

**UNIVERSIDADE DE BRASÍLIA
FACULDADE DE TECNOLOGIA
DEPARTAMENTO DE ENGENHARIA CIVIL E AMBIENTAL**

**MODELO BAYESIANO COMPLETO PARA ANÁLISE DE
FREQUÊNCIA DE CHEIAS COM INCORPORAÇÃO DO
CONHECIMENTO HIDRÁULICO NA MODELAGEM DAS
INCERTEZAS NA CURVA-CHAVE**

ANA LUISA NUNES DE ALENCAR OSORIO

ORIENTADOR: DIRCEU SILVEIRA REIS JUNIOR

**DISSERTAÇÃO DE MESTRADO EM TECNOLOGIA
AMBIENTAL E RECURSOS HÍDRICOS**

BRASÍLIA/DF: FEVEREIRO/2017

**UNIVERSIDADE DE BRASÍLIA
FACULDADE DE TECNOLOGIA
DEPARTAMENTO DE ENGENHARIA CIVIL**

**MODELO BAYESIANO COMPLETO PARA ANÁLISE DE
FREQUÊNCIA DE CHEIAS COM INCORPORAÇÃO DO
CONHECIMENTO HIDRÁULICO NA MODELAGEM DAS
INCERTEZAS NA CURVA-CHAVE**

ANA LUISA NUNES DE ALENCAR OSORIO

**DISSERTAÇÃO DE Mestrado submetida ao
DEPARTAMENTO DE ENGENHARIA CIVIL E AMBIENTAL DA
FACULDADE DE TECNOLOGIA DA UNIVERSIDADE DE BRASÍLIA
como parte dos requisitos necessários para a
obtenção do grau de Mestre em Tecnologia
Ambiental e Recursos Hídricos.**

APROVADA POR:

**Prof. Dirceu Silveira Reis Júnior, PhD (ENC-UnB)
(Orientador)**

**Prof.
Carlos Henrique Ribeiro Lima, PhD (UnB)
(Examinador Interno)**

**Prof. Eduardo Sávio Passos Rodrigues Martins, PhD (FUNCEME)
(Examinador Externo)**

BRASÍLIA/DF, 17 DE FEVEREIRO DE 2017.

FICHA CATALOGRÁFICA

OSORIO, ANA LUISA NUNES DE ALENCAR

Modelo Bayesiano Completo para análise de frequência de cheias com incorporação do conhecimento hidráulico na modelagem das incertezas na curva-chave [Distrito Federal] 2017. xvi, 161 p., 210 x 297 mm (ENC/FT/UnB, Mestre, Tecnologia Ambiental e Recursos Hídricos, 2017).

Dissertação de Mestrado – Universidade de Brasília. Faculdade de Tecnologia.

Departamento de Engenharia Civil e Ambiental.

1. Curva-chave

2. Inferência Bayesiana

3. Estudo de Frequência de Cheias

4. Incerteza

I. ENC/FT/UnB

II. Título (série)

REFERÊNCIA BIBLIOGRÁFICA

OSORIO, A. L. N. A. (2017). Modelo Bayesiano Completo para análise de frequência de cheias com incorporação do conhecimento hidráulico na modelagem das incertezas na curva-chave. Dissertação de Mestrado em Tecnologia Ambiental e Recursos Hídricos, Publicação PTARH.DM-196/17, Departamento de Engenharia Civil e Ambiental, Universidade de Brasília, Brasília, DF, 161 p.

CESSÃO DE DIREITOS

AUTOR: Ana Luisa Nunes de Alencar Osorio.

TÍTULO: Modelo Bayesiano Completo para análise de frequência de cheias com incorporação do conhecimento hidráulico na modelagem das incertezas na curva-chave.

GRAU: Mestre

ANO: 2017

É concedida à Universidade de Brasília permissão para reproduzir cópias desta dissertação de mestrado e para emprestar ou vender tais cópias somente para propósitos acadêmicos e científicos. O autor reserva outros direitos de publicação e nenhuma parte dessa dissertação de mestrado pode ser reproduzida sem autorização por escrito do autor.

Ana Luisa Nunes de Alencar Osorio
osorio.analuisa@gmail.com

À minha família

“Never give up and good luck will find you.”

Falcor - História sem fim (1984).

AGRADECIMENTOS

Agradeço, primeiramente, ao meu orientador Dirceu Reis, por todos os ensinamentos e conselhos que resultaram em melhorias nesse trabalho e na minha formação acadêmica. Agradeço a ele o desafio e a confiança depositada.

Agradeço também aos pesquisadores Jérôme Le Coz e Benjamin Renard do Instituto Nacional de Pesquisa em Ciência e Tecnologia da França (IRSTEA) pela atenção e constante disposição para discutir e dirimir dúvidas sobre o trabalho desenvolvido por eles em Análise de Incertezas de Curvas-Chaves.

Ao Engenheiro Cassio Rampinelli pelo apoio com as rotinas e programação no R que tornou possível a implementação do método apresentado nesse trabalho.

Aos professores Eduarno Nakano e Carlos Henrique Lima pelos ensinamentos em Inferência Bayesiana e apoio na verificação da abordagem proposta.

À equipe do DNIT, em especial aos chefes e amigos Edimarques Magalhães, Paula Passos, Galileu Santos, Fernando Fontes e Renato Amorim que me deram as condições necessárias para me dedicar ao mestrado e apoio nos momentos de sobrecarga.

À Beth Faber, Calvin Creech e Will Vetch do Corpo de Engenheiros dos Estados Unidos (USACE) e aos amigos Vanessa da Matta e Leonardo Zandonadi, pelos conselhos e comentários sobre o trabalho.

À minha família, e principalmente meus pais pelo amor e apoio incondicionais.

E finalmente ao meu marido, Felipe Castañon pelo apoio na otimização dos códigos desenvolvidos e pela paciência e amor constantes durante esse período de longas noites e reduzidos finais de semana. Sem ele eu não teria chegado até aqui.

RESUMO

A análise de frequência de cheias, utilizada na avaliação de riscos de inundação e no dimensionamento de estruturas hidráulicas, consiste no ajuste estatístico de uma distribuição teórica de probabilidade aos dados históricos de vazão. Na maioria das vezes, as vazões não são diretamente medidas em campo, mas estimadas a partir do histórico de cotas máximas de nível d'água e da curva-chave, que é uma função que tenta representar a relação estabelecida entre a cota de nível d'água e a vazão para a seção do rio em estudo. Como a curva-chave é apenas uma aproximação da real relação entre cota e vazão, os dados de vazão são altamente influenciados pela forma como essa curva é estabelecida e, portanto, sujeitos a incertezas. No caso da análise de cheias, essas incertezas podem ser ainda maiores, já que as vazões de pico são usualmente estimadas a partir da extrapolação da curva onde se tem pouca ou nenhuma informação disponível. A compreensão e melhor descrição dessas incertezas permite uma estimativa mais realista dos quantis estimados na análise de frequência de cheias. O uso da inferência Bayesiana é atraente nesses casos, pois o conhecimento das características hidráulicas do rio, assim como a modelagem das incertezas envolvidas, pode ser incorporado à análise de uma maneira natural, permitindo que se considere as incertezas nos dados de vazão na estimativa dos quantis.

Esse estudo propõe um Modelo Bayesiano Completo (BC), que incorpora o conhecimento hidráulico sobre o trecho do rio em questão na estimativa das incertezas inerentes ao ajuste da curva-chave, e considera essas incertezas na obtenção da curva de frequência de cheias. O modelo é aplicado em um caso sintético com parâmetros conhecidos e em um caso real na estação fluviométrica de Cáceres no rio Paraguai, com intuito de avaliar o impacto dessas incertezas em diferentes cenários de disponibilidade de dados e conhecimento sobre o trecho do rio. Os resultados obtidos mostram que o modelo proposto é adequado para estimativa da distribuição *a posteriori* dos parâmetros da curva-chave, da distribuição teórica de probabilidades e dos quantis e seus intervalos de credibilidade. O estudo demonstra que as incertezas da curva-chave não são desprezíveis e podem chegar até 70% do valor da incerteza global, e têm um impacto significativo na incerteza nos quantis de cheia. Além disso, os resultados mostram que a disponibilidade de dados de medição de descarga e cotas máximas anuais, assim como o conhecimento hidráulico sobre o trecho do rio podem alterar essas incertezas, indicando o importante papel dos levantamentos de campo na melhoria das estimativas.

ABSTRACT

Flood frequency analysis is developed to provide essential information for flood risk management and for the design of hydraulic structures. It consists of adjusting a theoretical probability distribution to a sample of streamflow data in order to relate flood magnitudes to their probability of exceedance. Daily streamflow records, however, are not actually measured in the field, but estimated based on daily measurements of water level in conjunction with the rating curve, a mathematical function that relates gauge readings with discharge values at a given station. As the rating curve is only an approximation of the real relationship between water levels and discharge values, daily streamflow data contain uncertainties. In flood frequency analysis, these uncertainties can be even larger because a relatively large portion of the data is estimated based on the extrapolation of the rating curve. The comprehension and quantitative assessment of these uncertainties are important to obtain a more realistic description of the uncertainties in flood quantile estimates. Bayesian inference is very attractive in these cases because it can easily incorporate the often imprecise knowledge available on the hydraulic behavior of the river into the flood frequency analysis, providing a natural way to not only evaluate the uncertainties in the streamflow sample, but also to consider these uncertainties in the estimated flood quantiles.

This dissertation develops a fully Bayesian model (BC) capable of incorporating imprecise knowledge on the hydraulic behavior of the river, when available, to estimate the uncertainties in the rating curve and to consider these uncertainties in the development of the flood frequency curve. The proposed model was applied in a synthetic case with known parameters and in a real case with data from a streamflow gauge located in the Paraguai River. These cases were used to evaluate the impacts of uncertainties in the rating curve in the flood quantile estimates considering different scenarios of record length, amount of discharge measurements and quality of information on the hydraulic behavior of the river. Results show the proposed fully Bayesian model provided adequate posterior distributions of parameters of the rating curve, the probability distribution of annual maximum floods, and flood quantiles. Estimated uncertainties in the rating curve were quite large, representing as much as 70% of the uncertainties in flood quantiles in extreme cases. Results also show that the amount of discharge measurements and prior knowledge on the hydraulic behavior of the river can have an impact on the uncertainties in the flood quantiles, indicating the potential role of field survey and discharge measurement in the reduction of uncertainties on flood quantile.

SUMÁRIO

RESUMO	VII
ABSTRACT	VIII
1 INTRODUÇÃO	1
2 OBJETIVOS	4
2.1 OBJETIVO GERAL	4
2.2 OBJETIVOS ESPECÍFICOS.....	4
3 REVISÃO BIBLIOGRÁFICA	5
3.1 ESTUDOS DE FREQUÊNCIA DE CHEIAS	5
3.1.1 Distribuição Generalizada de Valores Extremos (GEV).....	7
3.1.2 Estimativa do Intervalo de Confiança	10
3.2 OBTENÇÃO DE SÉRIES HISTÓRICAS DE VAZÃO	11
3.2.1 Curva-chave.....	11
3.2.2 Obtenção da Curva-chave.....	16
3.2.3 Extrapolação de Curvas-Chaves.....	17
3.3 ANÁLISE DE INCERTEZAS EM CURVAS-CHAVES	20
3.3.1 Fontes de Incertezas.....	20
3.3.2 Métodos Para Estimativa de Incertezas em Curvas-chaves.....	22
3.4 IMPACTO DAS INCERTEZAS DE CURVAS-CHAVES EM ESTUDO DE CHEIAS	23
3.5 INFERÊNCIA BAYESIANA.....	28
3.5.1 Aplicação em Análises de Incertezas em Curvas-chaves.....	29
3.5.2 Aplicação em Estudos de Frequência de Cheias	32
3.5.3 Aplicação em Estudos de Frequência de Cheias com incorporação de incertezas das Curvas-Chaves	34
3.6 SIMULAÇÕES MONTE CARLO POR CADEIAS DE MARKOV	37
3.6.1 Algoritmo Metropolis	38
3.6.2 Algoritmo Metropolis Hastings	39
3.6.3 Algoritmo Metropolis Adaptado (AM)	39
3.6.4 Verificação da Convergência	41
4 METODOLOGIA	44
4.1 ESTRATÉGIA DE AÇÃO	44
4.2 MODELO BAYESIANO COMPLETO DE INCORPORAÇÃO DE INCERTEZAS DA CURVA-CHAVE - BC	45

4.2.1	Função de Verossimilhança.....	47
4.2.2	Distribuições <i>a priori</i>	48
4.2.3	Distribuição <i>a posteriori</i> dos Parâmetros e dos Quantis de Cheia	52
4.3	MODELO BAYESIANO PARA GEV – BG.....	52
4.3.1	Função de Verossimilhança.....	53
4.3.2	Distribuições <i>a priori</i>	53
4.3.3	Distribuição <i>a Posteriori</i> dos Parâmetros e dos Quantis de Cheia.....	53
4.4	IMPACTO DAS INCERTEZAS DA CURVA-CHAVE NA ANÁLISE DE FREQUÊNCIA DE CHEIAS	54
4.4.1	Métrica para avaliação – Importância Relativa.....	54
4.4.2	Fatores Avaliados	56
4.5	ESTUDOS DE CASO.....	57
4.5.1	Caso com dados sintéticos.....	57
4.5.2	Caso Real - Cáceres.....	61
5	RESULTADOS E DISCUSSÃO.....	68
5.1	ESTUDO DE CASO SINTÉTICO	68
5.1.1	Resultados Gerais para o Modelo Bayesiano Completo (BC)	68
5.1.2	Resultados Gerais para o Modelo Bayesiano para GEV (BG).....	75
5.1.3	Análise da Importância Relativa.....	79
5.2	ESTUDO DE CASO – CÁCERES.....	88
5.2.1	Resultados Gerais para o Modelo Bayesiano Completo (BC)	88
5.2.2	Resultados Gerais para o Modelo Bayesiano para GEV (BG).....	94
5.2.3	Análise da Importância Relativa.....	97
6	CONCLUSÕES E RECOMENDAÇÕES	105
	REFERÊNCIAS BIBLIOGRÁFICAS	108
	APÊNDICE I – ROTINAS DOS MODELOS BC E BG	113
	APÊNDICE II – DISTRIBUIÇÕES A <i>POSTERIORI</i> COMPARATIVA DOS CENÁRIOS	

LISTA DE TABELAS

Tabela 3-1 - Fórmulas Hidráulicas clássicas para controles hidráulicos mais comuns (Le Coz <i>et al</i> 2014, modificada).....	15
Tabela 4-1 – Resumo das informações <i>a priori</i> dos parâmetros	50
Tabela 4-2 Distribuições <i>a priori</i> dos Parâmetros da GEV no modelo BG	53
Tabela 4-3 – Cenários com variação dos fatores avaliados e legenda adotada	57
Tabela 4-4 – Valor dos Parâmetros adotados para o caso sintético	58
Tabela 4-5 – Valores do parâmetro <i>a</i> para cada cenário	63
Tabela 5-1 – Valores estimados e reais dos parâmetros para o cenário sintético pelo modelo Bayesiano Completo (BC).....	72
Tabela 5-2 – Valores estimados e reais dos parâmetros para o cenário sintético pelo modelo Bayesiano da GEV (BG)	75
Tabela 5-3 - IG, IA e IR para os cenários PDH e PD-H – Estudo de Caso Sintético	81
Tabela 5-4 – IG, IA e IR para os cenários PDH e P-DH – Estudo de Caso Sintético.....	82
Tabela 5-5 Cálculo da Importância Relativa (%) para os cenários PD-H e PD-H- - Caso Sintético.....	87
Tabela 5-6 Cálculo da Importância Relativa (%) para os cenários PDH e P-DH – Caso Real	99
Tabela 5-7 - Cálculo da Importância Relativa (%) para os cenários PDH e PD-H – Caso Real	100
Tabela 5-8 - Cálculo da Importância Relativa (%) para os cenários PDH e PDH- (Caso Real)	101

LISTA DE FIGURAS

Figura 3-1 - Exemplo de Curva-chave com Dois segmentos (Boiten 2003, modificada)...	14
Figura 3-2 - Curva-chave com Laço na Passagem da Cheia (Boiten 2003, modificado) ...	15
Figura 3-3 – Extrapolação da Curva-Chave pelo método de Stevens (Sefione, 2002)	18
Figura 3-4 – Exemplo de Extrapolação pelo Método de Manning (Sefione, 2002).....	19
Figura 3-5 – Modelo de erros proposto por Kuczera (1996).....	25
Figura 3-6 Visão geral da estrutura proposta por Steinbakk <i>et al.</i> (2016). <i>D</i> representa os dados de resumo de descarga utilizados para estimativa dos parâmetros da curva-chave (θ). <i>H</i> representa a séries de cotas diárias. A partir de <i>H</i> e θ , são calculadas as vazões máximas anuais normalizadas pela área de drenagem (Y_i). β representa os parâmetros da distribuição de frequências, no caso , da GEV. <i>j</i> representa o número de iterações da GEV (3000) e <i>x_p</i> representa os quantis de cheia obtidos a partir da função do quantil $F - 1$	36
Figura 3-7 – Comparação entre sequências geradas pelo algoritmo Metropolis Hasting baseadas em distribuições propostas diferentes (Casella, 2010).	41
Figura 3-8 - Comparativo entre diferentes cadeias de Markov (Martinez & Martinez, 2002)	42
Figura 3-9 - Comportamento da média ao longo das iterações da cadeia, indicando sinais de convergência (esquerda) e não houve convergência(direita) (Rampinelli,2016)	42
Figura 4-1 - Fluxograma de Atividades.....	45
Figura 4-2 – Modelo Bayesiano completo proposto	46
Figura 4-3 – Procedimento de cálculo dos Intervalos de Credibilidade da incerteza global (IG) e (IA) pelo método Bayesiano completo (BC) e bayesiano apenas para GEV (BG) ..	55
Figura 4-4 Curva-Chave de parâmetros $a = 90$, $b = 100$ e $c = 1,67$ com dados gerados sinteticamente com erro de medição e de ajuste do modelo e eliminados para o cenário D- (pontos vermelhos)	60
Figura 4-5 Dados de cotas máximas anuais gerados sinteticamente com os dados removidos (pontos vermelhos) para o cenário H-, com indicação da maior cota máxima registrada no resumo de descarga (linha tracejada vermelha).....	60
Figura 4-6 – Dados de resumo de descarga (D) da Estação Fluviométrica de Cáceres com pontos vermelhos removidos para o cenário de disponibilidade reduzida de resumo de descarga (D-)	61

Figura 4-7 – Dados cotas máximas anuais da Estação Fluviométrica de Cáceres de 1974 a 2015 (H), máxima cota observada no resumo de descarga (linha tracejada) e pontos removidos para o cenário de disponibilidade reduzida H- (pontos vermelhos)	62
Figura 4-8 Inclinações instantâneas da lâmina d’água entre as referências de nível na hidrovia do rio Paraguai – Em vermelho o trecho onde se localiza a seção em estudo de inclinação de 2,9 cm/km Fonte: UFPR/ITTI,2015	64
Figura 4-9 – Perfis batimétricos medidos em 2011 (pontos vermelhos),2014 (Linha cheia azul) e 2015 (linha tracejada preta) com indicação de nível d’água provável da data de medição do ADCP. Fonte: Hidroweb (2016)	64
Figura 4-10 - Levantamento por ADCP – Fonte: UFPR/ITTI(2015).....	65
Figura 5-1 – Cadeias de Markov amostradas pelo algoritmo Adaptive Metropolis para 100.000 iterações pelo modelo Bayesiano Completo para todos os parâmetros do cenário sintético PDH das cadeias 00 (preto) e 01 (vermelho), valor médio da cadeia sem o aquecimento (linha cheia) e o valor real (linha tracejada) – Caso Sintético	69
Figura 5-2 – Resultado do valor de R médio (linha cheia) e com limite superior do intervalo de 95% de confiança (linha tracejada) obtido pelo teste de Convergência de Gelman e Rubin (1992) para o Cenário PDH – Caso Sintético.....	70
Figura 5-3 – Comparação entre as distribuições <i>a priori</i> (linha cheia) e <i>a posteriori</i> (linha tracejada) dos parâmetros com reta vertical do valor real dos parâmetros – Estudo de Caso Sintético.....	71
Figura 5-4 - Comparativos das curvas-chaves real (linha cheia) e média estimada (linha tracejada central) pelo método Bayesiano Completo (BC) com intervalo de credibilidade de 95% (hachura) e medições de descarga sintéticas com respectivos erros de medição (barra de erros) – Caso Sintético.....	72
Figura 5-5- Série de vazões máximas anuais com valores médios (Linha cheia), intervalo de credibilidade (hachura) e dados de vazões máximas reais (pontos) – Caso Sintético.....	74
Figura 5-6 – Distribuições <i>a posteriori</i> dos quantis de cheia para os tempos de retorno de 10, 25, 50 e 100 anos com indicação do intervalo de credibilidade (linha tracejada) e valor real do quantil (linha cheia) – Caso Sintético.....	74
Figura 5-7 – Curva de frequência de cheias estimada (linha tracejada) e real (linha cheia), com intervalo de credibilidade de 95% (hachura), com os dados de vazão gerados e intervalo de credibilidade de 95% representado pela barra de erros – Caso Sintético	75

Figura 5-8 – Cadeias estimadas para 40.000 iterações pelo modelo Bayesiano da GEV (BG) para todos os parâmetros do cenário sintético PDH das cadeias 00 (pretas) e 01 (vermelhas) com o valor médio da cadeia (linhas cheias) e o valor real (tracejadas) – Caso Sintético..	76
Figura 5-9 – Resultado do valor de R médio (linha cheia) e limite superior do intervalo de 95% de confiança (linha tracejada) obtido pelo teste de Convergência de Gelman e Rubin (1992) para o Cenário PDH com indicação dos limites de 1 e 1,2 – Caso Sintético.....	76
Figura 5-10 – Comparação entre as distribuições <i>a priori</i> (linha cheia) e <i>a posteriori</i> (linha tracejada) dos parâmetros com reta vertical do valor real dos parâmetros para o modelo BG – Caso Sintético	77
Figura 5-11 – Distribuições <i>a posteriori</i> dos quantis de cheia para os tempos de retorno de 10, 25, 50 e 100 anos com indicação do intervalo de credibilidade (linha tracejada) e valor real do quantil (linha cheia) para o modelo Bayesiano da GEV (BG) – Caso Sintético.....	78
Figura 5-12 - Intervalos IG e IA das curvas de frequência de cheias para BG (hachurado) e BC (linhas tracejadas) para o cenário PDH com vazões máximas geradas pelos parâmetros da curva-chave de maior valor da <i>posteriori</i> (pontos) – Caso Sintético	79
Figura 5-13 - Valor de IR para todos os cenários e tempos de retorno de 10, 25, 50 e 100 anos – Estudo de Caso Sintético.....	80
Figura 5-14 – Comparativo da importância relativa para PDH, P-DH, PD-H e P-D-H (Caso Sintético).....	80
Figura 5-15 - Comparativo das distribuições <i>a priori</i> (linha cheia) e <i>a posteriori</i> (linha tracejada) de todos os parâmetros dos cenários PDH (preto tracejado) e PD-H (vermelho tracejado) com indicação do valor real (linha vertical) – Caso Sintético.....	84
Figura 5-16 - Comparativo das distribuições <i>a priori</i> (linha cheia) e <i>a posteriori</i> (linha tracejada) de todos os parâmetros dos cenários PDH (preto tracejado) e P-DH (vermelho tracejado) com indicação do valor real (linha vertical) – Caso Sintético.....	85
Figura 5-17 – Comparativo das distribuições <i>a posteriori</i> dos quantis de tempo de retorno de 10 e 100 anos para os cenários PDH e PD-H com a indicação dos respectivos intervalos de credibilidade (linhas cheias verticais) – Caso Sintético.....	86
Figura 5-18 – Comparativo das distribuições <i>a posteriori</i> dos quantis de tempo de retorno de 10 e 100 anos para os cenários PDH e P-DH com a indicação dos respectivos intervalos de credibilidade (linhas cheias verticais) – Caso Sintético.....	86
Figura 5-19 Comparativo da importância relativa para variações de H para H- (Caso Sintético).....	86

Figura 5-20 – Cadeias de Markov para 100.000 iterações pelo modelo Bayesiano Completo para todos os parâmetros do cenário sintético PDH das cadeias 00 (preto) e 01 (vermelho), valor médio (linha cheia) e o valor real (linha tracejada) para a Estação de Cáceres.	89
Figura 5-21 - Resultado do valor de R médio (linha cheia) e limite superior do intervalo de 95% de confiança (linha tracejada) obtido pelo teste de Convergência de Gelman e Rubin (1992) para o Cenário PDH com indicação dos limites de 1 e 1,2 – Caso Real - Cáceres .	90
Figura 5-22 – Comparação entre as distribuições <i>a priori</i> (linha cheia) e <i>a posteriori</i> (linha tracejada) dos parâmetros para o modelo BC – Estudo de caso real em Cáceres	91
Figura 5-23 –Comparativos das curvas-chaves real (linha cheia) e estimada (linha tracejada central) pelo método Bayesiano Completo (BC) com intervalo de credibilidade de 95% (hachura) e medições de descarga observadas com erro de medição (barra de erros) – Caso Real - Cáceres.....	92
Figura 5-24 Série de vazões máximas anuais. Linha cheia corresponde aos valores médios e hachura com linha tracejada o intervalo de credibilidade de 95 % - Caso Real - Cáceres..	92
Figura 5-25 – Distribuições <i>a posteriori</i> dos quantis de cheia para os tempos de retorno de 10, 25, 50 e 100 anos com indicação do intervalo de credibilidade (linha tracejada) – Modelo BC – Caso Real	93
Figura 5-26 – Curva de frequência de cheias estimada pelo método bayesiano completo (BC) com intervalo de credibilidade de 95% para os dados da Estação de Cáceres.....	93
Figura 5-27 – Cadeias de Markov com 40.000 iterações pelo modelo Bayesiano da GEV (BG) para todos os parâmetros do cenário real PDH das cadeias 00 (preto) e 01 (vermelho), valor médio da cadeia sem o aquecimento (linha cheia) - Estudo de Caso – Cáceres	94
Figura 5-28 – Resultado do Teste de Convergência de Gelman e Rubin(1992) para o Modelo BG– Cenário PDH com indicação dos limites de 1 e 1,2 – Estudo de Caso – Cáceres	95
Figura 5-29 – Comparação entre as distribuições <i>a priori</i> (linha cheia) e <i>a posteriori</i> (linha tracejada) dos parâmetros para o modelo BC – Estudo de caso - Cáceres.....	95
Figura 5-30 – Distribuições <i>a posteriori</i> dos quantis de cheia para os tempos de retorno de 10, 25, 50 e 100 anos com indicação do intervalo de credibilidade (linha tracejada) – Modelo BG – Caso Real	96
Figura 5-31 – Intervalos IG e IA das curvas de frequência de cheias para BG (hachurado) e BC (linhas tracejadas) para o cenário PDH com vazões máximas geradas pelos parâmetros de valor de máxima <i>posteriori</i> (pontos) – Caso Real	97
Figura 5-32 – Valor de IR para todos os cenários e tempos de retorno de 10, 25, 50 e 100 anos – Estudo de Caso - Cáceres.....	98

Figura 5-33 – Comparativo da importância relativa para PDH, P-DH, PD-H e P-D-H – Estudo de Caso - Cáceres	98
Figura 5-34 - Comparativo das distribuições <i>a priori</i> (linha cheia) e <i>a posteriori</i> (linha tracejada) dos cenários PDH (preto) e P-DH (vermelho, preto quando <i>a priori</i> for a mesma) com indicação do intervalo HPD nas distribuições dos quantis (linhas verticais) – Caso Real	102
Figura 5-35 Comparativo das distribuições <i>a priori</i> (linhas cheias) e <i>a posteriori</i> (linhas tracejadas) dos cenários PDH (preto) e PD-H (vermelho) com indicação do intervalo HPD nas distribuições dos quantis (linhas verticais) – Caso Real	103
Figura 5-36 - Comparativo das distribuições <i>a posteriori</i> dos quantis de tempo de retorno de 10 e 100 anos para os cenários PDH e P-DH com a indicação dos respectivos intervalos de credibilidade (linhas cheias verticais) – Caso Real	104
Figura 5-37 - Comparativo das distribuições <i>a posteriori</i> dos quantis de tempo de retorno de 10 e 100 anos para os cenários PDH e PD-H com a indicação dos respectivos intervalos de credibilidade (linhas cheias verticais) – Caso Real	104
Figura 5-38 Comparativo da importância relativa para variações de H para H- (Caso Real)	104

1 INTRODUÇÃO

A análise de frequência de cheias tem como objetivo estimar os quantis de vazões, ou cheias de projeto para uma dada probabilidade de excedência ou tempo de retorno. A estimativa desses quantis é descrita na forma de uma curva de frequência de cheias que é a base da análise de riscos de inundação e definição das vazões de projeto necessárias, por exemplo, para o dimensionamento de estruturas como barragens, pontes e rodovias próximas a cursos d'água e elaboração de mapas de inundação.

