

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

QUANTIFICAÇÃO DE INCERTEZAS
ASSOCIADAS A VARIÁVEIS INTERVENIENTES
NA MODELAGEM HIDRODINÂMICA POR MEIO
DE MÉTODOS BAYESIANOS

Viviane Borda Pinheiro Rocha

Belo Horizonte
2019

Viviane Borda Pinheiro Rocha

**QUANTIFICAÇÃO DE INCERTEZAS ASSOCIADAS A
VARIÁVEIS INTERVENIENTES NA MODELAGEM
HIDRODINÂMICA POR MEIO DE MÉTODOS
BAYESIANOS**

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

Área de concentração: Hidráulica e Recursos Hídricos

Linha de pesquisa: Modelagem de Processos Hidrológicos

Orientador: Mauro da Cunha Naghettini

Coorientador: Luiz Rafael Palmier

Belo Horizonte

Escola de Engenharia da UFMG

2019

R672q

Rocha, Viviane Borda Pinheiro.

Quantificação de incertezas associadas a variáveis intervenientes na modelagem hidrodinâmica por meio de métodos bayesianos [recurso eletrônico] / Viviane Borda Pinheiro Rocha. – 2019.

1 recurso online (254 f. : il., color.) : pdf.

Orientador: Mauro da Cunha Naghettini.

Coorientador: Luiz Rafael Palmier.

Tese (doutorado) Universidade Federal de Minas Gerais, Escola de Engenharia.

Bibliografia: f. 245-253.

Exigências do sistema: Adobe Acrobat Reader.

1. Engenharia sanitária - Teses. 2. Recursos hídricos - Desenvolvimento - Teses. 3. Incerteza - Teses. 4. Teoria bayesiana de decisão estatística - Teses. 5. Método de Monte Carlo - Teses. 6. Markov, Processos de - Soluções numéricas - Teses. 7. Inundações - Teses. I. Naghettini, Mauro. II Palmier, Luiz Rafael. III. Universidade Federal de Minas Gerais, Escola de Engenharia. IV. Título.

CDU: 628(043)



UNIVERSIDADE FEDERAL DE MINAS GERAIS

Escola de Engenharia

Programa de Pós-Graduação em Saneamento, Meio Ambiente e Recursos Hídricos

Avenida Antônio Carlos, 6627 - 4º andar - 31270-901 - Belo Horizonte – BRASIL

Telefax: 55 (31) 3409-1882 - posgrad@desa.ufmg.br

http://www.smarh.eng.ufmg.br

FOLHA DE APROVAÇÃO

Quantificação de Incertezas Associadas A Variáveis Intervinentes Na Modelagem Hidrodinâmica por Meio de Métodos Bayesianos

VIVIANE BORDA PINHEIRO ROCHA

Tese defendida e aprovada pela banca examinadora constituída pelos Senhores:

Prof. MAURO NAGHETTINI

Prof. LUIZ RAFAEL PALMIER (COORIENTADOR)

Prof. FRANCISCO EUSTÁQUIO OLIVEIRA E SILVA

Prof. VÊBER AFONSO FIGUEIREDO COSTA

Prof. DIRCEU SILVEIRA REIS JÚNIOR

Prof. DANIEL HENRIQUE MARCO DETZEL

Aprovada pelo Colegiado do PG SMARH

Prof. Antonio Teixeira de Matos
Coordenador

Versão Final aprovada por

Prof. Mauro da Cunha Naghettini
Orientador

Belo Horizonte, 19 de dezembro de 2019.

AO FELIPE E AO ANDRÉ,
fontes inesgotáveis de apoio, compreensão e alegria.

AOS MEUS AMADOS PAIS, MÁRIO E NADINE,
pelo exemplo de paciência, amor incondicional e perseverança.

AGRADECIMENTOS

Agradeço aos meus amados pais e irmã, pelo apoio incondicional nos momentos difíceis e pelas alegrias compartilhadas. Hoje percebo como minhas trajetórias profissional e pessoal estão interligadas ao nosso relacionamento como família, o que me preenche recorrentemente de plena gratidão.

Ao meu avô Germán, que sei que estará sempre torcendo pelos meus passos, e às minhas amadas avós, cujos exemplos de alegria, constância e determinação não se desgastam com o tempo.

A Deus, sem O qual nada disso seria possível.

Ao meu amado Felipe e ao meu (ainda) pequenino André, por serem companheiros de vida e parceiros de tantas alegrias e aprendizados, e por compreenderem minha ausência em alguns momentos cruciais para conclusão deste trabalho.

Aos meus queridos amigos, que me proporcionam momentos de alegria e conversas cheias de aprendizado. Sinto que sou um pouco de cada um de vocês.

Ao meus caros orientador e coorientador, professores Mauro e Palmier, pelo apoio e paciência, além das incontáveis horas dedicadas à transmissão de conhecimento para concretização deste trabalho. A confiança que vocês sempre depositaram em mim certamente contribuiu para que eu chegasse até aqui, mesmo nos momentos difíceis do doutoramento.

Aos professores Veber e Francisco, pelas conversas esclarecedoras e pela contribuição a este trabalho. Desbravar o mundo da estatística Bayesiana foi mais agradável e menos árduo graças a vocês!

Aos amigos, colegas e funcionários do EHR, pelo convívio proporcionado ao longo de tantos anos, pelo auxílio em todo tipo de aperto tão comum à vida de doutorando, pelas conversas, sorrisos e cafés! Atravessar esse caminho cheio de incertezas foi mais suave tendo sua companhia agradável.

Ao Júlio e a sua equipe da secretaria do Programa de Pós-Graduação SMARH, pelo pronto atendimento e solicitude quanto às minhas inúmeras solicitações e indagações a respeito dos inevitáveis trâmites acadêmicos.

Aos engenheiros Marcelo de Deus, Alberto, Luiz Cesar, Vanessa e Lucas, da Gerência de Planejamento Energético da CEMIG, pelo interesse em relação ao trabalho e pelo fornecimento de informações indispensáveis à realização desta pesquisa.

Ao Éber, à Alice e ao Fernando, pela oportunidade de contribuir para um estudo realizado por sua equipe na CPRM.

À FAPEMIG – Fundação de Amparo à Pesquisa de Minas Gerais, pela concessão de bolsa de estudos.

RESUMO

As enchentes e inundações e seus efeitos sobre populações ribeirinhas têm constituído objeto de estudo da comunidade científica e técnica em todo o mundo. Na contramão de avanços no campo de aquisição e análise de dados e da modelagem numérica hidrológica e hidráulica, estimativas apontam que os danos associados às inundações tendem a aumentar. Diversos estudos realizados no setor de recursos hídricos nas últimas duas décadas defendem que uma das formas de se estimarem enchentes com maior precisão e de se compreenderem profundamente os modelos utilizados é a quantificação das distintas incertezas, ou erros, que afetam as previsões e predições estimadas pelos métodos e modelos numéricos voltados à síntese chuva-vazão e à propagação de cheias. Nesse sentido, esta pesquisa de doutoramento prospectou distintos métodos disponíveis para quantificação de incertezas oriundas das fontes de erros associadas a uma modelagem hidrodinâmica, aos dados de entrada que a alimentam e às saídas geradas, bem como para sua propagação sobre o resultado final, expresso sob a forma de hidrogramas, cotogramas, perfis de escoamento e mapas de inundação associados a enchentes importantes. Da pesquisa bibliográfica realizada, definiu-se a utilização de um esquema de abordagem Bayesiana, que consiste basicamente na atualização do conhecimento que se tem *a priori* acerca de parâmetros e dados necessários à modelagem à luz de informações observadas sobre o fenômeno modelado, como hidrogramas registrados no extremo de jusante de um trecho fluvial, ou manchas de inundação disponíveis. A adoção de métodos de cunho Bayesiano à modelagem hidrodinâmica, caracterizada por equações não-lineares de solução numérica, é viabilizada por meio da utilização de Métodos de Monte Carlo por Cadeias de Markov (MCMC), a fim de garantir a aproximação da chamada distribuição *a posteriori* dos parâmetros, dados e variáveis envolvidas na inferência, ou seja, obtida após a incorporação das informações observadas disponíveis sobre as cheias em determinado local. Para aplicação da metodologia, foi escolhido um estirão fluvial do rio São Francisco, entre a confluência com o rio Abaeté e a cidade de Pirapora, no estado de Minas Gerais, com disponibilidade de dados de vazão média diária de ponta-a-ponta, de 22 seções topobatimétricas e de um modelo digital de terreno com precisão horizontal de 1 m. Os resultados mostram que a metodologia adotada é adequada à quantificação das incertezas oriundas dos coeficientes de rugosidade de Manning na calha fluvial e na planície de inundação nesse tipo de modelo e abre uma série de possibilidades quanto a estudos futuros concernentes ao tema. O método garantiu também a obtenção de intervalos de credibilidade caracterizadores da incerteza em predições do modelo,

tais como hidrogramas e cotagramas na seção utilizada para a inferência Bayesiana e o controle de resultados.

Palavras-chave: Estimção de incertezas. Inferência Bayesiana. Modelos hidrodinâmicos. Métodos de Monte Carlo por Cadeias de Markov. Mapas de inundação probabilísticos.

ABSTRACT

Floods and their effects on riverine populations have been the subject of study of the scientific and technical community all over the world. Despite some improvements in modelling and data acquisition fields, estimates indicate that flood damage tends to increase. Several studies in the water sector over the past two decades have argued that one of the ways to estimate flooding more accurately and to understand the models used is to quantify the different uncertainties, or errors, that affect predictions as estimated by numerical methods and rainfall-runoff models and flood routing. In this sense, this doctorate research prospected different methods available for quantifying uncertainties associated with some parameters, data and variables involved in mathematical modeling of floods and their quantification and propagation on the final result, in this case expressed as hydrographs, as well as water surface elevation profiles and inundation maps. From the bibliographical inspection, it was defined the application of a Bayesian approach scheme, which basically consists of updating the prior knowledge about parameters and variables involved in modeling using observed information related to the phenomenon under consideration, such as flow and stage hydrographs registered in the downstream end of a river reach, and available flood inundation maps. The adoption of Bayesian methods for hydrodynamic modeling, characterized by nonlinear equations of numerical solution, is possible by means of Markov Chain Monte Carlo Methods (MCMC), in order to ensure the approximation of the posterior distribution of parameters, data and variables involved in the inference, that is, obtained after incorporating the observed information available on floods at a given location. For the application and validation of the methodology, a fluvial reach in the upper São Francisco river was selected, located between the Abaeté river outlet and the town of Pirapora, in Minas Gerais state, with daily mean flow data available in the upstream and downstream ends, 22 cross sections and a digital terrain model with horizontal accuracy of 1 m. The results show that the adopted methodology is adequate to quantify the uncertainties arising from the Manning's roughness coefficients in the channel and floodplains in this type of model and opens a series of possibilities for future studies concerning the subject. The method also ensured the estimation of credibility intervals that characterize the uncertainty in model predictions, such as flow and stage hydrographs in the section used for the Bayesian inference and control of results.

Keywords: Uncertainty estimation. Bayesian inference. Hydrodynamic models. Markov Chain Monte Carlo methods. Probabilistic inundation maps.

LISTA DE FIGURAS

Figura 1.1 – Estrutura de análise de incertezas com vistas à sua aplicação no gerenciamento integrado de riscos das inundações, proposta por Hall e Solomatine (2008).	36
Figura 3.1 – Representação esquemática de um sistema ambiental com seus componentes (adaptado de Liu e Gupta, 2007).	42
Figura 3.2 – Representação esquemática da aplicação do Teorema de Bayes à estimação de incertezas presentes na modelagem de sistemas ambientais, com ênfase sobre a incerteza paramétrica (adaptado de Vrugt, 2016).	56
Figura 3.3 – Diagrama explicativo do BaTEA e das premissas para sua dedução (adaptado de Kavetski <i>et al.</i> , 2003).	67
Figura 3.4 – Diagrama esquemático do método GLUE.	77
Figura 3.5 – Função adotada para estimação do fator de ponderação a ser adotado para cada célula do MDT que tiver sido inundada, seja nas simulações, seja na mancha observada, conforme proposição de Pappenberger <i>et al.</i> (2005). Nesse caso, a precisão do MDT era de 0,15 cm.	89
Figura 3.6 – Mapa probabilístico da cheia de 100 anos associado à incerteza do parâmetro CN, para trecho do rio Virgin, no Utah, Estados Unidos, elaborado no estudo de incertezas de Smemoe <i>et al.</i> (2007).	92
Figura 3.7 – Mapas determinístico (esq.) e probabilístico (dir.) modelados para um trecho do rio Dee, na Escócia. No segundo mapa, à direita, foram consideradas as incertezas associadas ao coeficiente de rugosidade de Manning na calha principal.	95
Figura 3.8 – Trajetórias percorridas por partículas fluidas em leitos e planícies fluviais.	102
Figura 4.1 – Elementos envolvidos numa estrutura Bayesiana para identificação de incertezas em uma modelagem hidrodinâmica, na forma de um procedimento de calibração.	115
Figura 4.2 – Fontes de incerteza e elementos associados à modelagem hidrodinâmica e ao esquema metodológico adotado considerados na presente pesquisa. Dentre os elementos cujas incertezas podem ser quantificadas, numerados de 1 a 6, os que estão destacados em preto correspondem às fontes de incertezas explicitamente tratadas pela metodologia adotada.	119
Figura 4.3 – Esquema de estimação da incerteza preditiva <i>a posteriori</i> para a informação Y utilizada na calibração de incertezas pelo método Bayesiano adotado, quando são utilizados hidrogramas ou cotagramas em determinada seção de um estirão fluvial.	138
Figura 5.1 – Localização do trecho fluvial em relação à rede hidrográfica local e regional.	148
Figura 5.2 – Seções transversais típicas do rio São Francisco no trecho considerado na pesquisa.	150
Figura 5.3 – Perfil transversal do rio São Francisco entre a foz do rio Abaeté e a do rio das Velhas, e indicação do trecho utilizado na modelagem hidrodinâmica.	152
Figura 5.4 – Posição das pontes e da estação fluviométrica situadas no rio São Francisco nas proximidades das cidades de Pirapora e Buritizeiro.	153
Figura 5.5 – Fotos que mostram o aspecto da vegetação dominante das margens do rio São Francisco no trecho considerado neste estudo.	154
Figura 5.6 – Fotos que mostram o aspecto da vegetação típicas das planícies de inundação do rio São Francisco no trecho estudado.	154
Figura 5.7 – Regime de vazões médias diárias registradas no rio São Francisco a jusante de Três Marias (CEMIG) e em Pirapora-Barreiro (CPRM).	155
Figura 5.8 – Hidrograma afluente e defluente ao trecho fluvial estudado, referentes à cheia de fevereiro de 1992 (esq.) e seu cotagrama estimado no extremo de jusante (dir.).	161
Figura 5.9 – Análise de frequência de máximos anuais de vazão média diária no posto fluviométrico de jusante, no rio São Francisco em Pirapora-Barreiro.	162

Figura 5.10 – Código em Matlab da <i>toolbox</i> do método DREAM para aproximação numérica da distribuição conjunta <i>a posteriori</i> considerando a função de verossimilhança Generalizada.	163
Figura 5.11 – Código elaborado em Matlab para uso do RAS Controller a fim de alterar os parâmetros n_c e n_p a cada iteração das cadeias de Markov.	164
Figura 5.12 – Organização dos principais cenários de inferência Bayesiana de incertezas considerados na pesquisa.	166
Figura 5.13 – Gráfico A (esq.): Valores médios dos resíduos ϵ_t em cada intervalo de tempo <i>versus</i> as vazões médias diárias observadas no posto fluviométrico de jusante; Gráfico B (dir.): percentis de 2,5%, 50% e 97,5% dos resíduos ϵ_t estimados para cada intervalo de tempo. Ambos os gráficos correspondem ao Cenário A1.	169
Figura 5.14 – Dispersão dos principais descritores estatísticos das 10.000 séries de resíduos $E\theta = Y - Y\theta$ obtidas <i>a posteriori</i> : Cenário A1.	170
Figura 5.15 – Quadrado dos resíduos ϵ_t medianos e correspondentes aos percentis 2,5% e 97,5% <i>versus</i> as vazões afluentes ao trecho fluvial estudado: Cenário A1.	173
Figura 5.16 – Dispersão dos principais descritores estatísticos das 10.000 séries de resíduos $\sigma_{\text{stat}} = \epsilon_t - \phi_1 \cdot \epsilon_{t-1}$ obtidas <i>a posteriori</i> : Cenário A1.	174
Figura 5.17 – Quantis empíricos de percentis notáveis colhidos por Δt nas séries de resíduos ϵ_t e os correspondentes quantis teóricos obtidos segundo as distribuições SEP e Normal padronizadas: Cenário A1.	176
Figura 5.18 – Histogramas de frequências relativas estimados para as amostras <i>a posteriori</i> dos parâmetros do modelo hidrodinâmico e da verossimilhança Generalizada: Cenário A1.	177
Figura 5.19 – Amplitude de níveis de água correspondentes às descargas em trânsito pela seção de Pirapora-Barreiro resultantes dos 10.000 pares nc, np obtidos <i>a posteriori</i> no Cenário A1.	179
Figura 5.20 – Diagramas de dispersão de valores extraídos das amostras construídas <i>a posteriori</i> : Cenário A1.	181
Figura 5.21 – Valores médios dos resíduos brutos e após remoção da correlação serial em relação à vazão observada no posto fluviométrico de jusante: Cenário A3.	183
Figura 5.22 – Quantis empíricos de percentis notáveis colhidos por Δt nas séries de resíduos ϵ_t e os correspondentes quantis teóricos segundo as distribuições SEP e N padronizadas: Cenário A4.	187
Figura 5.23 – Histogramas de frequências relativas estimados para as amostras <i>a posteriori</i> dos parâmetros do modelo hidrodinâmico e da verossimilhança Generalizada: Cenário A4.	188
Figura 5.24 – Diagramas de dispersão de valores extraídos das amostras construídas <i>a posteriori</i> : Cenário A4.	189
Figura 5.25 – Valores médios dos resíduos brutos e após remoção da correlação serial em relação à vazão observada no posto fluviométrico de jusante: Cenário B2.	190
Figura 5.26 – Quantis empíricos de percentis notáveis colhidos por Δt nas séries de resíduos ϵ_t e os correspondentes quantis teóricos segundo as distribuições SEP e N padronizadas: Cenário B2.	192
Figura 5.27 – Histogramas de frequências relativas estimados para as amostras <i>a posteriori</i> dos parâmetros do modelo hidrodinâmico e da verossimilhança Generalizada: Cenário B2.	193
Figura 5.28 – Histogramas de frequências relativas estimados para as amostras <i>a posteriori</i> dos parâmetros do modelo hidrodinâmico e da verossimilhança Generalizada: Cenário B3.	194
Figura 5.29 – Amplitude de níveis de água na seção de Pirapora-Barreiro resultantes da simulação hidrodinâmica dos 10.000 pares nc, np obtidos <i>a posteriori</i> no Cenário B2.	197
Figura 5.30 – Quadrado dos resíduos ϵ_t medianos e correspondentes aos percentis 2,5% e 97,5% <i>versus</i> as vazões afluentes ao trecho fluvial estudado: Cenário C1.	199

Figura 5.31 – Resíduos ϵt e medianos por intervalo temporal obtidos nos Cenários A1 a C1 plotados junto ao hidrograma observado na estação fluviométrica de referência.	200
Figura 5.32 – Quantis empíricos de percentis notáveis colhidos por Δt nas séries de resíduos ϵt e os correspondentes quantis teóricos segundo as distribuições SEP e N padronizadas: Cenário C3.	203
Figura 5.33 – Histogramas de frequências relativas estimados para as amostras <i>a posteriori</i> dos parâmetros do modelo hidrodinâmico e da verossimilhança Generalizada: Cenário C3.	205
Figura 5.34 – Histogramas de frequências relativas estimados para as amostras <i>a posteriori</i> dos parâmetros do modelo hidrodinâmico e da verossimilhança Generalizada: Cenário C4.	206
Figura 5.35 – Quantis empíricos de percentis notáveis colhidos por Δt nas séries de resíduos ϵt e os correspondentes quantis teóricos segundo as distribuições SEP e N padronizadas: Cenário D3.	210
Figura 5.36 – Ajuste da distribuição Gama às amostras do parâmetro nc inferidas nos Cenários A3 (esq.) e C3 (dir.).	216
Figura 5.37 – Intervalos de credibilidade a 95% associados à incerteza da predição de vazões (Cenários A3 e C3) ou de níveis de água (Cenários B2 e D3) no extremo de jusante do trecho fluvial estudado.	220
Figura 5.38 – Amplitude dos intervalos de credibilidade a 95% e sua relação com os valores observados das variáveis vazão (Cenários A3 e C3) ou nível de água (Cenários B2 e D3).	221
Figura 5.39 – Inspeção da variação de níveis de água na seção do rio São Francisco em Pirapora-Barreiro acarretada pela variação de vazões no intervalo de credibilidade a 95% da incerteza preditiva paramétrica (setas e valores em cinza) e total (setas e valores em preto).	226
Figura 5.40 – Inspeção da variação de vazões na seção do rio São Francisco em Pirapora-Barreiro causada pela variação de níveis de água no intervalo de credibilidade a 95% da incerteza preditiva paramétrica (setas e valores em cinza) e total (setas e valores em preto).	227
Figura 5.41 – Intervalo de credibilidade a um nível de 95% para os perfis de escoamento ao longo do trecho estudado, considerando-se somente a incerteza preditiva paramétrica: Cenário C3.	230
Figura 5.42 – Intervalo de credibilidade a um nível de 95% para os perfis de escoamento ao longo do trecho estudado, considerando-se a incerteza preditiva total: Cenário C3.	230
Figura 5.43 – Mapa de inundação probabilístico que contabiliza somente a incerteza preditiva associada à rugosidade, sobre a porção de montante da cidade de Pirapora, próximo à ponte da BR-356 sobre o rio São Francisco: Cenário A3.	231
Figura 5.44 – Mapa de inundação probabilístico que contabiliza somente a incerteza preditiva associada à rugosidade, sobre a porção de montante da cidade de Pirapora, próximo à ponte da BR-356 sobre o rio São Francisco: Cenário A3.	232

LISTA DE TABELAS

Tabela 4.1 – Limites inferior e superior, e tipo estipulado para as distribuições de probabilidades marginais <i>a priori</i> das variáveis latentes do modelo de resíduos.	126
Tabela 5.1 – Características das estações fluviométricas utilizadas.	149
Tabela 5.2 – Distribuições <i>a priori</i> e limites inferior e superior atribuídos aos parâmetros do vetor θ	159
Tabela 5.3 – Principais descritores estatísticos dos resíduos brutos e transformados: Cenário A1.	175
Tabela 5.4 – Descritores estatísticos dos parâmetros do modelo hidrodinâmico e da função de verossimilhança Generalizada: Cenário A1.	178
Tabela 5.5 – Principais descritores estatísticos dos resíduos brutos e transformados: Cenário A3.	183
Tabela 5.6 – Descritores estatísticos dos parâmetros do modelo hidrodinâmico e da função de verossimilhança Generalizada: Cenário A3.	184
Tabela 5.7 – Principais descritores estatísticos dos resíduos brutos e transformados: Cenário A4.	184
Tabela 5.8 – Descritores estatísticos dos parâmetros do modelo hidrodinâmico e da função de verossimilhança Generalizada: Cenário A4.	185
Tabela 5.9 – Principais descritores estatísticos dos resíduos brutos e transformados: Cenário B2.	190
Tabela 5.10 – Principais descritores estatísticos dos resíduos brutos e transformados: Cenário B4.	190
Tabela 5.11 – Descritores estatísticos dos parâmetros do modelo hidrodinâmico e da função de verossimilhança Generalizada: Cenário B2.	195
Tabela 5.12 – Descritores estatísticos dos parâmetros do modelo hidrodinâmico e da função de verossimilhança Generalizada: Cenário B3.	195
Tabela 5.13 – Principais descritores estatísticos dos resíduos brutos e transformados: Cenário C2.	202
Tabela 5.14 – Principais descritores estatísticos dos resíduos brutos e transformados: Cenário C3.	202
Tabela 5.15 – Principais descritores estatísticos dos resíduos brutos e transformados: Cenário C4.	202
Tabela 5.16 – Descritores estatísticos dos parâmetros do modelo hidrodinâmico e da função de verossimilhança Generalizada: Cenário C3.	204
Tabela 5.17 – Descritores estatísticos dos parâmetros do modelo hidrodinâmico e da função de verossimilhança Generalizada: Cenário C4.	204
Tabela 5.18 – Principais descritores estatísticos dos resíduos brutos e transformados: Cenário D3.	209
Tabela 5.19 – Descritores estatísticos dos parâmetros do modelo hidrodinâmico e da função de verossimilhança Generalizada: Cenário D3.	211
Tabela 5.20 – Percentis notáveis utilizados para elaboração dos intervalos de credibilidade a 95% associados à incerteza preditiva: Cenários A3 (esq.) e C3 (dir.).	222
Tabela 5.21 – Percentis notáveis utilizados para elaboração dos intervalos de credibilidade a 95% associados à incerteza preditiva: Cenários B2 (esq.) e D3 (dir.).	222

LISTA DE ABREVIATURAS E SIGLAS

APFM	<i>Associated Programme on Flood Management</i>
BaTEA	<i>Bayesian Total Error Analysis</i>
BMA	<i>Bayesian Model Averaging</i>
CEPED-UFSC	Centro Universitário de Estudos e Pesquisas sobre Desastres da Universidade Federal de Santa Catarina
CN	<i>Curve Number</i>
DEFRA	<i>UK Department of the Environment, Food and Rural Affairs</i>
DREAM	<i>Differential Evolution Adaptive Metropolis</i>
FDP	Função densidade de probabilidade
FEMA	<i>Federal Emergency Management Agency</i>
FIRM	<i>Flood Insurance Rate Maps</i>
FOA	<i>First Order Approximation method</i>
GLUE	<i>Generalized Uncertainty Estimation</i>
GSA	<i>Generalized Sensitivity Analysis</i>
GWP	<i>Global Water Partnership</i>
HPD	<i>Highest Probability Density</i> , ou intervalo de mais alta densidade de probabilidade
HSY	<i>Hornberger-Spear-Young method</i>
IFM	<i>Integrated Flood Management</i>

LH	Método <i>Latin Hypercube</i>
MC	Método de Monte Carlo
MCMC	Método de Monte Carlo por Cadeias de Markov
MDT	Modelo Digital de Terreno
MDE	Modelo Digital de Elevação
MVS	Método da Máxima Verossimilhança
NFIP	<i>National Flood Insurance Program</i>
ONU	Organização das Nações Unidas
RB	Risco de Bayes
SEP	<i>Skew Exponential Power distribution</i>
SIG	Sistema de Informações Geográficas
TIN	<i>Triangular Irregular Network</i>
TR	Período de retorno de uma cheia ou evento chuvoso
UE	União Europeia
USACE	<i>US Army Corps of Engineers</i> (Corpo de Engenheiros do Exército Americano)
USWRC	<i>U.S. Water Resources Council</i>
WMO	<i>World Meteorological Organization</i>
WFD	<i>Water Framework Directive</i>

LISTA DE NOTAÇÕES E SÍMBOLOS

1D, 2D, 3D Número de dimensões em que as variáveis características dos escoamentos são consideradas em uma modelagem hidráulica/hidrodinâmica

a_t Erro aleatório i.i.d. calculado no tempo t

$a_{\xi,t}$ Resíduos a_t padronizados

A_t Área molhada total ocupada pela água em determinada seção transversal, i.e., dada pela soma das áreas efetivas de escoamento, denotadas por A , e das áreas de armazenamento lateral, denotadas por Y

A Área molhada efetiva de escoamento em determinada transversal

Y Área de armazenamento lateral em determinada transversal

A_c, A_f Área molhada correspondente à calha principal em determinada seção transversal, e área molhada efetiva de escoamento nas planícies de inundação

$\bar{A}$ Soma das áreas ativas de escoamento na calha e na planície, em termos médios entre duas seções transversais consecutivas

B Contorno ou fronteira de um sistema ambiental

Operador de retardo de um modelo autoregressivo

C Coeficiente de perda de carga associado a estruturas hidráulicas

c Velocidade de propagação de uma onda de cheia, ou celeridade

d Número de parâmetros que compõem o vetor θ

D Profundidade média da calha principal ao longo de um estirão fluvial

D_{50} Diâmetro mediano das partículas que compõem o leito fluvial

$\mathbf{E}$	Vetor de resíduos entre as saídas observadas e as simuladas para um sistema
ε	Componentes do vetor $\mathbf{E}$, geralmente indexados temporalmente
$\mathbb{E}$	Valor esperado de uma variável aleatória
E_r e $\tilde{E}_r$	Elevações da linha de água simulada e extraída a partir do cruzamento da mancha de inundação com o modelo de terreno de base para a definição das áreas inundadas
F	Indicadores de desempenho baseados na extensão de inundação
$\mathfrak{S}$	Notação para um modelo representativo de um sistema ambiental
g e G	Número total de simulações nas quais o ponto i foi inundado e número total de simulações, respectivamente
$\tilde{h}$ e h	Alturas totais de chuva observada e verdadeira para determinado evento pluviométrico, indexados por z
i	Indexador dos vetores $\mathbf{X}$ e $\tilde{\mathbf{X}}$ e dos respectivos resíduos
	Indexador dos m estados do vetor paramétrico $\boldsymbol{\theta}$ usados na construção de amostras <i>a posteriori</i>
j	Indexador dos vetores $\mathbf{Y}$ e $\tilde{\mathbf{Y}}$ e dos respectivos resíduos
	Indexador de estados em uma cadeia de Markov
J	Quantidade de conjuntos de parâmetros e variáveis latentes sorteados para compor a incerteza preditiva associada a $\tilde{y}_p$
K	Fator de forma proposto por Beven e Binley (1992)
K_c, K_f	Condutância da calha principal e das planícies de inundação, respectivamente
k	Número de <i>lags</i> de uma amostra, aferidos em termos temporais

L, U	Limites inferior e superior de um intervalo de credibilidade, respectivamente
L_F	Função de verossimilhança informal baseada em uma das medidas F de desempenho
L_{pond}	Função de verossimilhança informal baseada em uma ponderação entre a extensão de inundação e as profundidades ao longo do tempo na seção transversal de jusante de um trecho fluvial
L_W e L_E	Verossimilhanças informais avaliadas por meio das larguras de topo e das elevações da linha de água, respectivamente, contabilizadas nas seções transversais usadas em uma modelagem hidráulica
m	Número de realizações do vetor paramétrico θ em um experimento MC ou MCMC que sejam utilizadas para construção da distribuição conjunta <i>a posteriori</i> de θ
	Fator multiplicativo definido em função do grau de meandrização de um canal, para adoção do método de Cowan
$\mathbf{m}$ ou $\boldsymbol{\varphi}$	Vetor de multiplicadores de chuva, composto por N_m parâmetros, denotados individualmente por m_z
N	Número de cadeias de Markov envolvido em um processo MCMC
n	Quantidade de valores componentes de vetores como $\mathbf{X}$ e $\mathbf{Y}$
	Coefficiente de rugosidade de Manning
n_c	Coefficiente de rugosidade de Manning na calha principal
n_p	Coefficiente de rugosidade de Manning nas planícies de inundação
n_b	Valor de base do coeficiente de rugosidade de Manning para aplicação do método de Cowan
n_1 a n_4	Fatores de correção aditivos para aplicação do método de Cowan

p_i	Probabilidade de um ponto no modelo digital de elevação do tipo TIN - <i>Triangulated Irregular Network</i> ser inundado
p_j	Probabilidade de inundação calculada para cada célula j do <i>grid</i> de um MDT
Q	Vazão em trânsito em um estirão fluvial
Q_c	Vazão transportada pela calha principal
q_l	Contribuição lateral ao escoamento por unidade de comprimento
$\bar{Q}_l$	Vazão incremental, ou lateral, média no intervalo Δt
r	Indexador de seção transversal
R	número total de seções transversais utilizado em um modelo hidráulico
R_h	Raio hidráulico em determinada seção transversal
S	Vetor de uma variável de estado de um sistema ambiental
S_f	Taxa de perda de energia mecânica do escoamento devido ao atrito exercido pelas paredes e pelo fundo do canal
$\bar{S}_f$	Taxa média de perda de carga por atrito entre duas seções transversais consecutivas
$\bar{S}_h$	Taxa de perda de carga devido a contrações ao escoamento e alterações de velocidade induzidas por estruturas hidráulicas
S_{fc}, S_{ff}	Taxa de perda de energia mecânica do escoamento estimada na calha principal e nas planícies de inundação, respectivamente
S_o	Declividade média do fundo de um leito fluvial ou canal
$\tilde{S}$ e S	Área inundada contabilizada a partir da mancha de inundação disponível e computada em cada simulação do modelo, respectivamente

$S_{obs,sim}$	Interseção entre as áreas inundadas observada e simulada, $\tilde{S}$ e S
$S_{\tilde{A}obs}$	Área inundada prevista pelo modelo que não foi inundada segundo a observação disponível (superestimação)
$S_{\tilde{A}sim}$	Área inundada não prevista pela simulação (subestimação)
t	Indexador do tempo
T	Número de estados, ou iterações, em cada cadeia de Markov de um processo MCMC
T^*	Número de iterações, ou estados, necessários para que o último parâmetro alcance a convergência em um processo MCMC
T_r	Tempo de subida de uma onda de cheia
$\bar{U}_l$	Velocidade média do escoamento associada à vazão lateral, avaliada em termos médios no intervalo Δt
U	Velocidade média do escoamento em determinada seção transversal
U_c, U_f	Velocidade média do escoamento em determinada seção transversal, avaliada na calha principal e nas planícies de inundação, respectivamente
Var	Variância de uma variável aleatória
x_c, x_f	Distância entre seções transversais consecutivas medidas na calha principal e nas planícies de inundação, respectivamente
x_e	Distância equivalente entre seções transversais consecutivas, estimada a partir de uma ponderação por áreas entre as distâncias x_c e x_f
$\mathbf{X}$	Vetor representativo de entradas a um modelo explicativo de um sistema ambiental, formado por valores x

$\tilde{\mathbf{X}}$	Vetor de dados observados relativos às entradas de um sistema, formado por valores $\tilde{x}$
w	Quantidade aleatória qualquer
$w_{i,j}$	Fator de ponderação binário (0 ou 1) para cálculo de p_j , associado a uma simulação i
W_r e $\tilde{W}_r$	Larguras de topo em cada seção transversal r , simulada pelo modelo e extraída a partir da mancha de inundação observada, respectivamente
$\bar{y}$	Valor médio das observações da variável de saída, $\tilde{y}$
$\mathbf{Y}$	Vetor representativo de saídas a um modelo explicativo de um sistema ambiental, formado por valores y
$\tilde{\mathbf{Y}}$	Vetor de dados observados relativos às saídas de um sistema, formado por valores $\tilde{y}$
$\tilde{\mathbf{Y}}_p$	Vetor de variáveis preditivas, denotadas por $\tilde{y}_p$
$\tilde{y}_{1-\alpha}$ $1 - \alpha$	Percentil da variável preditiva $\tilde{y}_p$ associado à probabilidade de não excedência $1 - \alpha$
z	Indexador de evento pluviométrico
Z, z	Cota altimétrica do fundo do canal fluvial, medido na calha principal e nas planícies de inundação, respectivamente
α	Parâmetro utilizado na construção de intervalos de credibilidade, representativo da probabilidade de excedência
$\overline{\beta_v}$	Fator de distribuição de velocidades, no caso, avaliado em termos médios entre seções transversais consecutivas
β_x	Vetor de parâmetros do modelo de erros de entradas à modelagem, $\mathbf{X}$

β_y	Vetor de parâmetros do modelo de erros das saídas fornecidas pela modelagem,
$\mathbf{Y}$	
ξ, β	Parâmetros da distribuição de probabilidades SEP padronizada
κ	Fator de ponderação <i>theta</i> , associado ao esquema numérico implícito do tipo <i>four-point</i>
$\mu_\xi, \sigma_\xi, c_\beta$ e ω_β	Parâmetros calculados a partir de ξ e β para compor a FDP da distribuição SEP padronizada
λ	Fator de ponderação entre as condutâncias da calha principal e das planícies de inundação
Ψ_0	Condições de contorno iniciais
θ	Vetor de parâmetros de um modelo representativo de um sistema ambiental
Θ	Espaço paramétrico, que contém o vetor θ
Ω	Informação prévia acerca de θ
σ	Vetor de desvios-padrão, composto por n valores σ_t
$\hat{\sigma}$	Vetor de desvios-padrão estimados a partir de erros de medição de observações disponíveis para $\tilde{\mathbf{Y}}$, composto por n valores $\hat{\sigma}_t$
σ_0 e σ_1	Intercepto e coeficiente angular, respectivamente, que caracterizam um modelo para o desvio-padrão, σ_t
σ_m^2 ou $\sigma_{\mathbf{m}}^2$	Variância dos multiplicadores de chuva ou de vazão (variável ou vetor)
ϕ_p	Coeficiente(s) do polinômio do processo autoregressivo de ordem p
η_t	Processo aleatório tal que $\eta_t \sim \mathcal{N}(0, \sigma^2)$, sendo σ constante, nesse caso
$\boldsymbol{\eta}_\varepsilon$	Vetor de variáveis latentes ou de parâmetros do modelo de resíduos e da

função de verossimilhança

ρ Função de perda

δ Estimador de θ

Parâmetro interno do método DREAM que controla os ajustes da distribuição norteadora dos saltos

ϑ Fração de quantidade de movimento de uma contribuição lateral que contribuiu para o rio principal

SUMÁRIO

AGRADECIMENTOS	5
RESUMO	7
ABSTRACT	9
LISTA DE FIGURAS	10
LISTA DE TABELAS	13
LISTA DE ABREVIATURAS E SIGLAS	14
LISTA DE NOTAÇÕES E SÍMBOLOS	16
SUMÁRIO	24
1 INTRODUÇÃO	26
2 OBJETIVOS.....	39
2.1 OBJETIVO GERAL.....	39
2.2 OBJETIVOS ESPECÍFICOS	39
3 REVISÃO DA LITERATURA	40
3.1 INTRODUÇÃO AO ESTUDO DAS INCERTEZAS EM MODELAGEM DE SISTEMAS AMBIENTAIS	40
3.1.1 Definição de incerteza e seus tipos: epistêmica e aleatória.....	40
3.1.2 A incerteza em modelagem de sistemas ambientais e suas fontes	41
3.2 FONTES DE INCERTEZA NA MODELAGEM HIDRÁULICA DE VAZÕES DE CHEIA E NA ELABORAÇÃO DE MAPAS DE INUNDAÇÃO.....	44
3.3 MÉTODOS DE CUNHO BAYESIANO COM POTENCIAL DE APLICAÇÃO À QUANTIFICAÇÃO DE INCERTEZAS NA MODELAGEM HIDRODINÂMICA	51
3.3.1 Métodos Bayesianos formais	53
3.3.2 O método BaTEA.....	66
3.3.3 Métodos pseudo-Bayesianos.....	73
3.4 TÉCNICAS DE AMOSTRAGEM UTILIZADAS EM APOIO AOS MÉTODOS DE AVALIAÇÃO DE INCERTEZAS	79
3.5 APLICAÇÃO DE MÉTODOS DE AVALIAÇÃO DE INCERTEZAS NA MODELAGEM HIDRÁULICA DESTINADA AO ESTUDO DE CHEIAS	84
3.6 CARACTERIZAÇÃO DO MODELO HIDRODINÂMICO ADOTADO	99
3.6.1 Introdução.....	99
3.6.2 Caracterização do módulo hidrodinâmico do modelo HEC-RAS: equações adotadas e método numérico para sua resolução	100
3.6.3 Incertezas inerentes à utilização do módulo hidrodinâmico do modelo HEC-RAS	106
3.6.4 Breve revisão de métodos para definição do coeficiente de rugosidade de Manning em canais	111
4 METODOLOGIA	114
4.1 INTRODUÇÃO	114
4.2 ESQUEMA METODOLÓGICO ADOTADO PARA QUANTIFICAÇÃO DE INCERTEZAS.....	117
4.2.1 Prescrição da função de verossimilhança	120
4.2.2 Seleção de distribuições de probabilidades a priori para os parâmetros estudados	123
4.2.3 Seleção de dados para atualização da informação a priori.....	126
4.2.4 Técnica de amostragem selecionada para aplicação do método Bayesiano de quantificação das incertezas	128
4.2.5 Análise das incertezas e resíduos obtidos a posteriori	131
4.2.6 Cenários de calibração	139
4.3 SELEÇÃO DE TRECHOS FLUVIAIS PARA APLICAÇÃO DO MÉTODO PROPOSTO	140
4.3.1 Análise de dados hidrométricos	141
4.3.2 Definição dos coeficientes de rugosidade de Manning.....	142
4.4 SIMULAÇÕES HIDRÁULICAS INICIAIS	143
4.4.1 Condições de contorno	144
4.4.2 Discretização espacial e temporal	144
4.4.3 Aquecimento do modelo.....	145
4.4.4 Avaliação da estabilidade numérica e dos modelos fisicamente plausíveis.....	146

5	APLICAÇÃO DA METODOLOGIA.....	147
5.1	CARACTERIZAÇÃO DO TRECHO FLUVIAL SELECIONADO E INFORMAÇÕES DISPONÍVEIS.....	147
5.2	SIMULAÇÕES HIDRÁULICAS INICIAIS E DEFINIÇÃO FORMAL DO TRECHO FLUVIAL MODELADO	155
5.3	QUANTIFICAÇÃO DAS INCERTEZAS NA MODELAGEM HIDRODINÂMICA DESTINADA AO ESTUDO DE CHEIAS NO TRECHO FLUVIAL AVALIADO.....	158
5.3.1	<i>Distribuições de probabilidades a priori dos parâmetros do modelo hidrodinâmico.....</i>	<i>158</i>
5.3.2	<i>Seleção de eventos de cheia para estudo de incertezas e composição de mapas probabilísticos</i>	<i>159</i>
5.3.3	<i>Acoplagem do HEC-RAS ao DREAM</i>	<i>162</i>
5.3.4	<i>Cenários de calibração das incertezas.....</i>	<i>164</i>
5.3.5	<i>Aplicação do esquema Bayesiano à calibração de incertezas</i>	<i>167</i>
5.3.6	<i>Quantificação da incerteza preditiva</i>	<i>218</i>
6	CONCLUSÕES	234
7	RECOMENDAÇÕES E TRABALHOS FUTUROS.....	240
	REFERÊNCIAS	245

1 INTRODUÇÃO

A avaliação de enchentes e inundações, e de seus efeitos sobre as populações ribeirinhas, tem constituído objeto de estudo de engenheiros hidrólogos e hidráulicos em todo o mundo. Os modelos numéricos de simulação hidráulica e hidrológica têm avançado sobremaneira nas três últimas décadas, progresso este promovido em parte pela aceleração da capacidade de processamento computacional de dados e pela possibilidade de aquisição de bases topográficas precisas em muitas partes do globo, sobretudo nos Estados Unidos e em alguns países europeus e asiáticos. Apesar das melhorias no campo técnico e de aquisição de dados, e de esforços de muitos governos e agências mundiais, federais e locais na prevenção e controle de enchentes e inundações, estudos apontam que os danos associados às mesmas continuam aumentando (MERWADE *et al.*, 2008; SAHOO e SREEJA, 2015, BRUIJN *et al.*, 2009).

No Brasil, estima-se que as perdas econômicas causadas por inundações atinjam valores entre US\$ 1 bilhão e US\$ 2 bilhões por ano (MILOGRANA, 2009). Em relatório elaborado pelo Centro Universitário de Estudos e Pesquisas sobre Desastres da Universidade Federal de Santa Catarina (UFSC - CEPED, 2016), os danos causados por eventos hidrológicos, destacando-se as inundações, somaram, em todo o Brasil, cerca de R\$ 17 bilhões, no período de 1995 a 2014, o que indica um prejuízo médio anual de R\$ 3,6 bilhões. Desse total, 37% concentram-se na região Sudeste e 27% na região Sul. Embora seja mencionado um aumento do prejuízo a partir dos anos 2000, o que também se observa em outras bases de dados ao redor do mundo, o referido instituto recomenda cautela quanto ao uso dessa afirmação, visto que importantes lacunas de dados para realização da estimativa foram constatadas nos anos mais distantes.

Por sua vez, nos países da União Europeia, levantamentos dão conta de que os danos associados às inundações atingiram, até 1980, € 90 bilhões, sendo que são estimados prejuízos anuais da ordem de € 20 bilhões em 2020 e de € 50 bilhões em 2050 em todo o território da UE (<https://ec.europa.eu/environment/water/infographics.htm>). Na Índia, Sahoo e Sreeja (2015) indicam que o dano anual médio decorrente das inundações, avaliado ao longo de 53 anos, foi de cerca de US\$ 330 milhões. Os autores estimam que há uma tendência de aumento desses números, visto que nos últimos 10 anos da série, entre 1996 e 2005, o dano anual médio foi de aproximadamente US\$ 775 milhões.

Os estudos compilados relevam em comum a estimativa de que os danos causados a habitações, à infraestrutura e aos setores produtivo e de serviços tenderão a aumentar, aspecto creditado à

Programa de Pós-graduação em Saneamento, Meio Ambiente e Recursos Hídricos da UFMG

impermeabilização crescente nas áreas urbanas (SAHOO e SREEJA, 2015), à ausência de manejo no âmbito das bacias hidrográficas, ao inchaço inevitável das cidades (WMO, 2009), e, mais recentemente, também associado ao fenômeno das mudanças climáticas (BRUIJN *et al.*, 2009; ARCADIS, 2012; ONU, 2015).

Inevitavelmente, nos últimos anos, muitas pesquisas voltaram-se à avaliação dos impactos das mudanças climáticas sobre a frequência e a magnitude de fenômenos meteorológicos, com o possível agravamento das estiagens, enchentes, nevascas e furacões em distintos pontos do globo. Outros dois fatos constantemente associados à problemática das inundações e eventos climáticos extremos referem-se ao crescimento exacerbado da população do planeta, estimada em 9 bilhões de pessoas em 2050 (ONU, 2018), e à mudança do padrão de ocupação humana no território. Em 2007, 50% da humanidade já se concentrava em áreas urbanas, e, em 2030, projeções apontam que esse número saltará para 60% (ONU, 2016).

A pressão do crescimento populacional sobre os ecossistemas terrestres e aquáticos, aliada ao advento das mudanças climáticas, explica o fato de o chamado gerenciamento integrado de inundações compor a agenda de uma das 17 Metas para o Desenvolvimento Sustentável até 2030 propostas pela Organização das Nações Unidas (ONU, 2015), notadamente a de Comunidades e Cidades Sustentáveis. Ademais, em função dos inúmeros desafios relacionados aos recursos hídricos dessa agenda, a ONU propôs que a década 2018-2028 seja a de ações e programas pela “Água para o Desenvolvimento Sustentável” (ONU, 2016), uma série de iniciativas em escala local e global para melhor aproveitamento, gerenciamento e disposição da água, em meio às quais se encontra o gerenciamento integrado das inundações. Nesse âmbito específico, a ONU destaca o papel do *Associated Programme on Flood Management* (APFM), desenvolvido pela *Global Water Partnership* (GWP) e a *World Meteorological Organization* (WMO), como uma plataforma de apoio técnico à implantação do gerenciamento integrado de inundações (IFM – *Integrated Flood Management*) em diversos países.

A WMO (2009) faz referência ao gerenciamento integrado de inundações como sendo um esquema de medidas estruturais e não estruturais destinadas ao controle de cheias e à mitigação de seus efeitos com vistas à maximização dos benefícios proporcionados pela planície de inundação, e à minimização das perdas de vidas humanas decorrentes das enchentes. Nesse contexto, atividades como a previsão mais precisa de áreas atingidas e a disseminação de informações pertinentes para a defesa civil, órgãos de planejamento e público em geral são

ações imprescindíveis à mitigação e à atenuação dos efeitos causados pelas inundações sobre as populações ribeirinhas em áreas urbanas e rurais, assim como sobre a economia local (MERWADE *et al.*, 2008).

Previamente ao enfoque dado pela ONU à questão das inundações, outras iniciativas já haviam sido tomadas na esfera política frente aos impactos das mudanças climáticas e à pressão do crescimento urbano sobre os ambientes aquáticos, fomentada em parte pela ocorrência de inundações de magnitude importante em algumas partes do globo, nos anos 1993-2008 (HALL e SOLOMATINE, 2008). No caso da União Europeia, por exemplo, é de 2000 a WFD – *Water Framework Directive* (EU, 2000), que estabelece o arcabouço legal sob a forma de conceitos, metas e diretrizes para alcance do uso sustentável dos corpos de água dentro do território comum aos países que a compõem, considerando-se o gerenciamento integrado dos recursos hídricos na escala de bacias hidrográficas. Na esteira da WFD, surgiu a *Floods Directive*, em 2007 (EU, 2007), que, utilizando conceitos e critérios extraídos da referida legislação-mãe, especifica e organiza as etapas para implantação de planos de gerenciamento de risco contra as inundações em todos os países da UE, com metas a serem atingidas até 2015, ano após o qual foram definidos ciclos de avaliação e revalidação dos planos concretizados.

Por sua vez, nos Estados Unidos, foi instituído pelo Congresso Nacional em 1968 o *National Flood Insurance Program* (NFIP), destinado a reduzir o impacto das inundações sobre a propriedade pública e privada, por meio da promoção do acesso a seguros contra enchentes e do estímulo à adoção de medidas para gerenciamento da planície de inundação e dos riscos associados às cheias (<https://www.fema.gov/national-flood-insurance-program>). Nesse sentido, são de incumbência da FEMA (*Federal Emergency Management Agency*) a manutenção, a atualização e a modernização dos *flood insurance rate maps* – FIRMs (MERWADE *et al.*, 2008), que consistem em mapas com o zoneamento da planície de inundação frente às cheias com 100 e 500 anos de período de retorno para milhares de comunidades e cidades ribeirinhas no país.

Permeando esse contexto, o setor de recursos hídricos vem experimentando uma série de avanços e melhorias dos modelos hidrológicos e hidráulicos utilizados e dos correspondentes dados de entrada, como informações topobatimétricas, séries temporais de precipitação, de nível de água e de vazão. Esse progresso, aliado à crescente necessidade de compreender, controlar e mitigar as inundações e seus efeitos, tem aberto novas frentes de pesquisa e trabalho

para a comunidade técnica e científica atuante no tema, dentre as quais, são de interesse para a presente discussão:

- O reconhecimento da equifinalidade, ou a possibilidade de obtenção de saídas dos modelos com qualidades de ajustamento comparáveis, a partir de diferentes e, às vezes, numerosas combinações possíveis entre os parâmetros que os governam (BEVEN, 2006), bem como da necessidade de propor alternativas para superá-la (TODINI, 2007), como a análise probabilística;
- A definição dos graus de refinamento e de complexidade necessários a um modelo para se ter uma resposta física razoável ou aceitável, do ponto de vista do realismo hidrológico e hidráulico (DI BALDASSARRE *et al.*, 2010), e que atendam às restrições financeiras e temporais e aos objetivos de cada estudo ou projeto;
- A necessidade de quantificação das incertezas, ou erros, associadas aos dados de entrada e de saída, aos parâmetros utilizados nas modelagens e aos próprios modelos (MERWADE *et al.*, 2008);
- O acesso quantitativo à forma como essas informações influenciam umas às outras e como cada uma contribui para a incerteza final (JUNG e MERWADE, 2012; 2015);
- A quantificação da incerteza preditiva, ou seja, dos erros envolvidos na predição de variáveis de saída estimadas pelos modelos (TODINI, 2007); e, finalmente,
- A atribuição de probabilidades aos resultados individuais fornecidos pelos modelos, em substituição à abordagem determinística, comumente adotada (MERWADE *et al.*, 2008; DI BALDASSARRE *et al.*, 2010).

Beven (2007) menciona que os recursos computacionais atuais e futuros permitirão a elaboração de modelos integrados de recursos hídricos em diversas escalas espaciais e temporais, transpondo as fronteiras institucionais e políticas. No entanto, o mesmo autor reconhece que o uso adequado de tais modelos em direção a predições e previsões mais precisas e confiáveis demandará o aprofundamento do conhecimento acerca das limitações e erros inerentes à representação por eles fornecida, o que só poderá ser alcançado por meio da inserção e da quantificação das incertezas em quaisquer estimativas realizadas com o uso da modelagem.

Segundo Camacho *et al.* (2015), estudos ambientais que apresentem e quantifiquem as incertezas envolvidas são cada vez mais relevantes para identificar as limitações dos modelos

utilizados e para fornecer, de maneira mais compreensível aos tomadores de decisão e ao público, como os resultados da modelagem ambiental são avaliados e usados para produzir a informação final de interesse.

No que tange ao uso da incerteza em engenharia de recursos hídricos, Krzysztofowicz (1999) defende que qualquer tomada de decisão racional nesse campo, a exemplo da operação de reservatórios e hidrovias e de sistemas de alerta contra inundações, deva ser respaldada na quantificação da incerteza total de variáveis hidrológicas e hidráulicas de interesse fornecidas pelos modelos numéricos. O mesmo autor sugere, ainda, que essa incerteza deva ser prescrita sob a forma de uma distribuição de probabilidades, condicionada ao conhecimento de *experts* e a informações disponíveis sobre o fenômeno modelado.

Ao longo da pesquisa de doutoramento ora apresentada, trabalhou-se com métodos e alternativas sequenciados de modo a responder, direta ou indiretamente, às questões dos excertos anteriormente colocados, focadas sobre a modelagem hidrodinâmica, com o objetivo final de gerar produtos de interpretação probabilística, como perfis de escoamento, hidrogramas defluentes e manchas de inundação. Seguindo o roteiro metodológico proposto, esperava-se, adicionalmente, melhor compreender o modelo matemático selecionado para propagação de cheias e os dados utilizados para essa finalidade. Para tanto, procedeu-se a uma ampla revisão de literatura pertinente ao tema, a fim de garantir uma seleção judiciosa dos procedimentos mais adequados aos dados de entrada (i.e., topografia, níveis de água e vazões) disponíveis em bacias localizadas no Brasil, ao modelo matemático selecionado e a seus parâmetros, e aos propósitos supracitados, bem como de possibilitar a adaptação de materiais e métodos à realidade das informações existentes.

Percebeu-se, durante a revisão bibliográfica, que a quantificação da incerteza associada a modelos de sistemas ambientais, dentre eles, aos modelos hidrodinâmicos, está intimamente ligada à teoria de probabilidades, de forma que os métodos abordados e adotados se concentraram naturalmente em um grupo de ferramentas de cunho probabilístico. Além disso, verificou-se que, dentre estas, as de abordagem Bayesiana apresentavam potencial de aplicação a situações em que há disponibilidade de informações sobre o fenômeno modelado, a exemplo de registros de monitoramento fluviométrico ou marcas de cheia e manchas de inundação. Tais dados prestam-se à atualização da incerteza estipulada *a priori*, por meio da utilização do Teorema de Bayes (VRUGT, 2016). O estudo de incertezas por meio de técnicas Bayesianas

torna-se, assim, um problema de calibração condicionado à informação existente (STEDINGER *et al.*, 2008; VRUGT, 2016). O avanço em relação à calibração tradicional de modelos, em que se buscam um ou alguns conjuntos paramétricos que otimizem determinada função-objetivo, é a obtenção de distribuições de probabilidades que caracterizem a incerteza, ou erro, inerente aos parâmetros do modelo. O desafio principal reside na prescrição de uma função de verossimilhança, uma expressão equivalente à função-objetivo, que seja fiel ao comportamento estatístico esperado para os resíduos da modelagem (SCHOUPS e VRUGT, 2010; HUTTON *et al.*, 2011), e a posterior checagem das premissas adotadas para os mesmos. Ademais, diversos estudos recentes se propuseram a expandir o estudo da incerteza dita paramétrica em direção à quantificação, via estrutura Bayesiana, das incertezas associadas a dados de entrada e de saída da modelagem (KAVETSKI *et al.*, 2003; 2006a; 2006b; VRUGT *et al.*, 2008; SCHOUPS e VRUGT, 2010; RENARD *et al.*, 2010; EVIN *et al.*, 2013).

Em grande compilação retrospectiva sobre os avanços e desafios da modelagem hidrológico-hidráulica destinada ao mapeamento de inundações, Merwade *et al.* (2008) defenderam que a quantificação das incertezas constituía um dos pontos centrais a ser estudado com urgência nos próximos anos pela comunidade de engenheiros de recursos hídricos. Passados 10 anos desse compêndio, verifica-se abundância de estudos nesse campo concentrados na síntese chuva-vazão (KAVETSKI *et al.*, 2003; 2006a; 2006b; STEDINGER *et al.*, 2008; VRUGT *et al.*, 2008; SILVA *et al.*, 2014), suprida por inúmeros exemplos desenvolvidos nos anos anteriores (vide, por exemplo, a compilação de BEVEN e BINLEY, 2014). A modelagem hidráulica, permanente ou não permanente, pelo contrário, tem apresentado número modesto de exemplos (CAMACHO *et al.*, 2015), agregados ao redor de alguns métodos de ampla utilização, como o GLUE – *Generalized Likelihood Uncertainty Estimation* (BEVEN e BINLEY, 1992), que, apesar da essência Bayesiana, constitui alvo de críticas devido principalmente à generalização das hipóteses acerca do comportamento estipulado para os resíduos da modelagem (KAVETSKI *et al.*, 2003; STEDINGER *et al.*, 2008; VRUGT *et al.*, 2008). Na pesquisa ora relatada, pretendeu-se explorar a viabilidade de utilização de técnicas Bayesianas distintas do GLUE para calibração da incerteza associada a variáveis e parâmetros da modelagem hidrodinâmica, em busca de sua melhor identificação. Além disso, buscou-se, por meio de indicadores e de gráficos, confirmar as hipóteses inicialmente tomadas acerca do comportamento estatístico dos resíduos da modelagem hidrodinâmica, questão não raro desconsiderada em estudos desse tipo (HUTTON *et al.*, 2011).

Em se tratando da modelagem hidrodinâmica destinada à obtenção de hidrogramas e cotogramas de cheia e de manchas de inundação, as incertezas devem-se a uma série de aspectos inerentes à estrutura e à parametrização dos modelos, à obtenção e ao formato dos dados de entrada e até mesmo às rotinas de sistematização de resultados. Dentre tais fatores, mencionam-se, a seguir, alguns de destaque:

- Vazão ou hidrograma de projeto ou de referência;
- Dados topográficos e batimétricos de alimentação dos modelos hidráulicos e seu tratamento para inserção na modelagem e para composição de resultados;
- Modelagem hidráulica: escalas espacial e temporal, parametrização, hipóteses, premissas e métodos adotados para resolução das equações que compõem sua estrutura; e
- Técnicas e algoritmos usados na geração de manchas de inundação, a partir da extração dos dados de saída do modelo hidráulico.

Jung e Merwade (2012; 2015) apontam diversas razões para se estimarem as incertezas devidas às variáveis e a informações envolvidas na modelagem hidrodinâmica, de forma individual e global, como a possibilidade de compreensão das variáveis ou dados que demandam maior aprimoramento em termos de aquisição, cálculo ou tratamento (a exemplo de dados topográficos e equações regionais de obtenção de quantis de cheias), a fim de reduzir seu peso na incerteza final.

Ademais, quando as incertezas são propagadas pela modelagem hidráulica (o que pode ser feito também acoplado à modelagem hidrológica), com o objetivo de se obter o resultado conjunto das mesmas sobre, por exemplo, a elaboração de manchas de inundação, podem ser atribuídos contornos probabilísticos à área atingida para um mesmo evento pluviométrico sobre a bacia contribuinte. Assim, dada uma determinada cheia, associada a um período de retorno TR, haverá regiões da planície de inundação com maiores ou menores chances de serem alcançadas do que outras, devido à presença de incertezas nos dados de entrada, parâmetros, equações e métodos que alimentam os modelos de previsão, simulação e propagação de cheias. As manchas de inundação assim traçadas, de cunho probabilístico (PAPPENBERGER *et al.*, 2005; 2006a; DI BALDASSARRE *et al.*, 2010; BEVEN *et al.*, 2011; JUNG e MERWADE, 2012), podem ser usadas para se priorizarem ações que reduzam o risco às inundações, o que seria impossível

caso houvesse somente um único contorno para a mesma cheia com caráter completamente determinístico (PAPPENBERGER *et al.*, 2005).

A abordagem probabilística aplicada sobre variáveis de saída da modelagem, a exemplo de manchas de inundação no caso dos modelos hidrodinâmicos, pode embutir em seu bojo, a depender dos procedimentos e técnicas adotados, a quantificação do reflexo das incertezas individuais de variáveis e parâmetros sobre o resultado final e a interferência relativa de umas sobre as outras (i.e., sensibilidade). Dessa forma, constitui ferramenta que permite que se compreendam, para as áreas ou trechos fluviais estudados, quais aspectos intervenientes têm maior peso sobre a definição de áreas atingidas e a interação entre os mesmos, e que se visualize o efeito isolado ou combinado das incertezas no produto final dos estudos (JUNG e MERWADE, 2012; 2015).

Além das aplicações mencionadas nos parágrafos anteriores, vislumbram-se outras possibilidades oriundas da quantificação das incertezas associadas a variáveis e parâmetros envolvidos na modelagem hidrodinâmica destinada ao estudo de cheias, bem como da incerteza final presente em variáveis de saída previstas por tais modelos. Compreender e quantificar a incerteza, ou erro final, sobre *outputs* como perfis de escoamento, hidrogramas, cotogramas, manchas de inundação, velocidades na planície de inundação e na calha principal, tempo de chegada do pico da onda de cheia e tempo de inundação, podem fornecer subsídios adicionais para atividades relacionadas ao gerenciamento do risco associado às inundações, tais como:

- Estimação de um intervalo, ou mesmo de uma distribuição de probabilidades, para o dano médio anual decorrente de inundações associadas a distintos períodos de retorno, em determinada bacia hidrográfica e/ou estirão fluvial, em face da incerteza total estimada em variáveis de saída (HALL e SOLOMATINE, 2008), como a velocidade máxima e a profundidade na planície de inundação e o tempo de inundação. Esse tipo de análise pode incorporar inclusive incertezas relacionadas aos métodos e parâmetros para quantificar os danos;
- Estimação, de forma semelhante à supracitada, de impactos socioambientais e de índices de vulnerabilidade associados às inundações (vide, por exemplo, ARCADIS, 2012, e SAHOO e SREEJA, 2015), de forma quantitativa ou semi-quantitativa;

- Cálculo do valor esperado de funções de utilidade (a exemplo do custo médio de uma decisão como não proceder a um alerta de inundação), obtido a partir da integração do produto de uma função de custo associado à decisão em questão pela distribuição de probabilidades de uma variável de saída de interesse (e.g., profundidade), conforme sugerido por Todini (2008). Segundo esse autor, a associação da incerteza da modelagem hidrodinâmica a funções que norteiem a tomada de decisão no gerenciamento de riscos das inundações pode reduzir a possibilidade de alarmes falsos ou de alarmes perdidos, ou seja, de decisões incorretas, em relação à situação em que são usadas previsões determinísticas fornecidas por esses modelos;
- Diferenciação das taxas de seguros contra enchentes, respaldada por mapas probabilísticos elaborados para uma cheia de referência, como aquela com 100 anos de período de retorno (MERWADE *et al.*, 2008; JUNG e MERWADE, 2012); e
- Estudo mais aprofundado e adoção de medidas estruturais e não estruturais para controle de cheias e mitigação de seus impactos que sejam robustas perante as incertezas consideradas (HALL e SOLOMATINE, 2008), ou seja, que ofereçam uma relação custo *versus* benefício (ou dano evitado) que seja pouco ou menos sensível à variabilidade inerente às saídas do modelo que respaldam seu cálculo, quando comparadas a outras alternativas.

Ampliando o horizonte de possibilidades vislumbradas pela análise de incertezas aplicada ao estudo das enchentes, Hall e Solomatine (2008) mencionam que o tema vai de encontro às necessidades do moderno e sustentável gerenciamento do risco associado às inundações, ao qual os governos e tomadores de decisão precisam responder com um portfólio variado de medidas estruturais e não estruturais. Para orientar estudos e pesquisas relacionados à tomada de decisão quanto às estratégias, planos e ao dimensionamento e à estruturação de medidas no contexto do gerenciamento integrado de riscos das inundações, Hall e Solomatine (2008) propuseram um sistema para incorporar a esse processo os principais aspectos da análise de incertezas.

Esses autores explicam que as incertezas têm sido consideradas de alguma forma desde os primórdios da engenharia de defesa contra inundações, ainda que sob a forma de coeficientes de majoração de resultados, como a borda livre adotada sobre níveis de água alcançados por cheias notáveis, ou com a adoção de equações conservadoras para dimensionamento de estruturas hidráulicas. A introdução da noção de risco, como produto entre impactos e

probabilidades, ao gerenciamento de inundações pavimentou o caminho para que as incertezas fossem tratadas de forma mais explícita.

Ademais, como a tomada de decisões quanto ao risco de inundações vem adquirindo maior complexidade, à medida que novos aspectos podem ser considerados (variações de longo termo no regime hidrológico das bacias hidrográficas, mudanças no regime das marés e elevação do nível médio dos mares, alterações de uso e ocupação do solo, efeitos de escala temporal e espacial na modelagem de processos ambientais), mais variáveis agregam incertezas na definição do risco, dos danos e custos associados às inundações e às possíveis medidas para sua mitigação e controle. A proposta de Hall e Solomatine (2008) fornece um resumo das potencialidades que a análise de incertezas pode oferecer e agregar ao gerenciamento de riscos associados às enchentes, e engloba as seguintes etapas, diagramadas na Figura 1.1 e resumidas na sequência:

- i. Definição dos objetivos e do escopo da análise de incertezas, em função das decisões às quais os resultados irão se prestar e das possíveis variáveis envolvidas no processo;
- ii. Identificação, definição e descrição de incertezas, a partir da listagem dos processos e fatores que podem influenciar as variáveis usadas para tomada de decisão;
- iii. Coleta de informações que auxiliem a quantificação das incertezas definidas previamente. Tal é o caso de observações e medidas de variáveis de interesse na bacia hidrográfica a ser estudada, sendo elas dados de entrada ou de saída dos modelos numéricos a serem usados, ou mesmo resultados de estudos anteriores que possam servir como referência;
- iv. Construção de funções matemáticas para quantificar as incertezas, a partir das informações coletadas na etapa anterior e da *expertise* de profissionais envolvidos na modelagem. Tais funções podem ser: (a) distribuições de probabilidades, (b) conjuntos de distribuições, ou (c) limites superior e inferior;
- v. Propagação das incertezas sobre as variáveis de saída e indicadores de interesse, como área inundada, danos, impactos socioambientais e o próprio risco associado a uma determinada inundação. Nessa etapa, deve ser usado algum método para se propagar e quantificar a variabilidade (i.e., erro) imposta aos *outputs* pelas incertezas definidas anteriormente;

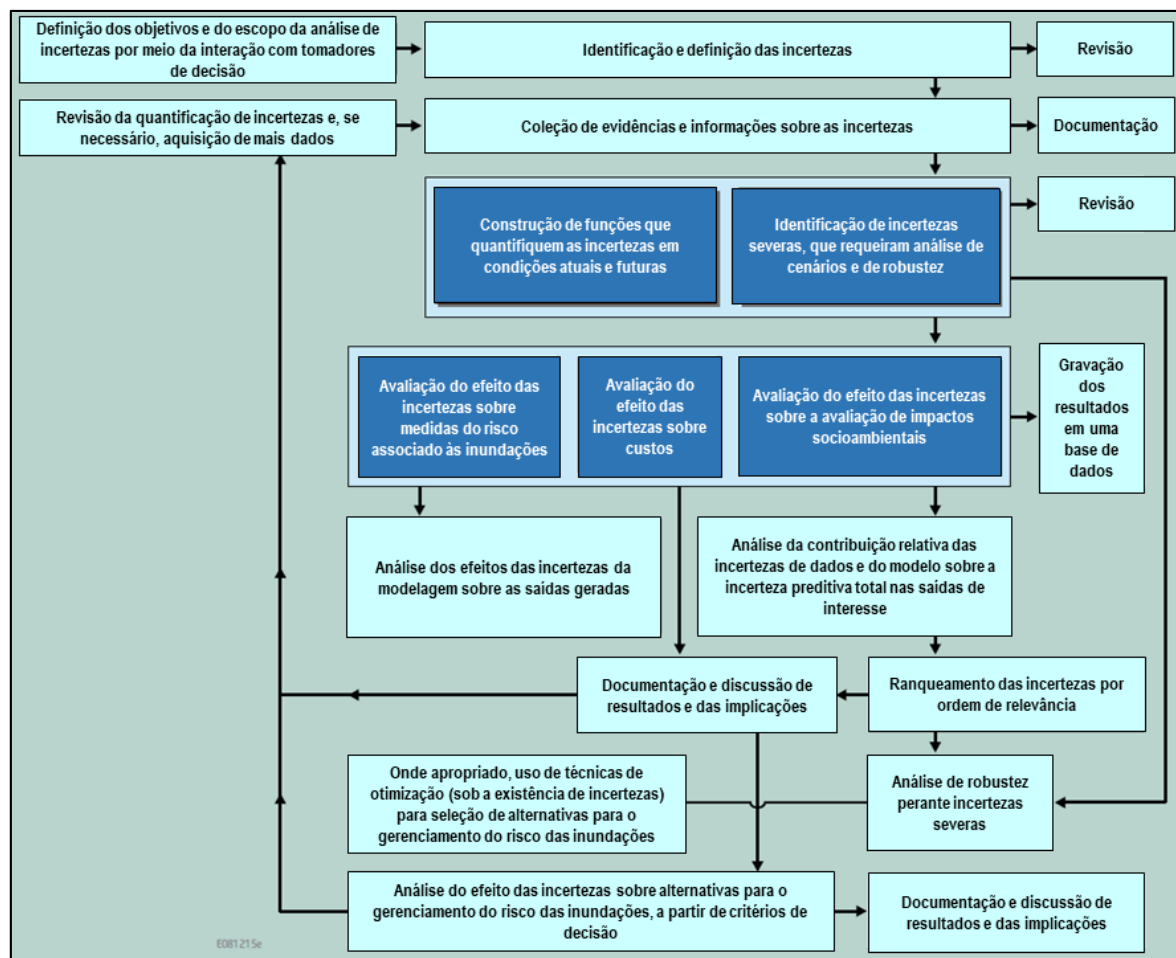


Figura 1.1 – Estrutura de análise de incertezas com vistas à sua aplicação no gerenciamento integrado de riscos das inundações, proposta por Hall e Solomatine (2008).

Fonte: tradução realizada a partir de figura extraída de Hall e Solomatine (2008).

- vi. Avaliação de sensibilidade dos *outputs* em função da incerteza nas variáveis e parâmetros considerados nas etapas anteriores. Existe uma série de métodos para se aferir a contribuição relativa de uma variável ou de conjuntos de variáveis, bem como de sua incerteza, sobre as saídas de interesse e, assim, de sua incerteza final, processo que auxilia no ranqueamento daquelas que devem ser finalmente utilizadas em determinada modelagem. Os resultados dessa etapa podem ser usados para refinar a fase anterior; e
- vii. Avaliação do efeito das incertezas sobre as alternativas (i.e., medidas estruturais e não estruturais) elencadas na tomada de decisão quanto ao gerenciamento do risco de inundações.

Cada uma dessas etapas deve ser documentada e revisada iterativa e permanentemente, à medida que novas informações sobre enchentes em certa bacia hidrográfica forem coletadas e que novas alternativas para controle e mitigação de seus efeitos emerjam como possíveis soluções locais ou regionais.

A inserção do estudo de incertezas no contexto mais amplo do gerenciamento de riscos impostos pelas enchentes trata-se de um processo complexo, multidisciplinar e que requer o envolvimento da defesa civil, dos setores usuários da água (indústria, agronegócio, comunidades rurais e urbanas, produção de energia hidrelétrica, empresas e autarquias de saneamento básico), e de instituições e empresas ligadas à gestão e ao gerenciamento das águas (agências e comitês de bacia, em se tratando do Brasil, por exemplo). Embora pareçam avançados para a realidade brasileira, propostas e esquemas como o de Hall e Solomatine (2008) fornecem uma perspectiva ampliada em que as incertezas sejam consideradas, permitindo que se vislumbrem aplicações a médio e longo prazo, inclusive em outras áreas ligadas ao setor de recursos hídricos, como navegação e abastecimento público.

Nos Estados Unidos e no Reino Unido, órgãos governamentais que estabelecem diretrizes para o gerenciamento do risco impostos pelas enchentes, tais como o USWRC – *U.S. Water Resources Council* e o DEFRA – *UK Department of the Environment, Food and Rural Affairs*, defendem que os tomadores de decisão, no que tange ao estabelecimento de medidas para mitigação e redução do risco às inundações, devem ter acesso às incertezas envolvidas nos estudos de definição dos riscos, benefícios, danos e impactos (HALL e SOLOMATINE, 2008), a fim de compreender a confiabilidade dos métodos e dados utilizados e das propostas indicadas, o que pode interferir no processo decisório quanto ao elenco e ao ranqueamento de opções a serem adotadas.

O tema da quantificação de incertezas associadas à modelagem hidráulico-hidrológica, e, mais especificamente, à modelagem hidrodinâmica, é desafiador devido à escassez de exemplos na vertente metodológica explorada neste trabalho, e a limitações de dados que permitam a calibração e a validação dos modelos probabilísticos adotados. Embora haja iniciativas que propuseram um esquema de avaliação de incertezas direcionado à realidade de cada estudo, finalidade e objetivo de uma análise desse tipo (PAPPENBERGER *et al.*, 2006b; BEVEN *et al.*, 2011), não há ainda um consenso quanto aos métodos que devem ser adotados, às variáveis que devem ser analisadas e aos produtos gerados, além da forma como a probabilidade deve ser

comunicada aos usuários finais da informação (BEVEN, 2007). O supracitado autor sugere cautela na utilização dos produtos gerados a partir de análises de incertezas e na generalização e extrapolação de resultados obtidos em determinada bacia hidrográfica para outros locais.

Tendo em vista esse panorama, a pesquisa ora relatada não pretende esgotar o tema, e tampouco explorar todos os desdobramentos que emergem da análise de incertezas, tais como aqueles sugeridos por Hall e Solomatine (2008), Todini (2007; 2008) e Arcadis (2012), mas lançar luz sobre elementos metodológicos que auxiliem o estabelecimento de diretrizes para realização de estudos futuros, sob as hipóteses de existência de dados observados sobre as inundações, como hidrogramas, cotogramas e marcas de cheia, e de utilização de distribuições de probabilidades para modelar as incertezas envolvidas na modelagem hidrodinâmica. A metodologia investigada neste trabalho dispõe-se a quantificar as incertezas inerentes à obtenção de perfis de escoamento e de hidrogramas e cotogramas de cheias de magnitude relevante, constituindo apenas uma parte de uma estrutura para gerenciamento dos riscos associados a inundações. Adicionalmente, os resultados obtidos mostram que o arcabouço metodológico aplicado pode também se prestar à compreensão do modelo hidrodinâmico utilizado, de suas limitações e capacidades, bem como do conteúdo das informações topobatimétricas e de monitoramento fluviométrico adotadas.

O corpo da presente tese encontra-se dividido em 6 capítulos, além deste de introdução. No Capítulo 2, são mencionados os objetivos geral e específicos. No Capítulo 3, é apresentada a revisão bibliográfica, a qual contém seis tópicos principais: (i) introdução ao estudo de incertezas na modelagem de sistemas ambientais; (ii) fontes de incertezas na modelagem hidráulica de cheias e na elaboração de mapas de inundação; (iii) métodos de cunho Bayesiano com potencial de aplicação na modelagem hidrodinâmica; (iv) técnicas de amostragem; (v) aplicação de métodos de avaliação de incertezas na modelagem hidráulica destinada ao estudo de cheias; e (vi) caracterização do modelo hidrodinâmico adotado. No Capítulo 4, são descritas as etapas metodológicas adotadas ao longo do trabalho. No Capítulo 5, é feita uma caracterização do trecho fluvial estudado, bem como são apresentadas as particularidades de aplicação da metodologia à presente situação e discutidos os resultados obtidos. Finalmente, no Capítulo 6 encontram-se as conclusões da pesquisa, e no Capítulo 7 são feitas recomendações para nortear eventuais trabalhos futuros concernentes ao tema.

2 OBJETIVOS

2.1 *Objetivo geral*

Esta tese de doutoramento tem por objetivo geral a avaliação, a seleção e a adequação de uma metodologia de cunho Bayesiano para avaliar e quantificar as incertezas associadas a variáveis, dados e parâmetros envolvidos na modelagem hidrodinâmica destinada ao estudo de cheias, com vistas ao desenvolvimento de possíveis aplicações, como a elaboração de mapas probabilísticos. Para confirmação da aplicabilidade da metodologia proposta, estipulou-se estudar pelo menos um trecho fluvial no Brasil.

2.2 *Objetivos específicos*

A fim de atingir o objetivo geral mencionado, esta pesquisa de doutoramento prevê os seguintes objetivos específicos:

- Pesquisar os diversos métodos utilizados para avaliação de incertezas aplicáveis à modelagem hidrodinâmica, com foco na abordagem Bayesiana, e propor um esquema metodológico adequado à realidade de informações comumente disponíveis no país;
- Definir os dados de entrada, variáveis e parâmetros da modelagem hidrodinâmica passíveis de terem suas incertezas ou erros quantificados pelo roteiro metodológico proposto;
- Avaliar o efeito da utilização de distintos tipos de informação disponíveis do fenômeno estudado sobre a quantificação das incertezas de variáveis e parâmetros;
- Proceder à verificação das hipóteses adotadas quanto aos resíduos gerados na modelagem hidrodinâmica, ou seja, avaliar se os mesmos corroboram o comportamento estatístico que embasa a aplicação de uma metodologia de abordagem Bayesiana;
- Quantificar e ilustrar o efeito da variabilidade (i.e., incerteza) dos referidos dados, variáveis e parâmetros sobre o resultado final, ou saídas do modelo, de forma individual e conjunta;
- Obter manchas de inundação e perfis de escoamento de cunho probabilístico para o(s) trecho(s) fluvial(is) selecionado(s); e
- Elaborar diretrizes para condução de estudos futuros de avaliação de incertezas sobre variáveis de saída associadas à modelagem hidrodinâmica, à luz da realidade brasileira em termos de aquisição de informações e de tomada de decisão quanto ao gerenciamento do risco de inundações.

3 REVISÃO DA LITERATURA

A revisão de literatura foi organizada de modo a fornecer um panorama acerca do conceito, das fontes e tipos de incertezas em modelagem matemática de sistemas ambientais, dos métodos disponíveis mais adotados para tratá-las quantitativamente. Na sequência, procede-se a uma compilação e ao aprofundamento de estudos nesse campo voltados à modelagem hidráulica de cheias, em regime permanente ou não permanente e seus desdobramentos, como a obtenção de manchas de inundação de cunho probabilístico.

3.1 ***Introdução ao estudo das incertezas em modelagem de sistemas ambientais***

3.1.1 **Definição de incerteza e seus tipos: epistêmica e aleatória**

A incerteza pode ser genericamente definida como o estado no qual não se tem conhecimento exato para se descreverem os componentes de um determinado sistema (HUTTON *et al.*, 2011). Ayyub e Klir (2006) ponderam que a incerteza constitui importante dimensão a ser considerada em análise de risco, uma vez que pode estar presente na definição de ameaças e modos de falha, dos danos às propriedades e à infraestrutura, dos impactos e suas magnitudes, bem como nas estratégias para mitigação do risco e nos critérios para análise de decisão.

A incerteza é comumente compartimentada em duas categorias (HALL, 2003; AYYUB e KLIR, 2006):

- Incerteza aleatória: também referida como variabilidade ou incerteza inerente (HALL, 2003), refere-se à variabilidade intrínseca de populações conhecidas, nas quais as observações ou medições podem ser aproximadas por distribuições de probabilidades (AYYUB e KLIR, 2006; HUTTON *et al.*, 2011) a fim de representá-la. Devido à sua natureza, comum aos fenômenos naturais, essa incerteza é de difícil redução. Um exemplo seria a variabilidade espacial e temporal das precipitações (HUTTON *et al.*, 2011) e das vazões em determinada seção fluvial. Ayyub e Klir (2006) mencionam como variáveis dotadas de incerteza aleatória as propriedades de resistência de materiais como o concreto e o aço, e os carregamentos sobre uma plataforma de exploração petrolífera causados pelas ondas do mar.

- Incerteza epistêmica: decorre do conhecimento limitado e incompleto de um sistema em estudo, e da inability de se compreendê-lo e descrevê-lo. É o caso das incertezas decorrentes de qualquer modelo que represente a realidade, dadas as inúmeras simplificações e hipóteses ao se partir do mundo real em direção a equações, métodos numéricos e modelagem reduzida, com vista a reproduzir a verdade. Ao contrário da incerteza aleatória, esse tipo de incerteza pode ser reduzido graças aos avanços no conhecimento dos fenômenos naturais, da quantificação/medição das variáveis envolvidas, do seu equacionamento físico e empírico, e dos métodos numéricos e recursos computacionais para sua resolução. De acordo com Ayyub e Klir (2006), a incerteza epistêmica, apesar de seu caráter subjetivo relativamente à incerteza aleatória, pode ser representada também por probabilidades, estimadas a partir do conhecimento prévio de especialistas, com potencial de refinamento à medida que informações e observações sobre o fenômeno ou sistema em estudo tornem-se disponíveis.

Segundo Todini (2007), a incerteza decorre do conhecimento inexato de alguma quantidade envolvida em uma modelagem matemática, e a forma mais compreensível de se expressar a informação disponível é por meio do uso de uma função densidade de probabilidade ou de uma função acumulada de probabilidades.

3.1.2 A incerteza em modelagem de sistemas ambientais e suas fontes

Um modelo que represente um sistema ambiental pode ser ilustrado esquematicamente pelo diagrama da Figura 3.1, do qual fazem parte seis componentes, conforme descrição a seguir (LIU e GUPTA, 2007).

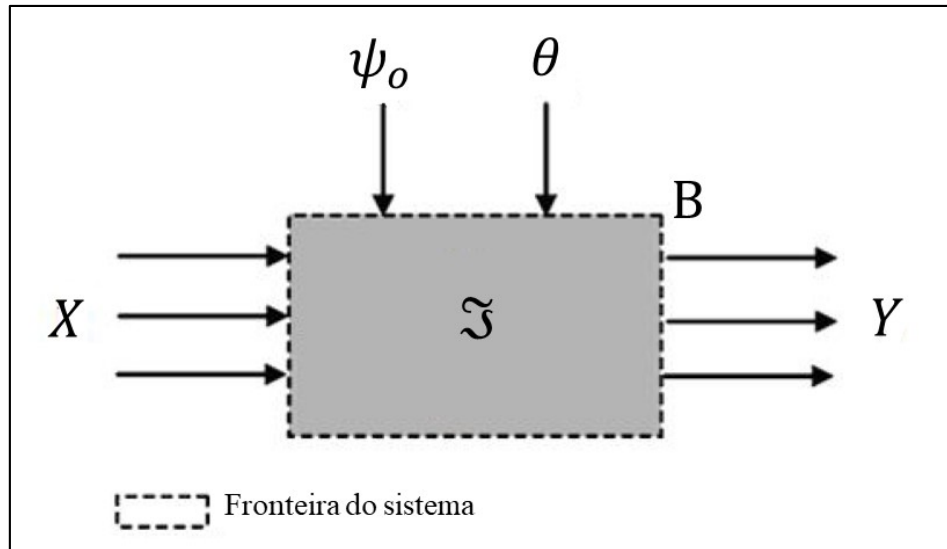


Figura 3.1 – Representação esquemática de um sistema ambiental com seus componentes (adaptado de Liu e Gupta, 2007).

Fonte: adaptado de Liu e Gupta (2007).

Na Figura 3.1, $\mathbf{B}$ é o contorno ou fronteira do sistema; $\mathbf{X} = \{x_1, \dots, x_n\}$ e $\mathbf{Y} = \{y_1, \dots, y_n\}$ são, respectivamente, vetores representativos das entradas e saídas ao sistema (fluxos de massa e/ou energia para dentro e para fora do sistema), de comprimento n , geralmente um indexador do tempo $t = \{1, \dots, n\}$, mas que também pode representar o espaço; Ψ_o são as condições iniciais; $\boldsymbol{\theta} = \{\theta_1, \dots, \theta_d\}$ é o vetor paramétrico do modelo, de comprimento ou dimensão d , e cujos componentes são considerados como invariáveis no tempo na maioria das aplicações; e $\mathbf{S} = \{s_1, \dots, s_n\}$ são variáveis de estado, armazenadas nas fronteiras do sistema e que evoluem juntamente com $\mathbf{Y}$ à medida que as equações que compõem o modelo $\mathfrak{T}$ são adotadas, condicionadas aos valores de $\mathbf{X}$ e $\boldsymbol{\theta}$.

O modelo $\mathfrak{T}$ pode ser considerado uma representação matemática formal da realidade e, como tal, automaticamente agregará incertezas epistêmicas. Hutton *et al.* (2011) descreveram três tipos de incertezas associadas a modelos que se propõem a explicar sistemas ambientais:

- Incerteza estrutural: deriva da representação matemática da realidade resultante das abstrações, hipóteses e simplificações aplicadas ao sistema ambiental em estudo para viabilizar a modelagem, bem como da representação numérica e da discretização espacial e temporal do modelo. Sendo assim, esse tipo de incerteza pode ser encontrada seja nas equações de $\mathfrak{T}$, seja através da representação da fronteira $\mathbf{B}$, e é considerada uma forma de incerteza epistêmica, a qual pode ser reduzida à medida que mais informações sobre o

sistema tornam-se disponíveis, possibilitando melhor compreensão acerca do mesmo e de sua representação. No entanto, os erros devidos ao conhecimento limitado da estrutura de um modelo nunca serão totalmente eliminados, já que é impossível conhecer a “verdade” total de um sistema. Adicionalmente, limitações em métodos de resolução numérica, em recursos computacionais e na aquisição de dados também acarretam erros dessa natureza, uma vez que levam a simplificações para representar a realidade. Finalmente, embora esse tipo de incerteza seja esperado e reconhecido de maneira intuitiva, atribuir expressões ou modelos de erros para quantificá-lo e compartimentá-lo em relação aos demais tipos de incertezas ainda é tarefa difícil e menos estudada;

- Incerteza paramétrica: é a incerteza atribuída aos valores dos parâmetros usados nas equações do sistema ambiental modelado, e que pode ter natureza aleatória e epistêmica. No segundo caso, os erros decorrem geralmente da discretização espaço-temporal dessas equações, resultando numa inabilidade do modelo em conciliar a escala das observações com seus parâmetros.

