionado a essa abordagem, uma vez que são necessárias centenas de milhares de iterações para a convergência. Desse modo, apesar de possível, a estimação de parâmetros de modelos a partir de assinaturas regionalizadas é limitada pelo custo computacional para implementação do ABC, dificultando a modelagem de múltiplas bacias, em estudos de grande escala.

Adicionalmente, ressaltam-se os significativos avanços encontrados na literatura referente às assinaturas hidrológicas e suas aplicações. Um número substancialmente menor de trabalhos é observado quando essas assinaturas são empregadas em conjunto a métodos ABC em estudos de modelagem hidrológica. Isso limita a comparação dos resultados obtidos.

Apesar disso, as conclusões aqui apresentadas, em geral, corroboram com a literatura recente, mas alguns aspectos metodológicos podem ser aprimorados e pontos não explorados neste estudo podem ser abordados. É interessante uma avaliação mais aprofundada da regionalização das assinaturas, empregando métodos mais robustos para identificação da região homogênea e seleção das estações. Diferentes bacias, escalas de tempo, modelos e assinaturas devem ser também avaliados, buscando tecer conclusões e recomendações mais assertivas a respeito da metodologia empregada e dos resultados obtidos.

Além disso, por meio da consideração de um modelo de erros, cujos parâmetros podem ser estimados simultaneamente aos do modelo hidrológico, é possível quantificar a incerteza total da modelagem (FENICIA *et al.*, 2018; KAVETSKI *et al.*, 2018; WILKINSON, 2013). No presente trabalho, a consideração deste modelo, além de tornar mais difícil a convergência das cadeias, não indicou alterações significativas dos resultados. Nesse contexto, recomenda-se que novas análises sejam feitas com o propósito de melhor avaliar o impacto, as vantagens e desvantagens da consideração de um resíduo simulado. É possível também realizar o pós-processamento das saídas do modelo determinístico, com o objetivo de quantificar a incerteza preditiva condicional (ROMERO-CUELLAR *et al.*, 2019).

Outra possibilidade é a consideração da incerteza com relação às vazões observadas, por meio da estimativa de um intervalo plausível para a curva-chave, resultando em uma faixa de valores para as assinaturas. Westerberg *et al.* (2020) adotaram este procedimento para calibração de um modelo no domínio do tempo e, dado o aumento da capacidade de predição, recomendam essa abordagem também para a calibração a partir de assinaturas.

Por fim, recomenda-se a consideração de novas assinaturas e de diferentes combinações. Ferramentas como o TOSSH (*Toolbox for Streamflow Signatures in Hydrology* – GNANN *et al.*, 2021), um *toolbox* para cálculo de diversas assinaturas usando o Matlab, podem ser incorporadas, dispensando a necessidade de desenvolvimento de funções específicas para cada assinatura.

REFERÊNCIAS

- A.L. HØJBERG; J.C. REFSGAARD. Model Uncertainty – parameter uncertainty versus conceptual models. *Water Science and Technology*, v. 52, n. 6, p. 177–186, 2005.
- AANDAHL, R. Zachariah *et al.* A model-based Bayesian estimation of the rate of evolution of VNTR loci in *Mycobacterium tuberculosis*. *PLoS Computational Biology*, v. 8, n. 6, 2012.
- ADDOR, N. *et al.* A Ranking of Hydrological Signatures Based on Their Predictability in Space. *Water Resources Research*, v. 54, n. 11, p. 8792–8812, 2018.
- ALBERT, Carlo. A Simulated Annealing Approach to Bayesian Inference. n. September, 2015. Disponível em: <<http://arxiv.org/abs/1509.05315>>.
- ALBERT, Carlo; KÜNSCH, Hans R.; SCHEIDEGGER, Andreas. A simulated annealing approach to approximate Bayes computations. *Statistics and Computing*, v. 25, n. 6, p. 1217–1232, 2014.
- ALBERT, Carlo; ULZEGA, Simone. Bayesian Parameter Inference for 1D Nonlinear Stochastic Differential Equation Models. n. October, p. 1–16, 2015. Disponível em: <<http://arxiv.org/abs/1509.05305>>.
- BAKER, David B. *et al.* A new flashiness index: Characteristics and applications to Midwestern rivers and streams. *Journal of the American Water Resources Association*, v. 40, n. 2, p. 503–522, 2004.
- BAUDET, C. *et al.* Cophylogeny reconstruction via an approximate bayesian computation. *Systematic Biology*, v. 64, n. 3, p. 416–431, 2015.
- BEAUMONT, Mark A. Approximate Bayesian computation. *Annual Review of Statistics and Its Application*, v. 6, p. 379–403, 2019.
- BEAUMONT, Mark A. Approximate Bayesian computation in evolution and ecology. *Annual Review of Ecology, Evolution, and Systematics*, v. 41, p. 379–406, 2010.
- BEAUMONT, Mark A.; ZHANG, Wenyang; BALDING, David J. Approximate Bayesian computation in population genetics. *Genetics*, v. 162, n. 4, p. 2025–2035, 2002.
- BECK, Hylke E.; DE ROO, Ad; VAN DIJK, Albert I.J.M. Global maps of streamflow characteristics based on observations from several thousand catchments. *Journal of Hydrometeorology*, v. 16, n. 4, p. 1478–1501, 2015.
- BENNETT, F. R. Gradient boosting machine assisted approximate Bayesian computation for uncertainty analysis of rainfall-runoff model parameters. *23rd International Congress on Modelling and Simulation - Supporting Evidence-Based Decision Making: The Role of Modelling and Simulation, MODSIM 2019*, n. December, p. 1063–1069, 2019.
- BENNETT, Kelly Louise *et al.* Historical environmental change in Africa drives divergence and admixture of *Aedes aegypti* mosquitoes: a precursor to successful worldwide colonization? *Molecular ecology*, v. 25, n. 17, p. 4337–4354, 2016.
- BEVEN, Keith. Prophecy, reality and uncertainty in distributed hydrological modelling. *Advances in Water Resources*, v. 16, n. 1, p. 41–51, 1993.
- BEVEN, Keith. *Rainfall-Runoff Modelling*. [S.l: s.n.], 2012.
- BEVEN, Keith; BINLEY, Andrew. The future of distributed models: Model calibration and uncertainty prediction. *Hydrological Processes*, v. 6, n. 3, p. 279–298, 1992.
- BIVAND, Roger S.; GÓMEZ-RUBIO, Virgilio; RUE, Håvard. Approximate Bayesian inference for spatial econometrics models. *Spatial Statistics*, v. 9, n. C, p. 146–165, 2014. Disponível em: <<http://dx.doi.org/10.1016/j.spasta.2014.01.002>>.
- BOSCARELLO, Laura *et al.* Regionalization of Flow-Duration Curves through Catchment Classification with Streamflow Signatures and Physiographic–Climate Indices. *Journal of Hydrologic*

Engineering, v. 21, n. 3, p. 05015027, 2016.

BULYGINA, Nataliya *et al.* Integrating different types of information into hydrological model parameter estimation: Application to ungauged catchments and land use scenario analysis. *Water Resources Research*. [S.l: s.n.], 2012

BULYGINA, Nataliya; GUPTA, Hoshin. Estimating the uncertain mathematical structure of a water balance model via Bayesian data assimilation. *Water Resources Research*, v. 45, n. 2, p. 1–20, 2009.

BUNN, Stuart E.; ARTHINGTON, Angela H. Basic principles and ecological consequences of altered flow regimes for aquatic biodiversity. *Environmental Management*, v. 30, n. 4, p. 492–507, 2002.

CHOWDHURY, A *et al.* Development of an Automatic Calibration Framework for Hydrologic Modelling Using Approximate Bayesian Computation. v. 10, n. 2, p. 119–126, 2016.

CHRISTOPHER, Jason D. *et al.* Flow parameter estimation using laser absorption spectroscopy and approximate Bayesian computation. *Experiments in Fluids*, v. 62, n. 2, p. 1–20, 2021. Disponível em: <<https://doi.org/10.1007/s00348-020-03122-2>>.

CHUECA, Luis J.; SCHELL, Tilman; PFENNINGER, Markus. Whole-genome re-sequencing data to infer historical demography and speciation processes in land snails: the study of two *Candidula* sister species. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, v. 376, n. 1825, p. 20200156, 2021.

COELHO, Caio A. S.; CARDOSO, Denis H.F.; FIRPO, Mári A. F. A seca de 2013 a 2015 na região sudeste do Brasil. *Climanálise*, n. 1, p. 55–61, 2016. Disponível em: <<http://climanalise.cptec.inpe.br/~rclimanl/revista/pdf/30anos/Coelhoetal.pdf>>.

CPRM. Mapa de declividade em percentual do relevo brasileiro.2010. Disponível em: <<http://www.cprm.gov.br/publique/Gestao-Territorial/Gestao-Territorial/Mapa-de-Declividade-em-Percentual-do-Relevo-Brasileiro-3497.html>> Acesso em: 03 de abr. de 2021.

CUI, Tao *et al.* Emulator-enabled approximate Bayesian computation (ABC) and uncertainty analysis for computationally expensive groundwater models. *Journal of Hydrology*, v. 564, n. December 2017, p. 191–207, 2018. Disponível em: <<https://doi.org/10.1016/j.jhydrol.2018.07.005>>.

DAVISON, Bruce *et al.* Parameter-state ensemble data assimilation using Approximate Bayesian Computing for short-term hydrological prediction. *Hydrology and Earth System Sciences Discussions*, n. September, p. 1–38, 2017.

DE PAIVA, Rodrigo Cauduro Dias *et al.* Advances and challenges in the water sciences in brazil: A community synthesis of the xxiii brazilian water resources symposium. *Revista Brasileira de Recursos Hídricos*, v. 25, p. 1–28, 2020.

DEL MORAL, Pierre; DOUCET, Arnaud; JASRA, Ajay. An adaptive sequential Monte Carlo method for approximate Bayesian computation. *Statistics and Computing*, v. 22, n. 5, p. 1009–1020, 2012.

DENG, Chao *et al.* Temporal variation and scaling of parameters for a monthly hydrologic model. *Journal of Hydrology*, v. 558, p. 290–300, 2018. Disponível em: <<https://doi.org/10.1016/j.jhydrol.2018.01.049>>.

DIAO, Wei *et al.* Multi-objective optimal operation of reservoir group in Jialing River based on DREAM algorithm. *Water Supply*, p. 1–14, 2021.

DORONINA, Olga A. *et al.* Autonomic closure for turbulent flows using approximate bayesian computation. *AIAA Aerospace Sciences Meeting, 2018*, n. 210059, 2018.

DROVANDI, Christopher C.; PETTITT, Anthony N. Likelihood-free Bayesian estimation of multivariate quantile distributions. *Computational Statistics and Data Analysis*, v. 55, n. 9, p. 2541–2556, 2011. Disponível em: <<http://dx.doi.org/10.1016/j.csda.2011.03.019>>.

DUTTA, Ritabrata *et al.* Interpretable pathological test for Cardio-vascular disease: Approximate

- Bayesian computation with distance learning. n. October, 2020. Disponível em: <<http://arxiv.org/abs/2010.06465>>.
- DUTTA, Ritabrata *et al.* Parameter estimation of platelets deposition: Approximate bayesian computation with high performance computing. *Frontiers in Physiology*, v. 9, n. AUG, 2018.
- DUTTA, Ritabrata; MIRA, Antonietta; ONNELA, Jukka Pekka. Bayesian inference of spreading processes on networks. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, v. 474, n. 2215, 2018.
- ENG, Ken *et al.* Predictability and selection of hydrologic metrics in riverine ecohydrology. *Freshwater Science*, v. 36, n. 4, p. 915–926, 2017.
- FAN, Y.; SISSON, S. A. Abc samplers. *arXiv*, n. February 2018, 2018.
- FAN, Ying *et al.* Hydrologic regulation of plant rooting depth. *Proceedings of the National Academy of Sciences of the United States of America*, v. 114, n. 40, p. 10572–10577, 2017.
- FENICIA, Fabrizio *et al.* Signature-Domain Calibration of Hydrological Models Using Approximate Bayesian Computation: Empirical Analysis of Fundamental Properties. *Water Resources Research*, v. 54, n. 6, p. 3958–3987, 2018.
- FENICIA, Fabrizio; KAVETSKI, Dmitri; SAVENIJE, Hubert H.G. Elements of a flexible approach for conceptual hydrological modeling: 1. Motivation and theoretical development. *Water Resources Research*, v. 47, n. 11, p. 1–13, 2011.
- FERNANDEZ VISENTINI, Alejandro *et al.* Inferring geostatistical properties of hydraulic conductivity fields from saline tracer tests and equivalent electrical conductivity time-series. *Advances in Water Resources*, v. 146, n. September, p. 103758, 2020. Disponível em: <<https://doi.org/10.1016/j.advwatres.2020.103758>>.
- GNANN, Sebastian J. *et al.* TOSSH: A Toolbox for Streamflow Signatures in Hydrology. *Environmental Modelling and Software*, v. 138, n. February, 2021.
- GUPTA, HOSHIN V.; WANEGER, T.; LIU, Y. Reconciling theory with observations: elements of a diagnostic approach to model evaluation. *HYDROLOGICAL PROCESSES*, v. 22, n. March 2008, p. 3802–3813, 2008. Disponível em: <<http://jamsb.austms.org.au/courses/CSC2408/semester3/resources/ldp/abs-guide.pdf>>.
- GUPTA, Hoshin V. *et al.* Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling. *Journal of Hydrology*, v. 377, n. 1–2, p. 80–91, 2009. Disponível em: <<http://dx.doi.org/10.1016/j.jhydrol.2009.08.003>>.
- HEUDORFER, B. *et al.* Index-Based Characterization and Quantification of Groundwater Dynamics. *Water Resources Research*, v. 55, n. 7, p. 5575–5592, 2019.
- HOPP, Luisa *et al.* The relevance of preferential flow in catchment scale simulations: Calibrating a 3D dual-permeability model using DREAM. *Hydrological Processes*, v. 34, n. 5, p. 1237–1254, 2020.
- HORNER, Ivan *et al.* Information content of snow hydrological signatures based on streamflow, precipitation and air temperature. *Hydrological Processes*, v. 34, n. 12, p. 2763–2779, 2020.
- HYNEK, Brian M.; BEACH, Michael; HOKE, Monica R. T. Updated global map of Martian valley networks and implications for climate and hydrologic processes. *Journal of Geophysical Research*, v. 115, n. E9, p. 1–14, 2010.
- JAISWAL, R. K.; ALI, Sohrat; BHARTI, Birendra. Comparative evaluation of conceptual and physical rainfall–runoff models. *Applied Water Science*, v. 10, n. 1, p. 1–14, 2020. Disponível em: <<https://doi.org/10.1007/s13201-019-1122-6>>.
- JAYAWARDENA, A. W. *Environmental and hydrological systems modelling*. [S.l: s.n.], 2013.
- JEHN, Florian U. *et al.* Using hydrological and climatic catchment clusters to explore drivers of

- catchment behavior. *Hydrology and Earth System Sciences*, v. 24, n. 3, p. 1081–1100, 2020.
- JEPSEN, S.; HARMON, T.; SHI, Y. Watershed model calibration to the base flow recession curve with and without evapotranspiration effects. *Water Resources Research*, 2016.
- JOWETT, Ian G.; DUNCAN, Maurice J. Flow variability in new zealand rivers and its relationship to in-stream habitat and biota. *New Zealand Journal of Marine and Freshwater Research*, v. 24, n. 3, p. 305–317, 1990.
- KAVETSKI, Dmitri *et al.* Signature-Domain Calibration of Hydrological Models Using Approximate Bayesian Computation: Theory and Comparison to Existing Applications. *Water Resources Research*, v. 54, n. 6, p. 4059–4083, 2018.
- LIMA, Fernando Neves; FERNANDES, Wilson; NASCIMENTO, Nilo. Joint calibration of a hydrological model and rating curve parameters for simulation of flash flood in urban areas. *RBRH*, v. 24, p. 539–547, 2019. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S2318-03312019000100222&tlng=en>.
- LIU, Y. R. *et al.* A Bayesian-based multilevel factorial analysis method for analyzing parameter uncertainty of hydrological model. *Journal of Hydrology*, v. 553, p. 750–762, 2017.
- LIU, Yuqiong *et al.* A wavelet-based approach to assessing timing errors in hydrologic predictions. *Journal of Hydrology*, v. 397, n. 3–4, p. 210–224, 2011. Disponível em: <<http://dx.doi.org/10.1016/j.jhydrol.2010.11.040>>.
- LU, Dan; RICCIUTO, Daniel; EVANS, Katherine. An efficient Bayesian data-worth analysis using a multilevel Monte Carlo method. *Advances in Water Resources*, v. 113, p. 223–235, 2018. Disponível em: <<https://doi.org/10.1016/j.advwatres.2018.01.024>>.
- MA, Qiumei *et al.* A censored shifted mixture distribution mapping method to correct the bias of daily IMERG satellite precipitation estimates. *Remote Sensing*, v. 11, n. 11, p. 1–25, 2019.
- MARIN, Jean Michel *et al.* Approximate Bayesian computational methods. *Statistics and Computing*, v. 22, n. 6, p. 1167–1180, 2012.
- MARJORAM, Paul *et al.* Markov chain Monte Carlo without likelihoods. *Proceedings of the National Academy of Sciences of the United States of America*, v. 100, n. 26, p. 15324–15328, 2003.
- MARTINEZ, Guillermo F.; GUPTA, Hoshin V. *Hydrologic consistency as a basis for assessing complexity of monthly water balance models for the continental United States*. *Water Resources Research*. [S.l.: s.n.], 2011
- MASSMANN, Carolina. Identification of factors influencing hydrologic model performance using a top-down approach in a large number of U.S. catchments. *Hydrological Processes*, v. 34, n. 1, p. 4–20, 2020.
- MCMILLAN, Hilary. Linking hydrologic signatures to hydrologic processes: A review. *Hydrological Processes*, v. 34, n. 6, p. 1393–1409, 2020.
- MCMILLAN, Hilary K. A review of hydrologic signatures and their applications. *Wiley Interdisciplinary Reviews: Water*, v. 8, n. 1, p. 1–23, 2021.
- MCMILLAN, Hilary K.; WESTERBERG, Ida K.; KRUEGER, Tobias. Hydrological data uncertainty and its implications. *WIREs Water*, v. 5, n. 6, p. 1–14, 2018.
- MCMILLAN, Hilary; WESTERBERG, Ida; BRANGER, Flora. Five guidelines for selecting hydrological signatures. *Hydrological Processes*, v. 31, n. 26, p. 4757–4761, 2017.
- MELO, Marília *et al.* *Avaliação Da Segurança Hídrica Para Abastecimento Público Na Região Metropolitana De Belo Horizonte: Estudo Da Crise Hídrica 2014-2015*. *Revista Brasileira de Climatologia*. [S.l.: s.n.], 2020
- MENDOZA, Pablo A. *et al.* How do hydrologic modeling decisions affect the portrayal of climate