Esse estudo geralmente decorre da avaliação estatística dos dados históricos das vazões no local ou região, o que torna a disponibilidade e qualidade deles fundamentais nesse processo. Entretanto, em virtude da difícil logística e custo (Beven e Hall, 2014; Mason *et al.*, 2016), em regra, a disponibilidade desses dados é reduzida e eles não são medidos diretamente. Na realidade, as vazões são estimadas a partir dos dados históricos de cota de nível d'água e da curva-chave, uma função que relaciona a cota do nível d'água no rio com a vazão. Enquanto os dados de cota de nível d'água são medidos diretamente e diariamente em campo, a curva-chave é estimada com base nas medições simultâneas de cota e vazão e/ou nas características hidráulicas do rio, na tentativa de descrever na forma de uma equação a relação entre nível e vazão na seção do rio sob estudo. Ressalta-se ainda, que pelo difícil acesso e alto risco para o operador, as medições de vazão raramente são realizadas em momentos de cheia extrema.

Em função disso, a estimativa das vazões fica sujeita a diferentes fontes de incertezas que podem estar vinculadas principalmente aos erros nas medições (nível e/ou vazão) (Despax *et al.*, 2016, Mcmillan *et al.*, 2010), à qualidade do ajuste da curva (Le Coz *et al.*, 2014; Morlot *et al.*, 2014) e sua extrapolação, quando necessária (Kuczera, 1996).

Da mesma forma, também potencializam as incertezas em relação à curva-chave, a desconsideração das peculiaridades da seção em estudo como: as variações em relação ao seu regime de referência (ex: remanso, histeresis) (Petersen-Overleir, 2006; Petersen-Overleir, 2009a; Mansanarez *et al.*, 2016), a alteração de suas características ao longo do tempo (ex: rugosidade, largura, declividade) (Jalbert *et al.*, 2011; Morlot *et al.*, 2014) e/ou ao longo da própria seção (Petersen-Overleir, 2005; Petersen-Overleir, 2008; Coxon *et al.*, 2015).

Portanto, para compreensão do grau de confiabilidade dos dados de vazão e análises geradas a partir deles, essas incertezas deveriam ser estimadas e incorporadas de alguma

forma nas análises (Le Coz *et al.*, 2014; Moyeed e Clarke, 2005; Steinbakk *et al.*, 2016). Nesse sentido, estudos recentes se propuseram a entender e estimar as incertezas da curva-chave adotando diferentes metodologias para diferentes condições hidráulicas (Kuczera, 1996; Petersen-Overleir, 2004; Petersen-Overleir e Reitan, 2005; Petersen-Overleir, 2006; Petersen-Overleir, 2008; Petersen-Overleir e Reitan, 2009a; Di Baldassarre *et al.*, 2012; Reitan e Petersen, 2008; Le Coz *et al.*, 2014).

Métodos utilizados nesses estudos, como de análise de sensibilidade (Lang *et al.*, 2010, Di Baldassari e Claps, 2011; Domeneghetti *et al.*, 2012), de regressão não linear (Venetis, 1970; Dymonch e Christian, 1972; Hershy, 1999; Clarke *et al.*, 2000), de estatística clássica (máxima verossimilhança) (Kuczera, 1996; Petersen-Overleir, 2004; Petersen-Overleir e Reitan, 2005, Petersen-Overleir e Reitan, 2009a; Petersen-Overleir, 2006; Petersen-Overleir, 2008; Di Baldassarre *et al.*, 2012) e mais recentemente de inferência Bayesiana (Kuczera, 1999; Moyeed e Clark, 2005; Macmillan *et al.*, 2010; Le Coz *et al.*, 2014) comprovaram que o impacto da incorporação das incertezas na série das vazões era considerável.

Dentre as diversas metodologias adotadas, o uso da inferência Bayesiana se destaca por tornar possível que o conhecimento prévio sobre os parâmetros que se busca estimar (função *a priori*) e suas incertezas sejam combinados com as observações (função verossimilhança) para obtenção da distribuição *a posteriori* desses parâmetros, ou de qualquer função dos mesmos. Assim, a distribuição *a posteriori* permite inferir sobre o valor mais provável do parâmetro e sua credibilidade.

Em estudos de cheias, os parâmetros de interesse são aqueles da equação da curva-chave e da distribuição teórica de probabilidades adotada para determinar a curva de frequência de cheias. Técnicas Bayesianas tornam possível que as incertezas nas estimativas da curva-chave discutidas anteriormente e da distribuição teórica de probabilidades (incerteza amostral e de ajuste da distribuição) possam ser empregadas para descrever as incertezas nas estimativas finais dos quantis de cheia.

O Teorema de Bayes permite considerar explicitamente os conhecimentos hidráulicos em determinado trecho de rio (ou a falta deles), o que pode resultar numa estimativa mais adequada da curva de frequência de cheias e de suas respectivas incertezas (Le Coz *et al.*, 2014).

Apesar de ser uma excelente estratégia para lidar com as incertezas, a abordagem Bayesiana só começou a ser empregada em estudos dessa natureza mais recentemente com

o advento de metodologias de simulações numéricas como as simulações Monte Carlo por Cadeia de Markov (ex. Gelman *et. al.*, 2014) e com o desenvolvimento dos computadores.

Entretanto, a sua aplicação na análise de incertezas em estudos de frequência de cheias ainda é limitada. O que se observa na literatura é uma análise separada das incertezas na estimativa dos parâmetros da curva-chave (Reitan e Petersen Overleir, 2008; Reitan e Petersen-Overleir, 2009; Petersen-Overleir e Reitan, 2009; Le Coz *et al.*, 2012) e da incerteza nos estudos de frequência de cheias (Coles e Tawn, 1996; Kuczera, 1999; Reis e Stedinger, 2005; O’Connel, 2005; Neppel *et al.*, 2010). Poucos estudos se propuseram a analisar de forma completa a incorporação das incertezas da curva-chave na estimativa dos quantis (Reitan e Petersen-Overleir, 2009; Osorio e Reis, 2016; Steinbakk, 2016).

Diante do histórico apresentado, verifica-se o restrito conhecimento que se tem em relação à consideração das incertezas da curva-chave na obtenção das curvas de frequência de cheia. Esse estudo visa, portanto, a proposição de um modelo com uso da Inferência Bayesiana de maneira mais completa que em estudos anteriores. Busca-se desenvolver uma metodologia de análise de frequência de cheias com a incorporação das incertezas desde as estimativas dos parâmetros da curva-chave; e entender a importância dessas incertezas em relação à incerteza total associada aos quantis.

Este trabalho é dividido em 4 partes principais: **Objetivos**, onde são delineados os objetivos principais e específicos do estudo com base no histórico apresentado, **Revisão Bibliográfica**, onde são apresentados os estudos já desenvolvidos sobre o tema com suas vantagens e limitações e a base teórica utilizada neste trabalho, **Metodologia**, onde é indicado os passos que foram seguidos para alcançar os objetivos propostos, **Resultados**, **Discussão** e **Conclusões**, onde se apresentam os principais resultados e conclusões obtidos.

2 OBJETIVOS

2.1 Objetivo Geral

Com base na importância da obtenção das curvas de frequência de cheias e descrição de suas incertezas na tomada de decisões no âmbito da análise de risco, esse estudo se propõe a avaliar o impacto da incorporação das incertezas da curva-chave considerando explicitamente as características hidráulicas da seção do rio na análise de frequência de cheias por meio da proposição de uma metodologia Bayesiana de estimativa dos parâmetros.

Objetiva-se a obtenção de um modelo completo que permita entender a importância dessas incertezas da curva-chave na curva de frequência de cheias.

2.2 Objetivos Específicos

Buscando alcançar o objetivo geral, foram delineados os seguintes objetivos específicos:

1) Desenvolver um modelo de inferência Bayesiana que possibilite incorporar o conhecimento hidráulico sobre o trecho do rio em questão na estimativa das incertezas inerentes ao ajuste da curva-chave e considerar essas incertezas na obtenção da curva de frequência de cheias;

2) Avaliar a contribuição das incertezas da curva-chave em relação à incerteza global dos quantis de cheia estimados empregando casos sintéticos e reais;

3) Analisar como o grau de informação sobre os aspectos hidráulicos do trecho do rio em questão, a quantidade disponível de medições de descarga e o tamanho da série histórica de cotas máximas anuais afetam a importância relativa das incertezas da curva chave nas incertezas totais dos quantis de cheia estimados.

3 REVISÃO BIBLIOGRÁFICA

3.1 Estudos de Frequência de Cheias

O estudo de frequência de cheias tem como objetivo relacionar a magnitude de eventos extremos com sua frequência de ocorrência. A curva de frequência de cheias obtida a partir desses estudos é insumo fundamental para análise de riscos de inundação. Além disso, os valores dos quantis obtidos a partir dessa curva também são utilizados para a determinação da vazão de projeto (associada a um determinado risco) de diversas estruturas hidráulicas como pontes e barragens.

Quando há disponibilidade de dados na região, esse estudo é desenvolvido a partir da série de valores de vazão obtidas nas estações fluviométricas de monitoramento. A partir desse estudo, obtêm-se os valores da vazão para diferentes probabilidades de ocorrência ou tempos de retorno, em que o tempo de retorno representa o inverso da probabilidade de determinado evento ser igualado ou superado.

Além das informações dessas séries, podem ser utilizados também indícios de cheias históricas que ultrapassaram determinada cota do nível d'água cuja vazão é conhecida como marcas de cheia em árvores ou casas, ou informações obtidas em outras estações próximas. Essas estratégias não serão abordadas aqui mas podem ser vistas em detalhe em Reis e Stedinger (2005) e Kuczera (1999).

Para obtenção da curva de frequência de cheias e previsão dos quantis, a esses dados históricos deve ser ajustada uma distribuição teórica de probabilidades que melhor descreva a amostra observada. Entretanto, não existe uma distribuição específica consensual que seja capaz de, sob quaisquer condições, descrever o comportamento de eventos máximos anuais de vazão (Naghetini e Pinto, 2007), o que torna a escolha dessa distribuição e a estimativa dos parâmetros populacionais desse modelo uma tarefa não muito simples.

As distribuições mais utilizadas podem ser distinguidas entre distribuições de valores extremos, como é o caso das distribuições Gumbel, Fréchet, Weibull, todas casos especiais da distribuição Generalizada de Valores Extremos (GEV), e as ditas não extremas, como a Exponencial, Pearson III, Log-Pearson III e Log-Normal de dois parâmetros.

Além das diferentes distribuições, existem diversos métodos que podem ser utilizados para a estimativa dos respectivos parâmetros. Os mais comuns são os métodos dos

momentos, o método da máxima verossimilhança e o método dos momentos-L (Naghetini e Pinto, 2007).

Entretanto, para incorporação de incertezas da curva-chave ou informações adicionais devido a conhecimento prévios do local na estimativa dos parâmetros e seu intervalo de confiança, esses métodos se tornam limitados (Kuczera, 1999; Reis e Stedinger, 2005). Nesse sentido, o uso da Técnica de Inferência Bayesiana, como alternativa à inferência estatística clássica, se mostra bastante vantajosa (Coles e Tawn, 1996; Reis e Stedinger, 2005; O'Connell, 2005; Neppel *et al.*, 2010).

De forma sucinta, essa técnica consiste na inferência estatística dos parâmetros a partir da função verossimilhança e dos conhecimentos *a priori* acerca desses parâmetros, permitindo, portanto, a introdução das incertezas em relação ao ajuste da curva-chave associado ao conhecimento que se tem das características geométricas e de escoamento do rio. O que se obtém a partir dessa inferência, além do valor mais provável do parâmetro, é a sua distribuição *a posteriori*. Como o enfoque desse trabalho é a análise das incertezas em estudos de cheias por análise Bayesiana, essa metodologia e sua aplicação em estudo de frequência serão melhor detalhadas no item 3.5.

Independente da técnica a ser aplicada, a confiabilidade da determinação dos parâmetros de uma dada distribuição de probabilidade também está intrinsecamente ligada ao tamanho, qualidade da amostra e sua representatividade.

Infelizmente, as séries de vazão não são longas e o processo de produção das mesmas é duvidoso em função das diversas incertezas envolvidas. Esse processo e suas incertezas serão descritas no Item 3.2. Dessa maneira, é necessário que os estudos de frequência sejam desenvolvidos com cuidado, conhecendo as limitações das técnicas aplicadas. Neste trabalho, optou-se por adotar a distribuição de valores extremos para o ajuste dos valores de vazão no método de análise Bayesiana proposto por ser largamente utilizada no mundo todo, embora outras distribuições possam ser empregadas.

Dessa maneira, nesse capítulo será descrita a distribuição Generalizada de Valores Extremos (GEV) adotada e dois métodos possíveis para a estimativa dos parâmetros e seus intervalos de confiança: Método de Máxima Verossimilhança e o Método de Máxima Verossimilhança Generalizada (Martins e Stedinger, 2000). O segundo método será apresentado em função da sua eficiência demonstrada por Martins e Stedinger (2000) quando comparado com os métodos de máxima verossimilhança, momentos e momentos L para determinados valores do parâmetro de forma da GEV e tamanhos reduzidos de amostra, conforme descrito a seguir.

3.1.1 Distribuição Generalizada de Valores Extremos (GEV)

Em estudo de frequência de cheias em que o objetivo é a modelagem dos máximos anuais observados, a Distribuição Generalizada de Valores Extremos (GEV) se mostra bastante atrativa. Essa distribuição incorpora as três formas assintóticas de valores extremos máximos (Frechet, Weibull e Gumbel) em uma única expressão.

Sua função de distribuição acumuladas, conforme descrita e melhor detalhada por Naghettini e Pinto (2007), pode ser descrita como:

$$F_X(x) = \exp \left\{ - \left[1 - \kappa \left(\frac{x - \xi}{\alpha} \right) \right]^{1/\kappa} \right\} \quad \kappa \neq 0 \quad (3.1)$$

$$F_X(x) = \exp \left\{ - \exp \left[- \frac{(x - \xi)}{\alpha} \right] \right\} \quad \kappa = 0 \quad (3.2)$$

em que κ (kapa), α (alpha) e ξ (csi) representam, nessa ordem, os parâmetros de forma, escala e posição. As distribuições de Frechet e Weibull ocorrem, respectivamente, para valores negativos e positivos de κ , respeitando os intervalos de x apresentados nas inequações (3.3), (3.4). Quando $\kappa = 0$, tem-se a distribuição de Gumbel, em que x pode assumir qualquer valor (3.5).

$$\text{Frechet} \rightarrow k < 0 \rightarrow \xi + \frac{\alpha}{\kappa} \leq x < +\infty \quad (3.3)$$

$$\text{Weibull} \rightarrow k > 0 \rightarrow -\infty < x \leq \xi + \frac{\alpha}{\kappa} \quad (3.4)$$

$$\text{Gumbel} \rightarrow k = 0 \rightarrow -\infty < x < +\infty \quad (3.5)$$

A função densidade da distribuição GEV é expressa por:

$$f_x(x) = \frac{1}{\alpha} \left[1 - \kappa \left(\frac{x - \xi}{\alpha} \right) \right]^{\frac{1}{\kappa} - 1} \exp \left\{ - \left[1 - \kappa \left(\frac{x - \xi}{\alpha} \right) \right]^{\frac{1}{\kappa}} \right\} \quad \kappa \neq 0 \quad (3.6)$$

$$f_x(x) = \frac{1}{\alpha} \exp \left\{ \frac{x - \xi}{\alpha} - \exp \left(\frac{x - \xi}{\alpha} \right) \right\} \quad \kappa = 0 \quad (3.7)$$

Os parâmetros da distribuição podem ser estimados por diferentes métodos como o método dos momentos (Stedinger *et al.*, 1993), momentos L (Stedinger *et al.*, 1993), de máxima verossimilhança (Stedinger *et al.*, 1993) e de máxima verossimilhança generalizada

(Martins e Stedinger, 2000). A partir dos valores desses parâmetros, o quantil (x_p) que representa a estimativa de vazão para uma dada probabilidade de excedência (p) para a distribuição GEV pode ser calculado conforme indicados nas Equações (3.8) e (3.9).

$$x_p = \xi + \frac{\alpha}{\kappa} [1 - (-\ln(p))^\kappa] \quad \xi \neq 0 \quad (3.8)$$

$$x_p = \xi + \alpha [1 - (-\ln(p))] \quad \xi = 0 \quad (3.9)$$

No caso do método de máxima verossimilhança, os parâmetros ótimos referem-se ao ponto onde se maximiza a função (log) de verossimilhança. A verossimilhança corresponde à função de densidade de probabilidade conjunta das observações de uma amostra, sendo, portanto, quando se assume independência entre os eventos, o produto das funções densidades de cada uma das observações. Uma das formas para obtenção desse vetor de parâmetros ótimo consiste em derivar e igualar a zero a função de verossimilhança (ou log da verossimilhança) em relação a cada parâmetro. Obtém-se, a partir dessa derivação, um sistema de equações cuja solução corresponde aos valores ótimos dos parâmetros.

No caso da distribuição GEV, a função log de verossimilhança pode ser descrita pela Equação (3.10).

$$\begin{aligned} \ln[L(\theta|x)] = & -n \ln(\alpha) \\ & + \sum_{i=1}^n \left[\left(\frac{1}{\kappa} - 1 \right) \ln \left[1 - \left(\frac{\kappa}{\alpha} \right) (x - \xi) \right] - \left[1 - \left(\frac{\kappa}{\alpha} \right) (x - \xi) \right]^{1/\kappa} \right] \end{aligned} \quad (3.10)$$

O sistema de equações obtido a partir da derivação dessa equação em relação aos parâmetros $\theta = (\xi, \alpha, \kappa)$ igualada a zero é, portanto:

$$\frac{1}{\alpha} \sum_{i=1}^S \left[\frac{1 - \kappa - (y_t)^{1/\kappa}}{y_t} \right] = 0 \quad (3.11)$$

$$-\frac{S}{\alpha} + \frac{1}{\alpha} \sum_{i=1}^S \left[\frac{1 - \kappa - (y_t)^{1/\kappa}}{y_t} \left(\frac{x_i - \xi}{\alpha} \right) \right] = 0 \quad (3.12)$$

$$\begin{aligned} -\frac{1}{\kappa^2} \sum_{i=1}^S \left\{ \ln(y_t) ([1 - \kappa - (y_t)^{1/\kappa}] \right. \\ \left. + \frac{1 - \kappa - (y_t)^{1/\kappa}}{y_t} \kappa \left(\frac{x_i - \xi}{\alpha} \right) \right\} = 0 \end{aligned} \quad (3.13)$$

$$y_t = \left[1 - \left(\frac{\kappa}{\alpha} \right) (x - \xi) \right] \quad (3.14)$$

Martins e Stedinger (2000) apontaram, entretanto, que o método de ajuste de máxima verossimilhança resulta em estimativas irrealistas de κ para amostras pequenas ($n < 50$). Nesse sentido, propuseram um método mais eficiente de estimativa por meio da introdução de restrição aos valores do parâmetro κ , pela junção da função verossimilhança com uma função *a priori* desse parâmetro, $\pi(\kappa)$. Assim, os valores de κ ficam limitados às faixas física e estatisticamente razoáveis, embasada na característica desse parâmetro em estudos de cheia. Esse método foi denominado: Estimativa de Máxima Verossimilhança Generalizada (GML).

O método consiste, portanto, na estimativa dos parâmetros que maximizem a função conjunta da verossimilhança com a função *a priori* (ou o logaritmo neperiano da função), correspondendo à multiplicação da função densidade de verossimilhança pela função $\pi(\kappa)$. Portanto, a função resultante que deve ser maximizada é a soma da Equação (3.10) com $\ln(\pi(\kappa))$. Para a escolha dessa função *a priori*, foi considerada a experiência na hidrologia, que indica valores de κ entre -0,3 e 0, e o exame da função densidade de probabilidade GEV que apresenta formas razoáveis para κ entre -0,3 e 0,3. Além disso, para κ entre -0,1 e -0,3 a GEV possui assimetria positiva, o que se adequa ao comportamento esperado de eventos de picos de cheia estritamente positivos (Martins e Stedinger, 2000).

A distribuição *a priori* atribui pesos para diferentes valores de κ dentro do intervalo permitido. Conforme demonstrado por Martins e Stedinger (2000), a combinação dessa distribuição com os dados observados da amostra do local produz estimativas mais precisas de eventos associados a diferentes tempos de retorno na faixa de interesse de κ , do que os métodos dos momentos, momentos-L e de máxima verossimilhança para amostras pequenas e médias.

Respeitando os limites supracitados de κ , a distribuição *a priori* aplicada no método é a distribuição Beta, descrita conforme a Equação (3.15):

$$\pi(\kappa) = \frac{(0,5 + \kappa)^{p-1} (0,5 - \kappa)^{q-1}}{B(p, q)}, \quad (3.15)$$

para valores de κ entre [-0,5, +0,5], com $p = 6$ e $q = 9$ e $B(p, q) = \Gamma(p)\Gamma(q)/\Gamma(p + q)$, com valor esperado de κ igual a -0,10 e variância de $(0,122)^2$.

3.1.2 Estimativa do Intervalo de Confiança

Quando se aplica a inferência estatística clássica por meio do método de máxima verossimilhança para estimativa dos parâmetros, o intervalo de confiança aproximado para os parâmetros da GEV e para os diferentes quantis de cheia pode ser calculado, conforme descrito por Coles (2001), pelo método Delta. Esse método é fundamentado na consideração da normalidade assintótica dos estimadores da máxima verossimilhança, considerando que o tamanho da amostra é grande o suficiente para que isso seja verdade. Dessa maneira, para n grande, tem-se que:

$$\hat{\theta} \sim N(\theta_i, \tilde{\varphi}_{i,i}), \quad (3.16)$$

em que $\tilde{\varphi}_{i,i}$ corresponde ao inverso da matriz de Informação de Fisher $I_o(\theta)$:

$$I_o(\theta) = \begin{bmatrix} -\frac{\partial^2}{\partial \theta_1^2} l(\theta) & \cdots & -\frac{\partial^2}{\partial \theta_1 \partial \theta_d} l(\theta) \\ \vdots & \ddots & \vdots \\ -\frac{\partial^2}{\partial \theta_d \partial \theta_1} l(\theta) & \cdots & -\frac{\partial^2}{\partial \theta_d^2} l(\theta) \end{bmatrix} \quad (3.17)$$

em que $l(\theta)$ representa o log da função de verossimilhança. O intervalo aproximado de confiança $(1 - \alpha)$ dos parâmetros pode ser calculado como:

$$\hat{\theta}_i \pm z_{\alpha/2} \sqrt{\tilde{\varphi}_{i,i}} \quad (3.18)$$

Sendo \varnothing uma função de θ , como é o caso do quantil x_p , o estimador de \varnothing_0 satisfaz :

$$\hat{\varnothing}_0 \sim N(\varnothing_0, V_\varnothing), \quad (3.19)$$

onde,

$$V_\varnothing = \nabla \varnothing^T V_\theta \nabla \varnothing, \quad (3.20)$$

em que V_θ corresponde à inversa da matriz hessiana dos parâmetros e

$$\nabla\phi = \left[\frac{\partial\phi}{\partial\theta_1}, \dots, \frac{\partial\phi}{\partial\theta_d} \right]^T, \quad (3.21)$$

,avaliado em θ_0 . Logo o intervalo de confiança para ϕ para confiança $(1 - \alpha)$ pode ser descrito como:

$$\hat{\phi} \pm z_{\frac{\alpha}{2}} \sqrt{V_{\phi}}, \quad (3.22)$$

3.2 Obtenção de séries históricas de vazão

Em função do alto custo e da difícil logística de se fazer medições diárias de vazão, os valores das séries históricas de vazão disponíveis nas estações fluviométricas são obtidos de forma indireta. Ou seja, as vazões não são medidas e sim, estimadas. Os registros diários das leituras das réguas instaladas são convertidos em vazão por meio da utilização da relação estabelecida entre a cota do nível d'água e a vazão descrita por curvas-chaves.

Este capítulo trata, portanto, da definição das curvas-chaves e as formas para sua obtenção que podem impactar nos valores de vazão das séries históricas produzidos, além das metodologias de extrapolação de curva-chave usualmente utilizadas.

Neste capítulo, focou-se nas características hidráulicas da curva-chave que servirão de subsídios para o modelo proposto nesta dissertação.

3.2.1 Curva-chave

A curva-chave de uma dada seção do rio tenta representar a relação existente entre os valores do nível d'água e vazão, sendo obtida, portanto, com base em medições simultâneas dessas variáveis.