Os parâmetros dos modelos agregam incertezas à modelagem devido a diversos fatores. Kavetski *et al.* (2003) e Lane (2005), por exemplo, citam que tais incertezas emergem do fato de que muitos modelos baseiam-se em parâmetros conceituais, ou seja, desprovidos de significado físico, cujas estimativas dependerão dos dados de entrada e saída utilizados e dos métodos de calibração. Mesmo parâmetros que podem ser obtidos a partir de propriedades físicas, a exemplo da permeabilidade do solo, acarretam erros, seja por problemas de medição, seja por representarem uma propriedade do sistema em âmbito local e em determinado intervalo de tempo. Além disso, a maioria dos parâmetros, seja qual for o tipo de modelo, se conceitual ou se fisicamente baseado, é tratada de forma estática no tempo. Tal hipótese, a depender dos objetivos de utilização de determinado modelo e de sua estrutura, pode ser considerada como válida, mas, em muitas situações, não corresponde à realidade do fenômeno modelado (HUTTON *et al.*, 2011).

Uma vez que a incerteza paramétrica pode resultar em erros significativos nas previsões de modelos (TODINI, 2007), a mesma tem recebido ênfase em estudos de quantificação de incertezas associadas à modelagem matemática de sistemas ambientais (HUTTON *et al.*, 2011). Um dos reflexos dessa valorização em detrimento das demais formas de erros é a superestimação das incertezas devidas aos parâmetros, que acabam por mascarar e embutir em

suas diversas formas de representação (como intervalos de confiança) as incertezas estruturais e de medição (KAVETSKI *et al.*, 2003; STEDINGER *et al.*, 2008).

- Incerteza de medição ou dos dados: refere-se à incerteza nos valores usados para definir as condições iniciais (Ψ_0), os *inputs* ou dados de entrada ($\mathbf{X}$), e as observações usadas para se avaliar as predições dos modelos (tanto de variáveis de estado, $\mathbf{S}$, quanto dos *outputs*, ou saídas, $\mathbf{Y}$), e pode ser de natureza aleatória e epistêmica. Esse tipo de incerteza resulta de erros em métodos, procedimentos e ferramentas de instrumentação/medição das variáveis de interesse, do descompasso entre a escala e a resolução espaço-temporal das observações e das informações requeridas ou previstas pelos modelos.

Vrugt (2016) defendeu que o conhecimento explícito dos erros de cada uma dessas fontes permitiria definir estratégias para melhoria dos modelos ou nos procedimentos de coleta e tratamento de dados de entrada e saída. Para Kavetski *et al.* (2003), a quantificação de incertezas oriundas dessas distintas fontes em modelagem ambiental constitui a chave para avanços nos procedimentos de calibração e validação desses modelos, e até mesmo na regionalização de parâmetros. Merwade *et al.* (2008) ressaltaram a importância de se obterem avanços na compreensão das incertezas associadas a diferentes fontes e aspectos em modelagem de cheias para mapeamento de inundações com o objetivo de se aumentar a confiabilidade nos procedimentos e métodos utilizados.

A quantificação das incertezas embutidas em cada uma dessas fontes tem sido objeto de diversos estudos ao longo dos últimos 30 anos, em diversas áreas do conhecimento, muitas vezes sob a forma de um problema de calibração (KAVETSKI *et al.*, 2003; HUTTON *et al.*, 2011; VRUGT, 2016).

3.2 Fontes de incerteza na modelagem hidráulica de vazões de cheia e na elaboração de mapas de inundação

De acordo com Merwade *et al.* (2008), as incertezas presentes em modelagem hidráulica de vazões de cheia devem-se às incertezas na vazão ou hidrograma de projeto ou de referência, nas elevações do terreno, consubstanciadas por bases topobatimétricas oriundas de distintas fontes, nas cotas altimétricas dos perfis de escoamento, e, em se tratando do traçado de manchas de inundação, nas técnicas utilizadas para mapeamento da área inundada. Tais fontes são detalhadas nos próximos parágrafos. Devido à falta de distinção geralmente encontrada na literatura do tema das incertezas, o termo modelagem hidráulica é utilizado aqui de forma

ampla, englobando a simulação por modelos matemáticos em regime permanente ou não permanente para obtenção de perfis de escoamento de cheias, e as correspondentes profundidades, áreas inundadas e velocidades de escoamento na calha e nas planícies, ao longo de trechos fluviais de interesse. Quando pertinente, haverá menção à modelagem hidrodinâmica de hidrogramas de cheia.

Uma importante fonte de incerteza reside na vazão ou hidrograma de projeto ou de referência utilizada na modelagem, cujos métodos de obtenção, selecionados principalmente em função da disponibilidade de dados fluviométricos e/ou pluviométricos e climatológicos, agregam inúmeras incertezas. Dentre as possíveis técnicas para seu cálculo, mencionam-se: (i) análise de frequência local ou regional; (ii) curvas de regressão regionais baseadas em características físicas das bacias; e (iii) modelagem chuva-vazão.

Em alguma etapa de qualquer um dos procedimentos supracitados, terá havido a necessidade de utilizar informações extraídas de uma ou mais curvas-chave, seja para calibrar os parâmetros do modelo hidrológico, seja para realizar a análise de frequência, ou mesmo para cálculo de vazões-índice que alimentam relações regionais para transferência da informação hidrológica. Sendo assim, a relação cota-descarga de seções fluviais de medição é uma importante fonte de incerteza que intrinsecamente permeia a modelagem hidráulica de vazões de cheia. Merwade *et al.* (2008) mencionam que um dos aspectos responsáveis pelos erros em curvas-chave é o fato de que, na maioria dos casos, estas são estabelecidas a partir de vazões médias e baixas, de forma que o ramo elevado é obtido por extrapolação ou por regressão a partir de poucos pontos. O método de extrapolação em si constitui outra fonte de incerteza (KUCZERA, 1996). Por sua vez, os dados das campanhas de medição de descarga líquida estão sujeitos a erros de medição sistemáticos e aleatórios (DOMENEGHETTI *et al.*, 2012).

A análise de frequência local, por sua vez, abarca incertezas originadas pelos distintos métodos utilizados para estimação dos parâmetros das distribuições de probabilidades avaliadas e pela escolha da distribuição candidata. Em adição, a extensão dos registros históricos disponíveis para seleção das amostras de máximos anuais ou parciais (geralmente de vazões médias diárias), a utilização de séries anuais ou parciais, e mudanças em tais séries causadas por alterações nas bacias hidrográficas ou por tendências climáticas são aspectos que introduzem incertezas ao processo em questão, e que naturalmente se apresentam também na análise de frequência em âmbito regional, uma vez que constituem etapas ou procedimentos a ela necessários.

Esta, por sua vez, e relações regionais baseadas em características físicas de bacias hidrográficas ou em vazões-índice acarretam incertezas adicionais às supracitadas, pois dependem da obtenção de atributos físicos (i.e., área de drenagem, declividade do talvegue principal, fator de compacidade, porcentagem de área impermeável) das áreas de contribuição para as quais se deseja transferir a informação hidrológica em questão, os quais são geralmente estimados por meio de mapas digitalizados ou de bases cartográficas georreferenciadas, em escalas variáveis.

Caso se opte pela modelagem chuva-vazão, os fatores que afetam a precisão da predição de hidrogramas ou vazões de referência, e que portanto atuam como fontes de incertezas, são inúmeros, dentre os quais se mencionam: (i) a origem e procedimentos de obtenção de dados pluviométricos e evaporimétricos e métodos para interpolação espacial e distribuição temporal das chuvas de projeto; (ii) as escalas de representação temporal e espacial das bacias hidrográficas de contribuição e dos processos que compõem o ciclo hidrológico; (iii) a estrutura e as hipóteses adotadas pelos modelos, bem como a definição dos parâmetros que os regem, a maioria dos quais de base conceitual; (iv) a equifinalidade de parâmetros, de maneira que várias combinações possíveis geram respostas com qualidade semelhante à luz de distintos indicadores de desempenho; (v) métodos para calibração dos modelos e as funções-objetivo adotadas, muitas das quais priorizam seja descargas de cheia, seja de estiagem, e cujos resultados devem ser usados com cautela a depender da vazão ou do período do regime de hidrológico a ser estudado; e (vi) a adoção de parâmetros para obtenção da resposta hidrológica a cenários climatológicos distintos daqueles utilizados em sua calibração.

Outro importante fator que introduz incertezas à modelagem de cheias, neste caso referente tanto à modelagem hidrológica para sua estimação quanto à hidráulica para sua propagação, refere-se aos dados topográficos e batimétricos de alimentação dos modelos hidrológicos e hidráulicos e seu tratamento para sua inserção nestes, cujos possíveis erros podem ser atribuídos à resolução horizontal e à precisão vertical das informações, aos distintos métodos para interpolação espacial dos dados brutos na geração de modelos digitais de terreno (MDTs) ou de elevação (MDEs), e aos critérios e métodos adotados para composição de tais superfícies a partir de dados de diversas fontes e escalas espaciais. De acordo com Merwade *et al.* (2008), esses aspectos podem interferir na modelagem e mapeamento de cheias de três maneiras. Primeiramente, podem afetar hidrogramas e vazões de referência quando utilizados em modelagem hidrológica. Em segundo lugar, acarretam flutuações nas elevações da linha de água

derivadas a partir de modelos hidráulicos, e, por fim, podem causar variações na extensão horizontal das manchas de inundação, a depender da disponibilidade das informações topobatimétricas e da forma como as mesmas forem tratadas.

Dentre os dados brutos, destacam-se curvas de nível, seções topobatimétricas e nuvens de pontos com coordenadas em três dimensões, oriundas de levantamentos de campo ou de perfilamento a laser, todos eles sujeitos a erros de medição sistemáticos e aleatórios, dependentes dos equipamentos utilizados em sua obtenção e dos métodos de leitura e de consistência da informação coletada.

Ademais, a depender do formato digital em que tais informações são armazenadas, deve-se proceder a uma transformação para alimentar os distintos modelos hidrológicos e hidráulicos, a qual dependerá do software utilizado (a exemplo do ArcGis e QGis) e do algoritmo aplicado na transformação de tais dados. Há ainda a possibilidade de que os dados disponíveis tenham sido levantados em distintos sistemas de coordenadas geográficas ou projetadas, e, para sua integração em um SIG (Sistema de Informações Geográficas), é necessário proceder à transformação da informação de um sistema para outro, procedimento cuja qualidade dependerá do método e do software adotado. Tais técnicas certamente agregam variabilidade aos resultados e à predição, associada à qualidade da informação topobatimétrica, e que geralmente não é tratada explicitamente em estudos de avaliação de incertezas.

Passando-se à modelagem hidráulica em si, há uma miríade de fatores que acarretam incertezas, ou erros, sobre as estimativas dos perfis de escoamento e, assim, sobre as profundidades e as cotas altimétricas alcançadas pela linha de água, e, conseqüentemente, sobre a largura e a extensão das áreas inundadas. Modelos hidráulicos são sensíveis à descrição geométrica da calha e da planície, aos parâmetros considerados, em especial, e predominantemente, à rugosidade do leito fluvial; ainda, à representação de estruturas hidráulicas, como pontes, bueiros e diques laterais, ao número de dimensões consideradas (1D, 2D ou 3D), e ao método numérico de resolução das equações que governam o escoamento.

Alguns autores defendem que um dos aspectos mais críticos na modelagem hidráulica deve-se à representação da calha e das planícies fluviais (ARONICA *et al.*, 1998; PAPPENBERGER *et al.*, 2005), o que inclui a qualidade da informação topobatimétrica (i.e., sua resolução espacial e precisão vertical), o espaçamento entre seções transversais (se 1D) e a configuração da malha

de cálculo (no caso de 2D ou 3D). Segundo Merwade *et al.* (2008), alimentar um modelo 2D com um modelo digital de terreno de qualidade inferior, ou seja, de baixa resolução, pode levar a resultados tão ou mais conservadores do que aqueles obtidos quando do uso de um modelo mais simples, unidimensional.

O coeficiente de rugosidade de Manning merece lugar de destaque dentre os fatores que trazem variabilidade aos resultados, por se tratar do principal parâmetro de modelos de escoamento livre uniforme e gradualmente variado largamente utilizados, e por apresentar vasta literatura acerca de sua quantificação, tendo constituído objeto da maioria dos estudos de avaliação de incertezas associadas à modelagem hidráulica de vazões e hidrogramas de cheia e ao seu mapeamento, muitas vezes em caráter exclusivo (PAPPENBERGER *et al.*, 2005; DI BALDASSARRE *et al.*, 2010; CAMACHO *et al.*, 2015). Apesar de haver métodos para seleção de seu valor, como aquele apresentado por Chow (1959), ou inúmeras tabelas e figuras para aferição visual (ARCEMENT e SCHNEIDER, 1989; HICKS e MASON, 1998; BRUNNER, 2016a), sua definição na maioria das aplicações permanece subjetiva, sujeita à experiência do engenheiro hidráulico. Somam-se a isso os inúmeros métodos disponíveis para ponderação da rugosidade global - vide, por exemplo, aqueles mencionados por USACE (BRUNNER, 2016a) - cujo valor depende da profundidade do escoamento. Adicionalmente, alguns autores mencionam a variabilidade sazonal da rugosidade, em função de mudanças no uso e na ocupação do solo na planície de inundação ao longo do ano ou de períodos maiores (DOMENEGHETTI *et al.*, 2012; BRUNNER, 2016b), como fator agregador de incertezas à definição de seu valor.

O número de dimensões consideradas na resolução das equações da continuidade e da conservação da quantidade de movimento também constitui relevante aspecto que acarreta variabilidade aos resultados. Ainda que prevaleçam estudos e trabalhos técnicos de propagação de cheias em modelos 1D, e assim, também os de quantificação de suas incertezas, prolifera nas esferas acadêmica e técnica a utilização de modelos 2D, em tese mais adequada aos casos em que se esperam efeitos importantes de remanso de confluências e de reservatórios e de amortecimento de cheias pelas planícies de inundação e lagoas marginais. No entanto, a confiabilidade dos modelos bidimensionais e sua propalada superioridade sobre os unidimensionais depende sobremaneira da qualidade dos dados de terreno e batimétricos, e da disponibilidade de informações para uma calibração adequada, como profundidade e velocidade em diversos pontos ou seções do estirão fluvial (MERWADE *et al.*, 2008), e não

somente cotagramas e hidrogramas nos extremos de jusante e de montante, geralmente considerados suficientes para o caso unidimensional. Werner (2004) concluiu que modelos 1D e 2D alimentados por dados topográficos de baixa resolução podem ter desempenho semelhante à situação de quando se utiliza esse tipo de informação com melhor qualidade se a calibração é realizada com registros de estações fluviométricas que não contenham cheias importantes. Nesse caso, as previsões para cheias de maior magnitude representam uma extrapolação da parametrização obtida, fator que agrega novas incertezas ao processo de simulação dos perfis de escoamento e de mapeamento das áreas inundadas, seja qual for o número de dimensões utilizadas na modelagem. Uma forma de contornar essa carência de informações sistemáticas sobre eventos maiores é utilizar marcas de cheia levantadas em campo ou mesmo o traçado de áreas inundadas estimadas a partir de fotos aéreas e imagens de satélite (STEPHENS *et al.*, 2012), técnicas essas que também constituem fontes de erros.

A modelagem 3D é reservada a casos mais específicos, aplicada geralmente a estruturas hidráulicas e aos trechos de sua influência a montante e a jusante, e raramente a estirões fluviais com extensão a partir de alguns quilômetros, devido à necessidade de maior precisão e à quantidade de dados para alimentação e calibração, bem como de maior esforço computacional para realizar as simulações. Desse modo, ainda são escassos os estudos de quantificação de incertezas oriundas da propagação de cheias em ambiente tridimensional, tal como realizado por Camacho *et al.* (2015) para os rios que aportam a uma baía na costa leste dos Estados Unidos e o estuário formado por eles.

Ainda que atualmente haja maior disponibilidade de dados topográficos mais precisos, e abundância de equipamentos potentes para realização de campanhas de medição de batimetria, marcas de cheia e de velocidades do escoamento na planície e na calha principal, fatores facilitadores do uso de modelos hidráulicos em ambiente 2D e 3D, a seleção do número de dimensões permanece um aspecto subjetivo, sujeito à análise e à *expertise* da equipe envolvida em determinado estudo ou projeto, num processo em que devem ser cotejados os objetivos da modelagem, as variáveis de saída de interesse, os recursos financeiros e computacionais disponíveis e o cronograma estipulado.

Todas as incertezas até aqui mencionadas referem-se à estrutura e à parametrização de modelos utilizados para simulação de perfis de escoamento de cheias em regime permanente ou não permanente, e aos dados que os alimentam. A representação de resultados sob a forma de mapas

de inundação, como é de interesse nesta pesquisa, traz em seu bojo possíveis novas fontes de erros. Segundo Merwade *et al.* (2008), essas incertezas devem-se principalmente aos procedimentos e algoritmos adotados por softwares acoplados aos modelos de simulação para processamento dos resultados neles obtidos e posterior mapeamento da área inundada, e às distintas resoluções espaciais e formatos digitais das bases topográficas usadas na modelagem hidráulica, os quais geralmente precisam ser adaptados para permitir o traçado da mancha.

Atualmente, há dois modelos digitais mais utilizados para representar superfícies tridimensionais em um SIG – Sistema de Informações Geográficas: (i) sob a forma de *raster*, de base matricial, em que o dado, no caso, as elevações do terreno ou as cotas de inundação, é armazenado no centro de cada célula, ou *pixel*, com formato quadrado, ou como (ii) TIN – *Triangular Irregular Network*, de base vetorial, em que a informação geográfica de interesse está posicionada em vértices conectados por uma rede de triângulos irregulares. É muito comum que as informações topobatimétricas e as das superfícies de escoamento simuladas sejam armazenadas e processadas nesses dois formatos, nem sempre os mesmos, e com distintas resoluções horizontais e precisão vertical, e o cruzamento de ambas demanda a transformação e interpolação de um para o outro, o que naturalmente agrega incertezas de difícil identificação e quantificação.

Ainda que a base topobatimétrica utilizada na modelagem hidráulica e na modelagem do terreno para posterior cruzamento de resultados e obtenção da área inundada seja a mesma, bem como seu formato digital, geralmente a forma como a informação geométrica deve ser tratada e repassada ao modelo hidráulico altera a precisão vertical e a resolução horizontal original. Portanto, uma vez que as simulações hidráulicas tenham sido realizadas, as cotas altimétricas de linha de água estarão disponíveis para cada seção transversal (se 1D) ou em cada nó da malha bi ou tridimensional (cuja representação depende do programa/modelo utilizado), e não necessariamente nos mesmos pontos cotados do terreno (no caso, o centro dos *pixels* de arquivos tipo *raster*, ou os vértices dos arquivos tipo TIN). Desse modo, é necessário realizar uma compatibilização entre ambas as superfícies (a do solo e a correspondente ao perfil de escoamento), procedendo a transformações e interpolações, cujos métodos e rotinas dependerão do software de pós-processamento da modelagem hidráulica e de georreferenciamento da mancha de inundação.

3.3 Métodos de cunho Bayesiano com potencial de aplicação à quantificação de incertezas na modelagem hidrodinâmica

A pesquisa bibliográfica acerca dos métodos disponíveis para quantificação de incertezas mostrou a amplitude de abordagens distintas que podem ser adotadas na modelagem de sistemas ambientais, cujas escolha e aplicabilidade dependem dos modelos utilizados, dos dados disponíveis, do conhecimento prévio acerca das variáveis envolvidas e dos objetivos esperados a partir da quantificação/avaliação de incertezas. Extensas compilações metodológicas podem ser encontradas em Morgan e Henrion (1990), Saltelli *et al.* (2004), Ayyub e Klir (2006), e Hutton *et al.* (2011). Ao se restringir o universo de aplicações à modelagem matemática dos escoamentos à superfície livre gradualmente variados, permanentes ou transientes, e direcionados a representar situações reais, verificou-se que o ferramental utilizado restringe-se a alguns métodos de cunho probabilístico e geralmente indireto, nos quais não é possível deduzir expressões analíticas para a variabilidade dos *outputs*, a exemplo de, no presente caso, hidrogramas e cotogramas em seções transversais e manchas de inundação para trechos fluviais.

Do cruzamento entre a bibliografia avaliada e as características dominantes dos modelos numéricos comumente adotados para a propagação de cheias por estirões fluviais, foram estabelecidos alguns critérios para aprofundamento da pesquisa bibliográfica e para a prospecção dos métodos com potencial de aplicação à quantificação das incertezas envolvidas, sejam elas de caráter estrutural, paramétrico e/ou de medição/observação.

Primeiramente, considerando-se que o objetivo principal foi a quantificação das incertezas em dados de entradas e parâmetros, e como essas se propagam sobre as saídas do modelo, os métodos a serem elencados e potencialmente eleitos para aplicação devem permitir que se ultrapasse a questão da compreensão conceitual do problema em direção à representação formal e matemática do conhecimento, ainda que limitado e incerto, acerca dos referidos elementos, conforme sugerem Hutton *et al.* (2011). Para tanto, a teoria de probabilidades tem sido o paradigma dominante para se representar a incerteza (HALL, 2003).

Em segundo lugar, as equações que governam os escoamentos à superfície livre não permanentes e gradualmente variados são de caráter não linear e não apresentam solução analítica para os casos reais, salvo a possibilidade de se adotarem hipóteses simplificadoras. Dessa forma, pressupõe-se que a propagação das incertezas das variáveis e parâmetros envolvidos nesse tipo de modelagem matemática sobre os *outputs* não possa ser deduzida de

forma analítica, de modo que métodos aproximativos, ou indiretos, precisem ser usados, baseados, por exemplo, em expansão por séries de Taylor ou na realização de uma grande quantidade de simulações do mesmo processo em análise via Métodos de Monte Carlo.

Por fim, uma das premissas centrais desta pesquisa é a de que o conhecimento prévio acerca dos parâmetros e variáveis possa ser atualizado a partir de experimentos, ou, no caso do tipo de problema em questão, de dados observados acerca do fenômeno que se pretende representar, prática corrente nas engenharias (AYYUB e KLIR, 2006). Para situações em que isso é possível, verifica-se que métodos de abordagem Bayesiana são indicados, além de serem os que prevalecem nos exemplos pesquisados quanto ao tema da obtenção de incertezas associadas à modelagem hidrodinâmica de hidrogramas de cheias e de escoamento permanente de vazões máximas em estirões fluviais existentes. Adicionalmente, espera-se que a atualização por dados observados inerente a esses métodos garanta uma melhor identificação das incertezas presentes em variáveis de entrada e associadas aos parâmetros do modelo selecionado, em comparação a outros métodos pesquisados com potencial de aplicação, e que esse aspecto contribua para fornecer intervalos de credibilidade mais realísticos para os *outputs*.

Sendo assim, foram elencados três métodos de abordagem Bayesiana em sua essência para serem tratados na revisão de literatura, considerados como potencialmente aplicáveis à quantificação de incertezas envolvidas na modelagem dos escoamentos livres, em regime permanente ou transiente, e à realidade de informações topobatimétricas e de monitoramento hidrológico comumente encontradas no Brasil, em termos de disponibilidade e discretização espaço-temporal. Conforme se verá, de maneira geral, não é possível obter soluções analíticas para a aplicação do Teorema de Bayes, sobre o qual todos esses métodos se baseiam, o que demanda a realização de milhares de repetições do experimento (i.e., modelagem) em questão. Por essa razão, no subcapítulo seguinte, serão apresentadas algumas técnicas de amostragem de apoio à utilização desses métodos.

A abordagem Bayesiana tem sido uma das formas mais amplamente utilizadas para avaliação e quantificação das incertezas associadas aos parâmetros e à predição na modelagem matemática de sistemas ambientais, com aplicações em hidrologia, hidráulica, geologia, hidrogeologia, limnologia e qualidade da água (HUTTON *et al.*, 2011; VRUGT, 2016). Partindo-se do Teorema de Bayes, surgiram diversos métodos com o intuito de estimar os erros inerentes à modelagem, dentre os quais podem ser citados os métodos Bayesianos formais,

pseudo-Bayesianos e variações como o BaTEA - *Bayesian Total Error Analysis* e o BMA - *Bayesian Model Averaging*. Este subcapítulo da revisão de literatura inicia-se pelos métodos Bayesianos formais aplicados ao estudo de incertezas, uma vez que seus aspectos centrais são relevantes à compreensão de todos os métodos de cunho Bayesiano adotados nesse campo. Os demais subitens tratam de algumas técnicas que constituem desdobramentos da aplicação do Teorema de Bayes ao tema abordado, vislumbradas inicialmente como possíveis métodos a serem adotados na pesquisa ora relatada.

3.3.1 Métodos Bayesianos formais

O Teorema de Bayes é representado pela seguinte equação:

$$P(B|A) = \frac{P(A|B) \cdot P(B)}{P(A)} \quad (3.1)$$

na qual:

$P(B|A)$ = probabilidade condicional de ocorrência do evento B dada a ocorrência do evento A;

$P(A|B)$ = probabilidade condicional de ocorrência do evento A dada a ocorrência do evento B;

$P(B)$ = probabilidade de ocorrência do evento B, definida *a priori*; e

$P(A)$ = probabilidade de ocorrência do evento A.

Segundo Naghettini e Pinto (2007), o Teorema de Bayes pode ser aplicado para a atualização de probabilidades previamente conhecidas ou estabelecidas acerca de um fenômeno B à luz de informações coletadas a respeito de uma variável ou fenômeno A que esteja a ele relacionado.

A Equação 3.1 pode ser adaptada ao estudo de erros ou incertezas na modelagem de sistemas ambientais, podendo incorporar as incertezas estruturais, de medição e paramétricas. O caso típico, no entanto, e mais encontrado na literatura, trata-se da aferição dessas últimas (HUTTON *et al.*, 2011). Sob tal perspectiva, o Teorema de Bayes será desenvolvido nos próximos parágrafos, seguindo-se equações e hipóteses apresentadas nos estudos de Kavetski *et al.* (2003), Hutton *et al.* (2011) e Vrugt (2016). A incorporação de outras fontes de erros será objeto de itens subsequentes. A notação apresentada em negrito neste texto representa matrizes ou vetores, enquanto que variáveis são nomeadas por letras em itálico e sem negrito.

Um modelo $\mathfrak{I}$ para um processo físico em avaliação pode ser representado pela seguinte expressão (VRUGT, 2016), utilizando-se a notação mostrada no diagrama da Figura 3.1:

$$\tilde{\mathbf{Y}} \leftarrow \mathfrak{Z}(\boldsymbol{\theta}, \tilde{\mathbf{X}}, \tilde{\boldsymbol{\Psi}}_0) + \mathbf{E} \quad (3.2)$$

na qual:

$\tilde{\mathbf{Y}} = \{\tilde{y}_1, \dots, \tilde{y}_n\}$ e $\tilde{\mathbf{X}} = \{\tilde{x}_1, \dots, \tilde{x}_n\}$ são vetores de dados observados, ou medidos,

relativamente às saídas e entradas ao modelo, ou simulador $\mathfrak{Z}$, respectivamente;

n representa o comprimento dos vetores sob avaliação, geralmente indexado em termos temporais, embora também possa representar o espaço;

$\boldsymbol{\theta} = \{\theta_1, \dots, \theta_d\}$ é o vetor de d parâmetros que regem o modelo $\mathfrak{Z}$, com valores contidos no espaço paramétrico $\boldsymbol{\Theta}$;

$\tilde{\boldsymbol{\Psi}}_0$ são as condições iniciais;

$\mathbf{E} = \{\varepsilon_1, \dots, \varepsilon_n\}$ é o vetor de resíduos, $\mathbf{E} = \tilde{\mathbf{Y}} - \mathbf{Y}$, os quais englobam erros de medição/observação em dados de entrada e de saída, além daqueles oriundos da estrutura do modelo $\mathfrak{Z}$ e de sua parametrização $\boldsymbol{\theta}$; e

$\mathbf{Y}$ é a resposta do modelo, ou seja, $\mathbf{Y} \leftarrow \mathfrak{Z}(\boldsymbol{\theta}, \tilde{\mathbf{X}}, \tilde{\boldsymbol{\Psi}}_0)$.

Adotando-se o formalismo Bayesiano expresso na Equação 3.1 à representação matemática em 3.2, tem-se que a distribuição de probabilidades *a posteriori* dos parâmetros do modelo pode ser obtida condicionando-se o comportamento espaço-temporal do modelo à resposta observada, ou medida, $\tilde{\mathbf{Y}}$, do sistema modelado, e a uma informação prévia acerca do vetor $\boldsymbol{\theta}$, denotada por $\boldsymbol{\Omega}$. Assim:

$$p(\boldsymbol{\theta}|\tilde{\mathbf{Y}}, \boldsymbol{\Omega}) = \frac{p(\tilde{\mathbf{Y}}|\boldsymbol{\theta}, \boldsymbol{\Omega}) \cdot p(\boldsymbol{\theta}|\boldsymbol{\Omega})}{p(\tilde{\mathbf{Y}}|\boldsymbol{\Omega})} \quad (3.3)$$

na qual:

$p(\boldsymbol{\theta}|\tilde{\mathbf{Y}}, \boldsymbol{\Omega})$ é a distribuição de probabilidades conjuntas *a posteriori* do vetor paramétrico, condicionada às observações do fenômeno modelado, denotadas por $\tilde{\mathbf{Y}}$;

$p(\boldsymbol{\theta}|\boldsymbol{\Omega})$ é a distribuição de probabilidades conjuntas *a priori* do vetor paramétrico;

$p(\tilde{\mathbf{Y}}|\boldsymbol{\theta}, \boldsymbol{\Omega}) \equiv L(\boldsymbol{\theta}, \boldsymbol{\Omega}|\tilde{\mathbf{Y}})$ denota a função de verossimilhança; e

$p(\tilde{\mathbf{Y}}|\boldsymbol{\Omega})$ é a distribuição preditiva *a priori*, obtida pela expressão a seguir:

$$p(\tilde{\mathbf{Y}}|\boldsymbol{\Omega}) = \int_{\boldsymbol{\Theta}} p(\tilde{\mathbf{Y}}|\boldsymbol{\theta}, \boldsymbol{\Omega}) \cdot p(\boldsymbol{\theta}|\boldsymbol{\Omega}) d\boldsymbol{\theta} = \int_{\boldsymbol{\Theta}} p(\tilde{\mathbf{Y}}, \boldsymbol{\theta}|\boldsymbol{\Omega}) d\boldsymbol{\theta}, \quad (3.4)$$

integrada ao longo de todo o espaço paramétrico, tal que $\boldsymbol{\theta} \in \boldsymbol{\Theta} \in \mathbb{R}^d$.

Verifica-se pelas Equações 3.3 e 3.4 que o conjunto $\boldsymbol{\Omega}$ permeia todas as etapas de cálculo, razão pela qual, e para simplificar as expressões que se seguem, essa variável será suprimida doravante neste texto. Por sua vez, como a integração em 3.4 é realizada sobre todo o intervalo paramétrico, seu resultado é uma constante que tem a função de normalizar o segundo membro da Equação 3.3 (FERNANDES e SILVA, 2017). Na prática, $p(\tilde{\mathbf{Y}}|\boldsymbol{\Omega})$ não é requerida para a

estimação *a posteriori*, uma vez que toda a inferência estatística sobre $p(\boldsymbol{\theta}|\tilde{\mathbf{Y}}, \boldsymbol{\Omega})$ pode ser feita a partir da distribuição não normalizada, como mostra a expressão a seguir (VRUGT, 2016):

$$p(\boldsymbol{\theta}|\tilde{\mathbf{Y}}) \propto L(\boldsymbol{\theta}|\tilde{\mathbf{Y}}) \cdot p(\boldsymbol{\theta}) \quad (3.5)$$

Em outras palavras, a informação que se tem *a priori* acerca dos parâmetros $\boldsymbol{\theta}$ é atualizada à luz de dados disponíveis $\tilde{\mathbf{Y}}$ sobre um fenômeno a eles associado, representativo do processo modelado por $\mathfrak{F}$. A distribuição *a posteriori* é um modelo distributivo ou amostra atribuído a um conjunto paramétrico $\boldsymbol{\theta}$, dado um vetor de informações $\tilde{\mathbf{Y}}$ usado para se avaliar a capacidade de aqueles parâmetros de reproduzir o comportamento do sistema relativamente a outros conjuntos $\boldsymbol{\theta}$ (STEDINGER *et al.*, 2008). Conjuntos paramétricos que acarretem erros menores apresentam maiores valores para a função de verossimilhança e, conseqüentemente, também para a distribuição *a posteriori*.

A aplicação do Teorema de Bayes à estimação da incerteza paramétrica é representada esquematicamente pela Figura 3.2. As demais fontes de incerteza presentes na modelagem de sistemas ambientais também aparecem nesse diagrama, a fim de se enfatizar que a prescrição de modelos para quantificá-las é igualmente importante, embora seja pouco comum e acarrete maior complexidade aos métodos Bayesianos para estimação de incertezas.

Kavetski *et al.* (2003) defendem o uso de métodos Bayesianos na calibração de modelos hidrológicos, no lugar de métodos tradicionais (i.e., por meio de funções objetivo ou métodos multicritério) associados a uma abordagem frequentista, por oferecerem em sua estrutura a identificação mais adequada e mais facilmente compreensível dos erros ou incertezas atrelados aos parâmetros da modelagem, sob a forma de intervalos de credibilidade, combinando dados observados de entrada e saída com o conhecimento prático e intuitivo dos estudiosos da área de conhecimento do modelo.

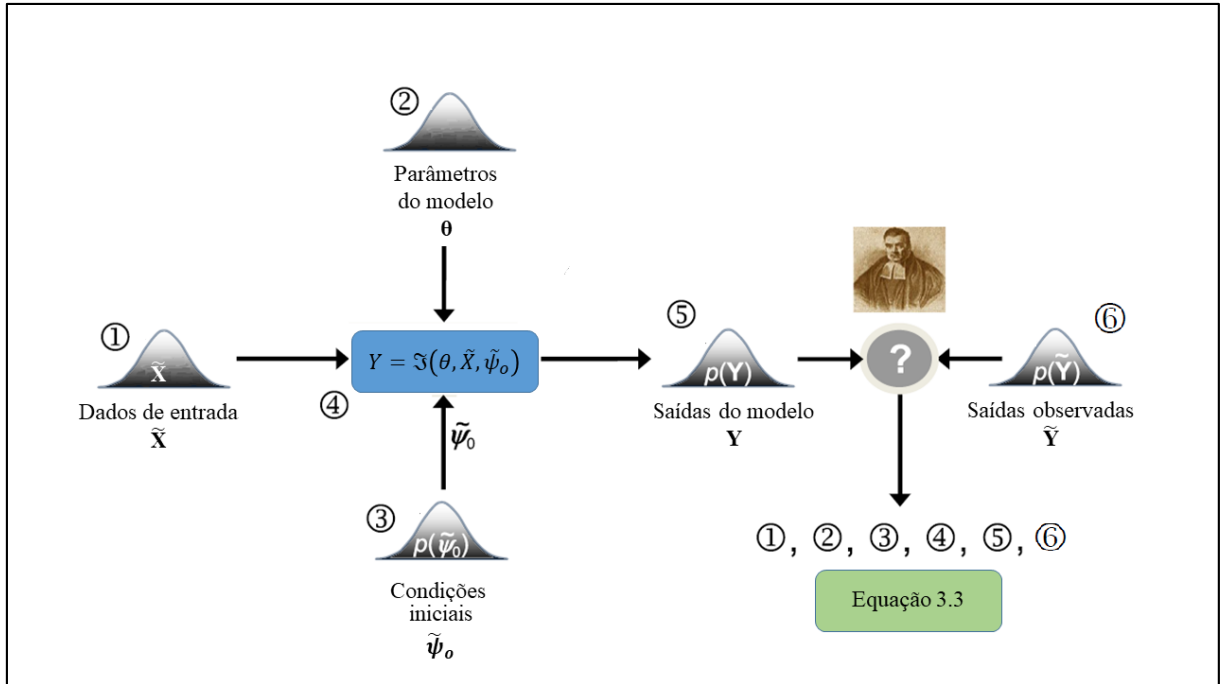


Figura 3.2 – Representação esquemática da aplicação do Teorema de Bayes à estimação de incertezas presentes na modelagem de sistemas ambientais, com ênfase sobre a incerteza paramétrica (adaptado de Vrugt, 2016).

Com base na Equação 3.5, para se implementar a abordagem Bayesiana a fim de se quantificar e avaliar as incertezas da modelagem, com foco sobre os parâmetros do modelo, são necessários 4 elementos básicos (HUTTON *et al.*, 2011):

- Definição da informação *a priori*, ou seja, das distribuições teóricas de probabilidade *a priori* para cada parâmetro do modelo;
- Escolha de um modelo de erros apropriado, composto pelas hipóteses acerca do comportamento estipulado para o vetor de resíduos **E**, bem como pela provável distribuição teórica de probabilidades que as representem e expressem;
- Dedução de uma função de verossimilhança apropriada, partindo-se do modelo de erros estipulado; e
- Dedução analítica da distribuição de probabilidades conjuntas *a posteriori*, quando possível, ou aplicação de métodos numéricos para sua estimação, o que ocorre na maioria das aplicações práticas em modelos de sistemas ambientais (HUTTON *et al.*, 2011; BOZZI *et al.*, 2015).

O significado e a obtenção de cada um desses elementos serão detalhados nos subitens a seguir.

3.3.1.1 Distribuições de probabilidades *a priori*

A distribuição *a priori* pode ser do tipo não informativa ou informativa. No primeiro caso, são utilizadas distribuições não informativas ou pouco informativas, que demonstrem conceitualmente o conhecimento vago que se tem a respeito dos parâmetros ou que traduzam a expectativa de que a informação fornecida pelos dados observados supere a informação *a priori* acerca dos parâmetros (FERNANDES e SILVA, 2017). Para tanto, diversos autores consideram a distribuição uniforme com limites superior e inferior que expressem o intervalo plausível para determinada variável (PAPPENBERGER *et al.*, 2005; VRUGT *et al.*, 2008; JUNG e MERWADE, 2012). Caso se queria definir a rigor uma distribuição não informativa que não seja imprópria, sob a hipótese de que θ seja ilimitado, pode-se utilizar alternativamente a distribuição Gama com parâmetro de escala de pequeno valor (FERNANDES e SILVA, 2017). Embora as distribuições Uniforme e Gama especificadas das formas mencionadas atribuam probabilidades iguais ou praticamente iguais para todas as possíveis realizações do vetor θ , não são consideradas estritamente não informativas. Nessa situação, deve-se utilizar a distribuição proposta por Jeffreys (1961 *apud* FERNANDES e SILVA, 2017), a qual deve ser definida de acordo com a função de verossimilhança selecionada (KAVETSKI *et al.*, 2006a).

No segundo caso, prescreve-se uma distribuição informativa, distinta daquelas supracitadas, que traduza o conhecimento prévio que se tem acerca da variável ou parâmetro avaliado. Podem ser aplicadas distribuições como a Normal (limitada ou não), a Beta, a Gama ou a Triangular, a exemplo daquelas adotadas em estudos como os de Kavetski *et al.* (2006a) e Bozzi *et al.* (2015).

A seleção do coeficiente de variação das distribuições *a priori* para os parâmetros a serem avaliados deve ser feita de maneira cuidadosa, como demonstram diversos estudos. Bozzi *et al.* (2015), por exemplo, concluíram que as propriedades estatísticas das saídas de um modelo hidráulico em regime permanente dependem sobremaneira do coeficiente de variação das variáveis de entrada, e admitem que geralmente essa informação não se encontra disponível, demandando sua seleção a partir de uma série de hipóteses, em um processo no qual a opinião de especialistas é importante.

3.3.1.2 A função de verossimilhança

A função de verossimilhança, $L(\theta, \Omega | \tilde{Y})$, é usada para sintetizar as distâncias entre a resposta simulada pelo modelo, denotada por $\mathbf{Y}$, uma vez que a pretensão é aproximar-se do verdadeiro

valor da variável modelada, e as observações correspondentes, $\tilde{\mathbf{Y}}$. Em outras palavras, constitui-se por uma expressão geralmente de cunho probabilístico que exprima as relações funcionais dos dados de saída em relação a $\boldsymbol{\theta}$, e, assim, o comportamento que se supõe para os resíduos, razão pela qual sua definição depende diretamente do chamado modelo de erros, ou *error model*, nas descrições de Kavetski *et al.* (2006a), Schoups e Vrugt (2010) e Hutton *et al.* (2011). Não é necessário que ela seja representada propriamente por uma função densidade de probabilidade – o que definirá se o método será Bayesiano formal ou informal (HUTTON *et al.*, 2011).

Uma hipótese muito adotada é a de que os resíduos sejam isentos de autocorrelação, situação na qual a verossimilhança de seu vetor, de comprimento n , e doravante representado sob a notação $\mathbf{E}(\boldsymbol{\theta}) = \tilde{\mathbf{Y}} - \mathbf{Y}(\boldsymbol{\theta}) = \{\varepsilon_1(\boldsymbol{\theta}), \dots, \varepsilon_n(\boldsymbol{\theta})\}$, pode ser expressa por (VRUGT, 2016):

$$L(\boldsymbol{\theta}|\tilde{\mathbf{Y}}) = f_{\tilde{y}_1}(y_1(\boldsymbol{\theta})) \times f_{\tilde{y}_2}(y_2(\boldsymbol{\theta})) \cdots \times f_{\tilde{y}_n}(y_n(\boldsymbol{\theta})) = \prod_{t=1}^n f_{\tilde{y}_t}(y_t(\boldsymbol{\theta})) \quad (3.6)$$

na qual $f_a(b)$ significa a função densidade de probabilidade de a avaliada em b .

A variável t indica que $\tilde{\mathbf{Y}}$, $\mathbf{Y}$, $\mathbf{E}$ e demais vetores são avaliados em termos temporais, seguindo-se a notação dominante em estudos de quantificação de incertezas aplicada à modelagem de sistemas ambientais. No entanto, a depender do tipo de problema e dos objetivos da avaliação, o conceito da Equação 3.6 pode ser adaptado à indexação em termos espaciais, ou seja, variáveis de saída e de estado, bem como os resíduos da modelagem, são analisados em determinadas posições do espaço modelado.

Sob a hipótese de que os resíduos sejam normalmente distribuídos com variância constante ou variável, ou seja, $\varepsilon_t(\boldsymbol{\theta}) \sim \mathcal{N}(0, \sigma_t^2)$, a Equação 3.6 apresenta a seguinte forma:

$$L(\boldsymbol{\theta}|\tilde{\mathbf{Y}}, \hat{\boldsymbol{\sigma}}^2) = \prod_{t=1}^n \frac{1}{\sqrt{2\pi\hat{\sigma}_t^2}} \cdot \exp \left[-\frac{1}{2} \cdot \left(\frac{\tilde{y}_t - y_t(\boldsymbol{\theta})}{\hat{\sigma}_t} \right)^2 \right] \quad (3.7)$$

Na Equação 3.7, $\hat{\boldsymbol{\sigma}} = \{\hat{\sigma}_1, \dots, \hat{\sigma}_n\}$ é um vetor de comprimento n com desvios-padrão estimados para os erros de medição das observações (VRUGT, 2016), donde se justifica a notação com acento circunflexo sobre o vetor e as correspondentes variáveis. Essa formulação admite erros de medição homoscedásticos (variância constante) ou heteroscedásticos (variância dependente da magnitude dos dados de saída).

Na condição em que se supõe a heteroscedasticidade dos resíduos, o vetor σ^2 pode ser inferido sob a estrutura Bayesiana juntamente com o vetor paramétrico, recebendo assim a denominação usual de variável *latente* (KAVETSKI *et al.*, 2006a; 2006b; SCHOUPS e VRUGT, 2010; VRUGT, 2016), e, conseqüentemente, constituindo também elemento caracterizado por incertezas.

Já para a situação de variância constante, σ^2 constante, esta pode ser inferida *a posteriori* por meio de seu estimador não-enviesado aplicado sobre as realizações do modelo, $\mathbf{Y}$, de modo a não aumentar a dimensionalidade do problema. Para essa finalidade, Stedinger *et al.* (2008) sugerem a utilização da expressão a seguir:

$$\hat{\sigma}^2 = \frac{1}{n} \cdot \sum_{t=1}^n (\tilde{y}_t - y_t(\boldsymbol{\theta}_{MVS}))^2 \quad (3.8)$$

na qual $\boldsymbol{\theta}_{MVS}$ representa o vetor paramétrico ótimo obtido pelo método de Máxima Verossimilhança e $y_t(\boldsymbol{\theta}_{MVS})$ é a correspondente saída do modelo no intervalo de tempo t .

Uma alternativa ao cálculo anterior é proposta por Kavetski *et al.* (2006a), que sugerem que σ^2 constante seja obtido após a estimação da distribuição *a posteriori* dos parâmetros do modelo e que cada elemento a ser usado no lugar de $y_t(\boldsymbol{\theta}_{MVS})$ seja resultante da média aritmética de todas as realizações simuladas para cada ordenada t .

Outra abordagem possível, quando se pretende avaliar a incerteza de σ^2 , mesmo sob a hipótese de homoscedasticidade, é tratá-la como variável latente, adotando-se uma distribuição de probabilidades marginal *a priori* para a mesma, de forma que esta seja inferida junto com o vetor $\boldsymbol{\theta}$.

Por razões de estabilidade numérica e simplicidade algébrica, é usual que se trabalhe com a função de log-verossimilhança, de maneira que a Equação 3.7, por exemplo, adquira o formato a seguir, sob a hipótese de que o vetor σ^2 seja conhecido previamente ao processo de inferência Bayesiana:

$$\ell(\boldsymbol{\theta}|\tilde{\mathbf{Y}}, \hat{\sigma}^2) = -\frac{n}{2} \cdot \log(2\pi) - \sum_{t=1}^n [\log(\hat{\sigma}_t)] - \frac{1}{2} \cdot \sum_{t=1}^n \left(\frac{\tilde{y}_t - y_t(\boldsymbol{\theta})}{\hat{\sigma}_t} \right)^2 \quad (3.9)$$

Essa expressão reduz-se à forma (KAVETSKI *et al.*, 2003):

$$\ell(\boldsymbol{\theta}|\tilde{\mathbf{Y}}) \propto -\left(\frac{n-1}{2}\right) \cdot \log\{\sum_{t=1}^n [\varepsilon_t(\boldsymbol{\theta})]^2\} \quad (3.10)$$

para o caso de os erros serem homoscedásticos, invariáveis em relação ao tempo e à variável simulada, $\mathbf{Y}$.

Da Equação 3.10 deriva a função-objetivo igual à soma dos quadrados das diferenças entre valores observados e simulados, amplamente utilizada para calibração de parâmetros de modelos de sistemas ambientais (KAVETSKI *et al.*, 2003; SCHOUPS e VRUGT, 2010; CHENG *et al.*, 2014) e que, portanto, considera que tais diferenças (i.e., resíduos) comportam-se sob as hipóteses de normalidade, homoscedasticidade e isenção de autocorrelação. Devido à sua simplicidade, a referida equação constitui uma das funções de verossimilhança mais aplicadas ao estudo de incertezas associadas a modelos ambientais, e, mais especificamente, na avaliação de incertezas envolvidas nas modelagens chuva-vazão e hidrodinâmica sob a perspectiva de uma abordagem Bayesiana formal (KAVETSKI *et al.*, 2003; VRUGT *et al.*, 2003; 2008; CAMACHO *et al.*, 2015).

Um dos grandes desafios da abordagem Bayesiana no estudo de incertezas é a correta prescrição do modelo de erros, e, conseqüentemente, da função de verossimilhança, uma vez que a maioria dos processos físicos modelados tem comportamento não linear e sua modelagem apresenta resíduos correlacionados, heteroscedásticos e não Gaussianos (KAVETSKI *et al.*, 2003; VRUGT *et al.*, 2009a; SCHOUPS e VRUGT, 2010). Dessa forma, a adoção da expressão 3.10, e mesmo de 3.9, pode acarretar a estimação incorreta da distribuição de probabilidades conjuntas *a posteriori* e também da incerteza preditiva, segundo alertam Kavetski *et al.* (2003), Vrugt *et al.* (2009a) e Hutton *et al.* (2011).

Em função desse impasse, foram realizados diversos estudos de avaliação de funções de verossimilhança que representassem adequadamente as condições anteriormente citadas para os resíduos, o que certamente agregou complexidade à estimação Bayesiana e a possibilidade de identificação mais adequada das incertezas associadas aos parâmetros da modelagem, inclusive com menores intervalos de credibilidade para estes, além da perspectiva de se representar, ainda que implicitamente, as incertezas estruturais e de medição de dados (SCHOUPS e VRUGT, 2010; HUTTON *et al.*, 2011).

A ideia de uma função de verossimilhança flexível, que acomode diferentes hipóteses quanto à não normalidade e heteroscedasticidade dos resíduos, parece interessante para modelos não

lineares. Bozzi *et al.* (2015) concluíram, por exemplo, que a propagação de incertezas a partir de distribuições *a priori* simétricas para o coeficiente de rugosidade e a vazão resulta em distribuições *a posteriori* preditivas com formas assimétricas para níveis de água modelados em regime permanente e uniforme, devido à não linearidade dominante nas equações que regem os escoamentos livres.

Como exemplo, menciona-se o caso em que se adota a premissa de que os resíduos $\mathbf{E}(\boldsymbol{\theta}) = \{\varepsilon_1(\boldsymbol{\theta}), \dots, \varepsilon_n(\boldsymbol{\theta})\}$ possuam correlação temporal (ou espacial, alternativamente), podendo-se explicitar esse comportamento ao se deduzir a log-verossimilhança. Supondo-se que os mesmos sigam um processo estocástico autoregressivo de ordem $p = 1$ e sejam normalmente distribuídos, tem-se que:

$$\varepsilon_t(\boldsymbol{\theta}) = \phi_1 \cdot \varepsilon_{t-1}(\boldsymbol{\theta}) + \eta_t, \text{ para } t = 1, \dots, n \quad (3.11)$$

em que:

ϕ_1 é o coeficiente autoregressivo de ordem 1; e

η_t é um processo aleatório tal que $\eta_t \sim \mathcal{N}(0, \sigma^2)$, com média nula e desvio-padrão σ constante, nesse caso.

Na Equação 3.11, o valor esperado é dado por $\mathbb{E}[\varepsilon_t(\boldsymbol{\theta})] = 0$, e a variância por $\text{Var}[\varepsilon_t(\boldsymbol{\theta})] = \sigma^2 / (1 - \phi_1^2)$, o que leva à seguinte formulação para a log-verossimilhança (SOROOSHIAN e DRACUP, 1980 *apud* VRUGT *et al.*, 2009a):

$$\begin{aligned} \ell(\boldsymbol{\theta}, \phi_1, \sigma^2 | \tilde{\mathbf{Y}}) = & -\frac{n}{2} \cdot \log(2\pi) - \frac{1}{2} \cdot \log \left[\frac{\sigma^{2n}}{(1-\phi_1^2)} \right] - \frac{1}{2} \cdot (1 - \phi_1)^2 \cdot \sigma^{-2} \cdot [\varepsilon_1(\boldsymbol{\theta})]^2 - \frac{1}{2} \cdot \sigma^{-2} \cdot \\ & \sum_{t=2}^n [\varepsilon_t(\boldsymbol{\theta}) - \phi_1 \cdot \varepsilon_{t-1}(\boldsymbol{\theta})]^2 \end{aligned} \quad (3.12)$$

Na expressão 3.12, as variáveis ϕ_1 e σ^2 são inferidas em conjunto com os parâmetros do modelo, razão pela qual seu posicionamento está à esquerda do vetor $\tilde{\mathbf{Y}}$ no 1º membro da equação. Além disso, essa função de log-verossimilhança pode ser adaptada para modelar resíduos heteroscedásticos e que sigam um processo estocástico autoregressivo de 1ª ordem (vide VRUGT, 2016, eq. 11).

A exemplo da Equação 3.12, foram desenvolvidas inúmeras outras, levando-se em conta, por exemplo, distintos graus de assimetria e curtose para os resíduos (SCHOUPS e VRUGT, 2010; SCHARNAGL *et al.*, 2015).

Apesar desses avanços, subsistem algumas ressalvas quanto à utilização do esquema Bayesiano para a avaliação de incertezas associadas aos parâmetros da modelagem de sistemas ambientais, na forma até agora apresentada. Em primeiro lugar, ainda que se trate de incorporar conceitualmente as incertezas estruturais e de medição de dados por meio da utilização de modelos de erros e de funções de verossimilhança mais elaborados, com a consequente inserção de variáveis latentes, não é possível quantificá-las separadamente mantendo-se a estrutura básica expressa nas Equações 3.3 e 3.5. A consequência é que esses dois tipos de incertezas acabam sendo propagados sobre as incertezas paramétricas, ou sobre as variáveis latentes que exprimem a incerteza sobre as previsões, ou saídas, de determinada modelagem. Assim, são modelos estatísticos adequados a circunstâncias nas quais se pretende avaliar a incerteza preditiva e o peso da incerteza correspondente aos parâmetros sobre a mesma (HUTTON *et al.*, 2011).

Outro ponto importante refere-se à necessidade de validação das hipóteses adotadas quanto aos resíduos, segundo alertam Kavetski *et al.* (2006a; 2006b) e Hutton *et al.* (2011). Isso pode ser realizado por meio de diagnósticos após a estimação Bayesiana, utilizando-se ferramentas como a plotagem de resíduos *versus* as saídas do modelo, gráficos quantil-quantil para se avaliar a normalidade dos resíduos, e de autocorrelação dos mesmos *versus* o *lag* temporal, como realizado por Thyer *et al.* (2009), Schoups e Vrugt (2010) e Silva *et al.* (2014).

Devido às questões supracitadas, nos últimos 15 anos foram empreendidos esforços no sentido de separar as fontes de incertezas paramétricas daquelas agregadas pelas medições de dados de entrada e de saída, e pela estrutura dos modelos matemáticos utilizados. Surgiram assim, abordagens como o BaTEA (*Bayesian Total Error Analysis*), método por meio do qual é possível avaliar separadamente e em conjunto as fontes de incertezas oriundas dos dados e dos parâmetros, assunto a ser tratado em próximo subcapítulo.

Outro aspecto que pode trazer dificuldades à estimação Bayesiana de incertezas é a quantidade de dados $\tilde{Y}$ disponível para sua aplicação (HUTTON *et al.*, 2011). Stedinger *et al.* (2008) demonstraram matematicamente o impacto do comprimento do vetor dessas informações sobre a importância atribuída aos vetores paramétricos θ distintos do vetor que promova a melhor resposta modelada, ou seja, sobre os valores da função de verossimilhança e da distribuição conjunta *a posteriori* apresentada por esses conjuntos, e até mesmo sobre a precisão das estimativas.

3.3.1.3 Representação da incerteza na estimação Bayesiana

Na abordagem Bayesiana, uma das hipóteses acerca dos parâmetros de um modelo qualquer é a de que estes são quantidades desconhecidas, as quais, por serem aleatórias, precisam ser descritas por uma distribuição de probabilidades (COSTA, 2015), descritivas do conhecimento (ou desconhecimento, como atestam alguns autores, como FERNANDES e SILVA, 2017) a seu respeito. Assim, a distribuição conjunta *a posteriori* refletiria o conhecimento acerca dos parâmetros θ após a incorporação dos dados observados sobre o fenômeno em estudo. Em contraposição, para a escola frequentista, as probabilidades representam a variabilidade intrínseca ao processo avaliado e estão relacionadas a frequências de longo termo e geralmente à necessidade de grandes amostras, sobretudo quando a teoria assintótica é utilizada (COSTA, 2015), o que, na maioria dos casos práticos, é impossível de se aplicar. Os parâmetros são tomados por quantidades fixas, ainda que desconhecidas, e que podem ser estimadas a partir de métodos como o de maximização da função de verossimilhança.

À luz da abordagem Bayesiana, uma vez que a distribuição *a posteriori* seja estimada, é possível estabelecer características do vetor paramétrico (COSTA, 2015), como estimadores pontuais e intervalos de credibilidade. Os parágrafos a seguir foram adaptados a partir de Fernandes e Silva (2017) e trazem, segundo estes autores, uma descrição simples dos conceitos envolvidos em estimação paramétrica Bayesiana, elaborada a partir das referências de Bernardo e Smith (1994), e Robert e Casella (2004).

De acordo com Bernardo e Smith (1994), a estimação Bayesiana trata de um problema de decisão, no qual, para se estimar um determinado parâmetro θ , é requerida a especificação de uma função de perda $\wp(\delta, \theta)$, que representa o erro, ou a penalidade, associada à escolha de δ como estimador de θ . Assim, sob essa ótica, busca-se um estimador que minimize o denominado risco de Bayes, definido por:

$$RB = \iint \wp(\delta, \theta) \cdot p(\tilde{Y}|\theta) dy d\theta \quad (3.13)$$

Nesse caso, $p(\tilde{Y}|\theta)$ representa a função de verossimilhança. Uma inversão na ordem de integração permite avaliar o estimador de θ em termos da perda esperada *a posteriori*, ou seja, o estimador δ_B de θ tal que seja válida a seguinte equação:

$$\delta_B = \min E[\wp(\delta, \theta)|\tilde{Y}] = \min_{\theta} \int_{\Theta} \wp(\delta, \theta) \cdot p(\theta|\tilde{Y}) dy \quad (3.14)$$

em que $p(\theta|\tilde{Y})$ é a distribuição de probabilidades marginais *a posteriori* do parâmetro θ .

A escolha da função de perda é feita de forma subjetiva. De acordo com Bernardo e Smith (1994), as principais funções de perda utilizadas na estimação de parâmetros são as seguintes:

- Quadrática: nesse caso tem-se que $\wp(\delta, \theta) = (\delta, \theta)^2$ e o estimador de Bayes para θ é a média *a posteriori* $\mathbb{E}[p(\theta|\tilde{Y})]$, admitindo-se que a média exista;
- Valor absoluto: nesse caso tem-se que $\wp(\delta, \theta) = |\delta, \theta|$ e o estimador de Bayes para θ é a mediana de $p(\theta|\tilde{Y})$, sob a hipótese de que a mediana exista; e
- Zero-um: nesse caso tem-se que $\wp(\delta, \theta) = \mathbf{1}_{(\delta \neq \theta)}$, onde $\mathbf{1}_{(a)}$ é a função indicadora, e o estimador de Bayes para θ é a moda de $p(\theta|\tilde{Y})$, desde que este descritor exista.

Conforme atestam Robert e Casella (2004), o cálculo do estimador δ pode prescindir de métodos de cálculo aproximativos, devido ao fato de a distribuição *a posteriori* $p(\theta|\tilde{Y})$ não apresentar de maneira geral uma forma analítica fechada, dificultando assim a integração direta da Equação 3.14. Os métodos apresentados no subitem seguinte, utilizados para se estimar $p(\theta|\tilde{Y})$, viabilizam a obtenção dos descritores δ supracitados.

Outra forma de se representar a incerteza presente em θ é a estimação por intervalos de credibilidade, o equivalente Bayesiano dos intervalos de confiança, obtidos quando se utiliza a abordagem frequentista no estudo de probabilidades. O intervalo de credibilidade de $100 \cdot (1 - \alpha)\%$ para um parâmetro θ é construído com base em sua distribuição marginal *a posteriori*, levando em conta a única amostra que foi de fato observada (o vetor $\tilde{Y}$, no caso). Como na escola Bayesiana o parâmetro é considerado um objeto aleatório, a interpretação dada ao intervalo de credibilidade é intuitiva, ou seja, é o intervalo no qual θ tem probabilidade $(1 - \alpha)$ de estar ali contido (FERNANDES e SILVA, 2017). Embora se faça um paralelo aos intervalos de confiança, os conceitos embutidos em sua obtenção atribuem-lhes uma interpretação distinta. Para a escola frequentista, o desempenho de estimadores e de seus correspondentes intervalos de confiança é abordado à luz da princípio de repetição da amostra, sob condições idênticas. Assim, a confiança de um intervalo, $100 \cdot (1 - \alpha)\%$, representa o percentual de intervalos que, construídos a partir de dados coletados sob condições rigorosamente idênticas às da amostra efetivamente disponível, conteriam o verdadeiro valor de θ (COSTA, 2015; FERNANDES e SILVA, 2017).

O intervalo de credibilidade deve ser construído a partir do conceito expresso a seguir:

$$\int_L^U p(w) dw = 1 - \alpha \quad (3.15)$$

Na Equação 3.15, w é uma quantidade aleatória, $p(w)$ é a sua distribuição de probabilidades (*a priori*, *a posteriori*, ou preditiva *a posteriori*), e L e U são, respectivamente, os limites inferior e superior do intervalo de credibilidade, o qual, na abordagem Bayesiana, representa o intervalo de mais alta densidade de probabilidade de conter w , a um nível $100 \cdot (1 - \alpha)\%$.

Para um quantil fixo α , não há um valor único para os limites L e U , mesmo que $p(w)$ seja unimodal. Dessa forma, adota-se o conceito do intervalo de mais alta densidade, traduzido do termo em inglês *Highest Probability Density* (HPD). De acordo com Bernardo e Smith (1994 *apud* FERNANDES e SILVA, 2017), um intervalo $I \subseteq \Omega$, onde Ω é o domínio de w , é o intervalo HPD, a um nível $100 \cdot (1 - \alpha)\%$ para w , com relação a $p(w)$, se:

$$P(w \in I) = 1 - \alpha \quad (3.16)$$

$$p(w_1) \geq p(w_2), \forall w_1 \in I \text{ e } w_2 \notin I \quad (3.17)$$

De acordo com Fernandes e Silva (2017), “o intervalo HPD é o mais curto dos intervalos que determinam a área abaixo da função densidade $p(w)$ igual a $(1 - \alpha)$ ”. Quanto mais curto for o intervalo HPD, mais certeza se tem sobre w .

A abordagem Bayesiana também fornece ferramental adequado para se estimar a incerteza associada à predição de valores futuros da variável $\mathbf{Y}$ pela modelagem sob avaliação, por meio da integração da distribuição conjunta *a posteriori* ao longo do espectro de valores do vetor paramétrico $\boldsymbol{\theta}$ (TODINI, 2007; STEDINGER *et al.*, 2008; SCHOUPS e VRUGT, 2010; CAMACHO *et al.*, 2015; SILVA *et al.*, 2017). Para tanto, utiliza-se a chamada distribuição preditiva *a posteriori*, representada pela expressão a seguir, cuja notação foi adaptada a partir das referências supracitadas para se adequar àquela adotada no presente texto:

$$f(\mathbf{Y}|\tilde{\mathbf{Y}}) = \int f(\mathbf{Y}|\boldsymbol{\theta}) \cdot p(\boldsymbol{\theta}|\tilde{\mathbf{Y}}) \quad (3.18)$$

na qual $p(\boldsymbol{\theta}|\tilde{\mathbf{Y}})$ denota a distribuição conjunta *a posteriori* dos parâmetros do vetor $\boldsymbol{\theta}$ e $f(\mathbf{Y}|\boldsymbol{\theta})$ representa o modelo em análise, $\mathfrak{Z}$. Em outras palavras, a resposta do modelo $\mathfrak{Z}$ deve ser

integrada ao longo de todas as possíveis realizações do vetor paramétrico estimadas após a observação dos dados $\tilde{\mathbf{Y}}$. De acordo com Silva *et al.* (2017), a aplicação da Equação 3.18 permite incorporar à predição de eventos futuros as incertezas associadas à variabilidade inerente ao processo físico modelado e ao procedimento de estimação de parâmetros, combinando incertezas de caráter aleatório e epistêmico às futuras realizações da variável $\mathbf{Y}$.

Todini (2008), por sua vez, define conceitualmente a incerteza preditiva como a probabilidade de qualquer valor futuro de um *output*, condicionada a todo o conhecimento e informação disponíveis até o momento presente, e agregados à definição da variável de interesse por meio de um processo contínuo de aprendizado.

Em se tratando da modelagem de sistemas ambientais, a exemplo dos modelos hidrodinâmicos, o cálculo analítico da integral na Equação 3.18 é inviável, dado o elevado número de parâmetros em $\boldsymbol{\theta}$, e a natureza não linear dos modelos utilizados. Desse modo, é prática comum o uso de técnicas numéricas aproximativas, como as que serão apresentadas no subcapítulo 3.4. Sob tal condição, a Equação 3.18 pode então ser avaliada de modo discreto por meio da expressão a seguir (adaptada de Camacho *et al.*, 2015):

$$f(\mathbf{Y}|\tilde{\mathbf{Y}}) = \sum_{i=1}^m f(\mathbf{Y}|\boldsymbol{\theta}_i) \cdot p(\boldsymbol{\theta}_i|\tilde{\mathbf{Y}}) \quad (3.19)$$

na qual m é o número de realizações do vetor paramétrico $\boldsymbol{\theta}$ utilizadas na estimação de sua distribuição conjunta *a posteriori*.

3.3.2 O método BaTEA

O método BaTEA (*Bayesian Total Error Analysis*) foi inicialmente elaborado por Kavetski *et al.* (2003) e propõe a separação formal e a quantificação das incertezas atreladas aos dados de entrada e aos parâmetros do modelo em uma estrutura Bayesiana, com o ideal de obtenção de um melhor ajuste da distribuição conjunta *a posteriori* dos parâmetros da modelagem, o que pode ser usado para uma predição com intervalos de credibilidade melhor definidos, sob a alegação de que a presença de erros, tanto nos dados de entrada quanto de saída dos modelos, introduz viés nos parâmetros. Segundo seus precursores, a não consideração dos erros em tais dados, seja por métodos tradicionais de calibração, ou mesmo sob a ótica Bayesiana, pode agregar um viés imprevisível às estimativas paramétricas e de suas respectivas incertezas, reduzindo a confiabilidade da reprodução do comportamento do sistema e da predição.

Outra vantagem defendida por seus idealizadores é a de que a distinção e a quantificação das distintas fontes de erros podem contribuir para reduzir a ambiguidade acerca das incertezas preditivas e dos parâmetros e melhorar os procedimentos de calibração/validação de modelos de sistemas ambientais.

Os próximos parágrafos demonstram a dedução das equações principais que compõem o método, extraídas e adaptadas a partir da referência de Kavetski *et al.* (2003) e acompanhadas pela Figura 3.3.

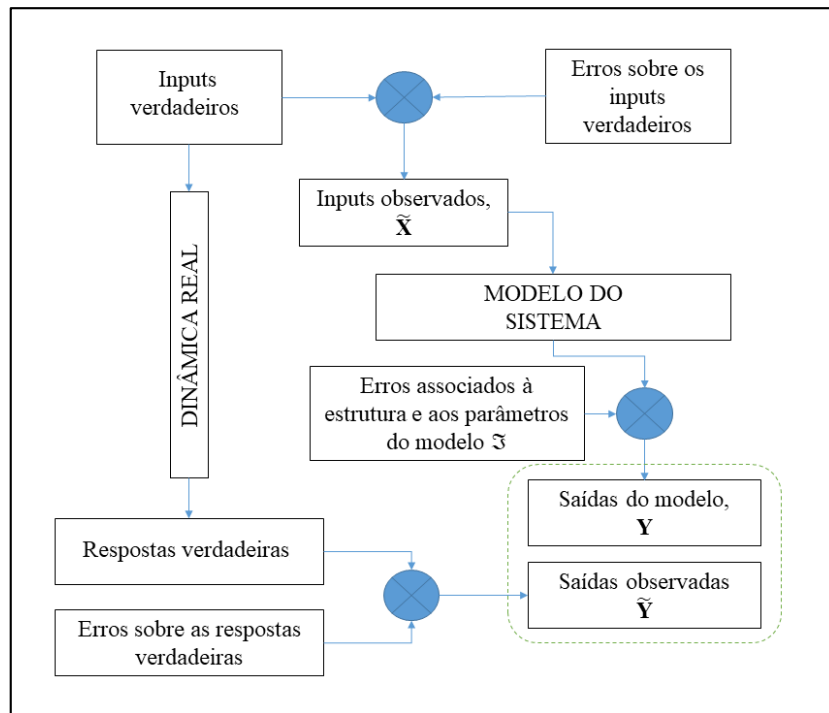


Figura 3.3 – Diagrama explicativo do BaTEA e das premissas para sua dedução (adaptado de Kavetski *et al.*, 2003).

Considerando-se que os vetores $\mathbf{X}$ (entradas) e $\mathbf{Y}$ (saídas) contêm erros de amostragem e de medição, presentes em seus pares observados $\tilde{\mathbf{X}}$ e $\tilde{\mathbf{Y}}$, e que $\mathbf{Y}$ agrega também erros inerentes aos parâmetros e à estrutura do modelo $\mathfrak{Z}$, o Teorema de Bayes expresso na Equação 3.3 pode ser reescrito da seguinte maneira (o conjunto Ω será suprimido à semelhança do que foi realizado nas equações a partir da 3.5):

$$p(\boldsymbol{\theta}|\tilde{\mathbf{X}},\tilde{\mathbf{Y}}) = \frac{p(\tilde{\mathbf{X}},\tilde{\mathbf{Y}}|\boldsymbol{\theta}) \cdot p(\boldsymbol{\theta})}{p(\tilde{\mathbf{X}},\tilde{\mathbf{Y}})} \quad (3.20)$$

na qual:

$p(\boldsymbol{\theta}|\tilde{\mathbf{X}}, \tilde{\mathbf{Y}})$ é a distribuição de probabilidades conjuntas *a posteriori* do vetor paramétrico, após a incorporação da informação representada pelas observações $\tilde{\mathbf{X}}$ e $\tilde{\mathbf{Y}}$;
 $p(\boldsymbol{\theta})$ é a distribuição de probabilidades conjuntas *a priori* do vetor paramétrico;
 $p(\tilde{\mathbf{X}}, \tilde{\mathbf{Y}}|\boldsymbol{\theta}) \equiv L(\boldsymbol{\theta}|\tilde{\mathbf{X}}, \tilde{\mathbf{Y}})$ denota a função de verossimilhança; e
 $p(\tilde{\mathbf{X}}, \tilde{\mathbf{Y}})$ é a distribuição preditiva *a priori* para as observações $\tilde{\mathbf{X}}$ e $\tilde{\mathbf{Y}}$.

Assim como realizado com o denominador da Equação 3.3, o termo $p(\tilde{\mathbf{X}}, \tilde{\mathbf{Y}})$ não precisa ser avaliado explicitamente e pode ser absorvido como constante de proporcionalidade. Dessa forma, segue que:

$$p(\boldsymbol{\theta}|\tilde{\mathbf{X}}, \tilde{\mathbf{Y}}) \propto L(\boldsymbol{\theta}|\tilde{\mathbf{X}}, \tilde{\mathbf{Y}}) \cdot p(\boldsymbol{\theta}) \quad (3.21)$$

Kavetski *et al.* (2003) propuseram então que o vetor $\mathbf{X}$ fosse tratado pelo mesmo status lógico que o vetor $\boldsymbol{\theta}$, a respeito do qual se tem somente um conhecimento prévio, a fim de que suas incertezas pudessem ser separadas em relação às deste e quantificadas. Sob tal perspectiva, a Equação 3.21 pode ser reescrita como:

$$p(\boldsymbol{\theta}, \mathbf{X}|\tilde{\mathbf{X}}, \tilde{\mathbf{Y}}) \propto L(\boldsymbol{\theta}, \mathbf{X}|\tilde{\mathbf{X}}, \tilde{\mathbf{Y}}) \cdot p(\boldsymbol{\theta}, \mathbf{X}) \quad (3.22)$$

Na Equação 3.22, $L(\boldsymbol{\theta}, \mathbf{X}|\tilde{\mathbf{X}}, \tilde{\mathbf{Y}}) \equiv p(\tilde{\mathbf{X}}, \tilde{\mathbf{Y}}|\boldsymbol{\theta}, \mathbf{X})$ pode ser interpretada como a probabilidade conjunta de se observarem os vetores $\tilde{\mathbf{X}}$ e $\tilde{\mathbf{Y}}$ dado o vetor $\boldsymbol{\theta}$ e os dados de entrada “reais”, $\mathbf{X}$ (KAVETSKI *et al.*, 2003).

A função de verossimilhança $L(\boldsymbol{\theta}, \mathbf{X}|\tilde{\mathbf{X}}, \tilde{\mathbf{Y}})$ pode ser subdividida em dois modelos probabilísticos, um para os verdadeiros *inputs*, $\mathbf{X}$, e o outro para os *outputs*, $\mathbf{Y}$. Assim, aquele que representa estes últimos incorpora as incertezas inerentes à estrutura e aos parâmetros do modelo, bem como os erros de medição nos dados de saída. Daí reside a principal diferença entre os métodos Bayesianos formais na maneira até aqui apresentada, geralmente focada sobre os parâmetros da modelagem, e o BaTEA, uma vez que, por meio deste, é possível separar as incertezas dos dados de entrada em relação aos demais componentes, com o objetivo primordial de garantir uma identificação mais adequada das incertezas paramétricas. À luz dessa premissa, a verossimilhança pode ser expressa como:

$$L(\boldsymbol{\theta}, \mathbf{X}|\tilde{\mathbf{X}}, \tilde{\mathbf{Y}}) \propto L(\tilde{\mathbf{X}}, \boldsymbol{\theta}, \mathbf{X}|\tilde{\mathbf{Y}}) \cdot L(\boldsymbol{\theta}, \mathbf{X}|\tilde{\mathbf{X}}) \quad (3.23)$$

Kavetski *et al.* (2003) propuseram então algumas simplificações, baseando-se nas seguintes hipóteses:

- $\tilde{\mathbf{X}}$ e $\tilde{\mathbf{Y}}$ são estatisticamente independentes, ou seja, $\tilde{\mathbf{Y}}$ depende somente dos erros de medição/observação que afetam as saídas do modelo, $\mathbf{Y}$, e dos verdadeiros valores das entradas, $\mathbf{X}$ (vide o esquema da Figura 3.3). Os erros que atuam sobre os dados de entrada não o fazem diretamente sobre as verdadeiras respostas do sistema e seus respectivos erros de medição/observação; e
- $\tilde{\mathbf{X}}$ é estatisticamente independente de $\boldsymbol{\theta}$, ou seja, os erros que interferem em $\mathbf{X}$ não são correlacionados aos parâmetros do modelo matemático $\mathfrak{T}$.

Sendo assim, a verossimilhança $L(\boldsymbol{\theta}, \mathbf{X} | \tilde{\mathbf{X}}, \tilde{\mathbf{Y}})$ torna-se tal que:

$$L(\boldsymbol{\theta}, \mathbf{X} | \tilde{\mathbf{X}}, \tilde{\mathbf{Y}}) \propto L(\boldsymbol{\theta}, \mathbf{X} | \tilde{\mathbf{Y}}) \cdot L(\mathbf{X} | \tilde{\mathbf{X}}) \quad (3.24)$$

Incorporando-se a Equação 3.24 à 3.22, chega-se à formulação básica do BaTEA, reescrita como:

$$p(\boldsymbol{\theta}, \mathbf{X} | \tilde{\mathbf{X}}, \tilde{\mathbf{Y}}) \propto L(\boldsymbol{\theta}, \mathbf{X} | \tilde{\mathbf{Y}}) \cdot L(\mathbf{X} | \tilde{\mathbf{X}}) \cdot p(\boldsymbol{\theta}, \mathbf{X}) \quad (3.25)$$

na qual:

$p(\boldsymbol{\theta}, \mathbf{X} | \tilde{\mathbf{X}}, \tilde{\mathbf{Y}})$ é a distribuição de probabilidades conjuntas *a posteriori* do vetor paramétrico e dos verdadeiros valores de $\mathbf{X}$, após a incorporação da informação representada pelas observações $\tilde{\mathbf{X}}$ e $\tilde{\mathbf{Y}}$;

$p(\boldsymbol{\theta}, \mathbf{X})$ é a distribuição de probabilidades conjuntas *a priori* do vetor paramétrico e do vetor $\mathbf{X}$;

$L(\boldsymbol{\theta}, \mathbf{X} | \tilde{\mathbf{Y}})$ refere-se à função de verossimilhança correspondente ao modelo de erros da saída $\tilde{\mathbf{Y}}$, o qual representa a incerteza associada à modelagem e aos dados de saída; e

$L(\mathbf{X} | \tilde{\mathbf{X}})$ é a função de verossimilhança oriunda do modelo de erros dos inputs $\tilde{\mathbf{X}}$ em relação aos seus verdadeiros valores, $\mathbf{X}$.

A partir da proposição teórica expressa em 3.24, os autores adaptaram-na à modelagem chuva-vazão de cunho conceitual com o intuito de quantificar separadamente as incertezas nos dados de entrada, representados, no caso, pelo vetor de alturas pluviométricas no intervalo de tempo considerado, e as devidas aos parâmetros do modelo (KAVETSKI *et al.*, 2006a). Para tanto, consideraram o conceito de multiplicadores de chuva aplicados aos eventos de chuva extraídos do vetor $\tilde{\mathbf{X}}$, denotado por:

$$\mathbf{X} \leftarrow g(\tilde{\mathbf{X}}, \boldsymbol{\varphi}) \quad (3.26)$$

Na expressão anterior, $\boldsymbol{\varphi}$ é o vetor de multiplicadores de chuva, denotado por $\mathbf{m} = \{m_z, z = 1, \dots, N_m\}$, em que m_z é o erro multiplicativo para a altura de chuva da z -ésima tormenta (ou evento). Esse modelo pressupõe que o erro deve ser aplicado de maneira individual sobre cada evento chuvoso, sob a forma:

$$\tilde{h}_z = m_z \cdot h_z, \quad (3.27)$$

em que $\tilde{h}_z$ e h_z são as alturas totais de chuva observadas e verdadeiras para o i -ésimo evento pluviométrico.

Considere-se também que $\{\boldsymbol{\beta}_x, \boldsymbol{\beta}_y\}$ sejam os vetores de parâmetros dos modelos de erros, e das correspondentes funções de verossimilhanças, estipulados para as entradas e saídas do modelo em questão. A partir da equação genérica do BaTEA em 3.25 e das hipóteses mencionadas nos três últimos parágrafos, Kavestki *et al.* (2006a) chegaram à seguinte formulação:

$$p(\boldsymbol{\theta}, \boldsymbol{\varphi}, \boldsymbol{\beta}_x, \boldsymbol{\beta}_y | \tilde{\mathbf{X}}, \tilde{\mathbf{Y}}) \propto L(\boldsymbol{\theta}, \tilde{\mathbf{X}}, \boldsymbol{\varphi}, \boldsymbol{\beta}_y | \tilde{\mathbf{Y}}) \cdot L(\boldsymbol{\beta}_x | \boldsymbol{\varphi}) \cdot p(\boldsymbol{\theta}) \cdot p(\boldsymbol{\beta}_y) \cdot p(\boldsymbol{\beta}_x) \quad (3.28)$$

na qual:

$L(\boldsymbol{\beta}_x | \boldsymbol{\varphi})$ representa a função de verossimilhança correspondente ao modelo de erros do vetor $\boldsymbol{\varphi}$ de multiplicadores m_z , a qual contém os parâmetros β_x . Os demais termos são análogos a seus pares na Equação 3.25.

Os autores explicam que as distribuições *a priori* $p(\boldsymbol{\beta}_x)$ e $p(\boldsymbol{\beta}_y)$ devem refletir o grau de conhecimento acerca dos modelos de erros pressupostos para as entradas e saídas. Se o usuário possui um bom conhecimento sobre possíveis erros de medição e de observação nos dados utilizados, a exemplo de informações sobre curvas-chave ou estimadas a partir de registros de pluviômetros, esse aspecto deve ser traduzido de alguma forma sobre tais distribuições. Outro ponto importante refere-se a $L(\boldsymbol{\beta}_x | \boldsymbol{\varphi})$, cujas probabilidades devem concentrar-se mais ao redor de um único valor à medida que os dados de entrada sejam mais precisos.

Os autores ressaltam que a separação das incertezas sob uma estruturação formal a partir do Teorema de Bayes resulta no acréscimo de variáveis latentes (GULL, 1989 *apud* KAVETSKI *et al.*, 2003) no processo de inferência, tais como a variância e a média dos multiplicadores. Além disso, se as incertezas nos dados de entrada e de saída forem muito amplas, a estimação

da variabilidade dos parâmetros torna-se questionável, principalmente se o modelo selecionado não for adequado ao sistema representado, em termos de sua estrutura.

Para testar a aplicabilidade do BaTEA à modelagem hidrológica, Kavetski *et al.* (2003) calibraram o modelo chuva-vazão ABC (FIERING, 1967 *apud* VOGEL e SANKARASUBRAMANIAN, 2003) a partir de métodos Bayesianos com hipóteses distintas acerca das fontes de erros e da função de verossimilhança: (i) incertezas presentes somente na resposta do modelo, e distribuição normal de tais erros para compor a função de verossimilhança; e (ii) incertezas nas entradas e saídas do modelo, com separação das funções de verossimilhança conforme proposta do BaTEA; no caso, foram empregadas distribuições normais com diferentes médias e variâncias para ambas. Para checar as premissas adotadas, aplicaram a metodologia a um estudo de caso hipotético, em que foram gerados eventos de chuva (entradas) “verdadeiros” a partir de um gerador estocástico (HENEKER *et al.*, 2001 *apud* KAVETSKI *et al.*, 2003). Cada um desses eventos recebeu um erro multiplicativo sorteado a partir de uma distribuição normal com média 1 e variância conhecida (daí a denominação multiplicadores de chuva para tais erros) a fim de emular as entradas observadas. A resposta do modelo foi calculada aplicando-se os eventos de chuva ao modelo com um conjunto de parâmetros pré-definido (o conjunto “real” de parâmetros), e a esta foi incorporado um modelo aditivo Gaussiano para obtenção da série de vazões observadas, $\tilde{Y}$. Assim, para a 2ª hipótese descrita, o vetor de entradas “observadas” $\tilde{X}$ (ou seja, sintéticas para este caso) foi obtido aplicando-se a seguinte formulação:

$$\tilde{h}_z = m_z \cdot h_z, \text{ com } m \sim \mathcal{N}(1, \sigma_m^2), \forall z, \quad (3.29)$$

em que σ_m^2 é a variância dos multiplicadores, nesse caso, conhecida e pré-fixada.

O vetor de dados de saída observados sintéticos $\tilde{Y}$ foi produzido a partir de:

$$\tilde{y}_t = y_t + \varepsilon_t, \text{ com } \varepsilon_t \sim \mathcal{N}(0, \sigma_t^2), \forall t, \quad (3.30)$$

com σ_t^2 conhecido e estabelecido pelos autores.

Os autores concluíram que a prescrição de um modelo adequado de erros para as entradas do modelo, por meio da inserção no Teorema de Bayes de uma função de verossimilhança para elas (no caso, assumindo-se um comportamento segundo uma $\mathcal{N}(1, \sigma_m^2)$, de acordo com a 2ª hipótese anteriormente descrita) resultou em intervalos de credibilidade para os parâmetros que

continham os valores pré-fixados e sua moda mais próxima a estes, ao contrário do que ocorreu ao se desconsiderarem os eventos de chuva como fontes de erro à modelagem – a 1ª hipótese supracitada e geralmente adotada em inúmeras aplicações práticas em calibração de incertezas paramétricas.

Ainda que o BaTEA represente um possível avanço no estudo de incertezas em modelagem de sistemas ambientais, é um método que impõe desafios e vários aspectos por elucidar. Kavetski *et al.* (2003) mencionam, por exemplo, que a escolha inadequada do modelo de erros dos dados de entrada, e assim de sua verossimilhança, pode acarretar a exploração inadequada do espaço paramétrico, com resultados semelhantes aos obtidos ao se ignorar a existência de incertezas nas entradas, premissa que ainda prevalece em estudos de incertezas em modelagem hidrológica e hidráulica. Nas avaliações contidas no referido trabalho, os autores constataram que o uso de uma distribuição uniforme como distribuição *a priori* para a variância dos multiplicadores de chuva e, genericamente, de distribuições não informativas com variância fixa (i.e., segundo uma distribuição do tipo Delta de Dirac), tem o potencial de transformá-los em meros graus de liberdade adicionais que, no mínimo, intensificam o esforço computacional da resolução numérica das Equações 3.22 ou 3.25 e, na pior das hipóteses, desestabilizam o algoritmo de inferência Bayesiana. Uma tentativa de solucionar esse problema foi proposta anos depois por esses estudiosos (KAVETSKI *et al.*, 2006a), ao sugerirem que a variância dos multiplicadores de chuva poderia ser utilizada como variável latente ao processo de inferência, com a atribuição de uma distribuição *a priori* vaga para a mesma – a Gama inversa, com parâmetros de forma e escala pré-definidos e ambos de valor positivo. Nessa proposição, a média dos multiplicadores, por sua vez, foi suposta como sendo igual à unidade, indicando ausência de viés.

Uma desvantagem do BaTEA é a introdução de variáveis latentes ao processo de inferência Bayesiana, referentes à inserção dos dados de entrada como fonte de incertezas e em quantidade proporcional ao tamanho da série temporal utilizada, o que agrega esforço computacional aos algoritmos de resolução numérica. Para contornar essa desvantagem, Kavetski *et al.* (2003) sugeriram a adoção de um multiplicador para cada evento de chuva, alternativamente à aplicação de um erro por altura de chuva no intervalo temporal considerado nas simulações. Essa proposta foi desenvolvida em maior detalhe anos depois, e discutida nos trabalhos de Kavetski *et al.* (2006a, 2006b), nos quais os autores exploraram o uso de distintas verossimilhanças e distribuições *a priori* para os parâmetros representativos dos multiplicadores de eventos de chuva.