- change impacts? *Hydrological Processes*, v. 30, n. 7, p. 1071–1095, 2016.
- MOGES, Edom *et al.* Review: Sources of hydrological model uncertainties and advances in their analysis. *Water (Switzerland)*, v. 13, n. 1, 2021.
- MOHAMMED, Safa A. *et al.* Impact of dataset size on the signature-based calibration of a hydrological model. *Water (Switzerland)*. [S.l: s.n.], 2021
- MONTANARI, A. Uncertainty of Hydrological Predictions. *Treatise on Water Science*, v. 2, p. 459–478, 2011.
- MOUELHI, Safouane *et al.* Stepwise development of a two-parameter monthly water balance model. *Journal of Hydrology*, v. 318, n. 1–4, p. 200–214, 2006.
- NASH, J. E.; SUTCLIFFE, J. V. River flow forecasting through conceptual models. Part I - A discussion of principles. *Journal of Hydrology*, v. 10, n. 3, p. 282–290, 1970.
- NEVES, Betânia Vilas Boas; RODRIGUES, Paulo César. Geoprocessamento como ferramenta no estudo de correlação entre a dinâmica da cobertura vegetal e evapotranspiração na bacia do Ribeirão Serra Azul - MG. *Revista Brasileira de Recursos Hídricos*, v. 12, n. 4, p. 87–102, 2007.
- NIJZINK, R. C. *et al.* Constraining Conceptual Hydrological Models With Multiple Information Sources. *Water Resources Research*, v. 54, n. 10, p. 8332–8362, 2018.
- NOBRE, Carlos A. *et al.* Some Characteristics and Impacts of the Drought and Water Crisis in Southeastern Brazil during 2014 and 2015. *Journal of Water Resource and Protection*, v. 08, n. 02, p. 252–262, 2016.
- NORTH, Gerald R.; ZHANG, Fuqing; PYLE, John. *Encyclopedia of Atmospheric Sciences: Second Edition*. [S.l: s.n.], 2014.
- NOTT, David J.; MARSHALL, Lucy; BROWN, Jason. Generalized likelihood uncertainty estimation (GLUE) and approximate Bayesian computation: What’s the connection? *Water Resources Research*, v. 48, n. 12, p. 1–7, 2012.
- NOURALI, Mahrouz. Comparison of Likelihood-Free Inference Approach and a Formal Bayesian Method in Parameter Uncertainty Assessment: Case Study with a Single-Event Rainfall–Runoff Model. *Journal of Hydrologic Engineering*, v. 26, n. 3, p. 05020049, 2021.
- OKKAN, Umut; FISTIKOGLU, Okan. Evaluating climate change effects on runoff by statistical downscaling and hydrological model GR2M. *Theoretical and Applied Climatology*, v. 117, n. 1, p. 343–361, 2014.
- PANIGRAHI, N.; DAS, B. S. Canopy Spectral Reflectance as a Predictor of Soil Water Potential in Rice. *Water Resources Research*, v. 54, n. 4, p. 2544–2560, 2018.
- PARK, Edward; LATRUBESSE, Edgardo M. Surface water types and sediment distribution patterns at the confluence of mega rivers: The Solimões-Amazon and Negro Rivers junction. *Water Resources Research*, v. 51, n. 8, p. 6197–6213, 2015.
- PERRIN, Charles; MICHEL, Claude; ANDRÉASSIAN, Vazken. Improvement of a parsimonious model for streamflow simulation. *Journal of Hydrology*, v. 279, n. 1–4, p. 275–289, 2003.
- PETERS, Gareth W. *et al.* Bayesian symbol detection in wireless relay networks via likelihood-free inference. *IEEE Transactions on Signal Processing*, v. 58, n. 10, p. 5206–5218, 2010.
- PINHEIRO, Viviane Borda; NAGHETTINI, Mauro; PALMIER, Luiz Rafael. Uncertainty estimation in hydrodynamic modeling using Bayesian techniques. *Revista Brasileira de Recursos Hídricos*, v. 24, 2019.
- PINHEIRO, VIVIANE; NAGHETTINI, MAURO. Calibração de um Modelo Chuva-Vazão em Bacias sem Monitoramento Fluviométrico a partir de Curvas de Permanência Sintéticas. *Revista Brasileira de Recursos Hídricos*, v. 15, n. 2, p. 143–156, 2010.

- PRASKIEVICZ, Sarah; LUO, Cehong. Unsupervised hydrologic classification of rivers: Watershed controls on natural and anthropogenic flow regimes, Alabama, USA. *Hydrological Processes*, v. 33, n. 8, p. 1231–1244, 2019.
- PRIETO, Cristina *et al.* Flow Prediction in Ungauged Catchments Using Probabilistic Random Forests Regionalization and New Statistical Adequacy Tests. *Water Resources Research*, v. 55, n. 5, p. 4364–4392, 2019.
- PRITCHARD, Jonathan K. *et al.* Population growth of human Y chromosomes: A study of Y chromosome microsatellites. *Molecular Biology and Evolution*, v. 16, n. 12, p. 1791–1798, 1999.
- RAGAB, Ragab *et al.* Application of Generalized Likelihood Uncertainty Estimation (GLUE) at different temporal scales to reduce the uncertainty level in modelled river flows. *Hydrological Sciences Journal*, v. 65, n. 11, p. 1856–1871, 2020. Disponível em: <<https://doi.org/10.1080/02626667.2020.1764961>>.
- REICHERT, Peter; AMMANN, Lorenz; FENICIA, Fabrizio. Potential and Challenges of Investigating Intrinsic Uncertainty of Hydrological Models With Stochastic, Time-Dependent Parameters. *Water Resources Research*, v. 57, n. 3, p. 1–28, 2021.
- RICHTER, Brian D. *et al.* A Method for Assessing Hydrologic Alteration within Ecosystems. *Conservation Biology*, v. 10, n. 4, p. 1163–1174, 1996.
- ROCHA, Paulo Cesar. Indicators of Hydrologic Alteration in the High Parana River Catchment: Human Interventions and Implications for Dynamic of the Fluvial Environment. *Sociedade & Natureza*, v. 22, n. 1, p. 191–211, 2010. Disponível em: <<http://www.seer.ufu.br/index.php/sociedadnatureza/article/view/9820>>.
- ROJAS, Rodrigo; FEYEN, Luc; DASSARGUES, Alain. Conceptual model uncertainty in groundwater modeling: Combining generalized likelihood uncertainty estimation and Bayesian model averaging. *Water Resources Research*, v. 44, n. 12, p. 1–16, 2008.
- ROMERO-CUELLAR, Jonathan *et al.* Hydrological post-processing based on approximate Bayesian computation (ABC). *Stochastic Environmental Research and Risk Assessment*, v. 33, n. 7, p. 1361–1373, 2019.
- RUBIN, Donald B. Bayesianly Justifiable and Relevant Frequency Calculations for the Applied Statistician. *The Annals of Statistics*, v. 12, n. 4, p. 1151–1172, 1984.
- SADEGH, MOJTABA; VRUGT, Jasper A. Approximate Bayesian Computation using Markov Chain Monte Carlo simulation: DREAM(ABC). p. 5375–5377, 2014.
- SADEGH, M.; VRUGT, J. A. Bridging the gap between GLUE and formal statistical approaches: Approximate Bayesian computation. *Hydrology and Earth System Sciences*, v. 17, n. 12, p. 4831–4850, 2013.
- SADEGH, Mojtaba *et al.* The stationarity paradigm revisited: Hypothesis testing using diagnostics, summary metrics, and DREAM(ABC). *Water Resources Research*, 2015a.
- SADEGH, Mojtaba *et al.* The stationarity paradigm revisited: Hypothesis testing using diagnostics, summary metrics, and DREAM(ABC). *Water Resources Research*, v. 51, p. 9207–9231, 2015b.
- SAHRAEI, Shahram; ASADZADEH, Masoud; UNDUCHE, Fisaha. Signature-based multi-modelling and multi-objective calibration of hydrologic models: Application in flood forecasting for Canadian Prairies. *Journal of Hydrology*, v. 588, n. May, p. 125095, 2020. Disponível em: <<https://doi.org/10.1016/j.jhydrol.2020.125095>>.
- SAWICZ, K. *et al.* Catchment classification: empirical analysis of hydrologic similarity based on catchment function in the eastern USA. *Hydrology and Earth System Sciences Discussions*, v. 8, n. 3, p. 4495–4534, 2011.
- SCHOUPS, Gerrit; VRUGT, Jasper A. A formal likelihood function for parameter and predictive

- inference of hydrologic models with correlated, heteroscedastic, and non-Gaussian errors. *Water Resources Research*, v. 46, n. 10, p. 1–17, 2010.
- SHAFII, Mahyar; TOLSON, Bryan. Optimizing hydrological consistency by incorporating hydrological signatures into model calibration objectives. *Water Resources Research*, 2015.
- SHENG, Sheng *et al.* Transferability of a Conceptual Hydrological Model across Different Temporal Scales and Basin Sizes. *Water Resources Management*, v. 34, n. 9, p. 2953–2968, 2020.
- SHIN, Mun Ju *et al.* A review of foundational methods for checking the structural identifiability of models: Results for rainfall-runoff. *Journal of Hydrology*, v. 520, p. 1–16, 2015. Disponível em: <<http://dx.doi.org/10.1016/j.jhydrol.2014.11.040>>.
- SISSON, S. A.; FAN, Y.; BEAUMONT, Mark A. (Org.). *Handbook of Approximate Bayesian Computation*. [S.l.: s.n.], 2019.
- SPRENGER, Matthias *et al.* *The Demographics of Water: A Review of Water Ages in the Critical Zone*. [S.l.: s.n.], 2019. v. 57.
- STOICA, R.S. *et al.* Shadow Simulated Annealing: A new algorithm for approximate Bayesian inference of Gibbs point processes. *Spatial Statistics*, v. 43, p. 100505, 2021. Disponível em: <<https://doi.org/10.1016/j.spasta.2021.100505>>.
- TAVARÉ, Simon *et al.* Inferring coalescence times from DNA sequence data. *Genetics*, v. 145, n. 2, p. 505–518, 1997.
- TODOROVIĆ, Andrijana *et al.* The 3DNet-Catch hydrologic model: Development and evaluation. *Journal of Hydrology*, v. 568, n. October 2018, p. 26–45, 2019. Disponível em: <<https://doi.org/10.1016/j.jhydrol.2018.10.040>>.
- TROIN, Magali *et al.* Uncertainty of hydrological model components in climate change studies over two nordic quebec catchments. *Journal of Hydrometeorology*, v. 19, n. 1, p. 27–46, 2018.
- TURNER, Brandon M.; VAN ZANDT, Trisha. A tutorial on approximate Bayesian computation. *Journal of Mathematical Psychology*, v. 56, n. 2, p. 69–85, 2012. Disponível em: <<http://dx.doi.org/10.1016/j.jmp.2012.02.005>>.
- UFV - CETEC - UFLA - FEAM. Mapa de solos do Estado de Minas Gerais. Belo Horizonte, Fundação Estadual do Meio Ambiente, 2010. 49p. Disponível em: <<http://www.feam.br/noticias/1/949-mapas-de-solo-do-estado-de-minas-gerais>> Acesso em: 03 de abr. de 2021.
- VRUGT, Jasper A. *et al.* Accelerating Markov chain Monte Carlo simulation by differential evolution with self-adaptive randomized subspace sampling. *International Journal of Nonlinear Sciences and Numerical Simulation*, v. 10, n. 3, p. 273–290, 2008.
- VRUGT, Jasper A. *et al.* Equifinality of formal (DREAM) and informal (GLUE) Bayesian approaches in hydrologic modeling? *Stochastic Environmental Research and Risk Assessment*, v. 23, n. 7, p. 1011–1026, 2009.
- VRUGT, Jasper A.; SADEGH, Mojtaba. Toward diagnostic model calibration and evaluation: Approximate Bayesian computation. *Water Resources Research*, v. 49, n. 7, p. 4335–4345, 2013.
- WAGENER, Thorsten; WHEATER, Howard; GUPTA, Hoshin V. *Rainfall-Runoff Modelling in Gauged Ungauged Catchments*. [S.l.]: Imperial College Press, 2004.
- WESTERBERG, I. K. *et al.* Calibration of hydrological models using flow-duration curves. *Hydrology and Earth System Sciences*, v. 15, n. 7, p. 2205–2227, 2011.
- WESTERBERG, Ida K. *et al.* Hydrological model calibration with uncertain discharge data. *Hydrological Sciences Journal*, v. 0, n. 0, 2020. Disponível em: <<http://dx.doi.org/10.1080/02626667.2020.1735638>>.

- WESTERBERG, Ida K. *et al.* Uncertainty in hydrological signatures for gauged and ungauged catchments. *Water Resources Research*, v. 52, n. 3, p. 1847–1865, 2016.
- WEYLAND, Mathias S. *et al.* Holistic View on Cell Survival and DNA Damage: How Model-Based Data Analysis Supports Exploration of Dynamics in Biological Systems. *Computational and Mathematical Methods in Medicine*, v. 2020, 2020.
- WILKINSON, Richard D.; TAVARÉ, Simon. Estimating primate divergence times by using conditioned birth-and-death processes. *Theoretical Population Biology*, v. 75, n. 4, p. 278–285, 2009. Disponível em: <<http://dx.doi.org/10.1016/j.tpb.2009.02.003>>.
- WILKINSON, Richard D. Approximate Bayesian computation (ABC) gives exact results under the assumption of model error. *Statistical Applications in Genetics and Molecular Biology*, v. 12, n. 2, p. 129–141, 2013.
- WLOSTOWSKI, Adam N. *et al.* Signatures of Hydrologic Function Across the Critical Zone Observatory Network. *Water Resources Research*, v. 57, n. 3, 2021.
- WU, Xia; MARSHALL, Lucy; SHARMA, Ashish. Quantifying input error in hydrologic modeling using the Bayesian error analysis with reordering (BEAR) approach. *Journal of Hydrology*, v. 598, n. January, p. 126202, 2021. Disponível em: <<https://doi.org/10.1016/j.jhydrol.2021.126202>>.
- YADAV, Maitreya; WAGENER, Thorsten; GUPTA, Hoshin. Regionalization of constraints on expected watershed response behavior for improved predictions in ungauged basins. *Advances in Water Resources*, v. 30, n. 8, p. 1756–1774, 2007.
- YAN, Ling; JIN, Jiming; WU, Pute. Impact of parameter uncertainty and water stress parameterization on wheat growth simulations using CERES-Wheat with GLUE. *Agricultural Systems*, v. 181, n. February, p. 102823, 2020. Disponível em: <<https://doi.org/10.1016/j.agsy.2020.102823>>.
- ZHANG, Zhenxing; BALAY, John W.; LIU, Can. Regional regression models for estimating monthly streamflows. *Science of the Total Environment*, v. 706, n. xxxx, p. 135729, 2019. Disponível em: <<https://doi.org/10.1016/j.scitotenv.2019.135729>>.

APÊNDICE I – DOMÍNIO TEMPO VERSUS DOMÍNIO DAS ASSINATURAS: ESCALA HORÁRIA

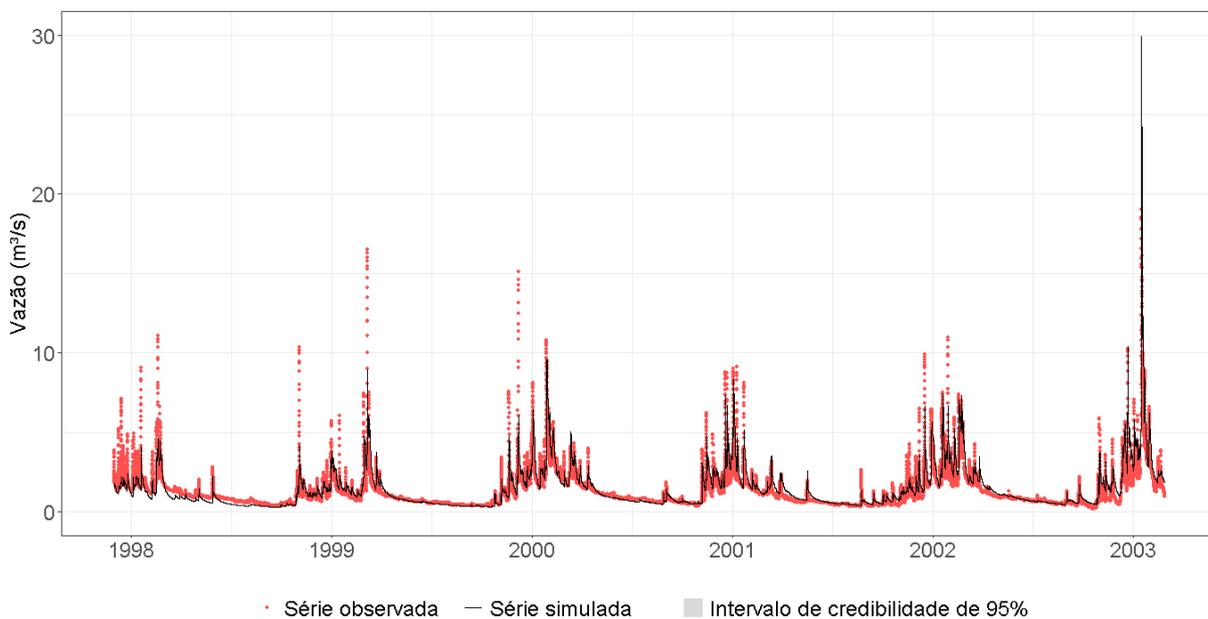


Figura I. 1 – Série temporal de vazões observadas e simuladas no domínio do tempo, para o período de calibração.