Com esses pares de pontos, objetiva-se obter uma curva que melhor se ajuste aos valores observados. Nesse sentido, diferentes técnicas podem ser utilizadas. Essas técnicas, podem ser simples, como métodos manuais, gráficos e de regressão (Jacon e Cudo, 1989), ou bastante complexas, como é caso de métodos de ajuste de parâmetros por máxima verossimilhança ou inferência Bayesiana (Le Coz *et al.*, 2014).

Dessa maneira, a estimativa da curva-chave é uma das grandes dificuldades no processo de produção das vazões. Além das incertezas inerentes aos dados observados, que

serão discutidos no item 3.3, as características geométricas e de escoamento do canal devem ser amplamente conhecidas para que a estimativa dos parâmetros da curva seja adequada.

A curva-chave pode ser representada de forma geral pela Equação (3.23).

$$Q = a(h - b)^c + \varepsilon \quad (3.23)$$

Nessa equação, a , b e c correspondem aos parâmetros que devem ser estimados com base nas características da seção de medição e os dados observados, e ε corresponde ao erro resultante das fontes de incerteza envolvidas no ajuste da curva. O parâmetro a corresponde a um coeficiente de escala que depende da geometria e rugosidade do rio, b a referência de nível de vazão nula e c é o coeficiente relacionado ao tipo de controle hidráulico.

A geometria do leito e da seção são definidas por três parâmetros principais: a declividades do fundo (S), a área molhada (A_m) e o raio hidráulico (R), calculado pela razão entre área molhada e perímetro molhado.

O quarto parâmetro, representativo da característica hidráulica do canal, é a rugosidade. Ela caracteriza a resistência do leito à passagem da água podendo variar ao longo da seção de acordo com a natureza física do material e sua granulometria no leito do rio e suas margens. A rugosidade é representada nas fórmulas de Manning (Eq. (3.24)) e Strickler (Eq.(3.25)) para escoamentos permanentes uniformes como n e K , respectivamente.

$$V = \frac{1}{n} R^{2/3} \sqrt{S} \quad (3.24)$$

$$V = KR^{2/3} \sqrt{S} \quad (3.25)$$

em que V , R e S correspondem à velocidade do escoamento, ao raio hidráulico e à declividade da linha d'água respectivamente.

A rugosidade pode ser obtida em campo, a partir das características observadas do leito e margens, do levantamento de perfis de velocidade e da geometria do leito. Na falta de levantamentos, ela pode ser estimada com base em tabelas indicativas de valores experimentais de rugosidade, como é o caso das apresentadas por Chow (1959).

Além das características geométricas e de rugosidade da seção, é muito importante conhecer o comportamento do escoamento do curso d'água na seção de medição e suas proximidades. Nesse contexto, o tipo de controle hidráulico e sua permanência devem ser analisados.

O controle hidráulico representa as características físicas do canal que governam a relação entre o nível d'água e a vazão em uma dada seção. Diz-se que o escoamento em uma seção limnimétrica está sob controle quando, em um trecho de geometria invariável, tem-se sempre uma mesma descarga para uma mesma cota. Existem dois tipos de controle

principais: o *controle de canal*, associado ao regime de escoamento uniforme, e a *seção de controle*, associada ao regime crítico (Boiten, 2003; Jaccon e Cudo, 1989).

No *controle de canal*, o fluxo é regulado, na maioria do tempo, pela geometria e rugosidade do canal. Esse tipo de controle é muito comum em rios de maior porte ou rios de planície. Nesse caso, as relações entre cota e vazão dependem das dimensões do leito, da rugosidade e da declividade do fundo, e podem ser descritas pelas Equações (3.24) e (3.25). Para que esse escoamento prevaleça, as condições devem ser favoráveis e a seção deve estar distante de singularidades.

Por outro lado, o controle tipo *seção de controle* ocorre próximo às singularidades do leito, como soleiras de vertedouros, mudanças bruscas de declividade, presença de bancos de areia, pontes ou outro tipo de obstrução em que o escoamento se torna crítico. Dessa maneira, o escoamento nas proximidades passa a ser não uniforme.

Os controles hidráulicos podem variar com a vazão, no espaço e no tempo, impactando significativamente a forma da curva. A existência de soleiras sucessivas ao longo do rio provoca alteração da curva-chave à medida que a vazão aumenta e cada controle vai sendo afogado, enquanto que instabilidades no leito do rio e deformações geométricas alteram a curva ao longo dos anos. Assim, a curva pode ter mais de um segmento para cada tipo de controle dependendo da faixa de vazão (Figura 3-1) e deve ser reavaliada constantemente para detecção das possíveis modificações nesses controles.

De acordo com Jaccon e Cudo (1989), um controle perfeito é um controle permanente no espaço e no tempo, como é o caso de saltos e cascatas em trechos rochosos e soleiras artificiais de vertedouros. Devido a esse fato, tais tipos de controle são considerados adequados para instalação de estações de medição.

Além dessa estabilidade, é importante considerar a sensibilidade da seção de medição para variações de vazão. A seção é mais sensível quanto maior for a variação de cota para uma pequena variação de vazão, o que ocorre em seções estreitas e profundas com escoamento fluvial. Quanto maior a sensibilidade, menor o impacto dos erros de leitura de cota que porventura acontecem no momento da medição, tendo em vista que uma grande variação da cota resultará em pequena variação na vazão estimada, tornando essas seções melhores para instalação da régua.

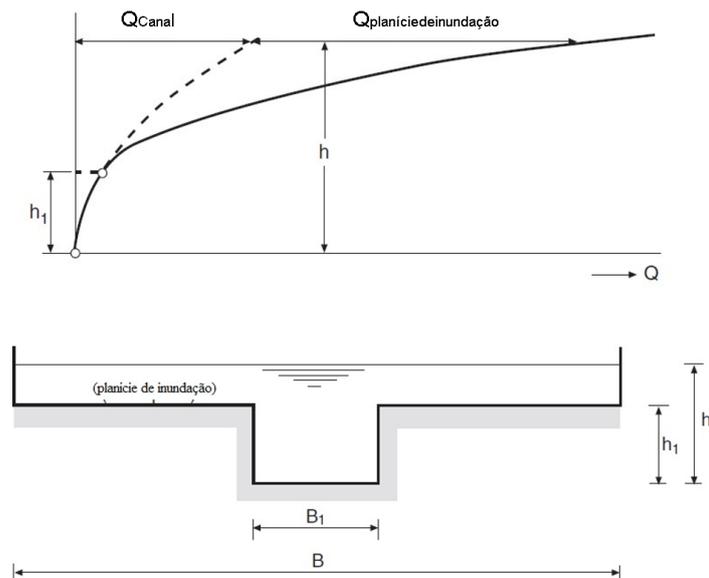


Figura 3-1 - Exemplo de Curva-chave com Dois segmentos (Boiten 2003, modificada)

Além dessa estabilidade, é importante considerar a sensibilidade da seção de medição para variações de vazão. A seção é mais sensível quanto maior for a variação de cota para uma pequena variação de vazão, o que ocorre em seções estreitas e profundas com escoamento fluvial. Quanto maior a sensibilidade, menor o impacto dos erros de leitura de cota que porventura acontecem no momento da medição, tendo em vista que uma grande variação da cota resultará em pequena variação na vazão estimada, tornando essas seções melhores para instalação da régua.

A Tabela 3-1 modificada de Le Coz *et al.* (2014) mostra como os parâmetros da equação da curva-chave (Eq. (3.23)) podem ser calculados para os controles mais usuais. Ressalta-se que no caso do controle de canal a parametrização proposta parte da premissa de que a largura do rio é suficientemente grande de forma que o raio hidráulico se aproxima da profundidade do rio. Nesse caso, o coeficiente c se aproxima de $5/3$, quando isso não for verdade esse coeficiente será inferior a esse valor.

Em todos os controles descritos anteriormente, o escoamento é considerado permanente e a curva-chave é biunívoca, ou seja, para cada valor de cota tem-se um único valor de vazão e vice-versa. Entretanto, essa relação nem sempre acontece. Ocorrem normalmente relações *unívocas*, onde se tem associada uma única vazão para cada cota, ou *não-unívocas*, nos casos em que ocorrem vazões diferentes para uma mesma cota. Essas relações e suas causas são melhor detalhadas no livro de Jacon e Cudo (1989). Graficamente, elas são observadas na curva na forma de um laço (ou histeresis) (

Figura 3-2).

Esse tipo de curva ocorre normalmente devido à passagem da cheia, às modificações temporárias do leito do rio, que aumentam ou reduzem a elevação do fundo, às modificações sazonais na vegetação das margens e aos efeitos de remanso. No caso da passagem da cheia, por exemplo, no momento da subida do valor da vazão os níveis d'água observados são menores que o valor de referência, produzindo valores de vazão inferiores aos reais, ao passo que, na descida do hidrograma, os níveis são maiores e a vazão obtida pela curva-chave de referência é maior que a real.

Como são vários fatores significativos que devem ser analisados ao se tentar estimar a relação existente em uma dada seção entre cota e vazão, é necessário um bom conhecimento hidráulico do trecho do rio. O conhecimento das características apresentadas e também de suas incertezas devem ser incorporados de alguma forma na definição da curva-chave a ser adotada.

No modelo proposto nesse trabalho é possível que tais características sejam incorporadas na análise para uma curva simples com um único segmento e sem variação em relação ao regime de referência e no tempo, sendo possível a implementação de melhorias no modelo para consideração de casos mais complexos.

Tabela 3-1 - Fórmulas Hidráulicas clássicas para controles hidráulicos mais comuns (Le Coz *et al* 2014, modificada)

Tipo de Controle	Suposições Ideais	a	b	c
Canal	Retângulo largo, escoamento permanente uniforme	$KB\sqrt{S}$	Elevação do leito	5/3
Seção (Retangular)	Vertedouro perpendicular ao escoamento sem efeito de remanso	$C_r B_r \sqrt{2g}$	Altura da Soleira	3/2
Seção (Triangular)	Vertedouro perpendicular ao escoamento sem efeito de remanso	$C_t t g \beta \sqrt{2g}$	Nível do Vértice Inferior	5/2
Seção (Orifício)	Vertedouro perpendicular ao escoamento sem efeito de remanso	$C_o A_o \sqrt{2g}$	Elevação do centro do orifício	1/2

K (coeficiente de Strickler), S (declividade) C_t e C_r (coeficiente de descarga do vertedouro), B (largura do canal ou do vertedouro), β (ângulo interno do vertedouro triangular), C_o (coeficiente de descarga do orifício) A_o (área do orifício), g (gravidade).

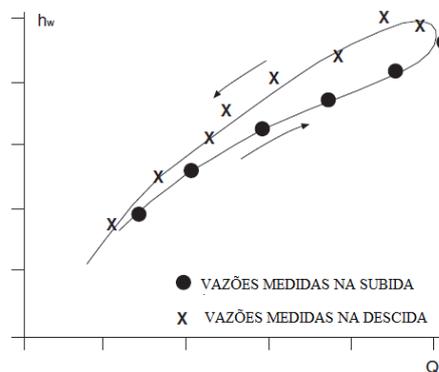


Figura 3-2 - Curva-chave com Laço na Passagem da Cheia (Boiten 2003, modificado)

3.2.2 Obtenção da Curva-chave

A curva-chave é obtida a partir das leituras simultâneas de vazão e de nível d'água realizadas ao longo do tempo na seção. Essas leituras são utilizadas para se estimar o traçado ou função que melhor se ajuste aos pares dessas medições. Essa estimativa pode não ser possível, considerando que irregularidades e instabilidades na seção podem não permitir que haja uma correspondência única entre os valores de cota e vazão.

Quando possível, a estimativa dos parâmetros é tradicionalmente feita de forma a minimizar os desvios entre os dados observados e os valores calculados por meio de uma regressão linear do log da equação da curva-chave.

Além desse ajuste tradicional, outras metodologias como a análise estatística dos dados por meio de regressão não linear (Venetis, 1970; Petersen-Overleir, 2004), análise hidráulica (Shmidt, 2002) e utilização de inferência Bayesiana (Kuczera, 1999; Le Coz *et al.*, 2014; Moyeed e Clark, 2005) também vem sendo utilizadas.

No caso do uso da regressão linear, o método dos mínimos quadrados pode ser aplicado para a estimativa dos parâmetros, de modo que a soma do quadrado da diferença entre o valor observado e o calculado seja a menor possível.

Assim, a equação da curva-chave (Eq. (3.23)) em forma potencial pode ser ajustada a uma reta aplicando o logaritmo em ambos os lados da equação:

$$\log Q = \log a + c \log (h - b) \quad (3.26)$$

A calibração dos parâmetros é obtida ajustando-se uma reta aos pares $(\log Q$ e $\log(h - b))$ de modo a minimizar, para a nuvem de pontos, o somatório dos quadrados dos desvios da variável dependente $(\log Q)$, conforme apresentado na Equação (3.27).

$$D = \sum_{i=1}^n (\log Q_i - (\log a + c \log (h - b)))^2 \quad (3.27)$$

A estimativa dos parâmetros da regressão também pode ser realizada pelo método da máxima verossimilhança apresentado no item 3.1.1, partindo-se da premissa de que os dados observados são normalmente distribuídos. Por outro lado, os parâmetros podem ser estimados por inferência Bayesiana, em que, além dos dados, também é levada em

consideração a informação *a priori* acerca dos parâmetros. Por ser o foco deste trabalho, ela será melhor descrita no item 3.5.1.

No caso da análise hidráulica são utilizados os modelos em que se busca representar o trecho do rio e suas características. Os parâmetros são então calibrados de forma que a diferença entre o nível simulado e observado seja a mínima possível (Schmidt,2002).

Após a estimativa dos parâmetros em qualquer um dos métodos, é importante verificar a validade física de cada um com base nas características da seção e do escoamento. Além disso, deve-se verificar os aspectos apresentados no item 3.2.1 quanto à estabilidade da curva-chave no tempo e sua variação ao longo da seção verificando a necessidade da consideração de mais de um segmento de curva.

3.2.3 Extrapolação de Curvas-Chaves

No caso de estudo de frequência de cheias, as cotas máximas são utilizadas na estimativa das vazões máximas. Entretanto, em virtude da dificuldade de se realizar medições no momento da cheia, as curva-chaves estimadas estão limitadas usualmente às medições inferiores aos valores observados de cotas máximas. Dessa maneira, a estimativa das vazões fica sujeita à extrapolação da curva-chave previamente estabelecida.

A extrapolação de uma curva consiste na prolongação criteriosa da mesma além da faixa de pontos registrados nas campanhas de campo. Esse procedimento depende do quanto se conhece/desconhece do comportamento hidráulico e características geométricas da seção e deve levar em conta que:

- A relação entre Q e H pode mudar significativamente a partir de uma dada cota em função de mudança de geometria, de controle hidráulico, rugosidade e comportamento do escoamento;
- A relação entre Q e H pode não ser única
- A curva pode se alterar ao longo do tempo ou em períodos diferentes no ano.

Existem algumas metodologias largamente utilizadas para extrapolação de curvas-chaves que partem de premissas em relação ao comportamento hidráulico da seção, como é o caso dos métodos de extrapolação logarítmica, de Stevens e de Manning (Jacon e Cudo,1989; Tucci e Silveira,1993) apresentados a seguir.

- a) Extrapolação Logarítmica

Nesse método, o princípio é aplicar na parte superior da curva o ajuste da expressão matemática exponencial (Eq. (3.23)). Plotam-se os dados observados em escala log x log e determina-se a reta que melhor se ajusta a eles, extrapolando a reta para o valor de cota máxima observada. Assume-se, portanto, que os valores dos coeficientes a e c são mantidos no trecho extrapolado.

Entretanto, conforme Tucci e Silveira (1993), esse método requer que sejam atendidas às seguintes condições: relação h/Q caracterizada por um bom alinhamento das medições de águas médias e altas sobre o papel log-log; medições existentes até uma cota suficientemente elevada para que a direção da reta seja bem definida; perfil transversal sem descontinuidade de forma nas cotas extrapoladas e controle de jusante permanente entre cotas médias e altas.

b) Método de Stevens

Esse método consiste na extrapolação com base na fórmula de Chézy, descrita da seguinte forma:

$$\frac{Q}{A\sqrt{R}} = C\sqrt{S} \quad (3.28)$$

em que Q , A , C , S e R representam, respectivamente, a vazão, a área molhada, o coeficiente de Chézy, a declividade e o raio hidráulico.

A componente $A\sqrt{R}$ corresponde ao fator geométrico e $C\sqrt{S}$ ao fator de declividade. Como esse segundo fator em condições estáveis costuma variar muito pouco com o aumento do nível d'água, a função $Q = f(A\sqrt{R})$ pode ser representada por uma reta que passa pela origem. A extrapolação dessa reta permite o cálculo da vazão a partir do conhecimento de $A\sqrt{R}$, que é função única e exclusivamente da cota, conforme apresentado na Figura 3-3.

Em linhas gerais, a partir da cota calcula-se o fator geométrico e, a partir desse valor, obtém-se o valor da vazão a partir da extrapolação da reta de $Q = f(A\sqrt{R})$.

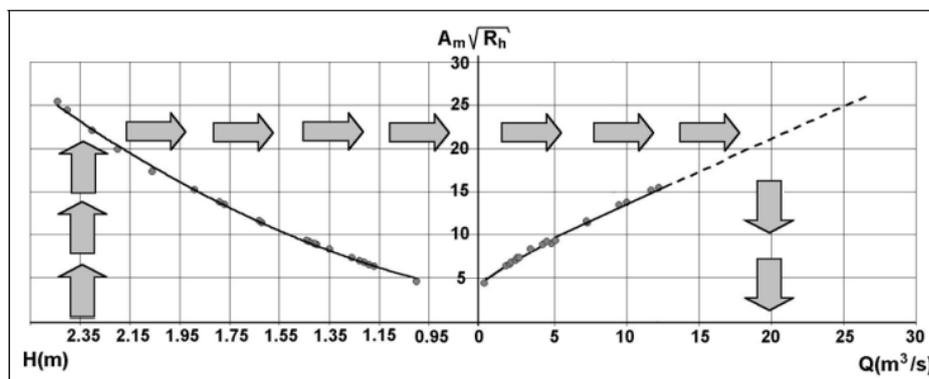


Figura 3-3 – Extrapolação da Curva-Chave pelo método de Stevens (Sefione, 2002)

c) Método de Manning

Esse método considera que o quociente da raiz da declividade da linha d'água pelo coeficiente de rugosidade de Manning tende a tornar-se constante para vazões altas (Tucci e Silveira, 1985).

$$K_m = \frac{\sqrt{S}}{n} = \frac{Q}{AR^{2/3}} \quad 3.29$$

Pelo fato do valor desse coeficiente (K_m) tender a se tornar constante, ou seja, independente da cota do nível d'água, ele é utilizado para efetuar a extrapolação para os dados não medidos. Primeiro, plotam-se os pontos de K_m em função das cotas H no intervalo de vazões medidas e, em seguida, verifica-se se existe alguma tendência assintótica que represente a estabilização do valor desse coeficiente.

Caso essa estabilização ocorra, o gráfico pode ser prolongado até os valores de cota que se deseja extrapolar e estima-se o valor de K_m . De posse do valor do coeficiente, calcula-se, com base na seção topobatimétrica do rio, as áreas e perímetros molhados para cada cota e obtém-se o valor da vazão a partir da Equação 3.29.

A Figura 3-4 exemplifica esse procedimento.

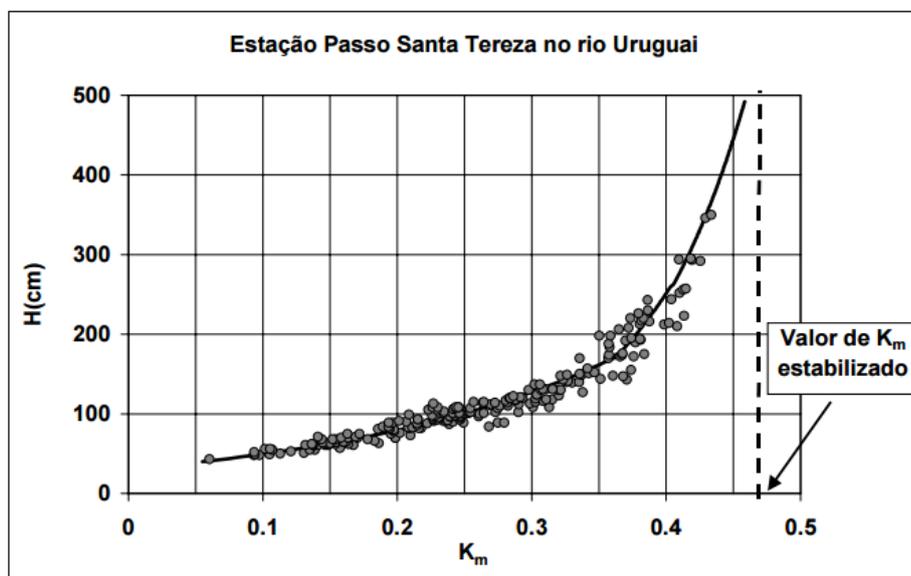


Figura 3-4 – Exemplo de Extrapolação pelo Método de Manning (Sefione, 2002)

A desvantagem desse método é que nem sempre essa assintoticidade da curva é observada. Dessa maneira, a incerteza inerente à estimativa de K_m é transferida para o trecho extrapolado da curva. Outro problema que pode ocorrer é a inversão dos valores de K_m ,

principalmente devido à presença de vegetação bastante densa nas margens, o que pode alterar completamente o gráfico e provocar erros graves na extrapolação (Sefione, 2002).

Os três métodos apresentados dependem de um certo conhecimento hidráulico da seção e que a condição de escoamento permanente uniforme seja mantida no trecho extrapolado.

Além desses métodos, a modelagem hidráulica também pode ser utilizada de forma a simular o comportamento do trecho do rio e as diferentes situações de alteração de controle e características hidráulicas e geométricas. Ressalta-se, porém, que por não haver medições de vazão alta, a calibração para alguns valores de vazão dependerá da calibração já realizada para os dados disponíveis e do conhecimento do comportamento do rio no trecho extrapolado.

No caso da aplicação de métodos estatísticos de estimativa de curva, como é o caso da inferência Bayesiana, uma forma de se fazer a extrapolação é por meio da extrapolação logarítmica, em que se assume que os parâmetros permanecem constantes para níveis de cota e vazão não observados. Entretanto, é importante avaliar se as vazões mais altas pertencem ao mesmo segmento de curva com parâmetros estimados para vazões menores. Caso isso não seja verdade, torna-se necessária a divisão da curva em segmentos, extrapolando-se, portanto, apenas o segundo segmento.

No trabalho proposto, considerou-se a extrapolação logarítmica em curva com apenas um segmento, sendo necessários estudos futuros mais aprofundados para consideração da extrapolação e suas incertezas.

3.3 ANÁLISE DE INCERTEZAS EM CURVAS-CHAVES

Um dos grandes problemas do uso dos dados de vazão é a incerteza envolvida na estimativa da curva-chave. No item anterior foi visto o quanto é importante o conhecimento das especificidades hidráulicas da seção de medição para melhor ajuste da curva. Porém, essas não são as únicas incertezas que podem afetar a obtenção da relação entre cota e vazão.

Nesse item serão tratadas as incertezas que impactam a estimativa da curva-chave e como essas incertezas foram avaliadas e abordadas em estudos disponíveis na literatura.

3.3.1 Fontes de Incertezas

Na avaliação de incertezas de curvas-chaves, é necessário diferenciar claramente as incertezas na relação cota-vazão definida para o regime de referência e as incertezas derivadas das medições instantâneas de descarga.

O primeiro tipo de incerteza refere-se aos erros de modelagem ou de desconhecimento da relação hidráulica entre as duas variáveis, partindo de funções que não representam fisicamente o comportamento do escoamento da seção. Esse erro se torna ainda mais importante em regiões onde os dados são escassos ou inexistentes, dificultando a extrapolação da curva para valores de vazão não medidos. Enquadra-se também nesse erro, o erro no ajuste, que representa o fato de que a formulação matemática não consegue representar perfeitamente os dados, mesmo que se disponha dos valores reais de medição.

A extrapolação da curva-chave é considerada a maior fonte de erros na estimativa dos valores de vazão (Kuczera, 1996; Clarke, 1999) e deve levar em consideração a estabilidade da seção, a geometria e o tipo de controle para evitar equívocos. Em casos controlados, a extrapolação é simplificada pelo uso da equação de Manning, em que se espera que a rugosidade não varie com o aumento do nível d'água. Como mencionado anteriormente, existem poucos métodos de extrapolação e em estatística clássica e Bayesiana ainda há uma lacuna de como realizar a extrapolação, e o que se faz é assumir que os parâmetros são os mesmos estimados para o segmento de curva dos dados medidos.

Os erros na extrapolação ocorrem normalmente devido à falta de conhecimento no momento de estimar a curva-chave em relação a alterações abruptas da relação entre nível d'água e vazão a partir de uma determinada cota. Mudanças na forma, no tipo de controle de jusante, na rugosidade, no tipo de vegetação, presenças de fluxos secundários ou *by-pass* também provocam alterações em sua forma.

Com relação aos erros nas medições instantâneas, podemos definir os erros na medição de vazão e na medição do nível d'água. Os erros na medida da vazão estão relacionados a três situações principais: erros instrumentais, humanos e locais; erros devido à não-estacionariedade da vazão na hora da medição e erros devido à alteração das condições hidráulicas em relação ao regime de referência por efeito, por exemplo, de remansos, alteração na vegetação ou no tipo de controle hidráulico, conforme já detalhado no item 3.2.1. A magnitude e o impacto desse erro se mostraram significativos em estudos anteriores (Di Baldassarre e Montanari, 2009; Di Baldassarre *et al.*, 2011; Despax *et al.*, 2016).

Por outro lado, os erros de medições de nível d'água, mais problemáticos em seções menos sensíveis, em que pequenos erros de leitura podem impactar no valor obtido para a

vazão, podem ser considerados negligenciáveis quando comparados com os erros nas medições de vazão (Le Coz *et al.*, 2014).

3.3.2 Métodos Para Estimativa de Incertezas em Curvas-chaves

Apesar da influência significativa das incertezas descritas anteriormente nos modelos hidrológicos e hidráulicos, elas na maioria das vezes não são consideradas (Mason *et al.*, 2016). Isso ocorre ou por desconhecimento do Hidrólogo em relação à produção dos dados hidrométricos, ou pelo reduzido número de pesquisas sobre métodos para quantificar essas incertezas e seus impactos (Petersen-Overleir e Reitan, 2008).