Ainda que o BaTEA agregue complexidade matemática e computacional à análise de incertezas, já que insere variáveis latentes ao processo, seus precursores defendem sua aplicação na tentativa de se reduzir o viés embutido nas estimativas paramétricas calculadas a partir da calibração realizada por meio de métodos Bayesianos clássicos, que não contabilizam o erro dos dados de entrada. Segundo eles, o viés ocorre mesmo que sejam utilizadas funções de verossimilhança que embutam a autocorrelação dos resíduos da modelagem e/ou sua heteroscedasticidade (Kavetski *et al.*, 2003). Em adição, sugerem que a estrutura teórica do método permite acomodar fácil e diretamente vários modelos probabilísticos de incertezas (i.e., distintas verossimilhanças), que representem cada uma ou algumas das diversas fontes de erros na modelagem (i.e., séries de entradas e saídas, parâmetros, estrutura do modelo). Mesmo que o processo de traçar hipóteses acerca dessas distintas origens agregue ainda subjetividade ao processo de calibração de parâmetros e de inferência sobre as incertezas envolvidas em modelagem, Kavetski *et al.* (2006a) defendem que o BaTEA, por permitir a inserção explícita de tais premissas, garante sua confirmação *a posteriori*, ao contrário de outros métodos, inclusive de cunho Bayesiano. Para eles, um dos maiores desafios do método é a estipulação de funções de verossimilhança representativas dos erros nos dados de entrada, e de distribuições *a priori* para seus parâmetros, visando agregar-lhes conhecimento técnico e significado físico. Para ilustrar essa dificuldade, os autores citam o caso de uma modelagem chuva-vazão conceitual em duas bacias com características físicas semelhantes e condições distintas quanto à disponibilidade de dados de entrada: em uma delas, há uma ampla e extensa rede de monitoramento hidrometeorológico, enquanto que, na outra, não há sequer um pluviômetro. Como representar matematicamente as incertezas dos registros pluviométricos em cada uma das bacias a fim de calibrar os parâmetros do modelo e as respectivas incertezas sob o formalismo Bayesiano?

Por fim, outro ponto a ser elucidado futuramente acerca do BaTEA é a sua aplicabilidade à modelagem hidrodinâmica, tema para o qual não foram encontrados estudos até o momento.

3.3.3 Métodos pseudo-Bayesianos

Embora os métodos Bayesianos formais pareçam promissores, já que utilizam medidas com esse cunho para descrever os resíduos da modelagem, os mesmos podem apresentar limitações em função da insuficiência de dados necessários à sua aplicação (HUTTON *et al.*, 2011) e à validação das hipóteses acerca do comportamento esperado para os erros. Em função dessas

dificuldades surgiu o GLUE – *Generalized Likelihood Uncertainty Estimation* (BEVEN e BINLEY, 1992), uma extensão do método GSA – *Generalized Sensitivity Analysis*, introduzido por Spear e Hornberger (1980 *apud* STEDINGER *et al.*, 2008). Sua proposta era garantir a quantificação de incertezas por meio de funções de verossimilhança informais, respaldadas no princípio da equifinalidade (BEVEN, 2006). Dentre tais medidas de desempenho, a função inversa da variância dos resíduos $\mathbf{E}(\boldsymbol{\theta}) = \tilde{\mathbf{Y}} - \mathbf{Y}(\boldsymbol{\theta}) = \{\varepsilon_1(\boldsymbol{\theta}), \dots, \varepsilon_n(\boldsymbol{\theta})\}$ é geralmente empregada, conforme proposição inicial de Beven e Binley (1992), tal que:

$$L(\boldsymbol{\theta}|\tilde{\mathbf{Y}}) = \left(\frac{\sum_{j=1}^n |\varepsilon_j(\boldsymbol{\theta})|^2}{n-2} \right)^{-K} \quad (3.31)$$

Na Equação 3.31, K é um parâmetro definido pelo usuário, chamado de fator de forma por Beven e Binley (1992), que lhe atribuíram inicialmente o valor da unidade. Se $K=0$, então cada simulação terá peso igual, e, à medida que $K \rightarrow \infty$, o peso prevalece sobre o conjunto de parâmetros $\boldsymbol{\theta}$ com melhor desempenho.

Outras métricas podem ser utilizadas à guisa de função de verossimilhança, a exemplo de critérios de aderência da resposta simulada em relação às observações disponíveis. Segundo seus idealizadores, a escolha dessa função está sujeita à subjetividade e influenciará, por consequência, os intervalos atribuídos ao vetor $\boldsymbol{\theta}$. Dessa forma, atestam que a expressão adotada precisa ser objeto de discussão e de justificativa.

Um critério de desempenho muito utilizado para otimização de parâmetros em modelagem de sistemas ambientais é mencionado por Stedinger *et al.* (2008) como função de verossimilhança informal utilizada com frequência junto ao GLUE aplicado a esses modelos, condição em que a equação original é adaptada para incorporar o conceito do fator de forma K :

$$L(\boldsymbol{\theta}|\tilde{\mathbf{Y}}) = \left[1 - \frac{\sum_{t=1}^n (\tilde{y}_t - y_t)^2}{\sum_{t=1}^n (\tilde{y}_t - \bar{\tilde{y}})^2} \right]^K \quad (3.32)$$

na qual $\bar{\tilde{y}}$ é a média amostral do vetor de observações $\tilde{\mathbf{Y}}$.

Outra função que pode ser usada é a apresentada por Beven e Freer (2001 *apud* STEDINGER *et al.*, 2008) como uma adaptação da expressão anterior:

$$L(\boldsymbol{\theta}|\tilde{\mathbf{Y}}) = \exp \left[-K \cdot \frac{\sum_{t=1}^n (\tilde{y}_t - y_t)^2}{\sum_{t=1}^n (\tilde{y}_t - \bar{\tilde{y}})^2} \right] \quad (3.33)$$

Jung (2011) testou a sensibilidade dos intervalos de credibilidade das áreas inundadas previstas pelo modelo hidráulico HEC-RAS, em dois trechos fluviais, a partir da utilização de duas equações distintas operando como função de verossimilhança informal na estrutura do GLUE:

$$L_W = \sum_{r=1}^R (W_r - \tilde{W}_r)^2 \quad (3.34)$$

$$L_E = \sum_{r=1}^R (E_r - \tilde{E}_r)^2 \quad (3.35)$$

nas quais:

W_r e $\tilde{W}_r$ são as larguras de topo em cada seção transversal r , simulada pelo modelo e extraída a partir da mancha de inundação observada para uma cheia excepcional, respectivamente; Similarmente, E_r e $\tilde{E}_r$ representam, para cada seção r , as elevações da linha de água simulada e extraída a partir do cruzamento da mancha de inundação com o modelo digital de terreno que serve de base para a definição das áreas inundadas;

R é o número total de seções transversais utilizado no modelo hidráulico; e

L_W e L_E são as verossimilhanças informais avaliadas por meio das larguras e elevações, respectivamente.

Outro exemplo de função de verossimilhança informal utilizada por Jung (2011) baseia-se na extensão da área inundada, conforme disposto na equação a seguir:

$$L_F = \left(\frac{S_{obs,sim}}{\tilde{S} + S - S_{obs,sim}} \right) \times 100 \quad (3.36)$$

em que:

$\tilde{S}$ é a área inundada observada, i.e., contabilizada a partir da mancha de inundação disponível e S é a área inundada computada em cada simulação do modelo;

$S_{obs,sim}$ representa a interseção entre as áreas inundadas observada e simulada; e

L_F é a função de verossimilhança informal baseada na extensão da área inundada.

Stedinger *et al.* (2008) mencionam ainda a possibilidade de se combinarem diversas funções de verossimilhança informais na tentativa de melhor identificar conjuntos paramétricos válidos.

O método GLUE pode ser resumido pelas seguintes etapas, conforme descrição de Vrugt *et al.* (2009a) e de Hutton *et al.* (2011), acompanhada pela representação esquemática adaptada de Mirzaei *et al.* (2015), conforme mostra a Figura 3.4:

- Especificação das distribuições *a priori* para cada parâmetro a ser considerado;

- Seleção de um método para amostragem dos parâmetros a partir das distribuições *a priori* e sua aplicação ao modelo, com posterior cômputo da função de verossimilhança informal escolhida para cada conjunto paramétrico;
- Definição de um limite para separação entre as respostas do modelo consideradas válidas ou não, que, na nomenclatura de Beven e Binley (1992), recebem a denominação de *behavioural* e *non-behavioural*. Pode-se utilizar tanto um valor de corte para a função de verossimilhança quanto uma porcentagem fixa em relação ao número total de simulações. Também pode ser adotado um critério de aderência das observações em relação às simulações, distinto da função de verossimilhança informal;
- Normalização dos valores de verossimilhança das simulações válidas para que sua soma se equipare à unidade; e
- Classificação dos conjuntos paramétricos e das respectivas probabilidades para obtenção da distribuição de probabilidades conjuntas das saídas do modelo, e seu uso para geração de intervalos de credibilidade.

Devido à flexibilidade de alguns desses procedimentos e à maior facilidade de adoção verificada em distintas áreas do conhecimento, quando comparado aos métodos Bayesianos formais, o GLUE foi extensamente aplicado à modelagem hidrológica e hidráulica, prevalecendo, nesse último caso, sobre as demais técnicas disponíveis em estudos de quantificação de incertezas. Blasone *et al.* (2008) atribuíram a preferência do GLUE em muitos estudos de modelagem de sistemas ambientais à sua simplicidade conceitual, à facilidade de implantação, e à possibilidade de absorver diferentes fontes de informação que podem ser combinadas a fim de definir a medida de verossimilhança.

Apesar dos diversos exemplos de aplicação, o GLUE tem sido alvo constante de críticas por parte de alguns grupos acadêmicos, a exemplo dos trabalhos conduzidos por Montanari (2005), Mantovan e Todini (2006), Stedinger *et al.* (2008) e Vrugt *et al.* (2009a). Dentre os aspectos destacados como desvantajosos devido ao seu potencial de conduzir à incorreta prescrição das incertezas e dos respectivos intervalos de credibilidade, em razão da subjetividade que os permeia, prevalecem: (i) a escolha da função de verossimilhança, que, devido ao seu caráter informal, geralmente não caracteriza adequadamente os resíduos da modelagem; (ii) a seleção do valor de K ; e (iii) a definição do limite de separação entre as saídas dos modelos válidas ou não. Adicionalmente, ainda que o GLUE permita inserir as incertezas de dados de entrada sob

o mesmo status das incertezas paramétricas, com a sujeição de ambas à mesma função de verossimilhança, tal procedimento pode acarretar sua identificação incorreta. Essa ressalva vale também para os métodos Bayesianos tradicionais, tratados no item 3.3.1, sob circunstâncias de utilização de funções de verossimilhança que descrevam incorretamente as características estatísticas dos resíduos.

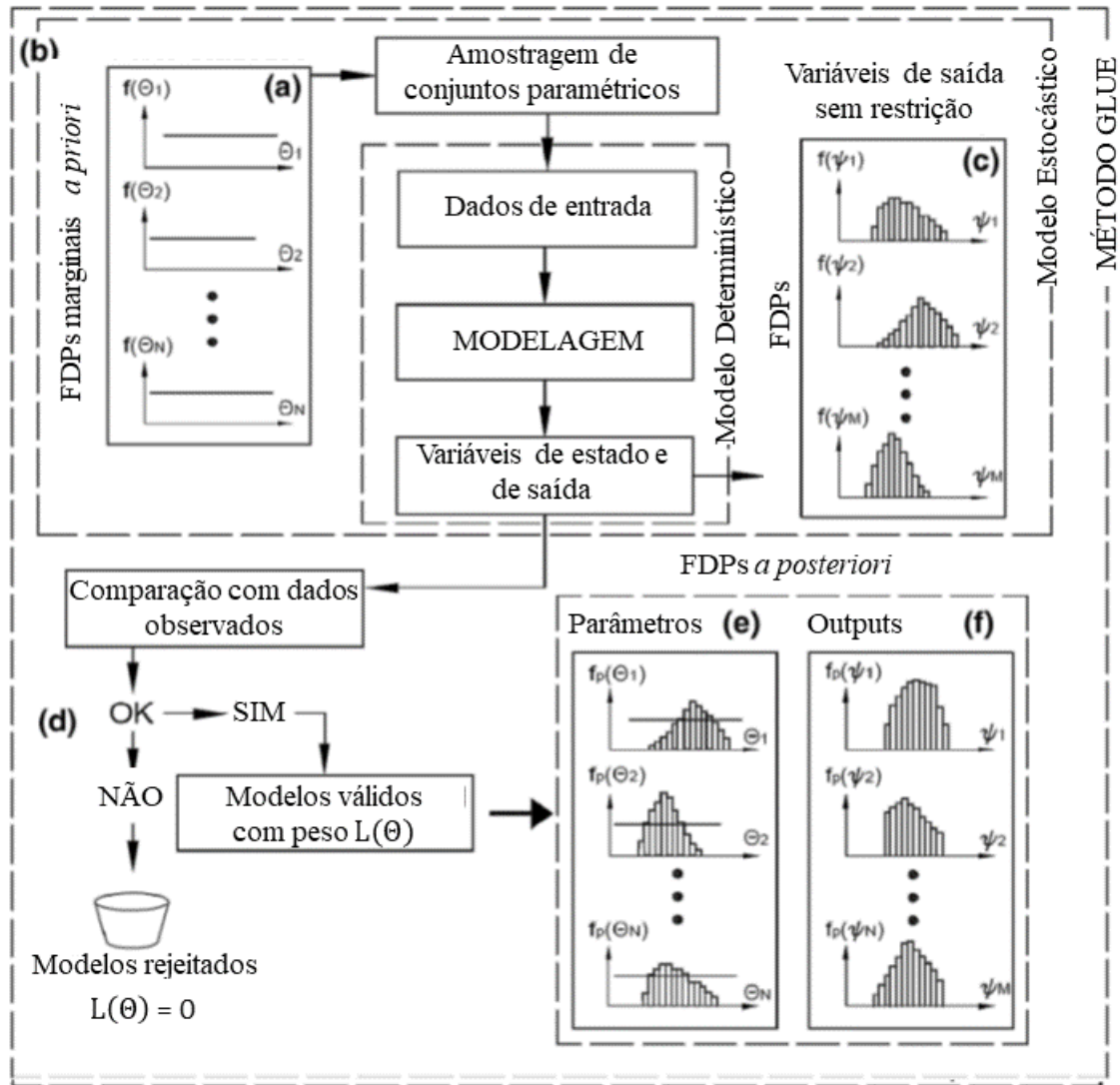


Figura 3.4 – Diagrama esquemático do método GLUE.

Fonte: adaptado de Mirzaei *et al.* (2015).

Stedinger *et al.* (2008), por exemplo, mostraram que o método pode gerar intervalos de credibilidade significativamente diferentes daqueles produzidos por procedimentos estatísticos amplamente aceitos, quando aplicados à calibração de um modelo chuva-vazão linear. Nesse caso, foi possível obter expressões analíticas para quantificar intervalos de confiança para as

predições do modelo sob a ótica da escola frequentista, os quais devem ser semelhantes aos intervalos de credibilidade obtidos por meio da análise Bayesiana sob a hipótese de utilização de funções de verossimilhança formais e uma amostra suficientemente grande de observações $\tilde{Y}$, ainda que as interpretações dadas a estes resultados sejam distintas por ambas as escolas (ZELLNER, 1971 *apud* STEDINGER *et al.*, 2008). Ao se depararem com resultados distintos mesmo para um modelo simples, os autores questionaram a aplicabilidade e confiabilidade do GLUE a problemas não lineares, para os quais é difícil ou até mesmo impossível prescrever expressões exatas para as predições. Em adição, demonstraram matematicamente que algumas das funções de verossimilhança informais mais utilizadas são incoerentes sob o ponto de vista estatístico também por não incorporarem o comprimento n do vetor $\tilde{Y}$ de observações, ao contrário do considerado pelas funções de verossimilhança formais.

Os referidos autores defenderam que um adequado modelo estatístico para descrever as incertezas, expresso por meio da função de verossimilhança, limitaria a subjetividade inerente ao GLUE, sobretudo no que tange aos critérios para seleção das saídas dos modelos consideradas como válidas em relação às não válidas. Em adição, esse cuidado garantiria uma correta distinção entre conjuntos paramétricos que impliquem melhor qualidade de ajuste daqueles que conduzem a ajustes ruins, os quais poderiam ser considerados aceitáveis quando avaliados à luz das funções de verossimilhança informais.

Outro aspecto de atenção é ressaltado por Kuczera e Parent (1998), que afirmaram que o GLUE, tradicionalmente acompanhado por métodos de Monte Carlo para amostragem do espaço multiparamétrico, pode demandar milhões de amostras e, mesmo assim, não explorar adequadamente a região de mais alta densidade de probabilidade. Tal desvantagem, segundo esses autores, estaria condicionada à incorreta prescrição da função de verossimilhança, à seleção das distribuições *a priori*, e à definição dos modelos considerados válidos. No entanto, nota-se que esse aspecto também pode ocorrer quando da utilização de métodos Bayesianos tradicionais, ainda que estejam menos sujeitos a escolhas subjetivas comparativamente ao GLUE, sobretudo caso sejam usados os referidos métodos de amostragem, conhecidamente pouco eficientes para amostragem de espaços com muitas dimensões associados a sistemas não lineares. Blasone *et al.* (2008) mostraram que é possível melhor explorar o espaço paramétrico na estrutura do GLUE utilizando-se métodos de Monte Carlo por Cadeias de Markov.

Ainda que tais aspectos possam conduzir a estimativas incorretas para as incertezas paramétricas e as entradas dos modelos, alguns estudos mostram que o GLUE pode produzir intervalos de credibilidade semelhantes aos fornecidos por métodos Bayesianos formais (VRUGT *et al.*, 2009a; JIN *et al.*, 2010), sendo recomendada a realização de análise de sensibilidade aos aspectos subjetivos supracitados a fim de contornar as limitações por eles impostas. Blasone *et al.* (2008) sugeriram uma calibração do fator de forma K e da fronteira de separação entre os cenários de simulação que conduzem a respostas consideradas como válidas daqueles que levam a respostas não válidas, a partir da maximização da quantidade de dados observados inseridos nos intervalos de credibilidade obtidos para as previsões. Stedinger *et al.* (2008) alertaram, no entanto, que esse procedimento pode penalizar a incerteza paramétrica, inflando-a de maneira irreal (não sustentada pelos dados existentes) em relação às demais fontes de incerteza envolvidas na modelagem

Montanari (2007) sugeriu que o GLUE deva ser compreendido mais como um método para análise ponderada de sensibilidade, por meio do qual seja possível avaliar o peso da incerteza de cada parâmetro sobre as incertezas paramétrica e preditiva, do que como uma técnica para quantificar e identificar adequadamente as incertezas, com o rigor estatístico que essa tarefa demanda.

Por fim, Kavetski *et al.* (2003) advertiram que o método em questão envolve todos os erros presentes em modelagem sob a estrutura implícita da função de verossimilhança informal, sem a necessidade de que sejam feitas e verificadas hipóteses acerca dos resíduos. A consequência é que todas as fontes de incerteza manifestam-se como incerteza paramétrica, superestimando-a, sob a justificativa da equifinalidade.

3.4 Técnicas de amostragem utilizadas em apoio aos métodos de avaliação de incertezas

Os métodos indicados nos itens anteriores demandam técnicas para propagar a incerteza em dados de entrada e parâmetros pelo modelo matemático em estudo. Em algumas raras circunstâncias, a depender da forma das distribuições de probabilidades *a priori*, da quantidade de parâmetros considerados e das funções de verossimilhança, é possível estimar uma expressão analítica para a distribuição conjunta *a posteriori*, em se tratando de métodos de cunho Bayesiano. No entanto, na maioria dos casos, em que geralmente há muitas dimensões envolvidas, é necessário utilizar técnicas numéricas de amostragem a fim de explorar o espaço

paramétrico para construir a distribuição conjunta *a posteriori* e quantificar as incertezas paramétricas e aquelas associadas a dados de entrada e de saída, e, por fim, a incerteza preditiva (HUTTON *et al.*, 2011).

Dentre as técnicas usadas, a mais simples trata-se do método de Monte Carlo (MC) convencional, por meio do qual os parâmetros são sorteados aleatória e independentemente a partir de suas distribuições de probabilidades marginais *a priori*, e assim utilizados para se estimar a distribuição *a posteriori*. Apesar de sua simplicidade, esse método não é eficiente para amostrar todo o espaço paramétrico, um impasse principalmente em problemas de muitas dimensões (i.e., parâmetros). Além disso, a depender do número de sorteios, o processo aleatório e com reposição não garante que o espaço amostral seja adequada e completamente explorado, o que é uma questão especialmente crítica quando se trabalha com distribuições de cauda pesada. Uma maneira de contornar essas desvantagens é aplicar testes para se avaliar a convergência da distribuição *a posteriori*, tais como os de Kolmogorov-Smirnov, Kuiper e Gelman-Rubin (PAPPENBERGER *et al.*, 2005; HUTTON *et al.*, 2011).

Por sua vez, o método *Latin Hypercube* – LH (McKAY *et al.*, 1979) pode ser considerado como mais eficiente em relação ao caso anterior, uma vez que divide os limites previstos para amostragem de cada parâmetro em intervalos disjuntos de igual probabilidade, a partir dos quais os parâmetros são amostrados aleatoriamente e sem reposição. Embora garanta que o espaço paramétrico seja explorado de maneira mais uniforme, essa técnica pode não amostrar adequadamente sua região de mais alta densidade de probabilidade (BLASONE *et al.*, 2008).

Os dois métodos supracitados podem ser utilizados também como procedimentos para quantificação de incertezas sem a estrutura Bayesiana, ou seja, sem a restrição imposta pela função de verossimilhança. Assim, os sorteios são realizados a partir das distribuições marginais estabelecidas para os parâmetros do modelo, de acordo com o conhecimento que se tem acerca de sua variabilidade, sendo os sucessivos conjuntos de parâmetros simulados no modelo em questão de forma determinística, a fim de se obter a resposta **Y** para cada um deles. Finalmente, pode-se avaliar o comportamento estatístico e a forma da distribuição de probabilidades da variável de saída. Ressalva-se, no entanto, que a aplicação dessa metodologia não fornece a incerteza *a posteriori* dos parâmetros, focada somente na propagação de sua variabilidade sobre os *outputs* e em quantificar a incerteza dos mesmos associada exclusivamente aos parâmetros. Eventualmente, dados de entrada podem ter sua incerteza

avaliada, sob o mesmo *status* lógico que os parâmetros, demandando igualmente a prescrição de uma distribuição de probabilidades marginais que modele sua variabilidade.

A maior deficiência da amostragem por simulações de Monte Carlo e pelo método *Latin Hypercube* é que ambas não usam a informação do processo de amostragem a cada iteração para garantir e melhorar a exploração da região de mais alta densidade de probabilidade (HUTTON *et al.*, 2011). Por essa razão, surgiu uma família de procedimentos conhecidos como métodos de Monte Carlo por Cadeias de Markov (MCMC), que consistem na geração de amostras dependentes a partir de cadeias de Markov. Tais cadeias constituem processos estocásticos dotados de memória, nos quais o próximo estado depende somente do estado atual, e que permitem a exploração do espaço paramétrico de maneira gradativa, o que pode contribuir para a redução das taxas de rejeição. Para que seja usada em algoritmos de amostragem, uma cadeia de Markov deve ser (FERNANDES e SILVA, 2017):

- Irredutível: o que significa que a cadeia é capaz de alcançar qualquer outro estado em um número finito de iterações com uma probabilidade maior que zero, independentemente de seu estado inicial;
- Aperiódica, ou seja, a cadeia não oscila entre um conjunto de estados em ciclos regulares; e
- Recorrente: para todo estado j , se o processo se inicia em j , ele retornará ao estado j , com probabilidade 1, em um número finito de iterações.

Uma cadeia de Markov com tais características é dita ergódica. O conceito central de todos os algoritmos de amostragem MCMC é obter uma amostra de $p(\boldsymbol{\theta}|\tilde{\mathbf{Y}})$ construindo-se uma cadeia de Markov ergódica que tenha as seguintes propriedades:

- A cadeia deve ter o mesmo conjunto de estados de $\boldsymbol{\theta}$;
- A cadeia deve ser de fácil simulação; e
- A distribuição de equilíbrio deve ser $p(\boldsymbol{\theta}|\tilde{\mathbf{Y}})$.

O código precursor dos métodos de amostragem do tipo MCMC é o algoritmo do passeio aleatório Metropolis (METROPOLIS *et al.*, 1953), com capacidade para criar cadeias com os supracitados requisitos, e que foi aprimorado em 1970 por Hastings (1970), o que deu origem ao algoritmo Metropolis-Hastings (MH), o primeiro a ser utilizado junto a um processo Bayesiano para inferência de incertezas (KUCZERA e PARENT, 1998).

Em linhas gerais, o código de Metropolis estipula que uma ou mais cadeias evoluam em paralelo, iniciadas em distintos pontos do espaço paramétrico, o qual é, na sequência, amostrado aleatoriamente a partir de uma distribuição de probabilidades multivariada que norteia os “saltos” a serem realizados, chamada de distribuição de referência. A cada novo “salto”, calcula-se a taxa entre a probabilidade do conjunto de parâmetros atual e a do conjunto anterior, a qual é comparada a um valor sorteado aleatoriamente de uma distribuição uniforme $\mathcal{U} \sim (0,1)$. Essa comparação define se o estado atual de cada cadeia deve ser usado para evoluir para o próximo passo, ou se este deverá ser feito a partir do estado anterior, e constitui uma das estratégias para melhor explorar o espaço paramétrico, em relação aos dois métodos anteriormente citados.

Considerando-se que a notação usada nas expressões a seguir sejam adaptadas para representar j estados do vetor paramétrico $\boldsymbol{\theta}$ e que a distribuição-alvo, ou proposta, seja a distribuição conjunta *a posteriori* $p(\boldsymbol{\theta}|\tilde{\mathbf{Y}})$, aproximada pela Equação 3.5 (página 31), então:

- Seja $p(\boldsymbol{\theta}_{t-1}|\tilde{\mathbf{Y}}) \cdot p(\boldsymbol{\theta}_p|\tilde{\mathbf{Y}})$ a probabilidade de o sistema encontrar-se no estado atual, $\boldsymbol{\theta}_{t-1}$;
- Seja $q(\boldsymbol{\theta}_{t-1} \rightarrow \boldsymbol{\theta}_p) \cdot q(\boldsymbol{\theta}_p \rightarrow \boldsymbol{\theta}_{t-1})$ a probabilidade condicional para se fazer um salto de $\boldsymbol{\theta}_{t-1}$ para $\boldsymbol{\theta}_p$,

Então, a probabilidade de aceitação para se fazer um salto do estado atual para o próximo, $p_{ac}(\boldsymbol{\theta}_{t-1} \rightarrow \boldsymbol{\theta}_p)$, é dada por (VRUGT, 2016):

$$p(\boldsymbol{\theta}_{t-1}|\tilde{\mathbf{Y}}) \cdot q(\boldsymbol{\theta}_{t-1} \rightarrow \boldsymbol{\theta}_p) \cdot p_{ac}(\boldsymbol{\theta}_{t-1} \rightarrow \boldsymbol{\theta}_p) = p(\boldsymbol{\theta}_p|\tilde{\mathbf{Y}}) \cdot q(\boldsymbol{\theta}_p \rightarrow \boldsymbol{\theta}_{t-1}) \cdot p_{ac}(\boldsymbol{\theta}_p \rightarrow \boldsymbol{\theta}_{t-1}) \quad (3.37)$$

Caso a distribuição de referência, ou norteadora dos saltos, for simétrica, então $q(\boldsymbol{\theta}_{t-1} \rightarrow \boldsymbol{\theta}_p) = q(\boldsymbol{\theta}_p \rightarrow \boldsymbol{\theta}_{t-1})$. Assim, a expressão 3.37 pode ser rearranjada segundo a forma a seguir:

$$\frac{p_{ac}(\boldsymbol{\theta}_{t-1} \rightarrow \boldsymbol{\theta}_p)}{p_{ac}(\boldsymbol{\theta}_p \rightarrow \boldsymbol{\theta}_{t-1})} = \frac{p(\boldsymbol{\theta}_p|\tilde{\mathbf{Y}})}{p(\boldsymbol{\theta}_{t-1}|\tilde{\mathbf{Y}})} \quad (3.38)$$

A fim de determinar se um salto em direção ao estado $\boldsymbol{\theta}_p$ deve ser aceito ou não, Metropolis *et al.* (1953) propuseram a seguinte relação:

$$p_{ac}(\boldsymbol{\theta}_{t-1} \rightarrow \boldsymbol{\theta}_p) = \min \left[1, \frac{p(\boldsymbol{\theta}_p)}{p(\boldsymbol{\theta}_{t-1})} \right] \quad (3.39)$$

O valor obtido por meio da Equação 3.39 é comparado então com um número aleatório, sorteado de uma distribuição uniforme $\mathcal{U} \sim (0,1)$. Sendo maior, a cadeia evolui para o estado $\boldsymbol{\theta}_p$. Senão, a evolução ocorre a partir do estado $\boldsymbol{\theta}_{t-1}$, e $\boldsymbol{\theta}_p$ é descartado.

Hastings (1970) estendeu o conceito da Equação 3.39 para as situações em que $q(\boldsymbol{\theta}_{t-1} \rightarrow \boldsymbol{\theta}_p) \neq q(\boldsymbol{\theta}_p \rightarrow \boldsymbol{\theta}_{t-1})$, ou seja, quando a distribuição de referência não for simétrica. Assim, o cerne do algoritmo conhecido por Metropolis-Hastings reside na seguinte expressão para obtenção da probabilidade de aceitação de um salto:

$$p_{ac}(\boldsymbol{\theta}_{t-1} \rightarrow \boldsymbol{\theta}_p) = \min \left[1, \frac{p(\boldsymbol{\theta}_p) \cdot q(\boldsymbol{\theta}_p \rightarrow \boldsymbol{\theta}_{t-1})}{p(\boldsymbol{\theta}_{t-1}) \cdot q(\boldsymbol{\theta}_{t-1} \rightarrow \boldsymbol{\theta}_p)} \right] \quad (3.40)$$

Vrugt *et al.* (2003) alertam que uma escolha inadequada da distribuição-alvo (ou seja, do produto das distribuições de probabilidades marginais definidas *a priori* pela função de verossimilhança) e da distribuição norteadora dos saltos, ou distribuição de referência, pode acarretar convergência lenta das cadeias devido à falta de informação *a priori* na região de mais alta densidade de probabilidade no espaço paramétrico e à possível dificuldade de sair de regiões de baixa densidade.

Uma importante característica do algoritmo de Metropolis-Hastings é que as distribuições são avaliadas apenas em termos das chamadas taxas de aceitação, ou seja, do quociente entre os produtos da distribuição-alvo pela distribuição de referência do estado atual e do anterior. Nesse aspecto reside a razão pela qual a constante de normalização da Equação 3.3, e detalhada na Equação 3.4, pode ser dispensada quando da utilização desse conjunto de algoritmos.

Robert e Casella (2004 *apud* FERNANDES e SILVA, 2017) apontam que o referido algoritmo alcança seu equilíbrio, após um grande número de iterações, tendo como distribuição estacionária a distribuição-alvo, $p(\boldsymbol{\theta}|\tilde{\mathbf{Y}})$. Uma vez alcançado o equilíbrio, todas as realizações do algoritmo constituirão amostras da distribuição de probabilidades conjuntas *a posteriori* de $\boldsymbol{\theta}$, permitindo a estimação de descritores estatísticos de cunho Bayesiano, tais como aqueles apresentados no item 3.3.1.3.

A fim de assegurar a convergência das cadeias, e de garantir a exploração do espaço paramétrico em direção à região de mais alta probabilidade da distribuição proposta, sobretudo para tratar problemas multiparamétricos, os métodos MCMC têm-se mostrado como campo profícuo de pesquisas nos últimos 20 anos (HUTTON *et al.*, 2011), evoluindo rapidamente a partir do algoritmo de Metropolis-Hastings. Alguns códigos que surgiram nesse período são o *Adaptive Metropolis* – AM (HAARIO *et al.*, 2001 *apud* VRUGT, 2016), para uma cadeia; o *Shuffled Complex Evolution Metropolis* – SCEM-UA (VRUGT *et al.*, 2003), e o *Differential Evolution Adaptive Metropolis* – DREAM (VRUGT *et al.*, 2009b), para várias cadeias evoluindo em paralelo. Vrugt (2016) faz uma análise retrospectiva e comparativa entre tais métodos, apontando as mudanças em relação aos primeiros algoritmos disponíveis.

Ainda que sejam um avanço em relação aos métodos de amostragem por MC e LH, os algoritmos do tipo MCMC também demandam a aplicação de testes para aferir a convergência e podem apresentar deficiências na exploração da distribuição-alvo em função da insuficiência de dados para atualização da informação *a priori* (HUTTON *et al.*, 2011).

3.5 Aplicação de métodos de avaliação de incertezas na modelagem hidráulica destinada ao estudo de cheias

No que tange à modelagem hidráulica destinada ao estudo de cheias, em regime permanente ou não permanente, prevalecem os estudos de quantificação de incertezas por métodos pseudo-Bayesianos, ou pela propagação da variabilidade dos parâmetros e dados de entrada, prescrita sob a forma de distribuições de probabilidades marginais, via simulações de Monte Carlo, nesse caso sem a atualização da informação *a priori*.

Muitos autores concentraram-se na análise de somente um parâmetro ou variável de cunho hidráulico ou hidrológico, como a vazão de projeto e o coeficiente de rugosidade de Manning (PAPPENBERGER *et al.*, 2005; SMEMOE *et al.*, 2007; ALI *et al.*, 2015), ou de conjuntos de até 3 atributos intervenientes na resposta de manchas de inundação, como topografia, vazão de pico e coeficiente de rugosidade de Manning (JUNG e MERWADE, 2012; 2015).

O modelo HEC-RAS – *River Analysis System*, concebido pelo *Hydrological Engineering Center* do Corpo de Engenheiros do Exército americano (BRUNNER, 2016a; 2016b), é utilizado à exaustão em diversos estudos de análise de sensibilidade e dos efeitos das incertezas inerentes aos dados (sobretudo topográficos) e parâmetros (principalmente o coeficiente de

rugosidade de Manning) comumente adotados em sua estrutura, a exemplo de Pappenberger *et al.* (2005; 2006a), Cook e Merwade (2009), Jung e Merwade (2012; 2015), e Ali *et al.* (2015), a maioria dos quais em ambiente unidimensional.

Jung e Merwade (2012) quantificaram as incertezas advindas de dados topobatimétricos, coeficiente de rugosidade de Manning e uma vazão de cheia notável sobre a incerteza preditiva envolvida no mapeamento de inundações e na determinação de perfis de escoamento, em conjunto e separadamente, por meio da realização de milhares de simulações de Monte Carlo acopladas ao método GLUE (BEVEN e BINLEY, 1992). Cada simulação de Monte Carlo representou uma combinação de parâmetros e variáveis, estimados aleatoriamente a partir de distribuições de probabilidades marginais *a priori* limitadas superior e inferiormente a faixas de valores plausíveis. Os resultados, expressos em termos dos intervalos de credibilidade a um nível de 95% para a área inundada, e elevação da linha de água e respectivas profundidades nas seções transversais, foram ponderados por meio de diversas medidas de verossimilhança informais, descritivas da qualidade do ajuste de cada saída aos valores paradigmáticos, representados seja por dados observados, seja por uma situação-modelo esperada, dependendo do trecho fluvial estudado e de sua disponibilidade de dados. As Equações 3.34 a 3.36, enunciadas às páginas 51 e 52, subitem 3.3.3, representam algumas das funções usadas.

Os autores avaliaram o efeito da função de verossimilhança sobre a função acumulada de probabilidades (FAP) preditiva obtida *a posteriori* e concluíram que, para os trechos fluviais selecionados, essa decisão interferia sobre a incerteza preditiva em distintos graus, a depender da variável considerada ou da combinação de todas as variáveis intervenientes. Nesse trabalho, foram considerados como válidas todas as simulações realizadas, sem que fosse utilizado um critério de corte, como sugere o GLUE. Algumas premissas e fontes de informações desse estudo são resumidas a seguir devido à sua relevância para a presente pesquisa:

- Utilização do modelo HEC-RAS em ambiente unidimensional, alimentado por seções transversais extraídas de um MDT com resolução horizontal de 1,5 m, em um caso, ou criadas a partir de levantamentos de campo associados a elevações retiradas de um MDT, para o outro trecho fluvial estudado. Ambos os estirões avaliados situam-se nos Estados Unidos;
- Realização de simulações sob a hipótese de escoamento permanente, utilizando: (i) a vazão de pico de uma cheia excepcional, estimada a partir da relação cota-descarga estabelecida

pelo USGS – U.S. *Geological Survey*, para um dos trechos, e (ii) a vazão da cheia com TR de 100 anos, calculada por meio de uma equação de regressão regional que a relaciona à área de drenagem, elaborada pelo programa computacional USGS *National Flood Frequency* – NFF (RIES e CROUSE, 2002 *apud* JUNG e MERWADE, 2012); e

- Uso das seguintes distribuições de probabilidades *a priori* para as variáveis, dados e parâmetros selecionados: (i) uniforme para o coeficiente de rugosidade de Manning, cujos limites inferior e superior foram definidos a partir de tabelas disponíveis na literatura para distintas tipologias de uso do solo; (ii) *t* de Student para as vazões de referência, com limites inferior e superior relativos aos quantis de 2,5% e 97,5%; e (iii) uniforme para a topografia, cujos valores limitantes basearam-se nas precisões verticais das respectivas bases topobatimétricas. O sorteio realizado sobre tais distribuições referia-se a uma porcentagem de variação a ser aplicada em relação ao valor central, e não a valores absolutos. No caso dos coeficientes de rugosidade, somente uma porcentagem era extraída de cada sorteio, posteriormente aplicada sobre as duas ou três tipologias de cada seção transversal, para todo o trecho fluvial avaliado.

A partir das análises realizadas, e considerando-se todas as funções de verossimilhança utilizadas, foram elaborados diversos produtos probabilísticos, como perfis de escoamento e manchas de inundação. Os autores concluíram que, em se tratando dos dois trechos fluviais estudados, as distintas funções de verossimilhança não acarretaram diferenças importantes (i.e., menores do que 5%) nos correspondentes intervalos de credibilidade associados à área inundada simulada pelo modelo HEC-RAS. No entanto, ressaltam que a generalização dessa conclusão demandaria a avaliação de diversos trechos fluviais, com distintas morfologias e diferentes fatores intervenientes na gênese de cheias.

O estudo conduzido por Jung e Merwade (2012) trata-se de um dos mais completos envolvendo o método GLUE para a modelagem hidráulica de escoamento permanente associado a cheias excepcionais, em que os autores propuseram-se compreender a parcela das variáveis consideradas como mais relevantes no modelo utilizado na composição da incerteza preditiva. No entanto, não foi feita uma análise do comportamento estatístico dos resíduos da modelagem, a fim de que fossem confirmadas as premissas que geralmente embasam a prescrição de funções de verossimilhança informais: a homoscedasticidade e a normalidade, bem como a ausência de correlação temporal (ou espacial, alternativamente). Além disso, a desconsideração de um modelo estatístico para os resíduos dificulta a obtenção da incerteza preditiva total, ou seja,

associada a todas as fontes de incerteza do modelo, ainda que de forma implícita. Assim, nesse caso, a incerteza preditiva ficou definida pelas incertezas associadas à topografia, à vazão de pico afluente e ao coeficiente de rugosidade, sem que houvesse uma verificação de outras possíveis fontes de incertezas por meio de uma função de verossimilhança formal construída a partir dos resíduos $\mathbf{E}(\boldsymbol{\theta}) = \tilde{\mathbf{Y}} - \mathbf{Y}(\boldsymbol{\theta}) = \{\varepsilon_1(\boldsymbol{\theta}), \dots, \varepsilon_n(\boldsymbol{\theta})\}$.

A fim de complementar o trabalho descrito anteriormente, Jung e Merwade (2015) realizaram uma análise de sensibilidade a partir do método FOA – *First Order Approximation* (BENJAMIN e CORNELL, 1970 *apud* JUNG e MERWADE, 2015) para avaliar o efeito da variância de cada variável, parâmetro ou dado utilizado nos modelos hidráulicos dos mesmos trechos fluviais estudados por esses autores em 2012 sobre a variância na saída do modelo, expressa nessa situação em termos de área inundada. Esse método baseia-se na expansão por séries de Taylor com truncamento após o termo de primeira ordem. Como as derivadas da área inundada em relação às variáveis supracitadas não têm solução analítica, foi necessário realizar milhares de simulações a fim de se obter uma curva empírica de relação entre a variável independente sob consideração e a variável dependente (i.e., área inundada).

Em paralelo, esses autores aplicaram o método HSY – *Hornberger-Spear-Young* (HORNBERGER e SPEAR, 1981 *apud* JUNG e MERWADE, 2015) para estimar a sensibilidade relativa entre as variáveis do modelo sobre o resultado final. Esse procedimento, considerado como um método do tipo GSA – *Global Sensitivity Analysis*, assemelha-se ao GLUE, com o acréscimo da aplicação do teste de Kolmogorov-Smirnov sobre as funções acumuladas de probabilidades elaboradas sobre as respostas classificadas como válidas e as não válidas a fim de se avaliar a distância máxima entre ambas e assim acessar quantitativamente a sensibilidade das saídas do modelo a um determinado parâmetro ou variável, relativamente às demais variáveis consideradas. Essa técnica foi aplicada para cada uma das três variáveis estudadas, fixando-se as outras duas em seus valores máximos e/ou mínimos.

Outros autores aplicaram o método GLUE para avaliar ao menos um fator agregador de incertezas à modelagem hidráulica (PAPPENBERGER *et al.*, 2005; ALI *et al.*, 2015).

Pappenberger *et al.* (2005) conduziram um amplo estudo de avaliação das incertezas relativas ao coeficiente de rugosidade de Manning na modelagem hidrodinâmica unidimensional, usando o aplicativo HEC-RAS. Foram estudados 2 trechos fluviais europeus, com distintas

disponibilidades de dados topobatimétricos e fluviométricos para construção dos modelos e atualização da informação *a priori* por meio do método GLUE. Foram avaliados diversos critérios para separação das respostas válidas em relação às não válidas, bem como medidas de desempenho e funções de verossimilhança que se adequassem à realidade de informações extraídas dos postos fluviométricos de jusante e/ou das manchas de inundação observadas. Para seleção das simulações hidrodinâmicas válidas, os autores propuseram uma expressão derivada a partir da medida de desempenho mostrada a seguir, baseada na comparação entre a área inundada prevista em cada simulação e a área inundada observada (HORRIT e BATES, 2001):

$$F_1 = \frac{S_{obs,sim} - S_{\tilde{obs}}}{\tilde{S}} \quad (3.41)$$

em que:

$S_{obs,sim}$ é a interseção entre a área inundada prevista em cada simulação, correspondente à vazão de pico da cheia, e a área inundada observada. Assim, essa variável representa a área inundada corretamente prevista pelo modelo;

$S_{\tilde{obs}}$ é a área inundada prevista pelo modelo que não foi observada (superestimação); e

$\tilde{S}$ é a área inundada observada, extraída no caso da mancha de inundação disponível.

A inovação de Pappenberger *et al.* (2005) em relação à Equação 3.41 residiu na multiplicação das áreas utilizadas nessa expressão por um fator ponderador atribuído a cada célula, ou *pixel* da mancha de inundação, tanto observada quanto oriunda das simulações. Tal fator varia entre 0 e 1, adquirindo o valor da unidade a partir de determinada profundidade de escoamento no pixel, definida em função da precisão do MDT (vide a Figura 3.5). Segundo os autores, essa profundidade limite refletiria a incerteza no MDT usado como base para quantificação das profundidades em toda a mancha de inundação.

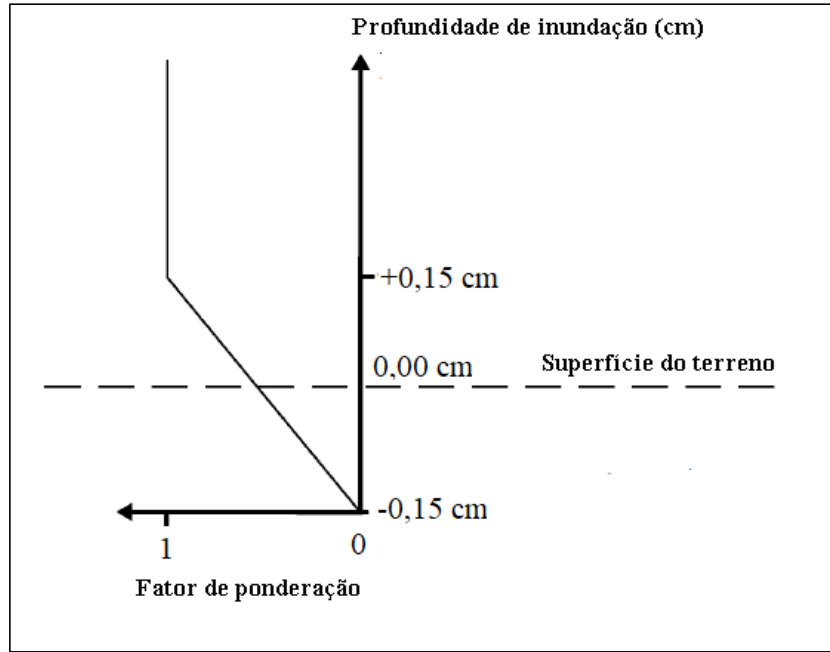


Figura 3.5 – Função adotada para estimativa do fator de ponderação a ser adotado para cada célula do MDT que tiver sido inundada, seja nas simulações, seja na mancha observada, conforme proposição de Pappenberger *et al.* (2005). Nesse caso, a precisão do MDT era de 0,15 cm.

Fonte: adaptado de Pappenberger *et al.* (2005).

Outro aspecto de destaque relativamente ao GLUE proposto por esses autores trata-se da função de verossimilhança informal finalmente utilizada, sob a premissa de se representar não só a extensão espacial da mancha de inundação, como também de se aferir a proximidade entre as vazões simulada e observada na seção de jusante, conforme sintetizado pela Equação 3.42:

$$L_{\text{pond}} = R \cdot \sum_{r=1}^R |E_r - \tilde{E}_r| + \left[1 - \frac{\sum_{t=1}^n (y_t - \tilde{y}_t)^2}{\sum_{t=1}^n (\tilde{y}_t - \bar{\tilde{y}})^2} \right] \quad (3.42)$$

na qual, coerentemente com a notação utilizada nas Equações 3.32 e 3.35, têm-se:

E_r e $\tilde{E}_r$ representam, para cada seção r , as elevações da linha de água simulada e extraída a partir do cruzamento da mancha de inundação com o MDT que serve de base para a definição das áreas inundadas;

R é o número total de seções transversais utilizado no modelo hidráulico;

y_t e $\tilde{y}_t$ são, no presente caso, as vazões do hidrograma de jusante simulado e observado, respectivamente, avaliadas em cada intervalo de tempo t ;

$\bar{\tilde{y}}$ é a vazão média do hidrograma observado; e

L_{pond} é o valor da função de verossimilhança informal avaliadas por meio das elevações e vazões.

Diferentemente de outros estudos que avaliaram o coeficiente de rugosidade de Manning com o GLUE, nessa situação foram adotados limites bem amplos para as distribuições *a priori* desse parâmetro, representadas pelo modelo Uniforme: 0,001 e 0,900, tanto para a calha, quanto para

as planícies de inundação, valores os quais, segundo os autores, comporiam a faixa de coeficientes de rugosidade fisicamente plausíveis indicados pelo módulo hidrodinâmico do HEC-RAS. No entanto, raramente são encontrados valores na literatura maiores do que 0,200 para esse parâmetro a fim de retratar a rugosidade de cursos de água naturais, sendo que os valores mais elevados caracterizam geralmente as planícies de inundação, ou estão associados à ocorrência de profundidades reduzidas na calha principal.

Um dos diversos testes relatados nesse artigo referiu-se à avaliação dos resultados ao se sortear cada uma das três rugosidades de cada seção transversal (i.e., um valor para a calha e um valor para a planície de inundação em cada margem) independentemente uma da outra em relação à opção de se sortear somente um único valor para a calha e outro para as planícies de inundação em ambas as margens, válidos para toda a extensão do trecho fluvial. Constatou-se que a segunda alternativa forneceu perfis de escoamento mais fisicamente plausíveis e garantiu maior estabilidade numérica às simulações hidrodinâmicas. Além disso, o sorteio independente dos coeficientes de rugosidade gerou uma superparametrização do problema: ao invés de 2 parâmetros sujeitos à inferência pseudo-Bayesiana, ter-se-iam $R \times 3$ parâmetros, sendo R o número de seções transversais usadas no modelo. No referido estudo, o teste realizado sob a hipótese de independência entre as rugosidades em cada seção e entre todas as seções não proporcionou nenhuma melhoria na modelagem trazida pela análise do GLUE e pela incorporação da informação $\tilde{Y}$ às incertezas. Nesse sentido, todas as simulações apresentaram desempenho e valores da função de verossimilhança semelhantes, avaliados por meio das Equações 3.41 e 3.42, respectivamente. Finalmente, parece mais provável que determinado nível de incerteza afete a seleção de todas as rugosidades ao longo de um trecho fluvial, ainda que diferenciadas entre algumas tipologias de solo e/ou de uso e ocupação, do que sejam atribuídas distintas incertezas a cada uma das rugosidades por seção transversal e entre todas as seções transversais, a menos que haja evidências para o contrário.

Apesar das tentativas de se condicionar o conhecimento *a priori* às informações disponíveis por meio de uma função de verossimilhança mais específica à modelagem hidrodinâmica voltada ao mapeamento de áreas inundáveis, os resultados apontaram a existência de centenas de combinações entre as rugosidades da calha e das planícies que apresentaram qualidade elevada à luz dos critérios adotados. Adicionalmente, verificou-se que em um dos estirões fluviais a função de verossimilhança (Equação 3.42) teve seus máximos em duas regiões bem distintas em termos de rugosidade da calha principal, o que foi creditado a uma mudança do regime de

escoamento. Finalmente, a incerteza *a posteriori* relativa ao coeficiente de rugosidade nas planícies manteve-se semelhante à sua incerteza prescrita *a priori*, em um dos trecho fluviais, sendo que, no outro trecho, o mesmo ocorreu também com a rugosidade na calha principal. Isso pode dever-se à magnitude das cheias utilizadas como paradigma de calibração das incertezas pelo método GLUE, que possivelmente não atingiram de maneira importante as planícies. Alternativamente, cogita-se a possibilidade de que o uso de funções de verossimilhança informais conduzam a uma identificação inadequada das incertezas paramétricas *a posteriori*.

Em outros casos, como no estudo de Smemoe *et al.* (2007), o efeito das incertezas foi avaliado exclusivamente sob a forma de mapas probabilísticos, variando-se somente uma variável de entrada ao modelo hidráulico, e realizando-se simulações de Monte Carlo com amostragem pelo método *Latin Hypercube* (McKAY *et al.*, 1979), sem que houvesse o condicionamento das saídas em relação a dados observados. Nesse trabalho, a variável avaliada como agregadora de incertezas foi o parâmetro CN – *Curve Number*, utilizado no método chuva-vazão do SCS – *Soil Conservation Service* (MISHRA e SINGH, 2003) e seu impacto sobre o mapeamento de áreas inundadas em ambiente 1D para a cheia de referência, com TR de 100 anos. Para tanto, foi adotada a hipótese de que esse parâmetro seguiria uma distribuição Normal ao redor do seu valor médio, estimado a partir da literatura, levando-se em conta o uso e a ocupação do solo na bacia estudada. O desvio-padrão para essa distribuição foi obtido a partir dos valores mínimo e máximo plausíveis para o CN, atribuídos pelos autores às condições I e III de umidade antecedente do solo avaliadas para o valor médio supracitado, considerando-se que este estaria associado à condição II. Nesse caso, o critério considerado na atribuição de probabilidades às área inundadas foi o seguinte:

$$p_i = \frac{g}{G} \times 100 \quad (3.43)$$

na qual:

p_i é a probabilidade de um ponto no modelo digital de elevação do tipo TIN – *Triangulated Irregular Network* – ser inundado;

g é o número total de simulações nas quais o ponto foi inundado; e

G é o número total de simulações.

A partir dos valores de p_i estimados para cada um dos pontos da malha TIN do terreno, foram traçados os contornos probabilísticos associados à variabilidade do CN. Na sequência, aferiu-se a sensibilidade de alguns descritores estatísticos da área inundada total em relação ao número total de simulações, que variou entre 100 e 1000. O mapa probabilístico elaborado variando-se

o parâmetro CN sob os critérios mencionados, e utilizando-se a Equação 3.39 tem o aspecto mostrado na Figura 3.6, para um total de 1000 simulações realizadas.

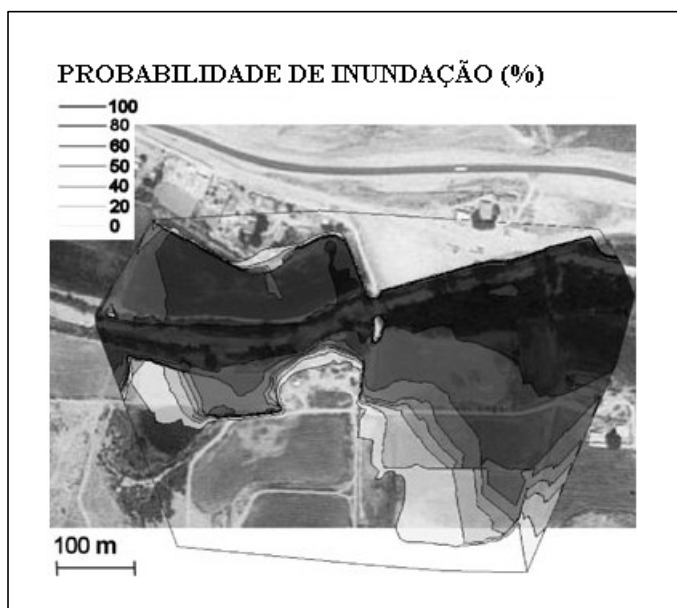


Figura 3.6 – Mapa probabilístico da cheia de 100 anos associado à incerteza do parâmetro CN, para trecho do rio Virgin, no Utah, Estados Unidos, elaborado no estudo de incertezas de Smemoe *et al.* (2007).

Fonte: adaptado de Smemoe *et al.* (2007).

Um exemplo muito interessante trata-se do estudo realizado por Di Baldassarre *et al.* (2010), no qual foram comparadas as abordagens determinística e probabilística para a elaboração do mapa de inundação associada à cheia com TR de 100 anos no rio Dee, na Escócia. Para o primeiro caso, foi utilizado um modelo 2D completo, fisicamente baseado, o TELEMAC 2D (GALLAND *et al.*, 1991 *apud* DI BALDASSARRE *et al.*, 2010), que, de acordo com os autores, seria adequado para representar a dinâmica de propagação e amortecimento de uma cheia importante sob a ótica determinística, mediante calibração do coeficiente de rugosidade de Manning na calha principal a partir de um evento de magnitude significativamente menor do que a do evento notável mencionado. O coeficiente de rugosidade nas planícies, por sua vez, foi fixado em 0,100, por ser considerado representativo de suas condições no trecho fluvial estudado. Para elaboração do modelo hidráulico 2D foi utilizado um MDT com resolução de 2 m e precisão vertical de 0,10 m.

Para seleção do valor ótimo da rugosidade de Manning na calha foi usada a seguinte medida de desempenho, comparando-se as sucessivas simulações em regime não permanente à mancha de

inundação estimada a partir de uma imagem de satélite disponível para o evento de calibração, associado a uma cheia com 2 anos de período de retorno:

$$F_2 = \frac{S_{obs,sim} - S_{\#obs}}{S_{obs,sim} + S_{\#obs} + S_{\#sim}} \quad (3.44)$$

em que:

$S_{obs,sim}$ é a extensão da área inundada corretamente simulada pelo modelo;

$S_{\#obs}$ é a área inundada prevista pelo modelo que não foi inundada segundo a observação disponível (superestimação); e

$S_{\#sim}$ é a área inundada não prevista pela simulação (subestimação).

Os valores de F_2 variam entre -1 e 1, sendo que a unidade indica um ajuste perfeito à área observada. Essa medida de desempenho garantiu ajustes adequados em estudos de modelagem para previsão de áreas inundadas e permite a comparação da resposta obtida em distintos trechos fluviais quanto a diversos eventos de cheia (HORRITT *et al.*, 2007). A calibração realizada sob a hipótese de aplicação da Equação 3.44 forneceu um valor de 0,04 para o coeficiente de rugosidade de Manning na calha principal, ressaltando-se que este provavelmente seria menor caso tivesse sido obtido a partir de uma cheia com magnitude mais próxima àquela de referência para o estudo. Em seguida, foi realizada a simulação em ambiente 2D, em regime permanente, para a vazão de pico da cheia de 100 anos, considerando-se a rugosidade recém-calibrada.

Já para a abordagem probabilística, por demandar a realização de milhares de simulações, foi reservado o uso do modelo LISFLOOD-LP (BATES *et al.*, 2013), caracterizado por ser 2D nas planícies e 1D na calha principal e por um esquema de resolução numérica mais simples do que o utilizado no modelo TELEMAC 2D. Para sua aplicação, inicialmente foi realizada uma análise de sensibilidade, variando-se os coeficientes de Manning na calha na faixa entre 0,020 e 0,100 e aferindo-se a correspondente medida F_2 (Equação 3.40), relativamente à cheia de TR 2 anos supracitada, sob a hipótese de escoamento transiente. Na planície, a rugosidade manteve-se fixa em 0,100, assim como feito na abordagem determinística. Os valores obtidos para o índice F_2 foram comparados àqueles calculados sob a hipótese de um perfil de escoamento linear entre os postos fluviométricos de controle a montante e a jusante, e todas as simulações que apresentaram maiores valores para F_2 foram consideradas como válidas para a estrutura do método GLUE. Considerou-se, assim, somente a incerteza paramétrica como fonte de erros associados à propagação da cheia de interesse. Da equação da medida F_2 foi derivada a função de verossimilhança informal utilizada na estrutura do GLUE, qual seja:

$$L_{F_2} = \frac{F_2 - \min(F_2)}{\max(F_2) - \min(F_2)} \quad (3.45)$$

Na equação anterior, $\max(F_2)$ e $\min(F_2)$ são, respectivamente, os valores máximo e mínimo da medida de desempenho F_2 (Equação 3.40) obtidos para o conjunto de simulações consideradas como válidas.

Finalmente, o modelo LISFLOOD-LP foi utilizado para realização de milhares de simulações variando-se o coeficiente de rugosidade de Manning na calha principal dentro da faixa obtida *a posteriori* por meio da Equação 3.45 e obtendo-se os correspondentes perfis de escoamento e áreas inundadas para a vazão de pico da cheia com TR de 100 anos. Para atribuição da probabilidade de inundação associada a essa cheia sobre cada célula do *grid* do MDT utilizado, foi adotada a seguinte equação:

$$p_j = \frac{\sum_i L_{F_2,i} \cdot w_{i,j}}{\sum_i L_{F_2,i}} \quad (3.46)$$

Em que $w_{i,j}$ vale a unidade se a célula tiver sido inundada na i -ésima simulação, e zero caso não tenha sido. Assim, cada célula do MDT recebeu um valor de probabilidade de inundação pela cheia de 100 anos, a qual se refere à incerteza preditiva associada exclusivamente à variabilidade da rugosidade na calha principal. O mapa probabilístico elaborado à luz de tais critérios é mostrado na Figura 3.7, ao lado do mapa determinístico correspondente.

Adicionalmente, foi elaborado um segundo mapa probabilístico do trecho considerado, desta vez agregando às incertezas *a posteriori* calibradas para a rugosidade na calha aquelas referentes à descarga de pico da cheia de 100 anos de período de retorno. Para tanto, foram usados diversos registros de vazões máximas anuais levantados em bacias hidrográficas do Reino Unido, extraídas da referência *Flood Estimation Handbook* (INSTITUTE OF HYDROLOGY, 1999 *apud* DI BALDASSARRE *et al.*, 2010), e reportados em Di Baldassarre *et al.* (2009), a partir dos quais se constatou que uma variação de $\pm 15\%$ seria adequada para explicar a variabilidade sobre o valor estimado para a vazão de pico dessa cheia notável, sem contabilizar as incertezas de medição. Assim, foram sorteados pares de vazão de pico e de rugosidade na calha para realização de novas simulações a fim de se avaliar o efeito combinado de suas incertezas sobre a predição da área atingida.

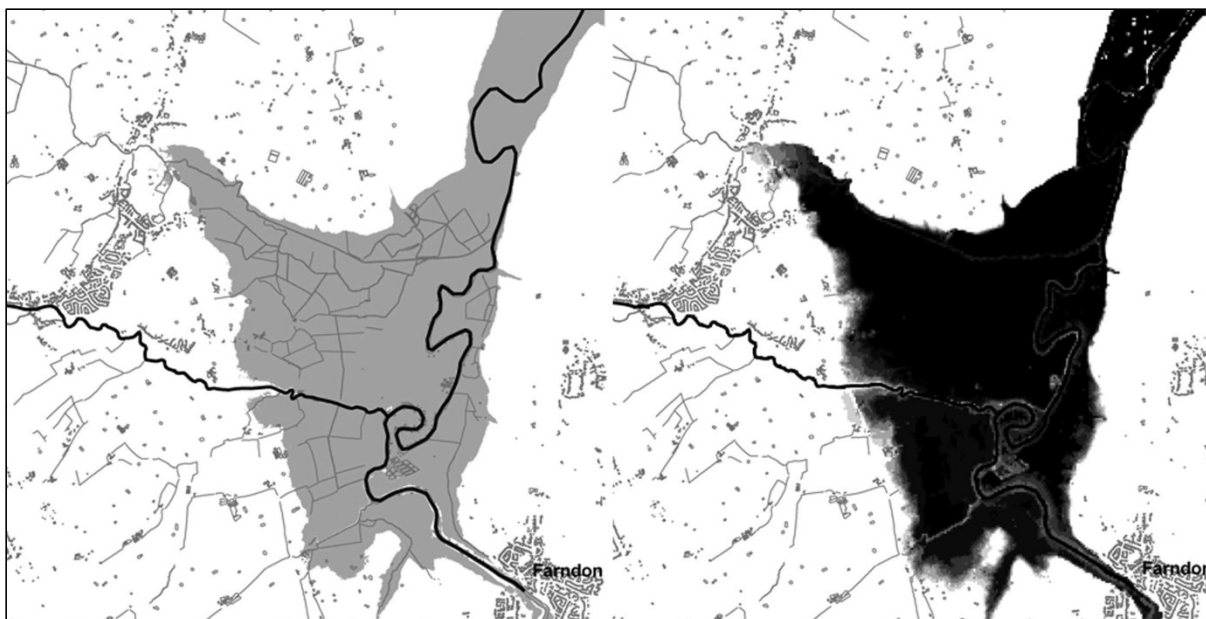


Figura 3.7 – Mapas determinístico (esq.) e probabilístico (dir.) modelados para um trecho do rio Dee, na Escócia. No segundo mapa, à direita, foram consideradas as incertezas associadas ao coeficiente de rugosidade de Manning na calha principal.

Fonte: adaptado de Di Baldassarre *et al.* (2010).

Por sua vez, Ali *et al.* (2015) realizaram uma análise de sensibilidade da incerteza paramétrica em uma modelagem hidrodinâmica unidimensional em face da variabilidade de precisão vertical e da resolução horizontal de distintas bases topográficas e os correspondentes modelos digitais de terreno, para um estirão do rio Johor, na Malásia. Primeiramente, calibraram e validaram o coeficiente de rugosidade de Manning na calha e nas planícies usando o modelo LISFLOOD-LP, que alia as modelagens 1D e 2D, a partir de hidrogramas e cotogramas observados no posto fluviométrico de montante no trecho fluvial de interesse para 2 cheias importantes, e considerando-se escoamento uniforme como condição de contorno de jusante. A mancha de inundação obtida para o primeiro evento foi considerada como paradigma para a estimação das incertezas, tendo sido elaborada sobre o MDT mais preciso disponível.

Na sequência, o evento de calibração foi simulado no HEC-RAS para 6 cenários distintos, construídos a partir de diferentes MDTs obtidos a partir das bases cartográficas existentes, com resoluções horizontais variando entre 1 m x 1 m e 90 m x 90 m. Em todos eles, os coeficientes de Manning na calha e na planície foram amostrados a partir de uma distribuição uniforme, com limites inferior e superior iguais a 0,020 e 0,080, e 0,030 e 0,100, respectivamente. O desempenho das simulações e a separação entre aquelas tidas por válidas em relação às não válidas foram aferidos a partir da soma das diferenças absolutas entre os níveis de água

máximos simulado e de referência (i.e., do modelo calibrado no LISFLOOD) em todas as seções transversais. A função de verossimilhança informal utilizada foi semelhante àquela correspondente à Equação 3.45. Por fim, foram elaborados os perfis de escoamento probabilísticos correspondentes a um nível de 95% de credibilidade, alcançados pelas vazões máximas ao longo do trecho fluvial, bem como gráficos relacionando as rugosidades das simulações válidas e os correspondentes valores do indicador de desempenho supracitado. Os autores constataram que todos os modelos, seja qual for o MDT de base, apresentaram maior sensibilidade à rugosidade da calha do que àquela da planície. Em adição, observaram que houve diferenças pouco significativas entre as rugosidades correspondentes à melhor performance em se tratando dos modelos elaborados a partir do MDT de melhor resolução, mesmo quando este foi reescalado para resoluções menores (i.e., escalas espaciais menores). No que tange aos perfis de escoamento de cunho probabilístico, houve diferença importante daqueles baseados nos MDTs de menor resolução horizontal em relação àqueles elaborados sobre o modelo hidrodinâmico de referência.

Embora os estudos até aqui relatados apresentem extensos resultados mostrando a sensibilidade de variáveis de saída, como profundidade de escoamento e área total inundada, principalmente à incerteza associada à rugosidade, não há uma análise dos resíduos da modelagem hidráulica ou hidrodinâmica que pudesse subsidiar e confirmar as hipóteses subjacentes às funções de verossimilhança utilizadas. Em adição, pouco se comenta sobre a incerteza paramétrica obtida *a posteriori*.

Um dos poucos exemplos encontrados até o momento retratando a aplicação de métodos Bayesianos formais à modelagem hidráulica e hidrodinâmica trata-se do estudo conduzido por Camacho *et al.* (2015). Em seu trabalho, foram avaliados três abordagens de cunho Bayesiano para avaliação de incertezas aplicadas a dois estudos de caso, sendo o primeiro um exemplo sintético unidimensional de propagação de cheia pelo método da onda difusiva, e o segundo uma modelagem hidrodinâmica 3D que serve de base para avaliação de processos estuarinos na baía de Weeks, no estado de Alabama, nos Estados Unidos.

Para ambas as situações, os três procedimentos adotados e comparados foram: método Bayesiano formal, com amostragem por simulações de (i) Monte Carlo, e de (ii) Monte Carlo por Cadeias de Markov; e o (iii) método GLUE, com utilização de uma função de verossimilhança informal.

No primeiro caso, os resíduos foram avaliados em termos das vazões defluentes no extremo de jusante, sendo que o hidrograma sintético “real” nesse local foi estimado a partir de um conjunto ideal dos parâmetros coeficiente de rugosidade de Manning, largura da seção transversal, sendo esta considerada prismática, e declividade do álveo, adicionando-se a cada simulação um erro aditivo Gaussiano com média nula e desvio-padrão igual a 10% da vazão defluente resultante. Os parâmetros supracitados constituíram objeto de inferência de incertezas, e receberam distribuições de probabilidades *a priori* do tipo uniforme. Para aplicação dos métodos (i) e (ii) apresentados no parágrafo anterior, foi utilizada a função de verossimilhança expressa na Equação 3.7, à página 35, considerando-se a variância dos resíduos constante, e sujeita também à inferência Bayesiana. Para essa variável latente, a distribuição *a priori* de Jeffreys (1961 *apud* FERNANDES e SILVA, 2017) foi utilizada, em que $p(\sigma_\varepsilon) \propto \sigma_\varepsilon^{-1}$.

Na aplicação do método pseudo-Bayesiano GLUE, foi usado o critério de Nash-Sutcliffe (NASH e SUTCLIFFE, 1970 *apud* CAMACHO *et al.*, 2015), cuja expressão é dada a seguir:

$$L(\boldsymbol{\theta}|\tilde{\mathbf{Y}}) = 1 - \frac{\sum_{t=1}^n (\tilde{y}_t - y_t)^2}{\sum_{t=1}^n (\tilde{y}_t - \bar{\tilde{y}})^2} \quad (3.47)$$

na qual $\bar{\tilde{y}}$ é a média amostral do vetor de observações $\tilde{\mathbf{Y}}$, no caso, composto por vazões defluentes no extremo de jusante do estirão sintético estudado.

Para o caso tridimensional, os resíduos foram avaliados de forma multivariada, em termos de níveis de água e de velocidades, considerando-se independência entre essas duas variáveis. Os primeiros foram levantados a partir de dados de um posto de monitoramento a montante da baía de Weeks, posicionado em um de seus principais cursos de água afluentes, enquanto que as segundas foram extraídas de registros sistemáticos de um posto situado na porção de jusante do estuário formado pelos rios que deságuam na baía. O parâmetro que constituiu objeto de quantificação de incertezas nessa condição foi exclusivamente a altura de rugosidade, em todos os métodos, considerada como um único valor em toda a extensão da malha representativa do sistema em estudo. Sua distribuição *a priori* foi Uniforme, com limites inferior e superior iguais a 0,001 e 0,150 m, respectivamente.

Para as simulações pelos métodos Bayesianos formais, a função de verossimilhança adotada foi a seguinte:

$$L(\boldsymbol{\theta}|\tilde{\mathbf{Y}}_1, \tilde{\mathbf{Y}}_2) = \left[\frac{1}{(2\pi\sigma_{\varepsilon_1}^2)^{n_1/2}} \exp\left(-\frac{1}{2\sigma_{\varepsilon_1}^2} \sum_{t=1}^{n_1} \varepsilon_{1,t}^2\right) \right] * \left[\frac{1}{(2\pi\sigma_{\varepsilon_2}^2)^{n_2/2}} \exp\left(-\frac{1}{2\sigma_{\varepsilon_2}^2} \sum_{t=1}^{n_2} \varepsilon_{2,t}^2\right) \right] \quad (3.48)$$

na qual:

$\tilde{\mathbf{Y}}_1$ e $\tilde{\mathbf{Y}}_2$ são os dados observados de nível de água (m) e de velocidade (m/s), respectivamente, os quais foram confrontados com os correspondentes valores simulados para obtenção dos resíduos ε_1 e ε_2 . Estes, por sua vez, compõem os vetores de comprimentos n_1 e n_2 . As variâncias dos resíduos, tomadas por constantes e denotadas por $\sigma_{\varepsilon_1}^2$ e $\sigma_{\varepsilon_2}^2$, foram inferidas a partir de seus máximos de verossimilhança após a aplicação de ambos os métodos anteriormente citados (a exemplo do mencionado junto à Equação 3.8, situada à página 35).

Já na situação de uso do GLUE, o critério de Nash-Sutcliffe foi adaptado para sua forma multivariada, tal como mostrado a seguir:

$$L(\boldsymbol{\theta}|\tilde{\mathbf{Y}}_1, \tilde{\mathbf{Y}}_2) = \left[1 - \frac{\sum_{t=1}^{n_1} (\tilde{y}_{1,t} - y_{1,t})^2}{\sum_{t=1}^{n_1} (\tilde{y}_{1,t} - \bar{\tilde{y}}_1)^2} \right] + \left[1 - \frac{\sum_{t=1}^{n_2} (\tilde{y}_{2,t} - y_{2,t})^2}{\sum_{t=1}^{n_2} (\tilde{y}_{2,t} - \bar{\tilde{y}}_2)^2} \right] \quad (3.49)$$

Na expressão anterior, é válida a notação utilizada na Equação 3.48, lembrando-se que variáveis com o símbolo $\tilde{}$ denotam dados observados, enquanto que aquelas expressas com o símbolo $\bar{}$ referem-se aos valores médios avaliados no tempo.

Em ambas os casos de aplicação (exemplo sintético em 1D e o exemplo real em 3D), verificou-se que as distribuições de probabilidades marginais *a posteriori* foram melhor definidas quando da aplicação do método Bayesiano formal, com forma e limites semelhantes para os dois procedimento de amostragem avaliados. Essa similaridade pode estar relacionada ao baixo número de parâmetros no caso da modelagem hidrodinâmica, em contraposição a outros modelos de sistemas ambientais, como aqueles do tipo chuva-vazão conceituais de base diária, para os quais os métodos MCMC poderiam apresentar superioridade comparativamente aos métodos MC na exploração do espaço multiparamétrico e, assim, na composição da distribuição de probabilidades conjuntas *a posteriori*.

Outro aspecto diferencial no trabalho de Camacho *et al.* (2015) reside no fato de que as funções de verossimilhança pautaram-se em uma avaliação detalhada do comportamento dos resíduos, realizada previamente à inferência Bayesiana, e levando-se em conta a variabilidade dos parâmetros em seus intervalos estipulados *a priori*. Para essa finalidade, foram usados gráficos de dispersão dos resíduos no tempo e de correlação serial, bem como histogramas, sempre avaliados em termos de seus valores médios.

Por fim, proliferaram nos últimos 20 anos avaliações para aferir o efeito das incertezas e da variabilidade nos dados topográficos oriundos de diversas fontes sobre a geração de mapas de inundação, comparando-se geralmente o desempenho da modelagem por meio de gráficos e pelo cálculo de algum indicador que relacione atributos de saída do modelo àqueles disponíveis em dados observados, como área inundada, larguras alcançadas nas seções transversais, e profundidade nas mesmas ou nos *grids* de Modelos Digitais de Terreno. Como exemplos, citam-se os trabalhos de Cook e Merwade (2009) e de Tsubaki e Kawahara (2013).

3.6 Caracterização do modelo hidrodinâmico adotado

3.6.1 Introdução

Para realização das simulações hidráulicas em regime não permanente foi selecionado o modelo HEC-RAS – *River Analysis System*, desenvolvido pelo *Hydrologic Engineering Center* do Corpo de Engenheiros do Exército Americano, em sua versão 5.0.5 (BRUNNER, 2016a; 2016b). Esse software permite a modelagem dos escoamentos a superfície livre em regime permanente e não permanente, em ambiente unidimensional, bidimensional ou combinando-se ambas as abordagens, a fundo móvel ou fixo. Além disso, trata-se de um modelo amplamente utilizado no meio técnico e acadêmico, nacional e internacional, com inúmeras aplicações no setor de recursos hídricos, muitas das quais estão documentadas em artigos científicos e relatórios de projeto, além de fóruns disponíveis na internet alimentados por seus usuários e por alguns de seus criadores.

A escolha do HEC-RAS para prosseguimento desta pesquisa baseou-se em sua flexibilidade, no sentido de ser possível adequar as abordagens – se 1D ou 2D, se permanente ou transiente, além de sua gratuidade, confiabilidade e robustez. Por fim, sua interface gráfica facilita a compilação, a visualização e a análise de resultados.

À medida que o(s) trecho(s) fluvial(is) para aplicação da metodologia foram selecionados, bem como avaliados seus dados disponíveis, priorizou-se a modelagem unidimensional, de maneira que os parágrafos seguintes referem-se a essa abordagem no que concerne ao número de dimensões consideradas no equacionamento do escoamento em canais e leitos fluviais.

O módulo não permanente do modelo HEC-RAS fundamenta-se na resolução numérica das equações de Saint-Venant, que se constituem nas equações do movimento dos fluidos em uma dimensão e adaptadas aos escoamentos à superfície livre, conforme proposição de Barré de

Saint-Venant em 1871. Algumas premissas relevantes para seu desenvolvimento são listadas a seguir, conforme apresentado por Chow *et al.* (1988):

- Escoamentos podem ser modelados em uma dimensão, isto é, a variação de velocidades e das demais variáveis hidráulicas pode ser avaliada na direção dominante do escoamento, geralmente a que coincide com o eixo longitudinal do canal;
- Os regimes de escoamento devem ser o uniforme ou o gradualmente variado, não apresentando alterações bruscas de velocidade e profundidade em distâncias reduzidas, de maneira que prevaleça a distribuição hidrostática de pressões e a aceleração vertical seja desprezível;
- A declividade do álveo deve ser pequena e o leito, fixo, de modo que os efeitos de deposição e erosão sejam pouco importantes;
- Coeficientes de atrito válidos para o escoamento permanente, uniforme e turbulento são aplicáveis, permitindo que os efeitos de resistência ao escoamento possam ser aproximados pelas equações de Chézy ou de Manning; e
- O fluido sob consideração deve ser incompressível e Newtoniano.

3.6.2 Caracterização do módulo hidrodinâmico do modelo HEC-RAS: equações adotadas e método numérico para sua resolução

Parte da explicação que segue foi extraída e adaptada do Manual de Referências Hidráulicas do HEC-RAS – versões 5.0 (BRUNNER, 2016a), a fim de elucidar os principais aspectos concernentes à solução das equações de Saint-Venant por esse modelo, e de apontar possíveis fatores agregadores de incertezas à modelagem hidrodinâmica por ele realizada destinada à propagação de cheias por canais e cursos de água naturais.

A base para a obtenção de hidrogramas e cotagramas ao longo de um canal à superfície livre ou estirão fluvial são as equações da continuidade e da conservação da quantidade de movimento deduzidas para um volume de controle infinitesimal, à luz dos critérios mencionados nos excertos anteriores. Assim:

$$\frac{\partial A_T}{\partial t} + \frac{\partial Q}{\partial x} - q_l = 0, \text{ e} \quad (3.50)$$

$$\frac{\partial Q}{\partial t} + \frac{\partial QU}{\partial x} + gA \cdot \left(\frac{\partial z}{\partial x} + S_f \right) = 0 \quad (3.51)$$

nas quais as seguintes variáveis podem ser quantificadas em uma seção transversal representativa do volume de controle:

A_t é a área molhada total ocupada pela água (CHOW *et al.*, 1988), dada pela soma das áreas efetivas de escoamento, denotadas por A , e das áreas de armazenamento lateral, i.e., nas quais a velocidade é considerada nula, a exemplo de lagoas marginais e regiões próximas aos aterros de pontes (m^2);

Q é a vazão em trânsito (m^3/s);

q_l é contribuição lateral ao escoamento por unidade de comprimento ($m^3/s.m$);

U é a velocidade média (m/s);

z é a cota altimétrica do fundo do leito fluvial ou do canal; e

S_f é a taxa de perda de energia mecânica do escoamento devido ao atrito imposto pelas paredes e pelo fundo do canal (m/m).

Na Equação 3.51, foram contabilizadas a força exercida pela gravidade como força de campo, e as forças de pressão e devidas ao atrito das paredes e do fundo atuantes sobre o líquido, como forças de superfície. Ainda nessa equação, o fluxo de quantidade de movimento atribuído à vazão lateral, q_l , ainda não é considerado, sendo adicionado mais à frente na dedução das expressões finais utilizadas pelo programa. Além disso, somente as áreas ativas, nas planícies (i.e., onde não houver áreas de armazenamento) e na calha, são nela contabilizadas.

No HEC-RAS, a perda de energia devido à rugosidade das paredes e do fundo tem sua declividade, denotada por S_f , estimada por meio da equação de Manning em determinada seção transversal, conforme segue:

$$S_f = \frac{Q|Q|n^2}{R_h^{4/3} A^2} \quad (3.52)$$

em que:

n é o coeficiente de rugosidade de Manning (cuja unidade, para assegurar que a equação 3.44 seja dimensionalmente homogênea, deve ser $s/m^{1/3}$, segundo BAPTISTA e COELHO, 2002); e

R_h é o raio hidráulico (m).

A modelagem 1D de uma onda de cheia em rios naturais impõe o desafio de se quantificar em apenas uma dimensão as distintas dinâmicas existentes entre a calha principal e as planícies de inundação, a depender da vazão em trânsito. Quando a descarga ultrapassa a capacidade da calha principal, inicia-se um escoamento lateral, em direção às planícies. À medida que a vazão e o nível de água aumentam, mais porções da planície transportam água, e geralmente ao longo

de um caminho mais curto do que aquele percorrido pela água na calha principal (vide a Figura 3.8), o que vai depender do grau de meandrização do leito fluvial. Finalmente, no ramo descendente do hidrograma, inicia-se um movimento das águas da planície em direção ao canal.

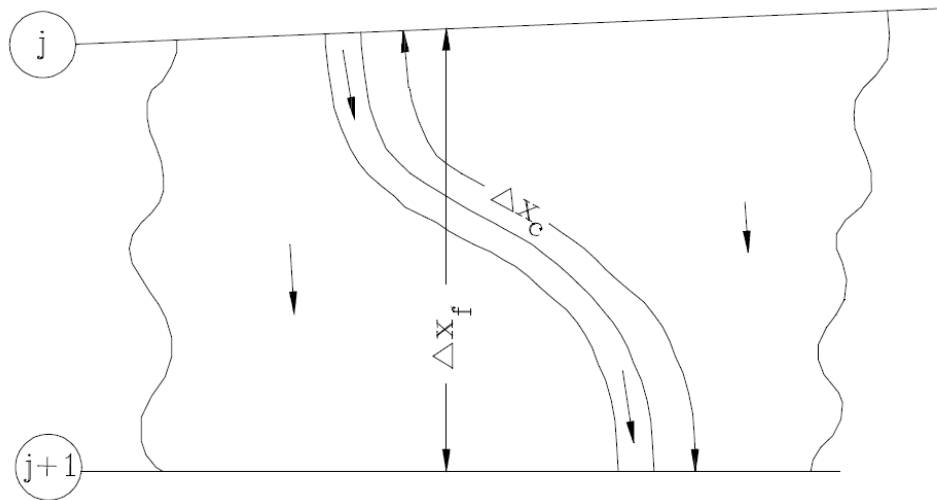


Figura 3.8 – Trajetórias percorridas por partículas fluidas em leitos e planícies fluviais.

Fonte: adaptado de Brunner (2016a).

À luz dessa interpretação, as Equações 3.50 e 3.51 precisaram ser manipuladas para melhor representar esse escoamento, que seria mais adequadamente modelado por 2 dimensões. Fread (1976) e Smith (1978), ambos citados por Brunner (2016a), por exemplo, dividiram o leito em dois canais distintos, e adaptaram as referidas expressões para cada um deles. Para simplificar o problema, assumiram que o nível de água nas planícies e na calha seria o mesmo (de maneira semelhante à considerada pelo modelo HEC-RAS), de modo que a transferência de quantidade de movimento entre tais porções do leito seria pouco significativa, e que a capacidade de transporte de ambas seria proporcional às condutâncias, ou fatores de condução, ou seja:

$$Q_c = \lambda Q \quad (3.53)$$

sendo λ dado por:

$$\lambda = \frac{K_c}{K_c + K_f} \quad (3.54)$$

nas quais:

Q é a vazão total transportada pelo curso de água (m³/s);

Q_c é a vazão total transportada pela calha principal (m³/s); e

K_c e K_f são as condutâncias da calha principal e das planícies de inundação, respectivamente (m^3/s);

A partir de tais hipóteses, e das igualdades 3.53 e 3.54, foi possível reescrever as Equações 3.50 e 3.51 da seguinte maneira:

$$\frac{\partial A_T}{\partial t} + \frac{\partial(\lambda Q)}{\partial x_c} + \frac{\partial[(1-\lambda)Q]}{\partial x_f} = q_l, \quad (3.55)$$

$$\frac{\partial Q}{\partial t} + \frac{\partial(\lambda^2 Q^2/A_c)}{\partial x_c} + \frac{\partial[(1-\lambda^2)Q^2/A_f]}{\partial x_f} + gA_c \cdot \left(\frac{\partial z}{\partial x_c} + S_{fc}\right) + gA_f \cdot \left(\frac{\partial z}{\partial x_f} + S_{ff}\right) = 0 \quad (3.56)$$

nas quais os subscritos c e f indicam, respectivamente, as variáveis hidráulicas e geométricas características da calha principal e da planície de inundação. Além disso:

$\frac{\partial z}{\partial x_c}$ é a declividade da calha principal (m/m); e

$\frac{\partial z}{\partial x_f}$ é a declividade da planície de inundação (m/m).

Para obtenção das variáveis vazão e profundidade por meio das Equações 3.55 e 3.56, foi desenvolvido um esquema numérico implícito baseado em diferenças finitas conhecido por *four-point* (CHOW *et al.*, 1988). As equações por diferenças finitas resultantes mantêm-se não lineares, demandando algum método iterativo para sua solução, como o de Newton-Raphson. A fim de proporcionar maior estabilidade à referida técnica numérica adotada sobre as equações estabelecidas até aqui, Barkau (1982 *apud* BRUNNER, 2016a) definiu um novo conjunto de expressões que se mostraram mais convenientes do ponto de vista computacional. Suas hipóteses principais foram: (i) a introdução do fator de distribuição de velocidades de Boussinesq, β_v , a fim de permitir a combinação dos termos convectivos, e (ii) a definição de uma trajetória equivalente para o escoamento, o que possibilitou a substituição dos termos relativos ao atrito por uma força equivalente (BRUNNER, 2016a). A versão de Barkau corresponde à solução e à abordagem utilizada pelo modelo HEC-RAS. Em termos de diferenças finitas, a equação da continuidade adquire o seguinte formato:

$$\Delta Q + \frac{\Delta A_c}{\Delta t} \cdot \Delta x_c + \frac{\Delta A_f}{\Delta t} \cdot \Delta x_f + \frac{\Delta Y}{\Delta t} \cdot \Delta x_f - \bar{Q}_l = 0, \quad (3.57)$$

em que:

ΔQ é o elemento diferencial de vazão, avaliado em termos médios ponderados entre duas seções transversais consecutivas e dois passos de tempo consecutivos segundo o esquema implícito *four-point* (m^3/s), cuja dedução pode ser conferida em referências como Chow *et al.* (1988) e Brunner (2016a);

A_c e A_f são as áreas ativas de escoamento, na calha e na planície, respectivamente, avaliadas em termos médios ponderados entre duas seções transversais consecutivas e dois passos de tempo consecutivos segundo o esquema implícito *four-point* (m^2);

Δx_c é a distância entre duas seções transversais consecutivas, contabilizada na calha principal (m);

Δx_f é a distância entre duas seções transversais consecutivas, avaliada na planície de inundação (m), sendo igual à média das distâncias entre as margens esquerda e direita das duas seções consideradas, em se tratando do HEC-RAS;

Y é a área das regiões inefetivas de escoamento, ou seja, que contribuem para a continuidade mas não para a quantidade de movimento (m^2), também avaliada em termos médios ponderados entre duas seções transversais consecutivas e dois passos de tempo consecutivos segundo o esquema implícito *four-point*; e

$\bar{Q}_l$ é a vazão incremental, ou lateral, média no intervalo Δt (m^3/s).