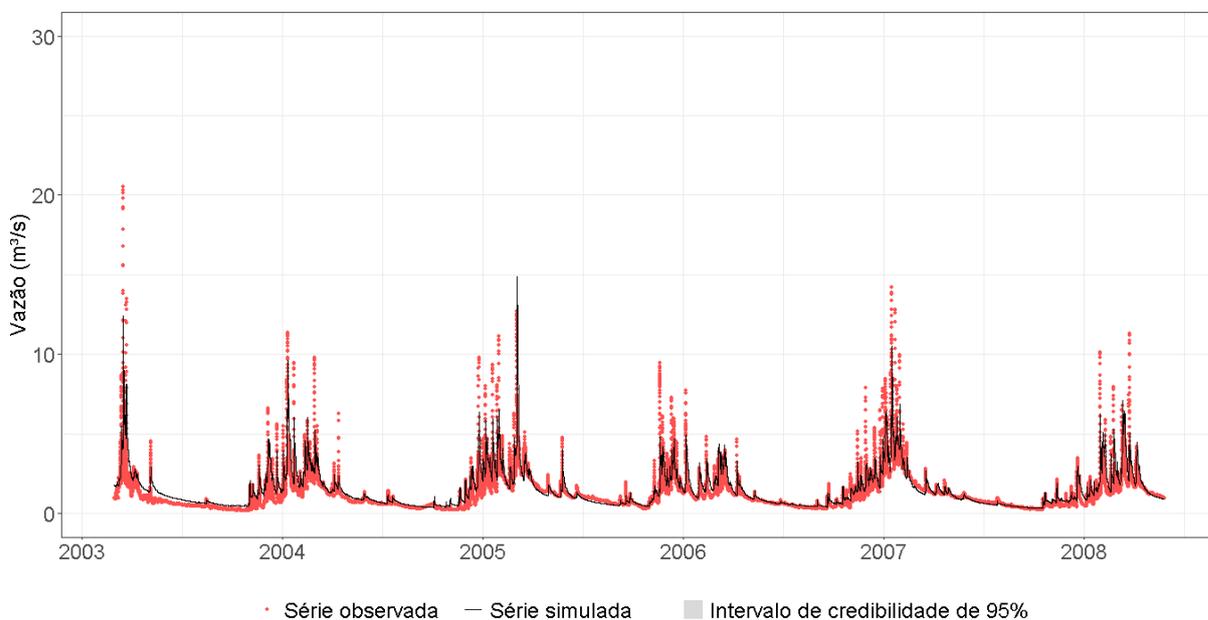


Figura I. 2 – Série temporal de vazões observadas e simuladas no domínio do tempo, para o período de validação.

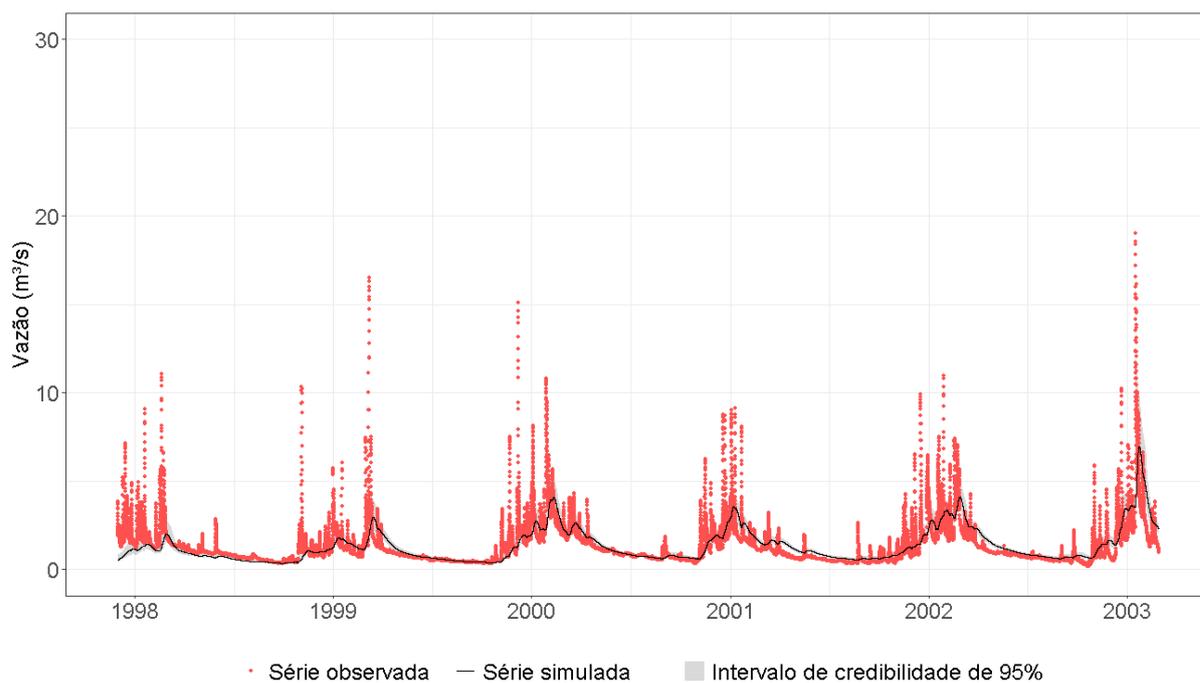


Figura I. 3 – Série temporal de vazões observadas e simuladas no domínio das assinaturas, para o período de calibração – q_{mean} .

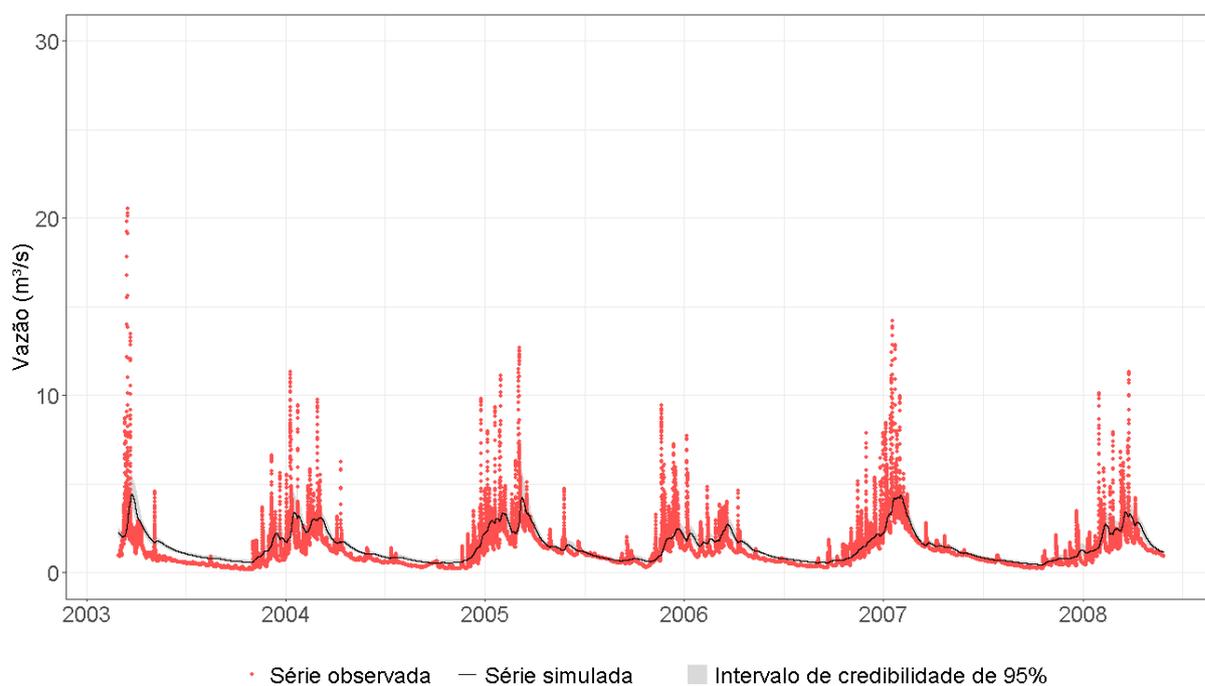


Figura I. 4 – Série temporal de vazões observadas e simuladas no domínio das assinaturas, para o período de validação – q_{mean} .

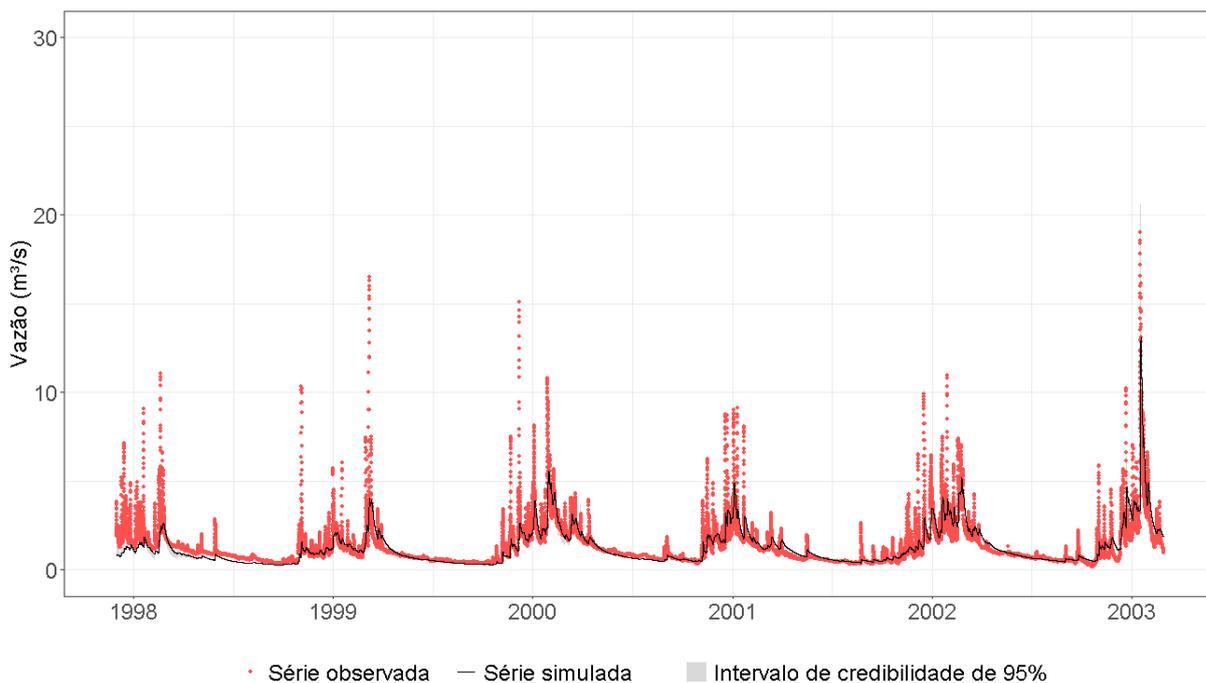


Figura I. 5 – Série temporal de vazões observadas e simuladas no domínio das assinaturas, para o período de calibração - P_{FDC} .

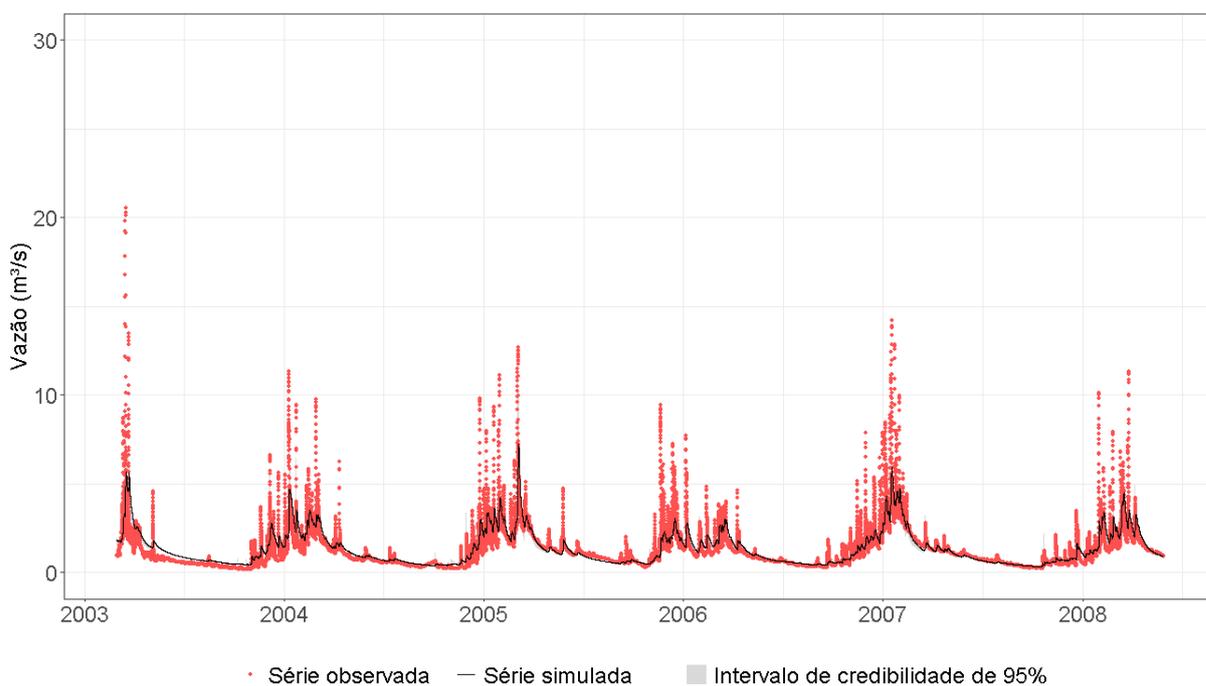


Figura I. 6 – Série temporal de vazões observadas e simuladas no domínio das assinaturas, para o período de validação - P_{FDC} .

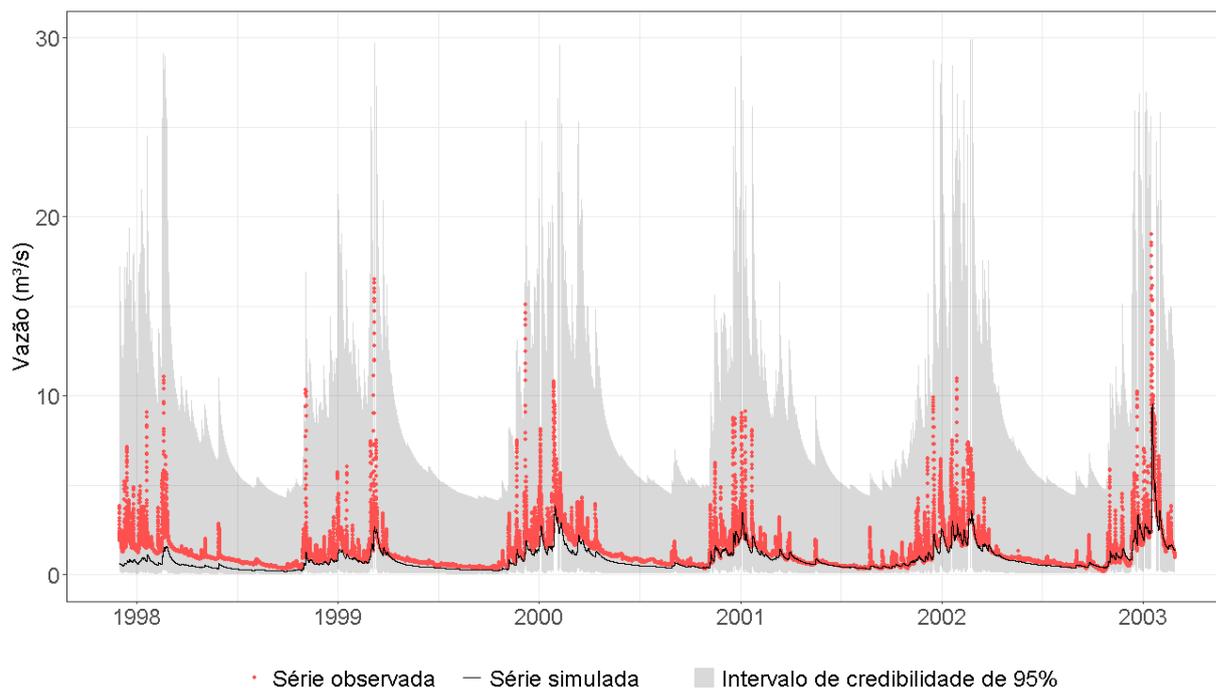


Figura I. 7 – Série temporal de vazões observadas e simuladas no domínio das assinaturas, para o período de calibração - S_{FDC} .