Não existe, atualmente, uma metodologia específica para avaliar essas incertezas. Todavia, muitos trabalhos foram desenvolvidos buscando estimá-las e avaliar seus impactos em estudos que usam como fonte de dados principal as séries de vazões estabelecidas com base na curva-chave tais como: calibração hidrológica e hidráulica; análises de seca e cheias; e dimensionamento de estruturas). Os métodos até hoje desenvolvidos para expressar as incertezas nas relações entre cota e vazão se dividem em quatro formas de abordagem principais.

A primeira abordagem baseia-se principalmente em análises de modelos hidráulicos e de sensibilidade dos parâmetros, fornecendo uma estimativa realista dos limites dos erros, como é o caso de estudos como de Shmidt (2002), Di Baldassarre e Montanari (2009) e Lang *et al.* (2010).

Di Baldassarre e Montanari (2009), por exemplo, investigaram as incertezas levando em consideração erros nos dados, nos procedimentos de interpolação e extrapolação dos dados medidos, nos efeitos de condições não uniformes de escoamento e no impacto da presença de vegetação na rugosidade do canal. Como os demais autores, eles concluem que os erros nas estimativas são importantes, porém, em função de não haver uma base estatística dessas incertezas, seus resultados não são diretamente aplicáveis.

A segunda abordagem refere-se a estudos que utilizam técnicas de regressão (Venetis, 1970; Clarke, 1999, Clarke *et al.*, 2000) e aplicação da função de verossimilhança levando em consideração curvas de segmento único (Petersen-Overleir, 2004), multi segmentadas (Petersen-Overleir e Reitan, 2005), com histeresis (Petersen-Overleir, 2006) e com vazões em seções com planície de inundação (Petersen-Overleir, 2008).

A terceira abordagem utiliza métodos de análise variográfica e análise dinâmica para a estimativa das incertezas e dos parâmetros em curvas-chaves não estacionárias. (Jalbert *et al.*,2011; Westenberg *et al.*,2011; Morlot *et al.*,2014)

Esses modelos, entretanto, possuem algumas desvantagens elencadas por Reitan e Petersen (2008) e Le Coz *et al.*(2014) no que diz respeito à probabilidade de obtenção de mais de uma solução de parâmetros (multimodalidade) e à desconsideração das características hidráulicas e geométricas do canal nas imediações da estação, assumindo que os pares de cota-vazão trazem toda informação necessária para a estimativa dos parâmetros.

A superação dessas limitações surgiu com um quarto tipo de abordagem para estimativa dessas incertezas. Essa abordagem consiste na aplicação da técnica de inferência Bayesiana, proposta pela primeira vez por Moyeed e Clark (2005) e melhor detalhada por Le Coz *et al.* (2014).

A Análise Bayesiana é fundamentada no teorema de Bayes, em que o conhecimento prévio do comportamento dos parâmetros (função *a priori*) é combinado com as observações (função verossimilhança) para obtenção da distribuição *a posteriori* desses parâmetros. Assim, as incertezas em relação aos conhecimentos hidráulicos podem ser contempladas na distribuição *a priori*, enquanto as incertezas do ajuste da curva podem ser inseridas como um parâmetro a ser estimado.

Em função das vantagens apontadas, a aplicação da análise Bayesiana foi adotada nesse trabalho para que fosse possível analisar o impacto da incorporação das incertezas no estudo de frequência de cheia. Tal método será melhor detalhado no item 3.5.

3.4 Impacto das Incertezas de Curvas-chaves em Estudo de Cheias

Estudos de frequência de cheias são desenvolvidos assumindo que as séries históricas de vazão não possuem erros. Porém, conforme já apresentado no item 3.3, essa prerrogativa não é verídica e os dados de vazão são fortemente influenciados por diferentes erros desde a medição dos dados à estimativa da curva-chave.

Como a estimativa das vazões de projeto normalmente é feita por meio do ajuste de uma distribuição teórica que melhor represente a frequência de ocorrência dos eventos aos dados das vazões máximas anuais, esse procedimento é afetado também pelas dificuldades em se escolher o modelo apropriado e pelas incertezas na estimativa dos parâmetros desse modelo.

Dessa maneira, a desconsideração dos erros pode falsear bastante os resultados, principalmente porque a maioria das vazões com tempos de retorno elevados são obtidas de trechos extrapolados das curvas-chaves que possuem, em função do desconhecimento, erros ainda mais expressivos (Lang *et al.*, 2010).

Durran (1990), Kuczera (1996), Clarke (1999), Reis e Stedinger (2005), Petersen-Overleir e Reitan (2009a), Di Baldassarre *et al.* (2012), Neppel *et al.* (2010), Lang *et al.* (2010), Montanari e Baldassarre (2012), Steinbakk *et al.* (2016) foram alguns dos autores que estudaram os possíveis impactos da aplicação dessas incertezas nas estimativas de cheias. A ideia principal desses estudos foi a aplicação de diferentes modelos de erro aos valores das vazões e a incorporação desses erros em métodos de inferência estatística clássica e Bayesiana para estimação dos parâmetros das distribuições de probabilidade escolhidas. Os intervalos obtidos com e sem a utilização do erro foram comparados para análise do impacto.

A distribuição de probabilidade adotada por quase todos os autores foi a de Valores Extremos (GEV), por ser mais simples de implementar (Kuczera, 1996). As exceções são Di Baldassarre *et al.* (2012) e Durran (1990), que adotaram, respectivamente, Gumbel (apenas um dos tipos da GEV) e Log-Normal, e Reis e Stedinger (2005), que utilizaram a distribuição Log Normal e Log Pearson III.

A base metodológica utilizada na maioria desses estudos foi a utilização de funções verossimilhança com a incorporação do erro e o cálculo dos intervalos de confiança dos quantis por modelos assintóticos (Kuczera, 1996; Di Baldassarre *et al.*, 2012, Petersen-Overleir e Reitan, 2009a). Nessa linha de estudos, as vazões das séries foram descritas em função das vazões reais desconhecidas somadas a um erro.

Kuczera (1996), por exemplo, propõe que esse erro seria proporcional ao valor da vazão menos o valor da maior vazão medida (Eq. (3.30)), denominada por ele como ponto de âncora.

$$w = q_1 + e_2(q - q_1) \quad (3.30)$$

A Figura 3-5 exemplifica esse modelo.

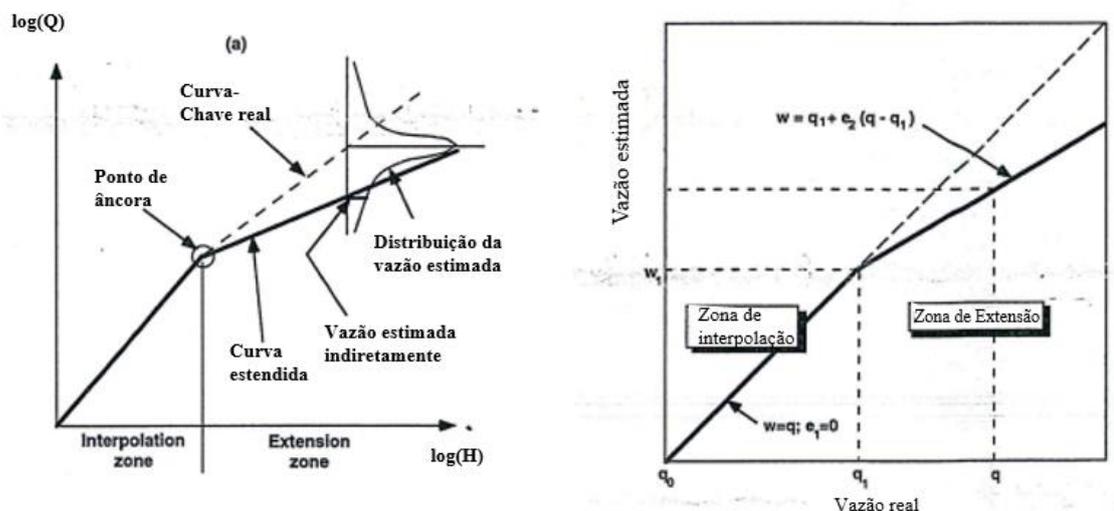


Figura 3-5 – Modelo de erros proposto por Kuczera (1996)

Modelos similares de erro foram utilizados por Di Baldassarre *et al.* (2009) e Neppel *et al.* (2010). O erro e_2 foi considerado uma variável aleatória positiva com valor esperado igual a 1, distribuído por uma função log normal com média 1 e desvio padrão conhecido, o qual era pré-definido pelos autores para comparação/avaliação dos resultados.

Os resultados das análises comparativas dos estudos citados foram muito próximos. Observou-se, em todos os casos, um aumento significativo dos intervalos de confiança dos valores dos quantis quando o erro foi incorporado às funções. Observou-se também que os erros causados pela extrapolação podiam corromper de forma substancial os valores das vazões de projeto (Kuczera, 1996). Porém, apesar de avaliarem o impacto de possíveis incertezas, esses métodos não investigaram a fundo a origem das incertezas e o conhecimento hidráulico do local de medição.

Além disso, apesar de apresentarem resultados válidos e interessantes, o uso de funções verossimilhança com incorporação de incertezas não se mostrou o mais vantajoso, como concluíram Kuczera (1996), Clarke (1999), Oveleir e Reitan (2009) e Reis e Stedinger (2005). Observou-se que na presença do erro, a estimativa dos parâmetros pela minimização da função log de verossimilhança apresentava múltiplos ótimos, resultando em soluções não únicas para os parâmetros. Além disso, a função ficava cada vez menos quadrática, o que tornava, conforme demonstrado nos estudos, a estimativa de incertezas por modelos assintóticos menos adequada (Kuczera, 1996; Clarke, 1999; Overleir e Reitan, 2009; Reis e Stedinger, 2005).

Neppel *et al.* (2010), Reis e Stedinger (2005) e Lang *et al.* (2010) propuseram, então, uma abordagem um pouco diferente. Utilizaram para estimativa dos parâmetros da distribuição de probabilidade uma metodologia Bayesiana, em que os erros pudessem ser introduzidos no modelo. Nesses modelos, o erro também foi considerado como proporcional ao valor da vazão, como proposto por Kuczera (1996).

No estudo de Neppel *et al.* (2010), as curvas-chaves foram estimadas por uma simulação hidráulica unidimensional com calibração do coeficiente de Manning a partir dos valores medidos e possíveis erros nas medições, sendo o resultado dessa análise de sensibilidade adotado como intervalo *a priori* do erro. Procedimento semelhante foi aplicado por Lang *et al.* (2010) em que se utilizou uma distribuição *a priori* triangular cuja base dependia dos resultados da análise de sensibilidade para o coeficiente de rugosidade.

Lang *et al.* (2010) propuseram a extrapolação da curva-chave por meio de modelo hidráulico unidimensional. Nesse modelo foi considerada a variação do coeficiente de rugosidade para obtenção da curva-chave com seu intervalo possível determinado por meio dessa análise de sensibilidade. O método foi então aplicado a 8 bacias no Mediterrâneo e os resultados com e sem a incorporação do erro foram comparados. A partir dessas análises observou-se que a distribuição *a posteriori* dos parâmetros de posição e escala da distribuição GEV são bastante sensíveis à incorporação do erro. Além disso, observou-se que a não consideração das incertezas leva a uma redução bastante otimista da incerteza final dos quantis de cheia. Os valores médios dos quantis também foram afetados ao se considerar os erros. A grande limitação desses estudos é que a incerteza da curva-chave depende da análise de sensibilidade dos dados pela modelagem hidráulica que requer um nível de conhecimento alto em relação ao trecho do rio, o que não é a realidade da maioria das estações estudadas.

Por sua vez, em sua análise, Reis e Stedinger (2005) compararam os resultados pela metodologia clássica MLE (máxima verossimilhança) e de inferência Bayesiana na estimativa das curvas de frequência de cheias para os seguintes casos: com e sem a incorporação dos erros, para dados observados e com a adição de informações históricas de cheias. A metodologia descrita pelos autores indicou o aumento do intervalo de credibilidades quando as incertezas foram incorporadas, demonstrando a viabilidade pertinência do uso da inferência Bayesiana para incorporação dos erros associados às vazões estimadas.

Modelo para estimativa conjunta dos parâmetros foi proposto apenas por Reitan e Petersen-Overleir (2009), Steinbakk *et al.* (2016) e Osorio e Reis (2016). O primeiro propôs

metodologia para combinação de duas funções verossimilhanças: um referente às incertezas da curva-chave e outra aos dados de vazão e sua variabilidade amostral e estimativa dos parâmetros dessas funções pelo método da máxima verossimilhança. O segundo e o terceiro, adotaram procedimento similar com o uso da inferência Bayesiana, em que as estimativas dos parâmetros da curva-chave e da curva de frequência de cheias eram feitas separadamente. Por se tratar do enfoque principal desse trabalho, esses estudos serão melhor descritos no item 3.5.3.

Nesses trabalhos, se concluiu que o impacto da incerteza também é significativo se não for devidamente considerado, mas que isso depende também da quantidade de dados de vazão disponível. Isso se deve ao fato de que em locais onde existem poucas informações, a incerteza amostral se torna mais importante que as incertezas associadas às curvas-chaves.

Apesar de mais completos que os demais em relação a incorporação das incertezas, algumas limitações podem ser apontadas. No primeiro, por se tratar de uma estimativa pelo método da máxima verossimilhança, em alguns casos o resultado não convergia para uma solução ótima como já havia sido demonstrado em estudos anteriores que adotaram essa metodologia. Já no segundo, o esforço computacional para geração das distribuições era muito grande (ver item 3.5.3) e a recomendação dos autores era de que uma estrutura Bayesiana hierárquica completa deveria ser desenvolvida para reduzir esse esforço. Por fim o terceiro, era um estudo preliminar utilizado como primeira tentativa de compreensão das incertezas para o desenvolvimento do modelo proposto nesse trabalho.

Em suma, os estudos, com exceção dos de Reitan e Overleir (2009), Steinbakk *et al.*(2016) e Osorio e Reis (2016), não avaliaram profundamente as fontes das incertezas. Nestes, os erros foram aplicados diretamente aos valores de vazões máximas utilizados nos estudos de frequência de cheias, sem uma avaliação detalhada da incerteza na estimativa dos parâmetros da curva-chave.

Constata-se, portanto, que a literatura carece de estudos que englobem completamente as questões referentes ao comportamento hidráulico e aos erros de medição, e que propaguem esses erros na estimativa dos parâmetros da distribuição a ser ajustada à amostra.

Estudos mais detalhados foram feitos analisando independentemente as incertezas na curva-chave, como é o caso do trabalho de Le Coz *et al.* (2014), porém, sem avaliar o impacto dessas incertezas em um estudo de frequência de cheias local.

Tendo em vista o que já foi feito sobre o assunto e as limitações envolvidas em cada estudo, e considerando o modelo detalhado proposto por Le Coz *et al.* (2014), nota-se a

pertinência da utilização de metodologias de aplicação de inferência estatística Bayesiana, nas quais é possível fazer uma análise mais completa, levando em conta os conhecimentos hidráulicos do local e as incertezas envolvidas.

Essa metodologia e suas aplicações para estimativas de parâmetros de curva-chave e de distribuições estatísticas utilizadas em estudos de frequência de cheias serão detalhadas no item 3.5 a seguir.

3.5 INFERÊNCIA BAYESIANA

A técnica de Inferência Bayesiana é uma ferramenta utilizada para estimativa dos parâmetros de determinada função ou distribuição condicionada aos valores observados e aos conhecimentos *a priori* acerca dos parâmetros. Por fornecer como resultado final a distribuição *a posteriori* de todos os parâmetros, ela também pode ser utilizada para predição da probabilidade de eventos futuros.

A especificação de determinada informação originada, por exemplo, do conhecimento do local, na forma de uma distribuição *a priori*, é considerada, simultaneamente, a maior vantagem e armadilha no uso dessa técnica (Coles, 2001). Ao mesmo tempo em que pode significar a adição de uma fonte relevante de informação, pode também introduzir certa subjetividade na análise, já que *a prioris* diferentes podem ser especificadas por diferentes especialistas para uma mesma situação.

Além do conhecimento *a priori*, é possível acrescentar as incertezas inerentes aos dados utilizados, o que torna essa metodologia bastante interessante para o trabalho em foco. Fundamentada no teorema de Bayes, calcula-se a distribuição *a posteriori* dos parâmetros que se deseja estimar com base na atualização dos conhecimentos *a priori* pela função verossimilhança dos dados observados (Eq. (3.31)).

$$p(\theta|x) = \frac{p(x|\theta)p(\theta)}{\int p(\theta)p(x|\theta)d\theta} \quad (3.31)$$

em que $p(\theta|x)$ representa a distribuição *a posteriori* dos parâmetros, $p(x|\theta)$ a função de verossimilhança e $p(\theta)$ a distribuição *a priori* de θ . O denominador é uma constante aplicada para normalizar a função de densidade *a posteriori*. Em casos onde o vetor de parâmetros é muito extenso, o cálculo do denominador se torna problemático, o que dificultou durante muito tempo o uso dessa estrutura (Coles, 2001; Le Coz *et al.*, 2014).

Com o advento de técnicas de simulação numérica como as simulações de Monte Carlo por Cadeias de Markov (MCMC), esse problema pôde ser contornado. Com o seu uso, a ser detalhado no item 3.6, não há necessidade do cálculo do denominador da função de Bayes. Dessa maneira, as soluções Bayesianas para os problemas de inferência estatística se tornaram mais populares.

A técnica MCMC, de forma geral, consiste na amostragem de uma sequência de parâmetros ou vetores de parâmetros, em que, para cada iteração, o valor do parâmetro é avaliado e comparado com o do passo anterior com base em uma regra de probabilidade, conforme será descrito posteriormente. A partir dessas iterações, obtém-se N vetores de parâmetros que convergem para distribuição *a posteriori* desses parâmetros.

Como discutido no item 3.3, o uso da inferência Bayesiana é pertinente e útil para análises de incertezas para obtenção dos parâmetros da curva-chave e na análise de frequência de cheias, e, conseqüentemente, utilizada para incorporar essas incertezas de forma mais completa no estudo de cheias, enfoque principal do trabalho.

Assim, nos próximos itens serão apresentados o histórico do uso da técnica em cada caso e as formulações básicas das funções verossimilhança e das distribuições *a priori*.

3.5.1 Aplicação em Análises de Incertezas em Curvas-chaves

A aplicação da técnica de Inferência Bayesiana em análise de incertezas em curvas-chaves é bastante vantajosa. Os diferentes erros descritos no item 3.3, inerentes à determinação da curva-chave, podem ser incorporados à análise. Além disso, também é possível introduzir os conhecimentos hidráulicos *a priori* para restringir o valor dos parâmetros da curva-chave às suas limitações físicas e evitar estimativas irreais desses parâmetros.

Os estudos mais específicos de análise de incerteza em curvas chave por Inferência Bayesiana focaram principalmente nas incertezas devido à falta de dados de medição para uma curva de apenas um segmento (Reitan e Petersen Overleir, 2008), no estabelecimento, incluindo o trecho extrapolado, de funções multi-segmentadas (Reitan e Petersen, 2009), e em curvas-chaves submetidas a efeitos de remanso (Petersen-Overleir e Reitan, 2009). Porém, a maioria dos trabalhos realizados não explicitou claramente as leis dos controles hidráulicos e os expoentes da equação encontrados muitas vezes não tinham base física interpretável (Le Coz *et al.*, 2012).

Nesse contexto, Le Coz *et al.* (2014) desenvolveram um método com o intuito de melhorar a interpretação dos controles hidráulicos e das incertezas nas medições. Ele é uma generalização do método de base hidráulica, desenvolvido por Lang *et al.* (2010), e do método de representação estatística das incertezas avaliadas por inferência Bayesiana e Simulações Monte Carlo por Cadeia de Markov. Ele permite a estimativa de curvas-chaves com mais de um segmento, considerando o erro das medições e do ajuste da curva.

Sua grande vantagem é a introdução das características hidráulicas. As informações dos controles podem ser introduzidas a partir do conhecimento do comportamento do escoamento para cada faixa de vazão. Esse conhecimento é introduzido na forma de uma distribuição *a priori*. Cada controle ou conjunto de controles que podem estar ativados possuem um conjunto de parâmetros que representam cada segmento da curva-chave.

Com base no que foi proposto por Le Coz *et al.* (2014), o teorema de Bayes para relação cota vazão pode ser descrito conforme a Equação (3.32).

$$\underbrace{p(\theta, \sigma_f | \tilde{H}, \tilde{Q})}_{a \text{ posteriori}} \propto \underbrace{p(\tilde{Q} | \theta, \sigma_f, \tilde{H})}_{\text{verossimilhança}} \underbrace{p(\theta, \sigma_f)}_{a \text{ priori}}, \quad (3.32)$$

em que \tilde{H} e \tilde{Q} correspondem aos pares de cota e vazão obtidos em campo, θ representa o vetor de parâmetros da curva-chave (a, b e c) e σ_f corresponde ao desvio padrão do erro remanescente. Esse erro representa a qualidade do ajuste do modelo que devido a sua forma matemática, não consegue representar perfeitamente a relação cota x vazão mesmo se os valores reais fossem conhecidos.

O valor da vazão real (\tilde{Q}) para cada valor de cota (\tilde{H}) pode ser representada como a soma entre o valor calculado pela curva-chave estimada com base no valor da cota ($f(\tilde{H}_i | \theta)$) e os erros oriundos de medição e do ajuste da curva-chave, conforme Equações (3.33) e (3.34) (Le Coz *et al.*, 2014).

$$\tilde{Q}_i = f(\tilde{H}_i | \theta) + \epsilon_i^f + \epsilon_i^q \quad (3.33)$$

$$\epsilon_i^f + \epsilon_i^q \sim N\left(0, \sqrt{\sigma_f^2 + u_{Q_i}^2}\right) \quad (3.34)$$

em que ϵ_i^f e ϵ_i^q correspondem ao erro remanescente e ao erro de medição da vazão, respectivamente. O desvio padrão do erro de medição é conhecido e depende do instrumento utilizado em campo, ao passo que o erro remanescente é desconhecido e introduzido como parâmetro a ser estimado no teorema de Bayes (Eq. (3.32)).

A equação de verossimilhança pode então ser descrita da seguinte forma (Le Coz *et al.*, 2014):

$$p(\tilde{Q}|\theta, \sigma_f, \tilde{H}) = \prod_{i=1}^N p_N\left(\tilde{Q}_i | f(\tilde{H}_i|\theta), \sqrt{\sigma_f^2 + u_{Q_i}^2}\right), \quad (3.35)$$

em que $p_N(z | m, s)$ é a função normal de densidade de probabilidade de média m e desvio padrão s , avaliada no ponto z .

Por fim, concluindo as formulações necessárias à análise Bayesiana para curvas-chaves, as informações *a priori* podem ser definidas conforme a Equação (3.36). Ressalta-se que, como se tratam de curvas segmentadas, para cada trecho de curva tem-se um vetor de parâmetros cujas *a priori* devem ser definidas. Ou seja, é calculado uma distribuição *a posteriori* para os parâmetros de cada segmento da curva. O número de segmentos da curva é dado de entrada para o modelo em conjunto com as informações *a priori* dos parâmetros de cada segmento.

$$p(\theta|\sigma_f) = p(\sigma_f) \prod_{j=1}^{Npar} p(\theta_j) \quad (3.36)$$

em que $p(\theta_j)$ representa a informação *a priori* de cada um dos parâmetros daquele segmento da curva-chave. No caso de aplicação de procedimento numérico para obtenção da distribuição *a posteriori* dos parâmetros, a saída final desse modelo são N valores de parâmetros e, conseqüentemente, N curvas-chaves possíveis com maior ou menor probabilidade de ocorrência. As análises dos intervalos de credibilidade são portanto, fáceis e imediatas.

De maneira similar, Reitar e Overleir (2009) propuseram método para estimativa de curvas-chaves multisegmentadas a partir de um modelo de regressão com erros multiplicativos e normalmente distribuídos (Eq.(3.37)),

$$\log(Q_i) = \sum_{j=1}^m I(b_{s,j-1} \leq b_i < b_{s,j}) \left(\log(a_j) + c_j(h_i - b_{0,j}) \right) + \varepsilon_i \quad (3.37)$$

em que j é o número do segmento dos m segmentos, $I(\cdot)$ representa a função indicador, resultando em 0 ou 1 caso o argumento seja falso ou verdadeiro, $b_{s,j}$ representa o ponto de mudança na curva de um segmento para o outro, e ε_i representa o ruído considerado independente e identicamente distribuído por uma log normal ($\varepsilon_i \sim N(0, \sigma^2)$).

Nesse modelo o número de segmentos (m) é desconhecido e não é estimado simultaneamente com os demais parâmetros. Testa-se o modelo para diferentes tamanhos de m e adota-se aquele com maior valor de probabilidade *a posteriori*. Esse modelo permite a incorporação de informação *a priori* em relação às características hidráulicas da seção, mas não leva em consideração as incertezas nas medições de vazão. Porém, apesar de subestimar a incerteza, é um modelo igualmente eficiente para estimativa dos parâmetros devendo apenas ser melhorado para consideração de modelos de erro mais completos.

Diante do exposto, adotou-se nesse trabalho a estrutura Bayesiana proposta por Le Coz *et al.* (2014) considerando curvas com apenas um segmento.

3.5.2 Aplicação em Estudos de Frequência de Cheias

A estrutura Bayesiana foi aplicada em estudos de cheias, por exemplo, por Coles e Tawn (1996), Kuczera (1999), Reis e Stedinger (2005), O'Connell (2005) e Neppel *et al.* (2010), comparando os resultados obtidos com métodos estatísticos clássicos de verossimilhança. Nesses estudos, à exceção de Coles e Tawn (1996), que trabalharam com precipitações, se avaliou também a incorporação de informações de enchentes históricas e o impacto da introdução de *prioris* mais informativas sobre os parâmetros.

Kuczera (1999) apresentou um método de inferência Bayesiana com a aplicação de simulações Monte Carlo e métodos de busca por gradientes para calcular a distribuição de probabilidade esperada e os limites de confiança para qualquer distribuição de frequência, empregando dados de vazão corrompidos. Os erros nos dados de vazão aplicados por ele poderiam estar associados tanto a erros na curva-chave quanto nas medições da vazão.

Seguindo a mesma linha, mas se utilizando de método numérico, Reis e Stedinger (2005) propuseram metodologia de inferência Bayesiana com simulações Monte Carlo por Cadeias de Markov (MCMC). O trabalho consistiu na avaliação de distribuições *a posteriori* de quantis de cheia, riscos de inundação e dos parâmetros para as distribuições Log normal e log Pearson III e na verificação das vantagens desse método quando comparado com a metodologia tradicional de máxima verossimilhança.

De acordo com Reis e Stedinger (2005), o uso de Análise Bayesiana para ajuste das distribuições de frequência de cheia é bastante atrativo por permitir: uma análise completa dos dados, incluindo as descrições apropriadas das incertezas em qualquer função dos parâmetros; uma vasta descrição das distribuições conjuntas das incertezas dos dados,

incluindo a magnitude dos eventos de pico de extrapolação de níveis máximos, e a incorporação dessas incertezas, sem dificultar os procedimentos numéricos.