Por sua vez, a equação da conservação da quantidade de movimento pode ser reescrita na forma a seguir:

$$\frac{\Delta(Q_c \Delta x_c + Q_f \Delta x_f)}{\Delta t \Delta x_e} + \frac{\Delta(\bar{\beta}_v U Q)}{\Delta x_e} + g \bar{A} \cdot \left(\frac{\Delta z}{\Delta x_e} + \bar{S}_f + \bar{S}_h \right) = \vartheta \frac{\bar{Q}_l \bar{U}_l}{\Delta x_e} e \quad (3.58)$$

em que:

as variáveis ΔQ_c , ΔQ_f , ΔQ e ΔU são avaliadas em termos médios ponderados seguindo-se o esquema implícito *four-point* e já tiveram seu significado elucidado em equações anteriores; $\bar{A}$ é a soma das áreas ativas de escoamento na calha e na planície, avaliada em termos médios entre duas seções transversais consecutivas (m^2);

$\bar{S}_f$ é a taxa média de perda de carga por atrito entre duas seções transversais consecutivas, estimada por meio da Equação 3.52 com todas as variáveis aferidas em termos médios entre duas seções consideradas (m/m);

$\bar{S}_h$ é a taxa de perda de carga devido a contrações ao escoamento e alterações de velocidade induzidas por estruturas hidráulicas, cuja estimativa no HEC-RAS é feita pela Equação 3.63 (m/m); e

$\vartheta \frac{\bar{Q}_l \bar{U}_l}{\Delta x_e}$ é o termo referente à adição de quantidade de movimento por algum tributário, desde que este tenha sido representado geometricamente no sistema fluvial estudado (m^3/s^2) no HEC-RAS, ou seja, com seções transversais. ϑ é a fração de quantidade de movimento que contribuiu para o rio principal, e $\bar{Q}_l$ e $\bar{U}_l$ são, respectivamente, a vazão e a velocidade média associada à contribuição lateral (m/s), ambas em termos médios entre dois passos de tempo consecutivos.

A declividade, ou taxa, da perda de energia pelo atrito é avaliada na Equação 3.58 em termos equivalentes, sendo calculada no HEC-RAS pela seguinte expressão:

$$g \bar{A} \bar{S}_f \Delta x_e = g \bar{A}_c \bar{S}_{fc} \Delta x_c + g \bar{A}_f \bar{S}_{ff} \Delta x_f e \quad (3.59)$$

em que:

os termos com o símbolo $\bar{}$ indicam que os mesmos são quantificados em termos médios entre duas seções transversais consecutivas, enquanto que os subscritos c e f continuam indicando variáveis hidráulicas e geométricas características da calha principal e da planície de inundação, respectivamente;

$\bar{S}_f$ é a perda de carga por atrito considerando-se toda a seção transversal, e quantificada na Equação 3.59 em termos médios entre duas seções transversais consecutivas, a exemplo das demais variáveis; e

Δx_e é a trajetória equivalente (m), dada pela expressão a seguir:

$$\Delta x_e = \frac{\bar{A}_c \Delta x_c + \bar{A}_f \Delta x_f}{\bar{A}} \quad (3.60)$$

O fator de distribuição de velocidades, avaliado em termos médios entre duas seções transversais consecutivas na Equação 3.58, por sua vez, pode ser definido por:

$$\bar{\beta}_v = \frac{(U_c^2 A_c + U_f^2 A_f)}{U^2 A} = \frac{(U_c A_c + U_f A_f)}{U A} \quad (3.61)$$

O que conduz ao termo convectivo na forma apresentada na Equação 3.58:

$$\Delta(\bar{\beta}_v U Q) = \Delta(U_c Q_c) + \Delta(U_f Q_f) \quad (3.62)$$

Por fim, $\bar{S}_h$ representa a declividade da perda de energia devido a forças exercidas pelas restrições ao escoamento criadas por estruturas como pontes, bueiros, comportas e vertedouros, a qual pode ser estimada pela seguinte expressão:

$$\bar{S}_h = C \cdot \frac{U^2}{2g\Delta x_e} \quad (3.63)$$

em que:

C é um coeficiente de perda de carga, definido em função do tipo de estrutura, da velocidade e de sua geometria, que, por simplicidade, é considerado como constante; e

O termo U^2 no caso deve ser avaliado entre duas seções transversais consecutivas nas proximidades da estrutura hidráulica em questão.

A adoção do esquema implícito de diferenças finitas baseado em soluções não lineares, a exemplo do *4-point scheme* para as Equações 3.57 e 3.58 implica na utilização de procedimentos iterativos, o que pode agregar esforço computacional e problemas de convergência na ocorrência de descontinuidades na geometria da leito fluvial (BRUNNER, 2016a). O HEC-RAS emprega a técnica proposta por Chen (1973) e Preissmann (1975), ambos citados por Brunner (2016a) para linearizar as equações da continuidade e da conservação da

quantidade de movimento avaliadas à luz das diferenças finitas, a fim de contornar tais dificuldades. Para solução do sistema linear de equações resultantes, as quais incluem também as condições de contorno e as iniciais, o modelo em questão utiliza um procedimento conhecido por *skyline storage scheme*. Detalhes adicionais podem ser conferidos no Manual de Referências Hidráulicas do *software* HEC-RAS (BRUNNER, 2016a).

3.6.3 Incertezas inerentes à utilização do módulo hidrodinâmico do modelo HEC-RAS

Os métodos, hipóteses, equações e coeficientes adotados para resolver as equações de Saint-Venant na estrutura do modelo HEC-RAS em ambiente unidimensional agregam incertezas de caráter epistêmico e aleatório à modelagem hidrodinâmica, muitas das quais ainda de difícil quantificação segundo uma abordagem probabilística, e, mais além, de cunho Bayesiano. Uma alternativa para a avaliação das incertezas epistêmicas associadas à seleção do modelo poderia ser a adoção do BMA – *Bayesian Model Averaging* (RAFTERY *et al.*, 2005), em cuja estrutura deveriam ser acoplados diversos modelos hidrodinâmicos, inclusive o HEC-RAS. Outra possibilidade que permitiria compreender as limitações e avaliar a variabilidade acarretada pelo arcabouço metodológico do modelo em questão é a realização de uma análise de sensibilidade de seus *outputs*, bem como de sua estabilidade e precisão em face da incerteza de alguns parâmetros e variáveis que compõem ou definem o conjunto das expressões 3.50 a 3.63. Por fim, algumas das fontes de incertezas podem ser tratadas e quantificadas seguindo-se uma metodologia Bayesiana formal ou informal, tal como apresentado nos subitens 3.3.1 a 3.3.3.

No item 3.2 foram percorridos os principais aspectos inerentes à modelagem hidráulica e hidrodinâmica e ao mapeamento de manchas de inundação que agregam incertezas de ordem epistêmica e aleatória à resposta simulada pelos modelos utilizados. No presente item, conhecidas as especificidades do HEC-RAS, é possível aprofundar-se nas fontes mais relevantes de incertezas em seu caso, as quais se encontram listadas nos próximos excertos, e discutidas na sequência. É importante que sua interferência sobre a precisão e a estabilidade do modelo seja avaliada, e, na medida do possível, que sua incerteza sobre os *outputs* seja quantificada.

- Discretização espacial e temporal adotada, ou seja, o espaçamento entre seções transversais e o intervalo de tempo definido para as simulações, sendo ambos utilizados para resolução do esquema numérico mencionado;

- Fator de ponderação do esquema numérico implícito;
- Erro máximo admitido (tolerância) e número máximo de iterações;
- Condições de contorno;
- Existência e representação de estruturas hidráulicas;
- Existência e representação de áreas inefetivas (i.e. de armazenamento);
- Condições iniciais; e
- Coeficiente de rugosidade de Manning.

Primeiramente, a discretização espacial e temporal, notadamente, os termos Δx e Δt , podem afetar a estabilidade numérica e a precisão no cálculo de vazões e níveis de água pelo módulo hidrodinâmico do modelo selecionado. Existem algumas diretrizes práticas que norteiam a seleção preliminar desses intervalos, cujo uso deve ser seguido por testes a fim de verificar a estabilidade, a convergência e a robustez do modelo. Assim, o intervalo máximo entre seções transversais consecutivas no modelo 1D pode ser definido pela seguinte relação (BRUNNER, 2016b):

$$\Delta x \leq \frac{cT_r}{65,6168} \quad (3.64)$$

em que:

Δx é dado em metros;

T_r é o tempo de subida da onda de cheia (s); e

c é a velocidade de propagação da onda de cheia (m/s).

Para grande parte dos cursos de água naturais, a velocidade da onda (ou celeridade), c , pode ser avaliada como sendo a velocidade de uma onda cinemática (CHOW, 1988), ou seja:

$$c = \frac{dQ}{dA}, \quad (3.65)$$

O equacionamento da expressão anterior depende da relação matemática entre a descarga e a profundidade e pode ser de difícil solução analítica. Uma maneira expedita para se obter a celeridade é aproximá-la à velocidade média no leito fluvial, U , multiplicada por um fator, conforme mencionado por Brunner (2016b). Assim, para rios naturais, a relação c/U é cerca de 1,5.

Outra forma de se aferir o espaçamento mínimo entre as seções pode ser avaliado pela expressão a seguir, também apresentada por Brunner (2016b) e atribuída a P. G. Samuel:

$$\Delta x \leq 0,04572 \cdot \frac{D}{S_o} \quad (3.66)$$

na qual:

Δx é dado em metros;

D é a profundidade média da calha principal ao longo do estirão fluvial considerado (m); e

S_o é a declividade média do fundo do leito fluvial ou canal (m/m).

O espaçamento máximo ideal indicado por equações como 3.64 e 3.66 pode ser alcançado por meio de levantamentos topobatimétricos, uma vez conhecido o leito fluvial em estudo. Devido a restrições de custo e de tempo, nem sempre é possível respeitar as condições estabelecidas por equações como as indicadas, de forma que pode ser utilizada a ferramenta de interpolação espacial disponível no HEC-RAS para que sejam criadas seções transversais artificiais a fim de garantir robustez e estabilidade ao modelo. Adicionalmente, é recomendado reduzir o intervalo máximo estimado pelas expressões supracitadas nas proximidades de estruturas ou obstáculos como pontes, degraus artificiais ou naturais, a montante de trechos mais íngremes e nas contrações e expansões significativas do leito fluvial.

Definido o intervalo espacial, é importante que seja avaliado o intervalo temporal mínimo para resolução numérica das equações de Saint-Venant, o que pode ser realizado por meio da equação a seguir, conhecida como *condição de Courant* (COURANT e FRIEDRICHS, 1948 *apud* CHOW *et al.*, 1988):

$$\Delta t \leq \frac{\Delta x}{c} \quad (3.67)$$

na qual c continua sendo a celeridade da onda de cheia (m/s), aferida pela expressão dQ/dA , ou aproximada por $c/U \cong 1,5$ em rios naturais.

Por fim, é importante ressaltar que, ainda que a relação 3.67 seja atendida, intervalos temporais e espaciais amplos podem acarretar dispersão numérica acentuada da onda de cheia, devido à adoção de valores médios no espaço e no tempo de diversas variáveis hidráulicas e geométricas no esquema implícito, comparativamente a intervalos menores que também obedeçam à condição de *Courant*. Sendo assim, quando for viável, é importante que seja avaliada a resposta

e a estabilidade do modelo a distintas combinações entre os intervalos Δx e Δt com potencial de utilização em determinada modelagem hidrodinâmica.

Em segundo lugar, o fator de ponderação *theta*, ou θ , doravante denominado κ a fim de diferenciar o termo em relação ao vetor de parâmetros do modelo hidrodinâmico, denotado por θ , define o valor da derivada espacial e interfere na posição da ordenada temporal no grid de resolução das equações da continuidade e da conservação da quantidade de movimento por diferenças finitas através do esquema implícito do tipo *four-point* (CHOW *et al.*, 1988; BRUNNER, 2016a). Seu valor varia entre 0,0 e 1,0. Valores no intervalo $0,5 < \kappa \leq 1,0$ garantem que teoricamente o método seja incondicionalmente estável; $\kappa = 0,5$ indica que este seja condicionalmente estável, a depender da discretização espaço-temporal utilizada, e instável se $\kappa < 0,5$ (BRUNNER, 2016a). No HEC-RAS, é possível estabelecer manualmente qual valor o parâmetro κ deverá assumir nas simulações, sendo o padrão igual à unidade. Brunner (2016b) sugere que sejam adotados valores entre 0,6 e 1,0, sendo que 0,6 confere maior precisão ao modelo, enquanto que 1,0 garante maior estabilidade numérica. A recomendação final é, uma vez definidos os intervalos Δx e Δt , que se testem a resposta, em termos de precisão, e a estabilidade do modelo elaborado iniciando-se pelo valor padrão de κ , e reduzindo-o em direção a 0,6, desde que o modelo permaneça estável.

Dois aspectos adicionais que podem interferir sobre a precisão, a estabilidade e a convergência do HEC-RAS são a tolerância e o número máximo de iterações a serem avaliadas a cada intervalo de tempo. O padrão estabelecido é de 0,006 m para a tolerância, em termos de profundidades, e de 20 iterações, no segundo caso, mas esses valores podem ser alterados manualmente pelo usuário. O erro máximo tolerado pode ser reduzido, a fim de aumentar a estabilidade do modelo, mas isso poderá aumentar o tempo necessário à simulação (BRUNNER, 2016b).

Adicionalmente, as condições de contorno estabelecidas a jusante no trecho fluvial modelado influenciam o trânsito da cheia ao longo de parte ou de todo o estirão considerado, a depender da declividade e da geometria representadas pelas seções transversais, bem como da faixa de variação das vazões. O módulo hidrodinâmico do HEC-RAS permite que sejam prescritas as seguintes condições de contorno a jusante: (i) cotograma ou níveis de água induzidos por estuários ou reservatórios; (ii) hidrograma, geralmente quando houver disponibilidade de informações de monitoramento hidrométrico; (iii) curva de descarga unívoca, sob a hipótese de

escoamento uniforme ou de existência de estrutura de controle hidráulico, como comportas ou vertedouros; e (iv) profundidade normal, situação válida para trechos com escoamento uniforme. A seleção da condição de contorno a jusante deve ser feita com cautela, em função das informações existentes de monitoramento sistemático de níveis e vazões, de marcas de cheia disponíveis, de operação de reservatórios, e de levantamentos topobatimétricos, além do grau de fidelidade de todos esses dados em representar o trecho fluvial sob estudo e suas condições de escoamento. Brunner (2016b) alerta que condições de contorno que causem mudanças abruptas na superfície da água, ou níveis de água muito baixos em relação à calha principal, podem acarretar oscilações na solução numérica do modelo, induzindo instabilidades e mesmo interrupção da modelagem.

A existência de estruturas hidráulicas transversais ou esconsas ao escoamento, como pontes, também pode interferir na precisão dos perfis de escoamento e do trânsito da cheia, além de contribuir para possíveis instabilidades numéricas. Em função do grau de constrição que acarretam ao escoamento, e da altura livre disponível, as mudanças de velocidade causadas pela estrutura podem ser muito abruptas, demandando maior número de seções transversais em seu trecho de inserção, maior discretização de sua curva de descarga, e, possivelmente, um intervalo temporal global menor para a simulação.

A representação das áreas inefetivas pelo modelo, isto é, de regiões da planície de inundação que contribuem com seus volumes de armazenamento para a continuidade, mas que têm velocidades muito baixas a ponto de não interferir na transferência de quantidade de movimento, também merecem atenção na modelagem hidrodinâmica. Sua existência pode causar descontinuidades na solução numérica, principalmente caso sejam atingidas as elevações a partir das quais tais áreas se tornam ativas, ou seja, quando passam a contribuir como corredor fluvial.

As condições iniciais, geralmente representadas por vazões mais baixas relativamente às descargas de cheia, devem ser selecionadas com cautela, devendo ser consistentes com a primeira ordenada do hidrograma de cheia afluente ao trecho fluvial estudado.

Por fim, o coeficiente de rugosidade de Manning, além de ser dotado dos aspectos que o caracterizam como uma das principais fontes de incerteza na modelagem hidrodinâmica, conforme mencionado no item 3.2, pode interferir também na estabilidade e precisão do

modelo. Quando subestimado, acarreta velocidades elevadas e até mesmo induz a ocorrência de escoamento supercrítico de maneira não realista. Por outro lado, sua superestimação pode causar atenuação excessiva do hidrograma de cheia. O subitem a seguir descreve alguns métodos e aspectos concernentes à definição deste parâmetro interveniente sobre a modelagem hidráulica e hidrodinâmica.

3.6.4 Breve revisão de métodos para definição do coeficiente de rugosidade de Manning em canais

A literatura para a definição dos coeficientes de rugosidade de Manning em cursos de água, sejam naturais ou construídos, é extensa, havendo diversas tabelas listando seus valores para canais e planícies de inundação em função do tipo de solo e/ou da cobertura vegetal ou material dominante como revestimento do leito fluvial. Arcement e Schneider (1989) mencionam que as características mais importantes na definição da rugosidade são: (i) o tipo e o tamanho das partículas dos materiais que compõem o leito e as margens do canal; e (ii) a forma do canal. Dalrymple e Benson (1984) apontam, adicionalmente: (iii) profundidade do escoamento; (iv) vegetação; e (v) alinhamento do canal.

Hicks e Mason (1998) explicam que há 3 métodos dominantes na definição dos coeficientes de rugosidade de Manning para canais. O primeiro deles é o método comparativo, em que a estimativa é feita a partir da comparação entre as características do material dominante do leito, vegetação e índices hidráulicos (a exemplo de R_h/D_{50} , relação entre o raio hidráulico e o diâmetro mediano das partículas do leito, segundo HICKS e MASON, 1998) com aquelas dos canais em que a rugosidade foi estimada a partir de dados de campo (i.e., vazão em trânsito e geometria das seções transversais no trecho avaliado), por meio da aplicação do método área-declividade (DALRYMPLE e BENSON, 1984). Geralmente, as estimativas disponíveis referem-se a condições ideais para estimação pelo supracitado procedimento, propícias ao escoamento uniforme, correspondendo a trechos retos, com declividade suave da linha de água e do fundo, e a vazões confinadas na calha principal. Dessa forma, uma ressalva à aplicação do método comparativo é a necessidade de o canal apresentar semelhanças quanto às condições morfológicas e hidráulicas de obtenção da rugosidade, o que nem sempre ocorre em canais naturais. Além das estimativas do coeficiente de Manning em diversos canais neozelandeses presentes na compilação de Hicks e Mason (1998), as tabelas e figuras de Chow (1959), Barnes (1967) e Dalrymple e Benson (1984) também se referem a valores estimados sob condições

prescritas que devem constituir ponto de atenção quando de sua utilização e extrapolação para cursos de água sob estudo.

A fim de possibilitar a estimação do coeficiente de rugosidade de Manning em trechos fluviais que podem não seguir as condições ideais supracitadas ou aquelas dominantes na obtenção dos coeficientes disponíveis na literatura, o segundo método mencionado por Hicks e Mason (1998) fundamenta-se na definição de um valor de base. Essa estimativa inicial pode ser obtida a partir do método comparativo ou tabelada segundo o material do leito, sendo a ela agregados fatores corretivos, a depender de características do canal, como variações entre as seções transversais, presença de obstruções, densidade da vegetação e grau de meandrização, conforme proposto inicialmente por Cowan (1956 *apud* HICKS e MASON, 1998) sob a forma da equação a seguir, válida para a calha principal:

$$n_c = (n_b + n_1 + n_2 + n_3 + n_4) \cdot m \quad (3.68)$$

em que:

n_b é o valor de base para um canal com escoamento predominantemente uniforme, em um trecho retilíneo, e regime preferencialmente subcrítico;

n_1 é o valor de correção para computar o efeito de irregularidades na superfície de revestimento do canal;

n_2 é o fator de correção devido às variações no tamanho e na forma das distintas seções transversais do canal;

n_3 é o fator de correção devido às obstruções presentes na calha principal;

n_4 é o fator incorporado pela presença de vegetação na calha;

m é um fator multiplicativo definido em função do grau de meandrização do canal; e

n_c é o valor do coeficiente de rugosidade de Manning adaptado às condições do canal sob consideração.

Chow (1959) apresenta valores tabelados para n_b , determinados a partir do material que compõe o leito ou reveste o curso de água, e estabelece uma série de critérios para aferir graus definidores dos fatores corretivos, n_1 a n_4 , e m . Arcement e Schneider (1989) estendem o supracitado trabalho, uma vez que tabelam valores de n_b a partir de diversos estudos, separando os tipos de canal por leito estável ou móvel (i.e., canais cujo leito seja em solo arenoso e com reposição constante de material), e apresentam um detalhamento na determinação dos graus para quantificação das correções na Equação 3.68. Além disso, estes autores sugerem que o método de Cowan possa ser adaptado às planícies de inundação. Sob tal hipótese, n_b continua sendo definido a partir do tipo de revestimento ou de solo que compõe a superfície das planícies, desde que se observe que, nessa situação, não há diferenciação quanto à mobilidade do leito

como no caso da calha principal. Em adição, os fatores corretivos são avaliados à luz de outros critérios, mais adequados a explicar possíveis variações na referida porção de uma seção transversal.

Cabe ressaltar que, independentemente da referência bibliográfica adotada na aplicação do método em questão, as faixas de variação dos fatores corretivos são amplas, principalmente na planície de inundação. Por essa razão, e devido ao fato de o coeficiente de Manning se tratar de um coeficiente empírico, a seleção de cada um dos componentes de n_c na Equação 3.68 está condicionada à existência de informações sobre o leito, a exemplo da curva granulométrica, à possibilidade de inspeção de campo para comprovação das características morfológicas e de composição do canal, bem como à expertise e ao julgamento dos profissionais envolvidos na quantificação de seus valores.

Há, por fim, uma terceira categoria de métodos, baseados em equações envolvendo características geométricas do material do leito, a exemplo de diâmetros notáveis obtidos a partir da curva granulométrica, bem como a declividade do álveo, o raio hidráulico e a velocidade média do escoamento para a faixa de vazões ou a descarga notável a serem consideradas para dimensionamento ou verificação de determinado canal. Nesse caso, geralmente o conjunto de equações é específico a uma determinada classe de materiais, como cascalhos, areias ou matacões, ou limitado a faixas de declividade, vazão e profundidade, devendo essas condições serem conhecidas previamente à sua aplicação. Citam-se os estudos de Simons e Richardson (1966), Bray (1979) e Griffiths (1981) como exemplos de equações empíricas para obtenção da rugosidade de Manning.

Seja qual for a referência adotada, deve-se avaliar cuidadosamente as condições sob as quais a rugosidade de Manning foi estimada para os trechos-modelo, e as etapas de aplicação prescritas em cada estudo. De modo semelhante, quando da utilização de equações empíricas, é preciso verificar antes as faixas de aplicação das variáveis independentes consideradas. A variabilidade de métodos, de valores tabelados para tipologias semelhantes, e de estudos disponíveis é aspecto que contribui para a incerteza na definição de um valor determinístico para esse coeficiente, além daqueles mencionados no item 3.2.

4 METODOLOGIA

4.1 Introdução

Neste capítulo, é feita uma descrição da sequência metodológica seguida na presente pesquisa direcionada à quantificação das incertezas associadas a parâmetros e variáveis envolvidas na modelagem hidrodinâmica voltada ao estudo de cheias, e de seu efeito sobre a incerteza preditiva. A seleção do método e dos critérios utilizados foi realizada a partir da revisão de literatura e da análise judiciosa dos métodos disponíveis em busca de sua potencial aplicação ao tipo de modelo pretendido, observando-se também as informações disponíveis para os trechos fluviais sob avaliação. Assim, o método escolhido será aqui melhor detalhado em relação ao seu conteúdo apresentado na Revisão de Literatura, de forma a mostrar as adaptações necessárias para a aplicação ao estudo de caso.

Na sequência, são mencionados os critérios e premissas adotados no processo de elaboração dos modelos hidrodinâmicos dos trechos fluviais pré-definidos, bem como na seleção e no tratamento de dados necessários para viabilizar o uso do método selecionado para quantificação das incertezas.

Conforme mencionado no capítulo anterior, optou-se pela utilização do aplicativo HEC-RAS - *River Analysis System*, em sua versão 5.0.5 (BRUNNER, 2016a; 2016b).

A Figura 4.1 relaciona os elementos passíveis de serem considerados em uma estrutura de cunho Bayesiano, formal ou informal, para identificação, sob a forma de um processo de calibração, das incertezas associadas a uma modelagem hidrodinâmica, a seus parâmetros e aos seus dados de entrada e de saída.

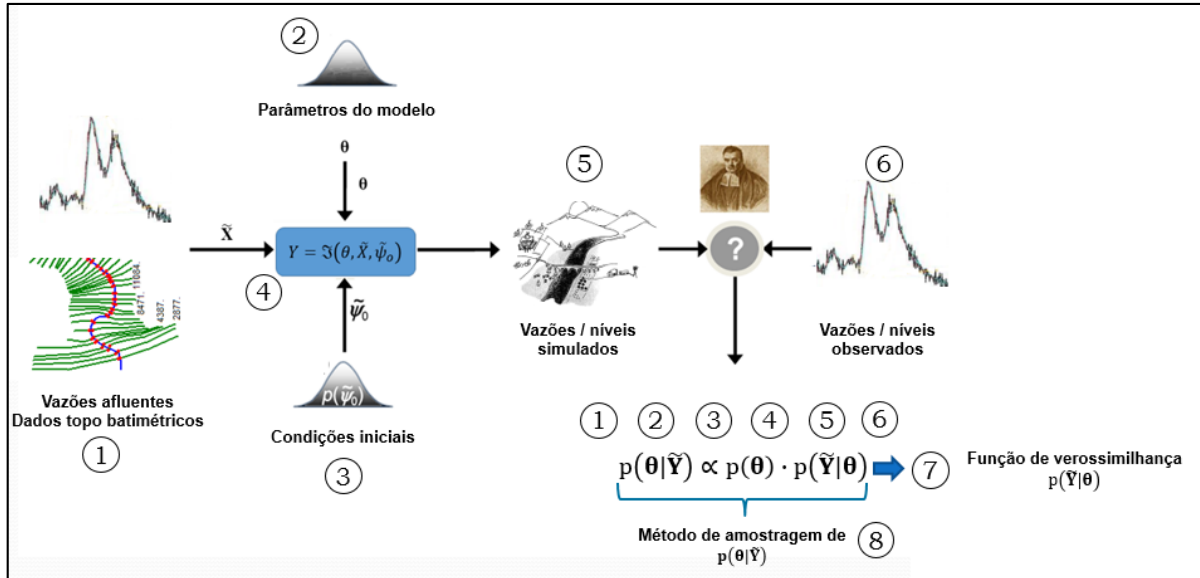


Figura 4.1 – Elementos envolvidos numa estrutura Bayesiana para identificação de incertezas em uma modelagem hidrodinâmica, na forma de um procedimento de calibração.

Os supracitados elementos são descritos a seguir, os quais estão numerados da mesma forma que aquela apresentada na Figura 4.1.

1) Dados de entrada: trata-se de hidrogramas afluentes aos estirões fluviais em estudo, bem como afluições de tributários, representadas por uma descarga constante ou por hidrogramas. Bases topobatemétricas, representadas por seções transversais quando a modelagem é feita em ambiente 1D, e dados cadastrais de estruturas como pontes e bueiros também constituem dados de entrada;

2) Parâmetros do modelo hidrodinâmico: no módulo unidimensional do modelo HEC-RAS, os parâmetros mais relevantes são os coeficientes de rugosidade de Manning. Outros parâmetros podem ser avaliados, a exemplo do fator de ponderação theta e do intervalo temporal das simulações;

3) Condições iniciais: são informações que representam o estado do sistema para dar início às simulações. Podem ser dados de vazão e/ou de nível de água em todas as seções transversais do trecho fluvial (em ambiente 1D), bem como níveis de água em áreas de armazenamento porventura representadas;

4) O módulo hidrodinâmico do modelo HEC-RAS em ambiente unidimensional, cujos principais aspectos foram tratados no subcapítulo 3.6;

5) Dados de saída do modelo, que podem ser, na presente situação: (i) hidrogramas e (ii) cotogramas nas seções transversais existentes; e (iii) manchas de inundação para todo ou parte dos estirões avaliados, elaboradas a partir dos níveis de água máximos alcançados durante determinado evento de cheia;

6) Dados observados do fenômeno modelado, ou seja, informações existentes sobre as cheias no trecho em estudo que possam ser comparadas com os *outputs*: marcas de cheia, manchas de inundação levantadas por imagem de satélite, e cotogramas e hidrogramas registrados em postos fluviométricos. Eventualmente, esses dados podem retratar alguma condição modelada como ideal ou verdadeira, a exemplo de uma mancha de inundação simulada por um modelo hidrodinâmico bidimensional (DI BALDASSARRE *et al.*, 2010);

As distribuições de probabilidades *a priori* também figuram como elementos necessários à aplicação de uma estrutura Bayesiana para quantificação de incertezas, e são utilizadas principalmente para descrever as incertezas estipuladas previamente para os parâmetros do modelo. A depender do método Bayesiano definido, tais distribuições podem ser usadas também para exprimir o conhecimento prévio acerca da variabilidade de dados de entrada e de saída da modelagem e das condições iniciais. Variáveis latentes porventura consideradas no processo, representativas do modelo de resíduos $\mathbf{E}(\boldsymbol{\theta})$, também podem ter suas incertezas quantificadas, prescindindo, sob tal hipótese, de distribuições *a priori*.

7) Função de verossimilhança adotada, elaborada a partir do modelo probabilístico hipotético válido para os resíduos $\mathbf{E}(\boldsymbol{\theta})$, ou de medidas de desempenho comumente utilizadas na modelagem hidrológica e hidráulica.

Ressalta-se que, na Figura 4.1, o Teorema de Bayes foi escrito em sua forma mais utilizada em estudos de quantificação de incertezas inerentes à modelagem ambiental, focada no conjunto paramétrico $\boldsymbol{\theta}$. No entanto, à medida que as outras fontes de incertezas podem ser explícita ou implicitamente contabilizadas, a notação da proporcionalidade tal como apresentada na figura em questão deve ser adaptada e expandida para inserir as variáveis representativas desses elementos adicionais à calibração das incertezas por um esquema Bayesiano.

8) Método de amostragem da distribuição conjunta *a posteriori*, adotado sob a hipótese de que não é possível obter soluções analíticas para a Equação 3.5 (página 31), cerne dos métodos Bayesianos.

A depender da estrutura adotada para inferência das incertezas por métodos Bayesianos, é possível quantificar as incertezas dos elementos 1 (dados de entrada), 2 (parâmetros do modelo), 3 (condições iniciais), 4 (estrutura do modelo), 5 (saídas do modelo) e 6 (dados, ou observações, de saída). O subcapítulo 4.2 elucida quais dentre tais elementos prevaleceram nas análises, e como foram considerados, enquanto que os subcapítulos 4.3 e 4.4 esclarecem aspectos relativos à seleção de trechos fluviais para aplicação da metodologia e às simulações hidráulicas iniciais, para compreensão aprofundada do modelo adotado previamente à realização das milhares de repetições necessárias à amostragem da distribuição conjunta *a posteriori*.

4.2 Esquema metodológico adotado para quantificação de incertezas

Ao longo da revisão de literatura, e à medida que as informações dos trechos fluviais pré-selecionados foram compiladas e analisadas, verificou-se que a aplicação de um método Bayesiano formal seria viável, mediante a adoção de determinados critérios, os quais são mencionados a seguir:

- Priorização da quantificação das incertezas associadas aos parâmetros do modelo HEC-RAS, notadamente aqueles cuja incerteza, ou variabilidade, possa ser explicada por uma distribuição de probabilidades, como é o caso dos coeficientes de rugosidade de Manning;
- As incertezas associadas aos dados de entrada e de saída, às condições iniciais e às demais fontes referentes à modelagem, refletidas juntamente com a incerteza paramétrica no vetor de resíduos $\mathbf{E}(\boldsymbol{\theta}) = \{\varepsilon_1(\boldsymbol{\theta}), \dots, \varepsilon_n(\boldsymbol{\theta})\}$, podem ser representadas em conjunto pela adoção de variáveis latentes ao processo de inferência Bayesiana, as quais precisam ser selecionadas idealmente de modo a caracterizar adequadamente o comportamento estatístico dos referidos resíduos da modelagem realizada. Adicionalmente, os resíduos $\mathbf{E}(\boldsymbol{\theta})$ podem ser avaliados tanto em termos temporais, quanto espaciais, a depender das observações $\tilde{\mathbf{Y}}$ disponíveis; e
- Avaliação das características estatísticas dos resíduos da modelagem e das incertezas obtidas *a posteriori* em função do tipo de informação usada como paradigma $\tilde{\mathbf{Y}}$ para calibração das incertezas (vide excerto 6 à página 94).

Embora a adoção do GLUE tenha se mostrado possível, foi dada prioridade à utilização de um método Bayesiano formal, devido às potencialidades dessa abordagem vislumbradas durante a revisão bibliográfica, à possibilidade de prescrever uma função de verossimilhança mais

adequada aos resíduos específicos da modelagem hidrodinâmica e ao número reduzido de estudos nesse campo envolvendo uma metodologia formal destinada a esse tipo de modelo.

Por outro lado, conforme os estudos avançavam, constatou-se que a separação das incertezas dos dados de entrada, a exemplo de vazões afluentes e dados topobatimétricos, em relação àquelas devidas aos parâmetros e demais fontes provavelmente demandaria maior quantidade de informações para compor os vetores $\tilde{Y}$, a fim de garantir a adequada identificação das incertezas, quando comparada à extensão disponível de registros que se prestariam a essa função, nos trechos fluviais pré-avaliados. Os exemplos de métodos Bayesianos formais utilizados para se proceder a essa separação entre as fontes de incertezas ainda se concentram ao redor de alguns modelos conceituais chuva-vazão, e valeram-se do uso de alguns milhares de vazões médias diárias registradas em estações de monitoramento como paradigma para calibração de incertezas, seja sujeitando todas as fontes de incertezas consideradas separadamente à mesma função de verossimilhança (VRUGT *et al.*, 2008), seja sob a estrutura do BaTEA (KAVETSKI *et al.*, 2006a; 2006b; RENARD *et al.*, 2010; EVIN *et al.*, 2013).

Embora seja provável que as fontes de incertezas e as próprias incertezas envolvidas em uma modelagem hidrodinâmica sejam menos numerosas e menores do que aquelas presentes em uma modelagem chuva-vazão conceitual, demandando menos dados $\tilde{Y}$ para sua quantificação, dificilmente uma cheia de magnitude importante a ser propagada por um estirão fluvial teria mais de uma centena de vazões que caracterizassem adequadamente essa porção do regime hidrológico em uma típica bacia hidrográfica brasileira com algumas dezenas de milhares de quilômetros quadrados. Hutton *et al.* (2011) alertam para a possibilidade de uma identificação inadequada das incertezas paramétricas caso a informação $\tilde{Y}$ seja insuficiente. Kavetski *et al.* (2006a), por sua vez, comentam sobre a necessidade de confirmação das hipóteses que embasam a prescrição das funções de verossimilhança elaboradas para os erros dos dados de entrada, quando da aplicação do BaTEA, por meio da análise de registros cada vez mais extensos dessas informações.

Adicionalmente, no que tange às incertezas agregadas à modelagem em função da precisão e da resolução de bases topobatimétricas, constatou-se que a incorporação desses erros agregaria esforço computacional importante ao sequenciamento da pesquisa, o que poderia inviabilizar as inúmeras verificações planejadas para avaliação da aplicabilidade da metodologia tal como ora delineada.

Finalmente, priorizou-se uma técnica de amostragem para aproximação numérica da distribuição conjunta *a posteriori*, $p(\theta|\tilde{Y})$, conhecida por DREAM - *Differential Evolution Adaptive Metropolis* (VRUGT *et al.*, 2009b), devido às vantagens preconizadas por seus idealizadores, a sua aplicação à modelagem de sistemas ambientais em diversas áreas do conhecimento e à vislumbrada facilidade de implantação e acoplamento ao HEC-RAS.

Feitas as ressalvas à aplicação de métodos Bayesianos à modelagem hidrodinâmica em questão, e delineadas as premissas para sua utilização na presente pesquisa, a Figura 4.2 foi elaborada a partir da figura anterior para representar graficamente as adaptações realizadas e os elementos considerados. Os subitens seguintes detalham esses aspectos.

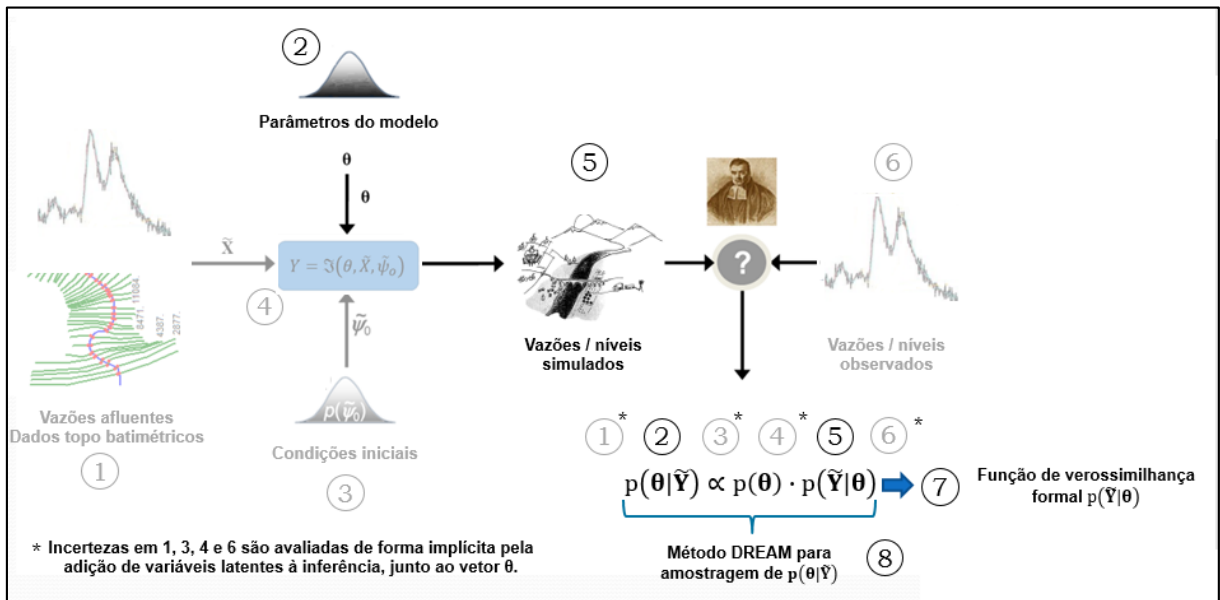


Figura 4.2 – Fontes de incerteza e elementos associados à modelagem hidrodinâmica e ao esquema metodológico adotado considerados na presente pesquisa. Dentre os elementos cujas incertezas podem ser quantificadas, numerados de 1 a 6, os que estão destacados em preto correspondem às fontes de incertezas explicitamente tratadas pela metodologia adotada.

Por fim, após a quantificação das incertezas dos parâmetros e variáveis latentes por meio da adaptação metodológica proposta nesta pesquisa, estipulou-se avaliar seu efeito sobre a predição de algumas variáveis de saída a partir da quantificação da incerteza preditiva, destacando-se: (i) a incerteza preditiva de ordem paramétrica, isto é, propagada através da modelagem sobre variáveis de saída em função das incertezas dos parâmetros inferidas *a posteriori*; e (ii) a incerteza preditiva total, ou seja, sobre as saídas do modelo devido a todas as fontes analisadas, ainda que parte delas seja representada de forma implícita pelas variáveis latentes.

4.2.1 Prescrição da função de verossimilhança

O módulo hidrodinâmico do modelo HEC-RAS baseia-se em um esquema implícito de diferenças finitas para a resolução numérica das equações de Saint-Venant, as quais envolvem derivadas parciais da vazão e da profundidade quanto ao tempo e ao espaço e relações não lineares entre variáveis hidráulicas e geométricas. Por essa razão, espera-se um comportamento não linear entre as variáveis de entrada (hidrograma prescrito no extremo de montante dos trechos estudados e afluições de bacias incrementais) e as condições de contorno e iniciais do sistema, e as variáveis de saída, expressas em termos de hidrogramas e níveis de água nas seções transversais dos estirões fluviais.

Dessa forma, estima-se que esse aspecto, aliado à presença de diversas fontes de incertezas que contribuem para a incerteza final, aferida por meio das diferenças $\mathbf{E}(\boldsymbol{\theta}) = \tilde{\mathbf{Y}} - \mathbf{Y}(\boldsymbol{\theta})$, confira a estas características de autocorrelação, heteroscedasticidade, assimetria e curtose. Para representar tais premissas matematicamente, adotou-se a função de verossimilhança, denotada por $L(\boldsymbol{\theta}|\tilde{\mathbf{Y}})$, conhecida pelo nome de Generalizada, conforme proposição de Schoups e Vrugt (2010), e assim chamada por sua capacidade de representar resíduos com tais características. Para exprimi-las, esses autores propuseram o seguinte modelo para os erros da modelagem $\mathbf{E}(\boldsymbol{\theta}) = \tilde{\mathbf{Y}} - \mathbf{Y}(\boldsymbol{\theta}) = \{\varepsilon_1(\boldsymbol{\theta}) \cdots, \varepsilon_n(\boldsymbol{\theta})\}$, denotados por $\varepsilon_t(\boldsymbol{\theta})$:

$$\Phi_p(\mathbf{B})\varepsilon_t = \sigma_t \cdot a_t \text{ com } a_t \sim \text{SEP}(0,1,\xi,\beta) \quad (4.1)$$

na qual:

$\Phi_p(\mathbf{B}) = 1 - \sum_{i=1}^p \phi_i \cdot B^i$ é um polinômio autoregressivo com p parâmetros ϕ_i ;

$\mathbf{B}$ é o operador de retardo ($B^i \varepsilon_t = \varepsilon_{t-i}$);

σ_t é o desvio-padrão no tempo t ; e

a_t é o erro aleatório i.i.d. com média zero, desvio-padrão unitário, descrito por uma distribuição conhecida por SEP (acrônimo dos termos em inglês *Skew Exponential Power*), com os parâmetros ξ e β associados, respectivamente, à assimetria e à curtose esperadas para esse tipo de erro.

O polinômio autoregressivo supracitado, $\text{AR}(p)$, computa a correlação serial possível de ocorrer em resíduos que apresentam comportamento não linear, a exemplo daqueles resultantes do uso de modelos chuva-vazão (VRUGT *et al.*, 2009a; SCHOUPS e VRUGT, 2010; EVIN *et al.*, 2013). Polinômios autoregressivos de médias móveis, denotados por $\text{ARMA}(p, q)$, também poderiam ser considerados. No entanto, conforme ponderam Schoups e Vrugt (2010), alguns estudos indicam que processos autoregressivos $\text{AR}(p)$ são suficientes e adequados para explicar

resíduos típicos de uma modelagem chuva-vazão. Embora não tenham sido encontrados trabalhos semelhantes com menções acerca da dependência serial dos resíduos correspondentes aos modelos hidrodinâmicos, tal hipótese também foi adotada na presente pesquisa.

A heteroscedasticidade, por sua vez, é considerada ao se assumir que o desvio-padrão dos resíduos, σ_t , aumenta linearmente com a vazão simulada, y_t (alternativamente, y_t também pode representar as elevações da linha de água ao longo do tempo na seção fluvial de saída, ou mesmo os níveis de água máximos alcançados em todas as seções transversais durante a cheia estudada – nesse segundo caso, a informação $\tilde{Y}$ tem caráter espacial). Formalmente:

$$\sigma_t = \sigma_0 + \sigma_1 \cdot y_t \quad (4.2)$$

Na Equação 4.2, σ_0 e σ_1 são, respectivamente, o intercepto e o coeficiente angular e constituem parâmetros do modelo de resíduos e da função de verossimilhança, podendo, portanto, ser inferidos junto aos parâmetros do modelo hidráulico, situação na qual se caracterizam como variáveis latentes.

Schoups e Vrugt (2010) defenderam que o uso de um modelo linear para o desvio-padrão das diferenças entre vazões simuladas por um modelo chuva-vazão e as correspondentes descargas observadas seria adequado devido à crescente incerteza no ramo superior de curvas-chave. Na presente pesquisa, espera-se encontrar o mesmo comportamento, devido às razões mencionadas. Ressalta-se, no entanto, que a condição expressa na Equação 4.2 pode não ser adequada quando se altera o tipo de observação a ser usada como paradigma $\tilde{Y}$, a exemplo dos níveis de água na seção de jusante ou das profundidades máximas atingidas ao longo de um estirão fluvial em decorrência de uma cheia notável. Portanto, a aplicabilidade da relação anterior, e do modelo de erros de maneira geral, foi acompanhada por diversas verificações quanto ao comportamento estatístico dos resíduos.

Finalmente, a distribuição de probabilidades SEP modela o ruído remanescente, denotado por a_t , obtido após a extração da autocorrelação e da heteroscedasticidade dos resíduos da modelagem $\varepsilon_t(\theta)$. Sua FDP – função densidade de probabilidade, adaptável a distintos graus de assimetria e de curtose, é dada pela expressão a seguir, desenvolvida por Schoups e Vrugt (2010) a partir da distribuição EP (acrônimo em inglês de *Exponential Power*), à qual foi conferida assimetria por meio de técnica proposta por Fernandez e Steel (1998):

$$p(a_t|\xi, \beta) = \frac{2\sigma_\xi}{\xi + \xi^{-1}} \cdot \omega_\beta \cdot \exp\left(-c_\beta \cdot |a_{\xi,t}|^{2/(1+\beta)}\right) \quad (4.3)$$

na qual os resíduos $a_{\xi,t}$ constituem os resíduos a_t padronizados por meio da expressão a seguir, de modo que o parâmetro β varie no intervalo $[-1,1]$:

$$a_{\xi,t} = \xi^{-\sin(\mu_\xi + \sigma_\xi \cdot a_t)} \cdot (\mu_\xi + \sigma_\xi \cdot a_t) \quad (4.4)$$

Nas Equações 4.3 e 4.4, os valores de μ_ξ , σ_ξ e c_β e ω_β são calculados como funções dos parâmetros representativos da assimetria, ξ ($\xi > 0$), e da curtose, β , conforme pode ser conferido no trabalho de Schoups e Vrugt (2010, Apêndice A, página 15). Casos especiais da SEP, a depender dos valores assumidos pelo parâmetro β , incluem as distribuições Uniforme ($\beta = -1$), Normal ($\beta = 0$) e de Laplace, ou exponencial dupla ($\beta = 1$).

Considerando-se as Equações 4.1 a 4.4, a função de log-verossimilhança Generalizada tem a seguinte forma, de acordo com dedução elaborada por Schoups e Vrugt (2010, Apêndice B, página 15):

$$\ell(\boldsymbol{\theta}, \boldsymbol{\eta}_\varepsilon|\tilde{\mathbf{Y}}) = n \cdot \log\left(\frac{2\sigma_\xi\omega_\beta}{\xi + \xi^{-1}}\right) - \sum_{t=1}^n \log(\sigma_t) - c_\beta \cdot \sum_{t=1}^n |a_{\xi,t}|^{2/(1+\beta)} \quad (4.5)$$

Na Equação 4.5, os resíduos $a_{\xi,t}$ e os valores de σ_t , σ_ξ e c_β e ω_β são computados conforme mostrado nas equações e parágrafos anteriores ou no trabalho de Schoups e Vrugt (2010, p. 15). Em adição, à semelhança da notação dada à Equação 3.12 (página 38), $\boldsymbol{\theta}$ é o vetor de parâmetros do modelo hidrodinâmico (i.e., coeficientes de rugosidade de Manning), e $\boldsymbol{\eta}_\varepsilon = \{\phi_p, \sigma_0, \sigma_1, \xi, \beta\}$ é o vetor de parâmetros do modelo de erros, representativos dos resíduos ε_t , e considerados como variáveis latentes ao processo de inferência Bayesiana, cujas incertezas serão estimadas conjuntamente com aquelas correspondentes a $\boldsymbol{\theta}$.

Ressalta-se que o modelo de resíduos de Schoups e Vrugt (2010) embute também um parâmetro correspondente a um viés multiplicativo presente nos *outputs* devido a possíveis erros sistemáticos existentes nos dados de entrada e de saída, bem como na estrutura e nos parâmetros dos modelos. No entanto, o mesmo não foi considerado na presente pesquisa devido à expectativa de que não haja esse tipo de erro atuante sobre o vetor $\mathbf{Y}$, quando comparado a $\tilde{\mathbf{Y}}$. Ademais, a estimação desse parâmetro, devido à forma como o mesmo está caracterizado no referido estudo, poderia compensar eventuais fontes de erros não representadas pela função de

verossimilhança Generalizada, ainda que tal representação ocorra de forma implícita por meio da adição das variáveis latentes mencionadas no parágrafo anterior.

Outro ponto importante a ser mencionado refere-se à complexidade agregada à inferência Bayesiana com o acréscimo de variáveis latentes, mais especificamente ao processo de amostragem numérica selecionado. Dessa forma, ao longo dos estudos, procedeu-se a diversas verificações quanto ao comportamento estatístico dos resíduos, a fim de corroborar as hipóteses inicialmente definidas para os mesmos quanto à autocorrelação, heteroscedasticidade, assimetria e curtose, e assim justificar a utilização de um modelo de erros e de uma função de verossimilhança com maior número de parâmetros, e, eventualmente, efetuar ajustes quanto às variáveis latentes de fato consideradas. Adicionalmente, buscou-se compreender, por meio dessas avaliações, as diferenças quanto à estrutura dos resíduos acarretadas pela utilização de distintos tipos de informações como paradigma $\tilde{Y}$ para calibração de incertezas. Tais análises, que refinaram a prescrição do modelo de erros e, assim, da função de verossimilhança a partir do modelo Generalizado, bem como fomentaram o estabelecimento de cenários distintos de calibração de incertezas pelo esquema Bayesiano, são detalhadas no subitem 4.2.5.

4.2.2 Seleção de distribuições de probabilidades *a priori* para os parâmetros estudados

4.2.2.1 Coeficientes de rugosidade de Manning

A rugosidade constitui o principal parâmetro de modelos hidráulicos e hidrodinâmicos unidimensionais cujas incertezas podem ser estimadas adotando-se uma abordagem probabilística, ou seja, atribuindo-se distribuições de probabilidades para quantificar a variabilidade aleatória e epistêmica associada à sua definição. No modelo HEC-RAS, a rugosidade é predominantemente representada em termos do coeficiente de Manning, embora possa ser tratada alternativamente como altura de rugosidade, abordagem menos utilizada.

Na presente pesquisa, a incerteza paramétrica associada à modelagem hidrodinâmica a ser inferida pelo método Bayesiano formal aplicado foi definida pela incerteza associada ao coeficiente de rugosidade de Manning, por ser esse um parâmetro utilizado na estimação da perda de carga entre seções transversais consecutivas e que, portanto, influencia a quantificação de vazões e de profundidades ao longo de um estirão fluvial e a marcha de propagação de cheias. A exemplo de estudos semelhantes de avaliação de incertezas em modelos hidráulicos, a expectativa foi a de que a incerteza atrelada ao coeficiente de rugosidade de Manning explicasse parte importante da incerteza embutida em seus *outputs*.

Devido à extensa bibliografia sobre o tema, aos distintos métodos para estimação do coeficiente de Manning, e à falta de elementos adicionais que possibilitassem a definição de probabilidades diferenciadas para determinados valores desse parâmetro (i.e. inspeção de campo ao longo dos estirões e em seções transversais típicas, seleção de coeficientes com um grupo de engenheiros hidráulicos com ampla experiência), pressupôs-se que o conhecimento *a priori* para o mesmo, tanto na calha quanto nas planícies, fosse traduzido pela distribuição uniforme, com limites inferior e superior correspondentes aos valores extraídos da bibliografia, atribuindo-se equiprobabilidade a quaisquer valores nesse intervalo. É importante ressaltar que, dentre os estudos de incertezas na modelagem hidráulica e/ou hidrodinâmica compilados na Revisão de Literatura, a escolha da distribuição uniforme para o coeficiente de Manning foi unânime, ainda que os critérios para definição dos limites utilizados sejam distintos entre tais trabalhos.

Para quantificação de valores característicos máximos e mínimos para os coeficientes de Manning ao longo dos trechos fluviais a serem estudados e que embasassem a construção de distribuições de probabilidades *a priori* do tipo uniforme, duas abordagens foram estipuladas:

- (i) Para cada seção transversal, separação da porção que constitui a calha principal, com atribuição de seus valores, e, para as planícies de inundação, definição de tipologias de uso e ocupação do solo ou de cobertura vegetal, para posterior estimação dos coeficientes em função de suas características. Assim, nesse caso, o vetor $\theta = \{n_1, \dots, n_d\}$ constitui-se por tantas d dimensões quantas forem as tipologias definidas por trecho fluvial estudado, somando-se a estas a rugosidade da calha principal; ou,
- (ii) Para cada seção transversal, separação entre calha principal e planícies de inundação, com definição de valores para essas duas porções, ao longo de todo o trecho fluvial. Sob tais circunstâncias, o vetor $\theta = \{n_c, n_p\}$ terá duas dimensões, sendo n_c a rugosidade de Manning na calha principal, e n_p a rugosidade de Manning em ambas as planícies (i.e., tanto na margem direita quanto na esquerda).

O tipo de abordagem adotado dependeu do nível de informações existentes em cada trecho fluvial, que garantisse maior ou menor diferenciação quanto às tipologias de uso, ocupação e cobertura vegetal, a exemplo do número de seções transversais em relação à extensão total do estirão, de imagens de satélite disponíveis para a área e de fotos registradas na ocasião dos

levantamentos topobatimétricos. Além disso, em trechos não urbanizados, dados sobre o material predominante no leito, nas margens e nas planícies foram considerados na definição das rugosidades de Manning, quando disponíveis.

4.2.2.2 Variáveis latentes ou parâmetros do modelo de resíduos

As variáveis latentes, que compõem o vetor $\boldsymbol{\eta}_\varepsilon = \{\phi_p, \sigma_0, \sigma_1, \xi, \beta\}$, são parâmetros estatísticos associados ao modelo probabilístico que define os resíduos da modelagem hidrodinâmica unidimensional e cujas incertezas foram inferidas juntamente com aquelas do vetor $\boldsymbol{\theta}$.

Nesse caso, sabendo-se da flexibilidade do referido modelo, em especial da distribuição $SEP(0,1, \xi, \beta)$ em representar os resíduos de uma modelagem após remoção da correlação serial e da heteroscedasticidade, atribuir distribuições *a priori* informativas para esses parâmetros não pareceu ser adequado ou possível, dado que distintos tipos de modelos e informações $\tilde{\mathbf{Y}}$ provavelmente acarretariam distintas estruturas de resíduos, e pouca analogia poderia ser feita de um estudo para outro. Dessa forma, foi estipulado o uso de distribuições de probabilidades *a priori* do tipo uniforme para as variáveis latentes envolvidas, com limites inferior e superior definidos segundo os seguintes critérios:

- Para os parâmetros da SEP, ξ e β , foram usados os respectivos domínios de validade indicados no estudo de Schoups e Vrugt (2010). Como $\xi > 0$, estipulou-se 0,1 como limite inferior e 10 como limite superior. À medida que os cenários de calibração de incertezas pelo esquema Bayesiano foram realizados, procedeu-se a avaliações com os resíduos a_t , notadamente com seus coeficientes de assimetria e de curtose (NAGHETTINI e PINTO, 2007, páginas 36 e 37), a fim de corroborar a seleção dos limites pré-estipulados e sua manutenção nos cenários seguintes. Ressalta-se que os parâmetros ξ e β , embora representativos da assimetria e da curtose das séries de resíduos a_t , não correspondem exatamente aos estimadores de seus coeficientes de assimetria e de curtose, de modo que sua análise deve ser de natureza apenas qualitativa;
- Para os parâmetros associados à heteroscedasticidade, σ_0 e σ_1 , foram feitas avaliações iniciais sobre o desvio-padrão de séries de resíduos ε_t geradas em caráter preliminar, levando-se em conta a variabilidade estipulada *a priori* para os coeficientes de rugosidade de Manning. Esse procedimento foi realizado para distintos tipos de informação $\tilde{\mathbf{Y}}$ e os correspondentes vetores simulados $\mathbf{Y}$ (i.e., seja com vazões defluentes ou níveis de água no

extremo de jusante dos trechos fluviais estudados, ou ainda, níveis de água máximos em todas as seções disponíveis);

- No caso dos parâmetros ϕ_p , associados ao modelo $AR(p)$ para explicar a correlação serial entre sucessivos resíduos ε_t de uma mesma série, os limites inferior e superior foram definidos como -1 e 1, respectivamente. Nessa situação, tratou-se de observar, após a realização de cada cenário de inferência Bayesiana de incertezas, se as raízes complexas da equação $\Phi_p(\mathbf{B}) = 0$ elaborada com as médias estimadas *a posteriori* dos parâmetros ϕ_p estariam fora do círculo unitário, a fim de garantir a estacionariedade do processo $AR(p)$ (MORETTIN e TOLOI, 2006). Nos cenários nos quais um modelo $AR(1)$ mostrou-se adequado e suficiente para explicar a dependência dos resíduos, então os limites finais utilizados *a priori* para identificação da incerteza do parâmetro ϕ_1 foram fixados em 0 e 1, respectivamente, conforme indicações constantes de Schoups e Vrugt (2010) e de Silva *et al.* (2014).

A Tabela 4.1 sintetiza os limites inferiores e superiores considerados para as variáveis latentes nas simulações de calibração Bayesiana de suas incertezas em conjunto àquelas dos coeficientes de rugosidade de Manning.

Tabela 4.1 – Limites inferior e superior, e tipo estipulado para as distribuições de probabilidades marginais *a priori* das variáveis latentes do modelo de resíduos.

Nome	Notação	Unidade	PDF <i>a priori</i>	Limite Inferior	Limite Superior
Desvio-padrão: intercepto	σ_0	m^3/s	UNIFORME	0	1000
Caso 1: $\mathbf{Y}$ e $\tilde{\mathbf{Y}}$ são vazões					
Caso 2: $\mathbf{Y}$ e $\tilde{\mathbf{Y}}$ são níveis de água		m		0	1
Desvio-padrão: inclinação	σ_1	–		0	1
Curtose	β	–		-1	1
Assimetria	ξ	–		0.1	10
Coeficientes de autocorrelação de p -ésima ordem *	ϕ_p	–		-1	1

* limites considerados desde que seja garantida a estacionariedade do modelo $AR(p)$

4.2.3 Seleção de dados para atualização da informação *a priori*

A utilização de métodos de cunho Bayesiano para estimação de incertezas prescinde de dados observados sobre o fenômeno modelado, que permitam a atualização do conhecimento prévio acerca dos parâmetros envolvidos e da variabilidade de variáveis latentes porventura consideradas no processo. Na modelagem hidrodinâmica, essa informação, que compõe o vetor

$\tilde{Y}$ na terminologia adotada neste trabalho, é geralmente de caráter temporal e associada às variáveis vazão em trânsito e às correspondentes cotas altimétricas de nível de água ao longo do tempo, disponíveis em seções que disponham de monitoramento sistemático.

Os registros históricos de cota e vazão nas estações fluviométricas existentes nos trechos fluviais pré-selecionados foram avaliados em busca de cheias importantes, com ocorrência de extravasamentos generalizados, e que tenham sido monitoradas a montante e a jusante. A hipótese estipulada previamente para sua adoção nesta pesquisa foi a de calibração das incertezas paramétricas e das demais fontes e a obtenção das correspondentes incertezas preditivas considerando-se cada cheia em separado, seja pelo uso de hidrogramas, seja de cotogramas. No entanto, outros esquemas podem ser aplicados, a exemplo de calibrações sucessivas realizadas com eventos de cheia de magnitude semelhante, em busca de maior refinamento das distribuições *a priori*, ou mesmo uma única calibração que se valha das referidas cheias formando somente um – e mais extenso – vetor de observações $\tilde{Y}$.

É imperioso destacar que as distribuições obtidas *a posteriori* sejam válidas para eventos de magnitude semelhante àquela das cheias utilizadas como paradigma de estimação de incertezas. As distribuições estimadas *a posteriori* podem, assim, constituir objeto de validação, em que intervalos de credibilidade construídos para a incerteza preditiva do modelo associada a outra cheia distinta daquela usada na calibração contenham seus valores observados.

Alternativamente, foram prospectados bancos de imagens de satélite disponíveis para as regiões de inserção dos trechos fluviais pré-selecionados em busca de manchas de inundação associadas a eventos de cheia de magnitude relevante. Nesse caso, a informação $\tilde{Y}$ de atualização do conhecimento *a priori* teria caráter espacial, caracterizada pelas cotas altimétricas dos níveis máximos alcançados nos locais das seções transversais disponíveis em decorrência da passagem de uma cheia de referência. A expectativa dessa tarefa foi a de que escalas distintas de informação (i.e., indexadas temporal ou espacialmente) induziriam diferentes estruturas dos resíduos da modelagem, de modo que seria importante aferir seu impacto sobre as incertezas paramétricas obtidas *a posteriori* e sobre a incerteza preditiva.

4.2.4 Técnica de amostragem selecionada para aplicação do método Bayesiano de quantificação das incertezas

Para obtenção numérica das distribuições conjunta e marginais *a posteriori* do vetor paramétrico θ e do vetor de variáveis latentes, η_{ϵ} , selecionou-se o método de amostragem do tipo MCMC conhecido por *Differential Evolution Adaptive Metropolis*, doravante denominado DREAM, e proposto por Vrugt *et al.* (2009b).

Seu algoritmo, a exemplo de outros baseados em múltiplas cadeias que evoluíram na esteira de seu precursor Metropolis-Hastings, usa um processo chamado evolução diferencial para realizar a evolução populacional com uma regra de seleção como a de Metropolis (Equação 3.39, página 60) para avaliar se pontos amostrados no espaço multiparamétrico devem ou não substituir seus estados atuais. O conceito de população aqui denota o conjunto formado pelos d parâmetros dos vetores θ e η_{ϵ} em um estado qualquer $t - 1$ caracterizado pela T -ésima iteração ao longo das N cadeias que evoluem em paralelo, ou seja, $\{\theta_{t-1}^1, \eta_{\epsilon,t-1}^1, \dots, \theta_{t-1}^N, \eta_{\epsilon,t-1}^N\}$. A escala e a orientação dos saltos são ajustadas continuamente a partir de parâmetros adicionais relacionados a distribuições de probabilidades ou a funções dependentes do número total de parâmetros d inferidos no processo Bayesiano. Tais algoritmos são indicados para lidar com problemas envolvendo parâmetros possivelmente correlacionados entre si, com distribuições-alvo multimodais e de caudas pesadas e com espaços multiparamétricos que possam apresentar inúmeras regiões de máximos locais. O uso de várias cadeias evita a convergência prematura e permite o uso de diversas estatísticas de teste para avaliação da convergência (VRUGT, 2016).

Os avanços do DREAM em relação a outros métodos de múltiplas cadeias baseados em algoritmos evolutivos referem-se: (i) à seleção de somente alguns parâmetros por estado $t - 1$ para realização dos saltos; (ii) à detecção de *outliers*, ou seja, de pontos atípicos em uma ou mais cadeias que possam acarretar a exploração inadequada do espaço multiparamétrico e comprometer o desempenho da simulação; e (iii) ao uso de um número menor de cadeias sem comprometer a convergência. Tais vantagens podem ser interessantes para lidar com a modelagem de sistemas ambientais não lineares envolvendo até dezenas de parâmetros (VRUGT, 2016). Outra vantagem do DREAM é a sua disponibilidade em linguagem do software Matlab (VRUGT *et al.*, 2008, 2009b), sob a forma de várias sub-rotinas que realizam as simulações em N cadeias de Markov simultâneas, avaliam sua convergência, e também fornecem diversos resultados sob a forma de gráficos e tabelas.

Cabe ao usuário definir alguns parâmetros internos do método e fornecer informações sobre o modelo ambiental em estudo. Dentre tais elementos, a seguir são mencionados os mais relevantes para o sequenciamento das simulações para aproximação de $p(\boldsymbol{\theta}, \boldsymbol{\eta}_\epsilon | \tilde{\mathbf{Y}})$ pelo DREAM:

- Número de cadeias de Markov, denotado por N . Vrugt (2016) menciona que sua quantidade pode ser menor do que os d parâmetros inferidos no processo, devido à amostragem de subespaços paramétricos. No entanto, deve-se garantir que $N \geq 2\delta + 1$, em que δ é um parâmetro interno que controla os ajustes da distribuição norteadora dos saltos, e tem o valor *default* de 3, sugerido em Vrugt (2016);
- Número total de amostras (ou ainda, realizações ou iterações, denotadas por T) de cada parâmetro em cada cadeia;
- Número e especificação de cada um dos parâmetros a serem amostrados ou inferidos, contabilizando-se os parâmetros do modelo ambiental em questão, no caso, o modelo hidrodinâmico, e os parâmetros do modelo de resíduos ou variáveis latentes consideradas;
- Prescrição da função de verossimilhança a ser utilizada;
- Limites mínimo e máximo para exploração de cada um dos parâmetros, e especificação de suas distribuições de probabilidades marginais estabelecidas *a priori* (i.e., distribuições Beta, Gama, Normal, Log-Normal, Uniforme ou, ainda, expressões definidas pelo usuário);
- Fornecimento do vetor de dados observados $\tilde{\mathbf{Y}}$ que servirá como paradigma de calibração das incertezas sob a ótica Bayesiana.

Finalmente, deve ser acoplada à rotina principal do DREAM uma sub-rotina que defina e simule o modelo $\mathfrak{S}$ para o sistema avaliado em questão, a fim de que suas milhares de realizações, ou *outputs*, sejam confrontadas com os dados observados a cada iteração em cada cadeia de Markov. Para tanto, foi elaborado um *script* também em Matlab para acionamento de comandos do HEC-RAS Controller (GOODELL, 2014; LEON e GOODELL, 2016), um conjunto de rotinas desenvolvidas pelo USACE para automatizar diversas etapas de modelagem no HEC-RAS, tais como entrada de dados, simulações sob distintos cenários de vazão e/ou geometria e plotagem de resultados. No presente caso, o comando principal acionado pelo *script* elaborado referiu-se à alteração dos coeficientes de rugosidade de Manning definidos em cada seção

transversal do modelo hidrodinâmico, em função dos valores amostrados nas cadeias de Markov.

Uma vez finalizado cada processo, ou cenário, de calibração Bayesiana, composto por milhares de simulações no HEC-RAS, o DREAM fornece várias saídas sob a forma de tabelas e gráficos que auxiliam na avaliação da qualidade da amostragem, da convergência do método em alcançar a distribuição-alvo estacionária (i.e., $p(\boldsymbol{\theta}, \boldsymbol{\eta}_\epsilon | \tilde{\mathbf{Y}})$, no presente caso), e na obtenção de informações subsidiárias de novas análises para continuidade da inferência Bayesiana, tais como:

- Gráficos com a evolução do valor amostrado para cada parâmetro, ao longo das N cadeias, ilustradas em paralelo;
- Amostras completas de cada parâmetro, ao longo de cada cadeia. Isso significa que, ao final do processo aproximativo, cada parâmetro terá $N \cdot T$ valores, ou amostras. Além disso, o DREAM devolve um arquivo de saída que combina os valores dos parâmetros ao longo de todas as cadeias. Assim, nesse arquivo, cada ciclo terá N valores dos parâmetros, extraídos da T -ésima iteração em cada cadeia. Por conceito, como cada cadeia de Markov parte de pontos distintos do espaço multiparamétrico e alcança a mesma distribuição estacionária (a distribuição-alvo, ou distribuição conjunta *a posteriori*), os resultados podem ser extraídos mesclando-se as amostras das cadeias, uma vez atingida a convergência;
- Valor da estatística de teste de Gelman-Rubin (1992) para cada parâmetro, a qual compara e utiliza em seu cálculo as variâncias de determinado parâmetro obtidas ao longo de todas as cadeias e entre as mesmas. Para que a convergência seja assegurada, deve-se garantir que seu valor seja igual ou menor do que 1,2, para cada um dos parâmetros (VRUGT, 2016; GILKS *et al.*, 1998). Embora outros testes possam ser utilizados, o de Gelman-Rubin é considerado como ferramenta de análise robusta e confiável para essa avaliação;
- Valor da taxa de aceitação, associada à probabilidade de aceitação pela regra de Metropolis (vide Equação 3.39, página 60), e calculada a cada $N \cdot T = 400$ amostras pelas sub-rotinas que compõem o DREAM. Vrugt (2016) sugere que taxas entre 15% e 30% indicam um desempenho adequado de métodos MCMC; e
- Arquivos de saída contendo as informações listadas a seguir:

- (i) Coeficientes de correlação linear médios entre pares de parâmetros, considerando-se as $N \cdot T$ amostras de cada um deles. Tais estimadores são importantes para se avaliar se há redundância na prescrição de parâmetros, sobretudo quanto às variáveis latentes;
- (ii) Para cada parâmetro, em cada cadeia: coeficientes de correlação serial para os lags 1, 5, 10 e 25. Valores muito elevados, acima de 0,80, para *lags* iguais ou maiores do que 5, indicam uma baixa taxa de diferenciação entre os valores amostrados em cada cadeia (VRUGT, 2016). Tais coeficientes auxiliaram na definição dos *lags* a serem saltados quando da elaboração das amostras de cada parâmetro para obtenção de suas distribuições marginais *a posteriori*;
- (iii) Descritores estatísticos dos parâmetros: média e desvio-padrão. São tomados os valores médios considerando-se as $N \cdot T$ amostras de cada um deles.

Os valores indicados pelos arquivos supracitados nos excertos (i) a (iii) devem ser tomados como referência aproximada para diagnóstico e análise de resultados, pois utilizam as $N \cdot T$ amostras totais de cada parâmetro. Uma vez comprovada a convergência, pelo critério de Gelman-Rubin, cabe ao usuário selecionar o tamanho final das amostras e construí-las para proceder à obtenção de descritores estatísticos e das distribuições marginais *a posteriori* dos parâmetros. Para tanto, deve-se excluir o período de *burn-in*, ou de aquecimento, necessário para que o último parâmetro alcance a convergência. Finalmente, é recomendável saltar k *lags*, como estratégia para garantir uma ampla amostragem dos parâmetros, em que o coeficiente de autocorrelação de ordem k seja menor ou igual a 0,80, considerando-se determinado parâmetro ao longo de suas cadeias.

4.2.5 Análise das incertezas e resíduos obtidos *a posteriori*

Uma vez compreendidos os critérios para realização das simulações MCMC pelo método DREAM, foi possível proceder à calibração Bayesiana das incertezas associadas aos parâmetros do modelo hidrodinâmico e às variáveis latentes previstas como adequadas para explicar os resíduos gerados na modelagem. Para tanto, foram traçados diversos cenários, a serem expostos no próximo subitem, a fim de se obter um panorama da aplicabilidade do esquema metodológico proposto à modelagem em questão.

Independentemente do cenário avaliado, em todos os casos foi importante analisar o comportamento conjunto e marginal inferido *a posteriori* dos parâmetros e variáveis latentes pré-definidos, com base nas saídas fornecidas pelas subrotinas do método DREAM, bem como aferir as características dos resíduos $\mathbf{E}(\boldsymbol{\theta}) = \tilde{\mathbf{Y}} - \mathbf{Y}(\boldsymbol{\theta}) = \{\varepsilon_1(\boldsymbol{\theta}) \cdots, \varepsilon_n(\boldsymbol{\theta})\}$ obtidos para todas as realizações do vetor paramétrico $\boldsymbol{\theta}$ obtido *a posteriori*, conforme explicado a seguir.

4.2.5.1 Avaliação estatística dos resíduos

De modo a garantir uma adequada parametrização da função de verossimilhança previamente definida, e de assegurar sua aderência aos resíduos da modelagem hidrodinâmica unidimensional, foram realizadas as análises descritas a seguir para todos os cenários de calibração ou inferência Bayesiana das incertezas dos parâmetros dos vetores $\boldsymbol{\theta}$ e $\boldsymbol{\eta}_\varepsilon$. Para tanto, foram estudadas as características estatísticas das séries de resíduos geradas a partir das diferenças $\mathbf{E}(\boldsymbol{\theta}) = \tilde{\mathbf{Y}} - \mathbf{Y}(\boldsymbol{\theta}) = \{\varepsilon_1(\boldsymbol{\theta}) \cdots, \varepsilon_n(\boldsymbol{\theta})\}$ obtidas para toda realização i do vetor paramétrico $\boldsymbol{\theta}$ obtido *a posteriori*. Ademais, as correspondentes séries de resíduos geradas após remoção da correlação serial e, posteriormente, da heteroscedasticidade, também foram avaliadas.

Ressalta-se que as avaliações ora descritas para os resíduos da modelagem hidrodinâmica constituíram caminho natural para definição de cenários de calibração das incertezas, uma vez que, feitas as primeiras simulações MCMC, percebeu-se a necessidade de realizar sucessivas alterações em busca da prescrição de variáveis latentes, dentre aquelas previamente definidas, que se ajustasse adequadamente aos dados $\tilde{\mathbf{Y}}$ e garantisse robustez à identificação das incertezas de $\boldsymbol{\theta}$ e do próprio vetor $\boldsymbol{\eta}_\varepsilon$. A análise dos resíduos constitui-se pelos seguintes procedimentos:

- Avaliação da correlação serial das séries de resíduos ε_t e $\Phi_p(\mathbf{B})\varepsilon_t = \sigma_t \cdot a_t$:

A ideia central dessa verificação foi definir a ordem p do modelo autoregressivo que caracteriza a correlação serial dos resíduos ε_t . Para tanto, iniciou-se o processo de inferência Bayesiana utilizando um modelo AR(1), e conferiu-se a remoção de correlação serial garantida pelas diferenças $\varepsilon_t - \phi_1 \cdot \varepsilon_{t-1}$, com ϕ_1 inferido nessa 1ª tentativa. Caso fosse detectada a existência de autocorrelação residual, então novas amostragens deveriam ser realizadas com modelos AR(p) de maior ordem.

A avaliação da correlação serial deu-se por meio de: (i) estimação dos coeficientes de autocorrelação amostral de 1ª ordem das séries de resíduos ε_t e $\Phi_p(\mathbf{B})\varepsilon_t = \sigma_t \cdot a_t$ (i.e., após

remoção de correlação serial até a ordem p), de modo a se verificar o grau de remoção garantido pelo modelo $AR(p)$, e a (ii) construção de correlogramas e correlogramas parciais das séries ε_t para se verificar a adequação dos modelos $AR(p)$ para modelar sua dependência serial.

- Análise da variância das séries de resíduos ε_t :

Nessa situação, a expectativa residuiu em se avaliar a necessidade do parâmetro σ_1 para explicar o desvio-padrão dos resíduos ε_t , ou seja, sua heteroscedasticidade. Assim, os primeiros cenários de inferência Bayesiana partiram do princípio de que um modelo de erros heteroscedásticos seria mais adequado à modelagem hidrodinâmica. A análise realizada quanto a esse aspecto foi essencialmente visual, realizada por meio da elaboração de gráficos relacionando os dados observados $\tilde{Y}$ (eixo x) com os correspondentes resíduos ε_t (eixo y), e dos primeiros (eixo x) com o quadrado dos segundos (eixo y).

- Quantificação de descritores estatísticos das séries de resíduos ε_t , $\Phi_p(\mathbf{B})\varepsilon_t = \sigma_t \cdot a_t$ e a_t , para cada um de seus intervalos temporais: média, mediana, moda, desvio-padrão, coeficiente de variação, coeficientes de assimetria e de curtose, e percentis 2,5%, 25%, 75% e 97,5%;
- Obtenção de histogramas para as séries formadas pelos percentis 2,5%, 25%, 50%, 75% e 97,5% dos sucessivos intervalos temporais das séries ε_t , $\Phi_p(\mathbf{B})\varepsilon_t = \sigma_t \cdot a_t$ e a_t , e cálculo das correspondentes frequências relativas e densidades de probabilidade empírica (i.e., dos supracitados resíduos); e
- Cálculo das probabilidades $p(a_t|\xi, \beta)$ pela Equação 4.3 (página 100), considerando-se a cada vez os valores de percentis 2,5%, 25%, 50%, 75% e 97,5% inferidos *a posteriori* para os parâmetros ξ e β , para quantis teóricos da variável a_t entre -5 e 5 (valores usualmente encontrados para os resíduos a_t oriundos da modelagem hidrodinâmica em diversos cenários testados).

Na sequência, foi feita a sobreposição das densidades empíricas, estimadas de acordo com o segundo excerto anteriormente descrito, com as densidades de probabilidade teóricas $p(a_t|\xi, \beta)$ calculadas conforme o parágrafo anterior, para cada um dos percentis de 2,5%, 25%, 50%, 75% e 97,5%. Essa análise permitiu obter um panorama geral da aderência da distribuição SEP aos resíduos após remoção da dependência serial e da heteroscedasticidade. O mesmo foi feito utilizando-se séries de resíduos a_t para determinadas realizações do vetor paramétrico θ e as

correspondentes probabilidades $p(a_t|\xi, \beta)$, com ξ e β extraídos de ao menos uma iteração i sorteada aleatoriamente, bem como correspondentes à iteração i que conduziu ao valor máximo da função de log-verossimilhança em cada cenário de inferência Bayesiana.

As análises descritas nos cinco excertos anteriores foram realizadas para todos os tipos de informações disponíveis utilizadas como paradigma $\tilde{Y}$ na calibração das incertezas.

4.2.5.2 Quantificação e avaliação da incerteza paramétrica e das variáveis latentes

Para construção das amostras de conjuntos paramétricos $\{\theta, \eta_e\}$ obtidas *a posteriori*, foram consideradas as seguintes premissas, as quais serviram como norteadoras dos parâmetros internos a serem definidos previamente às simulações MCMC:

- O período de *burn-in*, a ser excluído das amostras finais, deve ser definido como o número de iterações T^* utilizado pelo parâmetro que alcance a convergência por último, avaliada segundo a estatística de Gelman e Rubin (1992), devendo essa ser igual ou inferior a 1,2. Assim, podem ser geradas $N \cdot (T - T^*)$ iterações (ou amostras) dos parâmetros ao longo das N cadeias;
- Após remoção do período de *burn-in*, devem ser saltados k lags das amostras remanescentes de cada parâmetro a fim de compor as amostras finais, para garantir uma amostragem adequada de cada cadeia, representativa, e, ao mesmo tempo, extensa o suficiente para o total de parâmetros a serem inferidos. O valor de k pode ser definido como sendo aquele que assegure que a autocorrelação de ordem k para qualquer parâmetro, ao longo de qualquer cadeia, seja igual ou inferior a 0,80. Dessa maneira, as amostras finais a serem usadas para caracterização do comportamento *a posteriori* compõem-se por $m = N \cdot (T - T^*)/lag$ valores de cada parâmetro ou variável latente; e
- Os valores de T devem ser indicados na subrotina principal do DREAM de modo a garantir a criação de amostras de comprimento m , o qual permita a obtenção de descritores não enviesados para os parâmetros, bem como para sua incerteza e a incerteza preditiva. Como o número total de parâmetros deve flutuar ao redor de 3 e 8, a depender da confirmação das hipóteses traçadas para o modelo de resíduos e da correspondente função de verossimilhança, o ideal é que m oscile entre 5000 e 10000, seguindo-se recomendações constantes da compilação de VRUGT (2016).

Respeitadas tais premissas, ao final de cada cenário de simulações MCMC, obtiveram-se $m = N \cdot (T - T^*) / lag$ amostras para cada parâmetro ou variável latente, sobre as quais foram realizados os seguintes procedimentos:

- Estimação dos seguintes descritores estatísticos: média, mediana, moda, desvio-padrão, coeficiente de variação, coeficientes de assimetria e de curtose, e percentis 2,5% e 97,5%;
- Obtenção de histogramas, utilizando-se o intervalo de classes indicado pela fórmula de Sturges (1926 *apud* NAGHETTINI e PINTO, 2007);
- Construção das curvas de distribuição empírica, ordenando-se os quantis de forma crescente e calculando-se as correspondentes posições de plotagem pela fórmula de Weibull (NAGHETTINI e PINTO, 2007); e
- Em se tratando dos parâmetros do modelo hidrodinâmico, avaliação da aderência de algumas distribuições de probabilidade de variáveis contínuas por alguns testes de hipóteses, como o do Qui-Quadrado (NAGHETTINI e PINTO, 2007), o de Kolmogorov-Smirnov (CRUTCHER, 1975), o de Anderson-Darling (STEPHENS, 1979) e o de Filliben (NAGHETTINI e PINTO, 2007).

Finalmente, os descritores e gráficos elaborados conforme exposto nos excertos anteriores permitiram realizar a comparação entre as distribuições marginais *a posteriori* dos parâmetros dos vetores θ e η_e obtidas entre todos os cenários de calibração de incertezas previstos, auxiliando na decisão quanto à remoção de uma ou mais variáveis latentes em conjunto com as análises feitas sobre os resíduos, e na busca pela identificação mais robusta e adequada das incertezas da modelagem hidrodinâmica em questão.

4.2.5.3 Quantificação e análise da incerteza preditiva

No que tange à quantificação da incerteza preditiva *a posteriori*, realizada após o procedimento de calibração Bayesiana das incertezas paramétricas e das demais fontes em seus inúmeros cenários, foram construídos intervalos de credibilidade para futuras observações das variáveis $\tilde{Y}$ (sejam essas vazões, sejam níveis de água em determinadas seções ou no extremo de jusante dos trechos fluviais, ou, ainda, cotas máximas de inundação ao longo dos mesmos), doravante especificadas por $\tilde{y}_p$ e pelo vetor $\tilde{Y}_p$, associando-se a incerteza dos parâmetros do modelo hidrodinâmico à incerteza oriunda de outras fontes referentes à modelagem e à observação/medição de dados de entrada e de saída, representada pelas variáveis latentes. A

noção de intervalos de credibilidade sob a ótica Bayesiana foi apresentada de forma genérica na Equação 3.15 (página 41), válida para uma distribuição de probabilidades $p(w)$ de uma variável aleatória. Na condição ora prescrita, $p(w)$ deve equivaler à distribuição preditiva *a posteriori* de $\tilde{\mathbf{Y}}$, a qual pode ser obtida estendendo-se o conceito embutido na Equação 3.18 (página 42), elaborado inicialmente para $\mathbf{Y}$, sobre o qual foram computadas somente as incertezas paramétricas, considerando-se presentemente todas as fontes de incertezas avaliadas.

Sendo assim, foi necessário propagar as incertezas *a posteriori* paramétricas por meio do uso do modelo hidrodinâmico, e somá-las às incertezas das demais fontes, representadas pelas variáveis latentes, pois, nesse caso, não é possível traduzi-las sob a forma de parâmetros ou dados de entrada a serem inseridos no HEC-RAS. Sob tais hipóteses, os percentis da variável $\tilde{y}_p$, associados a uma probabilidade de não excedência $1 - \alpha$, doravante nomeados como $\tilde{y}_{1-\alpha}$, podem ser obtidos a partir da seguinte relação, adaptada de Schoups e Vrugt (2010), em sua forma vetorial:

$$P(\tilde{\mathbf{Y}}_p \leq \tilde{\mathbf{Y}}_{1-\alpha} | \tilde{\mathbf{X}}) = P\left([\mathbf{Y}(\tilde{\mathbf{X}} | \boldsymbol{\theta}) + \mathbf{E}(\boldsymbol{\eta}_e)]_{1 \dots J} \leq \tilde{\mathbf{Y}}_{1-\alpha} | \tilde{\mathbf{X}}\right) = 1 - \alpha \quad (4.6)$$

Na qual J conjuntos formados por parâmetros e variáveis latentes $\{\boldsymbol{\theta}, \boldsymbol{\eta}_e\}$ são sorteados aleatoriamente a partir da distribuição conjunta *a posteriori* correspondente às m amostras de comprimento d obtidas pela aplicação do método MCMC em questão, e construídas conforme mencionado nos subitens anteriores. Os valores de $\boldsymbol{\theta}$ são usados para modelar J séries temporais para as saídas do modelo, enquanto que os correspondentes valores de $\boldsymbol{\eta}_e$ são utilizados para gerar as respectivas séries de resíduos, denotadas na Equação (4.6) pelo vetor $\mathbf{E}(\boldsymbol{\eta}_e)$. Por fim, para cada sorteio em J , as ordenadas de ambas as séries geradas são somadas para cada passo de tempo t , a fim de representar a incerteza preditiva total. As J séries temporais resultantes correspondem a J observações da variável $\tilde{\mathbf{Y}}$ em cada intervalo de tempo, a partir das quais é possível calcular percentis preditivos $\tilde{y}_{1-\alpha}$ para cada intervalo temporal. Assim, por exemplo, utilizando-se $1 - \alpha = 0,975$ e $1 - \alpha = 0,025$, obtêm-se os percentis preditivos com 97,5% e 2,5% de não excedência, para cada intervalo temporal do hidrograma/cotagrama previsto, os quais, juntos, comporão o intervalo de credibilidade a um nível de 95%. Caso $\mathbf{E}(\boldsymbol{\eta}_e) = 0$, são obtidos percentis preditivos que incorporam somente a incerteza preditiva associada aos parâmetros do modelo hidrodinâmico, conforme conceituado inicialmente na Equação 3.18, à página 42.

Alternativamente, a expressão 4.6 pode ser utilizada para obtenção de percentis e intervalos preditivos associados ao nível de água máximo obtido nas seções transversais ao longo de um trecho fluvial em determinada cheia, sob a hipótese de que uma mancha de inundação tenha constituído o paradigma $\tilde{Y}$ para calibração das incertezas envolvidas em uma modelagem hidrodinâmica unidimensional.

O cômputo dos erros na expressão 4.6 demanda a geração de amostras independentes da distribuição $SEP(0,1, \xi, \beta)$, o que foi realizado a partir de um algoritmo explicado no estudo de Schoups e Vrugt (2010, página 5) e disponibilizado pelos autores sob solicitação. Uma vez calculadas J séries extraídas da SEP, é necessário multiplicar-lhes a variância σ_t e adicionar-lhes a correlação serial representada por ϕ_p , com as correspondentes variáveis latentes extraídas do conjunto $\{\theta, \eta_\epsilon\}$ de cada sorteio, seguindo-se o disposto na Equação 4.1 (página 99).