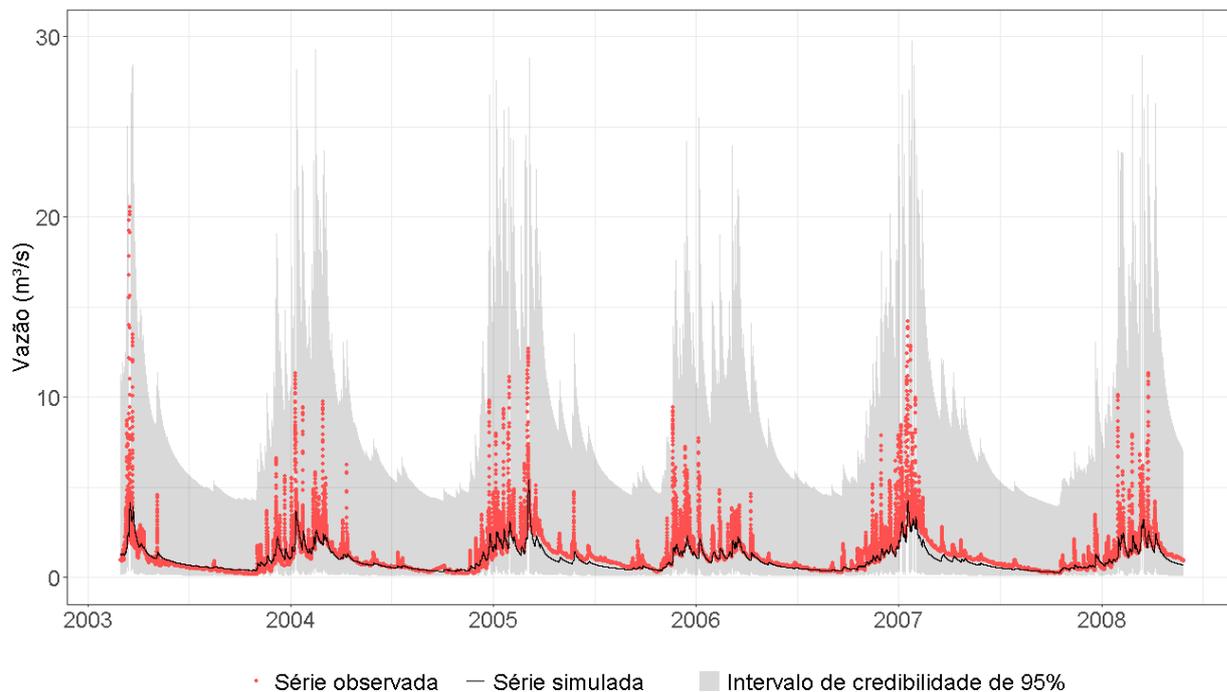


Figura I. 8 – Série temporal de vazões observadas e simuladas no domínio das assinaturas, para o período de validação - S_{FDC} .

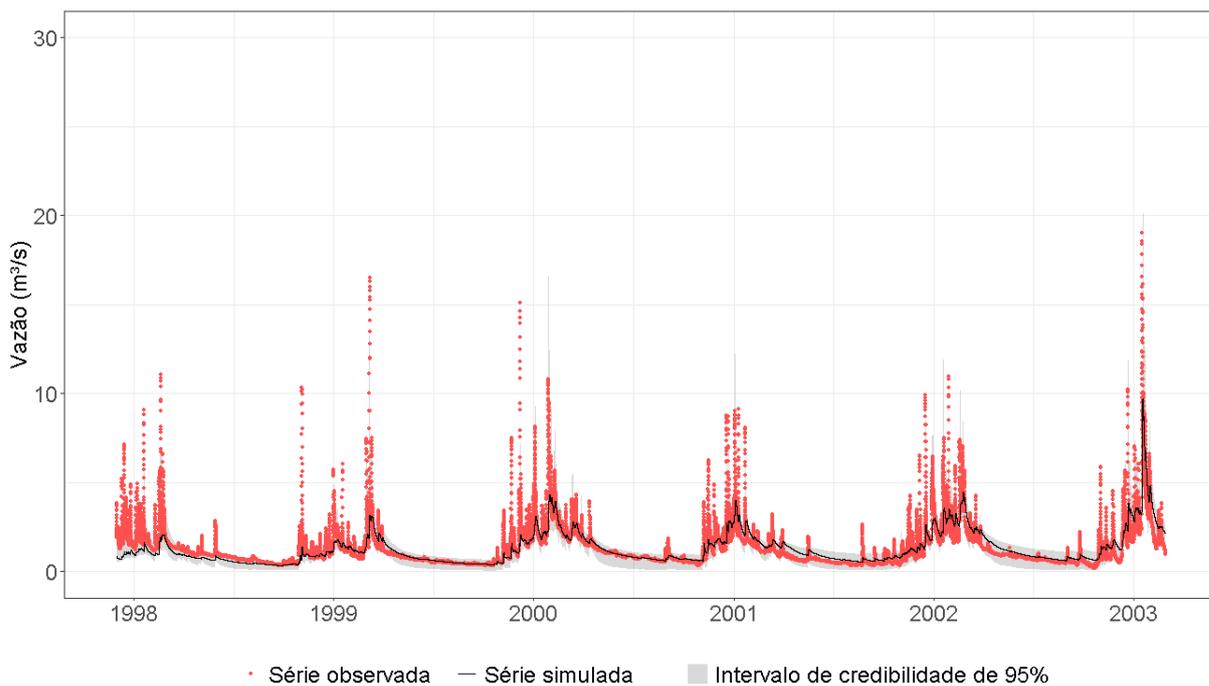


Figura I. 9 – Série temporal de vazões observadas e simuladas no domínio das assinaturas, para o período de calibração - c_a .

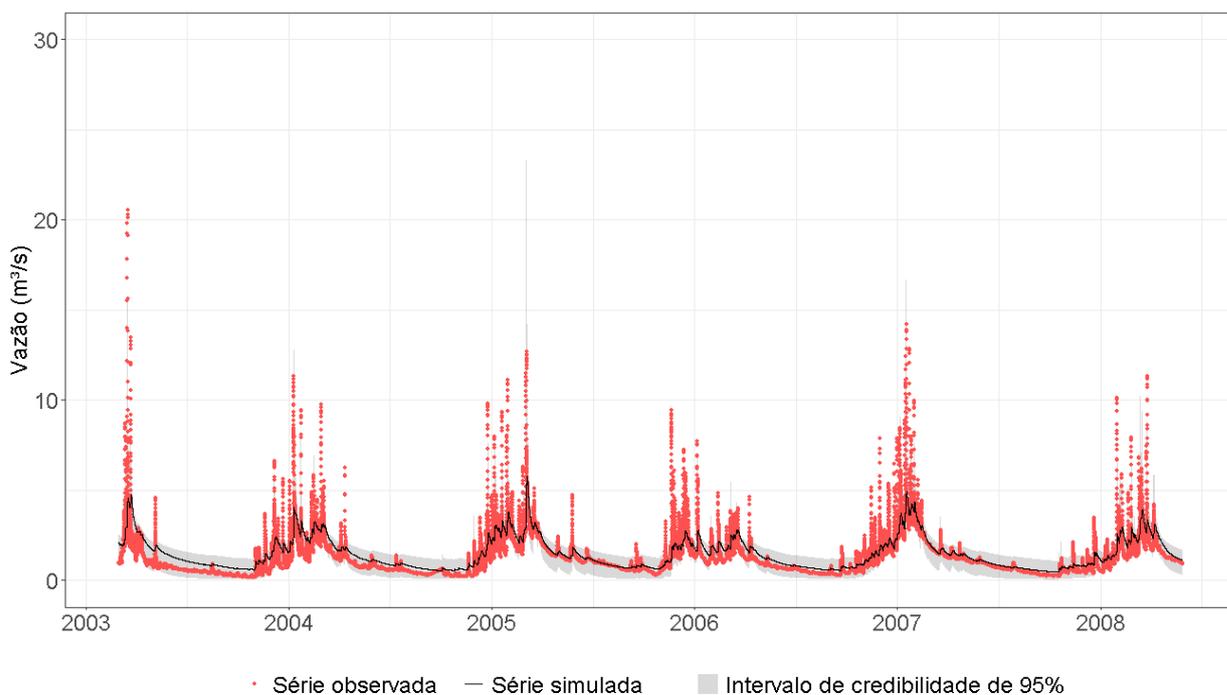


Figura I. 10 – Série temporal de vazões observadas e simuladas no domínio das assinaturas, para o período de validação - c_a .

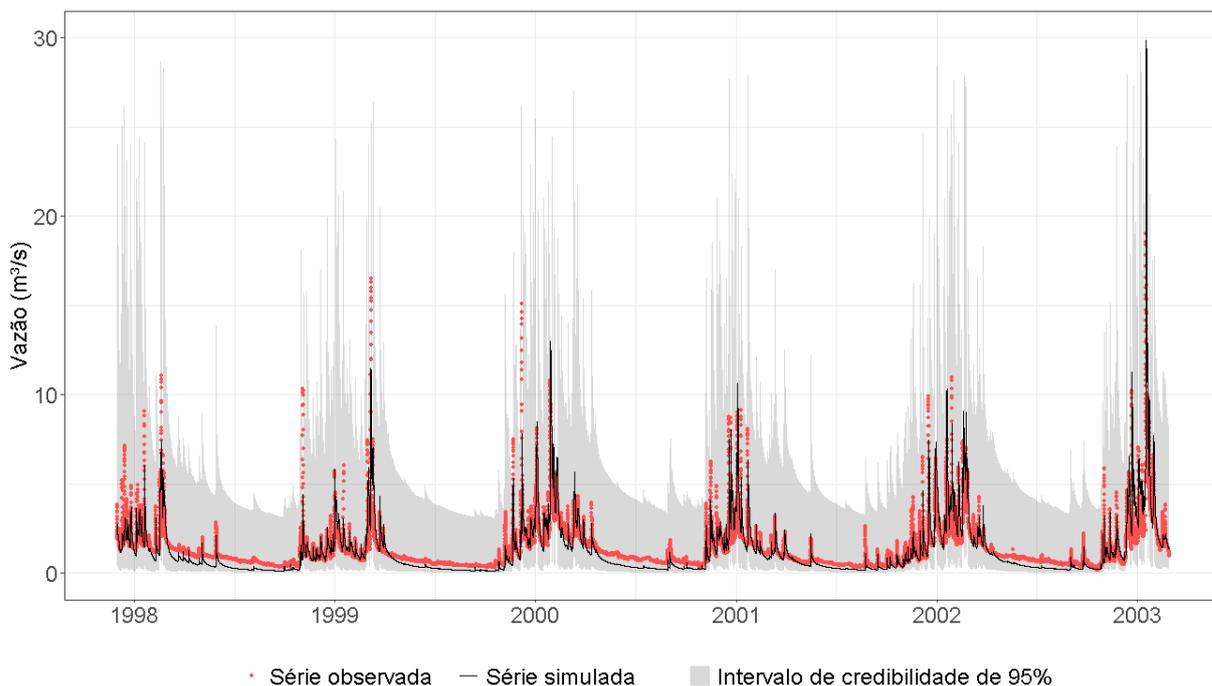


Figura I. 11 – Série temporal de vazões observadas e simuladas no domínio das assinaturas, para o período de calibração - F_L .

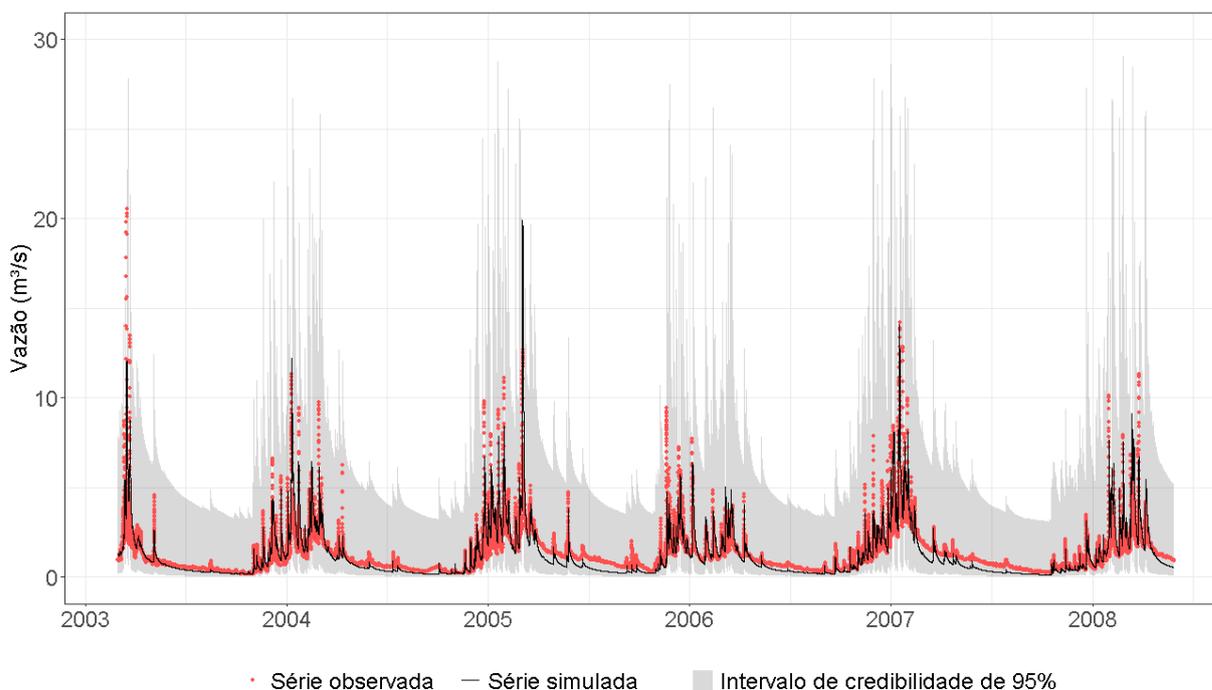


Figura I. 12 – Série temporal de vazões observadas e simuladas no domínio das assinaturas, para o período de validação - F_V .

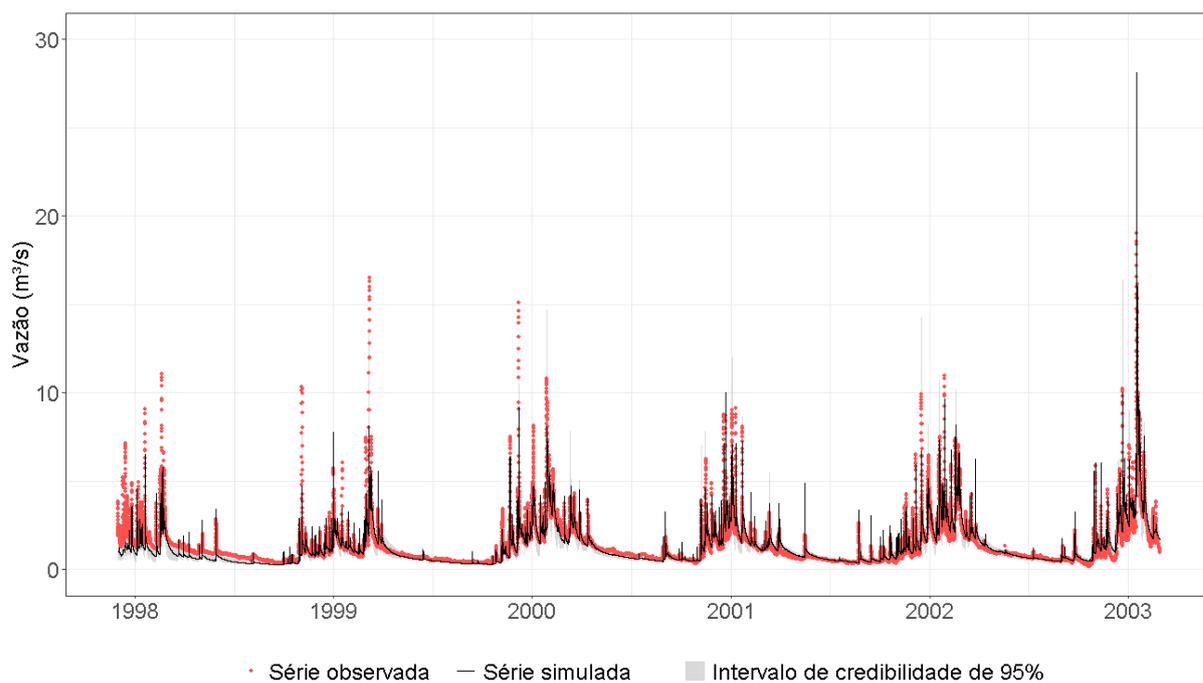


Figura I. 13 – Série temporal de vazões observadas e simuladas no domínio das assinaturas, para o período de calibração - $P_{FDC} + F_I$.

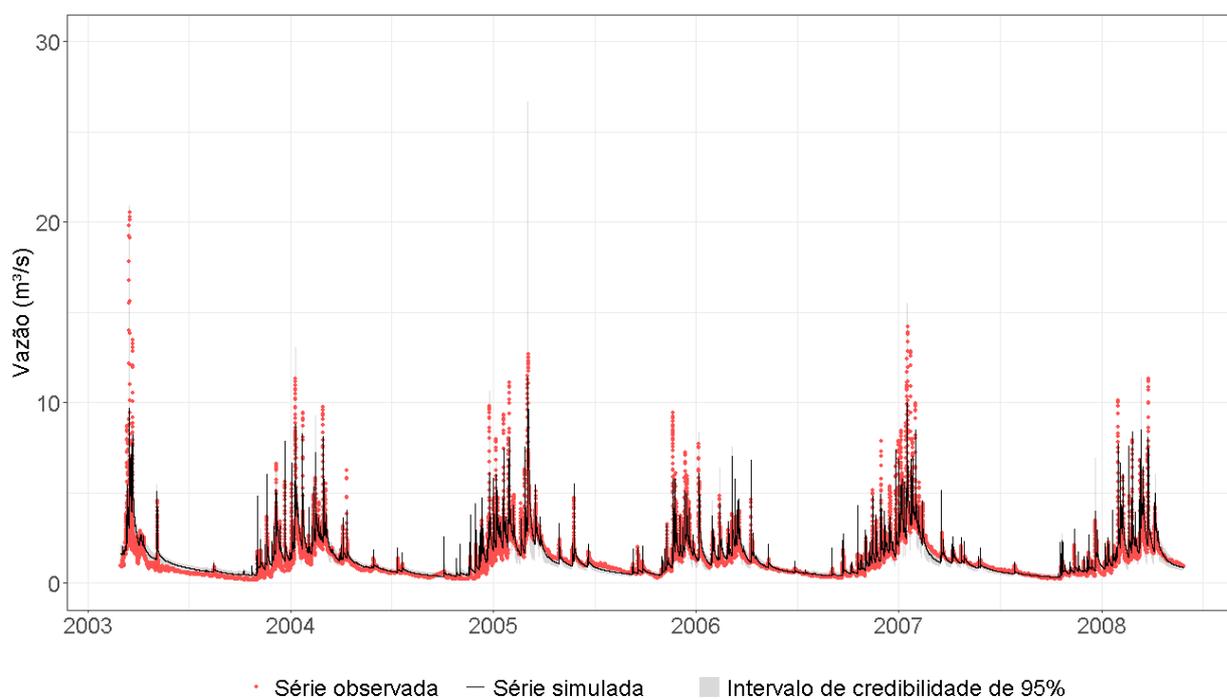


Figura I. 14 – Série temporal de vazões observadas e simuladas no domínio das assinaturas, para o período de validação - $P_{FDC} + F_I$.