Coles e Tawn (1996) demonstraram também, a partir de estudo de precipitações máximas, que o conhecimento e a incorporação de informação *a priori* em análise de eventos extremos é viável e que, se as especificações *a priori* forem tecnicamente fundamentadas, os benefícios de se usar análise Bayesiana em detrimento de análises de verossimilhança são consideráveis.

O'Connel (2005) foi um pouco mais além e propôs uma estrutura Bayesiana de estimativa não paramétrica das frequências de cheias. De acordo com o autor, o número restrito de distribuições disponíveis para os ajustes das vazões pode resultar em valores viesados dos quantis.

Todos os métodos desenvolvidos se mostraram bastante eficientes no cálculo dos parâmetros e da estimativa das incertezas, todavia, consideraram as incertezas nos valores das séries de vazão de forma ainda superficial, indicando a possibilidade de elaboração de um método que consiga englobar de maneira mais específica as incertezas dos parâmetros.

No caso do uso de inferência Bayesiana aplicado ao estudo de frequência de cheias, o foco é a estimativa dos parâmetros do modelo probabilístico adotado. A distribuição *a priori* desses parâmetros mostram o quanto se sabe sobre aquele parâmetro antes da observação dos dados.

Dessa maneira no teorema de Bayes os dados utilizados para o cálculo da verossimilhança passam a ser as vazões máximas anuais e os parâmetros a serem estimados são os parâmetros da distribuição de frequência de cheias adotada. No caso da distribuição GEV o teorema de Bayes pode ser descrito como:

$$\overbrace{p(\alpha, \xi, \kappa | Q_{max,i})}^{Posteriori} \propto \overbrace{p(Q_{max,i} | \alpha, \xi, \kappa)}^{Verossimilhança} \overbrace{p(\alpha)p(\xi)p(\kappa)}^{A\ Priori} \quad (3.38)$$

em que $p(\alpha)$, $p(\xi)$ e $p(\kappa)$ representam a distribuição *a priori* dos parâmetros que descreve o conhecimento anterior à análise dos dados $Q_{max,i}$. Essa *a priori* pode ser não informativa ou pode representar um conhecimento regional dos dados de vazão (Reis e Stedinger, 2005; Martins *et al.*, 2008).

O segundo termo representa a função verossimilhança dos parâmetros para o conjunto de dados de vazões máximas. Por fim, $p(\alpha, \xi, \kappa | Q_{max,i})$ representa a distribuição *a posteriori* dos parâmetros α , ξ , κ condicionados à observação de Q_{max} .

O procedimento apresentado não considera a incorporação de erro nas medições. Conforme apresentado no item 3.3, vários estudos propuseram a consideração do erro nas medições de vazão e na curva-chave em estudos de cheia por inferência Bayesiana. Porém, a maioria deles introduziu o erro diretamente na estrutura Bayesiana de frequência de cheias, sem efetivamente considerar as questões mais específicas sobre a origem do erro.

Nessa incorporação direta no estudo de frequência de cheias, basta que ele seja introduzido como parâmetro e que sua distribuição *a priori* seja fornecida de forma semelhante ao que foi proposto na incorporação das incertezas na estimativa das curvas-chaves.

3.5.3 Aplicação em Estudos de Frequência de Cheias com incorporação de incertezas das Curvas-Chaves

A aplicação da inferência Bayesiana para incorporação das incertezas das curvas-chaves em estudos de frequência de cheias ainda é muito recente. Poucos estudos se propuseram a utilizá-la para tentar compreender a importância dessas incertezas nos quantis de cheia.

Recentemente, Osorio e Reis (2016) e Steinbakk *et al.* (2016) propuseram estudos nesse sentido que começam a mostrar as vantagens do uso da técnica, a importância da consideração do máximo de fatores intervenientes no cálculo dos quantis e as melhorias ainda necessárias para se chegar a uma análise mais eficiente.

Em ambos os trabalhos se utilizou uma estrutura Bayesiana para a estimativa da distribuição *a posteriori* dos parâmetros da curva-chave. Em Osorio e Reis (2016), a estrutura adotada foi aquela proposta por Le Coz *et al.* (2014). Por outro lado, em Steinbakk *et al.* (2016) adotaram a estrutura proposta por Reitan e Petersen-Overleir (2009) de estimativa de curvas-chaves multisegmentadas. Ambas apresentadas no item 3.5.1. Nas duas estruturas se considera que a curva-chave se comporta conforme a Equação (3.23). A diferença é que em Reitan e Petersen Overleir (2009), o número de segmentos da curva é um parâmetro a ser estimado e as características hidráulicas não são consideradas na estimativa dos parâmetros.

Para a incorporação das incertezas da curva-chave nos quantis, primeiro são geradas as distribuições *a posteriori* dos parâmetros da curva-chave, em ambos os estudos, representadas por uma amostra de 10.000 vetores desses parâmetros gerados por simulações

Monte Carlo via Cadeia de Markov (MCMC). Em seguida cada conjunto desses vetores é aplicada em uma nova estrutura para estimativa dos quantis de cheia.

No caso de Osorio (2016), a distribuição *a posteriori* dos parâmetros da curva-chave, representada pelos conjuntos de parâmetros amostrados, foram utilizados para geração das séries de vazões máximas anuais. Em seguida, os parâmetros da GEV e quantis foram estimados por meio do método de Máxima Verossimilhança Generalizada (ver item 3.1.1). Para o cálculo da incerteza associada a esses quantis, o método Delta descrito no item 3.1.2 foi adotado para cada um dos conjuntos de parâmetros e quantis estimados. Por fim, o intervalo de confiança foi definido pela média de todos os intervalos de confiança obtido para cada um dos vetores. O objetivo desse estudo era apresentar uma metodologia, que tornasse possível a incorporação da incerteza da curva-chave nos quantis de cheia. Apesar de já demonstrar a importância da incorporação dessas incertezas, o estudo possui a limitação da metodologia proposta não estimar os parâmetros de forma conjunta, adotando diferentes metodologias em cada etapa, não ficando demonstrado, portanto, que a incerteza estimada de fato representa a incerteza final dos quantis.

Por outro lado, Steinbakk *et al.* (2016) utilizam uma segunda estrutura Bayesiana para estimativa dos parâmetros da curva de frequência de cheias, a exemplo do que foi apresentado em 3.5.2. Nesse sentido, para cada um dos 10.000 vetores, são gerados 3.000 vetores (valor de simulações adotado no estudo) dos parâmetros da GEV. A incerteza do quantil é derivada da amostra dos $10.000 \times 3.000 = 30.000.000$ de vetores. Além dessa estimativa, eles propõem uma metodologia para investigar separadamente a incerteza amostral e da curva-chave de forma a avaliar as contribuições dessas incertezas na incerteza combinada. Essa metodologia está ilustrada na Figura 3-6.

A estrutura apresentada na Figura 3-6 foi aplicada a 8 estações da Noruega com diferentes áreas de drenagem e disponibilidade de dados. O estudo concluiu que na maioria das estações a incerteza amostral era a maior responsável pela incerteza nos quantis, mas naquelas estações em que era necessária a extrapolação da curva-chave, a incerteza da curva-chave passava a ter papel mais significativo. Isso mostra que o resultado depende muito das características do local e do nível de informação disponível. Além disso, eles demonstraram, em duas estações específicas, o ganho de se aumentar o tamanho das séries de cotas máximas anuais e de se dispor de medições de vazões de cheia, o que reduz a incerteza no valor dos quantis caso não haja aumento da variabilidade amostral das vazões.

Mais uma vez, observa-se que a incorporação das incertezas da curva-chave é fundamental para uma análise completa e fidedigna da realidade local. Apesar de comprovar

esse fato de forma mais completa que em estudos anteriores, o trabalho de Steinbakk *et al.* (2016) ainda possui algumas limitações. Entre elas, destacam-se que: não foi considerado no modelo de estimativa dos parâmetros da curva-chave, a parcela do erro de medição implicando na subestimação da incerteza da curva-chave; foram consideradas *prioris* não informativas para todos os parâmetros sem incorporação de informação adicional que poderia contribuir na melhoria das estimativas, não sendo avaliada a importância dessa informação no resultado final; não foram consideradas as características hidráulicas dos locais, apenas os dados observados são utilizados na análise; a estrutura Bayesiana proposta estima os parâmetros separadamente exigindo um esforço computacional considerável.

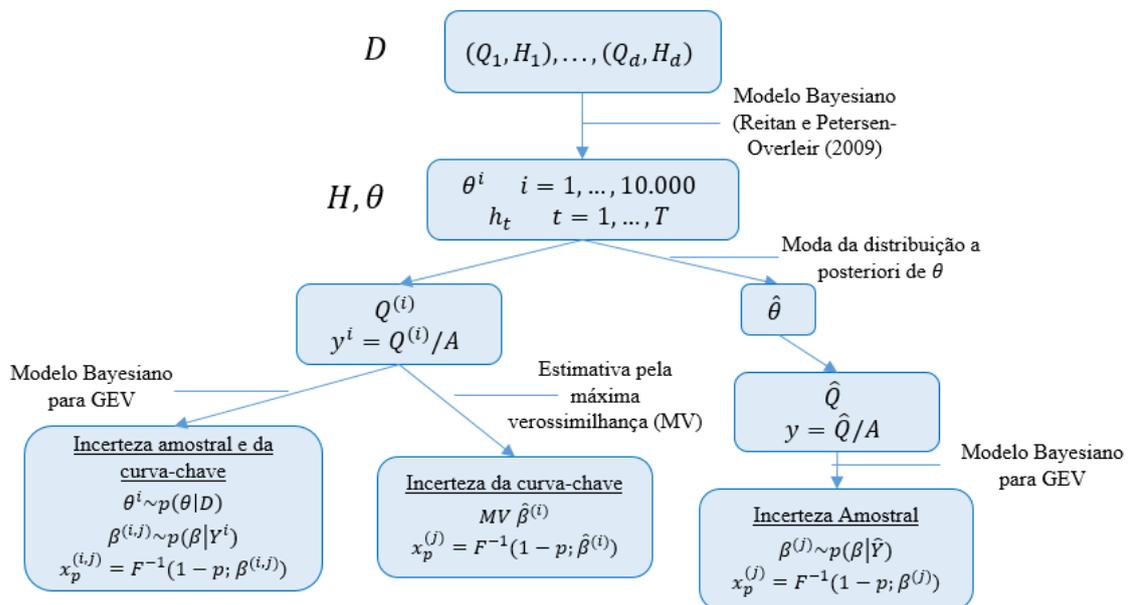


Figura 3-6 Visão geral da estrutura proposta por Steinbakk *et al.* (2016). D representa os dados de resumo de descarga utilizados para estimativa dos parâmetros da curva-chave (θ). H representa a séries de cotas diárias. A partir de H e θ , são calculadas as vazões máximas anuais normalizadas pela área de drenagem (Y_i). β representa os parâmetros da distribuição de frequências, no caso, da GEV. j representa o número de iterações da GEV (3000) e x_p representa os quantis de cheia obtidos a partir da função do quantil F^{-1}

Diante do exposto, verifica-se que ainda é necessária a proposição de uma estrutura Bayesiana que:

1. Estime os parâmetros, tanto da curva-chave como da distribuição de frequências, simultaneamente;
2. Incorpore as incertezas da curva-chave e informação *a priori*, incluindo a incerteza do ajuste do modelo e de medição;

3. Considere explicitamente as características hidráulicas da seção em análise.

Dessa maneira, este trabalho tenta vencer essa lacuna para uma estrutura de curva-chave com um único segmento conforme será apresentado no item 4.

3.6 SIMULAÇÕES MONTE CARLO POR CADEIAS DE MARKOV

No passado, a grande dificuldade na utilização de metodologias de inferência Bayesiana residia no cálculo da distribuição *a posteriori* em função da integral da constante normalizadora da fórmula de Bayes (Eq. (3.31)). No entanto, com o desenvolvimento dos computadores e de métodos numéricos de amostragem esse cálculo se tornou possível e mais acessível.

Um desses métodos numéricos é o de Simulações Monte Carlo por Cadeias de Markov. A simulação de Markov é utilizada quando não é possível (ou não é computacionalmente eficiente) amostrar θ diretamente da distribuição *a posteriori*. Ao invés disso, se produzem amostras iterativamente a partir de uma distribuição proposta, esperando-se que, a cada passo do processo, a distribuição de onde se obtém as amostras se assemelhe a $p(\theta|x)$.

As Cadeias de Markov em probabilidade são cadeias, conforme descrito por Gelman *et al.* (2014), que em determinadas condições possuem propriedades bastante interessantes para a determinação da distribuição *a posteriori*. A primeira refere-se à probabilidade de se encontrar em um determinado ponto da cadeia depender única e exclusivamente do ponto anterior, não importando como se chegou lá. A segunda, por sua vez, refere-se ao fato de que a probabilidade de passar de um ponto para outro se torna invariante e independe do ponto inicial de partida depois de um certo tempo dentro da cadeia. Para isso, o processo de amostragens deve ser simulado durante tempo suficiente para que a condição invariante, também chamada de estacionária, seja atingida e que esta represente adequadamente a distribuição *a posteriori* procurada.

Existem alguns algoritmos que aplicam esse processo. A grande diferença entre eles é como é desenvolvido o caminho aleatório dentro da cadeia. Dentre esses algoritmos destacam-se os algoritmos de Metropolis, Metropolis Hastings e Metropolis Adaptado.

Para verificação se a condição estacionária foi atingida, existem algumas metodologias de análise de convergência das cadeias que também serão melhor descritas neste item.

3.6.1 Algoritmo Metropolis

O algoritmo Metropolis propõe um caminho aleatório, em que são aplicadas regras de rejeição e aceitação para definir o passo seguinte na simulação e convergir para a distribuição procurada. O algoritmo, conforme descrito por Gelman *et al.* (2014), funciona da seguinte maneira:

- a) Define-se um ponto de partida θ^0 , em que $p(\theta^0|x) > 0$, a partir de uma distribuição inicial $p_0(\theta)$.

Para os passos seguintes $t=1,2, \dots, N$:

- b) Amostra um valor proposto para θ^* a partir de uma distribuição denominada proposta no tempo t , $q_t(\theta^*|\theta^{t-1})$. No caso do algoritmo Metropolis, essa distribuição deve ser simétrica e satisfazer a seguinte condição:

$$q_t(\theta_a|\theta_b) = q_t(\theta_b|\theta_a) \quad (3.39)$$

Que significa não existe preferência dentro da cadeia entre ir de a para b ou de b para a .

- c) Calcula a razão entre as *posteriors*:

$$\alpha(\theta^{t-1}, \theta^*) = \min \left[\frac{p(\theta^*|x)}{p(\theta^{t-1}|x)}, 1 \right] \quad (3.40)$$

- d) Define-se, então:

$$\theta^t = \begin{cases} \theta^* & \text{com probabilidade } \alpha \\ \theta^{t-1} & \text{com probabilidade } (1 - \alpha) \end{cases}$$

- e) Gera um valor aleatório u entre 0 e 1 a partir de uma distribuição uniforme e compara-se o valor de α calculado com o valor gerado:

Se $u \leq \alpha(\theta^{t-1}, \theta^*)$, então : $\theta^{t+1} = \theta^*$, aceita-se o valor proposto;

Caso Contrário, $\theta^{t+1} = \theta^{t-1}$, o valor proposto é rejeitado.

Ao final da simulação temos os valores aceitos obtidos de $\{\theta^1, \theta^2, \dots, \theta^N\}$ da distribuição *a posteriori*.

3.6.2 Algoritmo Metropolis Hastings

O algoritmo de Metropolis Hastings é uma evolução do algoritmo de Metropolis, no qual a distribuição proposta não precisa necessariamente ser simétrica. Nesse sentido, o valor de α é definido para reajustar o equilíbrio na cadeia da seguinte maneira (Gelman *et al.*, 2014):

$$p(\theta^{t-1}|x)q(\theta^{t-1}, \theta^*)\alpha(\theta^{t-1}, \theta^*) = p(\theta^*|x)q(\theta^*, \theta^{t-1})\alpha(\theta^*, \theta^{t-1}) \quad (3.41)$$

Se $p(\theta^{t-1}|x)q(\theta^{t-1}, \theta^*) > p(\theta^*|x)q(\theta^*, \theta^{t-1})$, então o movimento de θ^{t-1} para θ^* ocorre mais frequentemente. Assumindo que $\alpha(\theta^*, \theta^{t-1}) = 1$, que seria a maior probabilidade possível, calculamos o valor de $\alpha(\theta^{t-1}, \theta^*)$ a partir da Equação (3.42).

$$\alpha(\theta^{t-1}, \theta^*) = \min \left[\frac{p(\theta^*|x)q(\theta^*, \theta^{t-1})}{p(\theta^{t-1}|x)q(\theta^{t-1}, \theta^*)}, 1 \right] \quad (3.42)$$

De posse do novo valor de $\alpha(\theta^{t-1}, \theta^*)$, que permite o balanço entre as probabilidades, são seguidos os mesmos passos do algoritmo Metropolis, substituindo a fórmula da equação de α em (3.40) pela Equação (3.42).

3.6.3 Algoritmo Metropolis Adaptado (AM)

O Algoritmo Metropolis Adaptado (AM) proposto por Haario *et al.* (2001) é uma variação do algoritmo Metropolis convencional descrito acima. Nesse algoritmo a distribuição proposta é baseada na matriz de covariância *a posteriori* dos parâmetros. Essa matriz é calculada em cada iteração a partir das iterações anteriores. Assim, a distribuição proposta é atualizada com base no conhecimento adquirido até aquele momento sobre a distribuição *a posteriori*. Para obtenção do vetor de parâmetros proposto no passo i , por exemplo, amostra-se um valor de uma distribuição normal multivariada com média dada

pelos valores dos parâmetros no passo atual e matriz de variância-covariância C_i , descrita como:

$$C_i = \begin{cases} C_0 & ; \text{para } i \leq i_0 \\ s_d \text{Cov}((\theta_0, \dots, \theta_{i-1}) + s_d \varepsilon I_d) & ; \text{para } i > i_0 \end{cases} \quad (3.43)$$

em que, ε é um parâmetro de pequeno valor para garantir que a matriz C_t não se torne singular, I_d é a matriz identidade, e s_d corresponde a um parâmetro de escala que depende da dimensão d do vetor de parâmetros θ para garantir taxas de aceitação razoáveis para os valores propostos. Haario *et al.* (2001) sugere que s_d seja calculado como:

$$s_d = 2,4^2/d \quad (3.44)$$

A matriz de covariância na iteração $i + 1$ pode ser calculada de forma a satisfazer a seguinte equação:

$$C_{i+1} = \frac{i-1}{i} C_i + \frac{s_d}{i} (i \bar{\theta}_{i-1} \bar{\theta}_{i-1}^T - (i+1) \bar{\theta}_i \bar{\theta}_i^T + \theta_i \theta_i^T + \varepsilon I_d) \quad (3.45)$$

O cálculo da matriz de covariância da distribuição proposta requer a definição de um valor arbitrário inicial C_0 . Esse valor inicial pode ser obtido a partir das variâncias das distribuições *a priori* dos parâmetros. O algoritmo adaptado funciona da seguinte maneira:

- a) Inicializa $i = 0$;
- b) Calcule C_i a partir da Equação ((3.43);
- c) Gere o vetor de valores propostos θ^* , em que $\theta^* \sim N(\theta^i, C_i)$
- d) A razão de aceitação e o aceite da cadeia é realizado da mesma forma que no algoritmo de Metropolis Hastings.

Para valores propostos fora dos limites aceitáveis para o parâmetro a probabilidade de aceitação é considerada nula. De acordo com Marshall (2005), a maior vantagem dessa forma de abordagem é que o todo o vetor de parâmetros é atualizado simultaneamente. Isso reduz o tempo computacional e a complexidade sendo extremamente útil em caso em que os parâmetros são correlacionados.

Em virtude das vantagens desse algoritmo para vetores com grande número de parâmetros que possam apresentar correlação entre si, ele será adotado nesse trabalho para obtenção da distribuição *a posteriori* dos parâmetros da curva-chave e da distribuição de frequência de cheias.

3.6.4 Verificação da Convergência

A performance do algoritmo depende da taxa de aceitação/rejeição dos valores propostos, a qual depende, por sua vez, da distribuição proposta escolhida. Se na distribuição adotada os valores propostos são rejeitados com muita frequência, as simulações não conseguirão abranger o universo de valores possíveis e dificilmente convergirão. Por outro lado, se os valores são muito aceitos o processo levará muito tempo para convergir.

Busca-se, dessa maneira, uma distribuição que permita a avaliação dos valores possíveis da distribuição proposta em um tempo de simulação razoável. A Figura 3-7 mostra a comparação entre dois algoritmos baseados em distribuições propostas diferentes, na qual é possível observar a diferença entre os tempos de convergência.

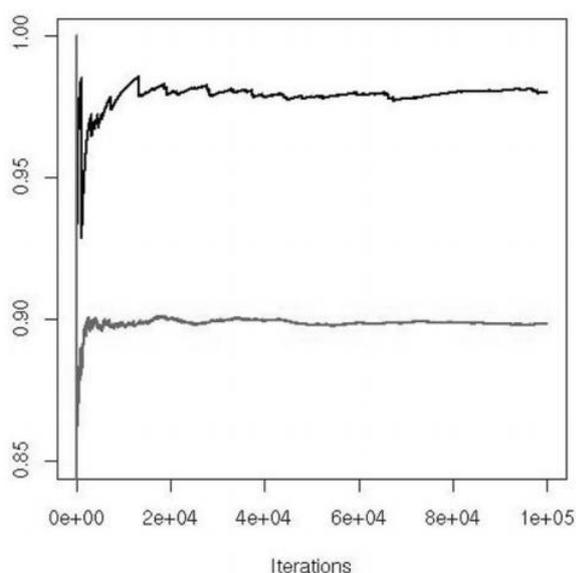


Figura 3-7 – Comparação entre sequências geradas pelo algoritmo Metropolis Hasting baseadas em distribuições propostas diferentes (Casella, 2010).

Em alguns casos, mesmo para muitas iterações, a cadeia não converge para a distribuição *a posteriori* procurada. Como as inferências são feitas com base na cadeia amostrada, é fundamental verificar a convergência da cadeia para validação dos resultados obtidos nas simulações MCMC.

A forma mais imediata de se verificar tal convergência é fazer uma avaliação visual do comportamento do traço da cadeia. A partir dessa análise visual, é possível avaliar o comportamento da cadeia e inferir o momento em que ela atinge estabilidade e passa a oscilar entre determinados valores extremos. O gráfico da Figura 3-8 demonstra como visualmente é possível obter conclusões acerca da convergência da cadeia.

Outra forma de avaliação visual da convergência consiste em plotar o comportamento da média dos valores dos parâmetros pelas iterações. Espera-se que em situações em que a cadeia convergiu, que a média se estabilize entorno de um valor único. A Figura 3-9 apresenta dois gráficos em situação de convergência e não convergência.

Além dos métodos visuais descritos, existem métodos mais assertivos para avaliação da convergência. Entretanto, a aplicação desses métodos está vinculada ao algoritmo utilizado para geração da cadeia (Martinez e Martinez, 2002). O método de Gelman e Rubin (1992) é uma exceção a essa regra e pode ser aplicado para qualquer tipo de algoritmo MCMC. Esse método foi escolhido para validação da convergência das cadeias obtidas pelo modelo proposto.

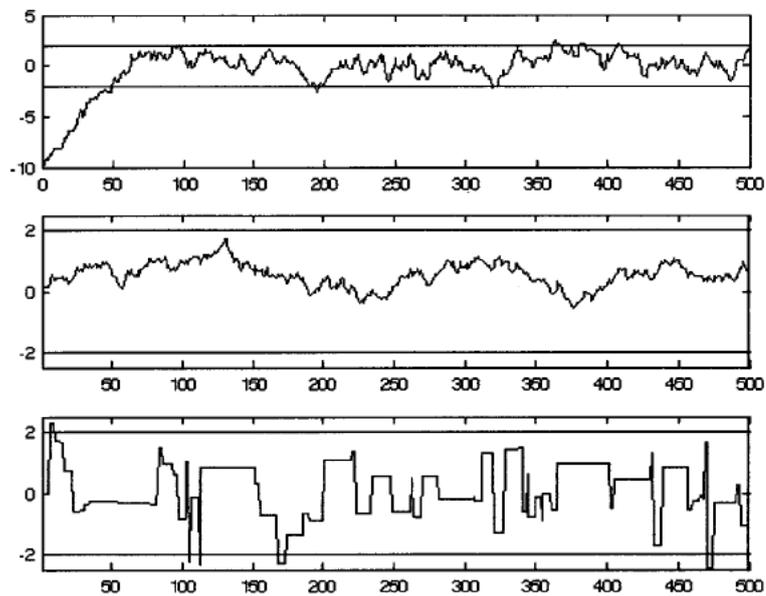


Figura 3-8 - Comparativo entre diferentes cadeias de Markov (Martinez & Martinez, 2002)

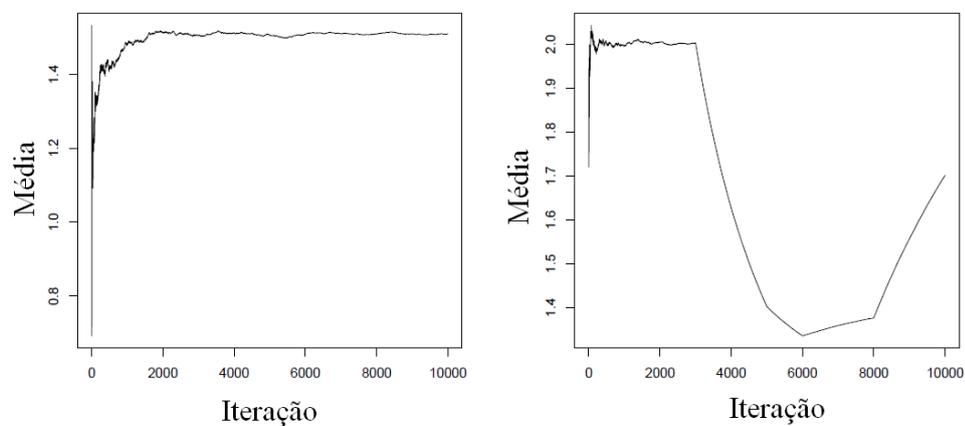


Figura 3-9 - Comportamento da média ao longo das iterações da cadeia, indicando sinais de convergência (esquerda) e não houve convergência(direita) (Rampinelli,2016)

O método se baseia na ideia de que a variância dentro de uma única cadeia será menor do que a variância nas sequências de cadeias combinadas, caso a convergência não tenha sido atingida (Martinez & Martinez, 2002).