A Figura 4.3 demonstra esquematicamente o processo de estimação da incerteza preditiva *a posteriori* obtida para as variáveis $\tilde{Y}$ passíveis de serem observadas em determinada seção de um estirão fluvial, considerando-se todas as fontes de incerteza quantificadas de forma explícita (i.e., paramétricas) ou implícita (i.e., demais fontes). À semelhança da expressão 4.6, a Figura 4.3 pode ser adaptada e interpretada a fim de mostrar o processo de obtenção da incerteza preditiva correspondente às cotas máximas alcançadas por uma cheia ao longo de um estirão fluvial quando manchas de inundação são usadas no processo de calibração de incertezas.

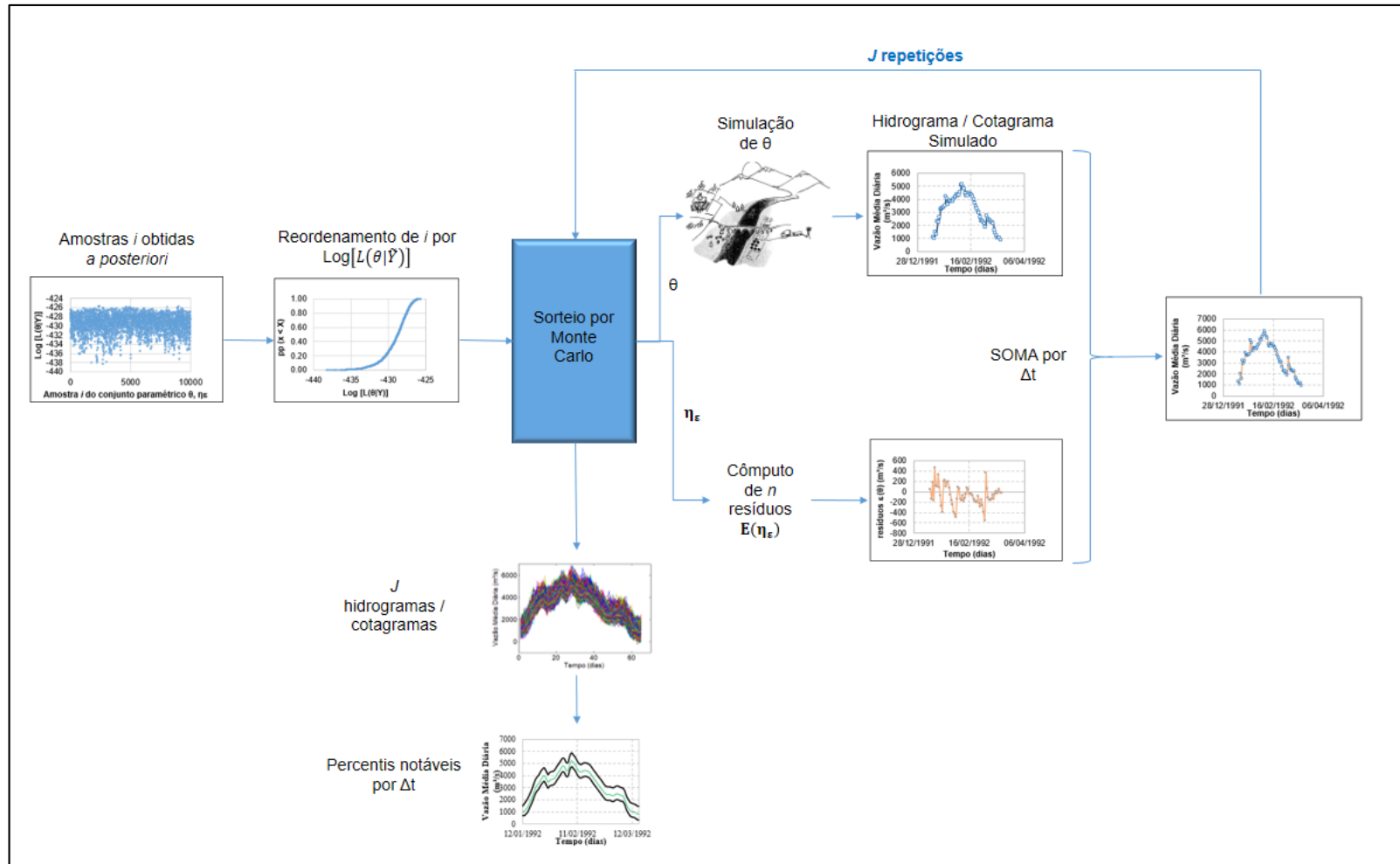


Figura 4.3 – Esquema de estimação da incerteza preditiva *a posteriori* para a informação $\tilde{Y}$ utilizada na calibração de incertezas pelo método Bayesiano adotado, quando são utilizados hidrogramas ou cotogramas em determinada seção de um estirão fluvial.

A Equação 4.6 pode ser aplicada ao mesmo evento de cheia utilizado como paradigma $\tilde{Y}$, seja sob a forma de cotagramas, seja de hidrogramas, de modo que se avalie a incerteza preditiva total contida em sua simulação no modelo hidrodinâmico selecionado e elaborado com as informações topobatimétricas e condições iniciais e de contorno disponíveis. Da mesma forma, também pode ser usada para se estimar a incerteza preditiva de outras cheias, com função de validação, desde que se estipule que a estrutura de erros nos hidrogramas e/ou cotagramas defluentes, ou, ainda, nas manchas de inundação, e a incerteza dos parâmetros e variáveis latentes mantenham-se válidas, o que parece razoável para eventos de magnitude semelhante àquela correspondente aos eventos utilizados na calibração de incertezas pelo esquema proposto. É importante ressaltar que se espera obter comportamentos distintos para os resíduos resultantes da calibração de incertezas da modelagem utilizando-se fontes distintas da informação $\tilde{Y}$, o que provavelmente impacta a obtenção das incertezas paramétrica e preditiva *a posteriori*.

Por fim, cabe lembrar que a avaliação do impacto da escolha de variáveis latentes e da informação $\tilde{Y}$ sobre a quantificação dos intervalos de credibilidade constitui, junto com as análises dos resíduos e das incertezas paramétricas e das variáveis latentes, ferramenta para definição e seleção final de cenários de calibração de incertezas pelo esquema Bayesiano, sob a estrutura metodológica proposta nesta pesquisa.

4.2.6 Cenários de calibração

Com base nos elementos apresentados nos subitens 4.2.1 a 4.2.5, foram predefinidos alguns cenários para testar e validar a metodologia proposta.

Ao longo dos estudos, percebeu-se a necessidade de estipular duas vertentes para estabelecimento de cenários de calibração das incertezas pelo esquema Bayesiano com amostragem pelo método DREAM, quais sejam:

- Quantidade de variáveis latentes necessárias à correta e suficiente prescrição da função de verossimilhança a partir da função Generalizada.

Com efeito, foram prescritos diversos cenários com a inserção/exclusão de uma ou mais variáveis latentes, a fim de conferir a necessidade de um modelo linear para o desvio-padrão dos resíduos ε_t e de um modelo autoregressivo para explicar sua dependência serial, bem como

garantir uma adequada aderência da distribuição SEP aos resíduos após remoção da autocorrelação e da heteros/homoscedasticidade.

Eventualmente, foram definidos cenários extras para ampliar ou reduzir os limites pré-estipulados para as distribuições de probabilidades marginais *a priori* das variáveis latentes, os quais foram apresentados no subitem 4.2.2.2.

- Tipo de variável componente do vetor de dados observados $\tilde{Y}$.

A ideia central ao estabelecer cenários em função do tipo de variável usada, notadamente níveis de água e vazões registrados em uma ou mais seções transversais ao longo dos trechos fluviais estudados, ou ainda, manchas de inundação observadas, foi avaliar o impacto dessa decisão sobre a quantificação das incertezas paramétrica e preditiva *a posteriori* e sobre o comportamento estatístico dos resíduos da modelagem hidrodinâmica.

4.3 Seleção de trechos fluviais para aplicação do método proposto

Ao longo dos dois primeiros anos desta pesquisa, realizou-se uma prospecção de trechos fluviais com informações suficientes para viabilizar a aplicação do método proposto. Avaliou-se a disponibilidade dos seguintes tipos de dados:

- Seções topobatimétricas, preferencialmente oriundas de levantamentos recentes, realizados nos últimos 10 anos;
- Fotos dos locais de levantamento das seções topobatimétricas, ou de caracterização geral do trecho, que auxiliassem na avaliação das condições da calha principal e das planícies de inundação, como material predominante no leito e nas margens, tipo e porte de vegetação e existência de obstruções e de áreas de escoamento inefetivo;
- Dados referentes à granulometria do leito e das margens;
- Cadastro topográfico de estruturas hidráulicas porventura existentes lateral ou transversalmente ao escoamento, levantado de preferência na mesma campanha das seções topobatimétricas;
- Imagens de satélite que contivessem manchas de inundação de cheias importantes;
- Modelo digital de terreno (MDT) com resolução horizontal mínima de 10 x 10 m para permitir o mapeamento das áreas inundadas;

- Registros de cota e de vazão de base diária pelo menos, com no mínimo 10 anos de extensão e que contivessem cheias importantes no trecho fluvial, em estações fluviométricas a montante e a jusante, e, preferencialmente, também em afluentes de maior porte; e
- Cotas altimétricas dos “zeros” (i.e., origens) das réguas de medição nas estações fluviométricas, sobretudo nas situadas a jusante do trecho, e, preferencialmente, obtidas na mesma campanha dos levantamentos topobatimétricos, ou informadas pelo órgão/entidade responsável pela operação dos postos de monitoramento.

Além dessas informações, registros fotográficos e reportagens sobre cheias importantes também foram procurados, em busca de possíveis marcas de cheia. Uma vez realizada a pré-seleção de estirões fluviais com base nos critérios supracitados, buscou-se detectar, de maneira qualitativa, os possíveis controles hidráulicos, por meio de imagens de satélite disponíveis no aplicativo GoogleEarth, a fim de nortear a escolha dos extremos de cada trecho, juntamente com o posicionamento dos postos de monitoramento com disponibilidade de registros de cotas e vazões.

4.3.1 Análise de dados hidrométricos

A análise dos dados hidrométricos disponíveis nos trechos pré-selecionados baseou-se nas seguintes avaliações preliminares:

- Verificação da existência de registros concomitantes de cota e vazão consistidos, em estações fluviométricas a montante e a jusante;
- Cálculo das áreas incrementais entre os postos existentes;
- Avaliação das cheias de maior porte nos postos disponíveis, a fim de utilizá-las como paradigma $\tilde{Y}$ de calibração de incertezas, seja em termos de cotas, seja em termos de vazões;
- Seleção dos máximos anuais de vazão média diária e posterior análise de frequência das respectivas amostras, para proporcionar uma comparação entre os eventos mais importantes;
- Plotagem concomitante dos registros completos de cota e vazão, juntamente com os dados consistidos de medição de descarga líquida, para avaliação da aderência dos pontos restituídos de vazão a partir das curvas-chave, e verificação do ramo de extrapolação e da posição das cheias de maior magnitude; e

- Transformação das ordenadas dos cotogramas dos eventos pré-selecionados nos postos situados a jusante dos trechos em cotas altimétricas, a partir dos valores das cotas altimétricas dos “zeros” (referências de nível consideradas como origem) das respectivas réguas linimétricas.

A existência dos dados mencionados no início deste subitem, aliada à extensão e à qualidade dos dados de monitoramento fluviométrico, bem como ao registro de cheias de maior magnitude, reduziu a quantidade de trechos fluviais para aplicação do esquema metodológico proposto a somente um estirão, situado no rio São Francisco, a jusante da usina hidrelétrica de Três Marias, no estado de Minas Gerais. Embora seja desejável que o trecho estudado não apresente regularização, uma vez que tal aspecto possa introduzir incertezas de difícil identificação, mesmo em se tratando de cheias de magnitude semelhante, nenhum outro estirão pré-selecionado apresentou o mesmo nível de informação que viabilizasse a aplicação da metodologia proposta.

4.3.2 Definição dos coeficientes de rugosidade de Manning

A partir de uma avaliação prévia das características da calha e de suas margens e das planícies de inundação, viabilizada por meio de fotos e de imagens de satélite do trecho finalmente selecionado, foram feitas estimativas acerca da incerteza dos correspondentes coeficientes de rugosidade de Manning, notadamente de limites máximos e mínimos que refletissem a variabilidade desse parâmetro, modelada pela distribuição uniforme. A manutenção desse modelo distributivo em relação às hipóteses inicialmente realizadas quanto à incerteza estipulada *a priori* para o coeficiente de Manning deveu-se à ausência de elementos que permitissem a diferenciação entre as probabilidades para distintos valores desse parâmetro, tal como inspeções de campo detalhadas ao longo do estirão fluvial escolhido e consulta a *experts* para aferir sua variabilidade.

Inicialmente, a partir da prospecção por fotos e imagens, definiu-se o critério, dentre os dois mencionados no subitem 4.2.2.1, para especificação de tipologias e/ou regiões das seções transversais às quais seriam atribuídos distribuições Uniformes com distintos limites.

Em segundo lugar, e sob a hipótese de que dificilmente um extenso trecho fluvial natural apresentaria as condições ideais de escoamento dos trechos cujos coeficientes de rugosidade de Manning foram estimados e reportados na literatura, foram previstos os seguintes critérios para

nortear a definição dos limites de exploração da incerteza associada a esse parâmetro, em cada tipologia e/ou porção do canal estipulados previamente:

- Estabelecimento de um valor de base, n_b , para aplicação do método de Cowan (1956 *apud* ARCEMENT e SCHENEIDER, 1989), válido para cada tipologia e/ou porção do canal, seguida por:
- Aplicação dos incrementos propostos por Cowan (1956 *apud* ARCEMENT e SCHENEIDER, 1989), n_1 a n_4 e m , segundo adaptação de Arcement e Schneider (1989), em seus valores mínimos, médios e máximos, para seleção do limite superior a ser adotado para a distribuição uniforme; e
- Comparação dos valores máximos estimados no excerto anterior com valores encontrados na literatura (CHOW, 1959; ARCEMENT e SCHENEIDER, 1989; HICKS e MASON, 1998; BRUNNER, 2016a), para verificação de sua plausibilidade física, e definição final do limite superior, para cada tipologia e/ou porção do canal.

Conforme se vê, buscou-se retratar a variabilidade intrínseca ao conhecimento limitado que se tem ao se determinar o coeficiente de rugosidade de Manning, mas respeitando-se limites físicos plausíveis, extraídos da bibliografia sobre o tema.

4.4 Simulações hidráulicas iniciais

Previamente ao estudo sistemático de incertezas, procedeu-se a uma série de simulações exclusivamente no ambiente do modelo HEC-RAS, a fim de compreender os controles hidráulicos representados pelas seções topobatimétricas existentes, proceder à montagem do trecho para a estimação das incertezas, aferir a sensibilidade ao coeficiente de rugosidade de Manning, bem como à discretização temporal adequada e às condições de contorno por jusante e por montante. O objetivo principal dessas verificações preliminares foi elaborar o modelo hidrodinâmico unidimensional a ser finalmente considerado como elemento para quantificação de incertezas segundo a proposta metodológica elaborada. Além disso, foram avaliados os elementos que podem acarretar incertezas de difícil quantificação por uma abordagem probabilística.

4.4.1 Condições de contorno

As seções topobatimétricas disponíveis, o trecho fluvial selecionado e as estações fluviométricas com informações com potencial para utilização no estudo foram posicionadas em uma base elaborada em SIG (Sistema de Informações Geográficas), a qual foi em seguida exportada para o HEC-RAS por meio de sua interface HEC-GeoRAS (ACKERMAN, 2011), a menos dos postos de monitoramento.

O modelo hidráulico elaborado preliminarmente, ou seja, contendo as seções transversais e o traçado do perfil do leito fluvial na calha principal, serviu de base para a compreensão do escoamento de ampla gama de vazões, avaliado em regime permanente e para uma faixa de coeficientes de rugosidade de Manning expeditamente definida. Assim, foi possível aferir a fidelidade do modelo em representar os controles hidráulicos detectados pelas imagens de satélite e fotos, a exemplo de corredeiras, aflúências de cursos de água de maior porte e estreitamentos naturais.

Detectados os controles hidráulicos relevantes, e posicionadas as estações fluviométricas cujos dados seriam usados como condições de contorno por montante, foi possível estabelecer os limites físicos para elaboração final do modelo hidrodinâmico.

É importante mencionar que dados de cotas e de vazões em estações fluviométricas existentes a jusante do trecho, embora pudessem constituir relevante informação a ser usada como condição de contorno de jusante (e.g., sob a forma de uma curva-chave ou de um hidrograma), não puderam ser utilizados dessa maneira. Isso se deve ao fato de que serviriam como paradigma $\tilde{Y}$ para calibração de incertezas da modelagem e dos dados, por meio de sua comparação com a resposta de milhares de simulações. Dessa forma, a condição de contorno de jusante mais provável seria a hipótese de escoamento uniforme, situação que demanda uma análise cuidadosa e possivelmente a inserção de novas seções a fim de complementar o trecho em estudo de modo a garantir a ocorrência desse regime.

4.4.2 Discretização espacial e temporal

Na sequência, definido o trecho fluvial a ser de fato modelado, condicionado por montante pela existência de registros de estações fluviométricas, e por jusante pela profundidade normal de escoamento (i.e., regime uniforme), foi possível aferir se a discretização espacial representada

pela distância entre as seções transversais disponíveis estava adequada e, posteriormente, definir o intervalo temporal para proceder às simulações em regime não permanente.

Assim, em primeiro lugar, foram aplicadas as relações 3.64 e 3.66 (páginas 84 e 85, respectivamente), ambas indicadas em Brunner (2016b), para avaliação da distância máxima ideal entre seções transversais consecutivas, a fim de evitar o efeito de difusão numérica pelo modelo (CHOW, 1959). Tais cálculos constituem aproximações, devido à incerteza inerente à definição das variáveis envolvidas em seu cômputo, a exemplo da celeridade da onda de cheia, da profundidade média na calha principal e da declividade média do álveo. Portanto, seus resultados devem ser usados com cautela e servir como indicadores da adequação dos dados topobatimétricos à situação a ser modelada, e de possíveis ajustes na quantidade e no espaçamento entre as seções transversais.

Em segundo lugar, definida a distância máxima por pelo menos uma das duas expressões supracitadas, aplicou-se a condição de *Courant*, considerando-se um valor máximo de 1,0 para a relação entre o intervalo espacial e a celeridade da onda de cheia (Equação 3.67, página 86). O valor obtido foi avaliado como sendo o máximo ideal a ser definido para as simulações de propagação hidrodinâmica das cheias selecionadas.

Em todas as verificações ora relatadas, as variáveis hidráulicas profundidade e velocidade, sendo esta última subsidiária de uma estimativa para a celeridade de uma onda de cheia, foram quantificadas em termos médios ao longo do estirão fluvial finalmente considerado, obtidas em simulações em regime permanente, para vazões que preenchessem completamente a calha principal, considerando-se valores médios para os coeficientes de rugosidade de Manning extraídos da literatura para canais naturais. O tempo de subida da onda de cheia foi aferido a partir dos eventos de cheia das estações fluviométricas de montante pré-selecionados como candidatos à utilização como dados de entrada no esquema de calibração de incertezas.

4.4.3 Aquecimento do modelo

Selecionados os eventos de cheia passíveis de serem utilizados como paradigma de calibração de incertezas, e de posse do trecho fluvial definido pelos procedimentos até aqui relatados, foi possível avaliar quais seriam as vazões iniciais dos hidrogramas afluentes a serem usados como condição de contorno, ou seja, as descargas para aquecimento do modelo e o número de

intervalos temporais ao longo dos quais estas deveriam ocorrer, a fim de evitar possíveis instabilidades numéricas.

4.4.4 Avaliação da estabilidade numérica e dos modelos fisicamente plausíveis

Uma vez que os intervalos espacial mínimo e temporal máximo foram estimados pelos procedimentos anteriores, é interessante que se avalie a resposta do módulo não permanente do HEC-RAS à variação do fator *theta*, nomeado por κ no presente trabalho. A recomendação de Brunner (2016b) é partir de $\kappa = 1,0$ em direção a $\kappa = 0,6$, o que pode melhorar a precisão das saídas fornecidas na modelagem, ao custo de ocorrência de possíveis instabilidades. Dessa forma, procedeu-se a uma série de testes a fim de verificar a sensibilidade dos resultados, notadamente as vazões e profundidades, à mudança nesse parâmetro associado ao esquema numérico de resolução das equações de Saint-Venant.

Outro aspecto avaliado quanto ao possível comprometimento da estabilidade numérica e das saídas da modelagem concerne à existência de pontes e à forma como as mesmas são inseridas e retratadas no modelo hidráulico. Portanto, procedeu-se à análise de sensibilidade de resultados após a inserção desse tipo de estrutura hidráulica no trecho fluvial estudado, bem como à busca de eventuais instabilidades, considerando-se nesse caso toda a amplitude de coeficientes de rugosidade de Manning estipulada *a priori*. Nessa condição, enfatizou-se a avaliação das profundidades nas seções a montante, a jusante e situadas sob a estrutura em questão. Outro ponto relevante avaliado foi a necessidade de inserção de seções transversais adicionais, no caso, interpoladas, para melhor retratar o perfil da linha de água nas proximidades da obra.

Em terceiro lugar, verificou-se a ocorrência de instabilidades numéricas e a sensibilidade das profundidades e velocidades médias nas seções transversais com áreas de escoamento inefetivo, a fim de definir se as mesmas deveriam ser consideradas no modelo hidrodinâmico em elaboração.

Finalmente, foram verificadas a plausibilidade e a sensibilidade das profundidades, vazões e velocidades médias de escoamento nas seções transversais utilizadas em função de diversas combinações entre os valores máximos, médios e mínimos para os coeficientes de rugosidade de Manning na calha principal e nas planícies e suas distintas tipologias a fim de garantir a convergência e a estabilidade do modelo ao longo das milhares de simulações pelo método DREAM.

5 APLICAÇÃO DA METODOLOGIA

Neste capítulo são explicados em detalhes todos os aspectos inerentes à aplicação do ferramental metodológico descrito ao estudo de caso, referente ao trecho fluvial selecionado à luz de uma série de critérios. Ademais, são apresentados os resultados obtidos e as constatações feitas a partir das análises propostas, com as devidas comparações com estudos semelhantes.

5.1 Caracterização do trecho fluvial selecionado e informações disponíveis

O estudo de caso refere-se a um trecho do rio São Francisco, entre a foz do rio Abaeté e a cidade de Pirapora, no estado de Minas Gerais, na amplitude de coordenadas UTM SIRGAS 2000 (zona 23S) 473.600 e 513.700 E e 8.005.150 e 8.087.000 N, perfazendo cerca de 100 km de extensão desse curso de água em sua alta porção, conforme mostra a Figura 5.1.

Esse trecho possui 2 estações fluviométricas a montante, operadas pela CEMIG – Companhia Energética de Minas Gerais, sendo uma no rio São Francisco e a outra no rio Abaeté, a 30 km e a 40 km do encontro de ambos os rios, respectivamente. No extremo de jusante, junto à ponte da BR-365, na área urbana de Pirapora, há outro posto, de operação pela CPRM – Serviço Geológico do Brasil, cujo histórico de cotas e vazões médias diárias foi considerado como a informação de atualização do esquema Bayesiano de estimação de incertezas, para cheias conhecidamente importantes no curso de água em questão. Para esse posto, a CPRM forneceu a cota altimétrica do “zero” do lance de réguas, a fim de permitir a obtenção de cotagramas construídos em termos de cotas altimétricas. A posição das estações fluviométricas em relação à rede hidrográfica de interesse do estudo é apresentada na Figura 5.1, assim como algumas de suas informações encontram-se sintetizadas na Tabela 5.1.

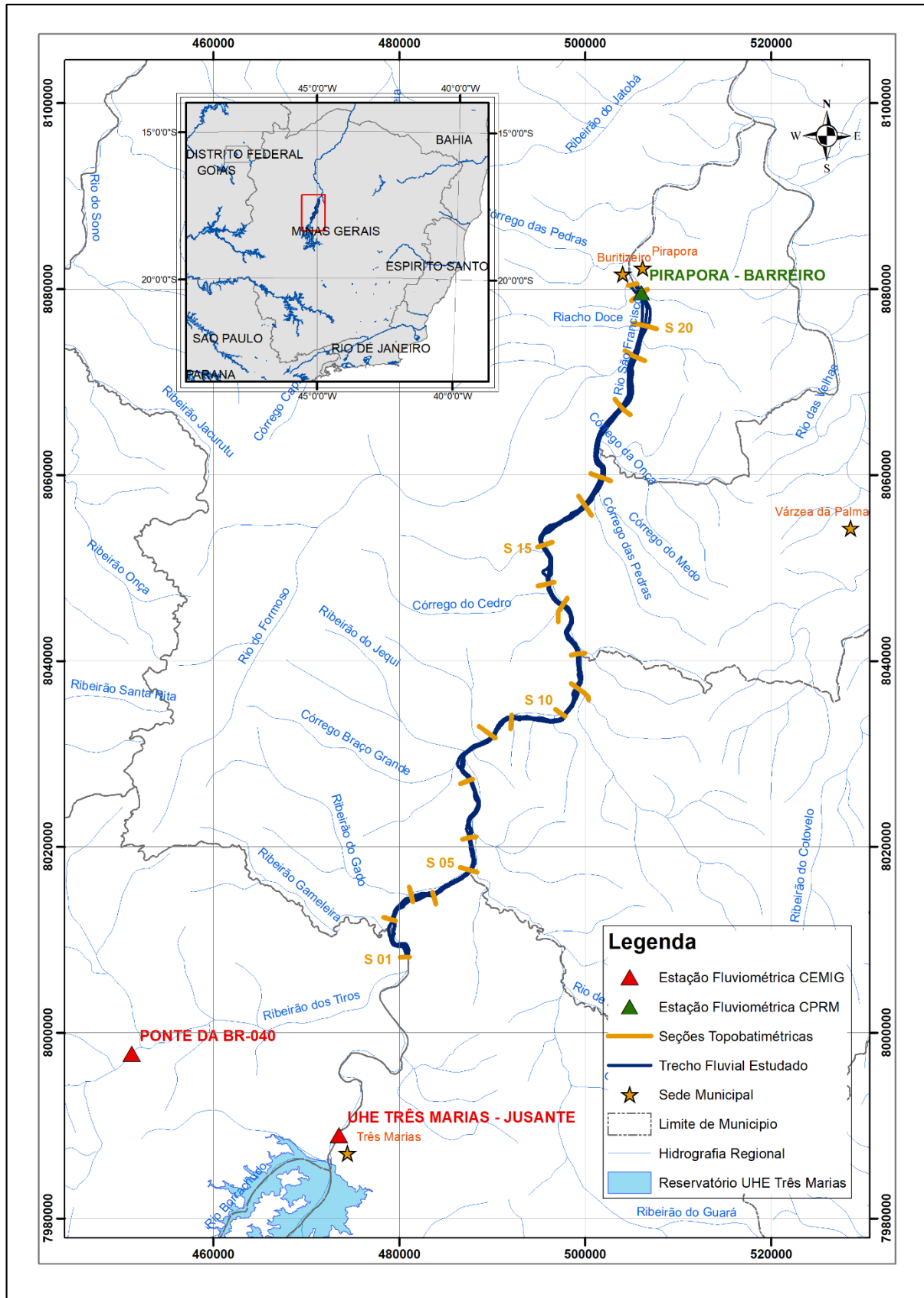


Tabela 5.1 – Características das estações fluviométricas utilizadas.

Código (ANA, 2009)	Nome	Operadora	Curso de Água	Coordenadas Geográficas SIRGAS2000		Período de Dados	Área de Drenagem (km²)
				Latitude	Longitude		
41020002	UHE Três Marias Jusante	CEMIG	rio São Francisco	18°11'13"O	45°15'10"S	01/1957 a 11/2017	50816
41090002	PCH Salto do Paraopeba Ponte BR 040	CEMIG	rio Abaeté	18°06'32"O	45°27'43"S	04/1963 a 11/2017	5186
41135000	Pirapora Barreiro	CPRM	rio São Francisco	17°22'09"O	44°56'36"S	06/1968 a 11/2016	62200

Por se tratar de um estirão de interesse direto da CEMIG, devido à existência da UHE Três Marias por montante, no município homônimo, há levantamentos topobatimétricos realizados em diversas ocasiões, e que perfazem a área desde a usina até a cidade de Pedras de Maria da Cruz, cerca de 280 km a jusante. Para o trecho finalmente selecionado neste estudo, havia disponibilidade de 22 seções transversais, levantadas em campanha de 2011, posicionadas conforme mostra a Figura 5.1. Para as avaliações hidráulicas iniciais realizadas para definição formal do estirão fluvial a ser representado, outras 7 seções transversais oriundas do mesmo levantamento foram necessárias.

Há também disponível um MDT de levantamento realizado pela CEMIG para o leito fluvial e planícies de inundação, ao longo do estirão entre a foz do rio Abaeté e a cidade de Pirapora, com resolução horizontal de 1 x 1 m, que engloba a calha principal e uma faixa com largura entre 1 e 2 km da planície de inundação em cada margem, mais larga próximo e ao longo da referida cidade, e mais estreita nas áreas de ocupação predominantemente rural.

A calha principal nessa porção do rio São Francisco apresenta largura média entre 200 e 500 m, e extensas planícies de inundação, com alguns quilômetros a partir de cada margem. A Figura 5.2 retrata algumas seções transversais típicas dentre as disponíveis e que foram utilizadas nas simulações no HEC-RAS.

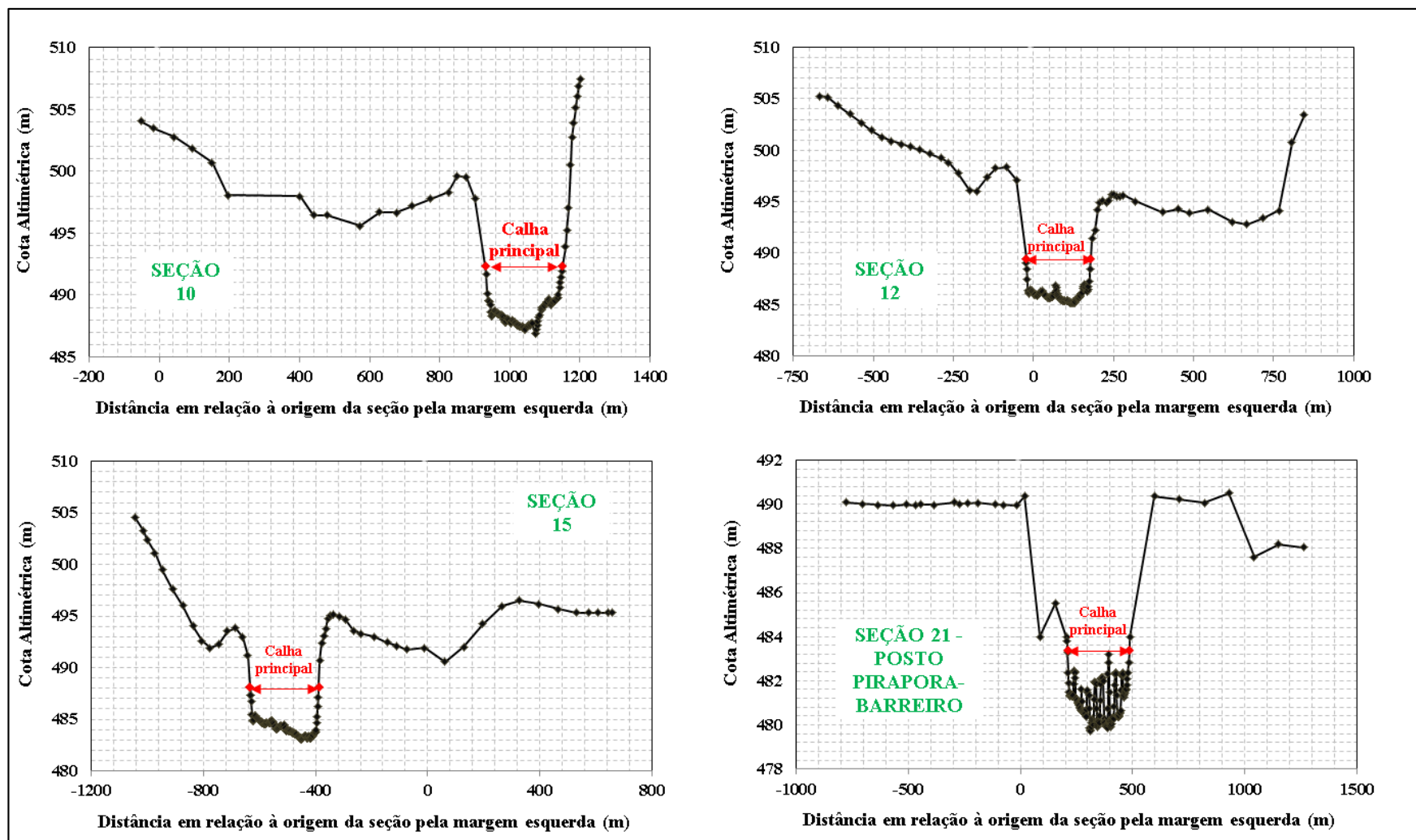


Figura 5.2 – Seções transversais típicas do rio São Francisco no trecho considerado na pesquisa.

A declividade média no trecho fluvial estudado é de 0,00018 m/m, estimada a partir das seções transversais utilizadas na pesquisa. A jusante da porção considerada neste estudo, na região central da cidade de Pirapora, há uma corredeira esculpida em um afloramento rochoso que opera como controle hidráulico transversal à calha fluvial. Sua declividade, de 0,005 m/m, promove a aceleração do escoamento e a ocorrência de escoamento crítico imediatamente a montante, a depender das vazões em trânsito no rio São Francisco e no seu importante afluente a jusante, o rio das Velhas. Esse tributário situa-se cerca de 30 km a jusante da cidade de Pirapora, e seu regime causa remanso até a região da corredeira, para ampla gama de vazões em ambos os cursos de água. O perfil transversal do rio São Francisco entre a foz do rio Abaeté e a do rio das Velhas ilustra essas mudanças de declividade e a entrada dos principais afluentes, condicionantes relevantes na definição do trecho final selecionado para a realização das simulações hidráulicas subsidiárias da avaliação de incertezas.

Na Figura 5.3 estão indicadas também as duas pontes existentes no estirão entre os rios Abaeté e das Velhas, ambas próximas às cidades de Pirapora e Buritizeiro e distantes entre si em cerca de 1500 m. Seu cadastro topográfico foi fornecido pela CEMIG, realizado junto com a campanha de levantamento das seções transversais consideradas. A estrutura de montante faz a travessia da BR-356 sobre o rio São Francisco, perfazendo aproximadamente 450 m de extensão. Cerca de 50 m a jusante estão instaladas as réguas linimétricas e o linígrafo da estação fluviométrica de Pirapora-Barreiro, cujos registros foram utilizados como referência para calibração das incertezas. A ponte de jusante, com 675 m de extensão e posicionada na região da corredeira, constituía uma travessia ferroviária, desativada na década de 1970 e atualmente usada como passagem de pedestres e bicicletas e veículos ciclomotores. As posições de ambas as obras-de-arte especiais e do posto fluviométrico de jusante são mostradas sobre imagem de satélite disponível no software Google Earth, na Figura 5.4.

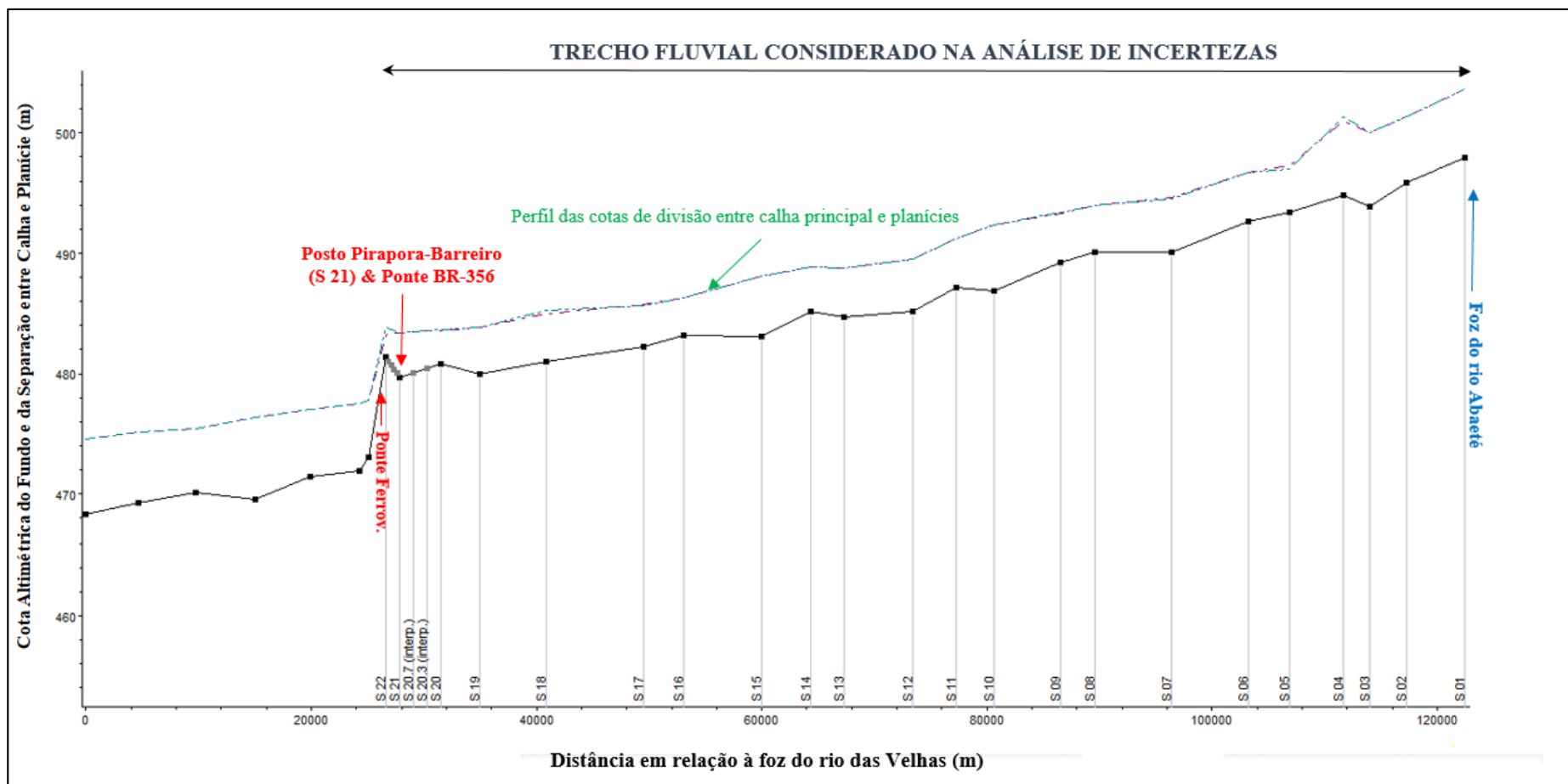


Figura 5.3 – Perfil transversal do rio São Francisco entre a foz do rio Abaeté e a do rio das Velhas, e indicação do trecho utilizado na modelagem hidrodinâmica.

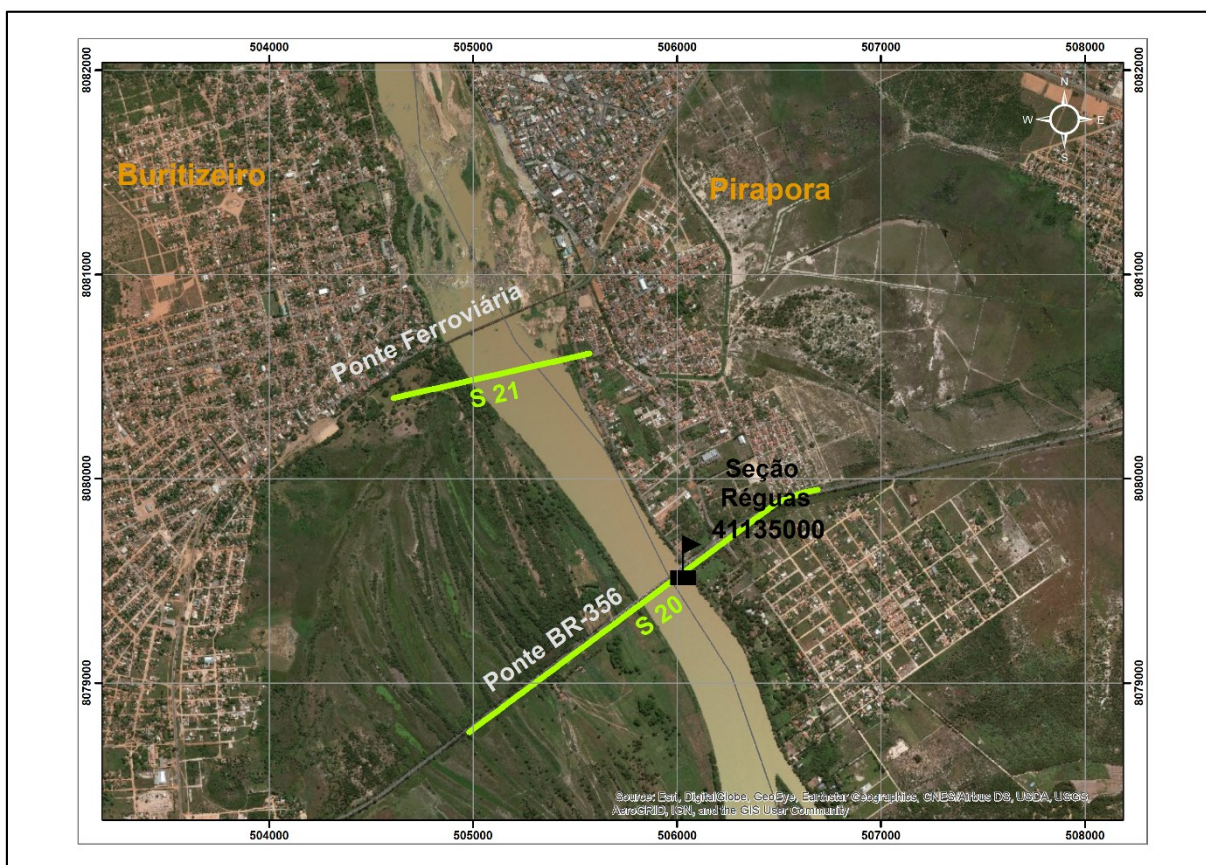


Figura 5.4 – Posição das pontes e da estação fluviométrica situadas no rio São Francisco nas proximidades das cidades de Pirapora e Buritizeiro.

O material constituinte do leito fluvial predominante nessa porção do rio São Francisco são as areias, segundo informações extraídas de estudos conduzidos por UFRGS-IPH (1998) e Roman (2004), a partir de dados hidrossedimentométricos coletados no próprio posto de Pirapora-Barreiro.

As planícies de inundação estendem-se por áreas de ocupação predominantemente rural ou sem intervenção antrópica, com prevalência de pastagens e vegetação arbustiva de pequeno e médio porte. A porção superior das margens caracteriza-se pela profusão de vegetação rasteira e arbustiva, juntamente com árvores de médio porte, formando a típica mata ciliar encontrada ao longo de todo o trecho, em uma faixa estreita em ambas as margens, geralmente com algumas dezenas de metros de largura. Algumas porções das planícies constituem áreas alagadiças, pantanosas, com vegetação principalmente rasteira, que operam como áreas de armazenamento lateral na ocasião de cheias. As fotos selecionadas e inseridas nas Figuras 5.5 e 5.6, obtidas na ocasião dos levantamentos topobatimétricos, mostram o aspecto dominante da vegetação das margens e de parte das planícies.



Figura 5.5 – Fotos que mostram o aspecto da vegetação dominante das margens do rio São Francisco no trecho considerado neste estudo.

Fonte: Tradicional Topografia e Engenharia, 2011.



Figura 5.6 – Fotos que mostram o aspecto da vegetação típicas das planícies de inundação do rio São Francisco no trecho estudado.

Fonte: Tradicional Topografia e Engenharia, 2011.

Os efeitos de regularização proporcionados pelo reservatório da UHE Três Marias atuam ao longo de todo o trecho estudado, detectáveis ao se inspecionar os registros de cotas e vazões nas estações fluviométricas do rio São Francisco a montante e a jusante do estirão. Na Figura 5.7, é possível visualizar o comportamento das vazões médias diárias nos referidos locais, marcado pela sazonalidade intranual típica do regime hidrológico na região Sudeste combinada à regularização visível em parte das cheias e, sobretudo, nas estiagens. De acordo com Pinheiro (2000), o rio São Francisco na porção estudada apresenta transbordamentos generalizados a partir de 3000 m³/s, descarga tomada como referência para seleção de eventos de cheia a serem utilizados como paradigma de calibração de incertezas.

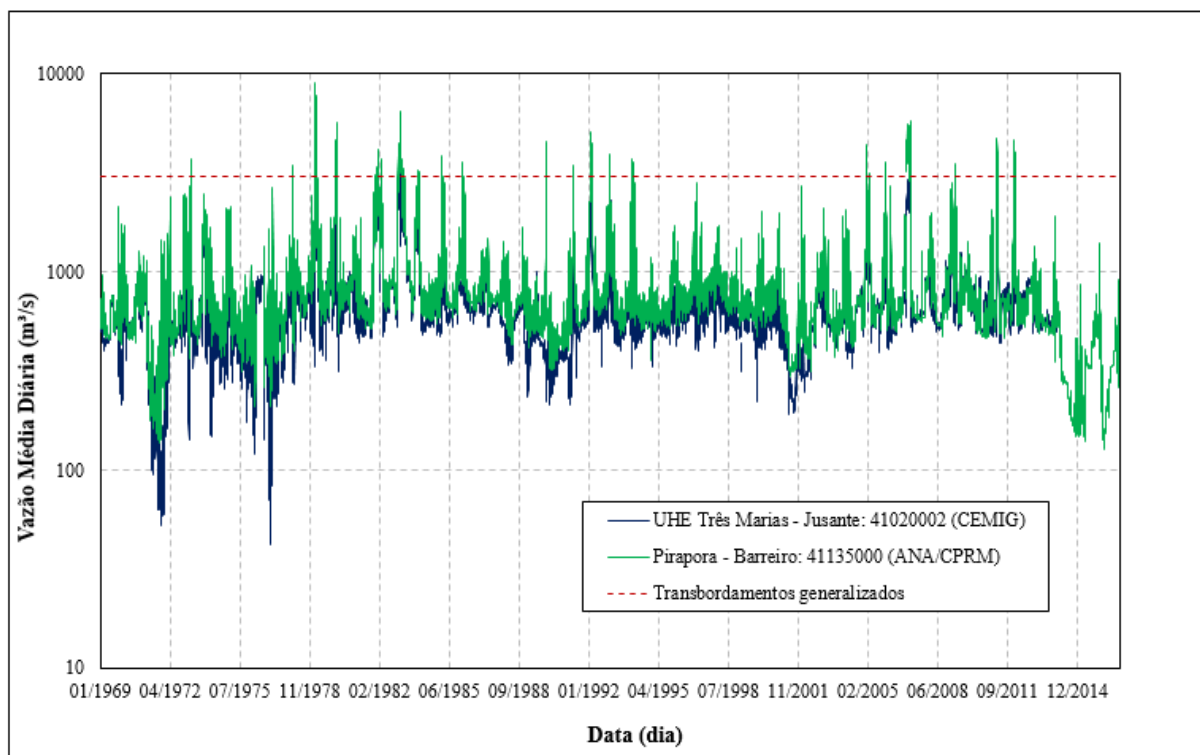


Figura 5.7 – Regime de vazões médias diárias registradas no rio São Francisco a jusante de Três Marias (CEMIG) e em Pirapora-Barreiro (CPRM).

5.2 Simulações hidráulicas iniciais e definição formal do trecho fluvial modelado

De posse dos dados mencionados, foi possível implantar um modelo hidráulico unidimensional preliminar do trecho selecionado, a fim de se verificar o regime de escoamento dominante, a presença e a influência de controles hidráulicos e sua representação pelas seções transversais disponíveis e decidir até qual seção o modelo deveria ser considerado. Essa avaliação deveu-se principalmente à expectativa de o regime ser predominantemente subcrítico, o que demanda a prescrição de uma condição de contorno por jusante. Embora a curva-chave do posto fluviométrico de jusante pudesse constituir uma excelente condição de contorno em outras circunstâncias, ela não foi aplicada nesse caso, já que um dos objetivos trata-se justamente de avaliar os erros entre as vazões e cotas observadas na correspondente seção fluvial e aquelas simuladas pelo modelo. Dessa forma, o trecho fluvial deveria ser representado pelo menos até o local ou trecho que atuasse como controle hidráulico da seção do referido posto, de maneira que o modelo preliminar foi construído até a foz do rio das Velhas para garantir a adequada definição do trecho a ser finalmente modelado.

Verificou-se, em face da informação expressa nas seções topobatimétricas disponíveis, que a corredeira situada na parte de montante do estirão ao longo da cidade de Pirapora opera como controle para vazões de até 5000 m³/s. Acima desse valor, a influência do remanso da confluência entre os rios São Francisco e das Velhas precisa ser levada em conta, e novas incertezas deveriam ser inseridas à inferência, a fim de representar o controle hidráulico e os efeitos impostos pelo importante tributário. Assim, uma vez que o objetivo principal se trata de verificar a aplicabilidade da metodologia com vistas à quantificação das incertezas dos parâmetros e dados de entrada e de saída, bem como da incerteza preditiva, definiu-se que o modelo seria finalizado imediatamente a montante dessa corredeira, considerando-se a propagação de vazões não superiores ao limite indicado.

Foi realizada uma série de simulações em escoamento permanente e não permanente a título de avaliação de sensibilidade às condições de contorno, à faixa de vazões, e aos limites pré-estipulados para os coeficientes de rugosidade. Algumas adaptações foram necessárias ao modelo definido entre a foz do rio Abaeté e a seção de controle hidráulico para o posto de Pirapora-Barreiro a fim de garantir a estabilidade numérica das simulações de propagação de cheias, mencionadas a seguir:

- Criação de seções estilizadas a jusante da seção de controle definida, a fim de garantir a ocorrência do regime crítico na mesma.

Esse trecho virtual foi acrescentado ao modelo, tendo uma declividade bastante acentuada, de 0,008 m/m, para assegurar o controle hidráulico do tipo crítico na seção a montante da corredeira. Essa alternativa mostrou-se adequada em função de duas razões: (i) o módulo hidrodinâmico do HEC-RAS não aceita uma condição de contorno crítica por jusante, como ocorre na realidade, e (ii) a utilização do regime uniforme como condição de contorno por jusante, que também demandaria a extensão virtual do trecho por jusante, acarretou instabilidades e interrupção da modelagem quando coeficientes de Manning muito elevados foram testados na calha ou nas planícies.

A cachoeira artificial inserida no modelo, ao contrário das outras condições testadas, mostrou-se adequada à variabilidade das rugosidades estipulada *a priori*, e garantiu a ocorrência do regime crítico ou próximo a ele na seção de controle, com variações no número de Froude nesse local de até $\pm 15\%$. Verificou-se que essa aproximação introduziu erros da ordem de alguns

centímetros comparativamente à condição de regime permanente com controle crítico por jusante, para vazões até 5000 m³/s.

Ressalva-se que essa região artificialmente criada no modelo para a quantificação de incertezas não é representada nos resultados e análises mostrados ao longo do texto.

- Os trechos dos rios Abaeté e São Francisco até sua confluência não foram considerados na modelagem hidrodinâmica.

Essa escolha foi realizada em função da expectativa de ser possível, em etapa avançada da pesquisa, separar as incertezas inerentes às vazões afluentes em relação às paramétricas e das demais fontes. Dessa forma, a propagação dos hidrogramas até sua junção agregaria complexidade à modelagem das incertezas, pois demandaria a análise de probabilidades conjuntas das vazões médias diárias nos dois cursos de água.

Sendo assim, as vazões médias diárias nas duas estações fluviométricas de montante foram somadas a fim de compor hidrogramas de cheia a serem propagados pelo estirão fluvial considerado. Em função das declividades e morfologia dos trechos a montante, e de sua extensão em comparação ao trecho modelado, considerou-se que a adição das vazões não acarretaria erros elevados.

- As afluições das bacias incrementais ao trecho estudado foram desconsideradas nas primeiras simulações de propagação de cheias.

Essa simplificação foi adotada devido à expectativa de que a estimação de cheias das bacias incrementais, nas quais não há monitoramento fluviométrico sistemático, poderia agregar erros à modelagem de magnitude maior do que a de suas áreas, que representam 11% da contribuição total até o extremo de jusante do trecho.

No entanto, à medida que foram realizados os primeiros cenários de inferência Bayesiana com a metodologia adotada, verificou-se a relevância de se considerarem tais afluições, questão a ser elucidada ao longo deste capítulo.

- As pontes rodoviária e ferroviária existentes no trecho estudado não foram inseridas no modelo hidrodinâmico para inferência das incertezas. O mesmo ocorreu para as áreas de armazenamento lateral.

Essa simplificação foi adotada em favor da estabilidade numérica do modelo, comprometida ao se fixarem valores muito elevados para as rugosidades, sobretudo na planície de inundação, o que comprometeria a realização das milhares de simulações necessárias à inferência de incertezas.

O efeito das pontes e das áreas de armazenamento foi avaliado, em regime permanente e não permanente, e constatou-se que sua presença causava o aumento de até 3 cm nas lâminas de água nas correspondentes seções e a montante.

- Seções interpoladas foram inseridas no extremo de jusante do trecho modelado, a fim de garantir a adequada modelagem do perfil de escoamento na região a montante e a jusante da seção do posto fluviométrico de Pirapora-Barreiro e a montante de seu controle hidráulico.

Adicionalmente, por meio das Equações 3.64 a 3.67 (páginas 84 a 86), verificou-se que o espaçamento entre as seções transversais disponíveis e as interpoladas era suficiente para proceder à modelagem hidrodinâmica das cheias pré-selecionadas. Finalmente, definiu-se que o intervalo de 3,0 horas seria suficiente para o cômputo das derivadas temporais do esquema numérico de resolução das equações de Saint-Venant no presente caso. As últimas aferições realizadas antes de se proceder à inferência das incertezas mostraram que a alteração do parâmetro de ponderação *theta*, ou κ , no intervalo $0,6 \leq \kappa \leq 1,0$ não impactou os valores fornecidos de vazão e nível de água no intervalo diário, usado na comparação de resultados e na obtenção dos resíduos da modelagem.

5.3 Quantificação das incertezas na modelagem hidrodinâmica destinada ao estudo de cheias no trecho fluvial avaliado

5.3.1 Distribuições de probabilidades *a priori* dos parâmetros do modelo hidrodinâmico

Em função das seções transversais disponíveis, e das correspondentes fotos e das imagens de satélite analisadas, foi possível separar o canal fluvial em duas porções ao longo do estirão estudado, seguindo-se a abordagem (ii) estipulada no subitem 4.2.2.1. Assim, ficaram definidos dois parâmetros do modelo hidrodinâmico cujas incertezas seriam inferidas, denotadas por $\theta = \{n_c, n_p\}$. Constatou-se que uma diferenciação mais detalhada, com definição de tipologias de vegetação e de uso e ocupação do solo, poderia ser viável mediante a inspeção em campo das seções transversais disponíveis e de seu entorno.

A partir dos critérios delimitados no subitem 4.3.2, foram estimados os limites inferior e superior da distribuição uniforme *a priori*, para n_c e n_p no trecho fluvial, conforme segue:

- Valor mínimo para a distribuição de n_c estimado como sendo o valor de base, n_b , indicado para calhas fluviais em leito móvel, ou seja, com prevalência de areias, supondo-se diâmetro mediano igual a 0,2 mm, segundo disposto na compilação de Arcement e Schneider (1989, p. 4);
- Valor mínimo para a distribuição de n_p estimado como sendo o valor de base, n_b apontado para planícies de inundação em solo firme, por ser um valor um pouco menor do que o correspondente às areias nesse caso; e
- Valores máximos de ambos os parâmetros: definidos a partir dos máximos sugeridos para calhas e planícies do manual de referências hidráulicas do HEC-RAS (BRUNNER, 2016a). Optou-se por essa referência já que a adoção do método de Cowan (1956 *apud* ARCEMENT e SCHNEIDER, 1989) aplicando-se os fatores aditivos e multiplicativos médios e máximos propostos por Arcement e Schneider (1989, páginas 7 e 9) mostraram-se irrealistas.

A Tabela 5.2 a seguir resume as distribuições *a priori* e os respectivos limites considerados para os parâmetros do modelo hidráulico, $\theta = \{n_c, n_p\}$. As características das distribuições de probabilidades marginais dos parâmetros do modelo de resíduos, $\eta_\varepsilon = \{\sigma_0, \sigma_1, \beta, \xi, \phi_1\}$, foram indicadas na Tabela 4.1.

Tabela 5.2 – Distribuições *a priori* e limites inferior e superior atribuídos aos parâmetros do vetor θ .

Nome	Notação	Unidade	PDF <i>a priori</i>	Limite Inferior	Limite Superior
Coefficiente de rugosidade de Manning na calha principal	n_c	–	UNIFORME	0,012	0,150
Coefficiente de rugosidade de Manning nas planícies de inundação	n_p	–		0,020	0,200

5.3.2 Seleção de eventos de cheia para estudo de incertezas e composição de mapas probabilísticos

Por meio da análise dos registros de cotas e vazões disponíveis, verificou-se que não há muitos eventos de cheia com vazões maiores do que 3000 m³/s resultantes da soma das descargas nos dois postos de montante, o que pode estar relacionado à regularização proporcionada pelo reservatório da UHE Três Marias, destinado também ao controle de cheias no rio São Francisco

a jusante da usina e em parte de sua média porção. Além disso, foram priorizados hidrogramas de cheia mais extensos, que proporcionassem maior comprimento para o vetor $\tilde{Y}$, e afluições que não ultrapassassem o limite superior de vazões condicionado pelo controle crítico por jusante, igual a 5000 m³/s.

Devido às razões mencionadas, foi selecionado o evento de cheia de fevereiro de 1992, uma das maiores registradas no alto rio São Francisco, para estimação de incertezas pelo método Bayesiano definido. A vazão de pico alcançada no posto de jusante, em Pirapora, foi de 5098 m³/s. A vazão de pico do hidrograma afluyente, estimado pela soma das vazões dos postos de montante, por sua vez, foi de 4577 m³/s. O período de dados considerado vai de 12/01/1992 a 16/03/1992, assim escolhido para assegurar estabilidade do modelo hidráulico e perfazer a subida e a descida completa do hidrograma de cheia. Tais aspectos são retratados na Figura 5.8, na qual também é mostrado o cotograma estimado no posto de jusante. Na Figura 5.9, verifica-se que o período de retorno da cheia em questão, estimado pela análise de frequência de máximos anuais de vazão média diária em Pirapora-Barreiro, é de cerca de 10 anos, corroborando a possibilidade de ocorrência de transbordamentos relevantes ao longo do estirão estudado.

As afluições das bacias não monitoradas foram estimadas para realização de alguns cenários de calibração de incertezas pelo esquema Bayesiano, a fim de se avaliar a sensibilidade da metodologia à sua inserção no modelo hidrodinâmico. Sua estimação foi realizada por meio dos seguintes procedimentos de regionalização:

- Definição de três áreas incrementais, denotadas por: (i) bacia do rio Formoso, (ii) bacia do rio de Janeiro, sendo estas duas correspondentes a afluentes ao trecho com maior relevância em comparação às demais áreas, e (iii) demais contribuições. As áreas de drenagem contabilizadas são, respectivamente, de 889 km², 1196 km² e 4246 km²;
- Seleção de um posto fluviométrico de referência para adimensionalização das vazões médias diárias pela vazão média de longo termo (Q_{MLT}). Para tanto, foi escolhido o posto do rio Abaeté em PCH Salto do Paraopeba - Ponte da BR-040 (41090002, indicado na Tabela 5.1), devido às semelhanças morfológicas, climáticas e de regime pluviométrico anual entre sua bacia de contribuição e as áreas não monitoradas;

- Estimativa da Q_{MLT} nas bacias sem dados, aplicando-se o estudo de regionalização de vazões médias nas bacias do Alto rio São Francisco conduzido por CPRM (2001); e
- Multiplicação das correspondentes vazões médias de longo termo estimadas pela série de vazões médias diárias adimensionalizadas pela Q_{MLT} do posto fluviométrico de referência.

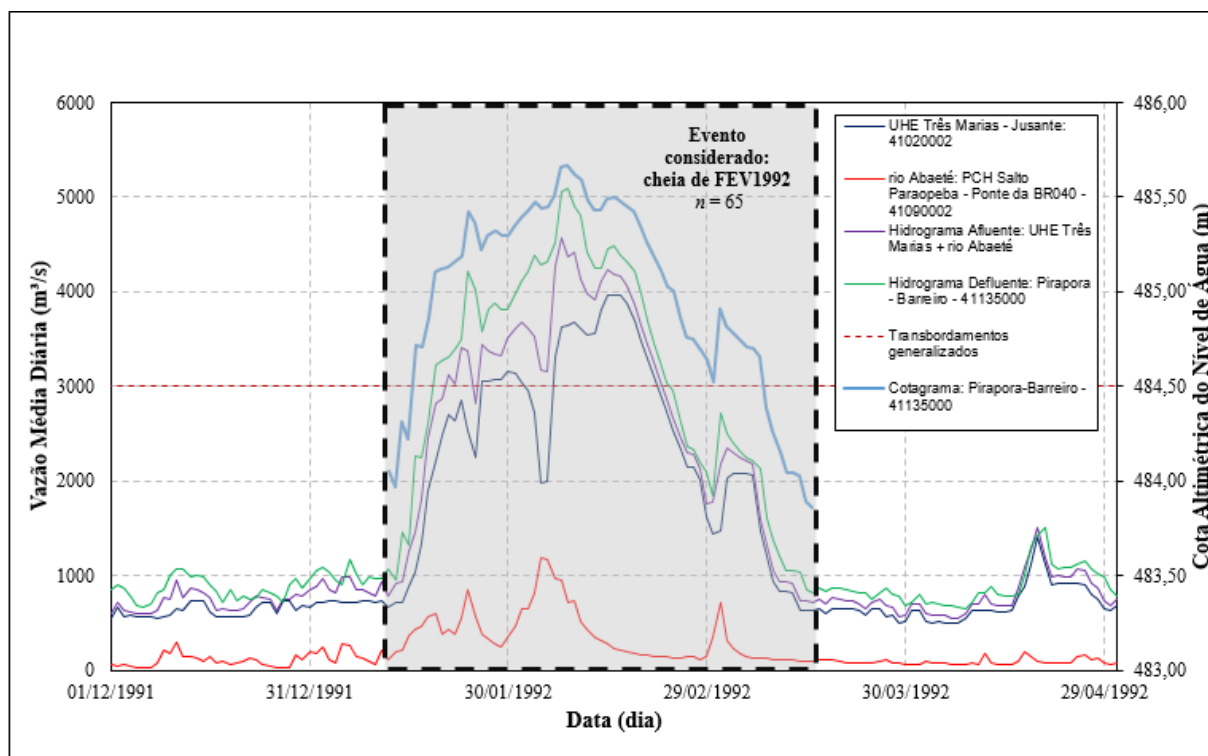


Figura 5.8 – Hidrograma afluente e defluente ao trecho fluvial estudado, referentes à cheia de fevereiro de 1992 (esq.) e seu cotagrama estimado no extremo de jusante (dir.).

Outras técnicas de regionalização, destinadas mais especificamente às vazões de cheia, foram testadas. No entanto, não se obtiveram valores verossímeis para as estimativas realizadas, de forma que se priorizou a aplicação da técnica baseada em vazões médias, mantendo-se as características do regime hidrológico expresso pelo posto selecionado como referencial para transferência da informação.

Os correspondentes hidrogramas da cheia de referência nas três áreas de contribuição mencionadas foram inseridos no modelo hidrodinâmico como afluências laterais, em alguns cenários de inferência Bayesiana, sendo adicionadas e propagadas junto com as vazões em trânsito pelo estirão principal a partir das seções de sua inserção. No caso das demais contribuições, caso (iii) ora tratado, suas afluências foram distribuídas uniformemente pelo trecho modelado.

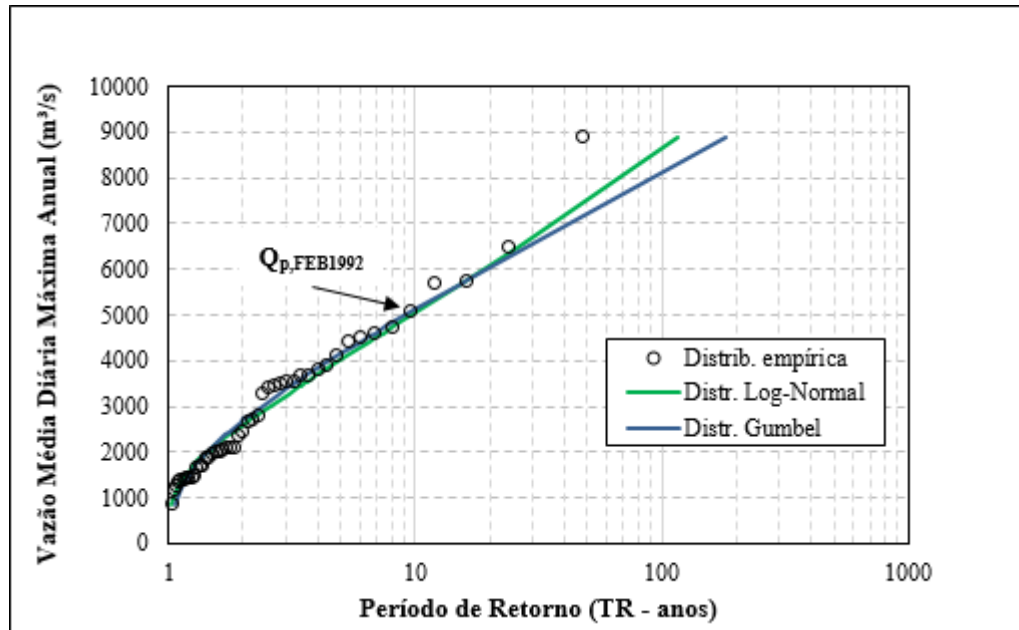


Figura 5.9 – Análise de frequência de máximos anuais de vazão média diária no posto fluviométrico de jusante, no rio São Francisco em Pirapora-Barreiro.

Finalmente, é importante mencionar que não foram encontradas imagens de satélite que retratassem manchas de inundação associadas às cheias de maior magnitude detectadas nos registros históricos das estações fluviométricas utilizadas, o que inviabilizou o uso de cotas máximas indexadas espacialmente como informação $\tilde{Y}$ para calibração de incertezas.

5.3.3 Acoplagem do HEC-RAS ao DREAM

O modelo hidrodinâmico elaborado à luz das informações e premissas mencionadas foi acoplado à *toolbox* de rotinas do método DREAM. Em sua sub-rotina principal, é necessário prescrever o número de parâmetros a ser inferido pelo esquema Bayesiano, a função de verossimilhança a se considerar e as distribuições *a priori* para cada um dos parâmetros do modelo, bem como das variáveis latentes consideradas. Além disso, foi elaborada uma rotina em Matlab que acionava o aplicativo HEC-RAS a cada iteração das cadeias de Markov a fim de acessar o modelo construído e alterar os parâmetros em avaliação, no caso, os coeficientes de rugosidade de Manning na calha principal e nas planícies, n_c e n_p .

O *script* principal do DREAM aplicado com a função de verossimilhança Generalizada é retratado na Figura 5.10. Por sua vez, a rotina elaborada em linguagem do Matlab para acessar e alterar o modelo hidrodinâmico montado no HEC-RAS a cada iteração T das N cadeias de Markov é apresentada na Figura 5.11.

<pre>%% 1. DREAM Core function parameters set-up: % Generalized Log-Likelihood applied to a hydrodynamic modelling case in 1D % environment using HEC-RAS:</pre>	
<pre>%% 1. DREAM Core function parameters set-up: DREAMPar.d = 7; % dimensions number (from the hydrodynamic model + latent variables) DREAMPar.delta = 3; DREAMPar.N = 8; DREAMPar.T = 1000; % number of evaluations in each chain DREAMPar.lik = 14; % global LV % GL Likelihood set-up</pre>	
<pre>%% 2. Parameters information: Par_info.initial = 'prior'; % index: 1 2 d+1 d+2 d+3 d+4 d+5 d+6 d+7 d+8 d+9 d+10 d+11 % name: nc np sig0 sig1 beta xi mul phil phi2 phi3 phi4 k lambda LV.fpar = [nan nan 0.1 0 0 1 0 0 0 0 0 0 1]; % Default values % for unestimated parameters set by the user parmin = [0.012 0.020 0 0 -1 0.1 0.000001 0 0 0 0 0 0.1]; % Lower Bounds parmax = [0.150 0.200 1000 1 10 10 0.0001 1 1 1 1 1 1]; % Upper bounds Par_info.boundhandling = 'fold'; % can try 'reflect', as alternative LV.idx_vpar = [1 2 3 4 5 6 8]; % Index parameters and latent variables Par_info.min = parmin(LV.idx_vpar); %Addressing lower bounds to parameters % and to latent variables for inference Par_info.max = parmax(LV.idx_vpar); % The same, but upper bounds Par_info.prior = { 'unifpdf(0.012,0.150)', 'unifpdf(0.020,0.200)', 'unifpdf(0,1000)', 'unifpdf(0,1)', ... 'unifpdf(-1,10)', 'unifpdf(0.1,10)', 'unifpdf(0,1)'};</pre>	
<pre>%% Bringing up the HEC-RAS Model to the Matlab scene! Func_name = 'HEC_Code';</pre>	
<pre>%% Load calibration data calibration_data = load('Q_41135000_24h.txt'); Meas_info.Y = calibration_data(1:end,1);</pre>	
<pre>%% Optional settings: options.modout = 'yes'; options.save = 'yes'; options.restart = 'yes'; options.diagnostics = 'yes'; % To do the in-chain convergence diagnostics options.parallel = 'no';</pre>	
<pre>%% Calling the main DREAM function, finally! [chain,output,fx] = DREAM(Func_name,DREAMPar,Par_info,Meas_info,options);</pre>	

Figura 5.10 – Código em Matlab da *toolbox* do método DREAM para aproximação numérica da distribuição conjunta *a posteriori* considerando a função de verossimilhança Generalizada.

```

% Function HEC_CODE: constitutes the model Y that generates response
% variables to serve as an input for the DREAM Code.

function Y = HEC_Code(x)

nc = x(1);
np = x(2);
global contador

%display('Roda HEC-RAS')
display(contador);

River_ID = 'SaoFrancisco';
Reach_ID = 'AbaeteVelhas';
ras_file_base = 'C:\hecras\TESE\tese.prj'
ras_file_dir = 'C:\hecras\TESE';

    % Calls the RASCONTROLLER
    h=actxserver('RAS505.HECRASCONTROLLER');
    % Opens the ras_project(j)
    h.Project_Open(ras_file_base);
    % Change the Manning's n using the 'nc' and 'np' parameters sorted in
    % the i-th DreamPar.T iteration
    h.Geometry_SetMann_LChR(River_ID, Reach_ID,Node_ID,np, nc, np, 'Error !!');
    % Shows the Current Plan
    h.CurrentPlanFile;
    % Computes the Current Plan hiding the window
    h.Compute_HideComputationWindow
    h.Compute_CurrentPlan(0,0);
    % Extracts results from the downstream end of the reach (z7=WS,
    % z8=outflow, z9=error message)

    h.Project_Save;
    %Get an array of HEC-RAS output variables and descriptions
    [z1,z2,z3,z4,z5,z6,z7,z8,z9] = ...
h.OutputDSS_GetStageFlow(River_ID,Reach_ID,'28921.58',0,0,0,0,z9);
    h.Project_Save; %Saves the project
    h.Project_Close;
    h.QuitRas;
    delete(h); % Deletes the i-th folder (due to space disk constraints)
    contador = contador + 1;
    Y = z8; % Assigns the z8 (outflow) from the txt as output for the function

```

Figura 5.11 – Código elaborado em Matlab para uso do RAS Controller a fim de alterar os parâmetros n_c e n_p a cada iteração das cadeias de Markov.

5.3.4 Cenários de calibração das incertezas

Conforme mencionado no subitem 4.2.6, à medida que os estudos foram realizados, verificou-se a necessidade de estabelecer uma série de cenários para aplicação do esquema proposto de calibração das incertezas associadas aos parâmetros do modelo hidrodinâmico e às variáveis latentes, representativas das demais fontes de incertezas inerentes ao processo de representação de sistemas ambientais por modelos matemáticos. Tais cenários foram estruturados de maneira lógica a fim de sistematizar as análises vislumbradas como necessárias para extrair conclusões

robustas acerca da aplicabilidade do método ao trecho fluvial selecionado e, eventualmente, a cursos de água de porte e morfologia semelhantes.

A organização utilizada na composição dos cenários de calibração considerou os seguintes aspectos:

- Definição da ordem do modelo autoregressivo $AR(p)$ pré-definido para modelar a dependência serial dos resíduos $\mathbf{E}(\boldsymbol{\theta}) = \tilde{\mathbf{Y}} - \mathbf{Y}(\boldsymbol{\theta}) = \{\varepsilon_1(\boldsymbol{\theta}) \cdots, \varepsilon_n(\boldsymbol{\theta})\} = \varepsilon_t(\boldsymbol{\theta})$. Os primeiros cenários simulados destinaram-se a aferir se um modelo de 1ª ordem seria suficiente, ou se modelos de maior ordem seriam necessários para explicar adequadamente a esperada correlação serial entre os resíduos obtidos *a posteriori*;
- Avaliação da relevância em se considerar o parâmetro σ_1 , destinado a explicar a suposta heteroscedasticidade pré-estipulada para os resíduos ε_t . Isso significa que alguns cenários foram definidos com a presença ou a ausência dessa variável latente, verificando-se seu impacto sobre a inferência das incertezas dos demais parâmetros, bem como sobre a incerteza preditiva e o comportamento dos supracitados resíduos;
- Avaliação da necessidade de fixação de determinados valores para os coeficientes ϕ_p . Esse aspecto adquiriu relevância à medida que se verificou que, em alguns cenários já simulados, os valores inferidos *a posteriori* para tais coeficientes acarretaram intervalos de credibilidade muito amplos para a incerteza preditiva;
- Impacto do tipo de informação $\tilde{\mathbf{Y}}$ usado como paradigma para calibração das incertezas, notadamente o cotograma e o hidrograma da cheia selecionada registrados na estação fluviométrica de jusante, nas proximidades da cidade de Pirapora; e
- Consideração ou não das afluições laterais, ou seja, das contribuições das bacias incrementais ao trecho fluvial considerado. Tal aspecto emergiu durante a realização dos primeiros cenários de simulação, tendo sido constatada a importância de os hidrogramas de vazões médias diárias estimados conforme exposto no subitem 5.2 serem incorporados às simulações hidráulicas.

Os cenários traçados a partir dos supracitados aspectos foram organizados conforme mostra o diagrama da Figura 5.12.

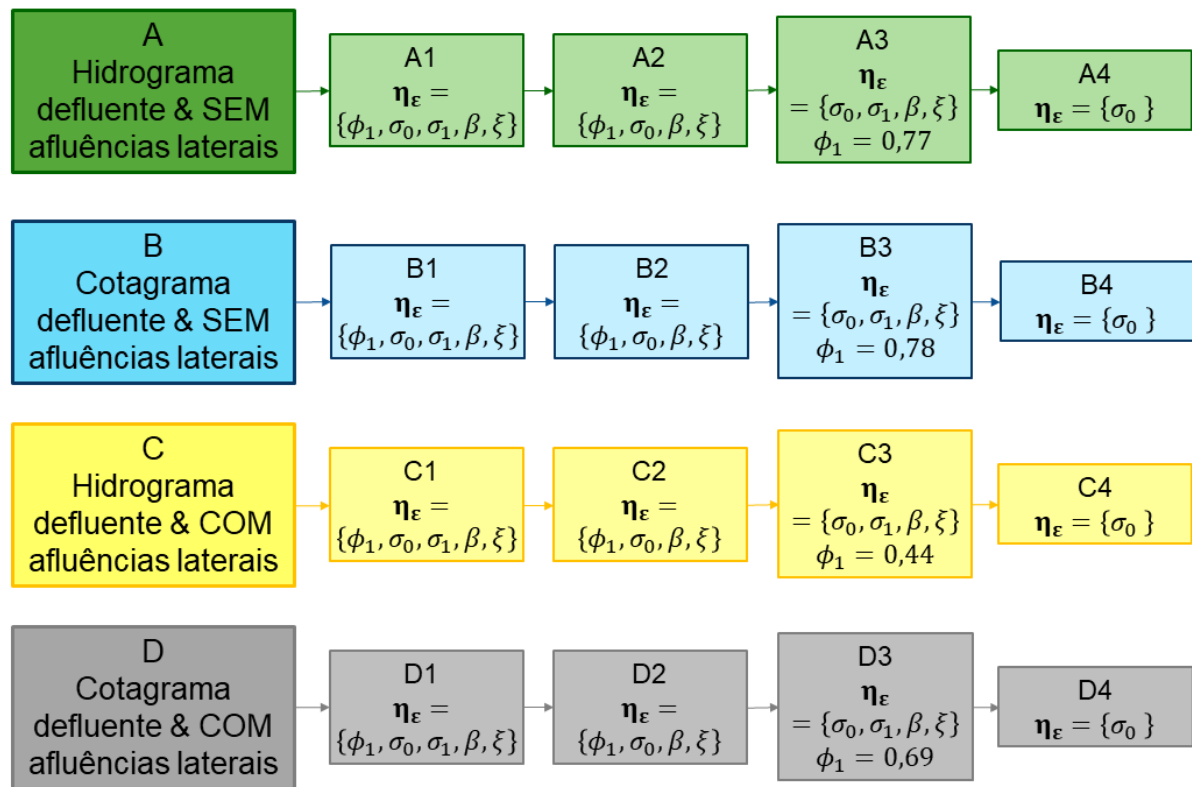


Figura 5.12 – Organização dos principais cenários de inferência Bayesiana de incertezas considerados na pesquisa.

A nomenclatura doravante utilizada no presente texto para referência aos cenários corresponde a uma letra, de **A** a **D**, seguida por um número, de **1** a **4**, conforme indicado no topo de cada caixa indicada no esquema da Figura 5.12.

Ademais, verifica-se que, na 1ª coluna à esquerda da referida figura, o eixo central de definição dos cenários em tipos, de **A** a **D**, organizou-se em função do tipo de informação $\tilde{Y}$ usada como paradigma para atualização do conhecimento prévio acerca da incerteza de variáveis e parâmetros, bem como da hipótese de contabilização ou não das contribuições laterais de bacias incrementais ao rio São Francisco no trecho considerado.

A 2ª coluna, da esquerda para a direita, concerne aos cenários nos quais previu-se a necessidade de prescrição de todas as variáveis latentes explicativas do modelo de resíduos inicialmente pressuposto. Ressalva-se que, no diagrama mostrado, não foram colocados cenários de avaliação da ordem p dos modelos $AR(p)$, visto que tais verificações foram feitas previamente à organização sistemática dos cenários conforme explicitado na figura. Ao longo dessas análises iniciais, constatou-se que a remoção da correlação serial por um modelo autoregressivo de 1ª

ordem seria suficiente no presente caso, daí a indicação do parâmetro ϕ_1 a partir da 2ª coluna do diagrama esquemático.

Adicionalmente, pela 3ª coluna de blocos vê-se que uma análise sistematicamente realizada em todos os tipos de cenários consistiu na avaliação da hipótese de homoscedasticidade para os resíduos. Na 4ª coluna, por sua vez, constam os cenários para cada tipo nos quais foi fixado um valor para ϕ_1 , por razões a serem percorridas no subitem seguinte. Os valores indicados nos blocos referem-se às médias obtidas para os coeficientes de correlação serial de 1ª ordem estimados para cada uma das séries de resíduos ε_t calculadas a partir das diferenças $\mathbf{E}(\boldsymbol{\theta}) = \tilde{\mathbf{Y}} - \mathbf{Y}(\boldsymbol{\theta})$, sendo $\boldsymbol{\theta}$ inferido *a posteriori* no cenário imediatamente anterior, considerando-se os tipos A a D.

Finalmente, na 5ª coluna estão os cenários de menor complexidade, em que somente σ_0 foi considerado como variável latente, sob as hipóteses de que os resíduos ε_t seriam isentos de dependência serial, homoscedásticos e Gaussianos, as quais são recorrentemente utilizadas em estudos de quantificação de incertezas, sob a estrutura de métodos Bayesianos, seja formais ou informais. Dessa forma, ainda que tais premissas não fossem verificadas nas séries de resíduos decorrentes da modelagem, julgou-se forçosa a verificação do efeito que as mesmas causariam sobre a forma e os limites das distribuições marginais *a posteriori* dos parâmetros n_c e n_p , e sobre a composição da incerteza preditiva.

5.3.5 Aplicação do esquema Bayesiano à calibração de incertezas

Delineados cada um dos componentes necessários à aplicação da metodologia proposta adaptados ao estudo de caso, e definidos os cenários de calibração das incertezas paramétricas e das demais fontes envolvidas na modelagem hidrodinâmica, é possível apresentar o panorama de resultados obtidos ao longo dos estudos.

No esquema numérico do tipo MCMC seguindo-se o método DREAM (VRUGT, 2009a), utilizaram-se sistematicamente $N = 8$ cadeias de Markov, cada uma com o mesmo número de iterações. O valor de N foi estipulado de maneira a atender ao preceito de que $N \geq 2\delta + 1$, sendo δ mantido como 3 como *default* do método (VRUGT, 2016). O número de iterações T para cada uma das cadeias variou de acordo com o período demandado individualmente para aquecimento em cada cenário, o que dependeu do número de dimensões d do problema, das condições de contorno adotadas no modelo, e do tipo de informação $\tilde{\mathbf{Y}}$ usado como paradigma.

As rotinas que compõem o DREAM no Matlab permitem gravar resultados parciais e analisá-los, de modo que as simulações por cenário fossem feitas por etapas, verificando-se inicialmente a quantidade de iterações T^* necessárias para que o último parâmetro atingisse a convergência pelo critério de Gelman-Rubin (1992). Constatada a necessidade de mais iterações, também avaliada em função do valor dos coeficientes de correlação serial de cada parâmetro/variável em cada cadeia para os $lags = 1, 5, 10$ e 25 , partia-se para o sequenciamento da mesma simulação a partir da iteração (ou estado das cadeias) imediatamente anterior.

De maneira geral, buscou-se utilizar o $lag\ k \geq 5$, condição que garantiu, em praticamente todos os cenários, que o coeficiente de correlação serial de todos os parâmetros e variáveis fosse igual ou menor do que $0,80$. Outro aspecto interveniente na definição da quantidade total de iterações em cada cenário foi a fixação do tamanho final das amostras obtidas *a posteriori*, ou seja, $m = [N \cdot (T - T^*)]/lag$, igual a 10.000 , seguindo-se as recomendações no trabalho de Vrugt (2016) para d dimensões do vetor $\{\theta, \eta_\epsilon\}$ entre 5 e 10 .

Nos próximos parágrafos, os resultados foram organizados de acordo com os tipos de cenários de simulação, de **A** até **D**.

5.3.5.1 Cenários tipo A: hidrograma de jusante como paradigma de calibração e ausência de afluições laterais como condição de contorno

A premissa básica que permeou as primeiras simulações em todos os tipos de cenários foi a de que o vetor de variáveis latentes seria composto por $\eta_\epsilon = \{\phi_p, \sigma_0, \sigma_1, \xi, \beta\}$, com p igual a 1 nas primeiras tentativas e posterior verificação da necessidade de modelos autoregressivos de maior ordem. Sendo assim, o Cenário A1 foi composto por $d = 7$ dimensões, considerando-se as 5 variáveis latentes e o vetor de parâmetros do modelo hidrodinâmico com 2 dimensões, caracterizadas pelos coeficientes de rugosidade de Manning na calha e nas planícies, ou seja, $\theta = \{n_c, n_p\}$.

Constatou-se que as 10.000 séries de resíduos da modelagem, $E(\theta) = \tilde{Y} - Y(\theta)$, estimadas a partir das diferenças em cada intervalo de tempo entre as vazões observadas em Pirapora-Barreiro e as correspondentes vazões simuladas pelo HEC-RAS, para cada realização i do vetor paramétrico $\theta = \{n_c, n_p\}$ obtido *a posteriori*, apresentaram média e desvio-padrão elevados, de $344\text{ m}^3/\text{s}$ e $298\text{ m}^3/\text{s}$, respectivamente, ambos avaliados em termos médios sobre todas as séries de resíduos obtidas. Tal condição reflete a ausência das afluições laterais na propagação da

cheia de fevereiro de 1992, e, de certa forma, mostra que há coerência física entre os dados de vazões médias diárias desse evento registrados nas estações fluviométricas de montante e de jusante, sem a presença de incrementais negativos. Esse aspecto é retratado na Figura 5.13, que mostra a prevalência de resíduos ε_t positivos em praticamente todos os intervalos temporais e toda a faixa de vazões observadas no posto de jusante. Em adição, as séries resultantes são assimétricas positivamente, com média do coeficiente de assimetria igual a 0,779, e platicúrticas, com coeficiente de excesso de curtose médio igual a -0,275. Verificou-se que poucas séries tiveram o excesso de curtose positivo, visto que o percentil 97,5% desse descritor estatístico valeu -0,047. Além disso, os coeficientes de correlação serial de 1ª ordem de cada uma das séries de resíduos ε_t oscilaram entre 0,70 e 0,81, com média 0,77, o que demonstra o elevado grau de dependência entre sucessivas diferenças das vazões observadas com as simuladas. Tais aspectos são demonstrados visualmente na Figura 5.14.