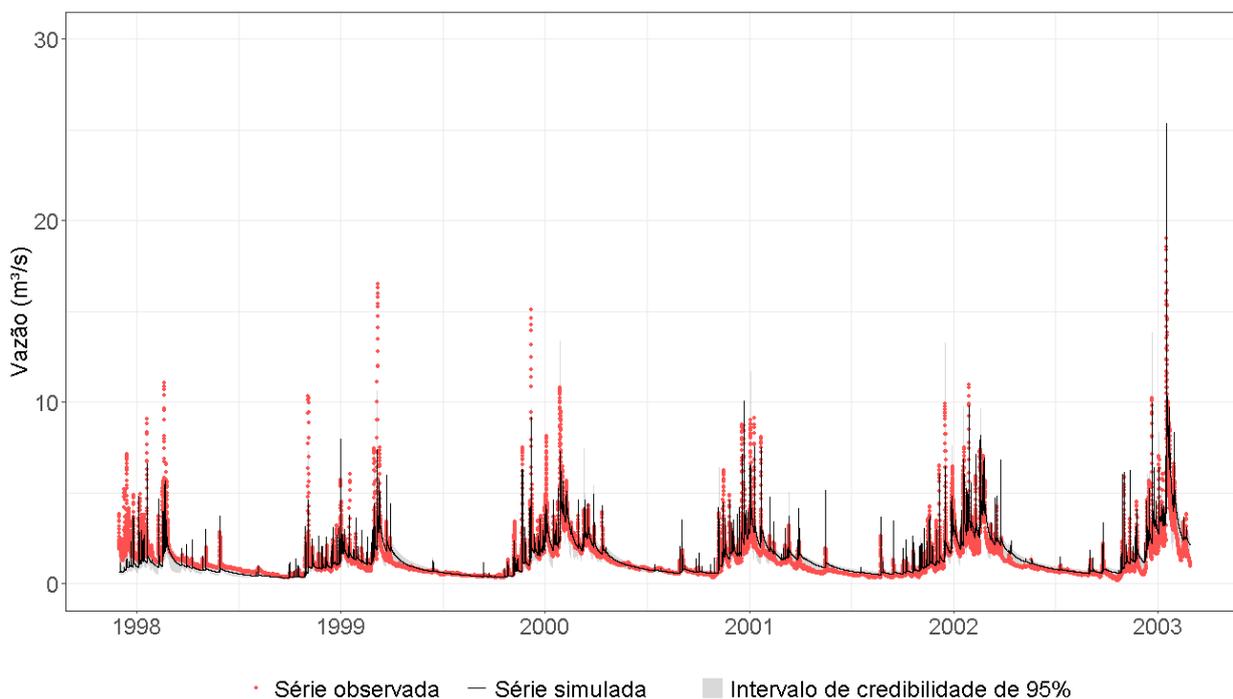


Figura I. 15 – Série temporal de vazões observadas e simuladas no domínio das assinaturas, para o período de calibração - $c_a + F_L$.

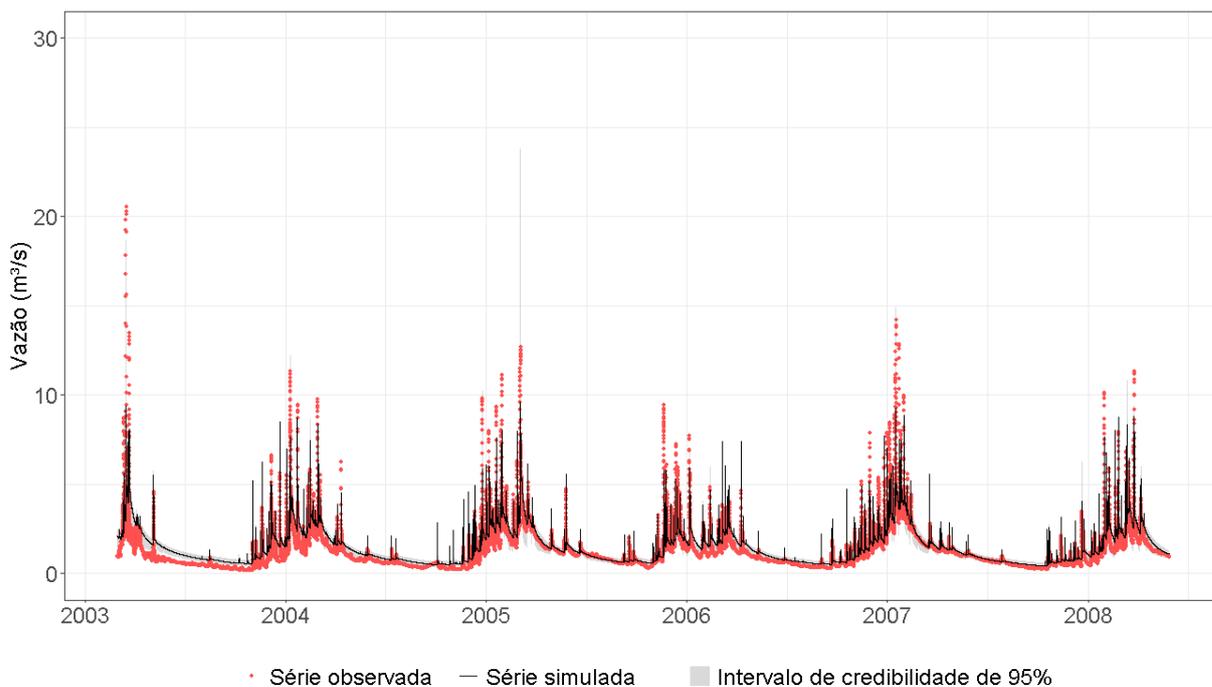
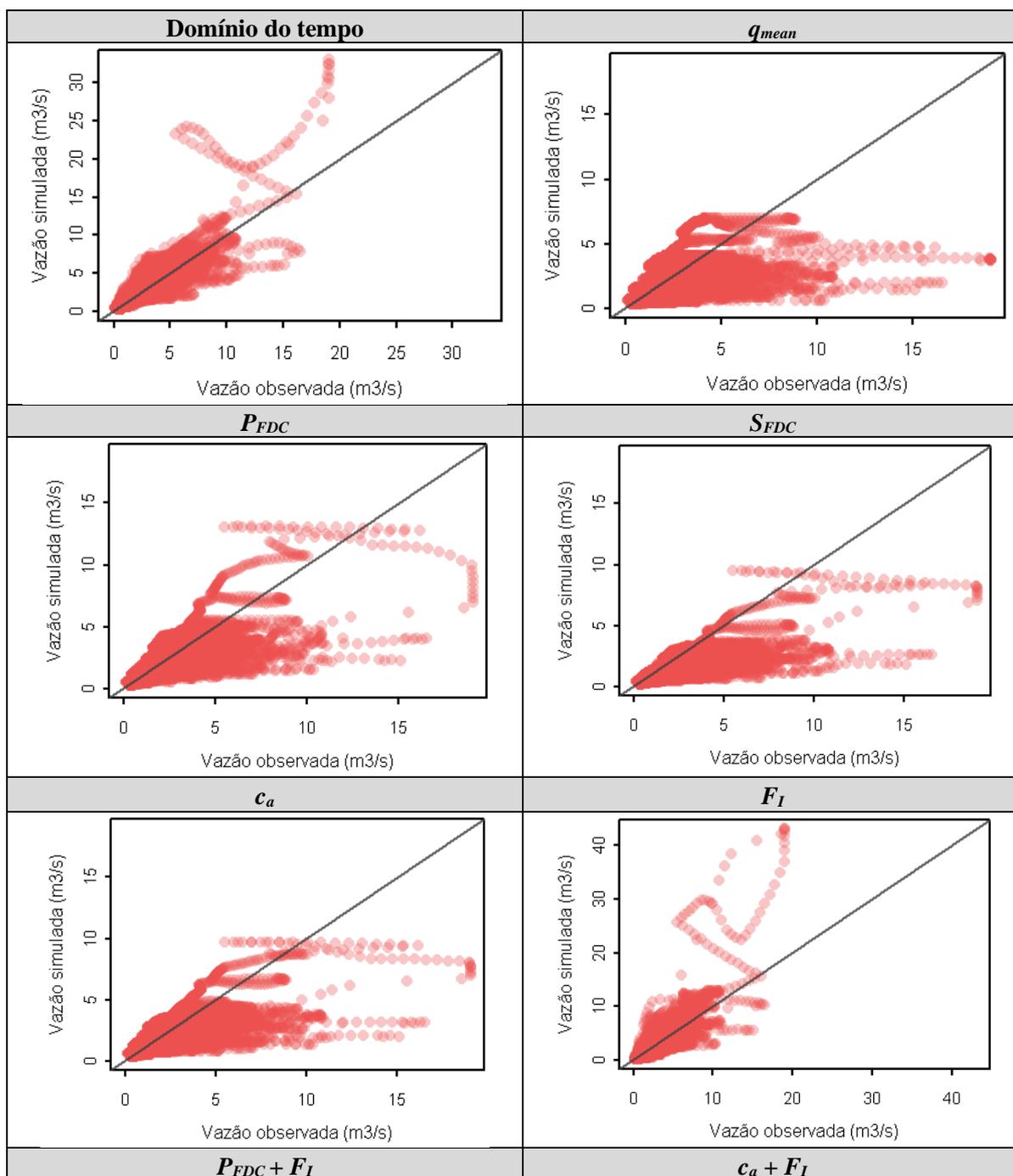


Figura I. 16 – Série temporal de vazões observadas e simuladas no domínio das assinaturas, para o período de validação - $c_a + F_L$.



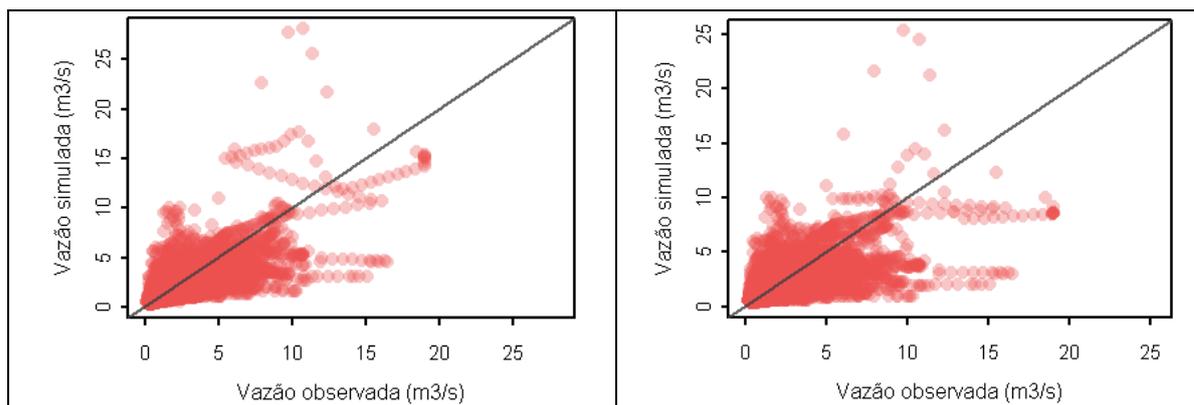
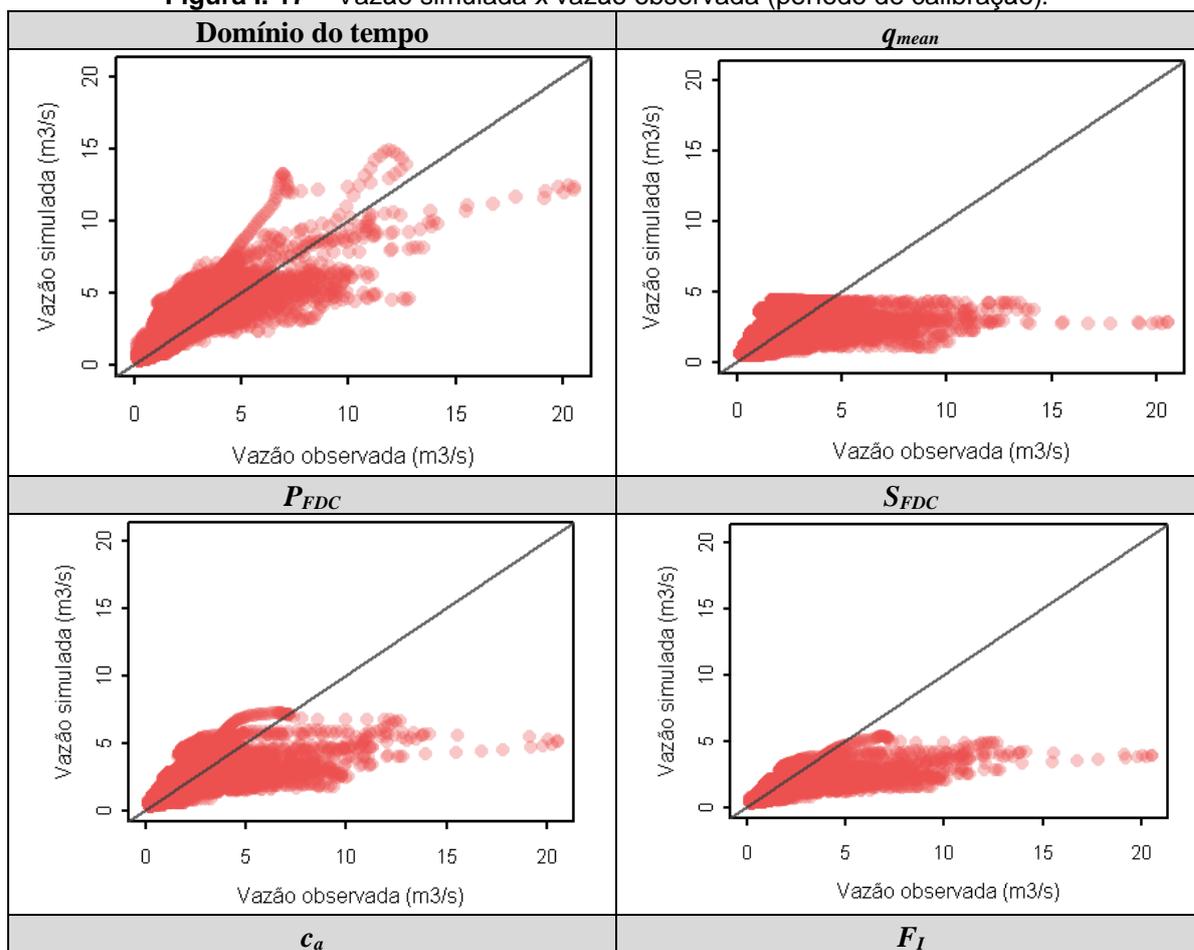


Figura I. 17 – Vazão simulada x vazão observada (período de calibração).



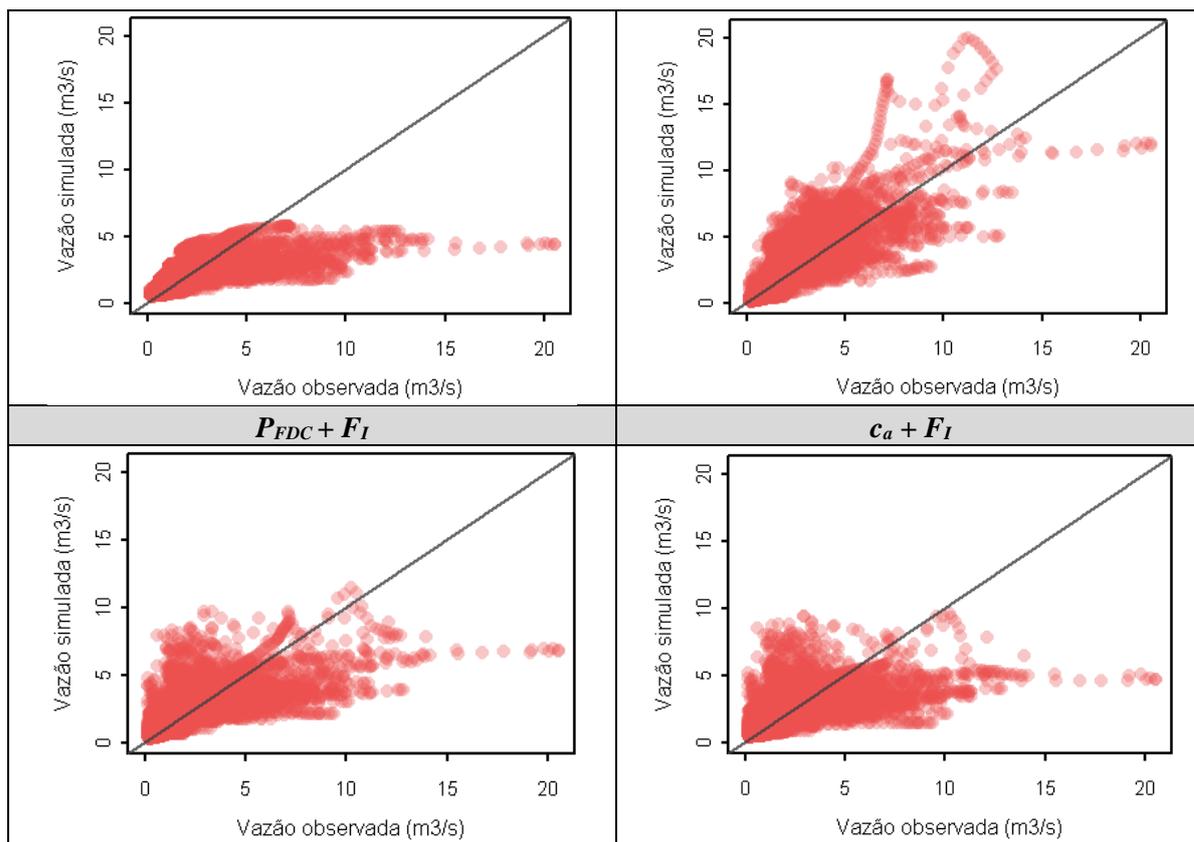


Figura I. 18 – Vazão simulada x vazão observada (período de validação).