O diagnóstico de Gelman-Rubin é realizado com base na simulação de múltiplas cadeias. Supondo que m cadeias de tamanho n tenham sido simuladas, com diferentes valores de parâmetros iniciais em que n corresponde ao total de valores da cadeia sem considerar o período inicial de aquecimento, calcula-se:

$$B = \frac{n}{m-1} \sum_{j=1}^m (\bar{\theta}_j - \bar{\bar{\theta}})^2, \text{ onde } \bar{\theta}_j = \frac{1}{n} \sum_{i=1}^n \theta_{ij}, \bar{\bar{\theta}} = \frac{1}{m} \sum_{j=1}^m \bar{\theta}_j \quad (3.46)$$

$$W = \frac{1}{m} \sum_{j=1}^m s_j^2, \text{ onde } s_j^2 = \frac{1}{n-1} \sum_{i=1}^n (\theta_{ij} - \bar{\theta}_j)^2 \quad (3.47)$$

A partir dos resultados das equações acima, é possível estimar a variância marginal *a posteriori* dos parâmetros como uma média ponderada de W e B , da seguinte maneira:

$$\widehat{Var}(\theta) = \frac{n-1}{n} W + \frac{1}{n} B \quad (3.48)$$

A convergência ao longo das iterações é monitorada a partir da estimativa do fator de redução \hat{R} . Esse fator de redução é calculado como:

$$\hat{R} = \sqrt{\frac{\widehat{Var}(\theta)}{W}} \quad (3.49)$$

O valor de \hat{R} reduz para 1 a medida que $n \rightarrow \infty$. Se o fator de redução é elevado (maior que 1,2), então temos razões para acreditar que a cadeia ainda não convergiu e que mais iterações devem ser realizadas para melhorar a convergência e garantir que a distribuição obtida se aproxima da distribuição procurada.

4 METODOLOGIA

Nesse capítulo está apresentada a metodologia proposta para alcançar os objetivos geral e específicos do trabalho. Primeiramente, é apresentada a estratégia de ação para implementação do modelo e para análise da importância relativa das incertezas da curva-chave. Na sequência, os passos descritos na estratégia são melhor detalhados nos itens 4.2 sobre o modelo Bayesiano Completo de incorporação de incertezas da Curva-Chave - BC, 4.3, sobre o modelo Bayesiano apenas para GEV, 4.4, sobre o impacto das incertezas da curva-chave na análise de frequência de cheias e 4.5, sobre os estudos de caso que serão utilizados para aplicação da metodologia proposta.

4.1 Estratégia de Ação

Tendo em vista o objetivo do estudo de avaliar o impacto da incorporação das incertezas da curva-chave na análise de frequência de cheias por meio da proposição de uma metodologia Bayesiana, foram adotados os passos apresentados no fluxograma da Figura 4-1. No fluxograma está apresentado também como a execução desses passos se relacionam com a conclusão de cada um dos objetivos específicos propostos no trabalho.

Primeiramente são definidas as incertezas que serão incorporadas no modelo, tanto da curva-chave, como do ajuste à distribuição de frequência e como essas incertezas serão traduzidas na forma de um modelo de erros que as represente (Passo 1). Em paralelo, define-se a função verossimilhança e a distribuição das informações *a priori* que melhor descrevem o problema em questão (Passos 2 e 3).

Com base nas informações supracitadas, desenvolve-se o Modelo Bayesiano Completo (Passo 4), que nesse trabalho é resolvido numericamente por meio da implementação do Algoritmo Metropolis Adaptado de simulações Monte Carlo por Cadeia de Markov (MCMC) (Passo 5 e 6).

O modelo é então aplicado a dois estudos de caso: um real com disponibilidade de dados e informações de campo acerca das características hidráulicas (Passo 6a); e um sintético onde todos os parâmetros são conhecidos e os dados e os erros associados a eles são gerados sinteticamente (Passo 6a e 6b). Com base nos resultados obtidos, é calculada a importância relativa (IR) das incertezas da curva-chave em relação à incerteza global para diferentes cenários de informação a priori (P e P-), disponibilidade de medição de descarga (D e D-) e disponibilidade de dados de cotas máximas anuais (H e H-) (Passos 7 e 8).

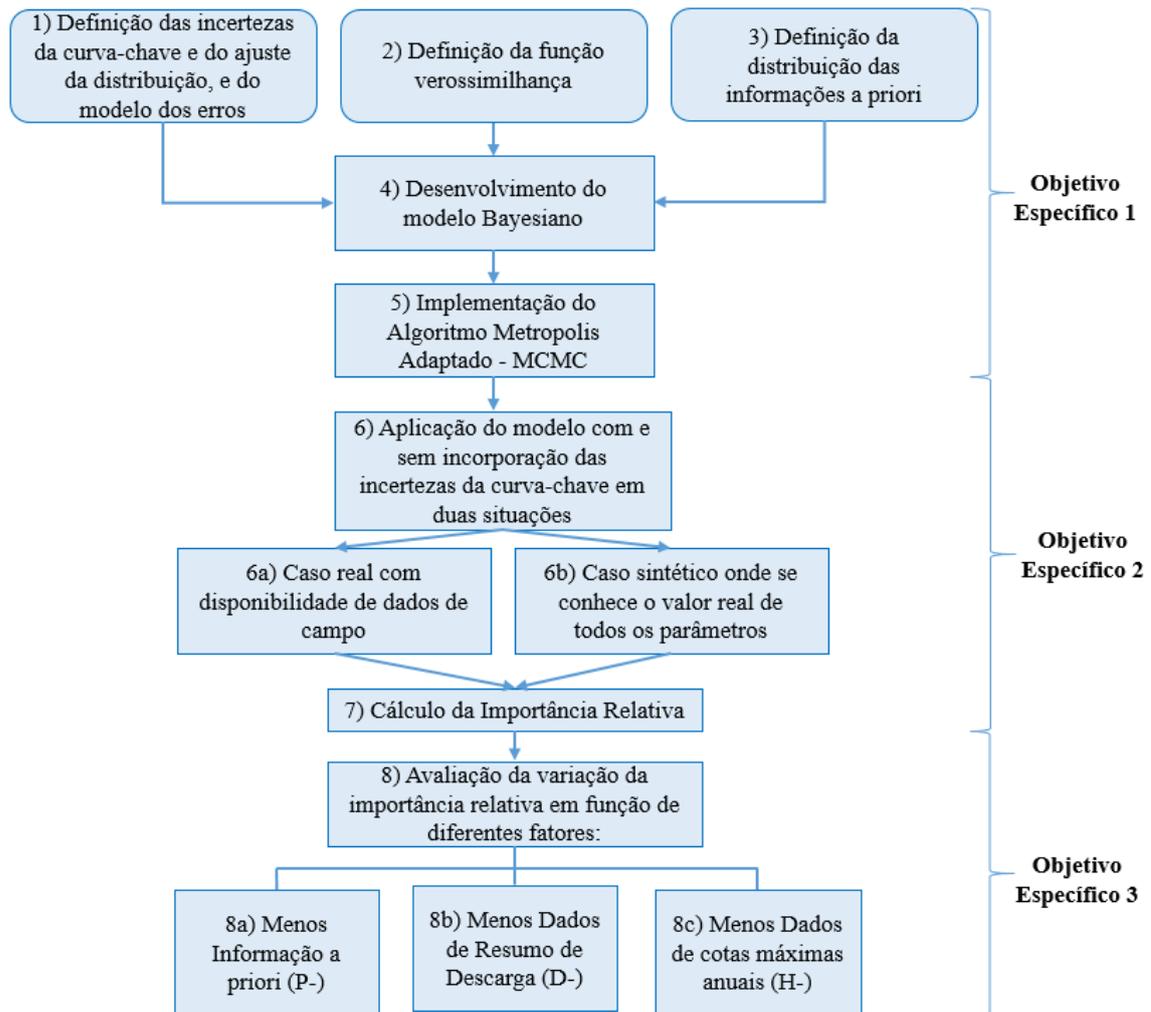


Figura 4-1 - Fluxograma de Atividades

4.2 Modelo Bayesiano Completo de Incorporação de Incertezas da Curva-Chave - BC

O modelo proposto neste trabalho tem o objetivo de realizar uma análise de frequência de cheias com incorporação das incertezas do ajuste de uma curva-chave aos dados de medição de descarga. A incorporação dessas incertezas é feita por meio de uma análise Bayesiana, em que os parâmetros da curva-chave são estimados de forma simultânea aos parâmetros da distribuição teórica de probabilidade de ocorrência de cheias. Conforme descrito anteriormente, a distribuição adotada neste trabalho para a definição da curva de frequência de cheias é a Generalizada de Valores Extremos (GEV).

O modelo proposto, como será apresentado com detalhes mais adiante, permite que se considere de forma explícita, por meio de distribuições *a priori* dos parâmetros da equação

da curva-chave, os erros de medição de descarga, relacionados com o tipo de aparelho empregado, e as estimativas, e respectivas incertezas, de características da seção hidráulica do rio, tais como a largura, rugosidade e declividade do trecho.

A estimativa simultânea dos parâmetros da equação da curva-chave e da distribuição GEV é realizada de forma empírica, por meio de Simulações Monte Carlo via Cadeias de Markov (MCMC). O algoritmo adotado no trabalho foi o Metropolis Adaptado (AM) apresentado no item 3.6.3. O resultado do modelo proposto é a geração de distribuições *a posteriori* de cada um dos parâmetros a partir das cadeias geradas, assim como a distribuição *a posteriori* dos quantis de cheia para cada tempo de retorno. Dessa distribuição, se calculam a curva de frequência de cheias e seus intervalos de credibilidade associados à todas as incertezas incorporadas ao modelo.

Para que o modelo possa ser empregado, são necessárias as seguintes informações: resumo de medição de descarga na seção de interesse, que consiste de pares de cota e vazão e seus respectivos erros de medição; as características da seção hidráulica como largura, rugosidade e declividade e suas incertezas para que sejam utilizadas na definição das distribuições *a priori* dos parâmetros da curva-chave (a , b e c); a série de cotas máximas anuais e as informações *a priori* dos parâmetros do modelo probabilístico adotado.

A Figura 4-2 apresenta de forma mais clara o modelo com seus dados de entrada e saída.

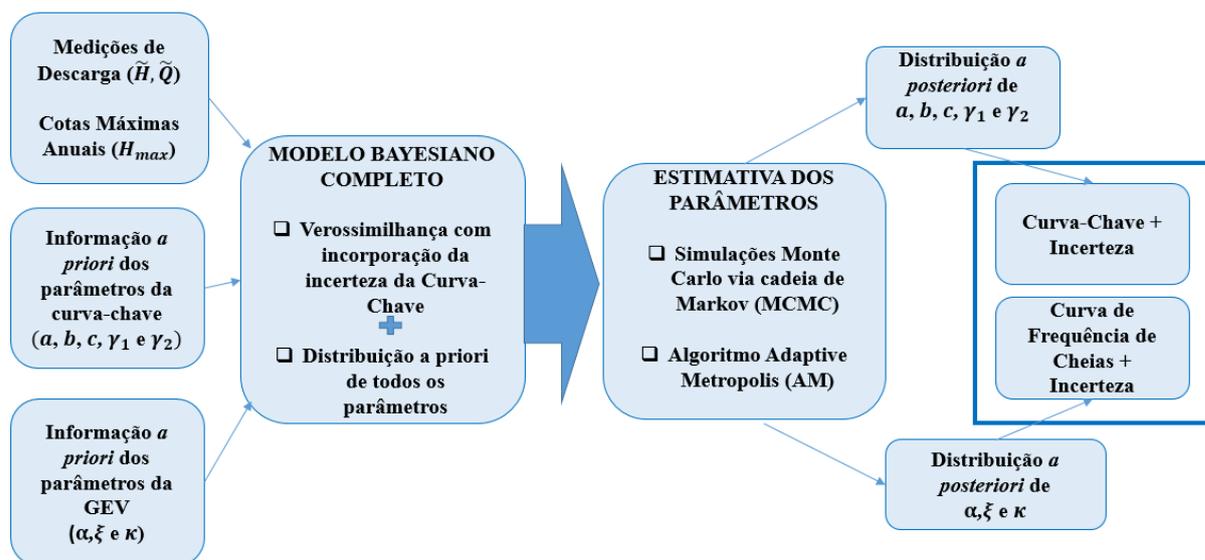


Figura 4-2 – Modelo Bayesiano completo proposto

Os itens a seguir apresentam detalhes sobre a formulação da função verossimilhança e das distribuições a priori, assim como sobre a estratégia de obtenção da distribuição *a posteriori* dos parâmetros e dos quantis de cheia.

4.2.1 Função de Verossimilhança

Com base nas formulações apresentadas nos itens 3.5.1 e 3.5.2, a distribuição *a posteriori* dos parâmetros da equação da curva-chave e da distribuição GEV pode ser descrita como sendo proporcional ao produto da função verossimilhança dos dados observados (primeiro termo da equação) com as distribuições *a priori* dos parâmetros (segundo termo da equação):

$$\begin{aligned} & p(a, b, c, \sigma_f, \alpha, \xi, \kappa, u_Q | \tilde{Q}, \tilde{H}, H_{max}, Q_{max}) \\ & \propto p(\tilde{Q}, Q_{max} | a, b, c, \sigma_f, \alpha, \xi, \kappa, \tilde{H}, H_{max}, u_Q) \cdot p(a, b, c, \sigma_f, \alpha, \xi, \kappa), \end{aligned} \quad (4.1)$$

em que a, b, c são os parâmetros da equação da curva-chave, σ_f corresponde ao desvio-padrão do erro do modelo da curva-chave, \tilde{Q} e \tilde{H} aos valores de cota e vazão medidos simultaneamente durante as medições de descarga, e H_{max} e Q_{max} às séries de cotas máximas anuais e vazões máximas anuais, respectivamente, sendo que a última depende de H_{max} , a, b, c e σ_f .

Como pode-se assumir que a probabilidade de ocorrência da vazão máxima anual (Q_{max}) é independente da probabilidade de se observar \tilde{Q} , a função verossimilhança pode ser separada em dois termos:

$$\begin{aligned} & p(\tilde{Q}, Q_{max} | a, b, c, \sigma_f, \alpha, \xi, \kappa, \tilde{H}, u_Q, H_{max}) \\ & \propto p(Q_{max} | a, b, c, \sigma_f, \alpha, \xi, \kappa, H_{max}) p(\tilde{Q} | u_Q, a, b, c, \sigma_f, \tilde{H}) \end{aligned} \quad (4.2)$$

Para o cálculo da probabilidade de ocorrência de \tilde{Q} (segundo termo), parte-se da premissa de que as medições de descarga seguem uma distribuição normal com média e desvio padrão conhecidos. A média é calculada pelo valor da vazão estimada $f(\tilde{H}_i | \theta)$ com base no vetor de parâmetros $\theta = a, b, c, \sigma_f$, enquanto o desvio padrão representa o erro total associado ao erro de medição (u_{Qi}) e ao erro de ajuste do modelo da curva-chave (σ_f). Dessa

maneira, a equação (4.3) descreve a probabilidade de ocorrência das medições de vazão dado que se conhece $a, b, c, \sigma_f, \tilde{H}$ e u_Q como sendo o produtório da probabilidade de ocorrências de cada medição de vazão i , partindo-se da premissa que os eventos são independentes.

$$p(\tilde{Q}|a, b, c, \sigma_f, \tilde{H}, u_Q) = \prod_{i=1}^N p_N\left(\tilde{Q}_i | f(\tilde{H}_i|\theta), \sqrt{\sigma_{f_i}^2 + u_{Q_i}^2}\right), \quad (4.3)$$

em que, $p_N(z | m, s)$ é a função normal de densidade de probabilidade com média m e desvio-padrão s , avaliada no ponto z . Seguindo a mesma lógica, define-se a verossimilhança dos valores de vazões máximas anuais, em que nesse caso assume-se que as vazões estimadas Q_{max} seguem uma distribuição GEV com parâmetros α, ξ, κ , conforme descrito na Equação (4.4),

$$p(Q_{max}|a, b, c, \sigma_f, \alpha, \xi, \kappa, H_{max}) = \prod_{i=1}^N f_{Q_{max_i}}(q_{max_i}) \quad (4.4)$$

em que:

$$f_{Q_{max_i}}(q_{max_i}) = \frac{1}{\alpha} \left[1 - \kappa \left(\frac{q_{max_i} - \xi}{\alpha} \right) \right]^{\frac{1}{\kappa} - 1} \exp \left\{ - \left[1 - \kappa \left(\frac{q_{max_i} - \xi}{\alpha} \right) \right]^{\frac{1}{\kappa}} \right\} \quad (4.5)$$

sendo:

$$Q_{max_i} = f(H_{max_i}|a, b, c) + \epsilon_i^f \quad (4.6)$$

$$\epsilon_i^f \sim N(0, \sigma_f^2) \quad (4.7)$$

O desvio-padrão do erro do modelo da curva-chave σ_f foi considerado nesse estudo como heteroscedástico, ou seja, sua variância depende do nível da vazão. Considera-se que o desvio-padrão do erro do modelo da curva-chave varia linearmente com a vazão, de forma que:

$$\sigma_f = \gamma_1 + \gamma_2 Q \quad (4.8)$$

Assim, γ_1 e γ_2 também são parâmetros estimados pelo modelo para o cálculo de σ_f para cada valor de vazão.

4.2.2 Distribuições *a priori*

Na inferência Bayesiana, a função verossimilhança é utilizada para atualizar o conhecimento a respeito dos parâmetros que se pretende estimar. Assim, parte-se

inicialmente desse conhecimento considerado anterior a obtenção dos dados coletados. Esse conhecimento deve estar associado a outras fontes de informações ou levantamentos de campo. No caso do modelo proposto, as informações acerca dos parâmetros $a, b, c, \sigma_f, \gamma_1, \gamma_2, \alpha, \xi$ e κ que se quer estimar devem ser descritas na forma de distribuições *a priori*.

Para os parâmetros da curva-chave, foi considerado que o conhecimento *a priori* em relação às características físicas da seção pode ser descrito por uma distribuição normal centrada no valor médio dependente de algum conhecimento adicional de campo ou de um especialista, com intervalo de confiança de 95% ($\pm\Delta$). Esse valor de Δ também pode ser descrito como duas vezes o valor do desvio padrão da distribuição normal.

A propagação dessas incertezas para o cálculo da média e desvio padrão dos parâmetros a, b e c , se dá, partindo-se da premissa que as variáveis são independentes, pela equação geral (4.9), (Singh *et al.*, 2007).

$$\sigma_y^2 = \sum_{i=1}^n \left[\left. \frac{\partial y}{\partial x_i} \right|_{x_i} \sigma_{x_i} \right]^2 = \sum_{i=1}^n \left[\frac{\partial y}{\partial x_i} \frac{\Delta_{x_i}}{2} \right]^2 \quad (4.9)$$

O parâmetro a , por exemplo, depende do valor da largura (B), declividade (S_0) e coeficiente de rugosidade (K), logo o seu valor dependerá do valor médio e da propagação das incertezas de cada um desses parâmetros. O Quadro 4-1 exemplifica esse cálculo para uma seção hidráulica fictícia.

O desvio-padrão do erro do modelo da curva-chave (σ_f), a ser estimado, está associado aos valores de γ_1, γ_2 e ao valor da vazão. Como essa incerteza é desconhecida, assume-se para esses parâmetros (γ_1, γ_2) *prioris* menos informativas.

No caso de γ_1 , sugere-se nesse trabalho uma distribuição uniforme de 0 a um limite superior ajustável dependendo da variância do valor das vazões estimadas na seção em estudo. O valor de γ_1 não pode ser maior do que o menor valor de vazão, para não gerar valores negativos de vazão.

Por outro lado, γ_2 deve representar um percentual da vazão. Dessa maneira, sugere-se uma distribuição uniforme de limites 0 e 1.

Ressalta-se, que esse modelo de erros em função de γ_1 e γ_2 foi baseada em estudos anteriores (Le Coz *et. al.*, 2014) e que é possível também considerar o parâmetro do erro como independente da vazão e estimá-lo como sendo um único parâmetro. Estudos mais

aprofundados a esse respeito devem ser desenvolvidos, e podem ser incorporados no método proposto nesse trabalho.

As distribuições *a priori* dos parâmetros da GEV foram definidas com base na literatura em estudos semelhantes de inferência Bayesiana para análise de frequência de cheias (Lang *et al.*,2010). Para facilidade do cálculo das posteriores, adotou-se no lugar do parâmetro de escala α , o parâmetro φ que corresponde à $\ln(\alpha)$.

Para os parâmetros de escala e posição, φ e ξ , foram assumidas *prioris* não informativas uniformes com limites de 1 e 10 para φ e -10.000 e 10.000 para ξ .

Por fim, a *priori* de κ foi definida com base na *priori* geofísica de Martins e Stedinger (2001) apresentada na Eq. (3.15), representada por uma distribuição beta entre os valores de -0,5 e 0,5 com média igual a -0,10 e variância igual a 0,122².

A Tabela 4-1 apresenta um resumo das informações *a priori* consideradas para cada um dos parâmetros no modelo proposto.

Tabela 4-1 – Resumo das informações *a priori* dos parâmetros

Parâmetro	Distribuição <i>a priori</i>
a	Normal com média e desvio padrão definidos com base nas características hidráulicas
b	Normal com média e desvio padrão definidos nas informações sobre a seção do rio
c	Normal com média e desvio padrão definidos com base no tipo de controle hidráulico
γ_1	Uniforme de 0 a max (valor que deve ser definido com base na magnitude das vazões)
γ_2	Uniforme de 0 a 1
φ	Uniforme de 1 a 10
ξ	Uniforme de -10.000 a 10.000
κ	Distribuição Beta(6,9) entre [-0,5;0,5]

Seção Hidráulica Fictícia

Declividade

	Medição	Confiança 95% (Δ)	σ
Nível d'água em 2 (m) - H2	102	1	0.50
Nível d'água em 1 (m) - H1	98	1	0.50
Distância entre 1 e 2 (m) - D	1000	100	50
Declividade (S)	0.004	0.00147	0.00073

Formula do erro:

$$\Delta S = 2S \sqrt{\left(\frac{\Delta D}{2D}\right)^2 + \left(\frac{\Delta H_1}{H_2 - H_1}\right)^2 + \left(\frac{\Delta H_2}{H_2 - H_1}\right)^2}$$

Rugosidade

Calibração de Modelo Hidráulico com base em medição de cota e vazão simultânea em campo

	média	ΔK
K	48	3
n	0.021	0.001

Largura

Medição em campo	média	ΔB
B(m)	150	10

Ponto de vazão nula

Estimado com base na seção hidráulica, refere-se a menor cota da seção hidráulica

	média	Δb
b(m)	230	0.5

Coefficiente c

Depende do controle hidráulico, no caso o controle é de canal

	média	Δc
c	1.67	0.05

Obs: Δ representa o intervalo de 95% de confiança dos dados

Cálculo do parâmetro a:

Para controle de Canal

$$a = KB\sqrt{S} \quad \Delta a = 2a \sqrt{\left(\frac{\Delta K}{2K}\right)^2 + \left(\frac{\Delta B}{2B}\right)^2 + \frac{1}{4}\left(\frac{\Delta S}{2S}\right)^2}$$

a	Δa
455.368	93.434

Prioris

Distribuição Normal	média	Confiança 95% (Δ)	σ
a	455.368	93.434	47.671
b	230	0.5	0.26
c	1.67	0.05	0.026

Quadro 4-1- Exemplo de Definição das informações *a priori* com base nas características hidráulicas da seção

4.2.3 Distribuição *a posteriori* dos Parâmetros e dos Quantis de Cheia

O algoritmo Metropolis Adaptado (AM) é utilizado para a aplicação de simulações Monte Carlo via Cadeia de Markov (MCMC). O algoritmo estima de forma numérica os parâmetros da curva-chave e da GEV. Assim, é gerada para cada parâmetro uma cadeia de Markov. Dessa cadeia, devem ser eliminados os valores iniciais referentes ao período de aquecimento e deve ser verificada sua convergência conforme métodos descritos no item 3.6.4. Em caso de convergência, a cadeia sem as iterações do aquecimento representa a distribuição *a posteriori* de cada um dos parâmetros.

Para obtenção da distribuição dos quantis de cheia, cada vetor de parâmetros da cadeia é utilizado para gerar os quantis com diferentes probabilidades de excedência, a partir da aplicação direta da equação (3.8), o que resulta em uma cadeia de igual número de iterações com valores dos quantis.

Com base nessa cadeia, calcula-se a curva de frequência de cheias média e seu intervalo de credibilidade de 95% que incorpora as incertezas da curva-chave. Ressalta-se que o resultado permite o cálculo de intervalos de diferentes níveis de credibilidade, mas para análise proposta nesse trabalho, optou-se pela utilização do intervalo de 95%.

As rotinas utilizadas no programa R, estão apresentadas no APÊNCIDE I.

4.3 Modelo Bayesiano para GEV – BG

O modelo Bayesiano para estimativa apenas dos parâmetros da GEV foi proposto para estimar os quantis sem a influência da incerteza da curva-chave. Dessa maneira, é utilizada apenas uma série de vazões máximas anuais e os parâmetros estimados passam a ser apenas α (ou no caso φ), ξ e κ . A função *a posteriori* passa a ser descrita como:

$$p(\alpha, \xi, \kappa | Q_{max}) \propto p(Q_{max} | \alpha, \xi, \kappa) \cdot p(\alpha, \xi, \kappa), \quad (4.10)$$

em que Q_{max} é estimada com base na série de cotas máximas anuais (H_{max}) e em um vetor único de parâmetros a , b e c . Nesse caso, adotam-se os parâmetros calculados no modelo Bayesiano Completo de maior valor da *posteriori*. Isso está descrito em detalhe no item 4.4.1.

4.3.1 Função de Verossimilhança

A função de verossimilhança no modelo BG é, portanto, similar à função do BC, sem a parcela referente à estimativa dos parâmetros da curva-chave, como apresentada na Equação (4.11) e (4.12).

$$p(Q_{max}|\alpha, \xi, \kappa) = \prod_{i=1}^N f_{Q_{max_i}}(q_{max_i}) \quad (4.11)$$

em que:

$$f_{Q_{max_i}}(q_{max_i}) = \frac{1}{\alpha} \left[1 - \kappa \left(\frac{q_{max_i} - \xi}{\alpha} \right) \right]^{\frac{1}{\kappa} - 1} \exp \left\{ - \left[1 - \kappa \left(\frac{q_{max_i} - \xi}{\alpha} \right) \right]^{\frac{1}{\kappa}} \right\} \quad (4.12)$$

4.3.2 Distribuições *a priori*

No modelo BG devem ser fornecidas as informações *a priori* apenas dos parâmetros da GEV. Foram utilizadas nesse modelo as mesmas distribuições adotadas no modelo bayesiano completo. Mais uma vez, substitui-se o parâmetro α pelo parâmetro φ . A Tabela 4-2 apresenta as distribuições *a priori* de cada parâmetro.