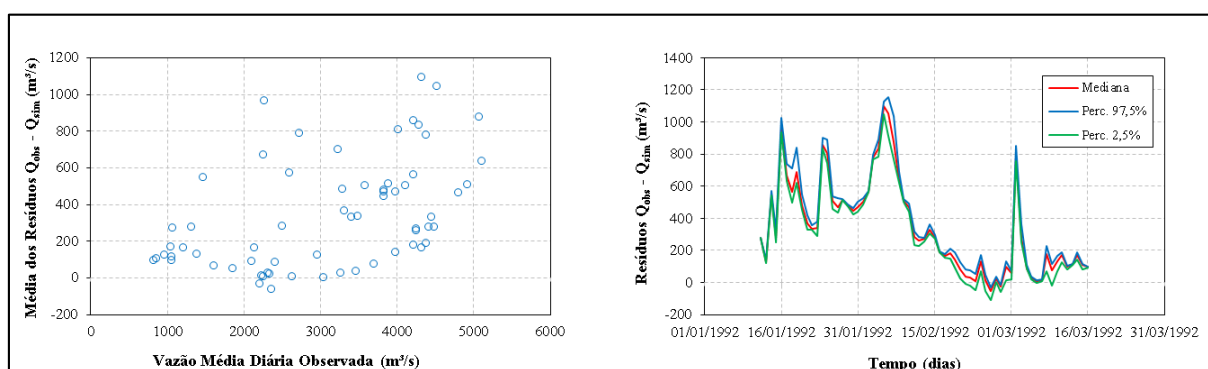


Figura 5.13 – Gráfico A (esq.): Valores médios dos resíduos ε_t em cada intervalo de tempo *versus* as vazões médias diárias observadas no posto fluviométrico de jusante; Gráfico B (dir.): percentis de 2,5%, 50% e 97,5% dos resíduos ε_t estimados para cada intervalo de tempo. Ambos os gráficos correspondem ao Cenário A1.

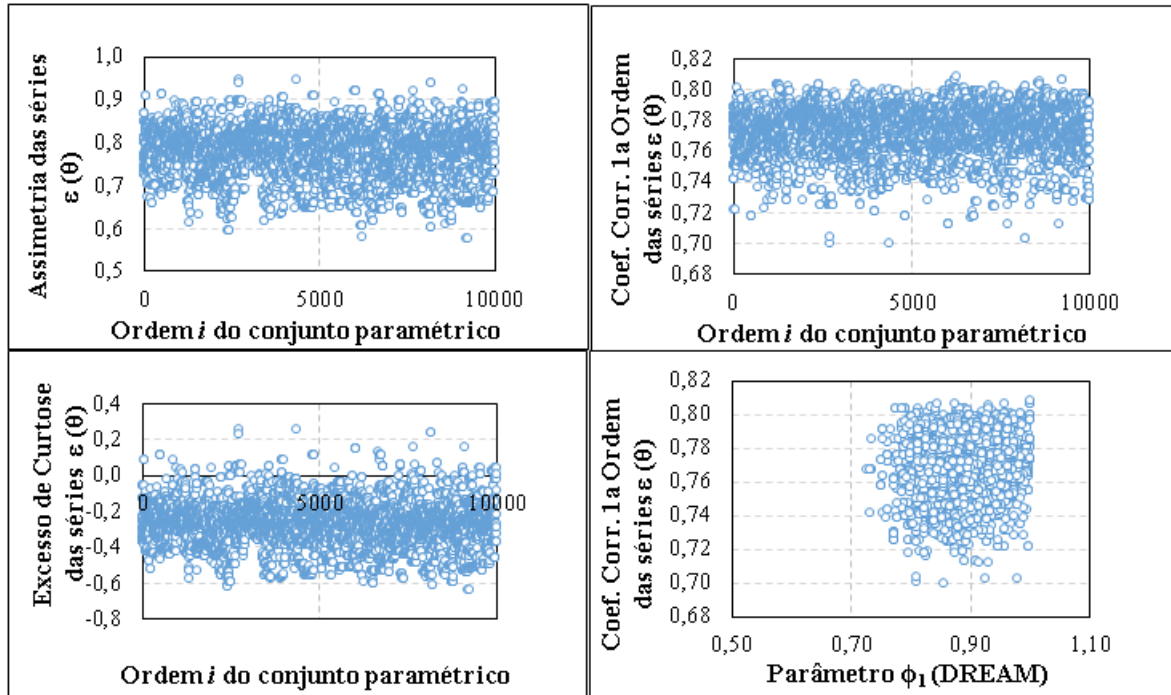


Figura 5.14 – Dispersão dos principais descritores estatísticos das 10.000 séries de resíduos $E(\theta) = \tilde{Y} - Y(\theta)$ obtidas *a posteriori*: Cenário A1.

Observou-se que o parâmetro ϕ_1 mostrou uma tendência de superestimar os coeficientes de correlação amostral de 1ª ordem calculados para essas séries. Enquanto a distribuição *a posteriori* desse parâmetro oscilava entre os percentis de 2,5% e 97,5% com valores de 0,725 e 1,00, respectivamente, estimativas do coeficiente de autocorrelação de 1ª ordem para as séries de ε_t indicavam valores ao redor de 0,700 e 0,809, correspondentes aos mesmos percentis, ou seja, com menor dispersão e valores consistentemente menores, aspecto retratado no 4º gráfico da Figura 5.14. Por alguma razão, o algoritmo DREAM superestimou a variabilidade amostral da correlação serial dos resíduos ε_t , embora suas estimativas englobem toda a faixa dos valores empíricos do correspondente coeficiente. Isso pode ter ocorrido devido ao número reduzido de observações, i.e., as n ordenadas de vazão observada no extremo de jusante do trecho estudado, em comparação ao número de dados observados utilizados na calibração de incertezas por esquemas Bayesianos em modelos chuva-vazão. Outra possível explicação é que valores mais elevados de ϕ_1 se devam à influência do ramo ascendente do hidrograma afluente, bastante pronunciado quando comparado à sua porção média e ao seu ramo descendente, aspecto que pode ter sido acentuado pela ausência das afluições laterais à simulação hidrodinâmica.

Por outro lado, os valores de ϕ_1 próximos a 1,0 acarretaram intervalos de credibilidade a 95% muito amplos (resultados mostrados no subitem seguinte), com amplitudes média e máxima de 6320 m³/s e 8557 m³/s, valores elevados comparados à vazão máxima registrada a jusante, de

5098 m³/s, além de produzir valores negativos para o percentil 2,5% de vazões já no ramo descendente do hidrograma. A geração de intervalos de credibilidade muito amplos foi relatada em outros estudos de inferência Bayesiana cujos dados observados foram séries temporais de vazão (SCHOUPS e VRUGT, 2010; SILVA *et al.*, 2014). Esse aspecto decorre do fato de que elevados coeficientes de correlação serial, ao serem incorporados na rotina de geração de resíduos, criam séries semelhantes a um passeio aleatório, aumentando gradualmente a amplitude dos intervalos de credibilidade, à medida que ocorre o desenvolvimento da cheia considerada.

Em adição, os elevados valores de ϕ_1 inferidos *a posteriori* pela aplicação do algoritmo DREAM provocaram a remoção da correlação serial das séries de resíduos ε_t a tal ponto que os coeficientes de correlação serial de 1ª ordem das séries de resíduos remanescentes, $\Phi_p(\mathbf{B})\varepsilon_t = \sigma_t \cdot a_t$, ou $\varepsilon_t - \phi_1 \cdot \varepsilon_{t-1}$ para AR(1), apresentaram valores persistentemente negativos, com média de -0,272, ainda que os mesmos não sejam estatisticamente distintos de zero, conforme testes realizados.

Todas essas evidências sugeriram a inferência Bayesiana e a posterior inspeção dos resíduos gerados sob a hipótese de um modelo AR(2) para explicar mais adequadamente a dependência temporal entre os resíduos desta modelagem, a fim de possibilitar a obtenção de intervalos de credibilidade mais realísticos, evitar a geração de correlação serial negativa nas séries de resíduos após a aplicação do modelo autoregressivo, e capturar mais fielmente a variabilidade dos coeficientes de correlação serial amostrais de 1ª e, presentemente, 2ª ordem.

Ainda que nesse cenário alternativo (resultados não mostrados) a média de ϕ_1 tenha se reduzido em relação ao seu valor no Cenário A1, a contabilização da correlação serial até o lag 2 produziu, novamente, valores muito amplos para o intervalo de credibilidade a 95%, com vazões negativas no limite inferior do ramo descendente do hidrograma.

A partir desses resultados, e sabendo-se da relevância do parâmetro ϕ_1 na composição da incerteza preditiva, partiu-se para a simulação de um cenário em que $\eta_\varepsilon = \{\phi_1, \sigma_0, \xi, \beta\}$, mantendo-se constante o valor de σ_1 , estimado a partir de outros cenários testados previamente, nos quais invariavelmente sua média tendia a 0,01. Novamente, percebeu-se uma distribuição *a posteriori* para ϕ_1 semelhante àquela produzida no Cenário A1, com média próxima de 0,9.

Ademais, inspecionou-se o aspecto visual do correlograma e do correlograma parcial da série formada pelos valores médios por Δt dos resíduos ε_t , de modo a se averiguar a possibilidade de utilização de um modelo ARMA(p,q), seguindo-se algumas recomendações disponíveis em MORETTIN e TOLOI (2006). O comportamento apresentado pelos resíduos em ambas as situações pareceu sugerir que modelos autoregressivos seriam de fato mais adequados para explicar a correlação serial nesse caso.

As constatações relatadas a respeito do parâmetro ϕ_1 tal como inferido no Cenário A1 mostraram a necessidade de fixação *a priori* de estimativas para o mesmo para sequenciamento da inferência das incertezas dos demais parâmetros e variáveis latentes, o que precipitou a concepção do Cenário A2. Antes que se prossiga à apresentação de seus resultados, algumas análises emanadas do presente cenário são de destaque.

Em primeiro lugar, verificou-se no Cenário A1 a relevância de se manter o parâmetro σ_1 para descrever a heteroscedasticidade dos resíduos ε_t , análise realizada por meio de inspeção visual em gráficos como os da Figura 5.13 e da Figura 5.15. Nesta última, elaborada a partir de recomendações de Wooldridge (2008), o quadrado dos resíduos ε_t avaliados em seus valores medianos, e nos seus percentis 2,5% e 97,5% por intervalo temporal, foram confrontados com a variável explicativa, considerada como sendo as ordenadas do hidrograma de montante no presente caso, sem as contribuições laterais.

Os gráficos da Figura 5.15 parecem sugerir que a variância dos resíduos no Cenário A1 não seja constante, apresentando valores maiores na faixa de vazões afluentes entre 2000 m³/s e 4000 m³/s. Coincidentemente, os transbordamentos da calha principal para as planícies de inundação acontecem na ocorrência dessas descargas, no trecho estudado do rio São Francisco, com pequenas lâminas de água na planície, maior interferência das rugosidades em relação às mesmas e pequenos raios hidráulicos nessa região do canal, o que pode causar descontinuidades numéricas na marcha da propagação da cheia e acarretar incertezas de difícil quantificação. No entanto, o modelo linear para o desvio-padrão dos resíduos, em função da magnitude das vazões simuladas, pode não ser o mais adequado para quantificá-lo. A mudança da Equação 4.2 (página 100), que exprime o modelo vigente para σ_t acoplado à função de verossimilhança Generalizada nas rotinas elaboradas por Vrugt (2016) para aplicação do método DREAM, demandaria alterações no *script* original proposto por esse autor. Embora isso seja possível, sugere-se que

esse aspecto constitua objeto de investigações futuras, a ser realizado idealmente sob a presença de um maior número de ordenadas n de vazões de cheia registradas.

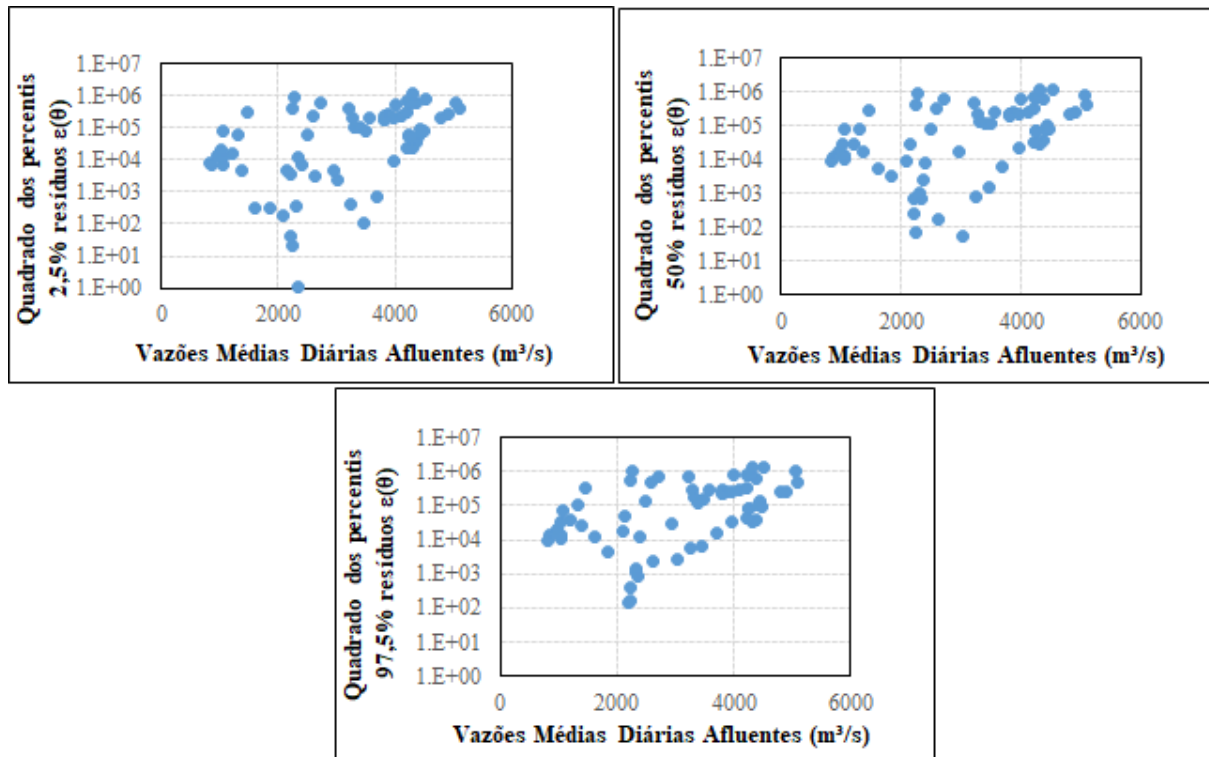


Figura 5.15 – Quadrado dos resíduos ε_t medianos e correspondentes aos percentis 2,5% e 97,5% *versus* as vazões afluentes ao trecho fluvial estudado: Cenário A1.

Em segundo lugar, ainda que as estimativas de ϕ_1 tenham sido superestimadas em relação aos seus correspondentes amostrais, verificou-se que a remoção da correlação serial, via diferenças $\varepsilon_t - \phi_1 \cdot \varepsilon_{t-1}$, reestruturou os resíduos resultantes ao redor de novos valores para as médias dessas séries, conferindo-lhes novo aspecto para a dispersão ao redor de tais valores centrais, e acentuando-lhes a assimetria e a curtose. Os principais descritores estatísticos das séries resultantes após aplicação do filtro autoregressivo de ordem 1 (i.e., remoção da correlação serial) são mostrados graficamente na Figura 5.16.

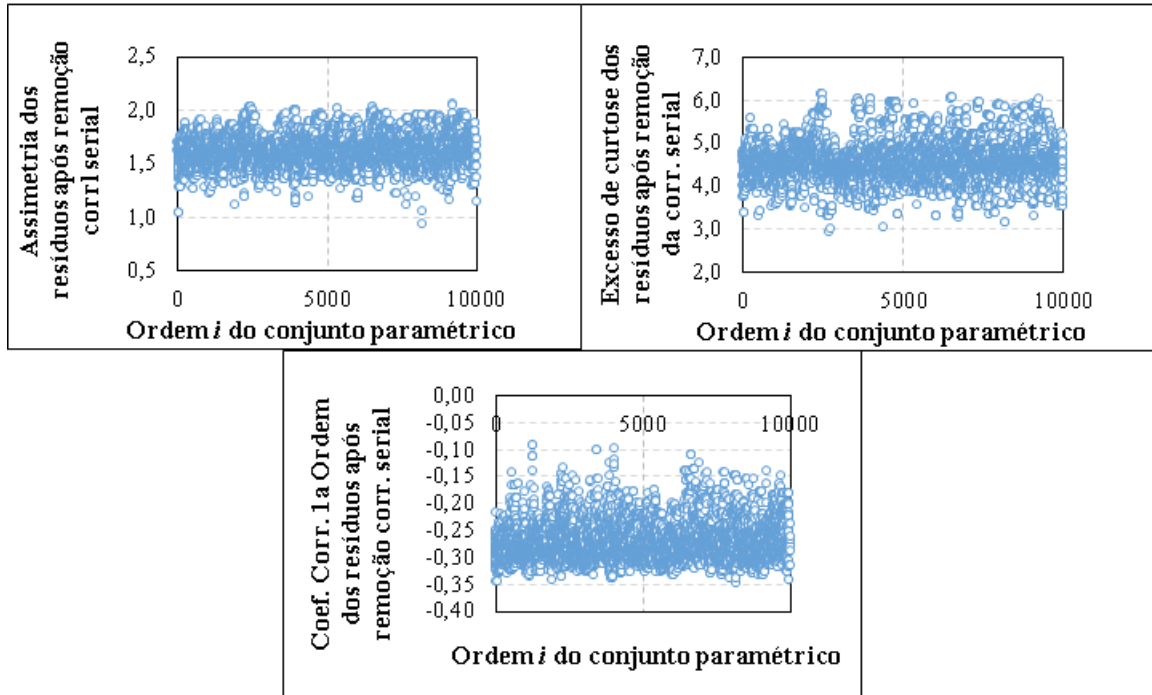


Figura 5.16 – Dispersão dos principais descritores estatísticos das 10.000 séries de resíduos $\sigma_t a_t = \varepsilon_t - \phi_1 \cdot \varepsilon_{t-1}$ obtidas *a posteriori*: Cenário A1.

Ao se dividir os resíduos remanescentes pelo desvio-padrão σ_t , procedendo-se à remoção dos efeitos da heteroscedasticidade, chega-se a séries de resíduos a_t , os quais, em tese, devem ter média nula, desvio-padrão unitário e ser independentes entre si e identicamente distribuídos. O exame dos descritores das séries a_t permite concluir que, no Cenário A1, a média das médias e dos desvios-padrão foram ligeiramente maiores do que os valores modelados pela SEP padronizada, respectivamente iguais a 0,223 e 1,177. Tal fato pode dever-se à persistência de valores elevados dos resíduos iniciais ε_t , que combinados às amostras *a posteriori* do vetor η_ε , não promoveram o alcance das características estatísticas ideais estipuladas. Os coeficientes de assimetria e de excesso de curtose, por sua vez, mantêm valores semelhantes aos das séries de resíduos previamente à divisão por σ_t , conforme era esperado.

Os descritores estatísticos das séries de resíduos ε_t , $\varepsilon_t - \phi_1 \cdot \varepsilon_{t-1}$ e a_t encontram-se sintetizados na Tabela 5.3.

Tabela 5.3 – Principais descritores estatísticos dos resíduos brutos e transformados: Cenário A1.

Tipo de resíduo avaliado	Média	Desvio-Padrão	Assimetria	Excesso Curtose	Coef. Corr. 1ª Ordem	Perc. 2,5%	Mediana	Perc. 97,5%
	(m³/s, se pertinente)		(adim.)			(m³/s, se pertinente)		
Resíduos ε_t (m³/s)	344	298	0,779	-0,275	0,772	-14,5	274	996
Resíduos $\varepsilon_t - \phi_1 \cdot \varepsilon_{t-1}$ (m³/s)	36,9	195	1,618	4,556	-0,272	-223	3,67	619
Resíduos a_t (adim.)	0,223	1,177	1,681	4,855	-0,297	-1,355	0,021	3,697

O terceiro aspecto a ser mencionado acerca dos resíduos deste cenário refere-se à verificação da aderência da SEP $(0,1,\xi,\beta)$ aos resíduos a_t , o que foi feito de maneira qualitativa por meio da inspeção visual entre quantis empíricos (i.e., amostrais) das séries formadas por diversos percentis notáveis calculados para cada Δt , e os correspondentes teóricos (i.e., explicados pela SEP padronizada) calculados a partir dos mesmos percentis dos parâmetros ξ e β . O mesmo foi feito para os valores de ξ e β extraídos da amostra i dentre as $m = 10.000$ amostras que acarretou o máximo da função de log-verossimilhança Generalizada (Equação 4.4, página 100), e os correspondentes quantis empíricos, ou seja, obtidos a partir das diferenças $\mathbf{E}(\theta) = \tilde{\mathbf{Y}} - \mathbf{Y}(\theta)$, sucedidas por $\varepsilon_t - \phi_1 \cdot \varepsilon_{t-1}$ e $a_t = (\varepsilon_t - \phi_1 \cdot \varepsilon_{t-1})/\sigma_t$, estimadas para os valores de ϕ_1 , σ_0 e σ_1 da mesma amostra i . O resultado é mostrado na Figura 5.17. A título de comparação, em todos os casos foram estimados também os quantis teóricos sob a hipótese de normalidade dos resíduos após remoção da dependência temporal e da heteroscedasticidade, ou seja, caso a_t seguisse uma distribuição $N(0,1)$.

Embora o ajuste não seja semelhante ao que se verifica ao se aplicarem algumas distribuições teóricas de probabilidades a amostras de variáveis hidrológicas aleatórias, a exemplo de vazões médias diárias máximas anuais ou vazões mínimas com 7 dias de duração, percebe-se a flexibilidade da SEP em modelar situações distintas da normalidade, representando amostras leptocúrticas e positivamente assimétricas, como o são as séries de resíduos a_t no cenário ora avaliado. Em contraposição, a distribuição Normal padronizada, com média nula e desvio-padrão unitário, mantém-se constante e não se ajusta à forma mais concentrada dos valores da amostra ao redor da média e à assimetria constatada. Por outro lado, a natureza não linear da modelagem hidrodinâmica parece gerar resíduos dificilmente modeláveis por equações matemáticas de alguns parâmetros, e provavelmente seria inútil aplicar testes estatísticos de aderência para conferir a adequabilidade da SEP aos resíduos a_t gerados. Além disso, cabe lembrar que os algoritmos MCMC buscam conjuntos paramétricos desejáveis, idealmente

situados na região de mais alta densidade de probabilidade conjunta, e não o(s) conjunto(s) ótimo(s). Dessa maneira, os parâmetros η_{ϵ} estimados *a posteriori* irão refletir a variabilidade das características estatísticas dos resíduos i.i.d. obtidos na modelagem hidrodinâmica, sem necessariamente apresentar um ajuste matemático mensurável através de métricas comumente utilizadas.

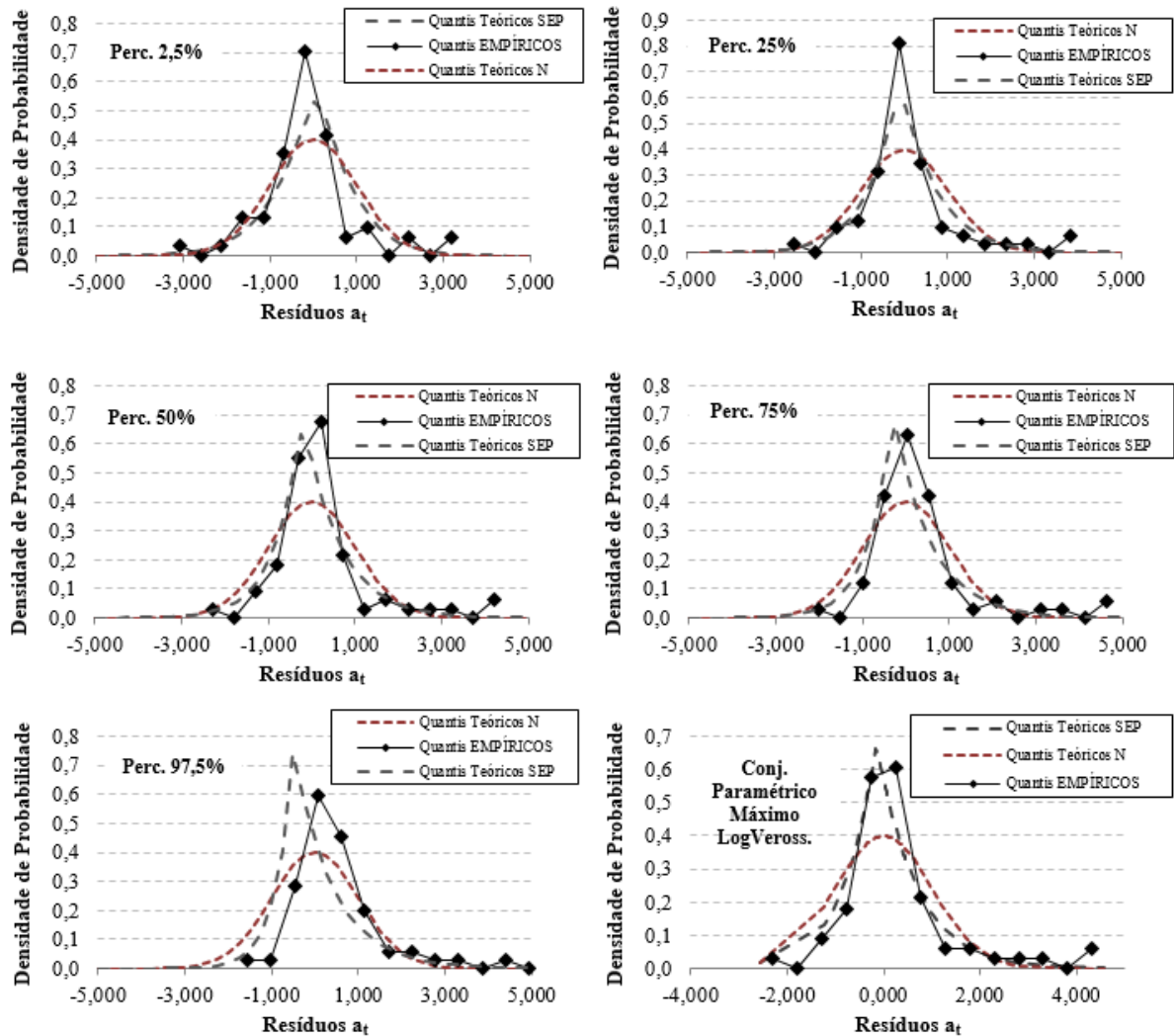


Figura 5.17 – Quantis empíricos de percentis notáveis colhidos por Δt nas séries de resíduos a_t e os correspondentes quantis teóricos obtidos segundo as distribuições SEP e Normal padronizadas: Cenário A1.

Por fim, são apresentadas na Figura 5.18 os histogramas de frequência relativa marginais estimados *a posteriori*, elaborados a partir dos 10.000 conjuntos paramétricos $\{\theta, \eta_{\epsilon}\}$ extraídos após a convergência da distribuição conjunta $p(\theta, \eta_{\epsilon} | \tilde{Y})$, ao lado dos correspondentes valores da função de log-verossimilhança. Alguns de seus descritores estatísticos encontram-se resumidos na Tabela 5.4.

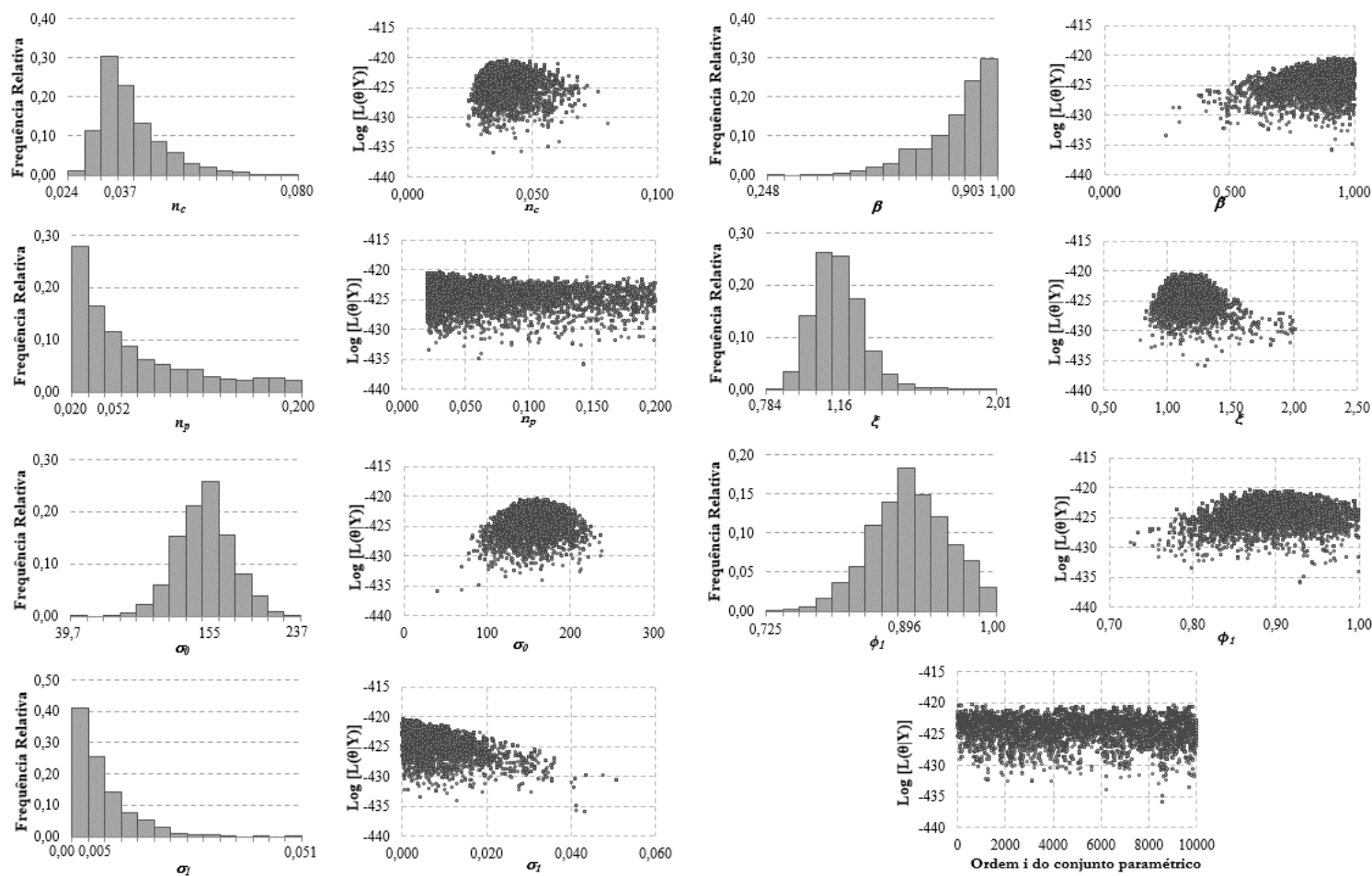


Figura 5.18 – Histogramas de frequências relativas estimados para as amostras *a posteriori* dos parâmetros do modelo hidrodinâmico e da verossimilhança Generalizada: Cenário A1.

Tabela 5.4 – Descritores estatísticos dos parâmetros do modelo hidrodinâmico e da função de verossimilhança Generalizada: Cenário A1.

Parâmetro	Média	Desvio-Padrão	CV	Moda	Mediana	Coefficiente de Assimetria	Quantil 2,5%	Quantil 97,5%
n_c	0,039	0,008	0,197	0,033	0,037	1,281	0,029	0,059
n_p	0,069	0,048	0,691	0,066	0,052	1,095	0,021	0,186
σ_0	155	24	0,152	160	155	0,041	109	202
σ_1	0,007	0,006	0,982	0,001	0,005	1,874	0,000	0,024
β	0,870	0,114	0,131	0,997	0,903	-1,281	0,589	0,996
ξ	1,169	0,142	0,121	1,157	1,156	1,311	0,948	1,476
ϕ_1	0,897	0,046	0,051	0,864	0,896	-0,162	0,804	0,983

A média do coeficiente de Manning na calha *a posteriori* foi de 0,039, com desvio-padrão de 0,008, denotando uma identificação mais precisa quando comparada àquela apresentada pela rugosidade nas planícies. Neste caso, o desvio-padrão foi de 0,048, 6 vezes maior do que na calha, com uma média *a posteriori* de 0,069.

A inspeção dos valores da função de log-verossimilhança dá conta de que seu máximo global ocorre para $n_p = 0,029$, com persistência de valores mais elevados correspondentes à faixa entre 0,020 e 0,050, aproximadamente. No mosaico de dispersão entre pares de parâmetros, exibido na Figura 5.20, percebe-se que à primeira metade da amostra *a posteriori* obtida para n_c correspondem valores de n_p em toda sua faixa estipulada *a priori*, mostrando dificuldade na identificação da rugosidade na planície. A esses valores de n_c provavelmente correspondem as maiores condutâncias na calha principal e, assim, maior capacidade de descarga, restando pouco a ser transportado pelas planícies de inundação, aspecto que poderia resultar em pouca diferenciação entre as rugosidades nessa porção do canal. Por outro lado, valores maiores de n_c , acima de suas medidas de tendência central, parecem combinar-se a valores menores de n_p , provavelmente a fim de compensar a redução da capacidade de descarga na calha principal. De todo modo, como a identificação *a posteriori* de n_p manteve seus limites originais, e com diferenciação menos acentuada da função de log-verossimilhança em relação à condição verificada para n_c , estima-se que a cheia utilizada não tenha tido magnitude suficiente para acarretar lâminas de água na planície a ponto de melhor identificar a rugosidade nessa região. De fato, o período de retorno do evento de 1992 é da ordem de 10 anos. Há registros de maiores enchentes, como a de 1979, mas sua utilização agregaria complexidade às condições de contorno consideradas no modelo e acarretaria incertezas adicionais. A Figura 5.19 mostra a

amplitude total de níveis de água máximos correspondentes às vazões em trânsito na seção das réguas linimétricas do posto de Pirapora-Barreiro simuladas para o conjunto de 10.000 amostras de $\{n_c, n_p\}$ obtidas *a posteriori* no Cenário A1. Observa-se que o maior nível de água atingido nesta seção, tomada como referência para aplicação do esquema Bayesiano de calibração de incertezas, inunda somente parte das margens além da calha principal, consideradas também como componentes da planície de inundação no presente caso. As regiões mais planas, cuja inundação alcançaria grandes extensões, não são afetadas pela cheia de 1992, pelo menos na referida seção. Embora outras seções tenham apresentado lâminas importantes em suas regiões mais amplas e abatidas da planície de inundação para algumas combinações de $\{n_c, n_p\}$, essa informação não pôde ser explícita e diretamente incorporada à metodologia utilizada devido à ausência de dados $\tilde{Y}$ nas mesmas para comparação com as saídas Y do modelo hidrodinâmico.

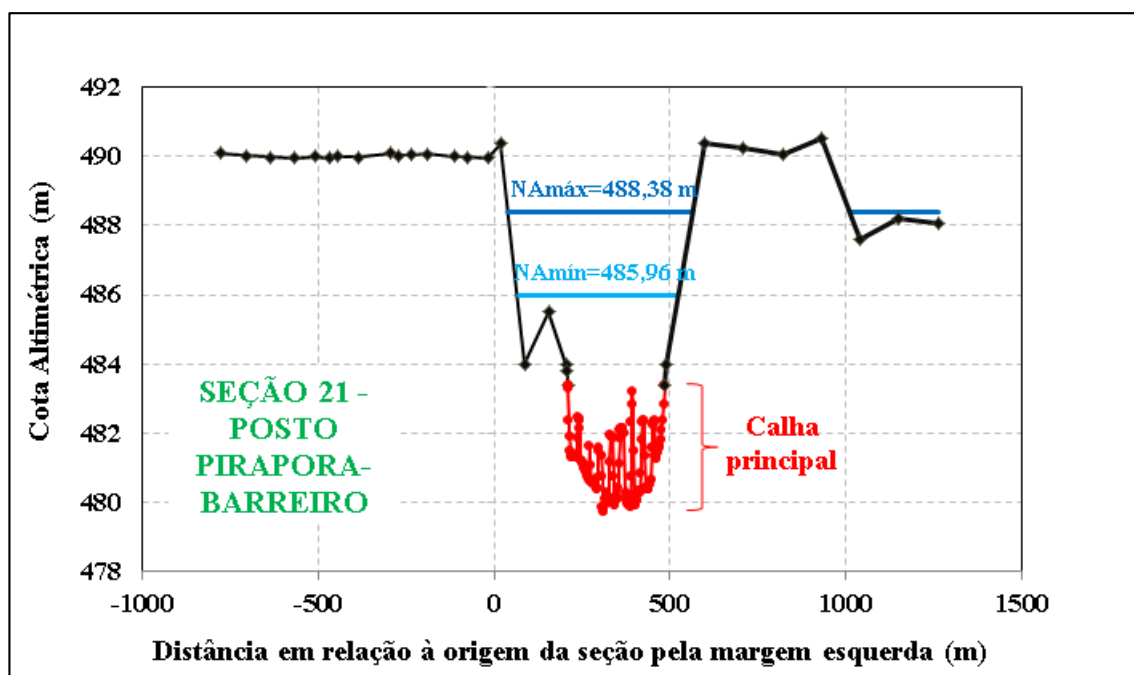


Figura 5.19 – Amplitude de níveis de água correspondentes às descargas em trânsito pela seção de Pirapora-Barreiro resultantes dos 10.000 pares $\{n_c, n_p\}$ obtidos *a posteriori* no Cenário A1.

Adicionalmente, na Figura 5.20 fica demonstrado o grau de correlação existente entre alguns parâmetros e variáveis. No caso de n_c e n_p , por exemplo, embora tenha se assumido que suas definições *a priori* seja independentes, o coeficiente de correlação linear das amostras *a posteriori* entre ambos é igual a -0,56, evidenciando uma possível compensação numérica de menores n_p 's para maiores n_c 's a fim de atingir o objetivo, ou paradigma, de calibração $\tilde{Y}$ a

jusante. Provavelmente, isso não ocorreria caso as rugosidades de Manning fossem selecionadas pelos métodos disponíveis na literatura, em que se busca plausibilidade física e semelhança com tipologias pré-definidas em fotos e tabelas e com materiais constituintes do leito fluvial.

De maneira semelhante ao caso de n_c , as variáveis latentes também tiveram seus limites significativamente estreitados em relação à condição *a priori*. Ressalta-se que o parâmetro σ_1 , com média próxima a 0, não pode ser desconsiderado, pois seus valores *a posteriori*, quando multiplicados pela vazão defluente e somados a σ_o , fornecem desvios-padrão médios entre 140 e 270 m³/s, para as vazões mais elevadas do hidrograma observado, e entre 120 m³/s e 230 m³/s, para as menores vazões. Alguns testes foram feitos considerando-se a variância constante ao longo do tempo, e observou-se que os valores *a posteriori* de σ_o compensam a ausência de σ_1 (resultados não mostrados – Cenário A2: vide Figura 5.12). De fato, ambas as variáveis estão correlacionadas oposta e linearmente em cerca de 40%, aspecto constatado visualmente por meio da inspeção da Figura 5.20. Além disso, as verificações com os resíduos fornecem indícios de que a hipótese de heteroscedasticidade parece mais correta no presente caso, ainda que os resultados indiquem a necessidade de um evento de cheia mais longo a fim de melhor identificar a heteroscedasticidade e a relevância de se testarem outros modelos, além do linear, para explicar o desvio-padrão desses resíduos. Por fim, vale lembrar que Schoups e Vrugt (2010) apontam coeficientes de correlação linear entre variáveis latentes e/ou parâmetros maiores ou iguais a 0,70 como possível indicador de superprescrição de parâmetros à inferência, o que não ocorreu em nenhum cenário avaliado.

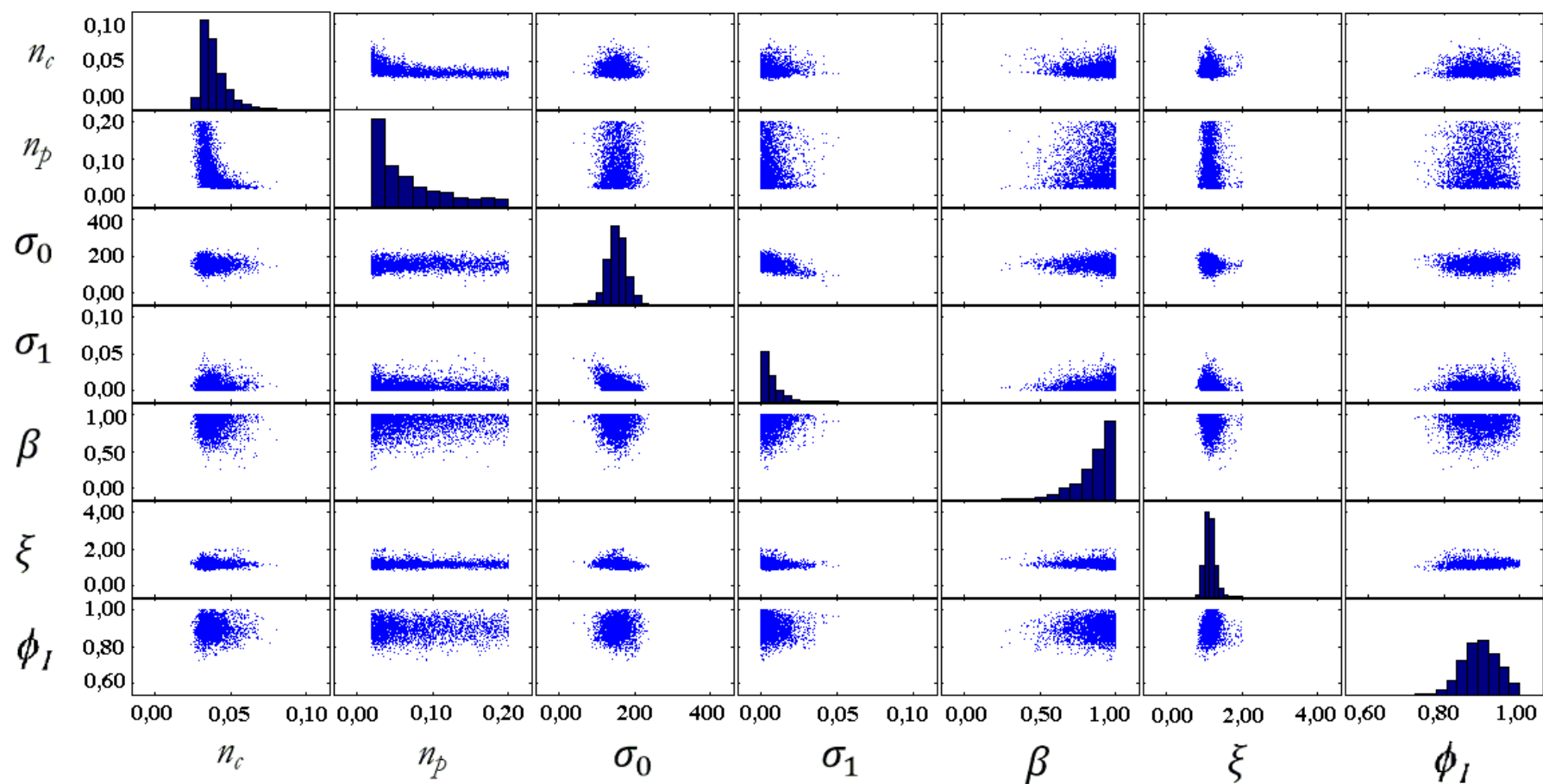


Figura 5.20 – Diagramas de dispersão de valores extraídos das amostras construídas *a posteriori*: Cenário A1.

A variável β , por sua vez, corrobora a aderência dos resíduos a_t a uma distribuição de Laplace, com sua média tendendo à unidade, conforme padronização da SEP apresentada em Schoups e Vrugt (2010). Finalmente, a amostra *a posteriori* inferida para ξ representou a assimetria dos resíduos i.i.d. Obviamente, outros modelos distributivos para os resíduos após remoção da dependência serial e da heteroscedasticidade poderiam ser ajustados, desde que se adaptassem à curtose e assimetria comprovadamente presentes nas séries de resíduos obtidas *a posteriori*.

O Cenário A3, demandado em função da superestimação predominante do parâmetro ϕ_1 no Cenário A1, considerou que o vetor de variáveis latentes seria composto de tal forma que $\eta_\epsilon = \{\sigma_0, \sigma_1, \xi, \beta\}$, com a incerteza de 6 variáveis sujeitas à inferência Bayesiana e ϕ_1 fixado em 0,77, correspondente ao valor médio estimado entre os coeficientes de correlação serial de 1ª ordem para as séries de resíduos ϵ_t do cenário que o antecedeu. Cogitou-se, ainda, em se estabelecer uma distribuição *a priori* para ϕ_1 mais restritiva do que nos Cenários A1 e A2, com limites inferior e superior próximos ao valor médio supracitado, extraídos a partir da inspeção das séries ϵ_t . No entanto, é provável que a distribuição marginal *a posteriori* desse parâmetro tenderia ao limite superior, em função do que foi observado nos cenários precedentes.

Os resultados foram muito semelhantes aos do Cenário A1, no que concerne ao comportamento estatístico das séries de todos os tipos de resíduos avaliados: (i) resíduos ϵ_t , (ii) resíduos $\epsilon_t - \phi_1 \cdot \epsilon_{t-1}$, ou seja, após remoção da dependência serial segundo o modelo autoregressivo mais cogitado, de ordem 1, e (iii) resíduos a_t , após remoção da heteroscedasticidade das séries $\epsilon_t - \phi_1 \cdot \epsilon_{t-1}$. Alguns elementos de destaque são mencionados a seguir e podem ser conferidos também na Tabela 5.5:

- Resíduos ϵ_t predominantemente positivos, com média e desvio-padrão elevados, assimetria positiva, embora não elevada (média = 0,780) e excesso de curtose negativo (média = -0,264). Em adição, tais séries de resíduos apresentaram valor do coeficiente de correlação serial de 1ª ordem coerente com o valor fixado *a priori* para ϕ_1 neste caso;
- A remoção da dependência temporal pelo modelo AR(1), com ϕ_1 fixo em 0,77, proporciona um rearranjo dos resíduos resultantes e avaliados em termos médios por Δt ao redor do zero (vide Figura 5.21), bem como acentua a assimetria positiva (média = 1,736) e muda a polaridade do excesso de curtose (média = 4,250). A média e o desvio-padrão prevaletentes

são minimizados, mas seguem positivos (médias iguais a 76,9 m³/s e 191 m³/s, respectivamente);

- A remoção da variância variável, a fim de que sejam obtidos resíduos padronizados (i.e., com média nula e desvio-padrão unitário) e i.i.d., produz séries a_t com média e desvio-padrão levemente distintos do ideal (valores médios iguais a 0,422 e 1,070, respectivamente). O coeficiente de correlação serial de 1ª ordem, por sua vez, indica um valor residual negativo, embora sem significância estatística, com média de -0,175, menor, portanto, do que a média do Cenário A1 nessa situação; e
- A aderência entre quantis empíricos e teóricos de percentis extraídos das séries a_t e dos parâmetros β e ξ apresenta aspecto semelhante ao do Cenário A1 (resultados não mostrados), revelando ajustes aceitáveis, avaliados qualitativamente, da SEP aos resíduos i.i.d. resultantes da modelagem hidrodinâmica em questão.

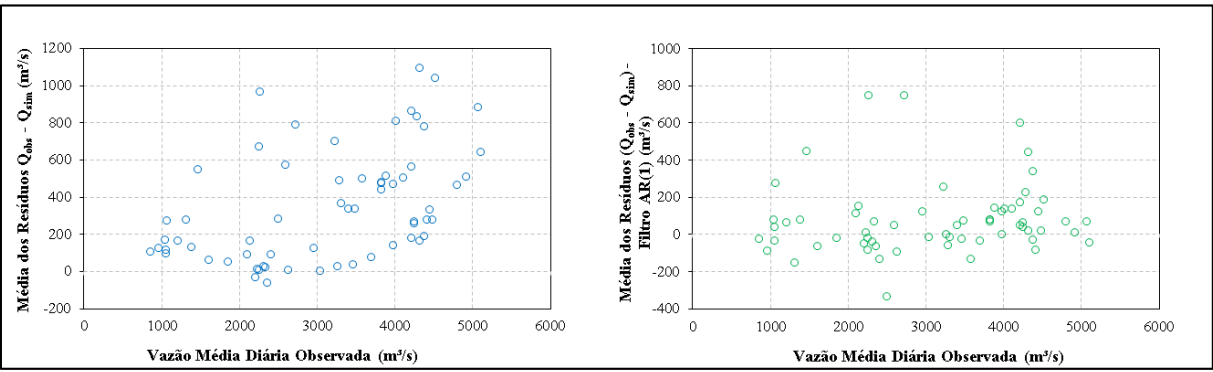


Figura 5.21 – Valores médios dos resíduos brutos e após remoção da correlação serial em relação à vazão observada no posto fluviométrico de jusante: Cenário A3.

Tabela 5.5 – Principais descritores estatísticos dos resíduos brutos e transformados: Cenário A3.

Tipo de resíduo avaliado	Média	Desvio-Padrão	Assimetria	Excesso Curtose	Coef. Corr. 1ª Ordem	Perc. 2,5%	Mediana	Perc. 97,5%
	(m³/s, se pertinente)		(adim.)			(m³/s, se pertinente)		
Resíduos ε_t (m³/s)	344	298	0,780	-0,264	0,771	-16,8	274	996
Resíduos $\varepsilon_t - \phi_1 \cdot \varepsilon_{t-1}$ (m³/s)	76,9	191	1,736	4,250	-0,175	-154	38,78	654
Resíduos a_t (adim.)	0,422	1,070	1,829	4,737	-0,208	-0,867	0,208	3,590

Ademais, os resultados quanto às distribuições marginais *a posteriori* dos parâmetros e variáveis foram muito parecidos àqueles extraídos do Cenário A1, em termos de forma, escala e posição, conforme pode ser conferido na Tabela 5.6.

Tabela 5.6 – Descritores estatísticos dos parâmetros do modelo hidrodinâmico e da função de verossimilhança Generalizada: Cenário A3.

Parâmetro	Média	Desvio-Padrão	CV	Moda	Mediana	Coefficiente de Assimetria	Quantil 2,5%	Quantil 97,5%
n_c	0,039	0,009	0,232	0,030	0,037	1,930	0,027	0,061
n_p	0,076	0,052	0,682	0,026	0,058	0,800	0,021	0,190
σ_0	164	30	0,185	225	164	0,180	107	227
σ_1	0,009	0,008	0,915	0,000	0,007	1,635	0,000	0,030
β	0,865	0,118	0,137	0,935	0,897	-1,189	0,562	0,996
ξ	1,064	0,133	0,125	0,955	1,054	1,154	0,855	1,370
ϕ_1	valor fixo = 0,77							

Todos esses aspectos, aliados à estimação do intervalo de credibilidade a 95% com limites inferior e superior fisicamente plausíveis (gráfico a ser mostrado no subitem 5.3.6), comparativamente aos produzidos no Cenário A1, conduzem à conclusão que a fixação de ϕ_1 não causou impactos relevantes à inferência das incertezas das demais variáveis e dos parâmetros envolvidos, proporcionando, ainda, a possibilidade de aplicação dos intervalos preditivos para a cheia utilizada na calibração em estudos futuros que demandem estimativas da incerteza preditiva total associada a cheias de magnitude semelhante.

Por fim, o Cenário A4 constitui-se pela inferência das incertezas do vetor paramétrico $\{\theta, \eta_\epsilon\} = \{n_c, n_p, \sigma_0\}$, a fim de se testar o impacto sobre a quantificação de incertezas sob a adoção de hipóteses acerca dos resíduos possivelmente inválidas para uma modelagem hidrodinâmica.

Os resultados, tanto em termos de resíduos, quanto das amostras obtidas *a posteriori* das variáveis envolvidas, revelam que ocorreram diferenças importantes em relação aos cenários anteriores do tipo A, conforme mostram a Tabela 5.7 e a Tabela 5.8, bem como a Figura 5.22 e a Figura 5.23.

Tabela 5.7 – Principais descritores estatísticos dos resíduos brutos e transformados: Cenário A4.

Tipo de resíduo avaliado	Média	Desvio-Padrão	Assimetria	Excesso Curtose	Coef. Corr. 1ª Ordem	Perc. 2,5%	Mediana	Perc. 97,5%
	(m³/s, se pertinente)		(adim.)			(m³/s, se pertinente)		
Resíduos ε_t (m³/s)	344	269	1,023	0,543	0,660	14,0	270	977
Resíduos $\varepsilon_t - \phi_1 \cdot \varepsilon_{t-1}$ (m³/s)	344	269	1,023	0,543	0,660	14,0	270	977
Resíduos a_t (adim.)	0,782	0,611	1,023	0,543	0,660	0,032	0,614	2,221

Tabela 5.8 – Descritores estatísticos dos parâmetros do modelo hidrodinâmico e da função de verossimilhança Generalizada: Cenário A4.

Parâmetro	Média	Desvio-Padrão	CV	Moda	Mediana	Coefficiente de Assimetria	Quantil 2,5%	Quantil 97,5%
n_c	0,020	0,006	0,300	0,014	0,019	1,956	0,012	0,035
n_p	0,107	0,053	0,491	0,028	0,106	0,047	0,024	0,195
σ_0	444	41	0,092	422	441	0,488	373	532
σ_1				valor fixo = 0,00				
β				valor fixo = 0,00				
ξ				valor fixo = 1,00				
ϕ_1				valor fixo = 0,00				

As seguintes conclusões podem ser destacadas em relação aos resíduos:

- Média e desvio-padrão das séries ε_t elevados, inalterados em relação aos cenários A1, A2 e A3, com médias iguais a 344 m³/s e 269 m³/s, respectivamente, para ambos os estimadores. A assimetria e o excesso de curtose, por sua vez, apresentaram valores maiores do que nos casos anteriores: médias de 1,023 e 0,543, respectivamente, sendo que as séries de resíduos desta vez têm forma leptocúrtica. As diferenças nesses dois coeficientes em relação às situações anteriormente simuladas provavelmente devem-se ao comportamento distinto das distribuições *a posteriori* do vetor paramétrico, em especial de n_c ;
- Os resíduos após padronização pelo desvio-padrão (i.e., com divisão de ε_t por σ_0) apresentaram valores médios para a média e o desvio-padrão iguais a 0,782 e 0,611, respectivamente, portanto, mais distantes dos parâmetros da distribuição padronizada SEP(0,1, ξ , β), que, no presente caso, reduz-se à distribuição Normal padrão, $\mathcal{N}(0,1)$. Em último caso, poderia ser testada uma distribuição Normal com parâmetros distintos dos indicados, desde que sujeitos à inferência Bayesiana em conjunto com o vetor $\{n_c, n_p, \sigma_0\}$. No entanto, o objetivo principal nessa situação é avaliar o impacto sobre a quantificação das incertezas com a adoção de premissas as mais simples possíveis quanto aos resíduos. Adicionalmente, as séries resultantes mantiveram assimetria e excesso de curtose positivos, o que evidenciava que o ajuste a uma distribuição Normal poderia ser inadequado; e
- Em terceiro lugar, verifica-se pela Figura 5.22 que o ajuste da distribuição Normal padrão aos resíduos ε_t/σ_0 é, em termos visuais, sensivelmente pior do que nos cenários anteriores, comportamento previsto em função da análise dos descritores estatísticos.

- No que diz respeito à incerteza paramétrica *a posteriori*, é possível constatar pela Figura 5.23 e pela
- Tabela 5.8 que o formato do histograma obtido para n_c e n_p é distinto daqueles dos cenários A1 a A3. No caso de n_c , houve uma tendência de concentrar os coeficientes de rugosidade de Manning numa faixa de valores menores do que nos demais casos avaliados até então, entre 0,012 e 0,035 (percentis 2,5% e 97,5%, respectivamente), e mediana de 0,019. Em paralelo, a julgar pelas frequências relativas, a distribuição da rugosidade nas planícies, n_p , manteve-se praticamente uniforme, com pouca diferenciação em relação à distribuição *a priori*. No diagrama de dispersão entre os valores dos parâmetros $\{n_c, n_p, \sigma_o\}$ (Figura 5.24), percebe-se pouca correlação entre os parâmetros n_c e n_p , sendo que toda a faixa *a posteriori* do segundo (com comportamento praticamente idêntico àquele definido *a priori*) é varrida pelos valores *a posteriori* do primeiro. Isso sugere que a incorreta prescrição da função de verossimilhança, nesse caso baseada nos quadrados das diferenças entre valores observados e simulados, gerou uma tendência ainda maior do que nos cenários anteriores em concentrar as vazões em trânsito na calha principal, correspondendo a menores rugosidade nessa porção, o que levou à menor necessidade de transporte pelas planícies. Em termos conceituais, pode-se assim inferir que a pouca informação sobre o hidrograma de fevereiro de 1992 remanescente nas planícies não foi suficiente para diferenciar o conhecimento estipulado *a priori* para a rugosidade nessa região do canal, aspecto que teve maior impacto sobre a inferência das rugosidades de Manning sob hipóteses inadequadas para modelar os resíduos resultantes.

Quanto a σ_o , verifica-se que seus valores foram maiores do que os identificados nos cenários A1 a A3, provavelmente devido à ausência de σ_1 , que, ao ser multiplicado pelas vazões simuladas, y_t , gera valores da mesma magnitude de σ_o , conforme verificado em cenários alternativos sem a presença da inclinação do modelo linear do desvio-padrão. No entanto, os valores de σ_o presentemente obtidos *a posteriori* superestimaram os valores de σ_t dos cenários anteriores, assim como os desvios-padrão de correspondentes séries ε_t .

Outro aspecto que merece destaque trata-se da pouca diferenciação dos valores da função de log-verossimilhança nesse cenário comparativamente àqueles computados nos casos anteriores, o que se constitui em outro indício da incorreta definição da mesma e das hipóteses acerca dos resíduos da modelagem que respaldam sua construção.

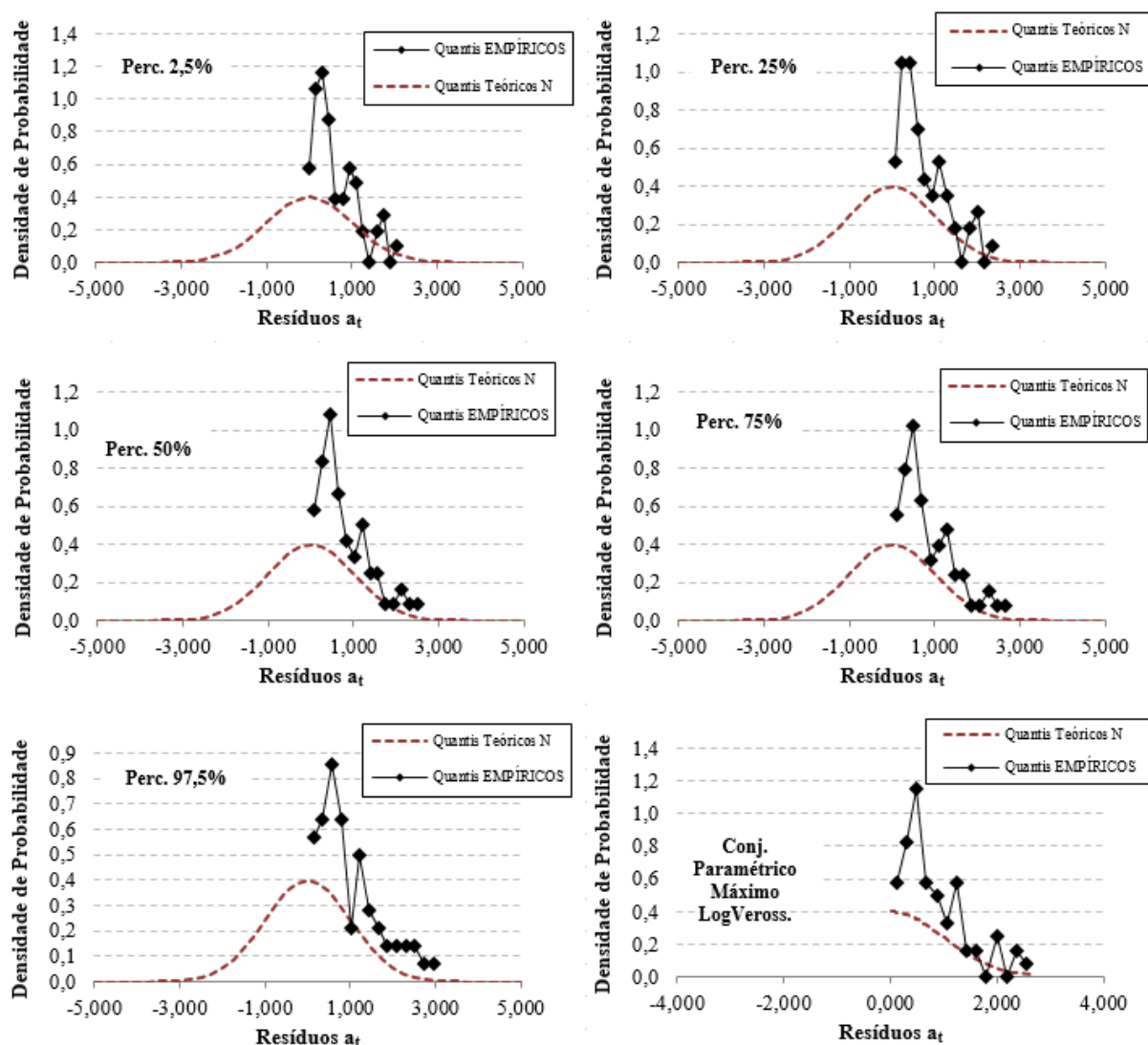


Figura 5.22 – Quantis empíricos de percentis notáveis colhidos por Δt nas séries de resíduos a_t e os correspondentes quantis teóricos segundo as distribuições SEP e N padronizadas: Cenário A4.

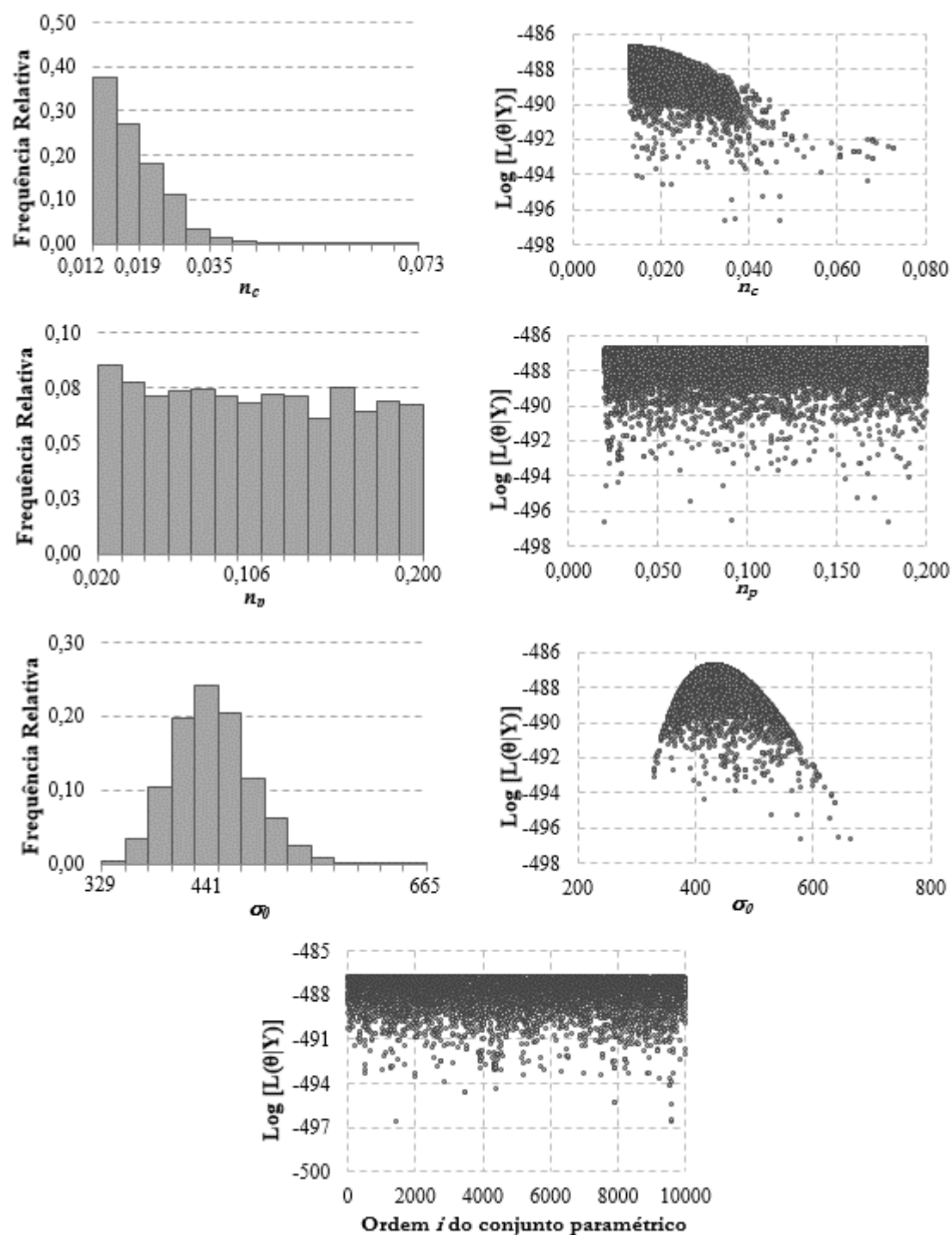


Figura 5.23 – Histogramas de frequências relativas estimados para as amostras *a posteriori* dos parâmetros do modelo hidrodinâmico e da verossimilhança Generalizada: Cenário A4.

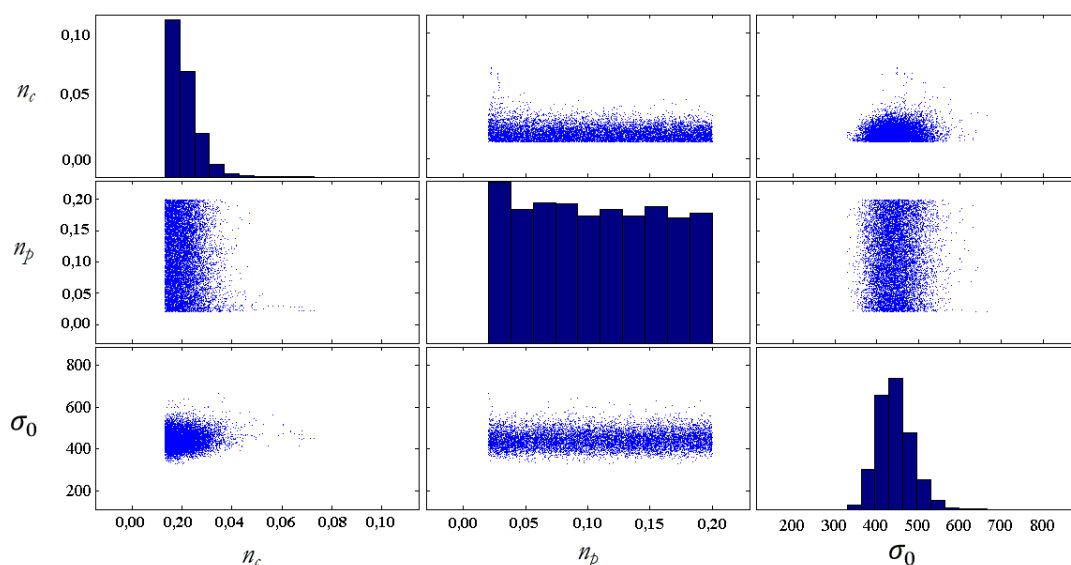


Figura 5.24 – Diagramas de dispersão de valores extraídos das amostras construídas a *posteriori*: Cenário A4.

5.3.5.2 Cenários tipo B: cotograma de jusante como paradigma de calibração e ausência de afluentes laterais como condição de contorno

Os resultados emanados das análises com hidrogramas suscitaram a inspeção detalhada do que poderia ocorrer com os resíduos e a identificação das incertezas caso a marcha temporal de níveis de água registrados na estação fluviométrica de jusante fosse utilizada como paradigma $\tilde{Y}$.

As análises iniciais, realizadas no Cenário B1, sob a premissa da necessidade de 7 variáveis e/ou parâmetros a serem inferidos no processo Bayesiano, mostraram a redundância de σ_1 para modelar o desvio-padrão, contribuindo para que os cenários subsequentes se baseassem na hipótese de homoscedasticidade. Portanto, os cenários cujos resultados serão apresentados são os de código B2 a B4.

De modo geral, os resíduos desses 3 cenários caracterizam-se pelos seguintes aspectos, sintetizados na Tabela 5.9 (Cenário B2) e na Tabela 5.10 (Cenário B4):

- Médias e desvios-padrão das séries de resíduos ε_t praticamente nulos, com suas médias um pouco maiores do que zero;
- Coeficientes de assimetria e de excesso de curtose positivos, com médias ao redor de 0,8 e 0,7, respectivamente, em termos de valores aproximados para os cenários apresentados;

- Forte correlação serial, com coeficiente de correlação linear de 1ª ordem oscilando entre 0,6 e 0,8 (valores aproximados, a fim de representar os 3 cenários em questão); e
- A remoção da dependência serial é, portanto, relevante nessa condição, e foi realizada nos Cenários B2 e B3. Além de promover independência aos resíduos, esse procedimento os reorganizou ao redor do zero de maneira homogênea (vide Figura 5.25). Ainda assim, isso não significa que os resíduos ε_t sejam heteroscedásticos, conforme se constata pela Figura 5.25.

Tabela 5.9 – Principais descritores estatísticos dos resíduos brutos e transformados: Cenário B2.

Tipo de resíduo avaliado	Média	Desvio-Padrão	Assimetria	Excesso Curtose	Coef. Corr. 1ª Ordem	Perc. 2,5%	Mediana	Perc. 97,5%
	(m, se pertinente)		(adim.)			(m, se pertinente)		
Resíduos ε_t (m)	0,012	0,153	0,830	0,682	0,780	-0,201	-0,030	0,342
Resíduos $\varepsilon_t - \phi_1 \cdot \varepsilon_{t-1}$ (m)	0,006	0,101	1,598	3,441	0,083	-0,128	-0,022	0,292
Resíduos a_t (adim.)	0,063	1,069	1,598	3,441	0,083	-1,371	-0,224	3,099

Tabela 5.10 – Principais descritores estatísticos dos resíduos brutos e transformados: Cenário B4.

Tipo de resíduo avaliado	Média	Desvio-Padrão	Assimetria	Excesso Curtose	Coef. Corr. 1ª Ordem	Perc. 2,5%	Mediana	Perc. 97,5%
	(m, se pertinente)		(adim.)			(m, se pertinente)		
Resíduos ε_t (m)	0,016	0,152	0,833	0,693	0,778	-0,196	-0,025	0,345
Resíduos $\varepsilon_t - \phi_1 \cdot \varepsilon_{t-1}$ (m)	0,016	0,152	0,833	0,693	0,778	-0,196	-0,025	0,345
Resíduos a_t (adim.)	0,100	0,980	0,833	0,693	0,778	-1,264	-0,166	2,221

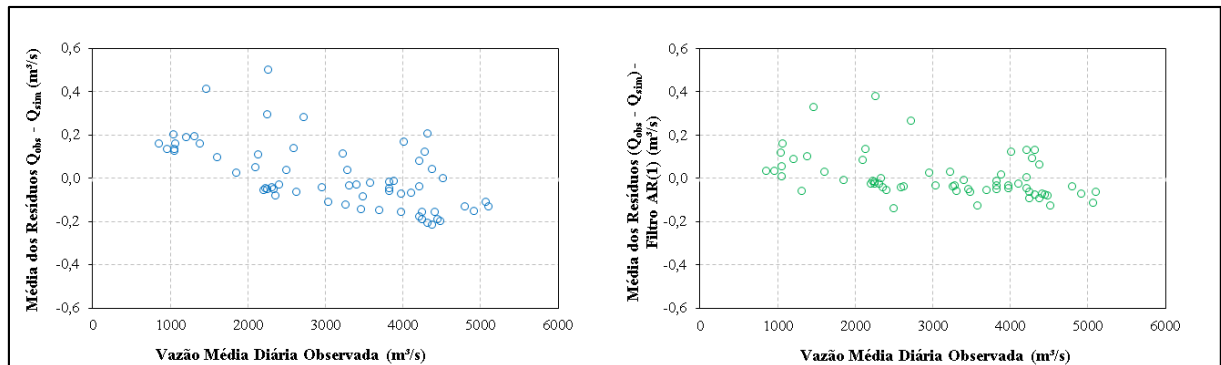


Figura 5.25 – Valores médios dos resíduos brutos e após remoção da correlação serial em relação à vazão observada no posto fluviométrico de jusante: Cenário B2.

No Cenário B2, observou-se que houve concentração de frequências relativas ao redor de dois valores para o parâmetro ϕ_1 : um pico maior, ao redor da moda da distribuição obtida *a posteriori*, de 0,532, e um pico menor, ao redor de 0,690, valor um pouco maior do que a mediana, de 0,620, tendo a dispersão da amostra ocorrido entre percentis 2,5% e 97,5% iguais a 0,489 e 0,827, respectivamente. Tal fato pode ter decorrido de que as soluções conjuntas buscadas pelo algoritmo DREAM foram atraídas para duas regiões do espaço multiparamétrico. Em adição, os coeficientes de correlação linear de 1ª ordem dos resíduos ε_t nesse cenário dispersaram-se entre os referidos percentis de 0,746 e 0,813, respectivamente, com mediana de 0,783. Percebe-se que, no presente caso, a variação de ϕ_1 foi mais ampla do que seus estimadores empíricos, emanados das séries ε_t . Outro aspecto que se sobressaiu nesse cenário foi a identificação ampla de ξ , associada à assimetria das séries de resíduos a_t . Essas duas especificidades fomentaram a concepção do Cenário B3, no qual o parâmetro ϕ_1 foi fixado em seu valor médio, igual a 0,78, e estimado a partir das séries ε_t do cenário imediatamente anterior. Nesse caso, ξ teve seus limites estreitados em relação à faixa definida *a priori*. Paralelamente, verificou-se que o impacto sobre as séries de resíduos ao se fixar ϕ_1 foi pequeno, de modo que os descritores estatísticos das séries ε_t , $\varepsilon_t - \phi_1 \cdot \varepsilon_{t-1}$ e a_t mantiveram-se muito semelhantes aos do Cenário B2. De modo semelhante, o ajuste entre quantis empíricos e teóricos para as amostras formadas por percentis notáveis de a_t em cada Δt e estimadas para os percentis equivalentes de ξ e β foi parecido entre os referidos cenários, conforme se pode constatar pela Figura 5.26 (relativa ao Cenário B2, mas representativa do ajuste também do Cenário B3). Cabe ressaltar que as amostras empíricas de a_t nesse tipo de cenário mostraram-se mais irregulares do que no caso dos cenários do tipo A, e o ajuste da SEP, avaliado de forma qualitativa e conceitual, pareceu adequado, mas com qualidade inferior àquela verificada nos cenários precedentes.

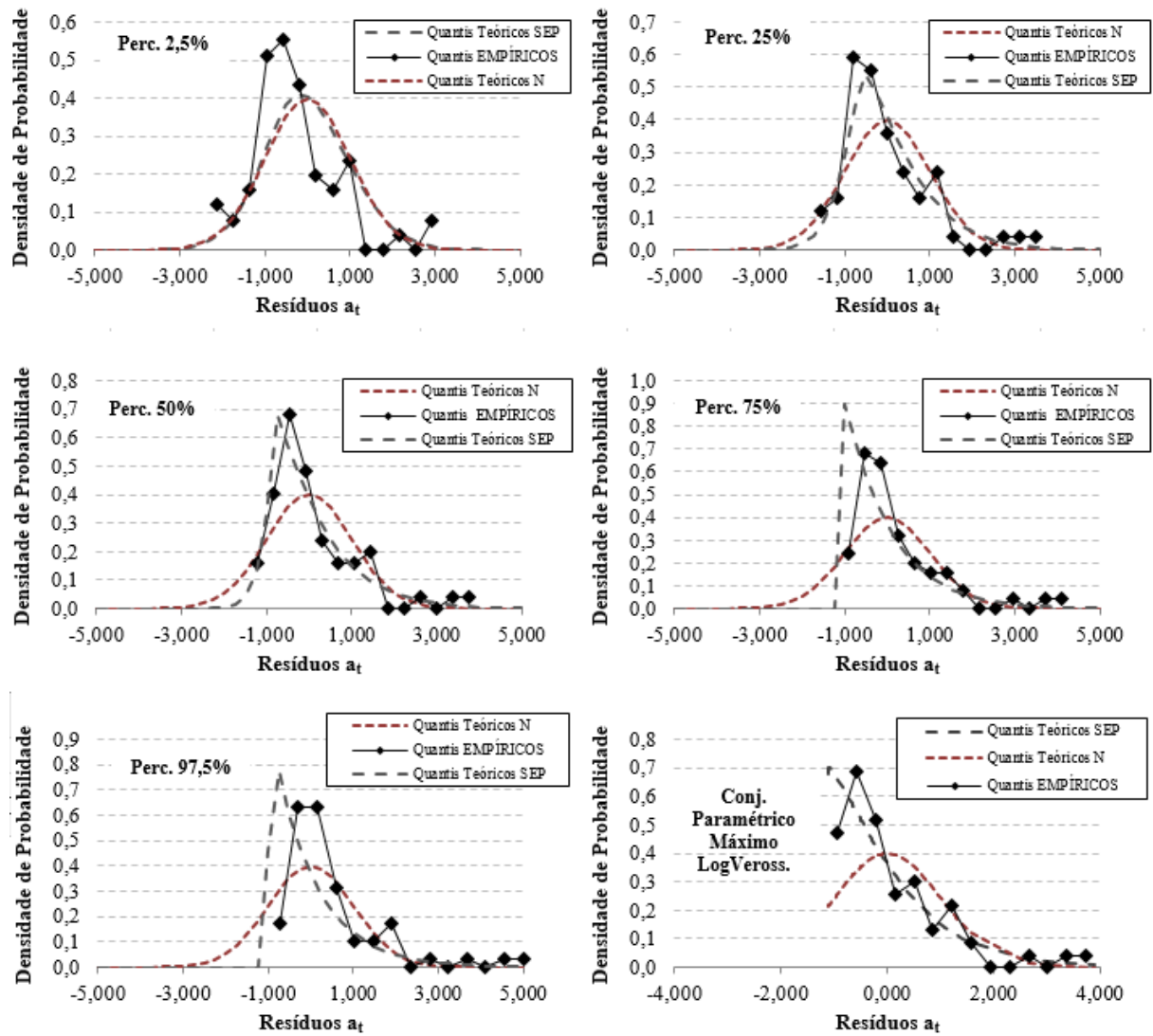


Figura 5.26 – Quantis empíricos de percentis notáveis colhidos por Δt nas séries de resíduos a_t e os correspondentes quantis teóricos segundo as distribuições SEP e N padronizadas: Cenário B2.

Ao se analisar os histogramas elaborados para as amostras construídas *a posteriori* dos cenários B2 e B3 (Figura 5.27 e Figura 5.28), e os descritores estatísticos das mesmas (Tabela 5.11 e Tabela 5.12), percebe-se que houve pouca diferença na identificação dos parâmetros n_c , n_p , σ_o e β . A exceção ocorreu para ξ , que no cenário B3 ficou restrito a uma faixa mais estreita em relação aos limites estabelecidos previamente, e mostrando prevalência de resíduos a_t assimétricos positivamente, em consonância aos coeficientes de assimetria estimados para tais séries.

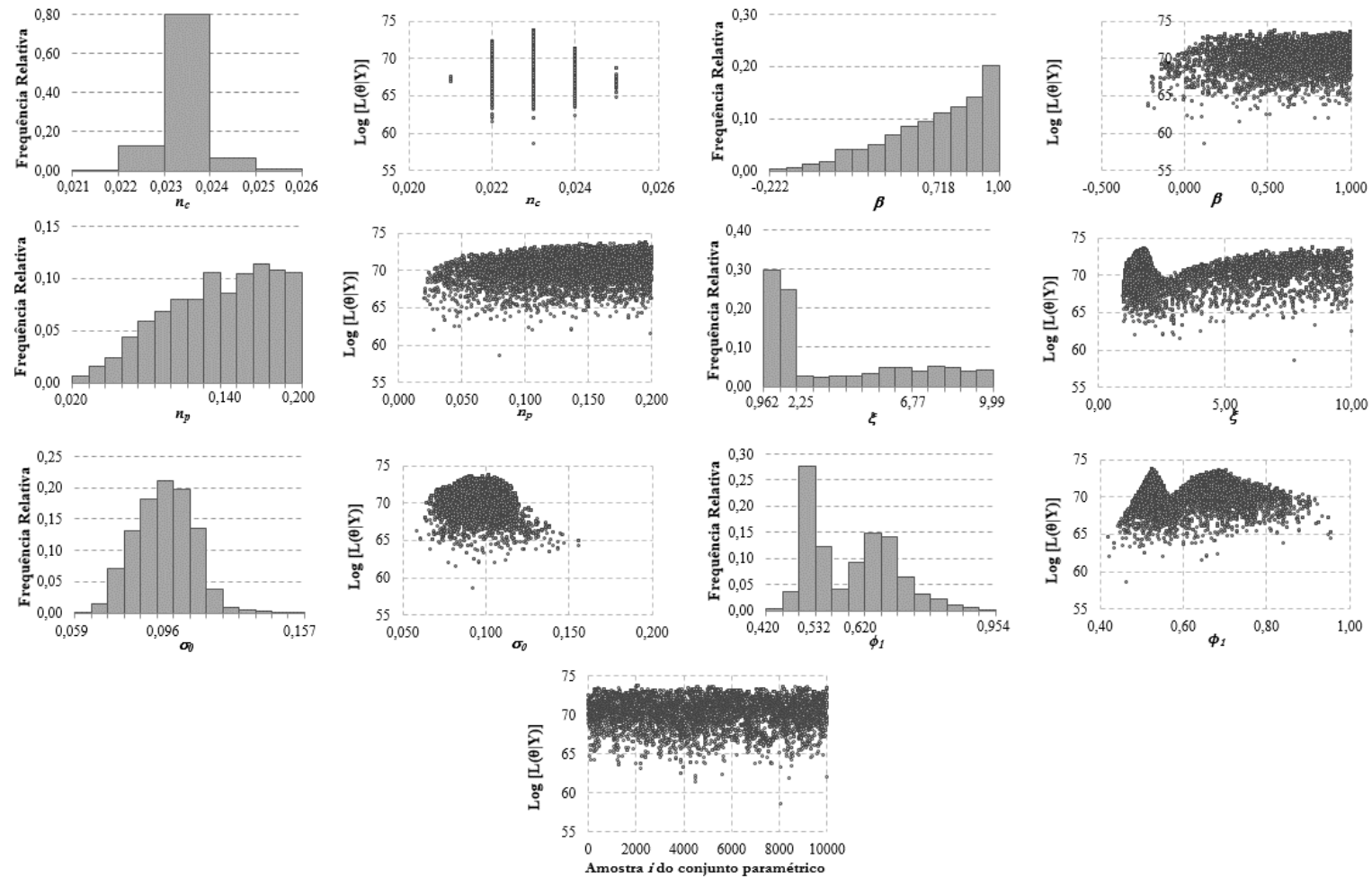


Figura 5.27 – Histogramas de frequências relativas estimados para as amostras *a posteriori* dos parâmetros do modelo hidrodinâmico e da verossimilhança Generalizada: Cenário B2.

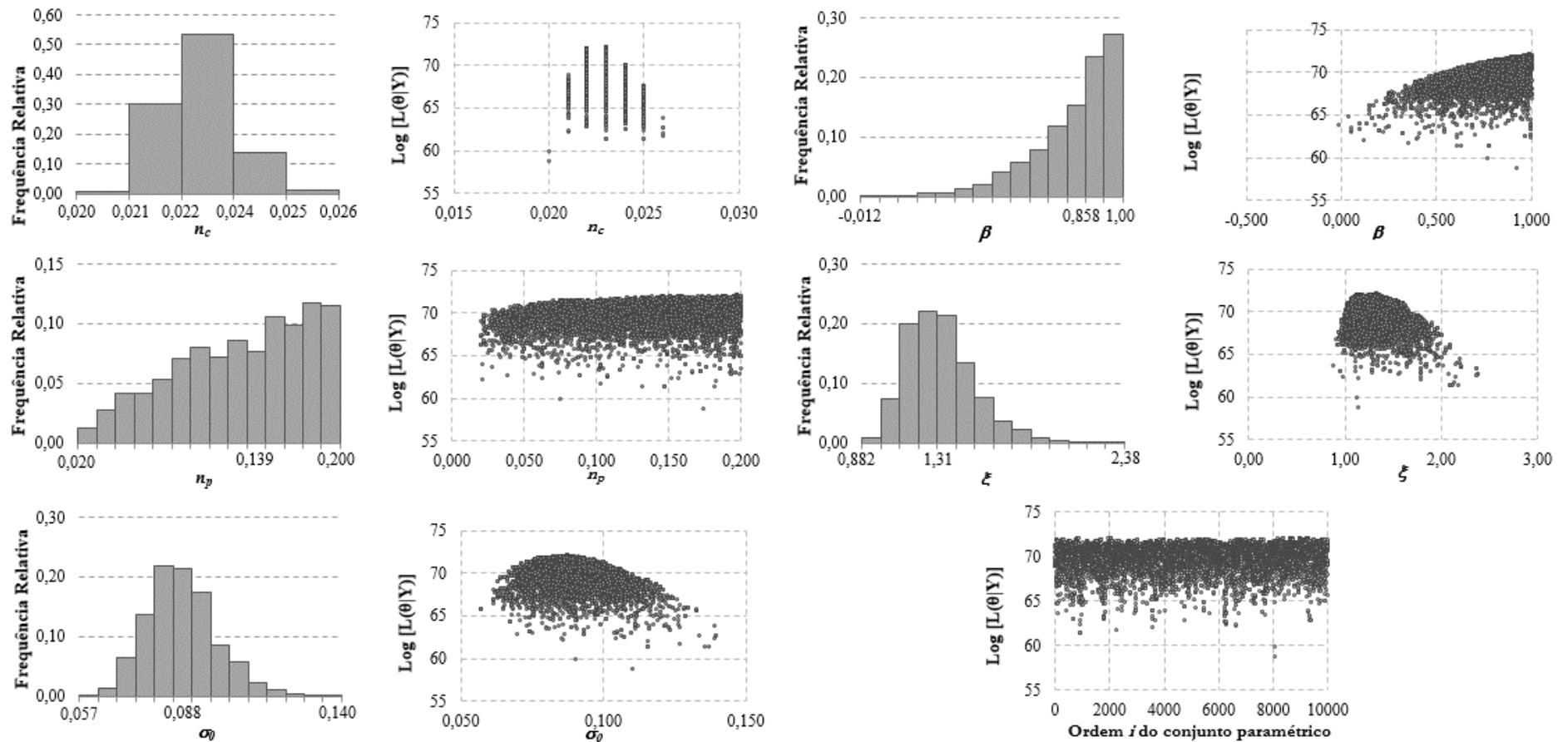


Figura 5.28 – Histogramas de frequências relativas estimados para as amostras *a posteriori* dos parâmetros do modelo hidrodinâmico e da verossimilhança Generalizada: Cenário B3.