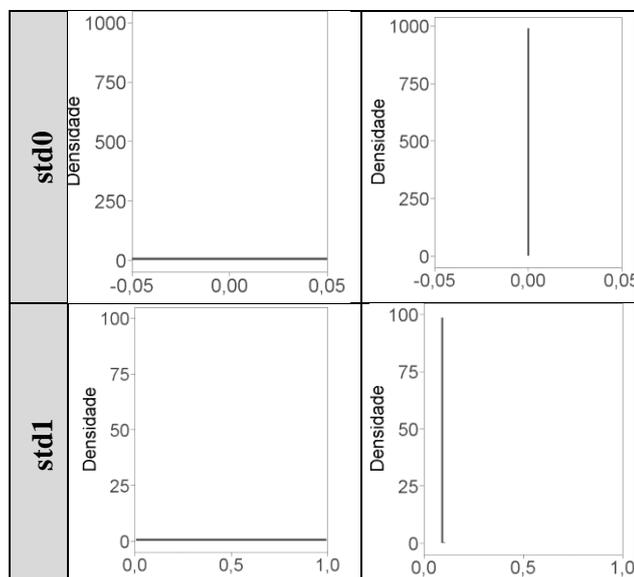


Figura I. 19 – Distribuição *a posteriori* dos parâmetros da função de verossimilhança generalizada, empregada no domínio do tempo.

APÊNDICE II – DOMÍNIO TEMPO VERSUS DOMÍNIO DAS ASSINATURAS: ESCALA MENSAL

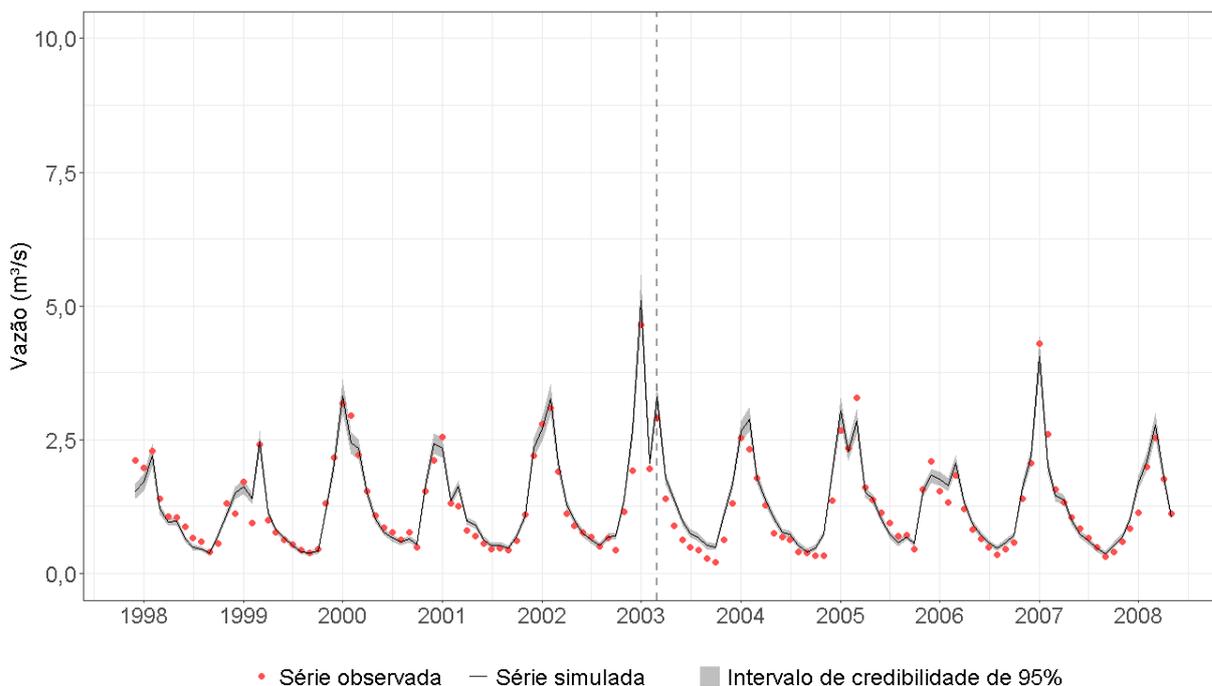


Figura II. 1 – Série temporal de vazões observadas e simuladas no domínio do tempo.

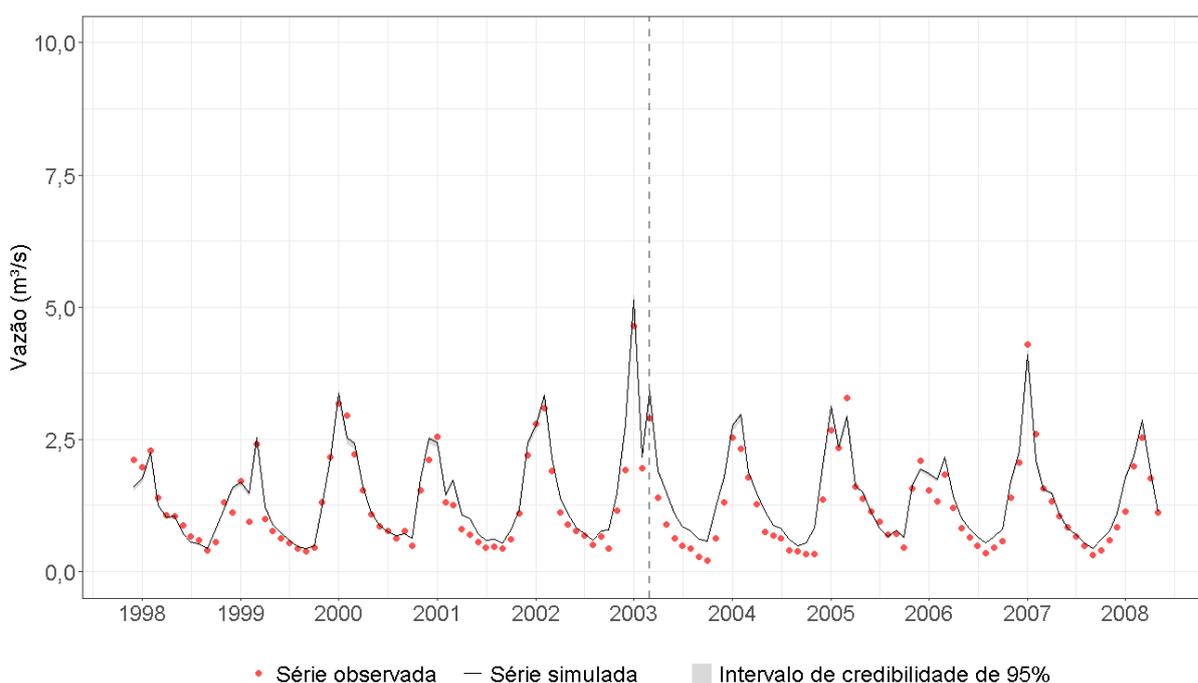


Figura II. 2 – Série temporal de vazões observadas e simuladas no domínio das assinaturas – q_{mean} .

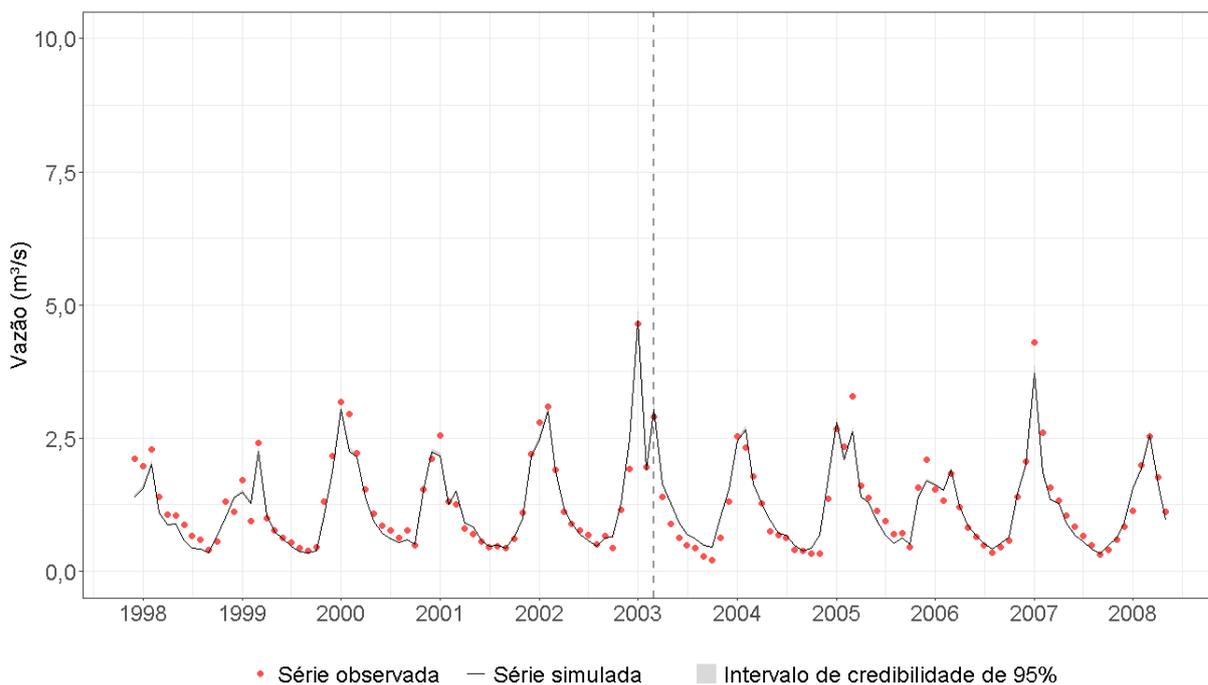


Figura II. 3 – Série temporal de vazões observadas e simuladas no domínio das assinaturas – P_{FDC} .

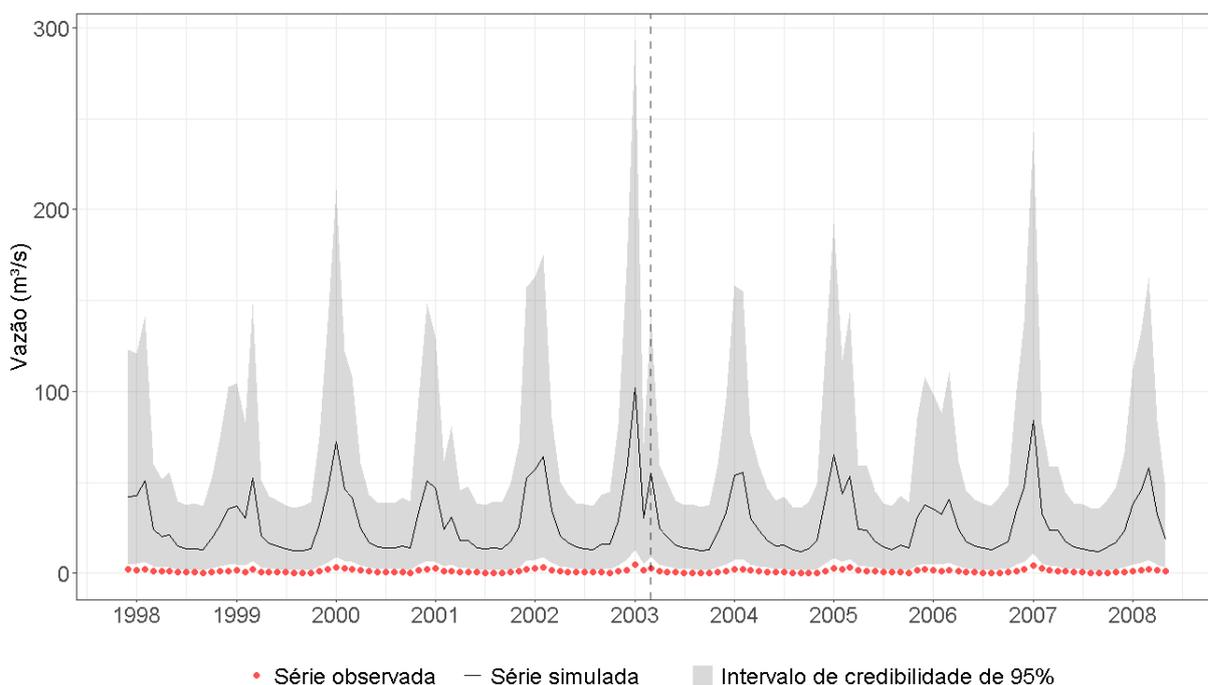


Figura II. 4 – Série temporal de vazões observadas e simuladas no domínio das assinaturas – S_{FDC} .

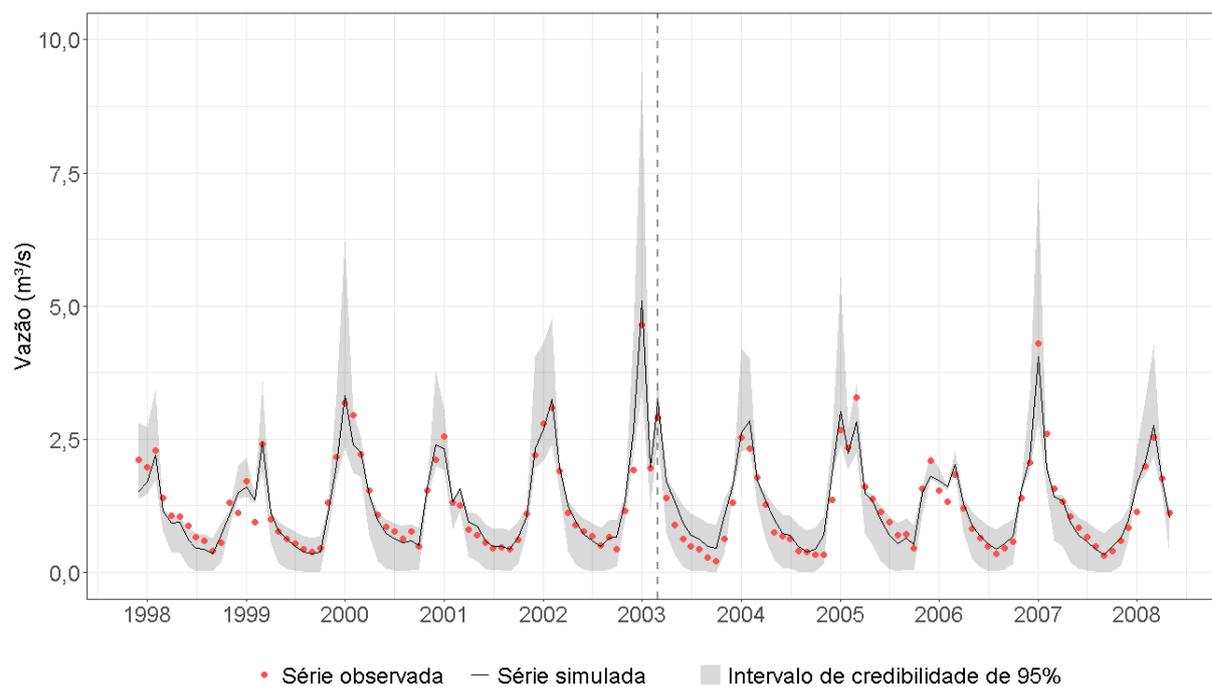


Figura II. 5 – Série temporal de vazões observadas e simuladas no domínio das assinaturas – c_a

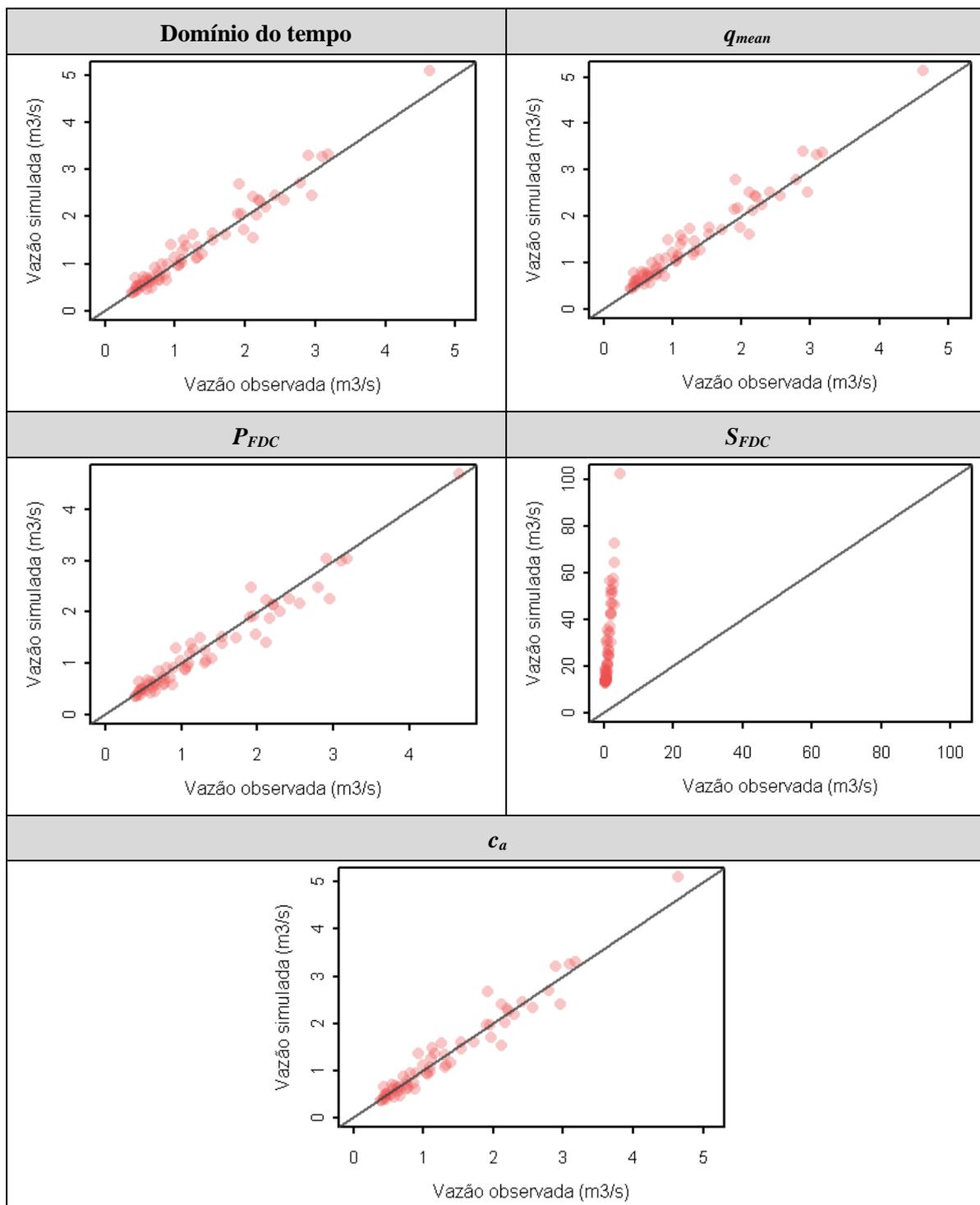


Figura II. 6 – Vazão simulada x vazão observada (período de calibração).