Tabela 4-2 Distribuições *a priori* dos Parâmetros da GEV no modelo BG

Parâmetro	Distribuição <i>a priori</i>
φ	Uniforme de 1 a 10
ξ	Uniforme de -10.000 a 10.000
κ	Distribuição Beta(6,9) entre [-0,5;0,5]

4.3.3 Distribuição *a Posteriori* dos Parâmetros e dos Quantis de Cheia

Da mesma maneira que foi realizado no modelo Bayesiano Completo, adotou-se o algoritmo Metropolis Adaptado (AM) para estimativa dos parâmetros. São gerados nesses caso, apenas 3 cadeias cuja convergência e período de aquecimento também devem ser verificados.

Para obtenção da distribuição dos quantis de cheia, cada vetor de parâmetros da cadeia é utilizado para gerar os quantis com diferentes probabilidades de excedência, a partir

da aplicação direta da equação (3.8), o que resulta em uma cadeia de igual número de iterações com valores dos quantis.

Com base nessa cadeia, calcula-se a curva de frequência de cheias média e seu intervalo de credibilidade de 95% sem a incorporação das incertezas da curva-chave.

As rotinas utilizadas no programa R para esse modelo também estão apresentadas no APÊNDICE I.

4.4 Impacto das Incertezas da curva-chave na análise de frequência de cheias

A avaliação do impacto das incertezas da curva-chave foi realizada nesse trabalho de forma a responder as seguintes questões:

1) Qual a importância da incerteza da curva-chave em relação à incerteza global nos quantis de cheia, que envolve também a incerteza na estimativa dos parâmetros da GEV dependente do tamanho da série de cotas máximas anuais?

2) Como é a variação dessa importância à medida que se reduz o conhecimento *a priori* dos parâmetros da equação da curva-chave, a disponibilidade de medições de descarga e o tamanho da série de cotas máximas anuais?

Para responder a essas questões, propôs-se o cálculo de uma métrica que tenta representar essa importância relativa em cenários que permitam avaliar os impactos para diferentes situações de disponibilidade de dados e informação *a priori*. Esses cenários estão apresentados em detalhe a seguir (Passos 7 e 8 da Metodologia proposta).

4.4.1 Métrica para avaliação – Importância Relativa

Conforme já mencionado, as incertezas na análise de frequência de cheia estão relacionadas ao desconhecimento da real relação entre cota e vazão (incerteza na curva-chave - IC) e ao reduzido número de cotas máximas anuais disponíveis para o ajuste de uma distribuição probabilística que represente os eventos e suas probabilidades de ocorrência adequadamente (incerteza amostral - IA).

Quando todas essas fontes de incerteza são consideradas, como ocorre na metodologia Bayesiana completa apresentada no item 4.2, o intervalo de credibilidade estimado para os quantis representa a incerteza global ($IG = IA + IC$), composta da incerteza amostral e da incerteza proveniente do ajuste da curva-chave.

Como o objetivo é verificar dentro dessa incerteza global, qual a parcela de contribuição da incerteza da curva-chave (Questionamento 1), é preciso estimar pelo menos uma das incertezas individualmente, IC ou IA. A estratégia adotada neste trabalho foi a de estimar a incerteza amostral (IA), para depois estimar IC, de forma indireta, empregando a estimativa de IG obtida pela abordagem Bayesiana completa.

Para que a incerteza amostral fosse estimada, é necessário remover a incerteza proveniente do ajuste da curva-chave, ou em outras palavras, é preciso definir os parâmetros da equação da curva-chave para só depois partir para o ajuste dos parâmetros da GEV. Optou-se então por adotar os parâmetros da curva-chave que maximizaram a distribuição *a posteriori* obtida com base na abordagem Bayesiana completa. Com base nessa estimativa dos parâmetros, denominada aqui de maxpost, obteve-se a série de vazões máximas anuais, que foi então empregada posteriormente na estimativa dos parâmetros da GEV, empregando o modelo Bayesiano apenas para estimativa dos parâmetros da GEV (BG), apresentado no item 4.3. A Figura 4-3 exemplifica esse procedimento.

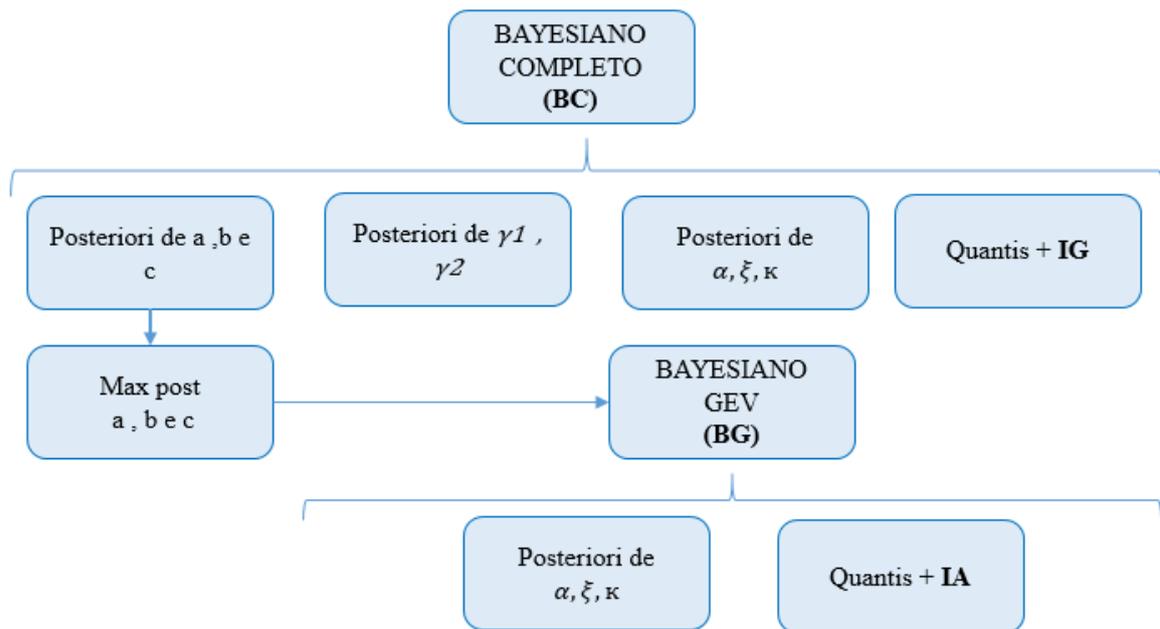


Figura 4-3 – Procedimento de cálculo dos Intervalos de Credibilidade da incerteza global (IG) e (IA) pelo método Bayesiano completo (BC) e bayesiano apenas para GEV (BG)

As estimativas baseadas em BC e BG permitem determinar os intervalos de credibilidade para todos os quantis de interesse. A questão passa a ser como representar IG e IA com base nos intervalos de credibilidade obtidos. Dessa maneira, propôs-se nesse trabalho, representar essa incerteza como sendo a razão entre a largura do intervalo de

credibilidade de 95% e a média da distribuição *a posteriori* dos quantis, conforme apresentado nas equações (4.13) e (4.14).

$$IG_{TR} = \frac{Q_{0,975,TR}^{BC} - Q_{0,025,TR}^{BC}}{\bar{Q}_{TR}^{BC}} \quad (4.13)$$

$$IA_{TR} = \frac{Q_{0,975,TR}^{BG} - Q_{0,025,TR}^{BG}}{\bar{Q}_{TR}^{BG}} \quad (4.14)$$

Em que $Q_{0,975,TR}^{BC}$, $Q_{0,025,TR}^{BC}$, $Q_{0,975,TR}^{BG}$ e $Q_{0,025,TR}^{BG}$ correspondem ao intervalo de maior densidade de probabilidade (HPD) da distribuição *a posteriori* dos quantis calculados para o tempo de retorno TR pelos métodos completo (BC) e apenas para GEV (BG).

Com ambos os intervalos estimados, a diferença entre eles representa, portanto, a parcela da incerteza da curva chave (IC). Dessa maneira, para avaliação do grau de importância dessa incerteza em relação a incerteza global, propôs-se uma métrica denominada como Importância Relativa (IR). Essa métrica nada mais é que o percentual da incerteza da curva-chave, calculada indiretamente pela incerteza global (IG) e incerteza amostral (IA), em relação à Incerteza Global (IG), conforme apresentado na Equação (4.15):

$$IR (\%) = \frac{IG - IA}{IG} \times 100 \quad (4.15)$$

4.4.2 Fatores Avaliados

A questão 2 levantada pode ser estudada a partir da comparação dos valores da importância relativa (IR) para diferentes cenários, baseados no grau de informação fornecida acerca dos parâmetros da curva-chave por meio das distribuições *a priori* adotadas, na quantidade de pares de valores de cota e vazão presentes no resumo de medições de descarga e no tamanho da série histórica de cotas máximas anuais. Os dois primeiros fatores afetam as incertezas provenientes do ajuste da curva-chave, enquanto o último está diretamente relacionado com a incerteza amostral no ajuste da distribuição GEV.

Espera-se que o valor da IR diminua à medida que se reduz a disponibilidade de dados de cotas máximas anuais, e aumente com a redução das informações *a priori* e disponibilidade de medições conjuntas de nível e vazão.

Com base nessa lógica, são propostas as variações de cenários apresentadas na Tabela 4-3.

Foram adotadas variações no nível de informação *a priori* (P para P-), na disponibilidade de dados de medições de descarga (D para D-) e na disponibilidade de dados de cotas máximas anuais (H para H-). Essas variações adotadas entre P e P-, D e D-, H e H- serão apresentadas separadamente nos estudos de caso.

Tabela 4-3 – Cenários com variação dos fatores avaliados e legenda adotada

Cenário	Mais dados de Medições de Descarga	Menos Dados de Medições de Descarga
	Mais dados de cotas máximas anuais	
<i>Priori +</i> Informativa	PDH	PD-H
<i>Priori -</i> Informativa	P-DH	P-D-H
	Menos dados de cotas máximas anuais	
<i>Priori +</i> Informativa	PDH-	PD-H-
<i>Priori -</i> Informativa	P-DH-	P-D-H-

4.5 Estudos de Caso

A métrica da importância relativa (IR) para os diferentes cenários definidos na Tabela 4-3 foram calculados em dois estudos de caso. O primeiro consiste de uma situação hipotética, com dados gerados de forma sintética, onde se conhecem todos os parâmetros envolvidos no problema. O segundo caso trata de uma situação real, baseada nos dados existentes para a estação fluviométrica de Cáceres, localizada no Rio Paraguai.

Os dados de cada um dos casos, bem como as premissas adotadas para aplicação do modelo, estão descritos nos itens a seguir.

4.5.1 Caso com dados sintéticos

Para análise dos resultados gerados pelo modelo e investigação do impacto das incertezas da curva-chave no estudo de frequência de cheias, o modelo foi aplicado a um caso com dados sintéticos em que todos os parâmetros eram conhecidos.

Primeiro, foram definidos todos os parâmetros a partir dos quais os valores sintéticos seriam gerados. A Tabela 4-4 apresenta o valor desses parâmetros.

Ressalta-se que nesse estudo de caso sintético, os valores dos parâmetros não guardam relação com nenhuma seção hidráulica conhecida, e os erros de ajuste do modelo também são aplicados sinteticamente.

Tabela 4-4 – Valor dos Parâmetros adotados para o caso sintético

Parâmetros Curva Chave		Parâmetros da GEV	
a	90	φ	5,70
b	100	ξ	1500
c	1,67	κ	-0,04
γ_1	25		
γ_2	0,10		
u_q	0,05 q		

A obtenção dos dados sintéticos consistiu primeiro na geração aleatória de valores de cota a partir de uma distribuição uniforme. Em seguida, a partir desses dados foram calculadas as vazões por meio da aplicação dos valores pré-definidos dos parâmetros da curva-chave indicados na Tabela 4-4. A essas vazões foram aplicados os erros aleatórios de medição e de ajuste do modelo, conforme Eq. (4.6). Os dados gerados de cota e vazão passam a ser o resumo de descarga (D) do caso sintético.

Para geração dos dados da série de cotas máximas anuais, primeiro foram gerados a partir de uma GEV, também de parâmetros pré-definidos (Tabela 4-4), valores de vazões máximas. Mais uma vez a esses valores foi aplicada a equação da curva-chave para obtenção das cotas máximas anuais H.

O procedimento descrito acima está detalhado no algoritmo apresentado no Quadro 4-2.

Os dados de resumo de descarga D e série de cotas máximas anuais H foram aplicados no modelo bayesiano completo (BC) e no modelo bayesiano da GEV (BG) para o cálculo da importância relativa (IR). Os mesmos dados D e H foram reduzidos em 50% ($n=100$ para $n=50$ e $m=50$ para $m = 25$) e aplicados para as duas informações *a priori* (P e P-) definidas no passo 3 do Quadro 4-2 para geração dos oito cenários de análise.

Além disso, esses mesmos resultados foram comparados aos valores pré-definidos para se verificar a validade das estimativas obtidas pelo modelo proposto. Os resultados para esse cenário podem ser observados no item 5.

Algoritmo para geração dos Dados Sintéticos

Primeiro gere **n** valores de cotas de uma distribuição UNIF(inf,sup)

Passo 1: Simule valores de medição de cota e vazão

1. for $j = 1, \dots, n$ **n=100**

(a) Gera os valores de vazão com base em h e nos parâmetros da curva-chave definidos

$$q_i = 90(h_j - 100)^{1,67}$$

(b) Calcula os erros de medição associados a cada valor de vazão

$$u_{qi} = \text{erro de medição (\%)} \times q_i$$

$$\sigma_{fi} = \gamma_1 + \gamma_2 q_i$$

$$\varepsilon_i \sim N(0, \sqrt{u_{qi}^2 + \sigma_{fi}^2})$$

(c) Calcula a vazão com erro

$$q_i^* = q_i + \varepsilon_i$$

2. Define $D = (h_1, q_1^*), \dots, (h_n, q_n^*)$

Passo 2: Simula as vazões máximas anuais

3. for $j = 1, \dots, m$ **m = 50**

(a) Gera X_j^* de uma GEV de parâmetros:

$$\alpha = 100 (\varphi = 5,70), \xi = 1000, \kappa = -0,04$$

(b) Calcula a série de cotas máximas anuais

$$Y_j^* = \frac{X_j^*}{90}^{1/1,67} + 100$$

4. Define $H = Y_1^*, \dots, Y_m^*$

Passo 3: Define as informações *a priori* dos parâmetros da curva chave

Parâmetro	Média	Desvio Padrão	
		P	P-
a	90	22,5	45
b	100	1	2
c	1,67	0,025	0,05

Quadro 4-2 - Algoritmo para geração de dados Sintéticos

As Figuras 4.4 e 4.5 apresentam os dados gerados sinteticamente.

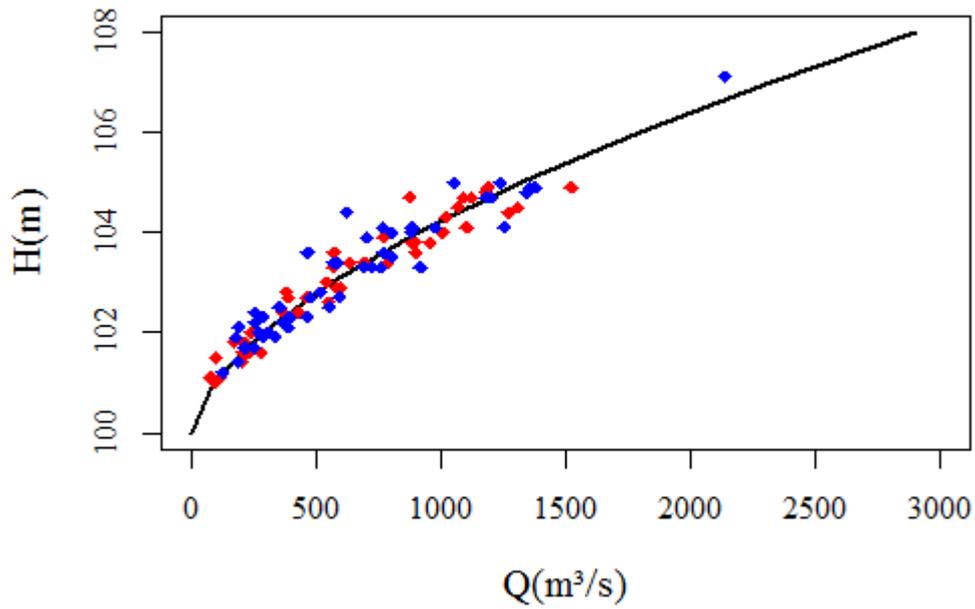


Figura 4-4 Curva-Chave de parâmetros $a = 90$, $b = 100$ e $c = 1,67$ com dados gerados sinteticamente com erro de medição e de ajuste do modelo e eliminados para o cenário D- (pontos vermelhos)

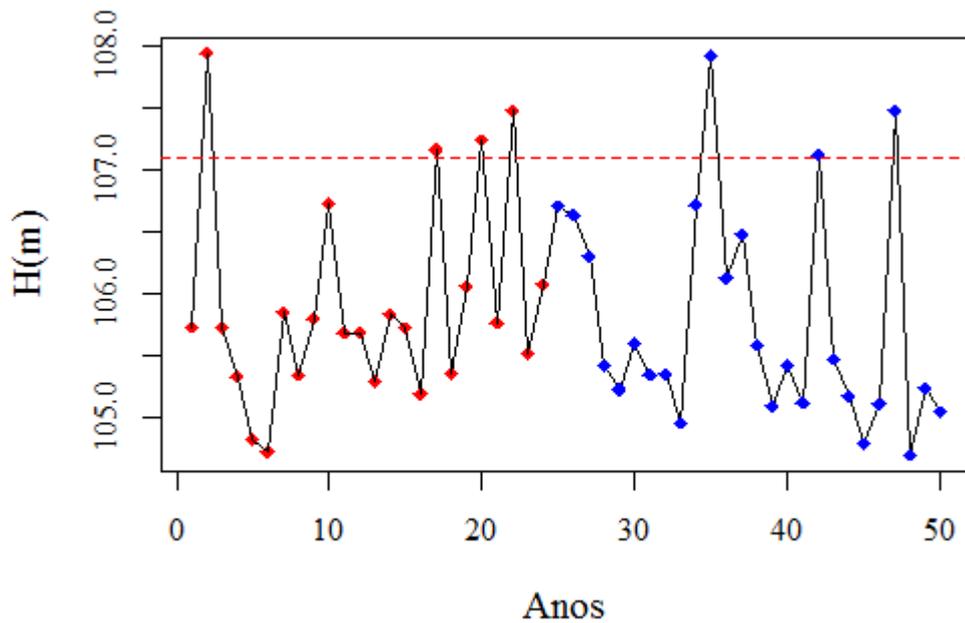


Figura 4-5 Dados de cotas máximas anuais gerados sinteticamente com os dados removidos (pontos vermelhos) para o cenário H-, com indicação da maior cota máxima registrada no resumo de descarga (linha tracejada vermelha)

4.5.2 Caso Real - Cáceres

Para aplicação do caso real, foi escolhida a estação fluviométrica de Cáceres (código: 66070004) localizada no rio Paraguai próxima a cidade de Cáceres. As medições de cota e vazão do resumo de descarga e cotas máximas anuais do posto foram analisadas e tratadas, e os dados finais após esse tratamento podem ser observados nos gráficos da Figura 4-6 e Figura 4-7. O tratamento consistiu na comparação das medições para uma mesma cota com o intuito de verificar possíveis inconsistências nas observações. Os dados originais da estação estão disponíveis no HidroWeb, banco de dados da Agência Nacional de Águas.

Para criação dos cenários de análise no estudo de caso real, foi proposta uma remoção dos dados de medições de cota e vazão e cotas máximas anuais. Os dados removidos estão indicados pelos pontos em vermelho nos gráficos das figuras 4.6 e 4.7.

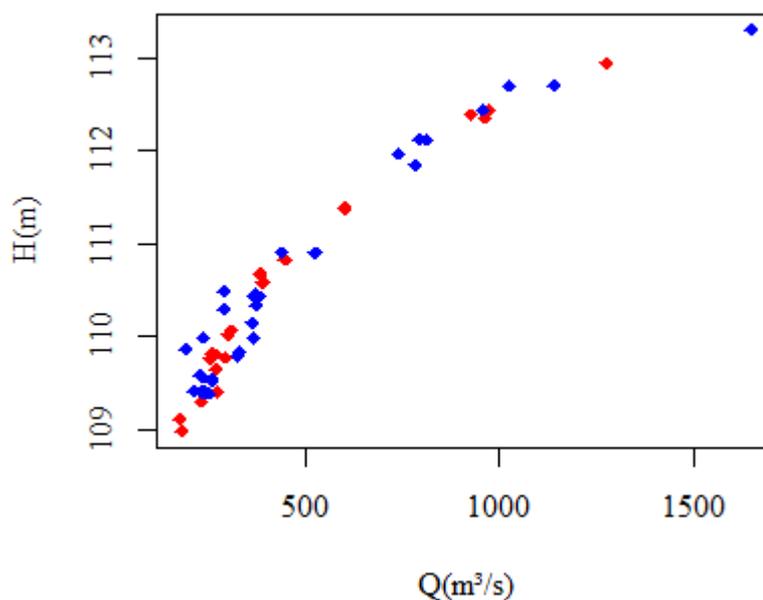


Figura 4-6 – Dados de resumo de descarga (D) da Estação Fluviométrica de Cáceres com pontos vermelhos removidos para o cenário de disponibilidade reduzida de resumo de descarga (D-)

Conforme pode ser observado nos gráficos das Figuras 4.6 e 4.7, a estação possui mais de 40 anos de medição de cotas diárias e 48 medições simultâneas de cota e vazão. Observa-se também o grau de extrapolação necessário para estimar as vazões máximas anuais, tendo em vista que a maior vazão medida representa uma cota de 113,3 m e, no período, 21 observações igualaram ou superaram esse valor.

As características hidráulicas da seção para definição das informações *a priori* aplicadas no modelo foram obtidas com base nos levantamentos de campo desenvolvidos

pela Universidade Federal do Paraná em 2014 (UFPR/ITTI, 2015). A Universidade levantou dados na região com o intuito de estudar a viabilidade da hidrovía do Paraguai. Dessa maneira, foram realizados levantamentos de batimetria, perfis instantâneos de nível d'água, medições de vazão, incluindo uma medição exatamente na seção da régua. Esses levantamentos subsidiaram a elaboração de um modelo hidráulico unidimensional da calha do rio devidamente calibrado e validado. O modelo desenvolvido pela universidade serviu como complemento para obtenção de mais informações acerca das características hidráulicas *a priori* na seção do rio.

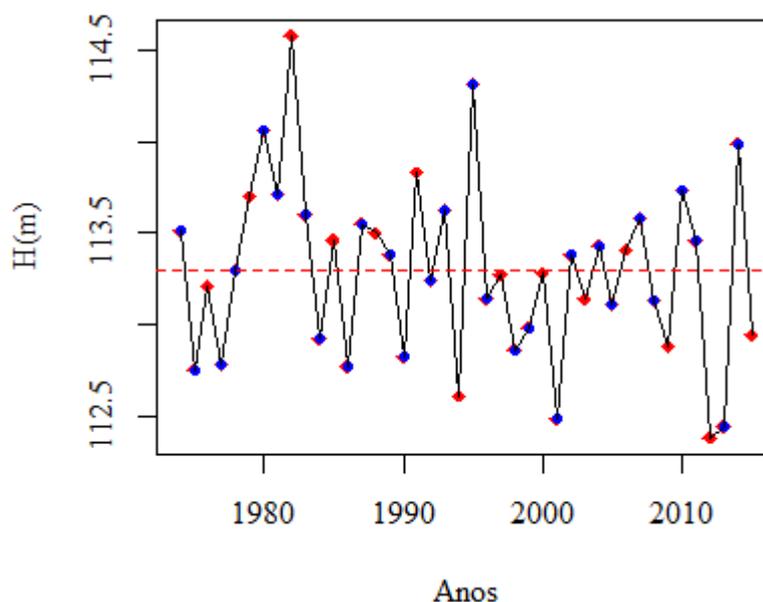


Figura 4-7 – Dados cotas máximas anuais da Estação Fluviométrica de Cáceres de 1974 a 2015 (H), máxima cota observada no resumo de descarga (linha tracejada) e pontos removidos para o cenário de disponibilidade reduzida H- (pontos vermelhos)

As medições de vazão foram realizadas pela Universidade com uso de ADCP (*Acoustic Doppler Current Profiler*). O ADCP mede a vazão indiretamente a partir dos perfis de velocidade e área da seção. Dessa maneira, o equipamento fornece além do valor da vazão na seção, os dados da batimetria de onde é possível extrair informações como largura e profundidade da seção. Esses dados foram disponibilizados e utilizados como entrada para aplicação do modelo de estimativa de curva-chave e curva de frequência de cheias Bayesiana proposto.

O modelo hidráulico serviu para estimativa do conhecimento *a priori* do coeficiente de Manning. O Manning adotado no modelo desenvolvido pela UFPR foi obtido após várias calibrações com base nas medições de vazão e nível simultâneas realizadas nas campanhas de campo.

O valor desse coeficiente associado às informações da largura da seção transversal obtida no banco de dados do Hidroweb (Figura 4-9) e estimada em campo pelo ADCP (Figura 4-10), e à declividade obtida com base no perfil de elevação de nível d'água (Figura 4-8), permitiram o cálculo do coeficiente a da curva-chave. O parâmetro b foi estimado também com base na seção transversal, enquanto que para o parâmetro c assumiu-se que o seção se comporta como um controle de canal.

Além disso, foram consideradas algumas simplificações. Considerou-se que:

- a curva possui apenas um segmento;
- ela não se alterou ao longo dos anos.

As informações de entrada *a priori* utilizadas para os parâmetros da curva-chave estão detalhadas no Quadro 4-3, enquanto as demais *a priori*s foram consideradas conforme Tabela 4-1.

O cenário real simulado foi realizado considerando os levantamentos de campo e a modelagem hidráulica para definição das informações *a priori*. Entretanto, esses dados nem sempre estão disponíveis e muitos dos parâmetros da curva-chave acabam sendo estimados sem o conhecimento das características hidráulicas e do escoamento.

Dessa maneira, buscou-se aplicar o modelo para o mesmo cenário real na estação de Cáceres, assumindo total desconhecimento da rugosidade do rio e aumento da incerteza associada a declividade do rio. A ficha com os novos parâmetros considerados está apresentada na Quadro 4-4.