Tabela 5.11 – Descritores estatísticos dos parâmetros do modelo hidrodinâmico e da função de verossimilhança Generalizada: Cenário B2.

Parâmetro	Média	Desvio-Padrão	CV	Moda	Mediana	Coefficiente de Assimetria	Quantil 2,5%	Quantil 97,5%
n_c	0,023	0,001	0,024	0,023	0,023	0,157	0,022	0,024
n_p	0,135	0,042	0,315	0,136	0,140	-0,409	0,047	0,196
σ_0	0,096	0,012	0,128	0,105	0,096	0,153	0,073	0,118
σ_1	valor fixo = 0,00							
β	0,659	0,266	0,404	0,713	0,718	-0,764	0,055	0,990
ξ	3,842	2,867	0,746	1,419	2,000	0,750	1,126	9,518
ϕ_1	0,619	0,098	0,159	0,532	0,620	0,460	0,489	0,827

Tabela 5.12 – Descritores estatísticos dos parâmetros do modelo hidrodinâmico e da função de verossimilhança Generalizada: Cenário B3.

Parâmetro	Média	Desvio-Padrão	CV	Moda	Mediana	Coefficiente de Assimetria	Quantil 2,5%	Quantil 97,5%
n_c	0,023	0,001	0,034	0,022	0,023	0,268	0,022	0,024
n_p	0,133	0,046	0,345	0,197	0,139	-0,421	0,039	0,197
σ_0	0,089	0,011	0,122	0,080	0,088	0,494	0,070	0,113
σ_1	valor fixo = 0,00							
β	0,813	0,159	0,195	0,970	0,858	-1,253	0,414	0,995
ξ	1,330	0,192	0,144	1,103	1,308	0,849	1,040	1,780
ϕ_1	0,780	0,000	0,000	0,780	0,780	1,000	0,780	0,780

Um fato muito interessante ocorreu com as distribuições *a posteriori* de n_c , em ambos os cenários: a incerteza desse parâmetro ficou confinada na faixa entre 0,021 (ou 0,020, no Cenário B3) e 0,026, com concentração dos valores amostrados em 0,022 e 0,023, com mediana e média em 0,023 nos dois casos. Esses valores são consistentemente menores do que aqueles verificados *a posteriori* nos cenários A1 e A3, o que induziria à obtenção de menores profundidades e níveis de água em relação àqueles simulados nos cenários precedentes, sob a hipótese de ocorrência de descargas em trânsito semelhantes. Por outro lado, a distância em relação aos valores médios desse parâmetro dos cenários A1 e A3 pode estar atrelada a uma possível inconsistência entre as cotas altimétricas das seções topobatimétricas e aquela da origem dos lances de réguas no posto fluviométrico de Pirapora-Barreiro. De fato, tais informações são oriundas de levantamentos topobatimétricos distintos, feitos em épocas distintas e por entidades diferentes, o que pode acarretar diferenças de medição. Por fim, ainda que a origem de tais dados fosse a mesma, de acordo com o formato da seção transversal no

referido posto de monitoramento, e de sua curva-chave, é possível que a variabilidade da rugosidade na calha principal, calibrada a partir de níveis de água, fosse de fato menor do que aquela estimada caso vazões fossem utilizadas, visto que o coeficiente de Manning afeta de formas distintas ambas as variáveis. No entanto, as medidas de tendência central deveriam ser mais próximas, ao contrário do que ocorreu, conduzindo à hipótese de inconsistência entre as bases topográficas disponíveis no estudo.

Adicionalmente, a rugosidade de Manning na planície, n_p , mostrou tendência contrária àquela constatada nos cenários A1 e A3, ou seja, com maiores frequências relativas em direção ao limite superior pré-estipulado para este parâmetro e medidas de tendência central maiores do que as estimadas nos referidos cenários precedentes. Isso pode se dever ao fato de que houve concentração de capacidade de descarga na calha, induzida pelos baixos valores de Manning *a posteriori* e também pelos níveis de água registrados (que podem ter sido subestimados devido à diferença entre as bases topográficas das seções e do zero das réguas de Pirapora-Barreiro), restando pouco para ser conduzido pelas planícies, e justificando a tendência, embora modesta, de rugosidades maiores nessa porção do canal. Por outro lado, ao contrário do que ocorreu naqueles casos, não há correlação linear entre as rugosidades na calha principal e nas planícies, de modo que a toda a amplitude de n_c correspondem valores de n_p em toda sua faixa *a priori*, o que indica também que não houve a diferenciação *a posteriori* desse parâmetro como ocorreu nos cenários tipo A. A Figura 5.29 mostra a amplitude de níveis de água máximos estimados na seção de referência de resultados considerando-se os 10.000 pares $\{n_c, n_p\}$ obtidos *a posteriori* no Cenário B2. À semelhança dos cenários A1 e A3, observa-se que de fato há pouca informação na parte mais extensa da planície de inundação a ser usada no esquema Bayesiano para calibração de incertezas, embora a amplitude verificada ao se utilizarem variáveis distintas como $\tilde{Y}$ sejam diferentes. Aspectos comparativos entre a utilização de níveis de água e vazões como paradigma $\tilde{Y}$ serão percorridos em maior detalhe nos subitens 5.3.5.6 e 5.3.6.

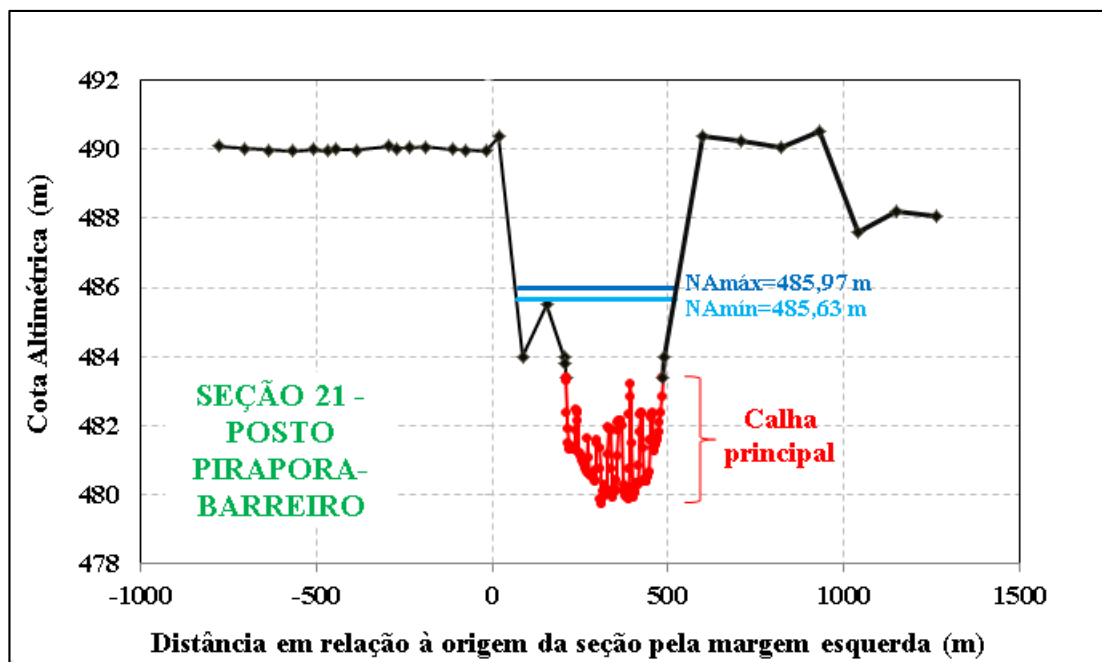


Figura 5.29 – Amplitude de níveis de água na seção de Pirapora-Barreiro resultantes da simulação hidrodinâmica dos 10.000 pares $\{n_c, n_p\}$ obtidos *a posteriori* no Cenário B2.

Finalmente, o Cenário B4 foi idealizado a fim de se verificar o que ocorreria com os parâmetros e variáveis latentes caso se adotassem as hipóteses clássicas quanto aos resíduos, as quais poderiam ser aplicáveis ao se alterar o tipo de informação usada como vetor $\tilde{Y}$.

Verificou-se, nesse caso, que os resíduos ε_t seguiram comportamento semelhante aos dos cenários B2 e B3, mas com menor variabilidade em termos de seus descritores estatísticos. Tal fato deveu-se provavelmente à acentuada identificação de n_c , que se restringiu exclusivamente aos valores 0,022 e 0,023, conduzindo a séries semelhantes de resíduos. Ademais, a exemplo do que ocorreu no cenário A4, os valores de σ_o foram “inflados” em relação aos seus valores obtidos *a posteriori* nos cenários precedentes.

Por outro lado, a ausência na modelagem de um ou mais parâmetros relacionados à dependência temporal entre diferenças sucessivas entre níveis de água observados e simulados, constatada como elevada no presente caso, gera intervalos de credibilidade preditivos para os níveis de água mais modestos do que no caso dos cenários B2 e B3 (resultados no subitem 5.3.6).

A identificação mais restritiva da variabilidade da rugosidade na calha e o estabelecimento de intervalos de credibilidade menos amplos podem parecer soluções atraentes para a calibração de incertezas pelo esquema Bayesiano. No entanto, a inspeção dos resíduos mostra que as

hipóteses inicialmente tomadas no cenário ora discorrido não se verificam, conforme constatado pela comparação entre quantis empíricos e teóricos dos resíduos $\varepsilon_t/\sigma_t = a_t$, de forma semelhante ao que ocorreu no Cenário A4. Estes mostraram média próxima de zero, desvio-padrão próximo de 1, mas assimetria e excesso de curtose positivos, ensejando que provavelmente a distribuição Normal padrão não se aplicaria adequadamente.

5.3.5.3 Cenários tipo C: hidrograma de jusante como paradigma de calibração e inserção de aflúências laterais como condição de contorno

Os resíduos predominantemente elevados, da ordem de algumas centenas de m³/s, e com correlação serial próxima de 1,0, suscitaram a necessidade de se avaliar a identificação de incertezas caso as aflúências das bacias incrementais ao trecho fluvial estudado fossem computadas no modelo hidrodinâmico. Sendo assim, à semelhança das situações anteriormente testadas, os cenários do tipo C partiram de hipóteses mais complexas acerca dos resíduos oriundos da modelagem, passando pela possibilidade de fixação *a priori* de ϕ_1 e de desconsideração de σ_1 , em direção às premissas de normalidade, independência e homoscedasticidade.

A análise dos resíduos ε_t no Cenário C1 revelou que a heteroscedasticidade não parece ser tão evidente, como havia ocorrido nos cenários do tipo A (vide Figura 5.30). Dessa forma, o Cenário C2 prestou-se à inferência de 6 variáveis e parâmetros, excluindo-se σ_1 . Em ambas as situações, constatou-se a tendência de os valores de ϕ_1 obtidos *a posteriori* superestimarem a dependência serial das diferenças entre vazões observadas e simuladas. Enquanto que os coeficientes de correlação amostral de 1ª ordem oscilavam entre 0,3 e 0,7 (valores aproximados a fim de representar simultaneamente os cenários C1 e C2), com média de 0,44 em ambos os cenários, os valores de ϕ_1 variavam na faixa entre 0,2 e 0,9, e média em torno de 0,66. Esse aspecto demandou o estabelecimento do Cenário C3, no qual se previu a fixação de ϕ_1 em seu valor médio estimado nos dois cenários anteriores, mencionado anteriormente, e o vetor de variáveis latentes igual a $\{\sigma_o, \beta, \xi\}$, a fim de se produzirem intervalos de credibilidade coerentes com a dependência serial das séries de resíduos obtidos.

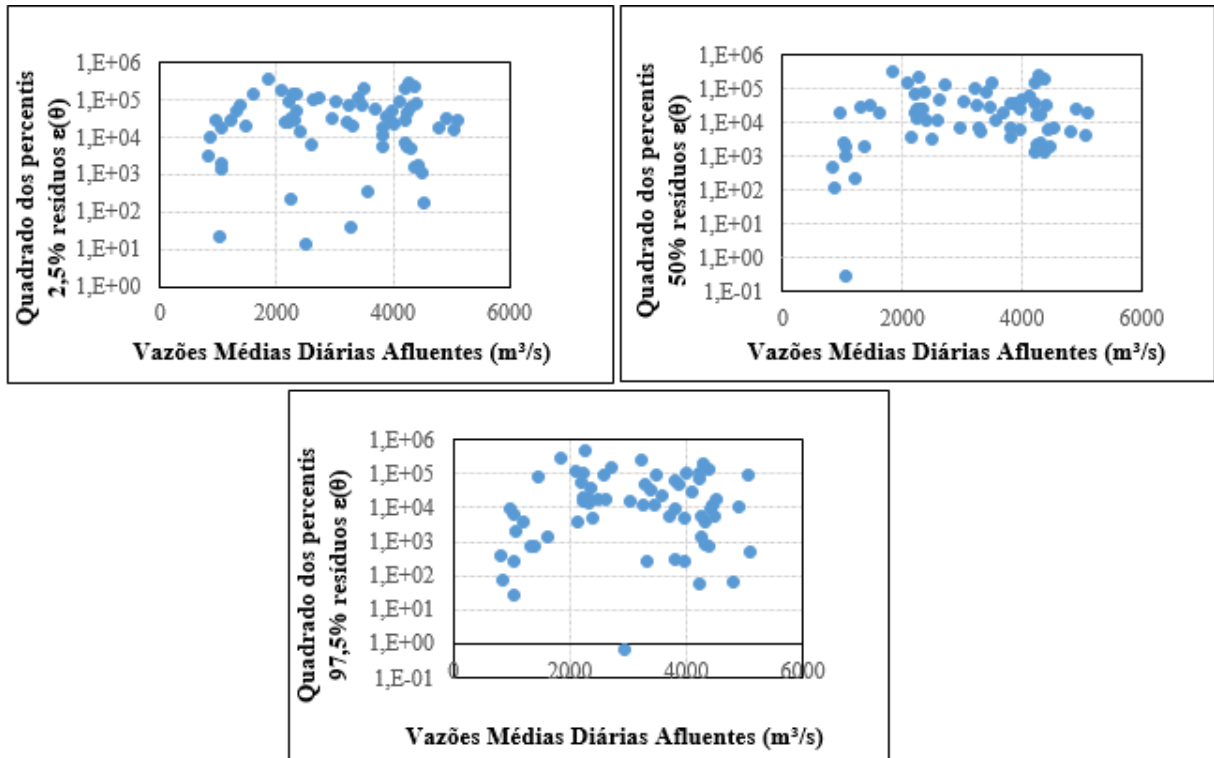


Figura 5.30 – Quadrado dos resíduos ε_t medianos e correspondentes aos percentis 2,5% e 97,5% *versus* as vazões afluentes ao trecho fluvial estudado: Cenário C1.

Conforme previsto ao longo das simulações dos cenários do tipo A, a adição e a propagação das afluições laterais ao hidrograma afluente principal poderiam de fato reduzir a dependência serial entre os resíduos das descargas observadas e simuladas, sobretudo no ramo ascendente da cheia de fevereiro de 1992, cujo início mostra-se bastante proeminente e para o qual parecem ter contribuído de maneira importante as bacias laterais, não regularizadas. Na Figura 5.31 são plotados os resíduos ε_t medianos por Δt calculados para os cenários A1 e C1 em relação ao hidrograma defluente monitorado no rio São Francisco em Pirapora-Barreiro.

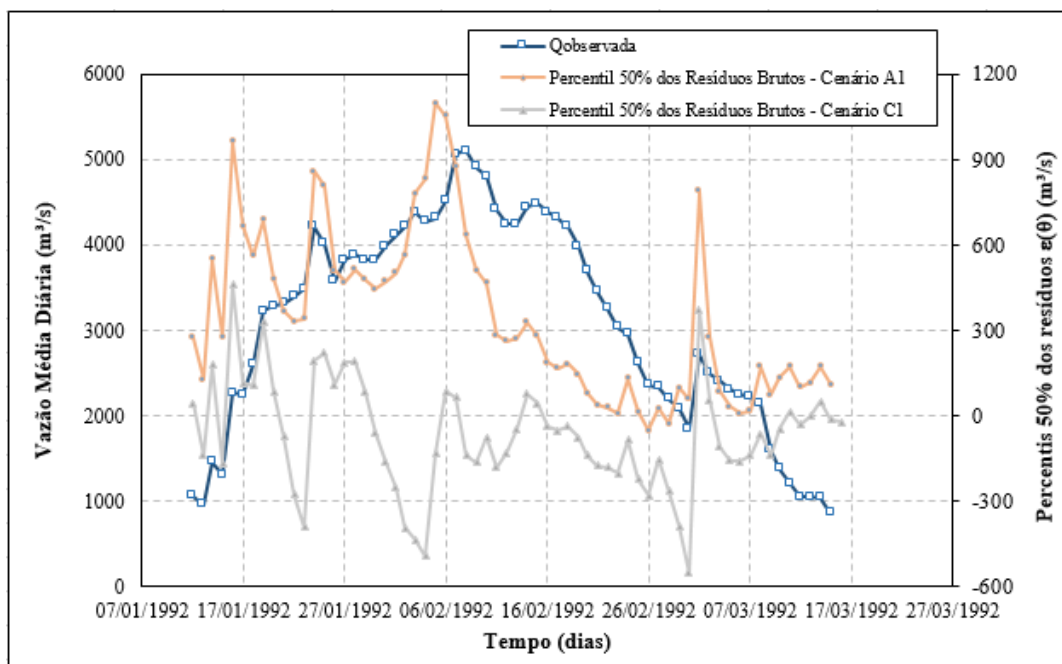


Figura 5.31 – Resíduos ε_t medianos por intervalo temporal obtidos nos Cenários A1 a C1 plotados junto ao hidrograma observado na estação fluviométrica de referência.

Observa-se que a redução dos resíduos de um tipo de cenário para o outro foi importante. Além disso, no Cenário C1, e de maneira geral, em todos os cenários do tipo C, os resíduos ε_t oscilam ao redor do zero, embora sua média tenha sido negativa, da ordem de algumas dezenas de m^3/s . Ademais, verifica-se em ambos os tipos de cenários que as diferenças maiores correspondem à 1ª parte da subida do hidrograma, bem como ao 1º, ao 2º e ao 3º picos registrados, indicados na Figura 5.31. Outro aspecto emanado pela referida figura é a prevalência de resíduos positivos nas porções inicial e final do hidrograma, bem como nos picos mais importantes, e negativos na porção média, de maiores vazões. Todas essas constatações podem estar atreladas a uma série de fatores, como: (i) incertezas nas estimativas das vazões de cheia nas bacias incrementais ao trecho, todas sem monitoramento linimétrico sistemático, sendo que tais descargas provavelmente têm maior relevância no ramo ascendente do hidrograma, enquanto que, em sua porção média, parece prevalecer o volume da cheia defluente da UHE Três Marias; (ii) discretização temporal inadequada para representar as cheias nas bacias menores, de algumas centenas de km^2 , nas quais os picos podem não ser capturados em termos de vazões médias diárias; (iii) discretização espacial adotada para compor as contribuições laterais; (iv) dificuldade de o modelo hidrodinâmico em reproduzir os efeitos naturais de translação e propagação da onda de cheia, seja devido ao esquema numérico disponível, ao espaçamento entre as seções transversais ou ao intervalo temporal adotado nas simulações, ainda que

selecionados segundo critérios consagrados; e, por fim, (v) diferenças entre a discretização temporal definida como adequada e suficiente à modelagem (igual a 3 horas) e aquela empregada no cômputo dos resultados, que precisa ser de 1 dia devido à disponibilidade das informações fluviométricas.

De maneira geral, os resíduos desse tipo de cenário tiveram as seguintes características:

- Resíduos ε_t : valores oscilando ao redor do zero, porém com média negativa, ao redor de -66,0 m³/s, já que o modelo hidrodinâmico tendeu a superestimar as vazões observadas; desvio-padrão menor do que nos cenários do tipo A, com médias desse descritor oscilando entre 180 m³/s e 200 m³/s, a depender do cenário; pequenos coeficientes de assimetria, próximos de zero, com médias variando entre 0,110 e 0,116 (cenários C1 e C2) e -0,075 e -0,171 (cenários C3 e C4);
- Resíduos após remoção da correlação serial: médias dessas séries continuam negativas, porém reduzidas a cerca de -20,0 m³/s (Cenários C1 e C2) e a -37,0 m³/s (Cenário C3). Nos Cenários C1 e C2, a exemplo de outros cenários, o coeficiente de correlação serial de 1ª ordem tornou-se negativo, embora sem significância estatística, com valores médios de -0,2. No Cenário C3, o valor mais modesto do parâmetro ϕ_1 acarretou menores valores residuais para este coeficiente, o qual apresentou média igual a -0,077. A assimetria e a curtose acentuaram-se, relativamente às séries de resíduos brutos, comportamento válido também para o Cenário C3, no qual os resíduos ε_t , praticamente simétricos, tornam-se simétricos positivamente ao serem transformados nesta etapa; e
- Resíduos a_t com média e desvio-padrão próximos de zero e da unidade, respectivamente. Esse aspecto também se verificou no Cenário A4, embora as médias desses descritores tenham se apresentado um pouco mais distantes dos parâmetros da distribuição Normal padronizada.

As Tabelas 5.13, 5.14 e 5.15 resumem os principais descritores das séries de resíduos brutos e transformados dos Cenários C2 (representativo também do Cenário C1, devido às diferenças pouco significativas entre ambos), C3 e C4, respectivamente.

Tabela 5.13 – Principais descritores estatísticos dos resíduos brutos e transformados: Cenário C2.

Tipo de resíduo avaliado	Média	Desvio-Padrão	Assimetria	Excesso Curtose	Coef. Corr. 1ª Ordem	Perc. 2,5%	Mediana	Perc. 97,5%
	(m³/s, se pertinente)			(adim.)		(m³/s, se pertinente)		
Resíduos ε_t (m³/s)	-66,7	201	0,116	0,655	0,438	-457,3	-69,7	352
Resíduos $\varepsilon_t - \phi_1 \cdot \varepsilon_{t-1}$ (m³/s)	-21,4	187	1,902	5,364	-0,221	-259	-48,88	515
Resíduos a_t (adim.)	-0,111	0,971	1,902	5,364	-0,221	-1,347	-0,254	2,675

Tabela 5.14 – Principais descritores estatísticos dos resíduos brutos e transformados: Cenário C3.

Tipo de resíduo avaliado	Média	Desvio-Padrão	Assimetria	Excesso Curtose	Coef. Corr. 1ª Ordem	Perc. 2,5%	Mediana	Perc. 97,5%
	(m³/s, se pertinente)			(adim.)		(m³/s, se pertinente)		
Resíduos ε_t (m³/s)	-66,3	188	-0,075	0,664	0,380	-453,0	-58,7	317
Resíduos $\varepsilon_t - \phi_1 \cdot \varepsilon_{t-1}$ (m³/s)	-37,3	174	1,235	3,073	-0,077	-293	-53,4	426
Resíduos a_t (adim.)	-0,196	0,912	1,235	3,073	-0,077	-1,542	-0,280	2,235

Tabela 5.15 – Principais descritores estatísticos dos resíduos brutos e transformados: Cenário C4.

Tipo de resíduo avaliado	Média	Desvio-Padrão	Assimetria	Excesso Curtose	Coef. Corr. 1ª Ordem	Perc. 2,5%	Mediana	Perc. 97,5%
	(m³/s, se pertinente)			(adim.)		(m³/s, se pertinente)		
Resíduos ε_t (m³/s)	-66,2	184	-0,171	0,768	0,353	-455	-52,1	319
Resíduos $\varepsilon_t - \phi_1 \cdot \varepsilon_{t-1}$ (m³/s)	-66,2	184	-0,171	0,768	0,353	-455	-52,1	319
Resíduos a_t (adim.)	-0,338	0,937	-0,171	0,768	0,353	-2,324	-0,266	1,631

No que concerne à aderência da SEP aos resíduos i.i.d., verificou-se comportamento semelhante aos obtidos nos cenários do tipo A, com ajuste visual aparentemente adequado, em se tratando dos Cenários C1 a C3. No caso do Cenário C4, manteve-se a tendência de a distribuição $N(0,1)$ não possibilitar o mesmo ajuste da SEP. No presente caso, constatou-se maior dificuldade em reproduzir o aspecto leptocúrtico na porção central das amostras formadas por quantis empíricos. A Figura 5.32 mostra o ajuste entre quantis empíricos e teóricos no caso do Cenário C3, por se tratar da situação em que a aderência pareceu, em termos conceituais e avaliada visualmente, de melhor qualidade.

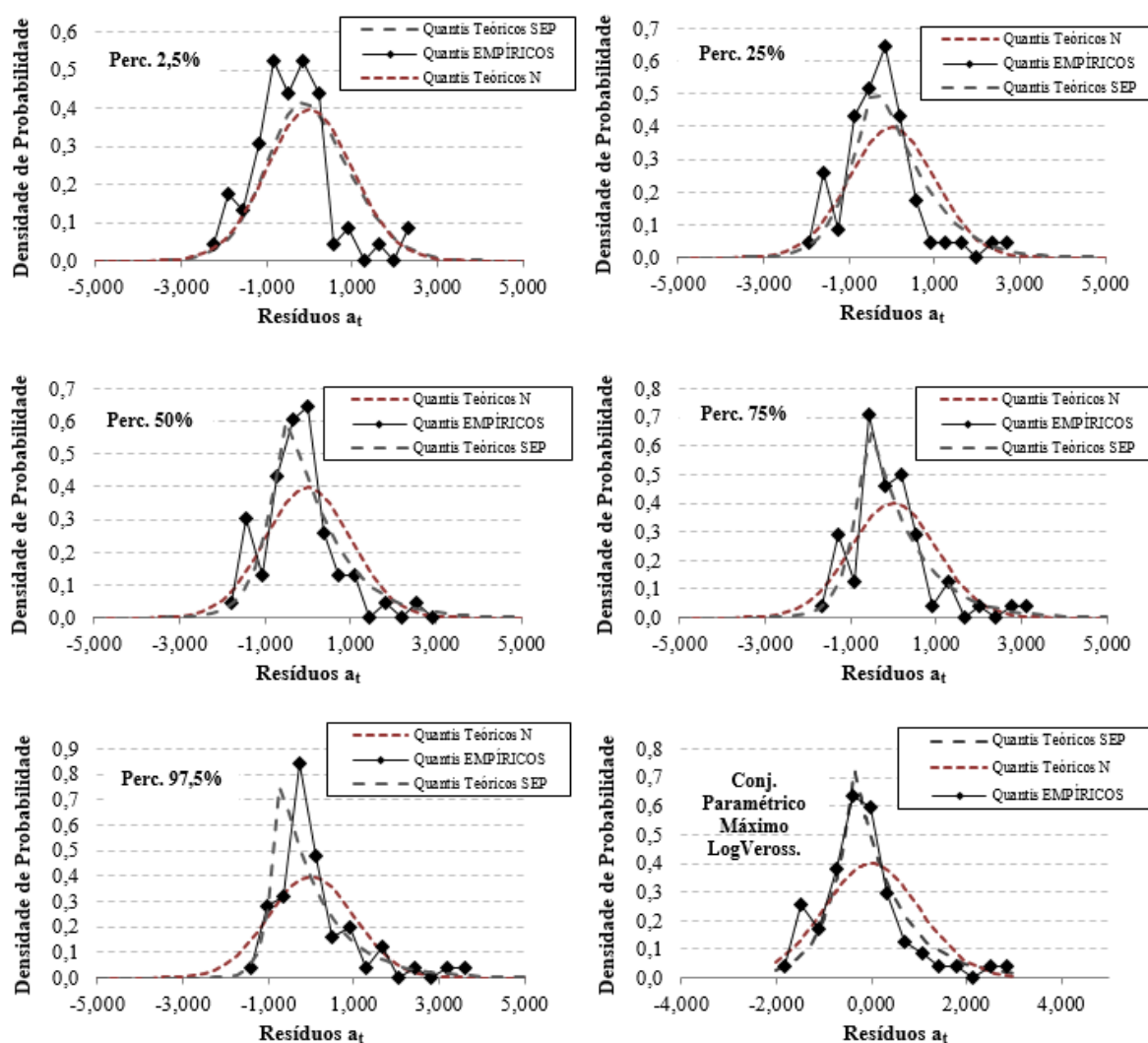


Figura 5.32 – Quantis empíricos de percentis notáveis colhidos por Δt nas séries de resíduos a_t e os correspondentes quantis teóricos segundo as distribuições SEP e N padronizadas: Cenário C3.

Finalmente, verificou-se que a identificação das incertezas das variáveis latentes e parâmetros do modelo hidrodinâmico não foi sensível às alterações conceituais realizadas entre os cenários em busca da inferência mais robusta e estatisticamente correta perante a mudança das condições de contorno da modelagem, isto é, adicionando-se as contribuições incrementais. A exceção ocorreu no Cenário C4, que, à semelhança dos cenários A4 e B4, baseia-se em hipóteses estatísticas não verificadas para os resíduos. As figuras 5.33 e 5.34 mostram o aspecto dos histogramas construídos a partir das 10.000 amostras do vetor $\{\theta, \eta_e\}$ obtidas *a posteriori* para os Cenários C3 - este considerado como o mais representativo dentre os de tipo C - e C4. Os principais descritores estatísticos extraídos dessas amostras para ambos os cenários estão sintetizados nas tabelas 5.16 e 5.17.

Tabela 5.16 – Descritores estatísticos dos parâmetros do modelo hidrodinâmico e da função de verossimilhança Generalizada: Cenário C3.

Parâmetro	Média	Desvio-Padrão	CV	Moda	Mediana	Quantil 2,5%	Quantil 97,5%
n_c	0,040	0,013	0,332	0,037	0,038	0,029	0,062
n_p	0,099	0,053	0,529	0,020	0,095	0,023	0,195
σ_0	192	20	0,105	181	191	157	237
σ_1	valor fixo = 0,00						
β	0,617	0,251	0,406	0,820	0,655	0,077	0,980
ξ	1,487	0,235	0,158	1,162	1,464	1,127	1,997
ϕ_1	valor fixo = 0,44						

Tabela 5.17 – Descritores estatísticos dos parâmetros do modelo hidrodinâmico e da função de verossimilhança Generalizada: Cenário C4.

Parâmetro	Média	Desvio-Padrão	CV	Moda	Mediana	Coefficiente de Assimetria	Quantil 2,5%	Quantil 97,5%
n_c	0,033	0,004	0,130	0,035	0,033	0,909	0,026	0,043
n_p	0,115	0,050	0,438	0,021	0,116	-0,088	0,027	0,195
σ_0	197	18	0,092	182	196	0,490	166	238
σ_1	valor fixo = 0,00							
β	valor fixo = 0,00							
ξ	valor fixo = 1,00							
ϕ_1	valor fixo = 0,00							

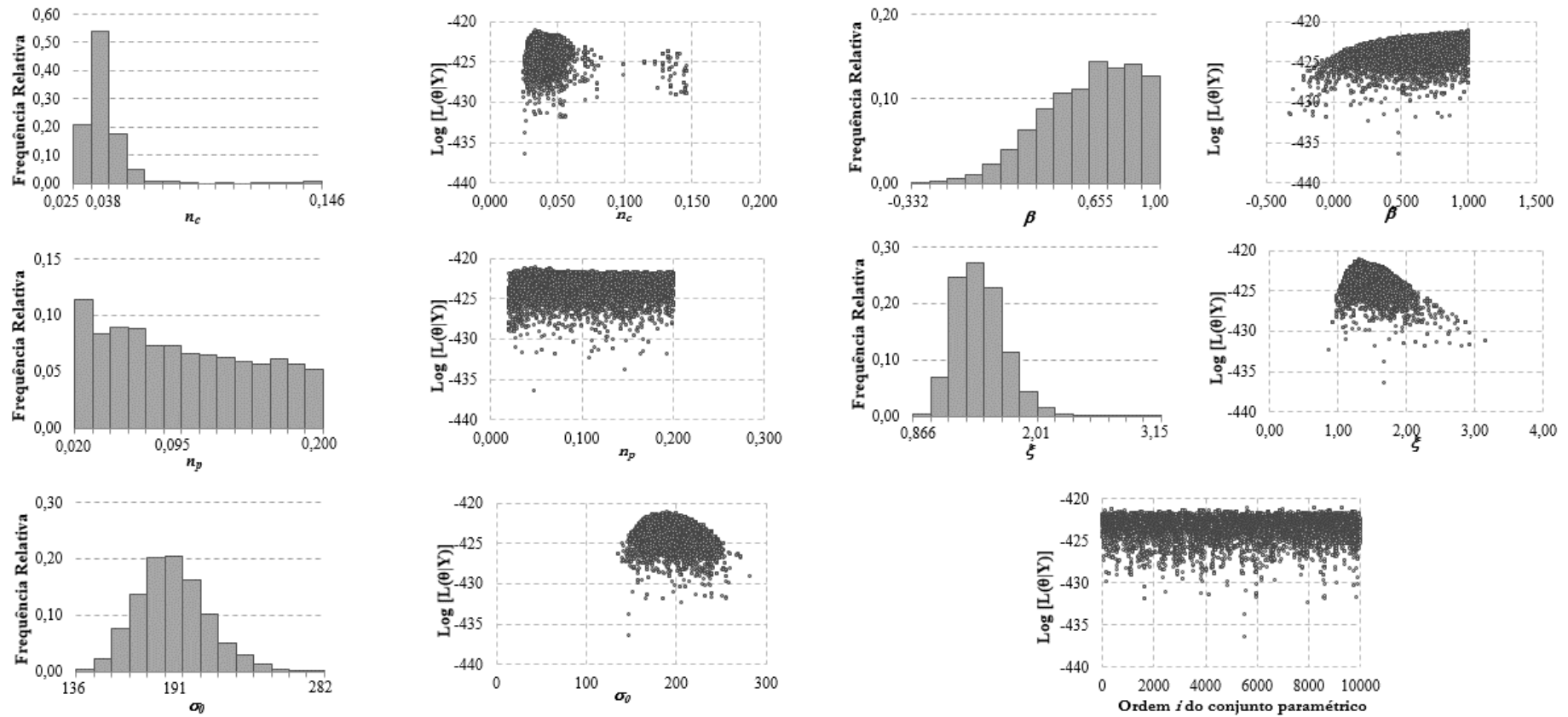


Figura 5.33 – Histogramas de frequências relativas estimados para as amostras *a posteriori* dos parâmetros do modelo hidrodinâmico e da verossimilhança Generalizada: Cenário C3.

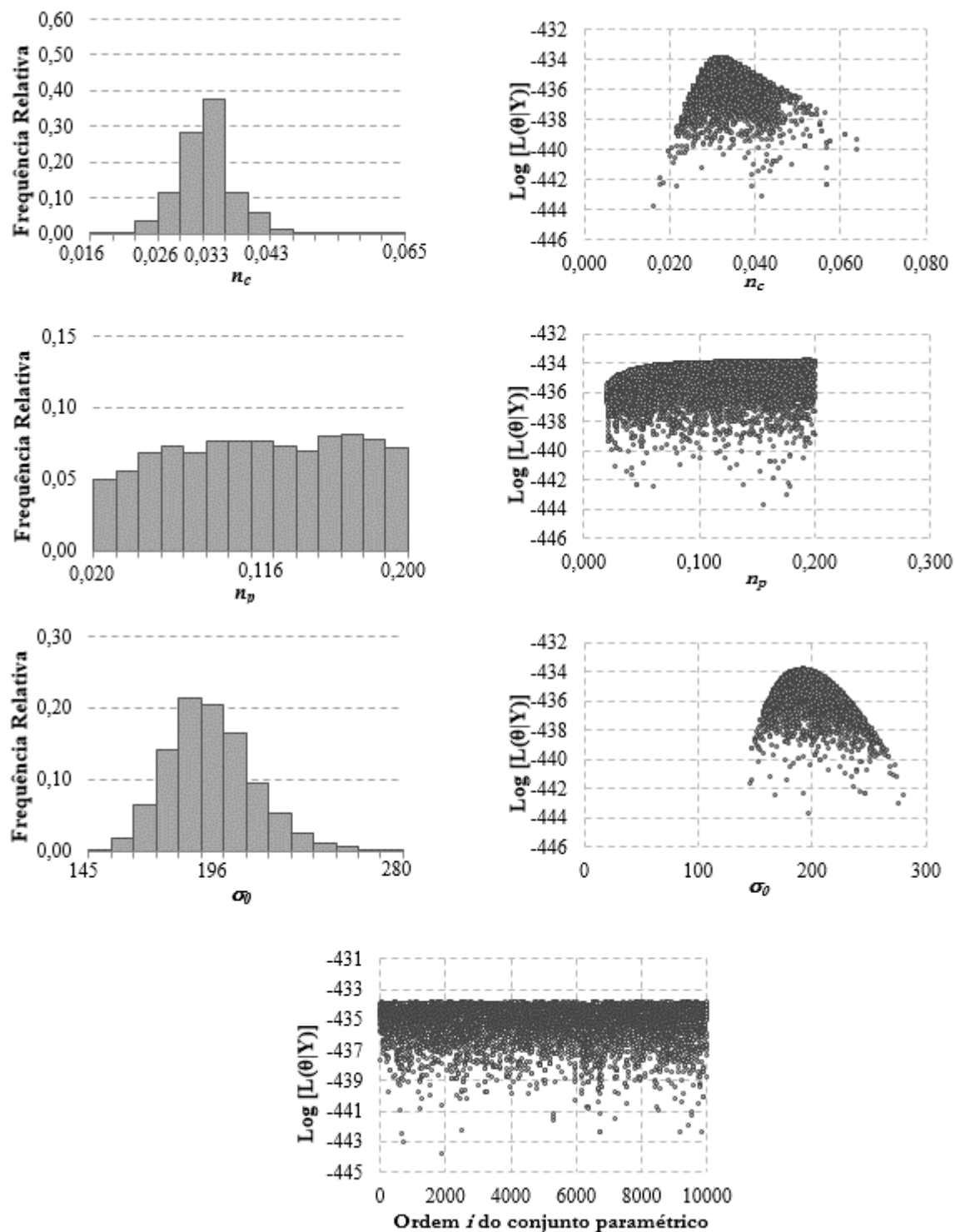


Figura 5.34 – Histogramas de frequências relativas estimados para as amostras *a posteriori* dos parâmetros do modelo hidrodinâmico e da verossimilhança Generalizada: Cenário C4.

Verifica-se que no Cenário C3 o parâmetro n_c desenvolveu-se ao longo de praticamente toda sua faixa estipulada *a priori*, porém concentrou 95% de seus valores no intervalo [0,029; 0,062], com mediana de 0,038 e média de 0,040. Esse aspecto, assim como a forma do histograma, assemelha-se àquele obtido para esse parâmetro no Cenário A3, tido como representativo daquele tipo de cenário. Tais constatações revelam a robustez da inferência Bayesiana apoiada sobre a escolha adequada da função de verossimilhança destinada a quantificar a incerteza paramétrica de uma modelagem hidrodinâmica unidimensional. Em outras palavras, a variabilidade, ou incerteza, inferida para os coeficientes de Manning na calha principal foi pouco sensível às alterações nas condições de contorno na presente situação, ou seja, adicionando-se ou não as afluições laterais. É importante ressaltar que o acréscimo de vazões ao trecho fluvial foi de cerca de 13%, em média, o que pode ter contribuído para essa constatação. De todo modo, tais achados são importantes na medida em que correspondem a resultados obtidos em outros estudos. Domeneghetti *et al.* (2012), por exemplo, verificaram que as rugosidades calibradas deterministicamente na calha principal do rio Pó, na Itália, apresentaram pouca variação para vazões próximas e maiores do que aquela correspondente ao débito de margens plenas, a partir da qual ocorrem os transbordamentos. Constatação semelhante pode ser depreendida das estimativas da rugosidade da calha principal na compilação de Hicks e Mason (1998), para trechos fluviais com vazões anuais médias de magnitude semelhante às do rio São Francisco na porção estudada.

Por outro lado, neste cenário, o comportamento obtido *a posteriori* para n_p foi distinto daquele observado no Cenário A3. Na presente situação, a identificação desse parâmetro ficou menos evidente, de forma que toda a faixa *a priori* de n_p varreu os valores menores de n_c inferidos *a posteriori*. Provavelmente, o transporte das vazões totais, contabilizadas as contribuições laterais, foi garantido principalmente pela capacidade de descarga da calha principal, fixadas as rugosidades na calha principal, restando pouca informação disponível que permitisse a diferenciação para as planícies de inundação,

Quanto às variáveis latentes, os valores obtidos *a posteriori* foram condizentes com o comportamento estatístico dos resíduos brutos, ε_t , e de a_t , nos Cenários C1 a C3, representando adequadamente o desvio-padrão dos primeiros (mesmo quando este foi assumido como variável, no Cenário C1, o que não foi comprovado), e a assimetria e a curtose positivas dos segundos. Além disso, sua inferência manteve-se robusta em face das alterações realizadas entre tais cenários.

Diferenças importantes quanto à identificação da rugosidade de Manning na calha ocorreram no Cenário C4 em relação aos três anteriores. Novamente, como no Cenário A4, este parâmetro teve sua faixa *a posteriori* com valores menores do que nos casos antecedentes, o que se reflete na média e nos percentis 2,5% e 97,5% (vide Tabela 5.17). De toda maneira, a provável subestimação de n_c em relação aos cenários C1 a C3 não foi tão acentuada como ocorreu nos cenários de tipo A. Finalmente, o parâmetro n_p sofreu pouca diferenciação em relação ao comportamento previamente estipulado, com leve tendência de acumular seus valores amostrados em direção ao seu limite superior. O desvio-padrão, por sua vez, apresentou formato e valores *a posteriori* semelhantes aos dos 3 cenários anteriores.

5.3.5.4 Cenários tipo D: cotograma de jusante como paradigma de calibração e inserção de aflúências laterais como condição de contorno

A sequência natural de realização dos cenários desse tipo foi análoga àquela utilizada nos demais cenários, partindo-se do princípio de que 7 variáveis latentes deveriam ter suas incertezas inferidas conjuntamente, e analisando-se a relevância e o impacto da manutenção/supressão e fixação de algumas delas.

Além disso, em função dos resultados encontrados nos cenários do tipo B, e da propagação de vazões incrementais, havia a expectativa de obtenção de rugosidades na calha principal ainda menores do que as estimadas naquelas condições.

De fato, no presente tipo de cenário, observou-se uma tendência geral em todos os casos de estimação de coeficientes de Mannings reduzidos, com médias ao redor de 0,020 e 0,021, o que acarretou reflexos sobre os descritores estatísticos das séries de resíduos brutos e transformados em relação àqueles dos cenários B.

No Cenário D1, constatou-se, a exemplo de outros cenários, que o parâmetro ϕ_1 tendeu a superestimar os valores dos coeficientes de correlação serial de 1ª ordem estimados para as séries formadas pelas diferenças sucessivas entre níveis de água observados e simulados. Naturalmente, isso acarretou limites pouco realistas para o intervalo de credibilidade preditivo para cotogramas no extremo de jusante do trecho estudado. Além disso, verificou-se maior variabilidade de n_c , comparativamente aos cenários homólogos (i.e., que se valeram de cotograma como paradigma de calibração), o que resultou em ampla incerteza preditiva devido somente à variação *a posteriori* desse parâmetro e de n_p .

Por tais razões, estipulou-se o Cenário D2 para avaliar se a desconsideração de σ_1 do processo de inferência poderia melhorar a identificação de n_c e de ϕ_1 , embora tenha se observado que os resíduos da modelagem sob as condições desse tipo de cenário pareciam apresentar a variância dependente da magnitude dos níveis de água simulados. No entanto, manteve-se a tendência do cenário anterior no que concerne a esses aspectos, de modo que se julgou importante considerar o Cenário D3 para inferência das incertezas do vetor $\{\theta, \eta_\varepsilon\} = \{n_c, n_p, \phi_1, \sigma_o, \sigma_1, \beta, \xi\}$. O Cenário D4, de forma parecida aos tipos de cenários anteriormente desenvolvidos, foi utilizado para inferência somente de 3 variáveis e/ou parâmetros, ou seja, $\{\theta, \eta_\varepsilon\} = \{n_c, n_p, \sigma_o\}$.

Conclui-se, nos Cenários D1 a D3, que os resíduos apresentavam as seguintes características, as quais podem ser representadas pelos resultados obtidos no Cenário D3 (vide Tabela 5.18):

- Resíduos ε_t denotam tendência da modelagem em superar os níveis de água observados na seção transversal de controle, embora pouco significativa, com médias da ordem de alguns centímetros e desvio-padrão ao redor de 0,13 m. Séries com assimetria positiva pouco acentuada e mesocúrticas, ou seja, com coeficiente de excesso de curtose negativo. Correlação serial importante, oscilando entre 0,4 e 0,9;
- Resíduos $\varepsilon_t - \phi_1 \cdot \varepsilon_{t-1}$ apresentam dependência serial residual, de valor negativo, embora sem valor estatístico. Removida a dependência serial, as séries de resíduos têm sua assimetria acentuada e tornam-se leptocúrticas; e
- Resíduos a_t apresentam média próxima de zero e desvio-padrão unitário, mantendo a assimetria e o excesso de curtose positivos.

Tabela 5.18 – Principais descritores estatísticos dos resíduos brutos e transformados: Cenário D3.

Tipo de resíduo avaliado	Média	Desvio-Padrão	Assimetria	Excesso Curtose	Coef. Corr. 1ª Ordem	Perc. 2,5%	Mediana	Perc. 97,5%
	(m³/s, se pertinente)		(adim.)			(m³/s, se pertinente)		
Resíduos ε_t (m³/s)	-0,008	0,127	0,106	-0,537	0,649	-0,239	-0,005	0,217
Resíduos $\varepsilon_t - \phi_1 \cdot \varepsilon_{t-1}$ (m³/s)	-0,001	0,097	1,158	2,191	-0,253	-0,137	-0,012	0,248
Resíduos a_t (adim.)	-0,007	0,977	1,158	2,193	-0,254	-1,374	-0,116	2,507

A aderência da distribuição SEP padronizada aos resíduos a_t , embora não pareça ter a qualidade visualizada nos cenários B1 a B3, mostrou-se adequada a reproduzir as caudas das amostras empíricas e seu formato leptocúrtico, conforme pode ser conferido na Figura 5.35.

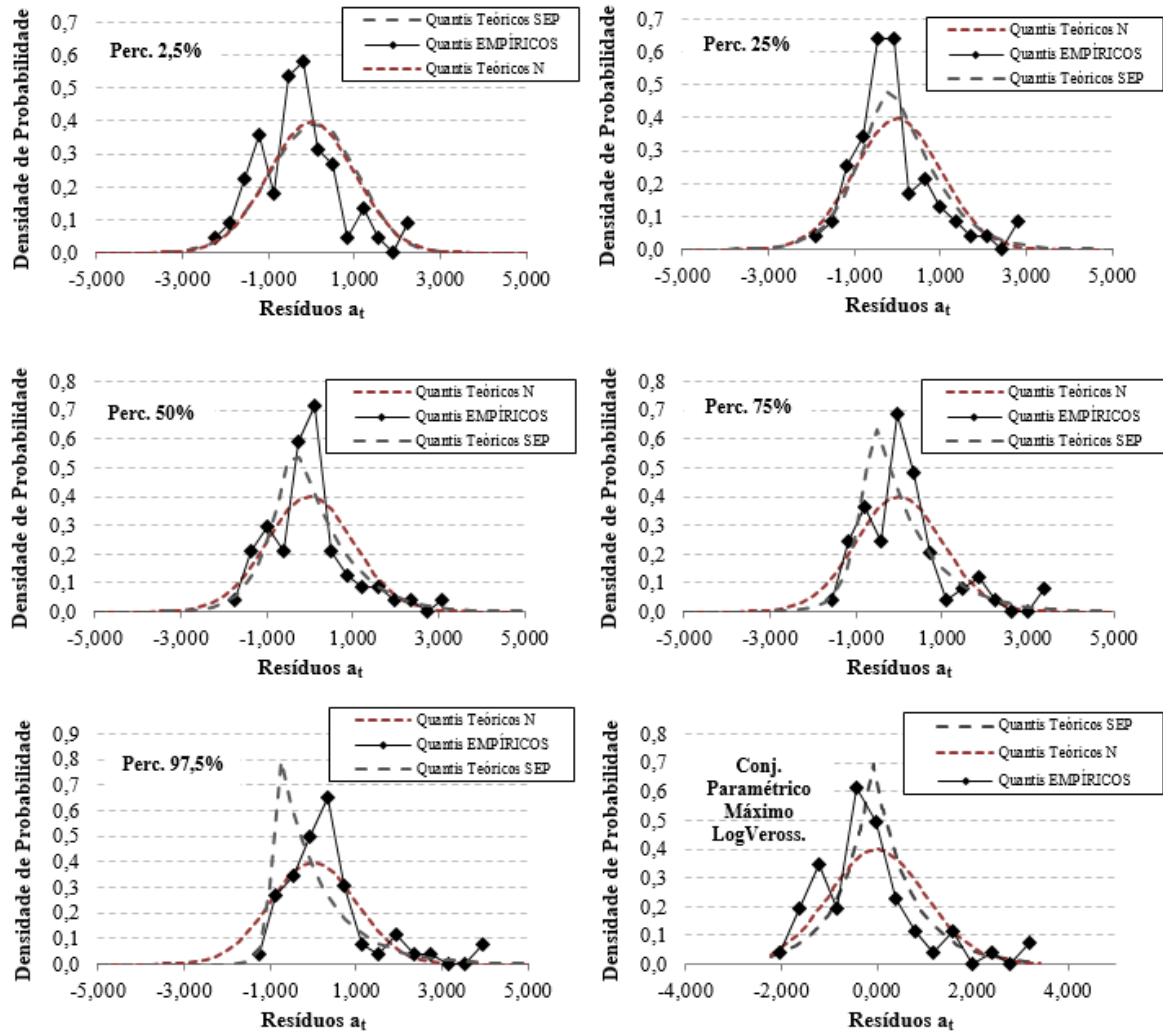


Figura 5.35 – Quantis empíricos de percentis notáveis colhidos por Δt nas séries de resíduos a_t e os correspondentes quantis teóricos segundo as distribuições SEP e N padronizadas: Cenário D3.

Um aspecto importante desses cenários refere-se à tendência geral em todos eles em subestimar os valores de n_c obtidos *a posteriori*, quando comparados àqueles estimados nos cenários B. Nestes, as médias das amostras de n_c variaram entre 0,022 e 0,023. Presentemente, tais medidas de tendência central oscilam entre 0,020 e 0,021, a depender do cenário. De fato, como se cogitou que haja alguma inconsistência entre os levantamentos que originaram as seções transversais usadas na modelagem e aquele que forneceu o zero das régua da estação fluviométrica de referência, era esperado que haveria uma compensação sobre as rugosidades

na calha principal, com sua redução em relação aos cenários em que as contribuições laterais não foram propagadas pelo modelo. De todo modo, a redução averiguada é coerente ao se considerarem as afluições laterais, premissa que parece mais correta do ponto de vista físico do sistema em estudo e da modelagem propriamente. Os principais descritores estatísticos das amostras construídas *a posteriori*, válidas para o Cenário D3, e representativas dos dois cenários anteriores, encontram-se sintetizadas na Tabela 5.19.

Tabela 5.19 – Descritores estatísticos dos parâmetros do modelo hidrodinâmico e da função de verossimilhança Generalizada: Cenário D3.

Parâmetro	Média	Desvio-Padrão	CV	Moda	Mediana	Coefficiente de Assimetria	Quantil 2,5%	Quantil 97,5%
n_c	0,020	0,0006	0,032	0,020	0,020	0,377	0,019	0,021
n_p	0,127	0,047	0,367	0,195	0,132	-0,298	0,036	0,197
σ_0	0,051	0,029	0,577	0,006	0,051	0,154	0,003	0,106
σ_1	0,0001	0,0001	0,5957	0,0002	0,0001	0,209	0,0000	0,0002
β	0,572	0,290	0,507	0,874	0,621	-0,537	-0,052	0,978
ξ	1,448	0,413	0,285	1,132	1,331	1,387	0,927	2,492
ϕ_1	valor fixo = 0,69							

Verifica-se, por outro lado, a mesma tendência dos cenários do tipo B em pouco diferenciar os valores de n_p , com um acúmulo, embora de pequena importância, dos valores amostrais próximos ao limite superior estipulado *a priori*. A identificação das incertezas das variáveis latentes variou pouco entre os cenários D1 a D3, e seus valores foram condizentes com as características emanadas a partir dos descritores estatísticos dos resíduos e do formato das distribuições de quantis empíricos associados a percentis notáveis.

Por fim, o Cenário D4 apontou, como nos tipos de cenários anteriores, uma identificação distinta das incertezas. No presente caso, tal aspecto ficou mais evidenciado para n_c , cuja amostra obtida *a posteriori* concentrou-se em um único valor, igual a 0,020. Trata-se, portanto, do único caso em que isso ocorreu no presente estudo. A análise dos resíduos mostrou, no entanto, que as hipóteses de homoscedasticidade e independência não se verificam, e tampouco a normalidade dos resíduos transformados.

5.3.5.5 Consolidação dos cenários avaliados

Os subitens anteriores detalharam os procedimentos metodológicos aplicados à definição e à análise dos cenários de inferência Bayesiana, buscando-se aferir a sensibilidade da identificação de incertezas às alterações realizadas entre os cenários e estabelecer diretrizes para obtenção de uma quantificação robusta.

Percebeu-se que uma das questões críticas da inferência voltada à modelagem hidrodinâmica residiu na identificação do parâmetro ϕ_1 , que na maioria dos cenários superestimou o grau de correlação serial entre os resíduos brutos. Embora a convergência tenha sido alcançada por essa variável latente, e pelos demais parâmetros e variáveis envolvidos em todos os casos, notou-se que a inferência das incertezas de ϕ_1 geralmente esteve atrelada a menores taxas de aceitação do método DREAM e à dificuldade de identificação das incertezas de outras variáveis latentes. Tais aspectos podem ter decorrido em função do elevado número de parâmetros a ser inferido, relativamente ao tamanho n do vetor de dados observados – fator que pode ter tido maior influência sobre a definição da dependência serial entre os resíduos.

A maneira encontrada para contornar esse impasse foi fixar o valor de ϕ_1 a partir de estimativas do coeficiente de correlação serial de 1ª ordem de sérios de resíduos ε_t obtidas em cenários anteriores, sob a hipótese de que um modelo autoregressivo de 1ª ordem seria suficiente para exprimir a dependência temporal entre os mesmos. Essa decisão pareceu adequada, visto que garantiu uma melhor identificação das incertezas *a posteriori* nos casos em que havia esse tipo de problema, estreitando o intervalo de credibilidade da incerteza preditiva de hidrogramas e cotogramas e acelerando a convergência do método MCMC utilizado.

A fixação *a priori* de ϕ_1 , a aderência da SEP aos resíduos a_t , a identificação de parâmetros e variáveis coerentes com descritores estatísticos dos resíduos e com as características físicas do canal no trecho estudado assim como a quantificação de limites plausíveis para a incerteza preditiva conduziram à seleção dos seguintes cenários como guias para estabelecimento de diretrizes e conclusões no sequenciamento desta pesquisa:

- Cenário A3, com $\{\theta, \eta_\varepsilon\} = \{n_c, n_p, \sigma_o, \sigma_1, \beta, \xi\}$ e $\phi_1 = 0,77$;

- Cenário B2, com $\{\theta, \eta_\epsilon\} = \{n_c, n_p, \phi_1, \sigma_o, \beta, \xi\}$. Nesse caso, embora a amostra *a posteriori* de ϕ_1 tenha ampliado a variabilidade dos coeficientes de correlação serial de 1ª ordem dos resíduos ϵ_t , a média inferida foi mais baixa do que a média daquele estimador e os extremos de sua distribuição *a posteriori* apresentaram frequências relativas muito baixas, de forma que os valores correspondentes tiveram pouco impacto na composição do intervalo de credibilidade da incerteza preditiva;
- Cenário C3, com $\{\theta, \eta_\epsilon\} = \{n_c, n_p, \sigma_o, \beta, \xi\}$ e $\phi_1 = 0,44$; e
- Cenário D3, com $\{\theta, \eta_\epsilon\} = \{n_c, n_p, \sigma_o, \sigma_1, \beta, \xi\}$ e $\phi_1 = 0,69$.

5.3.5.6 Inspeção da incerteza do coeficiente de rugosidade de Manning na calha principal

A análise das amostras dos parâmetros n_c e n_p obtidas *a posteriori* permitiu concluir que, de maneira geral, as incertezas do primeiro foram adequadamente identificadas, enquanto que as do segundo mantiveram seus limites e frequências praticamente iguais às da distribuição uniforme prevista inicialmente. Verificou-se também que, nos cenários onde houve alguma diferenciação *a posteriori* de n_p , este operou mais como um elemento matemático de compensação da condutância estimada na calha principal do que como um parâmetro ao qual possa ser atribuído algum significado físico, de forma semelhante ao que pode ocorrer em uma calibração feita de maneira determinística.

Tais constatações levam a crer que seria necessário utilizar uma cheia de maior período de retorno, que houvesse causado transbordamentos mais importantes, a fim de garantir que mais informações sobre a planície de inundação pudessem ser utilizadas para calibrar a incerteza de n_p . Sob a condição de não haver registro de cheias que se prestassem a esse fim, estimar um único valor ou faixa mais estreita do que a presentemente usada para n_p , valendo-se de comparações de fotos e imagens de satélite e dados coletados em campo com valores da literatura consagrada no tema, parece ser ainda o procedimento mais adequado.

Por outro lado, as incertezas de n_c inferidas *a posteriori* mostraram-se suficientemente robustas e consistentes quando o hidrograma de fevereiro de 1992 foi utilizado como paradigma de calibração, permitindo que algumas diretrizes possam ser definidas a partir das distribuições estimadas por meio dos Cenários A3 e C3, a fim de nortear possíveis avaliações futuras quanto à incerteza preditiva de ordem paramétrica em rios naturais de porte e material constituinte

semelhantes aos do rio São Francisco na porção considerada neste trabalho, considerando-se cheias de magnitude semelhante. Alguns comentários fazem-se relevantes:

- Os descritores estatísticos em ambos os casos são muito semelhantes entre si, apresentando variação máxima de $\pm 0,002$;
- Os percentis 2,5% e 97,5% (iguais a 0,027 e 0,029, e 0,061 e 0,062, respectivamente, para os cenários A3 e C3 correspondem a limites mínimos e máximos para a rugosidade de Manning indicados na literatura para canais naturais de médio de grande porte, a exemplo de Chow (1959, p. 113) e BRUNNER (2016a, p. 112). Em um experimento realizado por CPRM (2016, comunicação eletrônica), em que este parâmetro foi avaliado de forma independente por um grupo de 25 consultores e especialista e de 37 estudantes de engenharia civil em 8 trechos fluviais em bacias com centenas a milhares de km², verificou-se que os valores mínimos e máximos são semelhantes aos obtidos no presente procedimento de inferência de incertezas;
- Os valores mínimos em ambas as amostras *a posteriori* (0,023 e 0,025, respectivamente, para os cenários A3 e C3) coincidem com os valores de base para aplicação do método de Cowan indicados em Arcement e Scheneider (1989, p. 4) para leitos fluviais móveis, com material arenoso predominante, e diâmetros medianos dos grãos a partir de 0,6 mm e 0,8 mm, respectivamente;
- Em adição, os valores médios das amostras *a posteriori*, iguais a 0,039 e 0,040 (cenário A3 e C3, respectivamente), correspondem aos valores médios definidos nas amostras do experimento supracitado, especialmente naquelas correspondentes aos rios de maior porte e com material constituinte do leito semelhante àquele do rio São Francisco no trecho estudado. Além disso, esses mesmos valores são encontrados na referência de Hicks e Mason (1998, p. 74, 148, 160) em canais naturais com vazões anuais médias e materiais de fundo semelhantes aos do presente caso;
- Os valores das amostras obtidas *a posteriori* que são maiores do que o percentil 97,5%, na faixa entre 0,061 (ou 0,062) e 0,146, correspondem a limites máximos indicados na literatura. No entanto, são válidos para canais com seções muito irregulares (CHOW, 1959), e com presença importante de vegetação na calha principal (BRUNNER, 2016a), o que não é o caso do trecho fluvial estudado;

- Ambas as amostras *a posteriori* ora consideradas como referenciais apresentaram assimetria positiva, de forma semelhante ao encontrado no experimento de CPRM (2016, comunicação eletrônica).

A fim de consolidar o conhecimento adquirido *a posteriori* para n_c , foi avaliado o ajuste de algumas distribuições de probabilidade de variáveis contínuas sobre as amostras obtidas, por meio da utilização dos testes de hipóteses do Qui-Quadrado (NAGHETTINI e PINTO, 2007), de Kolmogorov-Smirnov (CRUTCHER, 1975), de Anderson-Darling (STEPHENS, 1979) e de Filliben (NAGHETTINI e PINTO, 2007). Para ambas as amostras, nenhuma das distribuições teóricas elencadas previamente apresentou ajuste adequado à luz dos quatro testes de aderência aplicados, a um nível de significância de 5%, embora em termos visuais algumas delas pareciam apropriadas a modelar a incerteza do referido parâmetro. Inspeccionando-se possíveis causas para esse fato, verificou-se que às estatísticas de teste correspondem valores muito baixos, geralmente de caráter assintótico para amostras com n elevado (no presente caso, $n = 10.000$), que facilmente seriam ultrapassados pelas estatísticas amostrais.

Sendo assim, a seleção de uma distribuição teórica dentre as candidatas valeu-se de critérios visuais e conceituais. Nesse sentido, estipulou-se a distribuição Gama para modelar a incerteza aleatória e epistêmica associada à rugosidade de Manning na calha principal do rio São Francisco em seu trecho entre a UHE Três Marias e a cidade de Pirapora. Além de ter se ajustado bem às duas amostras de referência, conforme se pode conferir pelas figuras a seguir, esta distribuição tem seus dois parâmetros, denotados por χ (escala) e ω (forma), facilmente estimáveis a partir dos momentos amostrais de 1ª e 2ª ordem (vide, por exemplo, NAGHETTINI e PINTO, 2007, p. 231), tendo ainda a vantagem de ser bastante flexível a variáveis aleatórias contínuas e positivas.

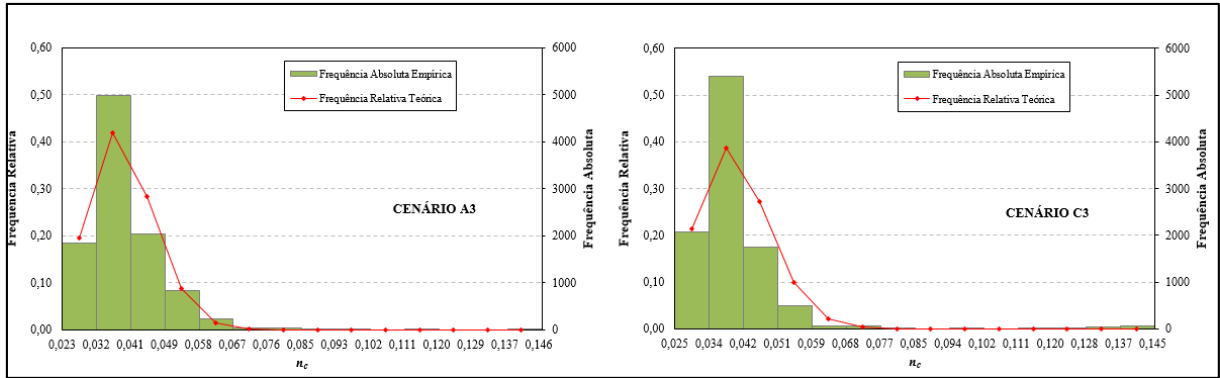


Figura 5.36 – Ajuste da distribuição Gama às amostras do parâmetro n_c inferidas nos Cenários A3 (esq.) e C3 (dir.).

Assim, caso se deseje avaliar a incerteza preditiva associada ao coeficiente de rugosidade de Manning na calha principal de um trecho fluvial que guarde semelhanças quanto ao porte e ao material constituinte do leito em relação ao estirão estudado, sem que tenha sido realizada uma calibração por esquema Bayesiano com dados locais, é possível estimar uma distribuição de probabilidades que explique a incerteza associada ao referido parâmetro, a partir do seguinte roteiro:

- Definir o valor médio para a rugosidade de Manning na calha, denotado por $\bar{n}_c$, a partir dos valores de base disponíveis em referências como Chow (1959), Hicks e Mason (1998), e Brunner (2016a), considerando-se características do material constituinte do leito, a existência de vegetação nas margens e no fundo, e seu tipo, e o grau de irregularidades nas seções e diferenças quanto à forma entre as seções que compõem o trecho em análise;
- Estimar o desvio-padrão para $\bar{n}_c$, considerando-se um coeficiente de variação, CV, entre 30% e 50%. Valores desse estimador próximos ao limite superior indicado representam maior incerteza quanto à estimativa da rugosidade. O desvio-padrão pode, então, ser quantificado, a partir da fórmula $s_{n_c} = CV \cdot \bar{n}_c$ (Equação 5.1), sendo $CV \sim [0,3; 0,5]$.
- Quantificar os parâmetros da distribuição Gama:

$$\hat{\chi} = \frac{s_{n_c}^2}{\bar{n}_c} \text{ e } \hat{\omega} = \frac{\bar{n}_c^2}{s_{n_c}^2} \quad (5.2 \text{ e } 5.3)$$

A partir da função inversa da distribuição Gama, e dos parâmetros estimados da forma sugerida, é possível calcular quantis de rugosidade da calha para diversas probabilidades. Por fim, pode-se proceder à amostragem de valores pelo método de Monte Carlo, com sorteio a partir da distribuição estimada, realizando-se na sequência milhares de simulações no modelo hidráulico

ou hidrodinâmico utilizado. Recomenda-se que o coeficiente de rugosidade de Manning seja definido, por tipologia de vegetação e de solo, a partir do método comparativo, valendo-se de fotos, imagens de satélite e inspeção de campo, ou aplicando-se o método de Cowan, quando possível. Essa abordagem, variando-se a rugosidade somente na calha, e fixando-a nas planícies de inundação, para avaliar a incerteza preditiva devido à primeira, foi utilizada nos estudos de Di Baldassarre *et al.* (2010) e de Ali *et al.* (2015).

Finalmente, menciona-se que nenhuma dentre as distribuições de probabilidades de variáveis contínuas avaliadas mostrou-se adequada para modelar as amostras de n_c obtidas *a posteriori* nos cenários B2 e D3, nos quais foi usado o cotograma da cheia de referência para calibração das incertezas. Como os valores desse parâmetro variaram muito pouco, concentrando-se em ambos os casos em 2 valores (0,022 e 0,023, e 0,020 e 0,021, respectivamente), cogitou-se que uma distribuição de variáveis discretas, como a hipergeométrica, poderia ser mais adequada para explicar sua (pouca) variabilidade. Relembra-se que se esperava encontrar faixas mais estreitas para n_c quando da utilização de níveis para a calibração Bayesiana, já que, segundo a equação de Manning, as vazões são mais sensíveis do que as profundidades à variação deste parâmetro. No entanto, a expectativa era a de que os valores médios *a posteriori* de n_c seriam semelhantes aos estimados quando da calibração pelo hidrograma. As diferenças encontradas foram atribuídas a possíveis inconsistências entre as bases topográficas das seções transversais e do zero das réguas linimétricas em Pirapora-Barreiro. De fato, toda a faixa de n_c nos cenários B2 e D3 resumiu-se a valores correspondentes a estimativas mínimas da rugosidade em canais naturais encontradas nas consagradas referências já mencionadas ao longo do texto. Por essas razões, priorizou-se a análise das amostras de n_c estimadas pela calibração com hidrogramas, a fim de extrair diretrizes para a generalização de uma distribuição que possa ser utilizada em outros estudos.

As recomendações ora realizadas devem ser usadas com cautela, lembrando-se que novas calibrações sob um esquema Bayesiano devem ser conduzidas em trechos fluviais de áreas de contribuição e morfologias semelhantes, para cheias de distintos períodos de retorno, de modo a possibilitar a generalização de diretrizes. Por fim, havendo disponibilidade de registros de cotas e vazões, é possível usá-los para: (i) calibração da incerteza de n_c por um esquema Bayesiano, utilizando-se como FDP *a priori* a distribuição Gama com os parâmetros estimados conforme indicado, a fim de averiguar se é possível estreitar os limites estipulados, e/ou (ii)

validar a distribuição adotada, verificando-se se o intervalo de credibilidade associado à incerteza preditiva paramétrica abrange o hidrograma disponível.

5.3.6 Quantificação da incerteza preditiva

No presente estudo, a incerteza preditiva, isto é, associada à predição das variáveis cota e vazão na extremidade de jusante do trecho fluvial estudado, foi quantificada em todos os cenários principais mostrados na Figura 5.12, a fim de se avaliar sua sensibilidade às diferentes hipóteses definidoras dos mesmos e de servir, por fim, como critério auxiliar na definição daqueles mais adequados à modelagem realizada.

A maneira presentemente utilizada para se estimar a incerteza preditiva, seja de ordem paramétrica, seja agregando todas as fontes de incerteza associadas à modelagem e aos dados de entrada e de saída, foi construindo-se intervalos de credibilidade a um nível de 95%, por meio da aplicação da expressão 4.6 (página 115) e dos critérios mencionados no subitem 4.2.5.3.

Os resultados são mostrados a seguir em termos gráficos e tabulares, para os cenários selecionados como representativos das incertezas, os quais foram apresentados e justificados no subitem 5.3.5.5. A seguinte convenção deve ser adotada em sua análise:

- IC Incerteza Preditiva Paramétrica: representa o intervalo de credibilidade a 95% elaborado quando se propagam somente as incertezas paramétricas pelo modelo hidrodinâmico, ou seja, associadas aos pares $\{n_c, n_p\}$;
- IC Incerteza Preditiva Parcial: representa o intervalo de credibilidade a 95% construído a partir da soma das vazões ou níveis obtidos em cada Δt ao se propagarem as incertezas paramétricas aos erros aleatórios e homo/heteroscedásticos. Tais erros foram estimados pela multiplicação $\sigma_t \cdot a_t$, com σ_t e a_t calculados a partir das variáveis latentes σ_0 , σ_1 (esta somente no Cenário B2), β e ξ extraídas de um conjunto paramétrico $\{\theta, \eta_\epsilon\}$ amostrado pelo método de Monte Carlo, sendo finalmente adicionados às vazões/cotas simuladas a partir do correspondente par de $\{n_c, n_p\}$.

O referido intervalo é chamado de parcial, pois ainda falta considerar a dependência serial para se concluir a construção dos intervalos preditivos que contabilizem a incerteza total, ou seja, oriunda de todas as fontes de incertezas previstas. Embora não seja um intervalo recomendado

para estabelecimento de quantis preditivos de vazões ou níveis, ele é mostrado nas figuras a fim de proporcionar uma analogia com a situação em que intervalos preditivos fossem construídos sem o cômputo da correlação serial entre os resíduos.

- IC Incerteza Preditiva Total: representa o intervalo de credibilidade a 95% obtido quando se somam as vazões/níveis que incorporam a incerteza dita parcial, em cada Δt , aos componentes estocásticos produzidos pelos modelos AR(1), ou seja, $\varepsilon_t = \phi_1 \cdot \varepsilon_{t-1} + \sigma_t \cdot a_t$, com ϕ_1 fixo ou extraído do mesmo conjunto paramétrico $\{\theta, \eta_\varepsilon\}$ sorteado por Monte Carlo que deu origem ao hidrograma/cotograma que contabiliza a chamada incerteza parcial.
- Os índices *obs* e *sim*, associados às variáveis vazão (Q) e nível (NA), indicam, respectivamente, valores observados ou simulados pelo modelo hidrodinâmico;
- A sigla “MVS” refere-se aos hidrogramas ou cotogramas simulados no modelo a partir do conjunto paramétrico $\{n_c, n_p\}$ que conduziu ao maior valor da função de log-verossimilhança em cada cenário;
- A sigla AIC refere-se ao termo “amplitude do intervalo de credibilidade”, expressa em metros ou m³/s, a depender da variável utilizada como paradigma; e
- A letra *H* indica a profundidade correspondente ao nível de água simulado ou observado, estimada para permitir o cálculo de relações percentuais no caso do uso de cotogramas.

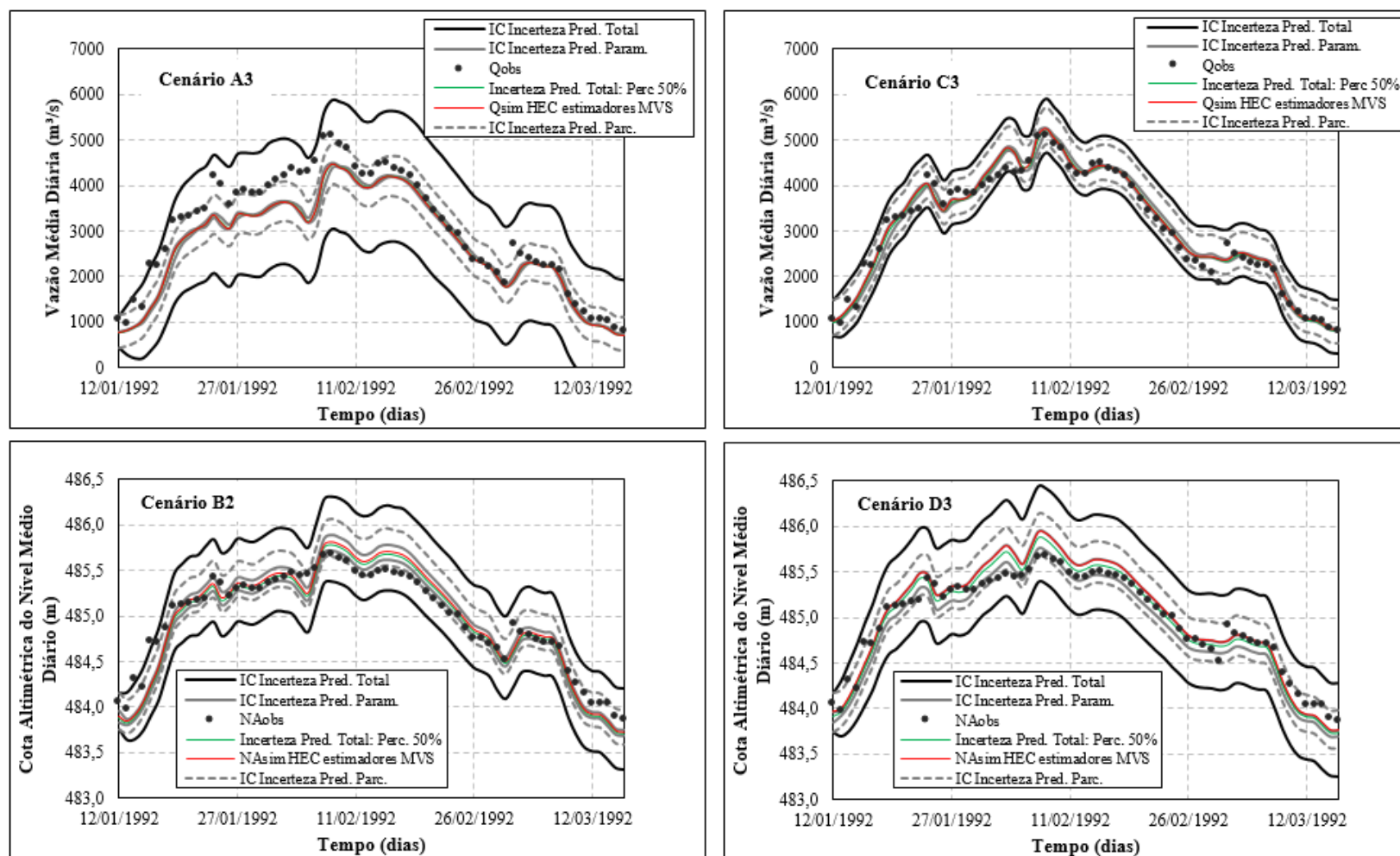


Figura 5.37 – Intervalos de credibilidade a 95% associados à incerteza da predição de vazões (Cenários A3 e C3) ou de níveis de água (Cenários B2 e D3) no extremo de jusante do trecho fluvial estudado.

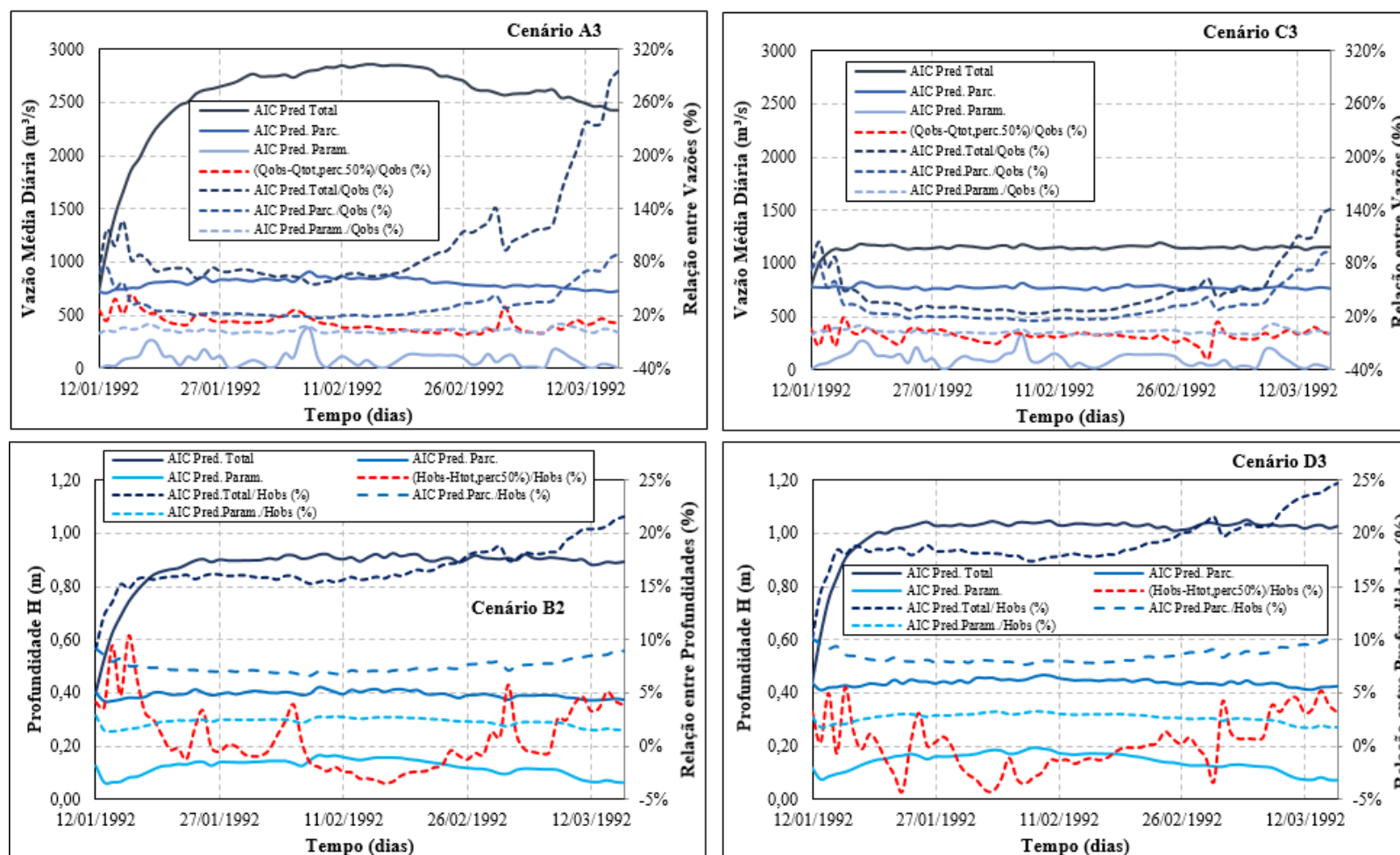


Figura 5.38 – Amplitude dos intervalos de credibilidade a 95% e sua relação com os valores observados das variáveis vazão (Cenários A3 e C3) ou nível de água (Cenários B2 e D3).

Tabela 5.20 – Percentis notáveis utilizados para elaboração dos intervalos de credibilidade a 95% associados à incerteza preditiva: Cenários A3 (esq.) e C3 (dir.).

Vazão ou Indicador de Referência	Perc. 2,5% (m³/s)	Mediana (m³/s)	Perc. 97,5% (m³/s)	AIC (m³/s)	AIC/Q _{obs} %
Incerteza Preditiva Paramétrica: Cenário A3					
Q _{inicial,obs} = 1063 m³/s (Mín AIC & Mín AIC/Q _{obs})	787	787	787	0,00	0,00%
Q _{pico,obs} = 5098 m³/s	4380	4463	4480	99,5	1,95%
Q _{final,obs} = 821 m³/s	722	724	731	9,18	1,12%
Máx AIC: Q _{07/02/1992} = 5061 m³/s	3958	4193	4326	368	7,28%
Máx AIC/Q _{obs} : Q _{09/03/1992} = 1601 m³/s	1477	1528	1641	164	10,24%
Média AIC (m³/s)			90,2		
Média AIC/Q _{obs}			3,21%		
Incerteza Preditiva TOTAL: Cenário A3					
Q _{inicial,obs} = 1063 m³/s (Mín AIC)	431	781	1152	721	67,83%
Q _{pico,obs} = 5098 m³/s (Mín AIC/Q _{obs})	3061	4448	5863	2803	54,97%
Q _{final,obs} = 821 m³/s (Máx AIC/Q _{obs})	-490	725	1934	2425	295,29%
Máx AIC: Q _{14/02/1992} = 4447 m³/s	2702	4114	5557	2855	64,20%
Média AIC (m³/s)			2544		
Média AIC/Q _{obs}			100,49%		

Vazão ou Indicador de Referência	Perc. 2,5% (m³/s)	Mediana (m³/s)	Perc. 97,5% (m³/s)	AIC (m³/s)	AIC/Q _{obs} %
Incerteza Preditiva Paramétrica: Cenário C3					
Q _{inicial,obs} = 1063 m³/s (Mín AIC & Mín AIC/Q _{obs})	1019	1019	1019	0,00	0,00%
Q _{pico,obs} = 5098 m³/s (Máx AIC)	5162	5239	5262	100,2	1,97%
Q _{final,obs} = 821 m³/s	840	841	846	6,03	0,73%
Máx AIC/Q _{obs} : Q _{09/03/1992} = 1601 m³/s	1623	1676	1820	197	12,33%
Média AIC (m³/s)			99,2		
Média AIC/Q _{obs}			3,55%		
Incerteza Preditiva TOTAL: Cenário C3					
Q _{inicial,obs} = 1063 m³/s (Mín AIC)	707	988	1490	783	73,66%
Q _{pico,obs} = 5098 m³/s (Mín AIC/Q _{obs})	4726	5207	5893	1167	22,88%
Q _{final,obs} = 821 m³/s (Máx AIC/Q _{obs})	327	814	1488	1161	141,36%
Máx AIC: Q _{24/02/1992} = 2951 m³/s	2487	2972	3689	1202	40,72%
Média AIC (m³/s)			1149		
Média AIC/Q _{obs}			47,61%		

Tabela 5.21 – Percentis notáveis utilizados para elaboração dos intervalos de credibilidade a 95% associados à incerteza preditiva: Cenários B2 (esq.) e D3 (dir.).

Nível de Água ou Indicador de Referência	Perc. 2,5% (m)	Mediana (m)	Perc. 97,5% (m)	AIC (m)	AIC/H _{obs} %
Incerteza Preditiva Paramétrica: Cenário B2					
NA _{inicial,obs} = 484,05 m (Máx AIC/H _{obs})	483,83	483,89	483,96	0,13	2,94%
NA _{pico,obs} = 485,67 m (Máx AIC)	485,72	485,81	485,89	0,17	2,82%
NA _{final,obs} = 483,86 m (Mín AIC)	483,69	483,72	483,75	0,06	1,49%
Mín AIC/H _{obs} : NA _{14/01/1992} = 484,31 m	483,87	483,90	483,93	0,06	1,37%
Média AIC (m)			0,12		
Média AIC/H _{obs}			2,30%		
Incerteza Preditiva TOTAL: Cenário B2					
NA _{inicial,obs} = 484,05 m (Mín AIC & Mín AIC/H _{obs})	483,75	483,87	484,15	0,40	9,22%
NA _{pico,obs} = 485,67 m	485,38	485,79	486,30	0,92	15,43%
NA _{final,obs} = 483,86 m (Máx AIC/H _{obs})	483,31	483,70	484,21	0,89	21,53%
Máx AIC: NA _{17/02/1992} = 485,45 m	485,25	485,64	486,17	0,93	16,12%
Média AIC (m)			0,87		
Média AIC/H _{obs}			16,71%		

Nível de Água ou Indicador de Referência	Perc. 2,5% (m)	Mediana (m)	Perc. 97,5% (m)	AIC (m)	AIC/H _{obs} %
Incerteza Preditiva Paramétrica: Cenário D3					
NA _{inicial,obs} = 484,05 m	483,85	483,91	483,97	0,12	2,71%
NA _{pico,obs} = 485,67 m (Máx AIC & Máx AIC/H _{obs})	485,73	485,82	485,92	0,19	3,27%
NA _{final,obs} = 483,86 m (Mín AIC)	483,69	483,73	483,76	0,07	1,72%
Mín AIC/H _{obs} : NA _{13/03/1992} = 484,04 m	483,85	483,88	483,92	0,07	1,68%
Média AIC (m)			0,14		
Média AIC/H _{obs}			2,65%		
Incerteza Preditiva TOTAL: Cenário D3					
NA _{inicial,obs} = 484,05 m (Mín AIC & Mín AIC/H _{obs})	483,74	483,91	484,18	0,44	10,10%
NA _{pico,obs} = 485,67 m	485,37	485,85	486,41	1,04	17,38%
NA _{final,obs} = 483,86 m (Máx AIC/H _{obs})	483,26	483,73	484,28	1,02	24,69%
Máx AIC: NA _{05/03/1992} = 484,74 m	484,23	484,70	485,28	1,05	20,84%
Média AIC (m)			1,00		
Média AIC/H _{obs}			19,11%		

Dentre as constatações realizadas por meio da análise dos intervalos de credibilidade, destacam-se as seguintes, as quais são demonstradas pelas figuras e tabelas fornecidas:

- A incerteza preditiva total, assim chamada por incorporar todas as fontes de erros da modelagem e dos dados, é mais ampla sob a hipótese de utilização de vazões como paradigma de calibração, quando avaliada em termos relativos (vide os gráficos da Figura 5.37).