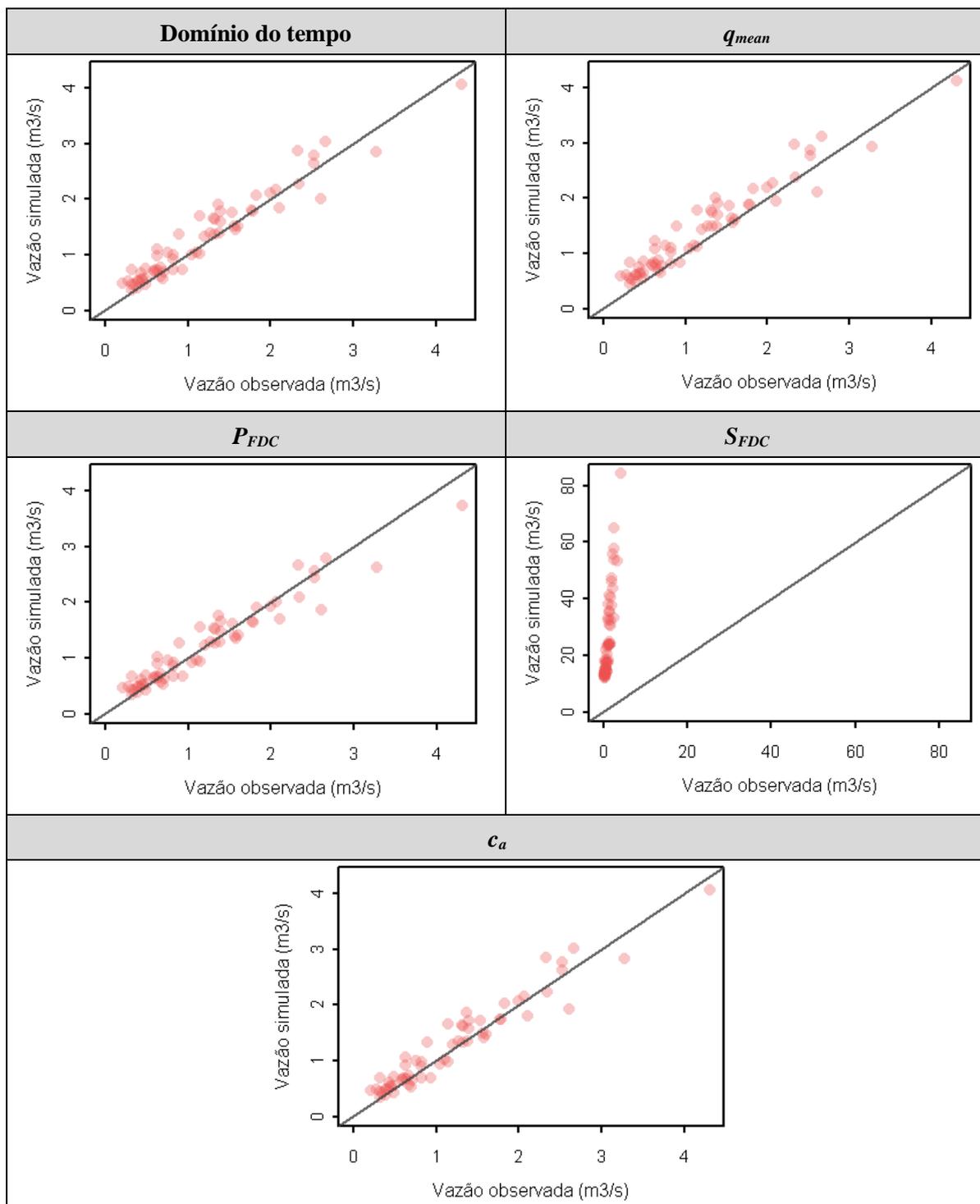


Figura II. 7 – Vazão simulada x vazão observada (período de validação).

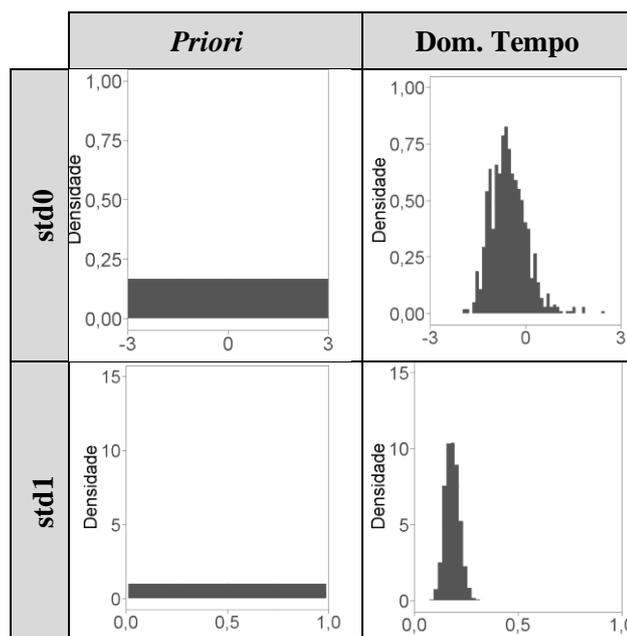


Figura II. 8 – Distribuição *a posteriori* dos parâmetros da função de verossimilhança generalizada, empregada no domínio do tempo.

APÊNDICE III – ASSINATURAS INFERIDAS A PARTIR DE REGISTROS REGIONAIS

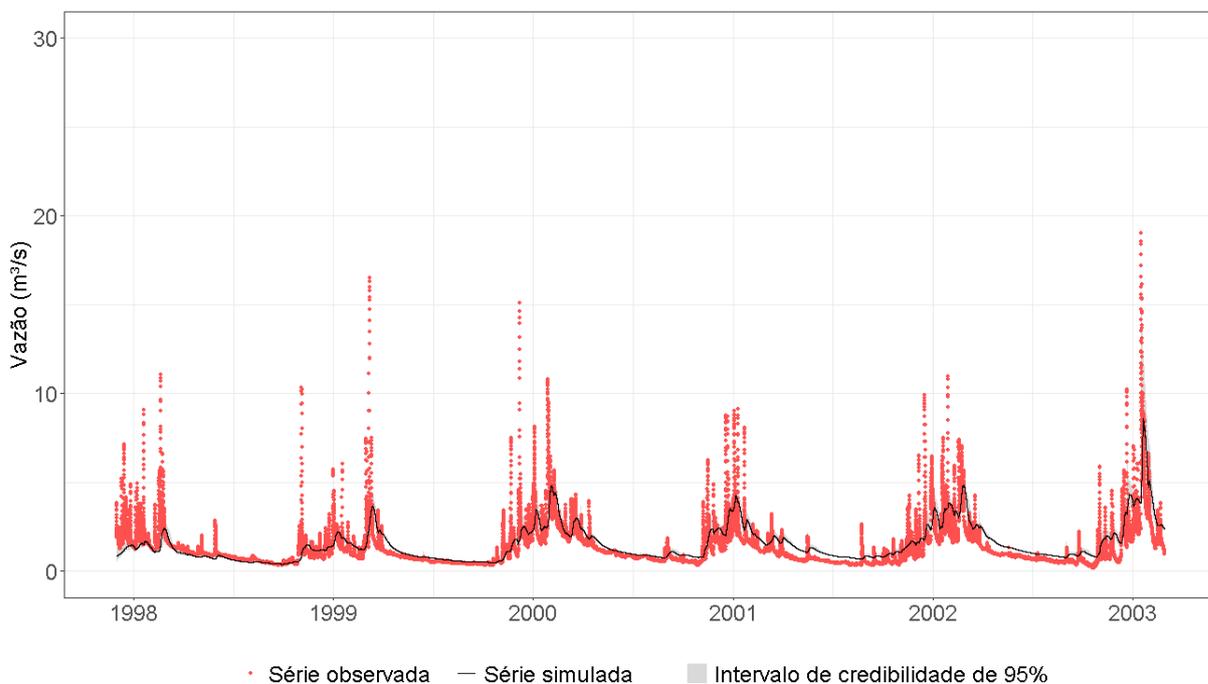


Figura III. 1 – Série temporal de vazões observadas e simuladas no domínio das assinaturas, para o período de calibração – q_{mean} .

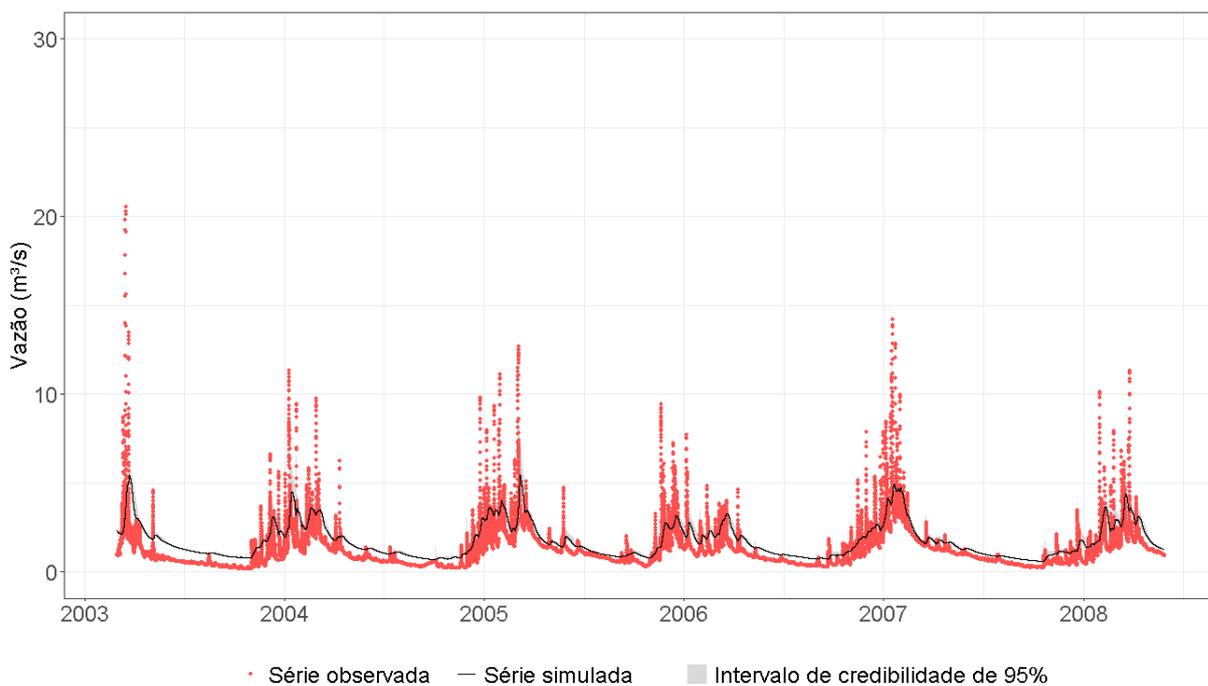


Figura III. 2 – Série temporal de vazões observadas e simuladas no domínio das assinaturas, para o período de validação – q_{mean} .

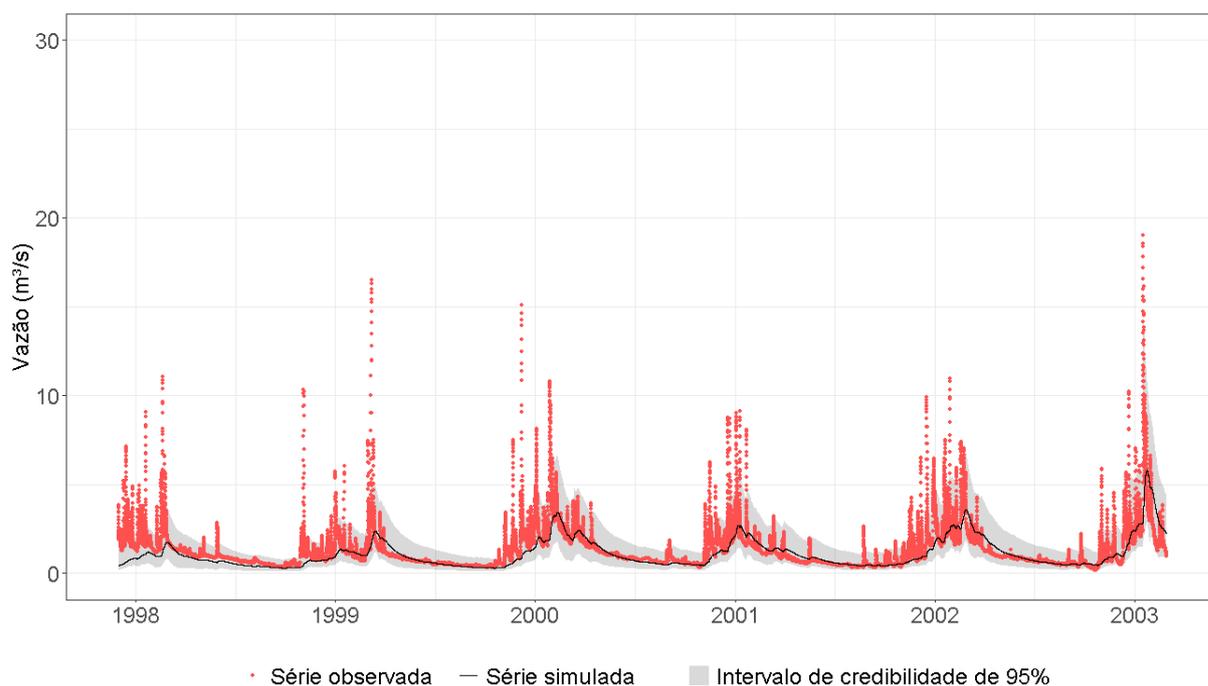


Figura III. 3 – Série temporal de vazões observadas e simuladas no domínio das assinaturas, para o período de calibração – P_{FDC} .

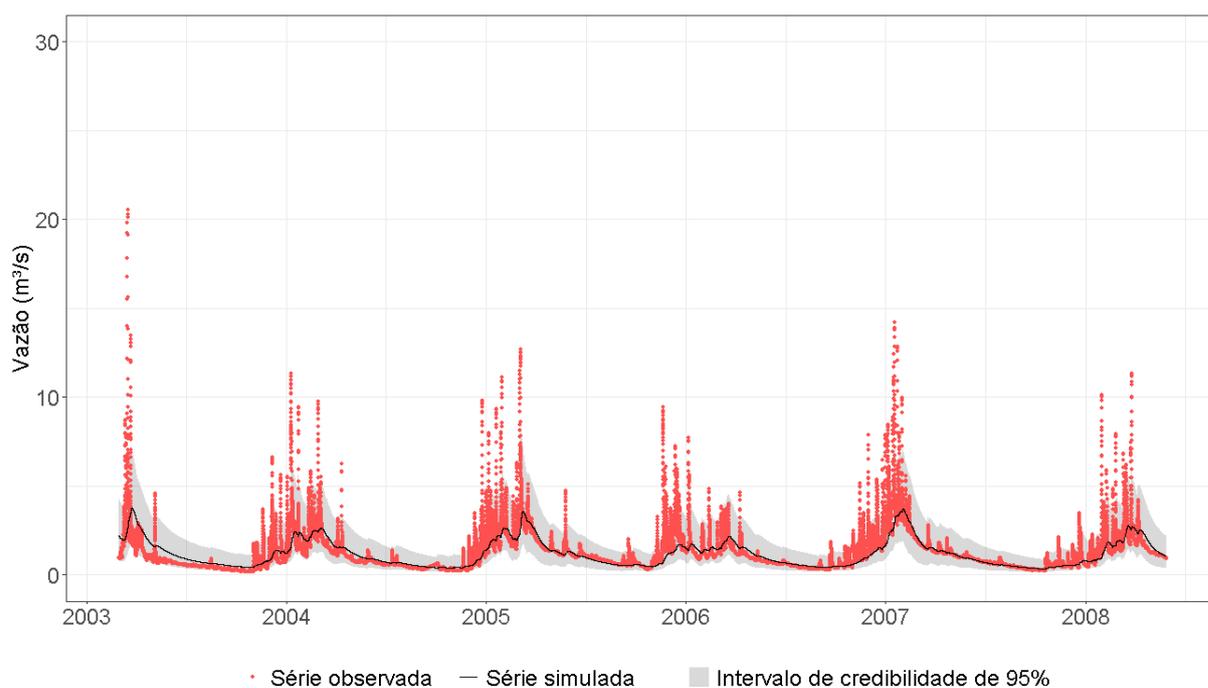


Figura III. 4 – Série temporal de vazões observadas e simuladas no domínio das assinaturas, para o período de validação – P_{FDC} .

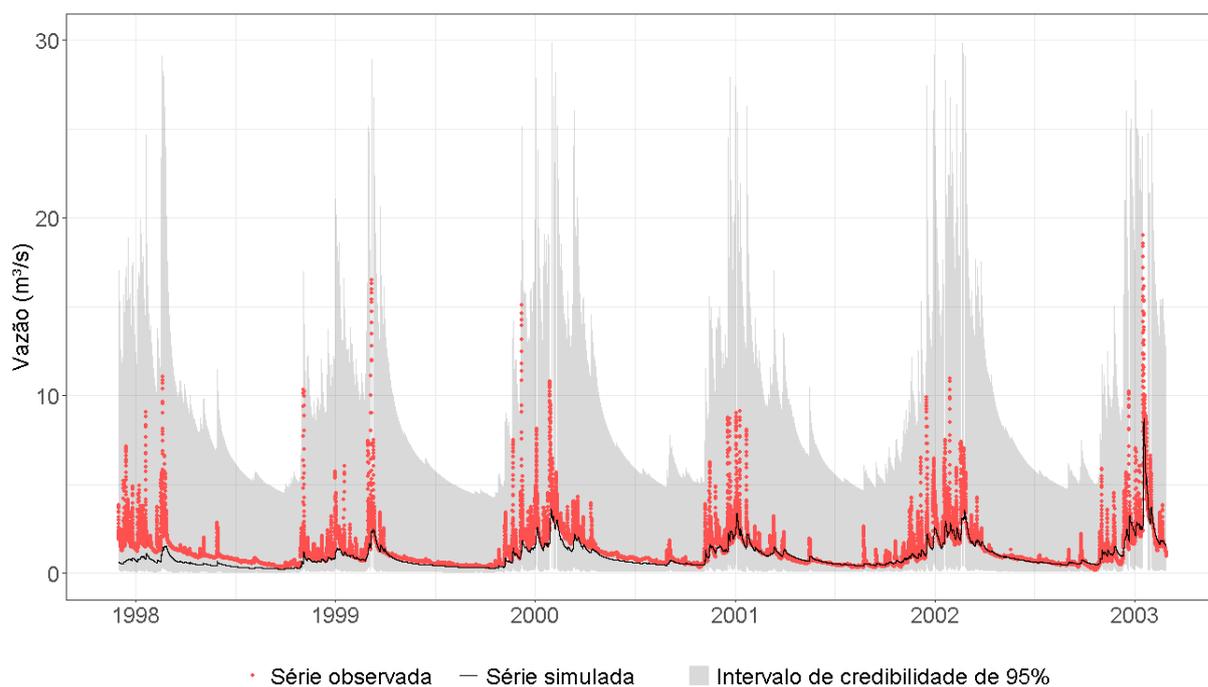


Figura III. 5 – Série temporal de vazões observadas e simuladas no domínio das assinaturas, para o período de calibração – S_{FDC} .

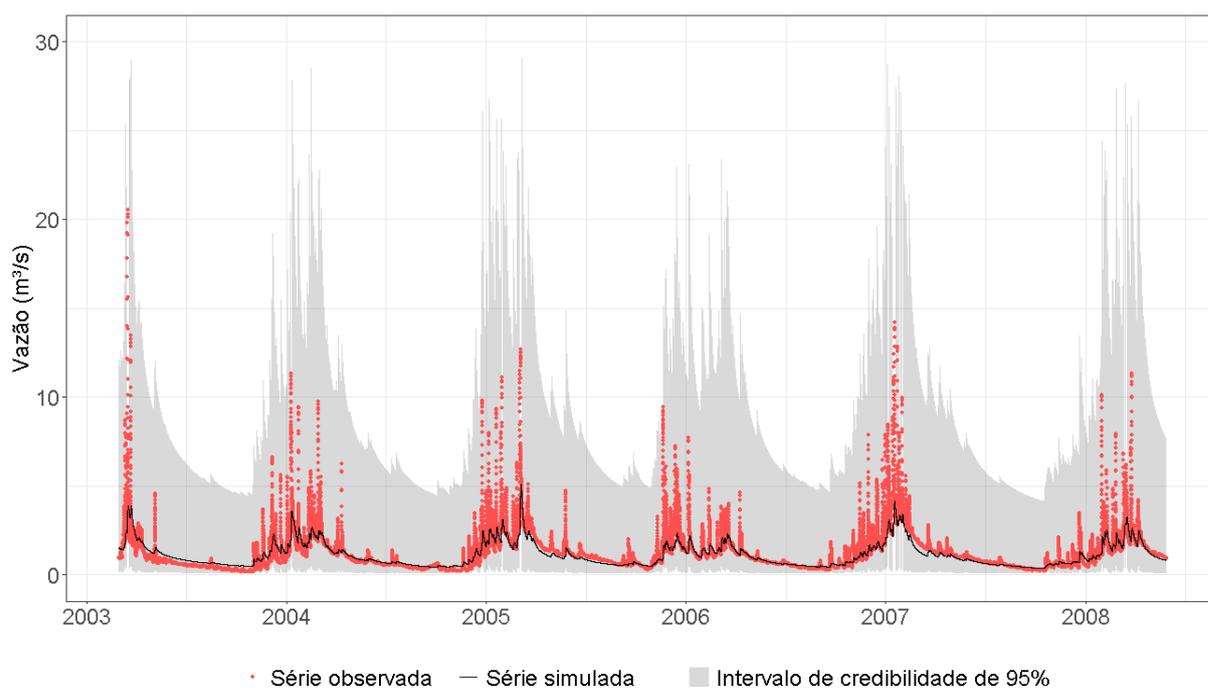


Figura III. 6 – Série temporal de vazões observadas e simuladas no domínio das assinaturas, para o período de validação – S_{FDC} .

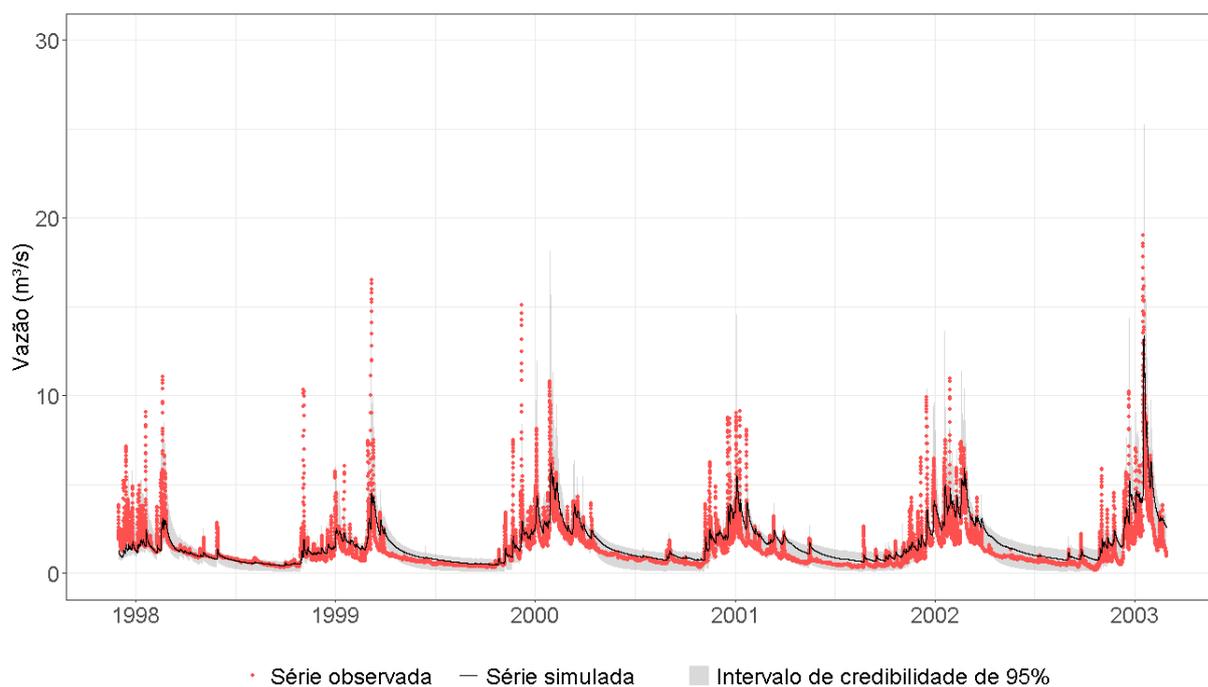


Figura III. 7 – Série temporal de vazões observadas e simuladas no domínio das assinaturas, para o período de calibração – C_a .

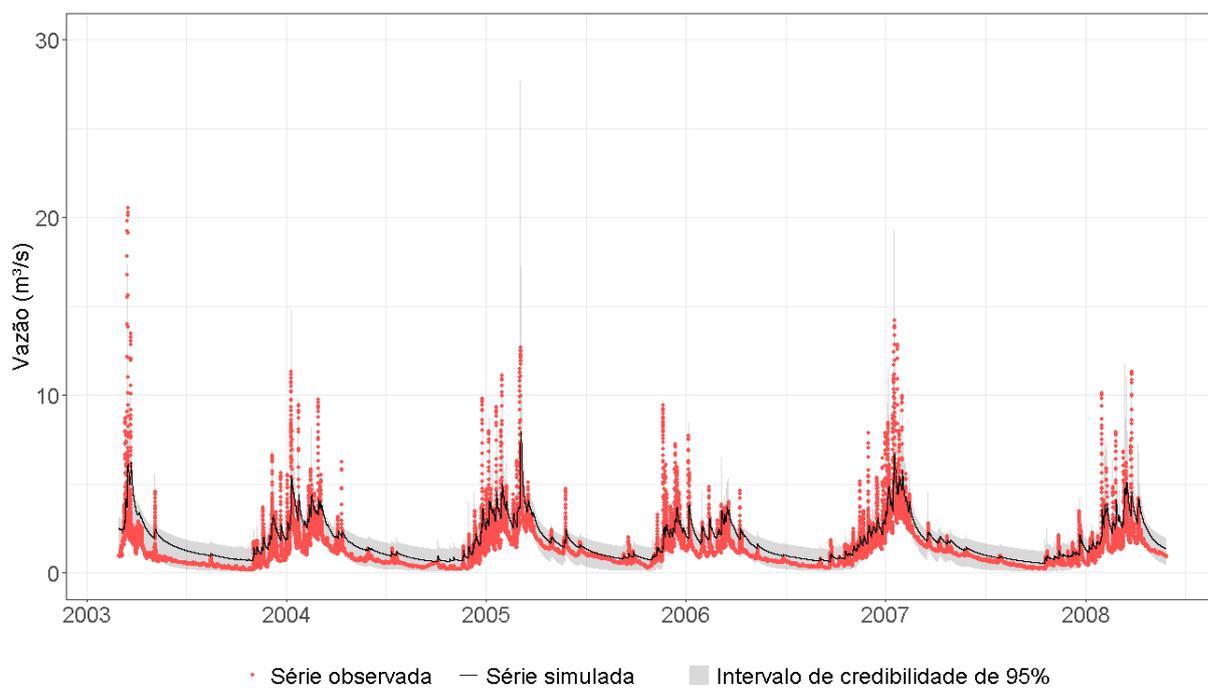


Figura III. 8 – Série temporal de vazões observadas e simuladas no domínio das assinaturas, para o período de validação – c_a .

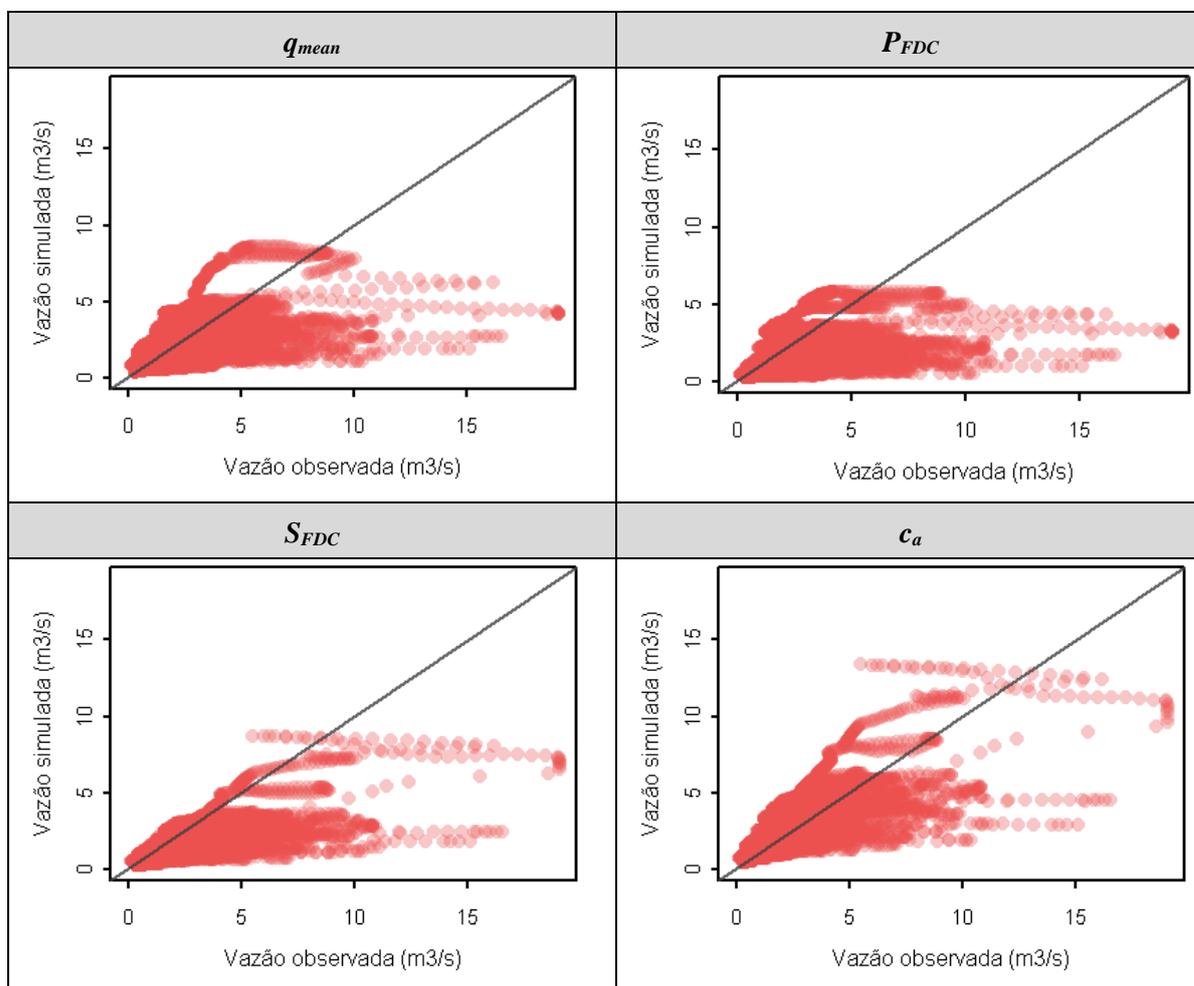


Figura III. 9 – Vazão simulada x vazão observada (período de calibração).

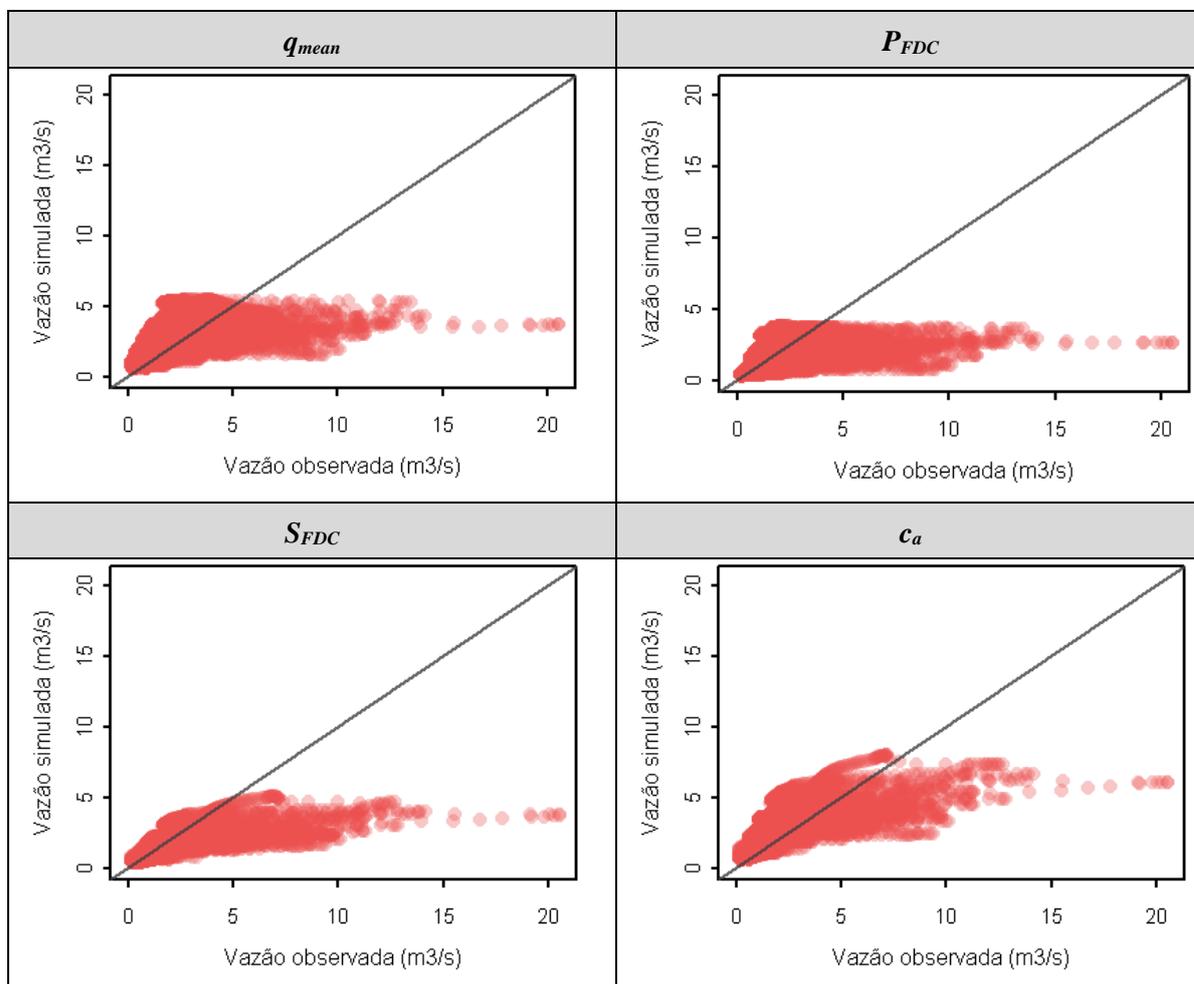


Figura III. 10 – Vazão simulada x vazão observada (período de validação).