Isto pode dever-se à relação entre vazões e cotas, definida pela equação de Manning e pela geometria da seção transversal de referência e das seções do trecho fluvial estudado, de forma que alterações na rugosidade têm maior sensibilidade sobre a vazão.

Outro fator que pode ter contribuído para esse comportamento, em conjunto ao aspecto mencionado anteriormente, concerne aos distintos graus de assimetria, curtose, correlação serial e homos/heteroscedasticidade apresentados pelos resíduos ao se alterarem as variáveis usadas como informação $\tilde{Y}$, ainda que todos tenham sido modelados pela função de verossimilhança Generalizada.

Ademais, cabe lembrar que se cogitaram possíveis inconsistências entre os dados topobatimétricos e a cota altimétrica da origem das réguas linimétricas do posto de jusante, as quais parecem ter acarretado uma tendência de subestimação de níveis de água em relação à geometria do trecho fluvial estudado. Tal aspecto também influenciou a estrutura estatística dos resíduos da modelagem gerados pela diferença entre os níveis de água no lugar das vazões, propiciando intervalos de credibilidade mais estreitos.

- Ao se utilizar o hidrograma como paradigma de calibração, a incerteza preditiva total reduz-se de maneira importante sob a hipótese de propagação dos hidrogramas das bacias incrementais.

Ainda que tais hidrogramas tenham sido estimados por meio de uma técnica de regionalização, o que, em princípio, agregaria mais incertezas ao processo, o resultado final, avaliado à luz dos critérios selecionados, foi de redução da incerteza final no extremo de jusante do trecho estudado. Isso pode ter ocorrido em função da mudança nas características estatísticas dos resíduos, visto que os mesmos aproximaram-se do zero, além de terem apresentado uma redução importante em seu grau de dependência temporal, em relação ao Cenário A3. De todo modo, o procedimento de transferência de informação hidrológica utilizado é robusto e simples

e baseia-se na definição de semelhanças morfológicas, climáticas e físicas entre as bacias com e sem dados de monitoramento, atribuindo confiabilidade às estimativas realizadas.

- Em todos os cenários, mesmo entre aqueles que não foram selecionados como representativos, a incerteza preditiva de ordem paramétrica é significativamente menor do que a incerteza preditiva total.

Ao contrário do que se supunha inicialmente nesta pesquisa, a incerteza paramétrica, representada na modelagem hidrodinâmica 1D pelos coeficientes de rugosidade de Manning, não constitui a fonte mais relevante de incertezas. Outras fontes contribuem sobremaneira para a variabilidade de predições expressas em termos de cotas e vazões geradas pelo modelo usado, como erros de medição/observação e de natureza epistêmica em vazões afluentes e nas seções transversais utilizadas. Erros epistêmicos e de medição relacionados às vazões e cotas no posto de jusante, consideradas neste estudo como a informação $\tilde{Y}$, também estão embutidos de maneira implícita nas diferenças $\tilde{Y} - Y(\theta)$, contribuindo para a incerteza final (SCHOUPS e VRUGT, 2010). Ademais, outras fontes de erros, relacionados à estrutura da modelagem no HEC-RAS, de caráter principalmente epistêmico e que também contribuem para a incerteza final, foram avaliadas igualmente sob a forma de resíduos modelados pela função de verossimilhança Generalizada. Dentre tais fontes de incertezas, algumas tiveram seus efeitos individualmente quantificados de maneira não sistemática ao longo do trabalho, como a representação de pontes, o intervalo temporal das simulações e o fator de ponderação *theta*, e mostraram-se menos importantes do que as de origem paramétrica.

- A consideração da correlação serial é importante na estimação da incerteza preditiva total.

Um dos aspectos vislumbrados como dos mais relevantes trata-se da contribuição da correlação serial das séries de resíduos na composição da incerteza preditiva total, conforme destacado nos gráficos da Figura 5.37. Por outro lado, a propagação das cheias das bacias incrementais pelo modelo proporcionou, no caso estudado, uma redução da dependência serial entre os resíduos avaliados em termos de vazão, o que, por sua vez, diminuiu o peso desse fator ao se estimar a incerteza preditiva total (Cenário C3). De toda maneira, a incorporação da correlação serial dos resíduos pelo esquema metodológico adotado parece mais adequada à elaboração de intervalos preditivos associados a uma modelagem hidrodinâmica, em busca de limites realísticos para as variáveis de saída e que mostrem os efeitos combinados das diversas fontes de incertezas.

- Os intervalos preditivos, são positivamente assimétricos, embora tal aspecto não seja evidente em termos visuais. Esse comportamento praticamente unânime dos resíduos foi adequadamente reproduzido na geração dos intervalos de credibilidade, o que foi constatado seja por intervalo temporal, seja ao longo de cada uma das séries de resíduos brutos geradas *a posteriori*; e
- Por fim, verifica-se também pelos gráficos da Figura 5.37 que a maioria das vazões defluentes observadas encontra-se dentro do intervalo de credibilidade a 95%, o que mostra a capacidade do modelo hidráulico elaborado com as informações topobatimétricas e fluviométricas disponíveis em reproduzir o comportamento da seção fluvial de jusante. Entretanto, observa-se uma limitação do modelo em caracterizar o ramo ascendente do hidrograma, com tendência a subestimar as vazões.

Uma avaliação adicional realizada no quesito da incerteza preditiva concerne à verificação da variação dos níveis de água acarretada pela variabilidade das vazões no intervalo de credibilidade paramétrico e total a 95%, quando estas foram usadas como paradigma. Analogamente, avaliaram-se os efeitos sobre as vazões sob a hipótese de variação dos níveis em seus intervalos preditivos. Para tanto, os valores extremos (i.e., percentis 2,5% e 97,5%) referentes à vazão/nível de pico do hidrograma/cotagrama que estão indicados na Tabela 5.20 e na Tabela 5.21, respectivamente para os Cenários C3 e D3, foram lançados sobre a curva-chave do rio São Francisco em Pirapora-Barreiro a fim de se estimar os correspondentes níveis e descargas em intervalo médio diário. Ressalta-se que essa aferição presta-se a uma análise qualitativa de como ambas as variáveis se relacionam, pois os valores foram estimados, sendo indicados na Figura 5.39 e na Figura 5.40 em termos aproximados. Além disso, a curva-chave da seção de referência altera-se a cada simulação com distintos valores das rugosidades. Outro ponto a se observar é que os cenários usados na presente avaliação são aqueles nos quais as aflúncias laterais foram consideradas.

A variação dos níveis de água entre seus limites preditivos a 95% provoca maior alteração sobre as vazões do que a variação destas quando estimada pelo correspondente intervalo preditivo. Segundo a curva-chave, um acréscimo de 1,04 m no nível máximo da seção de referência acarretaria um aumento de vazão da ordem de 4500 m³/s, valor quase 4 vezes maior do que a incerteza associada à vazão de pico, de cerca de 1200 m³/s.

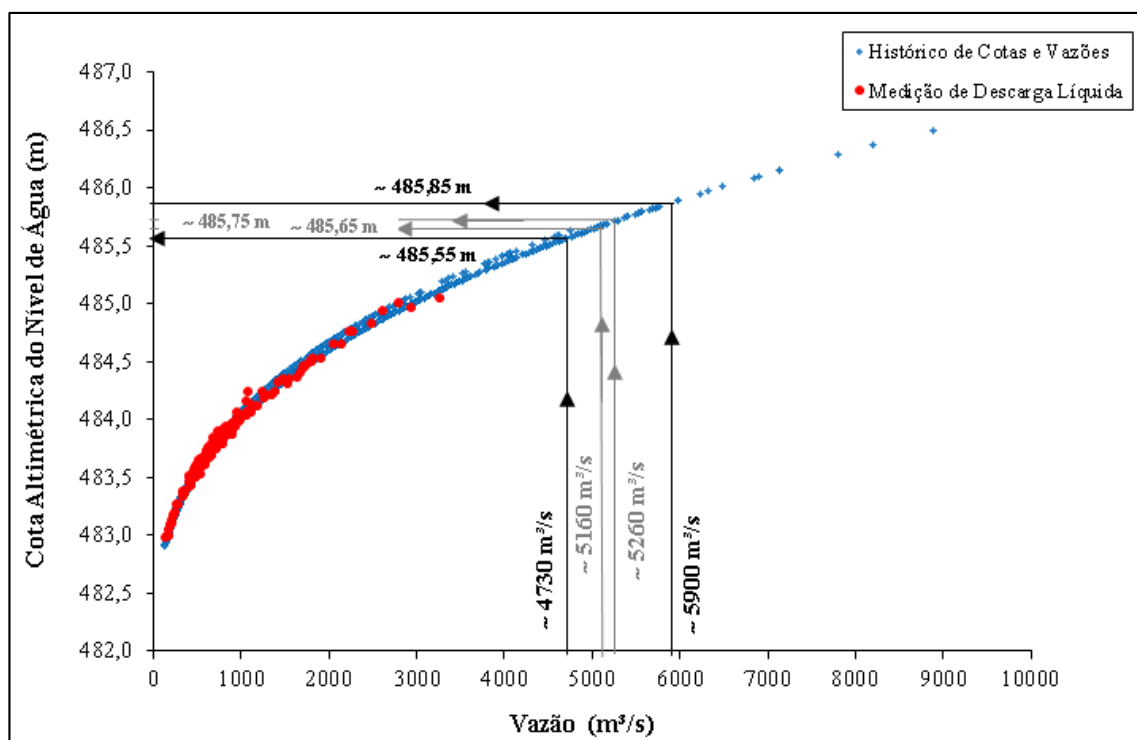


Figura 5.39 – Inspeção da variação de níveis de água na seção do rio São Francisco em Pirapora-Barreiro acarretada pela variação de vazões no intervalo de credibilidade a 95% da incerteza preditiva paramétrica (setas e valores em cinza) e total (setas e valores em preto).

Por sua vez, a variação das vazões em seus limites preditivos, extraídos de seu intervalo de credibilidade a 95%, causa um impacto sobre os níveis menor do que aquele previsto no respectivo intervalo de credibilidade construído para o cotograma. Um aumento de 1200 m³/s na vazão de pico em Pirapora-Barreiro causaria, segundo sua curva-chave, uma lâmina de água 0,30 m maior, estimativa 3,5 vezes menor do que a incerteza total associada aos níveis de água máximos na referida seção.

Dessa forma, constata-se que, apesar de a incerteza preditiva associada à variável nível ser relativamente menor do que a incerteza preditiva referente às vazões, quando ambas são comparadas aos valores observados de seus correspondentes paradigmas, a variabilidade causada sobre as vazões, caso se usem os níveis como referência, é muito maior do que a própria incerteza dessas.

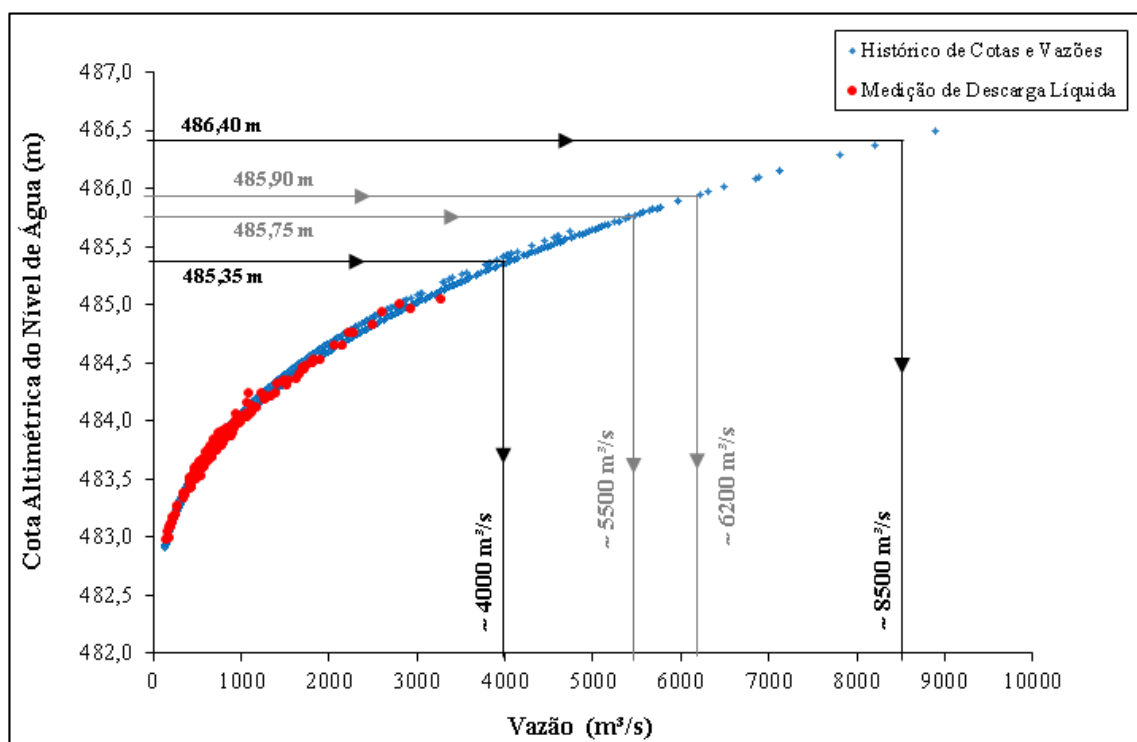


Figura 5.40 – Inspeção da variação de vazões na seção do rio São Francisco em Pirapora-Barreiro causada pela variação de níveis de água no intervalo de credibilidade a 95% da incerteza preditiva paramétrica (setas e valores em cinza) e total (setas e valores em preto).

Sendo assim, recomenda-se que, no trecho fluvial estudado, à luz dos critérios e dados utilizados, sejam consideradas as vazões e seus respectivos intervalos de credibilidade preditivos e incertezas paramétricas obtidas *a posteriori* como indicadores da incerteza associada à modelagem hidrodinâmica 1D para estudos futuros e para o estabelecimento de diretrizes a partir dos resultados ora mostrados. Justificativa adicional a essa escolha refere-se à possível inconsistência de natureza topográfica mencionada ao longo dos resultados.

Adicionalmente, havendo consistência entre dados topográficos, a calibração das incertezas pelo esquema Bayesiano utilizando-se um cotagrama de jusante poderia ser usada em estudos hidráulicos nos quais a variável nível de água é o objeto principal de projeto, a exemplo da avaliação da segurança fornecida por diques laterais para proteção contra cheias. Como se viu, a faixa preditiva será mais ampla, a favor da segurança, comparativamente à situação em que se usam vazões para esta calibração. Obviamente, critérios técnicos e econômicos devem ser levados em conta para seleção de alternativas, sobretudo sob a possibilidade de adoção de valores de referência associados a percentis mais elevados.

No que tange aos mapas de inundação e perfis de escoamento probabilísticos, foram utilizados dois critérios para sua composição. Para sua elaboração, foram consideradas as incertezas de variáveis e parâmetros obtidas *a posteriori* no Cenário C3, definido como o mais adequado, confiável e robusto dentre aqueles selecionados como referenciais, devido aos aspectos explorados ao longo dos itens 5.3.5 e 5.3.6. Os critérios são mencionados a seguir:

- Contabilização somente da incerteza preditiva paramétrica:
 - (i) Sorteio de 10.000 pares de parâmetros $\{n_c, n_p\}$ pelo método de Monte Carlo a partir da distribuição conjunta do vetor $\{\theta, \eta_\epsilon\}$ estimado *a posteriori*, seguido pela simulação de cada par no modelo hidrodinâmico;
 - (ii) Armazenamento dos níveis máximos em cada seção transversal, obtidos em cada uma das correspondentes simulações hidrodinâmicas;
 - (iii) Ordenação dos níveis máximos, por seção transversal, e posterior cálculo de seus percentis notáveis;
 - (iv) Elaboração dos perfis de escoamento probabilísticos a partir dos percentis notáveis de 2,5%, 50% e 97,5% estimados em todas as seções transversais; e
 - (v) Construção do mapa de inundação probabilístico a partir dos percentis notáveis de 2,5%, 50% e 97,5% computados nas últimas 6 seções de jusante (inclusive as interpoladas).
- Contabilização da incerteza preditiva total (procedimento descrito no capítulo de metodologia):
 - (i) Separação dos percentis notáveis de 2,5%, 50% e 97,5% associados à incerteza preditiva total correspondente à vazão de pico do hidrograma de cheia;
 - (ii) Simulação dos respectivos percentis no modelo HEC-RAS em regime permanente, utilizando-se o par $\{n_c, n_p\}$ que conduziu ao valor máximo da função de verossimilhança utilizada na inferência Bayesiana; e
 - (iii) Uso dos níveis obtidos em cada seção transversal do trecho fluvial estudado para compor os perfis de escoamento probabilísticos, e nas últimas 6 seções de jusante para montagem do mapa de inundação probabilístico.

Ressalta-se que ambos os procedimentos descritos apresentam limitações, de forma que os resultados fornecidos devem ser usados com cautela. Em ambos os casos, parte-se do pressuposto de que a estrutura de resíduos dos níveis de água seja semelhante àquela correspondente às vazões no extremo de jusante do trecho fluvial estudado, sem que tenha sido possível confirmar tal hipótese, principalmente em se tratando de toda a extensão do estirão fluvial. No segundo caso, não é possível usar o modelo hidrodinâmico, pois as vazões de referência foram extraídas diretamente do intervalo de credibilidade associado à incerteza preditiva total e seria impraticável traduzi-la em termos de rugosidades. A mancha probabilística construída nessa situação é meramente indicativa e sinaliza a necessidade de se utilizarem manchas observadas como paradigma de calibração Bayesiana, caso se queira quantificar a incerteza preditiva total associada à variável extensão da inundação.

Por outro lado, nas circunstâncias em que a informação $\tilde{Y}$ utilizada na calibração de incertezas foram os cotagramas registrados no extremo de jusante do trecho fluvial estudado, não foram traçados os correspondentes mapas de inundação e perfis de escoamento probabilísticos. Para tanto, seria necessário estimar os percentis notáveis associados à variável nível de água em todas as seções transversais disponíveis, por meio da aplicação da Equação 4.6 em cada uma delas, sob a hipótese de semelhança na estrutura de resíduos em todas as seções em relação à seção de jusante, ainda que não fosse possível confirmá-la. Devido à seleção das incertezas obtidas *a posteriori* no Cenário C3, que se valeu de vazões como paradigma $\tilde{Y}$, para elaboração de produtos concernentes à incerteza preditiva paramétrica e total, tal procedimento não constituiu escopo da presente pesquisa. Por fim, a dificuldade de confirmação de hipóteses quanto aos resíduos ao se utilizar uma variável distinta daquela cuja incerteza preditiva desejasse estimar mostra, mais uma vez, a necessidade e a importância da adoção de manchas de inundação como paradigma de calibração de incertezas para atender ao objetivo de obtenção de mapas de inundação e perfis de escoamento probabilísticos sob a metodologia utilizada.

Os resultados são mostrados nas Figuras 5.41, 5.42, 5.43 e 5.44.

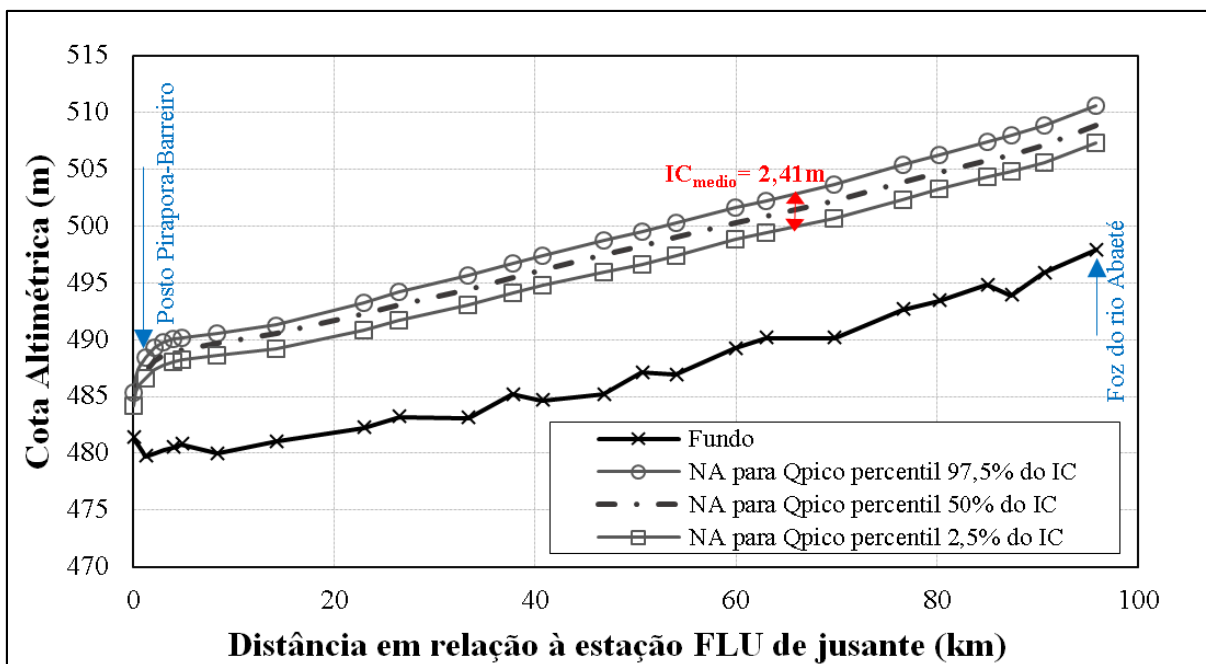


Figura 5.41 – Intervalo de credibilidade a um nível de 95% para os perfis de escoamento ao longo do trecho estudado, considerando-se somente a incerteza preditiva paramétrica: Cenário C3.

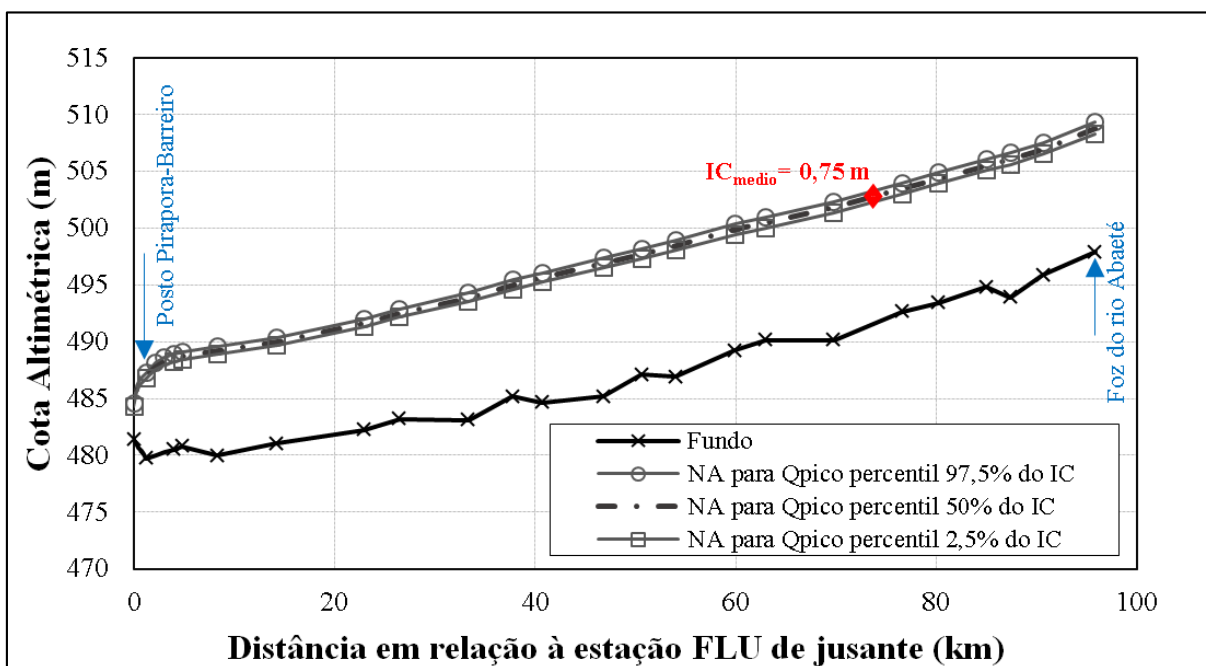


Figura 5.42 – Intervalo de credibilidade a um nível de 95% para os perfis de escoamento ao longo do trecho estudado, considerando-se a incerteza preditiva total: Cenário C3.

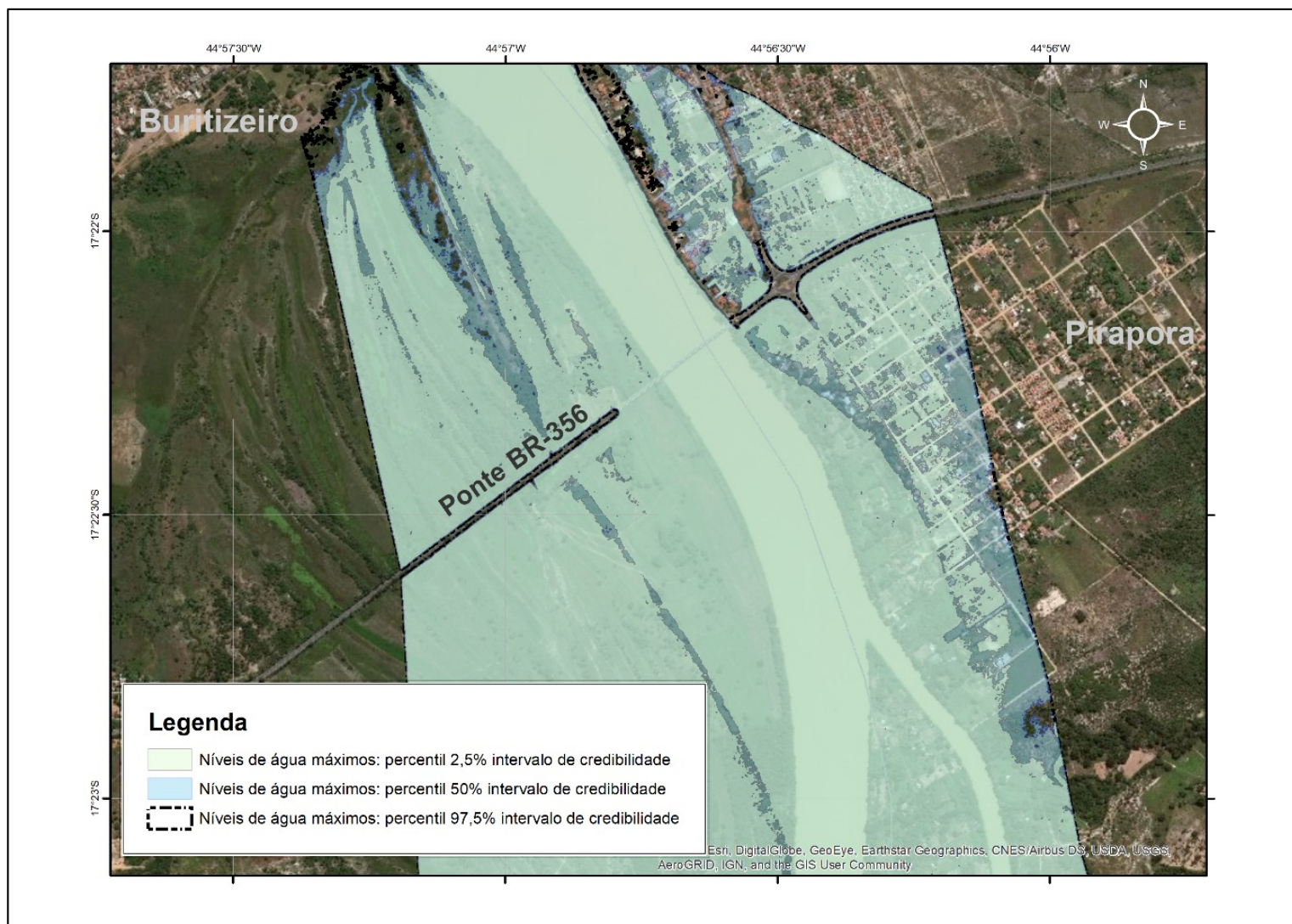


Figura 5.43 – Mapa de inundação probabilístico que contabiliza somente a incerteza preditiva associada à rugosidade, sobre a porção de montante da cidade de Pirapora, próximo à ponte da BR-356 sobre o rio São Francisco: Cenário A3.

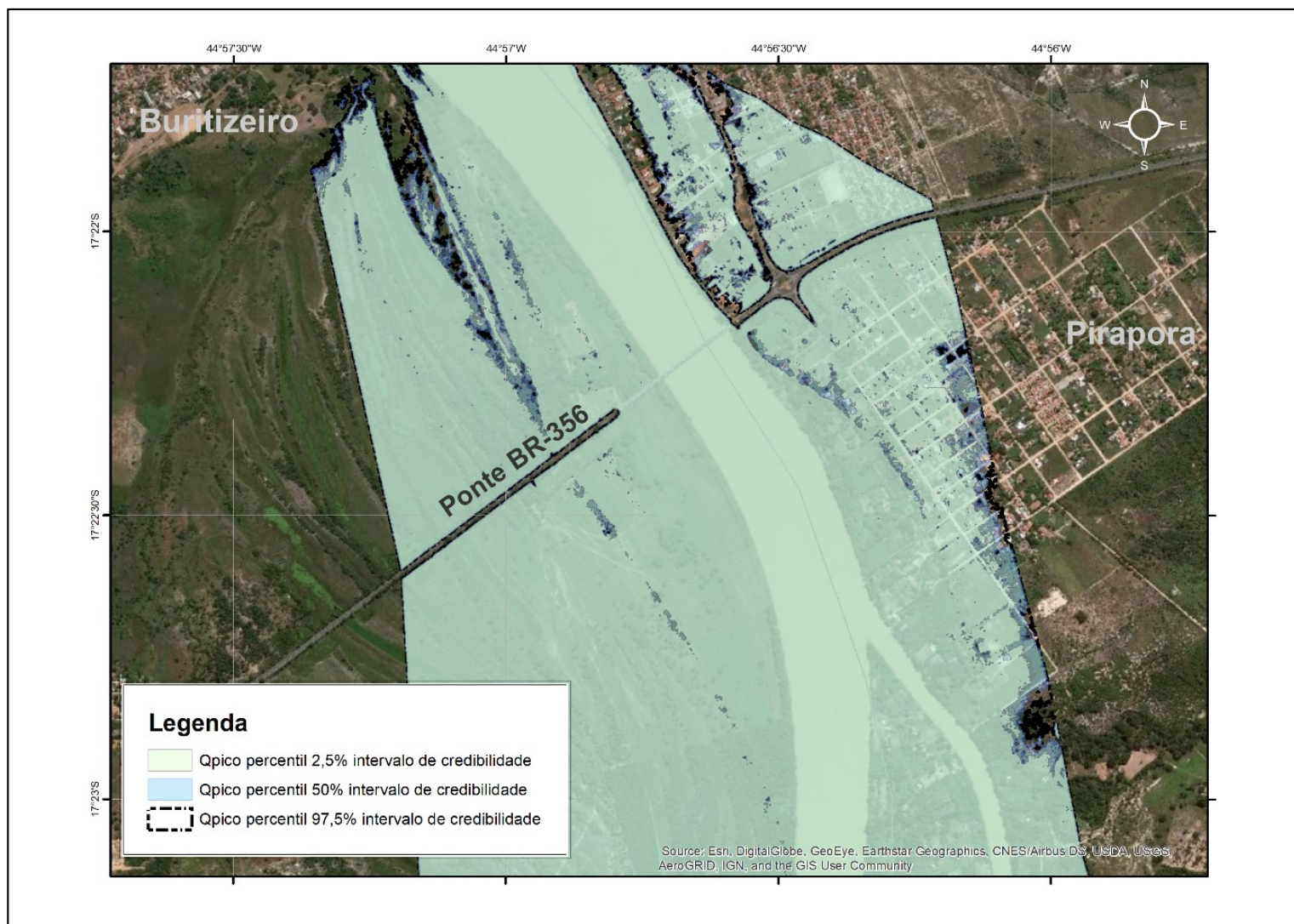


Figura 5.44 – Mapa de inundação probabilístico que contabiliza somente a incerteza preditiva associada à rugosidade, sobre a porção de montante da cidade de Pirapora, próximo à ponte da BR-356 sobre o rio São Francisco: Cenário A3.

Observa-se que a extensão da inundação na porção de jusante do trecho fluvial estudado, na área urbana de Pirapora a montante da corredeira, é sensível à incerteza paramétrica inferida *a posteriori*, mesmo que em termos de vazões e de níveis sua variabilidade pareça ter efeito reduzido. Esse impacto importante provavelmente advém da morfologia da planície de inundação na região mapeada, na qual pequenas variações de lâmina de água produzem importantes mudanças na largura de topo das seções e na extensão da inundação.

A parte a jusante da corredeira não foi estudada por fazer parte do trecho do rio São Francisco com interferência relevante do remanso causado pelo rio das Velhas. Verifica-se, em função da forma do vale fluvial na área mapeada e da ampla variação da rugosidade, que mesmo pequenas alterações de vazão podem induzir grandes mudanças na largura alcançada pela lâmina de água.

Por fim, vale ressaltar que as manchas de inundação probabilísticas expostas na Figura 5.43 e na Figura 5.44 foram traçadas sobre o MDT de maior resolução horizontal disponível, que mostrou ter cobertura limitada para representar a totalidade da planície de inundação em muitos locais ao longo do trecho levantado. Uma alternativa para se contornar tal aspecto seria proceder à compatibilização de dados oriundos de bases topográficas de distintas fontes na planície de inundação, embora tal procedimento agregue incertezas adicionais de difícil quantificação pelo esquema adotado neste estudo.

6 CONCLUSÕES

Os estudos realizados no âmbito desta pesquisa mostraram como é possível estimar as incertezas associadas à modelagem hidrodinâmica e aos dados que a alimentam, a partir de um esquema Bayesiano formal, em que o conhecimento *a priori* acerca da variabilidade de parâmetros e variáveis associadas aos resíduos é atualizado com informações observadas de alguma variável descritiva do fenômeno em questão, no caso, as vazões e níveis monitorados no extremo de jusante do trecho fluvial selecionado, situado na alta porção do rio São Francisco. Dentre as conclusões extraídas a partir das análises e resultados realizados, destacam-se as seguintes:

- O esquema metodológico proposto adaptou-se adequadamente à realidade de informações topobatimétricas e fluviométricas comumente encontradas em bacias hidrográficas brasileiras de médio e grande porte que possibilitam uma modelagem hidrodinâmica unidimensional.

Nesse sentido, foi necessário proceder a algumas mudanças na metodologia idealizada inicialmente, valendo-se de uma série de procedimentos de análise minuciosa dos resíduos resultantes da modelagem em questão. As adaptações referem-se ao número de variáveis latentes utilizadas, o qual variou em função do tipo de informação usada como paradigma de calibração de incertezas e das condições de contorno adotadas nos cenários de simulação, notadamente aquelas relacionadas à consideração ou não de afluentes das bacias incrementais ao trecho fluvial estudado.

- A adoção de uma função de verossimilhança com mais parâmetros, também chamados de variáveis latentes, agregou complexidade e maior esforço computacional às simulações para aproximação da distribuição conjunta *a posteriori*, uma vez que a convergência do método MCMC adotado foi alcançada mais tardiamente à medida que mais variáveis eram adicionadas ao processo. Por outro lado, tais variáveis foram extremamente relevantes à inferência Bayesiana, uma vez que representaram características estatísticas esperadas para os resíduos de uma modelagem matemática não linear, como o é a modelagem hidrodinâmica.

De fato, as várias análises feitas sobre os resíduos brutos e transformados, no caso, removendo-se a correlação serial e extraíndo-se a variância, seja esta variável ou não, revelaram distintos graus de assimetria e de excesso de curtose, bem como dependência serial relevante. Tais

características não são comumente consideradas na prescrição de um modelo estatístico dos resíduos que sirva de suporte para a construção da função de verossimilhança quando da adoção de métodos Bayesianos para quantificação de incertezas associadas a uma modelagem matemática de um sistema ambiental e aos dados de entrada e de saída utilizados, conforme se constatou na revisão de literatura (KAVETSKI *et al.*, 2003; STEDINGER *et al.*, 2008; VRUGT, 2008; 2009a; SCHOUPS e VRUGT, 2010; DI BALDASSARRE *et al.*, 2010; HUTTON *et al.*, 2011; BEVEN *et al.*, 2011; JUNG e MERWADE, 2012; CAMACHO *et al.*, 2015; VRUGT, 2016).

Constatou-se, de maneira semelhante ao explorado em outros estudos de quantificação de incertezas com funções de verossimilhança distintas daquela tradicionalmente usada, a qual se vale das hipóteses de homoscedasticidade, normalidade e independência dos resíduos, que a identificação das incertezas paramétricas é sensível ao modelo de resíduos e à função de verossimilhança. Como, então, saber *a priori* qual seria a função de verossimilhança mais correta a ser adotada em um estudo de quantificação de incertezas? Seguindo-se as recomendações de Kavetski *et al.* (2003), de Schoups e Vrugt (2010) e de Hutton *et al.* (2011), é importante que sejam realizadas diversas avaliações sobre os resíduos a fim de se confirmarem as hipóteses nas quais se baseiam o modelo de resíduos e a função de verossimilhança, sugestões estas que embasaram profundamente a presente pesquisa.

- Adicionalmente, a introdução de variáveis latentes associadas ao modelo de resíduos e à função de verossimilhança adotada, denominada Generalizada (SCHOUPS e VRUGT, 2010), permitiu a identificação de incertezas distintas daquelas de origem paramétrica. Trata-se de erros aleatórios e epistêmicos associados à estrutura do modelo, bem como oriundos da medição/obtenção/estimação de dados de entrada e de saída à modelagem (seções topobatimétricas, vazões afluentes e defluentes). Separar todas essas fontes de incertezas ainda é um desafio (KAVETSKI *et al.*, 2006a; SCHOUPS e VRUGT, 2010; HUTTON *et al.*, 2011), devido à falta de informações para avaliá-las de forma compartimentada e para elaborar funções de verossimilhança para cada uma delas, a depender da metodologia adotada.

Nesta pesquisa, inicialmente se vislumbrou a possibilidade de proceder à separação das incertezas paramétricas daquelas associadas às vazões afluentes, por meio do método BaTEA, preconizado por Kavetski *et al.* (2003). No entanto, à medida que os estudos foram

desenvolvidos, percebeu-se que a quantidade de informações sobre as cheias no estirão fluvial selecionado, com vetores de comprimento n menores do que uma centena, provavelmente seria pequena para permitir a calibração tal como sugerido pelos referidos autores. Cogitava-se, ainda, aplicar um esquema pseudo-Bayesiano amplamente utilizado em modelagem hidrológica e hidráulica, o anteriormente mencionado GLUE (BEVEN e BINLEY, 1992), a fim de compartimentar as incertezas do hidrograma afluente, da topografia e dos coeficientes de Manning, sujeitos à mesma função de verossimilhança. As severas críticas ao GLUE encontradas em algumas referências (KAVETSKI *et al.*, 2003; TODINI, 2007; STEDINGER *et al.*, 2008; VRUGT *et al.*, 2009a) bem como os resultados alcançados com a adoção do esquema Bayesiano formal contribuíram para que aquela consagrada metodologia fosse posta de lado. Afinal, as incertezas *a posteriori* de todos esses componentes da modelagem poderiam ser erroneamente identificadas ao serem avaliadas à luz de um modelo incorreto de resíduos, e a quantificação da incerteza preditiva poderia ser igualmente comprometida.

Por fim, a introdução de variáveis latentes ao processo de inferência Bayesiana, que representem a natureza dos resíduos da modelagem hidrodinâmica, mostrou-se apropriada para quantificar as demais fontes de incertezas além da paramétrica, ainda que de forma implícita e compacta. Esse aspecto foi de suma importância na composição da incerteza preditiva associada às vazões e níveis de água no extremo de jusante do trecho fluvial estudado no rio São Francisco, à medida em que foi possível propagar as incertezas paramétricas pelo HEC-RAS hidrodinâmico e adicionar em seguida as incertezas das demais fontes, consubstanciadas nas variáveis latentes.

Por outro lado, incertezas de ordem epistêmica, referentes à prescrição e à seleção do modelo hidrodinâmico, não puderam ser tratadas explicitamente por meio do esquema proposto. Algumas dentre tais incertezas tiveram seus efeitos avaliados como uma análise de sensibilidade sobre as vazões e níveis de água ao longo do estirão fluvial de interesse, e mostraram-se pouco importantes perante as incertezas quantificadas formalmente. Outras abordagens, a exemplo do BMA – *Bayesian Model Averaging* (DUAN *et al.*, 2007; VRUGT e ROBINSON, 2007), podem ser futuramente utilizadas para compreensão desse tipo de erro, embora seus resultados provavelmente não sejam facilmente comparáveis àqueles obtidos com a metodologia presentemente aplicada.

- A distribuição SEP padronizada, tal como proposta por Schoups e Vrugt (2010) caracterizada pela média nula e por desvio-padrão unitário, e curtose e assimetria representados por mais dois parâmetros, respectivamente, β e ξ , mostrou-se flexível para quantificar os resíduos i.i.d. resultantes da modelagem 1D no HEC-RAS, após remoção da dependência temporal e da hetero/homoscedasticidade.

Esse comportamento foi especialmente interessante ao se alterarem as condições de contorno das simulações hidrodinâmicas, inserindo-se ou não as afluições laterais ao trecho estudado, e o paradigma de calibração de incertezas, se cotograma ou hidrograma no posto de Pirapora-Barreiro. Percebeu-se a adequação da SEP às características de assimetria e de curtose das séries de resíduos resultantes em todos os casos. Ressalva-se que a avaliação da aderência dessa distribuição aos resíduos i.i.d. foi de caráter qualitativo, por inspeção visual do ajuste entre quantis empíricos e teóricos, e conceitual, observando-se a representatividade garantida pelos quatro parâmetros da SEP padronizada aos descritores estatísticos das séries de resíduos a_t . Constatou-se que, tanto os resíduos brutos gerados pelas diferenças entre vazões/níveis de água observados e simulados, quanto os resíduos transformados, dificilmente seriam aderentes a uma distribuição de probabilidades teórica, o que foi atribuído à não-linearidade que é intrínseca à modelagem hidrodinâmica.

- O método MCMC utilizado, denominado DREAM (VRUGT *et al.*, 2009b), alcançou a convergência da distribuição conjunta *a posteriori* em todos os cenários apresentados ao longo da tese, corroborando com as vantagens sugeridas em Vrugt *et al.* (2009b) e Vrugt (2016). Verificou-se que a convergência foi mais lenta nos cenários de inferência de maior número de parâmetros, 7 ou 8, no máximo. Por outro lado, percebeu-se também dificuldade de identificação de uma ou mais variáveis latentes quando o parâmetro ϕ_1 estava sujeito à inferência. Tal fato, aliado à superestimação deste parâmetro em alguns cenários, condicionou a sua fixação para sequenciamento de novas simulações de inferência.
- Embora haja algumas diferenças quanto aos resíduos obtidos em cada tipo de Cenário (nomeados por A, B, C e D), observou-se similaridade por tipo quanto à forma, posição e escala entre as distribuições obtidas *a posteriori* para os coeficientes de rugosidade de Manning na calha principal e nas planícies (n_c e n_p), à exceção dos cenários em que as hipóteses de normalidade, homoscedasticidade e independência foram adotadas. Tal aspecto demonstra robustez da inferência realizada perante algumas alterações realizadas a fim de

melhorar a identificação de alguns dentre os demais parâmetros envolvidos e de proporcionar a construção de intervalos de credibilidade preditivos realistas.

Nesse quesito, destacou-se a identificação *a posteriori* de n_c , cujos percentis 2,5% e 97,5% e a média mostraram-se coerentes a valores mínimos, máximos e usuais encontrados na literatura para canais fluviais como o rio São Francisco na porção estudada. Ademais, a forma da distribuição empírica inferida assemelha-se às distribuições avaliadas em um estudo conduzido pela CPRM (2016, comunicação eletrônica), a partir da consulta a *experts* sobre esse parâmetro em 8 trechos fluviais situados em Minas Gerais.

Em contraposição, as amostras obtidas *a posteriori* para n_p mostraram que, no presente caso, este parâmetro teve função primordialmente compensatória aos valores de n_c , sem que tenha sido possível atribuir plausibilidade física aos valores mais frequentes das distribuições empíricas. Credita-se esse comportamento à pouca informação em termos de vazões ou cotas nas planícies de inundação contida na cheia utilizada como paradigma.

- Ressalva-se que, ao contrário do esperado, a utilização de cotograma como paradigma $\tilde{Y}$ de calibração de incertezas não produziu estimativas de medidas de tendência central da rugosidade na calha semelhantes àqueles produzidos quando o hidrograma da cheia de fevereiro de 1992 foi usado.

As distribuições *a posteriori* para esse parâmetro, nos cenários dos tipos B e D, mostraram valores sempre inferiores aos obtidos nos de tipos A e C. Aventa-se que essa subestimação sistemática esteja associada às possíveis diferenças quanto à precisão dos levantamentos topográficos que deram origem às seções transversais utilizadas, fornecidas pela CEMIG, e ao “zero” das réguas linimétricas do posto fluviométrico de Pirapora-Barreiro, operado pela CPRM. Dessa forma, é provável que cotogramas elaborados a partir da cota altimétrica disponível para esta origem das réguas subestimem sistematicamente os níveis de água quando comparados com as cotas altimétricas de fundo da seção transversal na estação fluviométrica em questão.

Outro ponto de destaque nesse quesito concerne ao desvio-padrão reduzido das distribuições *a posteriori* para n_c quando os cotogramas foram usados, comparativamente à situação revelada pelos hidrogramas. Provavelmente, isto se deve ao fato de que as vazões são mais sensíveis às variações da rugosidade do que os níveis de água, de forma que, para atender à faixa de variação

do cotograma da cheia de 1992, a rugosidade deveria estar confinada entre limites muito próximos, ao contrário do que ocorreu quando o hidrograma foi o paradigma.

Além disso, ainda que a incerteza *a posteriori* de n_c calibrada com o cotograma tenha sido menor, constatou-se, por meio de uma análise simplificada sobre a curva-chave do posto de jusante, que o impacto causado sobre as vazões nesse local seria maior do que a variabilidade produzida sobre estas devido à variação de n_c quando inferido pelo uso do hidrograma. Tal aspecto está atrelado à forma como essas variáveis relacionam-se na equação de Manning e à morfologia da seção de avaliação de resultados e do trecho a jusante.

- Ao contrário do que se esperava inicialmente, a incerteza preditiva associada à variabilidade *a posteriori* dos parâmetros do modelo hidrodinâmico 1D foi pequena perante aquela produzida pelas fontes de erros representadas pelas variáveis latentes.

Por um lado, tal aspecto pode dever-se à posição da seção transversal do posto fluviométrico de comparação de resultados, cujo controle hidráulico para a faixa de vazões estudada é do tipo crítico e se situa 3,6 km a jusante. Por outro, é possível que as incertezas acarretadas pelos erros de medição/estimação de vazões afluentes, das seções topobatimétricas e das vazões defluentes de fato sejam de maior relevância do que as primeiras.

Embora não tenha sido viável quantificar a parcela da incerteza preditiva total atribuída a cada uma das outras fontes de erros mencionadas, por meio do método Bayesiano formal adotado, os resultados extraídos demonstram que a definição correta da incertezas associadas às predições de cotas e vazões em um modelo hidrodinâmico deve contabilizá-las em conjunto às paramétricas, ainda que representadas de forma implícita.

- Os mapas e perfis de escoamento probabilísticos elaborados no contexto deste trabalho são indicativos de possíveis desdobramentos e produtos a serem contabilizados quando a incerteza preditiva é adequadamente quantificada. Cabe lembrar que, para sua elaboração, a condição mais correta teria sido a inferência de incertezas pela adoção de uma ou mais manchas de inundação observadas, a fim de resguardar e corroborar a estrutura de resíduos associada às profundidades obtidas ao longo do estirão fluvial.

7 RECOMENDAÇÕES E TRABALHOS FUTUROS

Dentre os métodos pré-selecionados para compor o arcabouço metodológico de cunho Bayesiano, ao longo desta pesquisa verificou-se a viabilidade de aplicar somente aquele denominado como esquema formal, com utilização de uma função de verossimilhança adequada a resíduos de modelagem com comportamento heteroscedástico, não Gaussiano e com autocorrelação. Além disso, verificou-se, por meio de numerosas simulações do modelo hidráulico elaborado, a necessidade de limitar as vazões propagadas pelo trecho considerado, devido à existência de mais de um controle hidráulico na seção cujos dados serviram como informação de atualização do conhecimento *a priori* dos parâmetros do modelo e da função de verossimilhança.

As conclusões extraídas das simulações e resultados mostraram que há uma vertente inteira a ser pesquisada, concernente à utilização de métodos de cunho Bayesiano à quantificação de incertezas associadas a modelos hidráulicos/hidrodinâmicos e aos seus dados de entrada e de saída. Sendo assim, vislumbram-se possíveis verificações a fim de confirmar a aplicabilidade de tais técnicas como ferramenta para estimar e quantificar as incertezas de uma modelagem hidrodinâmica em trechos fluviais com a disponibilidade de informações comumente encontradas no Brasil ao alcance do meio acadêmico. Sugerem-se as seguintes frentes de trabalho e diretrizes para verificações futuras:

- Validação das incertezas obtidas neste estudo, realizando-se amostragem pelo método de Monte Carlo dos conjuntos paramétricos $\{\theta, \eta_{\epsilon}\}$ *a posteriori*, priorizando-se os do Cenário C3 (calibrados com hidrograma e sob a consideração das afluências laterais). Os correspondentes valores dos pares (n_c, n_p) devem ser simulados no modelo hidrodinâmico, e às vazões calculadas devem ser adicionados os resíduos $\mathbf{E}$, conforme conceituação da Equação 4.6 (página 115).

É importante que seja selecionada uma cheia cuja vazão de pico tenha ordem de grandeza semelhante àquela da cheia de referência usada na calibração, respeitando-se o limite de 5000 m³/s, condicionado pela controle hidráulico a jusante. Espera-se que os intervalos de credibilidade elaborados contenham os dados observados $\tilde{\mathbf{Y}}$ associados à cheia selecionada para validação das incertezas.

- Refinamento das incertezas paramétricas e do modelo de resíduos por atualização Bayesiana, ou seja, por meio da utilização de novas cheias de magnitude semelhante à de fevereiro de 1992. Esse tipo de avaliação foi sugerido por Hall e Solomatine (2008). O mesmo pode ser feito com cheias de maior período de retorno, à medida que as séries de monitoramento tornem-se mais extensas.

Alternativamente, é possível utilizar a informação de duas ou mais cheias notáveis, com períodos de retorno da mesma ordem de grandeza, para compor o vetor $\tilde{Y}$, a fim de alimentar a calibração de incertezas pelo esquema proposto em um único cenário. Tal análise pode fornecer a sensibilidade das distribuições *a posteriori* ao tamanho n do vetor de dados observados.

- Recomenda-se que estudos futuros de calibração de incertezas em cursos de água naturais pelo esquema Bayesiano sugerido priorizem o uso de hidrogramas observados como paradigma, e que considerem a estimação das contribuições laterais nas bacias não monitoradas. Os limites dos intervalos de credibilidade preditivos associados às vazões defluentes mostraram-se mais realistas nessa condição, bem como seu reflexo sobre as variações de nível de água na seção de referência.

Hidrogramas defluentes preditivos obtidos por meio da metodologia aplicada presentemente, que incorporem a incerteza paramétrica e das variáveis latentes, podem assim ser utilizados em estudos de balanço hídrico, de propagação de cheias e de assoreamento em reservatórios, bem como no dimensionamento de seus órgãos extravasores.

A avaliação de outras variáveis de saída, como velocidades e níveis de água, tanto localmente quanto ao longo de um estirão fluvial, é recomendada sob as hipóteses supracitadas incorporando-se somente a incerteza paramétrica do modelo hidrodinâmico, devido às limitações de se traduzir a incerteza preditiva total sob a forma de outras variáveis distintas daquela utilizada na calibração Bayesiana.

Por outro lado, caso se queira definir diretamente a incerteza sobre a predição de níveis de água em determinada seção transversal, a calibração por meio do uso do cotograma é recomendada, lembrando-se de que provavelmente serão produzidas estimativas mais conservadoras. Tal diretriz aplica-se, por exemplo, ao dimensionamento de estruturas hidráulicas como diques laterais de proteção contra cheias, situados nas proximidades da seção cujos dados forem usados como paradigma de calibração.

Por sua vez, para o mapeamento de planície de inundação, caso se queira avaliar a incerteza preditiva total sobre as áreas atingidas, recomenda-se fortemente que seja considerada ao menos uma mancha de inundação observada como informação $\tilde{Y}$ a ser usada como paradigma de calibração.

Ainda no que tange à seleção do tipo de variável a ser usada como informação $\tilde{Y}$, uma avaliação pertinente a ser realizada concerne à adoção de ao menos duas variáveis na composição da função de verossimilhança, a exemplo de cotas e vazões em determinada seção transversal ou em distintas seções, de modo semelhante ao realizado por Camacho *et al.* (2015) – vide Equação 3.48 à página 74. Tal abordagem provavelmente teria impacto sobre a forma e os limites estimados para as distribuições *a posteriori* e poderia até mesmo melhorar a identificação de tais incertezas mesmo com a utilização de somente um evento notável. Por outro lado, a hipótese de eventuais inconsistências sistemáticas em níveis de água, tais como aquelas constatadas ao longo desta pesquisa, e em vazões deve ser averiguada previamente para garantir a correta aplicação desses registros à metodologia proposta.

Por fim, as estimativas obtidas a partir dos intervalos de credibilidade preditivos para níveis de água e vazões, respeitadas as condições e limitações discorridas anteriormente, podem ser utilizadas para se quantificar o impacto da incertezas paramétricas e de outras fontes sobre a estimativa de danos associados às inundações, e sobre impactos de ordem ambiental. Por sua vez, os mapas probabilísticos podem ser utilizados na composição de mapas de vulnerabilidade e de risco final associado a enchentes de determinado período de retorno.

- A aplicação do método GLUE pode parecer uma alternativa atraente para a separação das incertezas paramétricas e de dados de entrada, como seções transversais e vazões de pico afluentes. No entanto, o uso de medidas de desempenho como função de verossimilhança informal pode conduzir a estimativas incorretas das incertezas marginais e dos intervalos de credibilidade preditivos, segundo comprovado em estudos como os de Kavetski *et al.*, (2003), Stedinger *et al.* (2008) e Vrugt *et al.* (2009a). Nesse sentido, é relevante que sejam conduzidos trabalhos de comparação entre as incertezas quantificadas pelo GLUE e por métodos Bayesianos formais.
- Um método Bayesiano formal vislumbrado para possibilitar a separação entre as incertezas paramétricas e as oriundas de erros de medição associados às estimativas de vazões afluentes, inclusive as incrementais, e defluentes, no caso da modelagem hidrodinâmica, é o

BaTEA (KAVETSKI *et al.*, 2003). Para tanto, é importante que sejam prescritas funções de verossimilhança e suas respectivas variáveis latentes para as fontes de erros as quais se pretende compartimentar, e que se comprovem as estruturas de seus resíduos. Métodos como esse parecem viáveis à medida que haja registros de cheias de maior duração, proporcionando o aumento do número de dados a serem usados como informação $\tilde{Y}$. Dados adicionais também podem ser necessários para comprovar o suposto comportamento dos resíduos, a exemplo de campanhas de medição de descarga líquida e obtenção de marcas de cheia excepcionais.

- Caso se queira explorar as incertezas associadas a cheias excepcionais no trecho fluvial estudado, a exemplo da cheia de fevereiro de 1979, é forçoso que o modelo hidrodinâmico unidimensional elaborado seja complementado com inserção de seções até o rio das Velhas, bem como neste afluente e no trecho que se segue a jusante de sua confluência. Certamente, isso agregará complexidade às avaliações e esforço computacional, ao mesmo tempo que possibilitará a verificação de possíveis alterações na identificação da incerteza paramétrica e de outras fontes em face de alterações no controle hidráulico da seção transversal do posto fluviométrico de referência e do trecho a montante.
- A relevância das incertezas relacionadas às bases topobatimétricas utilizadas pode ser acessada por meio de uma análise de sensibilidade dos resultados encontrados na pesquisa, como intervalos de credibilidade dos cotogramas e hidrogramas, aos erros presentes em dados de levantamento topobatimétrico, atrelados, por exemplo, à precisão das medidas, em abordagem semelhante à do estudo de Ali *et al.* (2015).
- A avaliação da incerteza paramétrica sobre a predição de hidrogramas de cheia de períodos de retorno moderados (10 ~ 50 anos) em estirões fluviais de porte e morfologia semelhantes aos do rio São Francisco no trecho considerado pode ser feita de maneira preliminar considerando-se as diretrizes estabelecidas no subitem 5.3.5.6, previamente a uma calibração via método Bayesiano formal, seguida pela propagação das incertezas pelo modelo hidrodinâmico.
- A calibração de incertezas pelo método Bayesiano formal adotado pode ser avaliada considerando-se a aplicação de um modelo autoregressivo aos resíduos padronizados, após a remoção da heteroscedasticidade, com inferência conjunta de $\boldsymbol{\eta}_\epsilon = \{\phi_p, \sigma_0, \sigma_1, \xi, \beta\}$, seguindo-se proposição de Evin *et al.* (2013). Segundo esses autores, a padronização prévia pela variância resolveria eventuais problemas relacionados à quantificação da incerteza

preditiva quando os componentes de ϕ_p produzissem séries de resíduos brutos semelhantes a um passeio aleatório, conforme se observou em alguns cenários realizados na presente pesquisa.

Nos cenários nos quais a heteroscedasticidade dos resíduos foi constatada, o modelo utilizado na presente pesquisa para descrever σ_t foi representado por uma equação linear (Equação 4.2, página 100), na qual o desvio-padrão dos resíduos brutos varia de modo crescente segundo as vazões simuladas. Sugere-se que, em estudos futuros que se valham do esquema metodológico proposto aplicado à modelagem hidrodinâmica, tal modelo seja adaptado a fim de representar melhor a dispersão dos resíduos observados. Para tanto, recomenda-se que dois modelos sejam testados, formados respectivamente por: (i) duas equações lineares sequenciais, ambas crescentes, sendo os correspondentes limite superior e inferior representados pela vazão de transbordamento, a exemplo do débito de margens plenas; e (ii) duas equações lineares sequenciais, tendo a descarga de transbordamento por limite de aplicação entre ambas; no entanto, a primeira equação seria crescente, enquanto a segunda, decrescente, de modo a reproduzir o comportamento verificado em cenários como os de tipo A (vide, por exemplo, a Figura 5.15, à página 153). Embora tais considerações agreguem complexidade à metodologia, na forma de variáveis latentes adicionais, é interessante que se avalie seu impacto sobre as incertezas paramétricas obtidas *a posteriori*, e sobre a incerteza preditiva.

REFERÊNCIAS

- ACKERMAN, C. T. HEC-GeoRAS – GIS Tools for Support of HEC-RAS using ArcGIS: User's Manual. Davis, CA: USACE, Institute for Water Resources, HEC, 2011.
- ALI, A. M.; SOLOMATINE, D. P.; DI BALDASSARRE, G. Assessing the impact of different sources of topographic data on 1-D hydraulic modelling of floods. *Hydrology and Earth System Sciences*, v. 19, p. 631-643, 2015.
- AGÊNCIA NACIONAL DE ÁGUAS. Inventário das Estações Fluviométricas. Brasília: ANA; SGH, 2ª ed., 2009.
- ARCADIS. Working Group Floods(CIS) resource document “Flood Risk Management, Economics and Decision Making Support”. European Union, oct. 2012.
- ARCEMENT, G.J.; SCHNEIDER, V.R. Guide for Selecting Manning's Roughness Coefficients for Natural Channels and Flood Plains - Water Supply - Paper 2339. Los Angeles, California: U. S. Geological Survey, 1989.
- ARONICA, G.; HANKIN, B.; BEVEN, K. Uncertainty and equifinality in calibrating distributed roughness coefficients in a flood propagation model with limited data. *Advances in water resources*, v. 22, n. 4, p. 349-365, 1998.
- AYYUB, B. M.; KLIR, G. J. Uncertainty Modeling and Analysis in Engineering and the Sciences. Chapman and Hall/CRC, 2006.
- BAPTISTA, M. B.; COELHO, M. M. L. P. Fundamentos de Engenharia Hidráulica. Belo Horizonte: Editora UFMG/Escola de Engenharia da UFMG, 2002, 440 p.
- BARNES, H. H. Roughness Characteristics of Natural Channels. US Geological Survey Water-Supply Paper n. 1849. Washington: US Government Printing Office, 1967, 219 p.
- BATES, P.; TRIGG, M.; NEAL, J.; DABROWA, A. LISFLOOD-LP: User's Manual – Code Release 5.9.6. Bristol: School of Geographical Sciences, University of Bristol, 2013, 49 p.
- BERNARDO, J.; SMITH, A. Bayesian Theory. New York: John Wiley and Sons, 1994, 586 p.
- BEVEN, K. A manifesto for the equifinality thesis. *Journal of hydrology*, v. 320, n. 1-2, p.18-36, 2006.
- BEVEN, K. Towards integrated environmental models of everywhere: uncertainty, data and modelling as a learning process. *Hydrology and Earth System Sciences*, v. 11, n. 1, p. 460-467, 2007.
- BEVEN, K.; BINLEY, A. The future of distributed models: model calibration and uncertainty prediction. *Hydrological processes*, v. 6, n. 3, p. 279-298, 1992.
- BEVEN, K.; BINLEY, A. GLUE: 20 years on. *Hydrological processes*, v. 28, n. 24, p. 5897-5918, 2014.

BEVEN, K.; LEEDAL, D.; MCCARTHY, S. Framework for assessing uncertainty in fluvial flood risk mapping. Flood Risk Management Research Consortium – FRMRC – Research Report SWP1.7, jul. 2011.

BLASONE, R.S.; VRUGT, J.A.; MADSEN, H.; ROSBJERG, D.; ROBINSON, B.A.; ZYVOLOSKI, G.A. Generalized likelihood uncertainty estimation (GLUE) using adaptive Markov Chain Monte Carlo sampling. *Advances in Water Resources*, v. 31, n. 4, p. 630-648, 2008.

BOZZI, S.; PASSONI, G.; BERNARDARA, P.; GOUTAL, N.; ARNAUD, A. Roughness and discharge uncertainty in 1D water level calculations. *Environmental Modeling & Assessment*, v. 20, n. 4, p.343-353, 2015.

BRAY, D. I. Estimating average velocity in gravel-bed rivers. *Journal of the Hydraulics Division, ASCE*, vol. 105, n. HY9, p. 1103-1122, 1979.

BRUIJN, K.; KLIJN, F.; ÖLFERT, A.; PENNING-ROWSELL, E.; SIMM, J.; WALLIS, M. Flood risk assessment and flood risk management: An introduction and guidance based on experiences and findings of FLOODsite (an EU-funded Integrated project). In: KLIJN, F., Editor. *Deltares | Delft Hydraulics*, Delft, the Netherlands, feb. 2009.

BRUNNER, G.; CEIWR-HEC. HEC-RAS - River Analysis System Hydraulic Reference Manual – version 5.0. US Army Corps of Engineers, Institute for Water Resources, 2016a.

BRUNNER, G.; CEIWR-HEC. HEC-RAS - River Analysis System User's Manual – version 5.0. US Army Corps of Engineers, Institute for Water Resources, 2016b.

CAMACHO, R.A.; MARTIN, J.L.; MCANALLY, W.; DÍAZ-RAMIREZ, J.; RODRIGUEZ, H.; SUCSY, P.; ZHANG, S. A comparison of Bayesian methods for uncertainty analysis in hydraulic and hydrodynamic modeling. *Journal of the American Water Resources Association*, v. 51, n. 5, p. 1372-1393, 2015.

CHENG, Q.B.; CHEN, X.; XU, C.Y.; REINHARDT-IMJELA, C.; SCHULTE, A. Improvement and comparison of likelihood functions for model calibration and parameter uncertainty analysis within a Markov chain Monte Carlo scheme. *Journal of hydrology*, v. 519, p. 2202-2214, 2014.

CHOW, V. T. *Open-channel hydraulics*. New York: McGraw-Hill, 1959, vol. 1.

CHOW, V. T.; MAIDMENT, D. R.; MAYES, L. W. *Applied Hydrology*. International edition, McGraw-Hill, Inc. 1988.

COOK, A.; MERWADE, V. Effect of topographic data, geometric configuration and modeling approach on flood inundation mapping. *Journal of Hydrology*, v. 377, n. 1, p. 131-142, 2009.

COSTA, V. A. F. Estimacoo indireta de quantis de enchentes extremas a partir de modelos chuva-vazoo com emprego conjunto de um gerador estocstico de precipitacoo diria, anlise Bayesiana e distribuicoes limitadas superiormente. 197 p. Tese (Doutorado em Saneamento,

Meio Ambiente e Recursos Hídricos) - Departamento de Hidráulica e Recursos Hídricos, Universidade Federal de Minas Gerais, Belo Horizonte - MG, 2015.

CPRM – Serviço Geológico do Brasil. Regionalização de vazões das sub-bacias 40 e 41 – Alto São Francisco. Volume II – Regionalização de vazões médias. Relatório Final – Volume II. Convênio 015/2000 ANEEL – 013/CPRM/2000. Belo Horizonte: Superintendência Regional de Belo Horizonte, 2001.

CPRM – Serviço Geológico do Brasil. Estimativa do coeficiente de Manning para aferição da incerteza de vazões de cheias calculadas por métodos indiretos. Belo Horizonte: CPRM – Superintendência Regional de Belo Horizonte, 2016.

CRUTCHER, H. L. A note on the possible misuse of the Kolmogorov-Smirnov test. *Journal of Applied Meteorology*. v. 14, n. 8, p.1600-1603, 1975.

DALRYMPLE, T.; BENSON, M. A. Measurements of Peak Discharge by the Slope-Area Method. *Techniques of Water Resources Investigation of the US Geological Survey – Chapter 2, Book 3 – Applications of Hydraulics (TWI 3-A2)*. US Government Printing Office, 1984.

DI BALDASSARRE, G.; LAIO, F.; MONTANARI, A. Design flood estimation using model selection criteria. *Physics and Chemistry of the Earth, Parts A/B/C*, v. 34, n. 10-12, p. 606-611, 2009.

DI BALDASSARRE, G.; SCHUMANN, G.; BATES, P.D.; FREER, J.E.; BEVEN, K.J. Flood-plain mapping: a critical discussion of deterministic and probabilistic approaches. *Hydrological Sciences Journal – Journal des Sciences Hydrologiques*, v. 55, n. 3, p. 364-376, 2010.

DOMENEGHETTI, A.; CASTELLARIN, A.; BRATH, A. Assessing rating-curve uncertainty and its effects on hydraulic model calibration. *Hydrology and Earth System Sciences*, v. 16, n. 4, p. 1191-1202, 2012.

DUAN, Q.; AJAMI, N. K.; GAO, X.; SOROOSHIAN, S. Multi-model ensemble hydrologic prediction using Bayesian model averaging. *Advances in Water Resources*, v. 30, n. 5, p. 1371-1386, 2007.

EU - EUROPEAN UNION. Directive 2000/60/EC of the European Parliament and of the Council of 23 October 2000 establishing a framework for Community action in the field of water policy. *Official Journal of the European Communities*, L327, dec. 2000.

EU - EUROPEAN UNION. Directive 2007/60/EC of the European Parliament and of the Council of 23 October 2007 on the assessment and management of flood risks. *Official Journal of the European Union*, L288, nov. 2007.

EVIN, G.; KAVETSKI, D.; THYER, M.; KUCZERA, G. Pitfalls and improvements in the joint inference of heteroscedasticity and autocorrelation in hydrological model calibration. *Water Resources Research*, v. 49, p. 4518-4524, 2013.

FERNANDES, W.; SILVA, A. T. Introduction to Bayesian Analysis of Hydrologic Variables. In: NAGHETTINI, M., Editor. *Fundamentals of Statistical Hydrology*. Springer, 2017. Chp. 11, p. 497-536.

FERNANDEZ, C.; STEEL, M. J. F. On Bayesian modeling of fat tails and skewness. *Journal of the American Statistical Association*, v. 93, n. 441, p. 359–371, 1998.

GELMAN, A.; RUBIN, D.B. Inference from iterative simulation using multiple sequences. *Statistical science*, v. 7, n. 4, p. 457-472, 1992.

GILKS W.R.; RICHARDSON S.; SPIEGELHALTER D., Editors. *Markov chain Monte Carlo in practice*. Boca Raton, Florida: CRC press, 1998.

GRIFFITHS, G. A. Flow resistance in coarse gravel bed rivers. *Journal of the Hydraulics Division, ASCE*, vol. 107, n. HY7, p. 899-918, 1981.

GOODELL, C. *Breaking the HEC-RAS Code: A User's Guide to Automating HEC-RAS*. Portland, Oregon: H2ls, 2014.

HALL, J. W. Handling uncertainty in the hydroinformatic process. *Journal of Hydroinformatics*, v. 5, n. 4, p. 215-232, 2003.

HALL, J; SOLOMATINE, D. A framework for uncertainty analysis in flood risk management decisions. *International Journal of River Basin Management*, v. 6, n. 2, p. 85-98, 2008.

HASTINGS, W. Monte Carlo sampling methods using Markov chains and their application, *Biometrika*, v. 57, n. 1, p. 97-109, 1970.

HICKS, D. M.; MASON, P. D. *Roughness Characteristics of New Zealand Rivers: A Handbook for Assigning Hydraulic Roughness Coefficients to River Reaches by the "Visual Comparison" Approach*. Wellington: NIWA and Water Resources Survey Publications, LLC, 1998, 326 p.

HORRITT, M. S.; DI BALDASSARRE, G.; BATES, P. D.; BRATH, A. Comparing the performance of a 2-D finite element and a 2-D finite volume model of floodplain inundation using airborne SAR imagery. *Hydrological Processes*, v. 30, n. 20-21, p. 2745-59, 2007.

HORRITT, M.S.; BATES, P.D. Effects of spatial resolution on a raster based model of flood flow. *Journal of Hydrology*, v. 253, n. 1-4, p. 239-249, nov. 2001.

HUTTON, C.J.; VAMVAKERIDOU-LYROUDIA, L.S.; KAPELAN, Z.; SAVIC, D.A. *Uncertainty Quantification and Reduction in Urban Water Systems (UWS) Modelling: Evaluation Report*. 92 p. 2011.

JIN, X. L.; XU, C. Y.; ZHANG, Q.; SINGH, V. P. Parameter and modeling uncertainty simulated by GLUE and a formal Bayesian method for a conceptual hydrological model. *Journal of Hydrology*, v. 383, n. 3-4, p. 147-155, 2010.

JUNG, Y.; MERWADE, V. Uncertainty quantification in flood inundation mapping using generalized likelihood uncertainty estimate and sensitivity analysis. *Journal of Hydrologic Engineering*, v. 17, n. 4, p. 507-520, 2012.

JUNG, Y.; MERWADE, V. Estimation of uncertainty propagation in flood inundation mapping using a 1-D hydraulic model. *Hydrological Processes*, v. 29, p. 624-640, 2015.

JUNG, Y. Uncertainty in Flood Inundation Mapping, 2011, 116 f. Thesis (Doctorate of Philosophy – Ph.D., Civil Engineering), Purdue University, West Lafayette, Indiana, 2011.

KAVETSKI, D.; FRANKS, S.W.; KUCZERA, G. Confronting input uncertainty in environmental modelling. Calibration of watershed models. In: DUAN, Q.; GUPTA, H. V.; SOROOSHIAN, S.; ROUSSEAU, A. N.; TURCOTTE, R., Editors. *Calibration of Watershed Models*. American Geophysical Union, 2003. Vol. 6, Chp. 4, p. 49-68.

KAVETSKI, D.; KUCZERA, G.; FRANKS, S.W. Bayesian analysis of input uncertainty in hydrological modeling: 2. Application. *Water Resources Research*, v. 42, n. 3, mar. 2006a.

KAVETSKI, D.; KUCZERA, G.; FRANKS, S.W. Bayesian analysis of input uncertainty in hydrological modeling: 1. Theory. *Water Resources Research*, v. 42, n. 3, mar. 2006b.

KRZYSZTOFOWICZ, R. Bayesian theory of probabilistic forecasting via deterministic hydrologic model. *Water Resources Research*, v. 35, n. 9, p. 2739-2750, 1999.

KUCZERA, G. Correlated rating curve error in flood frequency inference. *Water Resources Research*, v. 32, n. 7, p. 2119-2127, 1996.

KUCZERA, G.; PARENT, E. Monte Carlo assessment of parameter uncertainty in conceptual catchment models: the Metropolis algorithm. *Journal of Hydrology*, v. 211, n. 1-4, p. 69-85, nov. 1998.

LANE, S. N. Roughness – time for a re-evaluation? *Earth Surface Processes and Landforms*, v. 30, n. 2, p. 251-3, 2005.

LIU, Y.; GUPTA, H. V. Uncertainty in hydrologic modeling: Toward an integrated data assimilation framework. *Water Resources Research*, v. 43, n. 7, 2007.

LEON, A.S.; GOODELL, C. Controlling HEC-RAS using MATLAB. *Environmental Modelling & Software*, v. 84, p. 339-348, oct. 2016.

MANTOVAN, P.; TODINI, E. Hydrological forecasting uncertainty assessment: Incoherence of the GLUE methodology. *Journal of hydrology*, v. 330, v. 1-2, p. 368-381, 2006.

McKAY, M. D.; BECKMAN, R. J; CONNOVER, W. J. A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics*, v. 21, n. 2, p. 239-245, 1979.

MERWADE, V.; OLIVERA, F.; ARABI, M.; EDLEMAN, S. Uncertainty in flood inundation mapping current issues and future directions. *Journal of Hydrologic Engineering*, v. 13, n. 7, p. 608-620, 2008.

METROPOLIS, N.; ROSENBLUTH, A.; ROSENBLUTH, M.; TELLER, A.; TELLER, E. Equation of state calculations by fast computing machines. *J Chem Phys*, v. 21, n. 6, p. 1087-1092, 1953.

MILOGRANA, J. Sistemática de auxílio à decisão para a seleção de alternativas de controle de inundações urbanas. 316 p. Tese (Doutorado em Tecnologia Ambiental e Recursos Hídricos) - Departamento de Engenharia Civil e Ambiental, Universidade de Brasília, Brasília - DF, 2009.

MIRZAEI, M.; HUANG, Y. F.; EL-SHAFIE, A.; SHATIRAH, A. Application of the generalized likelihood uncertainty estimation (GLUE) approach for assessing uncertainty in hydrological models: a review. *Stochastic environmental research and risk assessment*, v. 29, n. 5, p. 1265-1273, 2015.

MISHRA, S.K.; SINGH, V. Soil Conservation Service Curve Number (SCS-CN) Methodology. Springer, 2003.

MONTANARI, A. Large sample behaviors of the generalized likelihood uncertainty estimation (GLUE) in assessing the uncertainty of rainfall-runoff simulations. *Water resources research*, v. 41, n. 8, 2005.

MONTANARI, A. What do we mean by ‘uncertainty’? The need for a consistent wording about uncertainty assessment in hydrology. *Hydrological Processes*, v. 21, n. 6, p. 841-5, 2007.

MORETTIN, P. A.; TOLOI, C. *Análise de Séries Temporais*. São Paulo: Blucher, 2ª Ed., 2006, 564 p.

MORGAN, M. G.; HENRION, M. *Uncertainty: A Guide to Dealing with Uncertainty In Quantitative Risk and Policy Analysis*. New York: Cambridge University Press, 1990.

NAGHETTINI, M.; PINTO, E. J. A. *Hidrologia Estatística*. CPRM – Serviço Geológico do Brasil, 2007, 561 p.

ONU – United Nations. Water Action Decade (2018 – 2028): United Nations Secretary-general’s plan. Dec. 2016. Disponível em https://wateractiondecade.org/wp-content/uploads/2018/03/UN-SG-Action-Plan_Water-Action-Decade-web.pdf. Acesso em 19 de setembro de 2019.

ONU - United Nations. President of the General Assembly: summary on the Water Action Decade. New York. Apr. 2018. Disponível em <https://wateractiondecade.org/wp-content/uploads/2018/04/Summary-Water-Decade-23-April.pdf>. Acesso em 19 de setembro de 2019.

ONU - United Nations. Infográfico: Sustainable Cities: why they matter. Nov. 2016. Disponível em <https://www.un.org/sustainabledevelopment/wp-content/uploads/2016/08/11.pdf>. Acesso em 19 de setembro de 2019.

ONU - United Nations. Transforming our world: the 2030 agenda for sustainable development. A/RES/70/1. Oct. 2015. Disponível em <https://sustainabledevelopment.un.org/content/documents/21252030%20Agenda%20for%20Sustainable%20Development%20web.pdf>. Acesso em 19 de setembro de 2019.

PAPPENBERGER, F.; BEVEN, K.; HORRITT, M.; BLAZKOVA, S. Uncertainty in the calibration of effective roughness parameters in HEC-RAS using inundation and downstream level observations. *Journal of Hydrology*, v. 302, n. 1, p. 46-69, 2005.

PAPPENBERGER, F.; MATGEN, P.; BEVEN, K. J.; HENRY, J. B.; PFISTER, L.; FRAIPONT, P. Influence of uncertain boundary conditions and model structure on flood inundation predictions. *Advances in Water Resources*, v. 29, n. 10, p. 1430-1449, 2006a.

PAPPENBERGER, F.; HARVEY, H.; BEVEN, K.; HALL, J.; MEADOWCROFT, I. Decision tree for choosing an uncertain analysis methodology: a wiki experiment. *Hydrological Processes*, v. 20, p. 3793-3798, 2006b.

PINHEIRO, M. C. Modelação matemática do escoamento e caracterização geométrica da morfologia fluvial: subsídios para a navegação e a operação de reservatórios. 212 p. Tese de Doutorado. Rio de Janeiro: COPPE/UFRJ, 2000.

RAFTERY, A. E.; GNEITING, T.; BALABDAOUI, F.; POLAKOWSKI, M. Using Bayesian model averaging to calibrate forecast ensembles. *Monthly Weather Review*, v. 133, n. 5, p. 1155-1174, 2005.

RENARD, B.; KAVETSKI, D.; KUCZERA, G.; THYER, M.; FRANKS, S. W. Understanding predictive uncertainty in hydrologic modeling: The challenge of identifying input and structural errors. *Water Resources Research*, v. 46, n. 5, 2010.

ROBERT, C. P.; CASELLA, G. Monte Carlo Statistical Methods. New York: Springer, 2nd ed., 2004, 645 p.

ROMAN, V. Análise do equilíbrio morfológico em rios regularizados: modelagem matemática unidimensional do transporte de sedimentos no rio São Francisco. Dissertação (Mestrado em Saneamento, Meio Ambiente e Recursos Hídricos) - Departamento de Hidráulica e Recursos Hídricos, Universidade Federal de Minas Gerais, Belo Horizonte - MG, 2004.

SALTELLI, A.; TARANTOLA S.; CAMPOLONGO F.; RATTO. M. Sensitivity Analysis in Practice: a Guide to Assessing Scientific Models. John Wiley & Sons, 2004.

SAHOO S. N.; SREEJA P. Development of flood inundation maps and quantification of flood risk in an urban catchment of Brahmaputra River. *ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part A: Civil Engineering*, v. 3, n. 1, 2015.

SCHARNAGL, B.; IDEN, S.C.; DURNER, W.; VEREECKEN, H.; HERBST, M. Inverse modelling of in situ soil water dynamics: Accounting for heteroscedastic, autocorrelated, and non-Gaussian distributed residuals. *Hydrology and Earth System Sciences Discussions*, v. 2, p. 2155-2199, feb. 2015.

SILVA, F. E.; NAGHETTINI, M.; FERNANDES, W. Avaliação bayesiana das incertezas nas estimativas dos parâmetros de um modelo chuva-vazão conceitual. RBRH – Revista Brasileira de Recursos Hídricos, v. 19, n. 4, p. 148-159, 2014.

SILVA, A. T.; PORTELA, M. M.; NAGHETTINI, M.; FERNANDES, W. A Bayesian peaks-over-threshold analysis of floods in the Itajaí-açu river under stationarity and nonstationarity. Stochastic environmental research and risk assessment, v. 31, n. 1, p. 185-204, 2017.

SIMONS, D. B.; RICHARDSON, E. V. Resistance to flow in alluvial channels. Physiographic and Hydraulic Studies of Rivers - Geological Survey Professional Paper, 422-J. U.S. Government Printing Office, Washington, 1966.

SMEMOE, C. M.; NELSON, E. J.; ZUNDEL, A. K.; MILLER, A. W. Demonstrating floodplain uncertainty using flood probability maps. Journal of the American Water Resources Association, v. 43, n. 2, p. 359-371, 2007.

SCHOUPS, G.; VRUGT, J.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, 2010.

STEDINGER, J.R.; VOGEL, R.M.; LEE, S.U.; BATCHELDER, R. Appraisal of the generalized likelihood uncertainty estimation (GLUE) method. Water Resources Research, v. 44, n. 12, 2008.

STEPHENS, M. A. The Anderson-Darling statistic. Technical Report n. 31. Stanford: Stanford University, Department of Statistics, 1979.

STEPHENS, E. M.; BATES, P. D.; FREER, J. E.; MASON, D. C. The impact of uncertainty in satellite data on the assessment of flood inundation models. Journal of Hydrology, v. 414, p. 162-173, 2012.

THYER, M.; RENARD, B.; KAVETSKI D.; KUCZERA, G.; FRANKS, S. W.; SRIKANTHAN, S. Critical evaluation of parameter consistency and predictive uncertainty in hydrological modeling: A case study using Bayesian total error analysis. Water Resources Research, v. 45, n. 12, 2009.

TRADICIONAL TOPOGRAFIA E ENGENHARIA. Relatório técnico conclusivo do levantamento topobatimétrico realizado a jusante da barragem Três Marias, nos rios São Francisco e Abaeté no município de Três Marias – MG. 2011.

TODINI, E. A model conditional processor to assess predictive uncertainty in flood forecasting. International Journal of River Basin Management, v. 6, n. 2, p. 123-37, 2008.

TSUBAKI, R.; KAWAHARA, Y. The uncertainty of local flow parameters during inundation flow over complex topographies with elevation errors. Journal of Hydrology, v. 486, p. 71-87, 2013.

UFSC – CEPED: Universidade Federal de Santa Catarina: Centro Universitário de Estudos e Pesquisas sobre Desastres. Relatório de Danos Materiais e Prejuízos Decorrentes de Desastres Naturais no Brasil: 1995-2014. Florianópolis: CEPED UFSC, 2016.

UFRGS-IPH – Universidade Federal do Rio Grande do Sul: Instituto de Pesquisas Hidráulicas. Diagnóstico das Condições Sedimentológicas dos Principais Rios Brasileiros. Rio de Janeiro: ELETROBRÁS, 2ª Ed., 1998, 100 p.

VOGEL, R. M.; SANKARASUBRAMANIAN, A. Validation of a watershed model without calibration. *Water Resources Research*, v. 39, n. 10, 2003.

VRUGT, J.A.; GUPTA, H.V.; BOUTEN, W.; SOROOSHIAN, S. A Shuffled Complex Evolution Metropolis algorithm for optimization and uncertainty assessment of hydrologic model parameters. *Water resources research*, v. 39, n. 8, 2003.

VRUGT, J. A.; ROBINSON B. A. Treatment of uncertainty using ensemble methods: Comparison of sequential data assimilation and Bayesian model averaging. *Water Resources Research*, v. 43, n. 1, 2007.

VRUGT, J.A.; TER BRAAK, C.J.; CLARK, M.P.; HYMAN, J.M.; ROBINSON, B.A. Treatment of input uncertainty in hydrologic modeling: Doing hydrology backward with Markov chain Monte Carlo simulation. *Water Resources Research*, v. 44, n. 12, dec. 2008.

VRUGT, J.A.; TER BRAAK, C.J.; GUPTA, H.V.; ROBINSON, B.A. Equifinality of formal (DREAM) and informal (GLUE) Bayesian approaches in hydrologic modeling? *Stochastic environmental research and risk assessment*, v. 2, n.7, p.1011-1026, 2009a.

VRUGT, J.A.; TER BRAAK, C.J.F.; DIKS, C.G.H.; ROBINSON, B.A.; HYMAN, J.M.; HIGDON, D. 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, 2009b.

VRUGT, J.A. Markov chain Monte Carlo simulation using the DREAM software package: Theory, concepts, and MATLAB implementation. *Environmental Modelling & Software*, v. 75, p. 273-316, 2016.

WERNER, M. G. A comparison of flood extent modelling approaches through constraining uncertainties on gauge data. *Hydrology and earth system sciences*, v. 8, n. 6, p. 1141-1152, 2004.

WOOLDRIDGE, J. M. Introdução à Econometria: uma Abordagem Moderna. São Paulo, SP: Cengage Learning, 2008, 848 p.

WMO – World Meteorological Organisation. Integrated Flood Management – concept paper. Geneva, 2009. Disponível em https://www.floodmanagement.info/publications//concept_paper_e.pdf. Acesso em 19 de setembro de 2019.