O intervalo de confiança do coeficiente de Manning adotado foi definido com base na faixa de valores possíveis de rugosidade para um rio de calha natural. Em um rio, o coeficiente pode variar de 0,02 até 0,10. Dessa maneira o coeficiente de variação do parâmetro a passou de 0,2 (cenário anterior) para 0,5 conforme apresentado na Tabela 4-5.

O objetivo desse cenário é tentar avaliar, mesmo que de forma preliminar o impacto da redução das informações iniciais *a priori* na incerteza da curva de frequência de cheias. Esse novo cenário será denominado como Cenário P-, enquanto o cenário original será denominado P.

Tabela 4-5 – Valores do parâmetro a para cada cenário

Cenário	Média a	Desvio	
		Padrão a	CV
Cenário P	63.3	13.3	0.21
Cenário P-	39.5	21.1	0.53

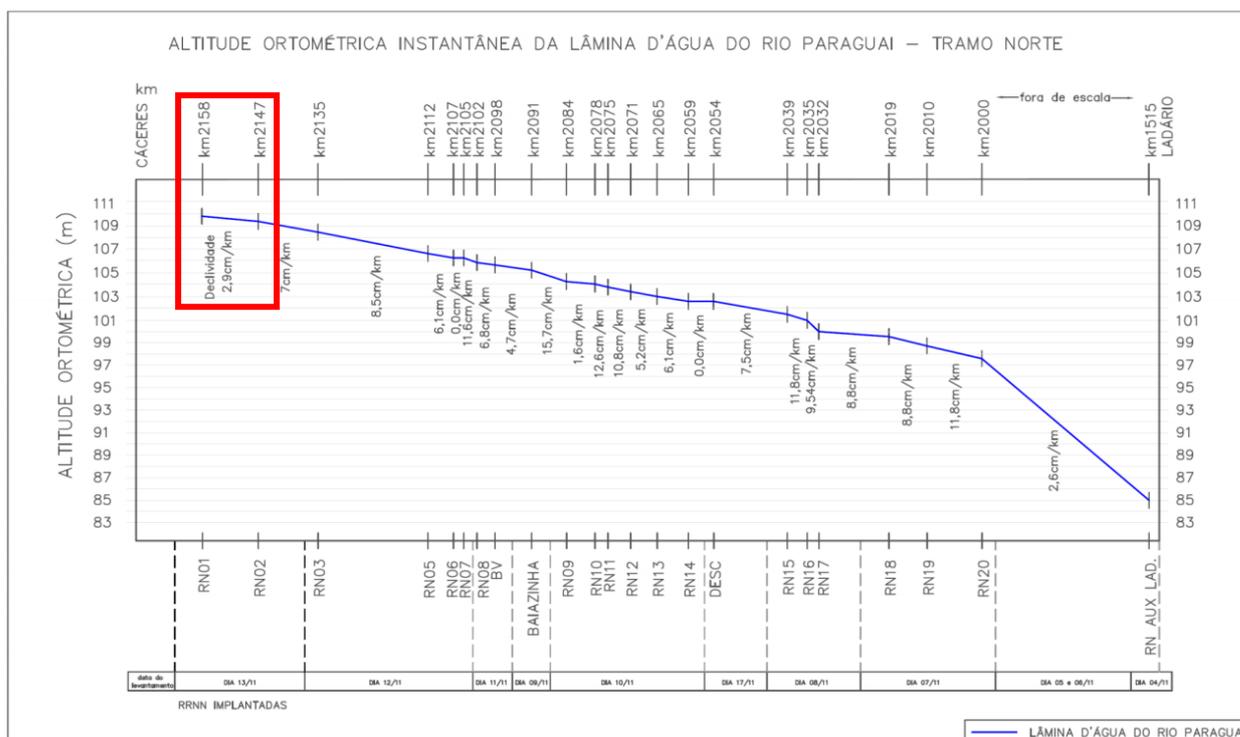


Figura 4-8 Inclinações instantâneas da lâmina d'água entre as referências de nível na hidrovia do rio Paraguai – Em vermelho o trecho onde se localiza a seção em estudo de inclinação de 2,9 cm/km Fonte: UFPR/ITTI,2015

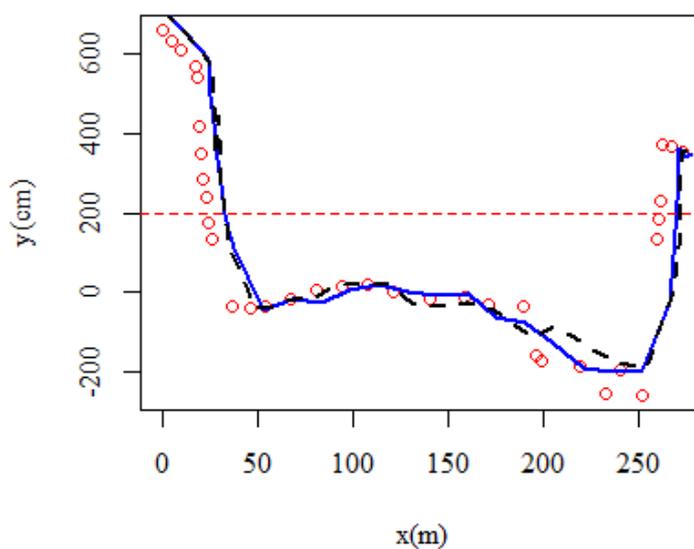


Figura 4-9 – Perfis batimétricos medidos em 2011 (pontos vermelhos),2014 (Linha cheia azul) e 2015 (linha tracejada preta) com indicação de nível d'água provável da data de medição do ADCP. Fonte: Hidroweb (2016)

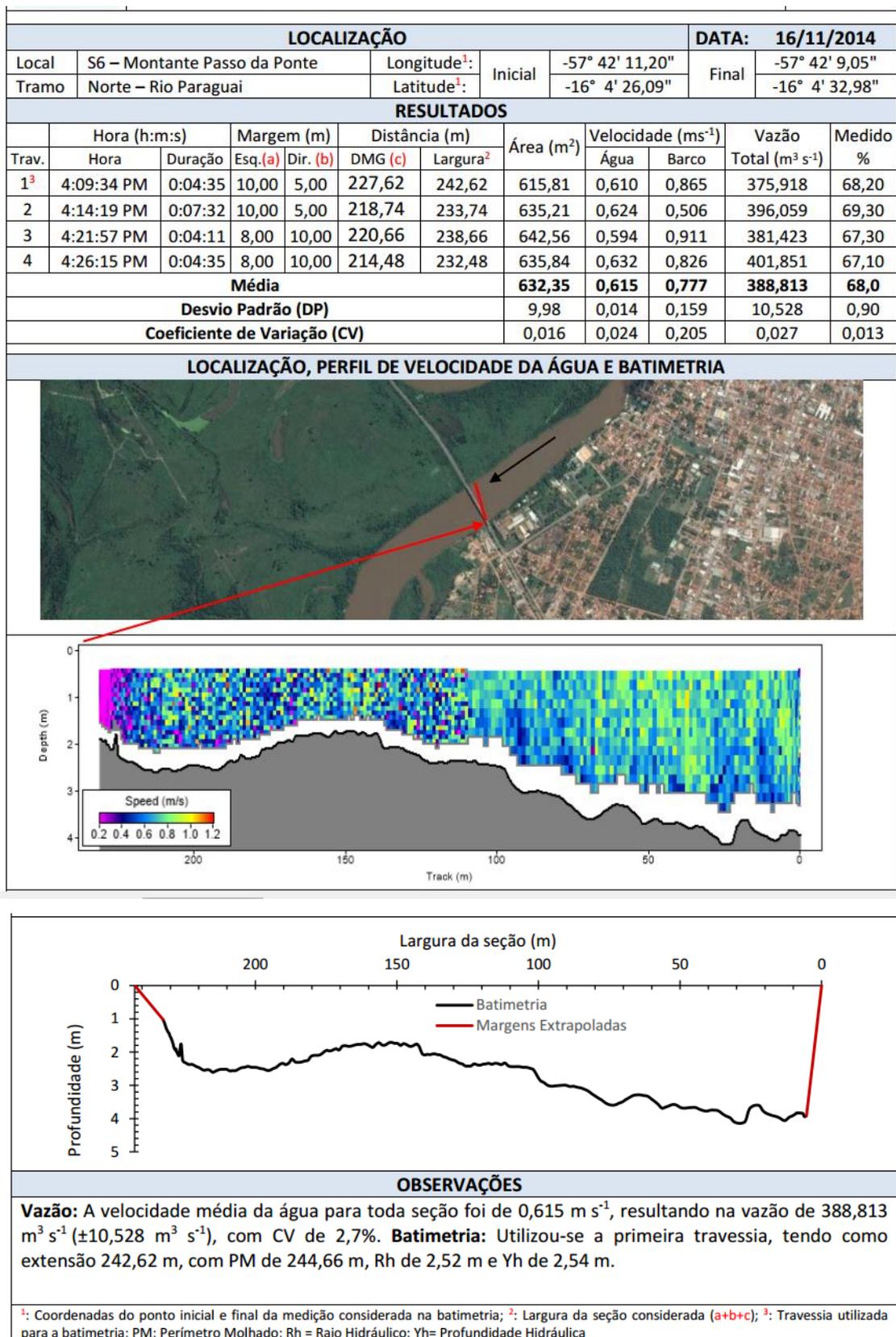


Figura 4-10 - Levantamento por ADCP – Fonte: UFPR/ITTI(2015)

Quadro 4-3 – Informações *a priori* dos parâmetros da curva-chave com base nas características hidráulicas da seção e suas incertezas

Estação de Cáceres			
<u>Declividade</u>			
	Medição	Confiança 95% (Δ)	σ
Nível d'água em 2 (m) - H2	110.2	0.2	0.4
Nível d'água em 1 (m) - H1	109.86	0.2	0.4
Distância entre 1 e 2 (m) - D	11500	10	20
Declividade (S)	0.0000296	0.00002	0.00005
Formula do erro:			
$\Delta S = 2S \sqrt{\left(\frac{\Delta D}{2D}\right)^2 + \left(\frac{\Delta H_1}{H_2 - H_1}\right)^2 + \left(\frac{\Delta H_2}{H_2 - H_1}\right)^2}$			
<u>Rugosidade</u>			
Calibração de Modelo Hidráulico com base em medição de cota e vazão simultânea em campo			
	média	ΔK	
K	48	3	
n	0.021	0.001	
<u>Largura</u>			
Medição em campo	média	ΔB	
B(m)	242	10	
<u>Ponto de vazão nula</u>			
Estimado com base na seção hidráulica, refere-se a menor cota da seção hidráulica			
	média	Δb	
b(m)	106.2	0.5	
<u>Coefficiente c</u>			
Depende do controle hidráulico, no caso o controle é de canal			
	média	Δc	
c	1.67	0.05	
Obs: Δ representa o intervalo de 95% de confiança dos dados			
<u>Cálculo do parâmetro a:</u>			
Para controle de Canal			
$a = KB\sqrt{S}$	$\Delta a = 2a \sqrt{\left(\frac{\Delta K}{2K}\right)^2 + \left(\frac{\Delta B}{2B}\right)^2 + \frac{1}{4}\left(\frac{\Delta S}{2S}\right)^2}$		
	a	Δa	
	63.2	26.7	
<u>Prioris</u>			
	média	Confiança 95% (Δ)	σ
Distribuição Normal			
a	63.2	26.7	13.3
b	106.2	0.5	0.25
c	1.67	0.05	0.025

Quadro 4-4 - Parâmetros adotados para o Cenário P- com realce nos valores alterados em relação ao Cenário P

Estação de Cáceres - Cenário menos informativo			
<u>Declividade</u>			
	Medição	Confiança 95% (Δ)	σ
Nível d'água em 2 (m) - H2	110.2	0.4	0.8
Nível d'água em 1 (m) - H1	109.86	0.4	0.8
Distância entre 1 e 2 (m) - D	11500	100	200
Declividade (S)	0.0000296	0.00005	0.00010
Formula do erro:			
$\Delta S = 2S \sqrt{\left(\frac{\Delta D}{2D}\right)^2 + \left(\frac{\Delta H_1}{H_2 - H_1}\right)^2 + \left(\frac{\Delta H_2}{H_2 - H_1}\right)^2}$			
<u>Rugosidade</u>			
Calibração de Modelo Hidráulico com base em medição de cota e vazão simultânea em campo			
	media	ΔK	
K	30	20	
n	0.033	0.05	
<u>Largura</u>			
Medição em campo	média	ΔB	
B(m)	242	10	
<u>Ponto de vazão nula</u>			
Estimado com base na seção hidráulica, refere-se a menor cota da seção hidráulica			
	média	Δb	
b(m)	106.2	0.5	
<u>Coefficiente c</u>			
Depende do controle hidráulico, no caso o controle é de canal			
	média	Δc	
c	1.67	0.05	
Obs: Δ representa o intervalo de 95% de confiança dos dados			
<u>Cálculo do parâmetro a:</u>			
Para controle de Canal			
$a = KB\sqrt{S}$	$\Delta a = 2a \sqrt{\left(\frac{\Delta K}{2K}\right)^2 + \left(\frac{\Delta B}{2B}\right)^2 + \frac{1}{4}\left(\frac{\Delta S}{2S}\right)^2}$		
	a	Δa	
	39.5	42.1	
<u>Prioris</u>			
Distribuição Normal	média	Confiança 95% (Δ)	σ
a	39.5	42.1	21.1
b	106.2	0.5	0.25
c	1.67	0.05	0.025

5 RESULTADOS E DISCUSSÃO

Nesse capítulo estão apresentados os principais resultados obtidos a partir da aplicação dos modelos Bayesiano Completo (BC) e Bayesiano para estimativa dos parâmetros da GEV (BG) no cenário completo (PDH) em ambos os estudos de caso. Em seguida, são apresentados os cálculos das incertezas globais e amostrais e o valor da importância relativa para os oito cenários simulados (PDH, P-DH, PD-H, P-D-H, PDH-, P-DH-, PD-H-P-D-H-) e a discussão sobre os resultados obtidos.

5.1 Estudo de Caso Sintético

5.1.1 Resultados Gerais para o Modelo Bayesiano Completo (BC)

Os dados gerados sinteticamente em conjunto com as distribuições *a priori* dos parâmetros foram inseridos no modelo BC para obtenção da distribuição *a posteriori* de todos os parâmetros (da curva-chave e da GEV) e quantis. Os resultados foram analisados cuidadosamente para verificação das estimativas obtidas. Primeiramente, foram analisadas as cadeias de Markov amostradas pelo algoritmo Metropolis Adaptado apresentadas na Figura 5-1.

As características da cadeia permitiram verificar, primeiro visualmente, uma possível convergência para todos os parâmetros do caso específico estudado. As cadeias foram plotadas com a indicação do valor médio e o valor real dos parâmetros para dois pontos de partida diferentes (cadeia 00 e cadeia 01).

Para confirmação da convergência do modelo de forma mais assertiva, o teste de convergência de Gelman e Rubin (1992) foi aplicado aos resultados das cadeias apresentadas (00 e 01). O gráfico do teste e resultados finais dos valores do fator R podem ser observados na Figura 5-2. Os resultados apresentados referem-se ao cenário completo PDH.

Conforme apresentado no Item 3.6.4, a convergência é considerada satisfatória quando o valor do fator R se aproxima da unidade. Observa-se, na Figura 5-2, que tanto o valor médio como o limite superior do intervalo de confiança de 95% convergem para valores abaixo de 1,2 a partir da iteração 50.000. Esse resultado nos leva a crer que, adotando um aquecimento de 50.000 iterações, as cadeias representem adequadamente a distribuição *a posteriori* dos parâmetros.

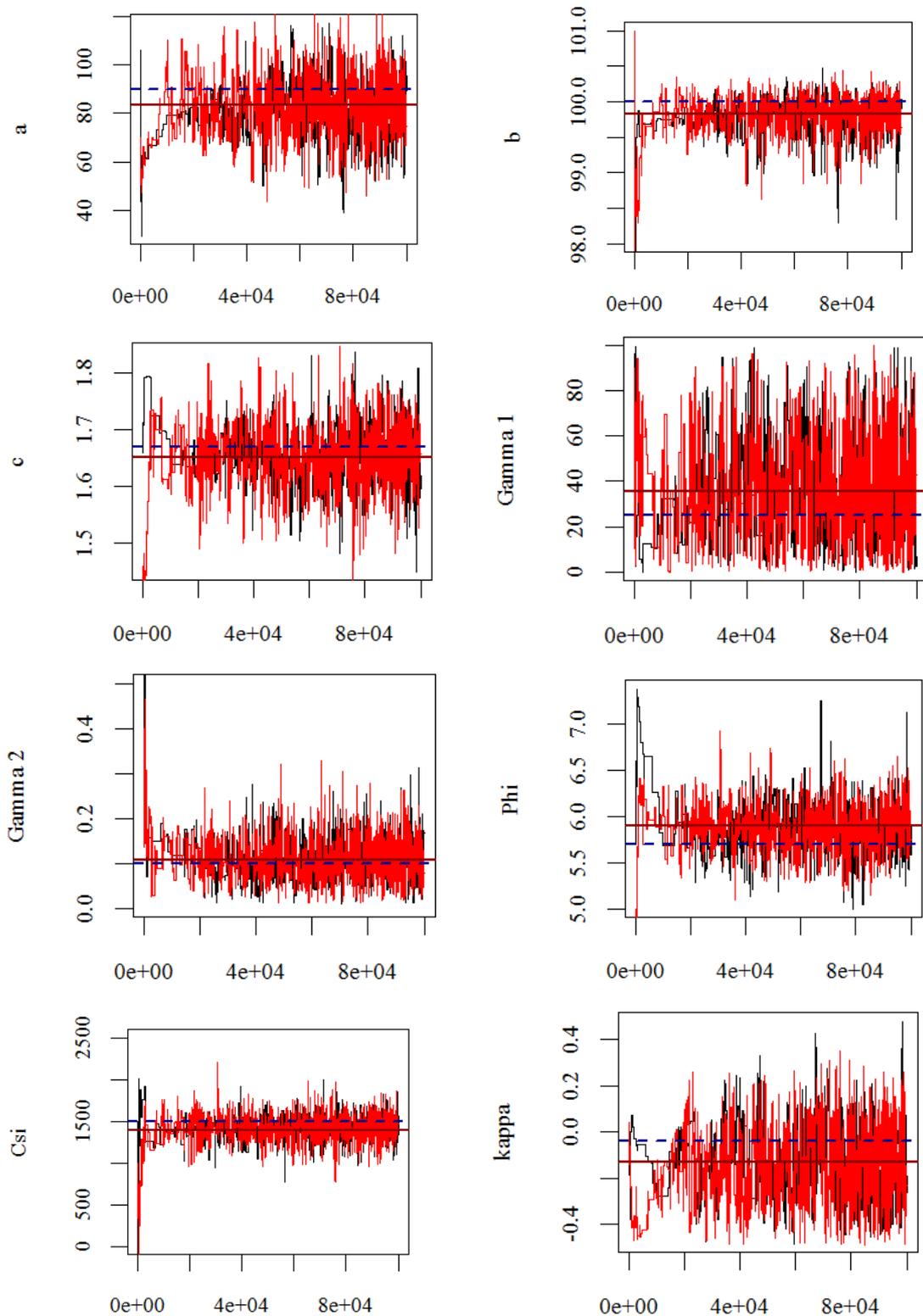


Figura 5-1 – Cadeias de Markov amostradas pelo algoritmo Adaptive Metropolis para 100.000 iterações pelo modelo Bayesiano Completo para todos os parâmetros do cenário sintético PDH das cadeias 00 (preto) e 01 (vermelho), valor médio da cadeia sem o aquecimento (linha cheia) e o valor real (linha tracejada) – Caso Sintético

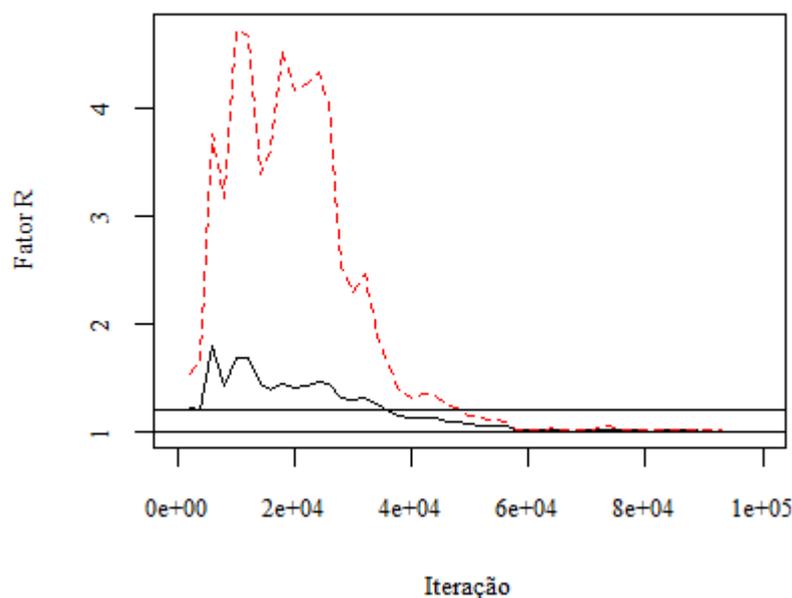


Figura 5-2 – Resultado do valor de R médio (linha cheia) e com limite superior do intervalo de 95% de confiança (linha tracejada) obtido pelo teste de Convergência de Gelman e Rubin (1992) para o Cenário PDH – Caso Sintético

Na Figura 5-3, foram plotados os gráficos das distribuições *a priori* e *a posteriori* para cada um dos parâmetros. Neles é possível verificar a atualização da distribuição *a priori* pela função verossimilhança dos dados, e a respectiva distribuição *a posteriori* estimada.

A partir dos gráficos da Figura 5-3, é possível observar o nível de desconhecimento *a priori* de todos os parâmetros. Esse nível de incerteza *a priori* é propagado na incerteza *a posteriori*, principalmente quando não há medições de campo que atualizem essa informação e reduzam as incertezas. Verifica-se também pelos gráficos que para os parâmetros c e κ , os dados da amostra não parecem condizer com a informação que se tem *a priori* sobre esses parâmetros, fazendo com que as distribuições *a posteriori* apresentem desvios-padrões maiores.

Para verificação dos resultados do modelo, as estimativas da média da distribuição de cada parâmetro, o conjunto de parâmetros de máximo valor da *posteriori* (maxpost) e suas incertezas foram comparadas com os valores arbitrados para esses parâmetros na geração dos dados sintéticos. Esse resultado pode ser observado na Tabela 5-1.

A partir da distribuição dos parâmetros da curva-chave (a , b , c , γ_1 e γ_2) são geradas as distribuições de vazões para qualquer valor de cota de nível d'água. A curva-chave apresentada na Figura 5-4 indica o valor médio e intervalo de credibilidade de 95% dessas distribuições de vazão para cotas de nível d'água entre as elevações 101 e 109. Para fins de

comparação e verificação da metodologia, plotou-se também nesse gráfico a curva-chave real.

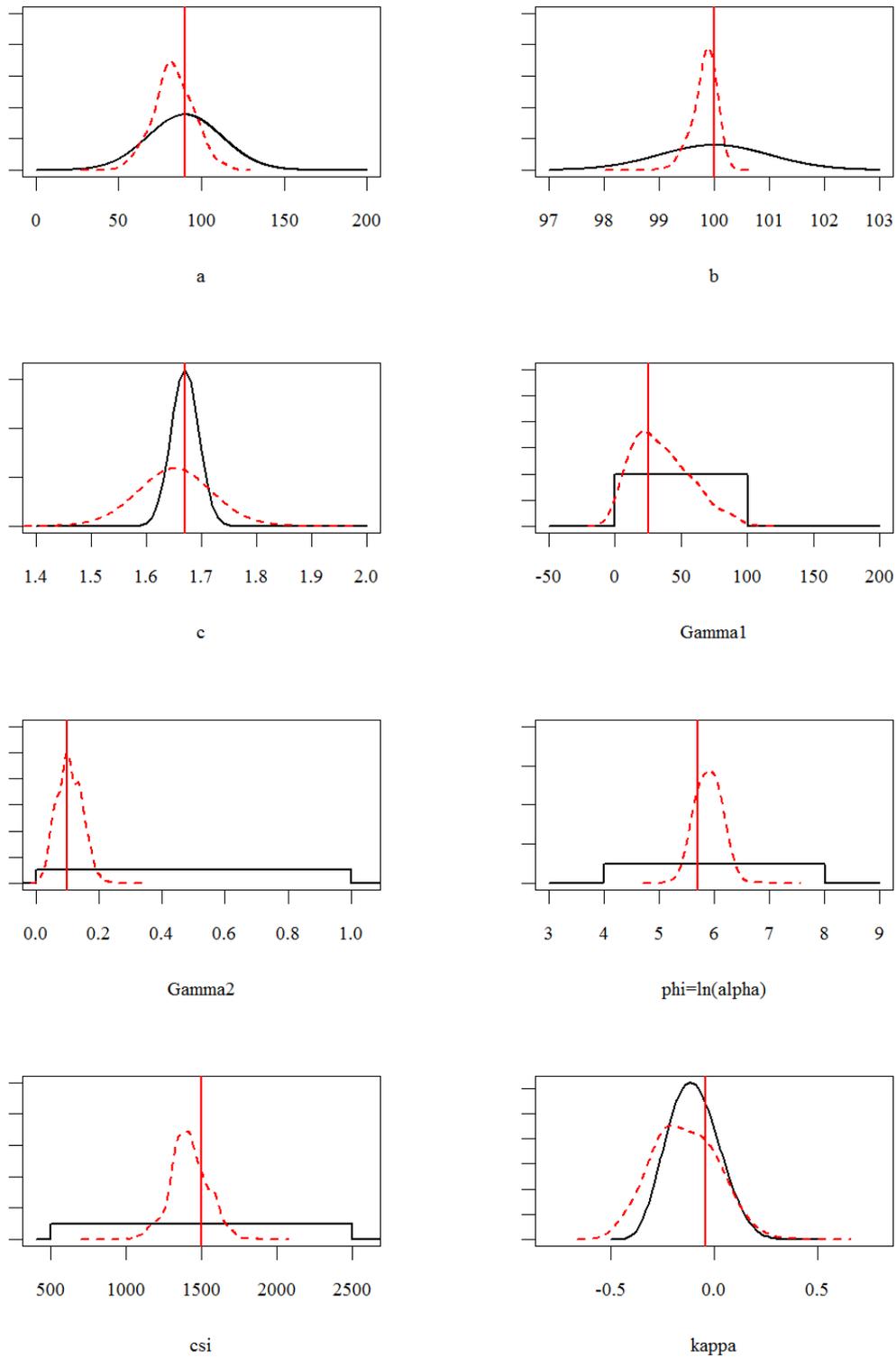


Figura 5-3 – Comparação entre as distribuições *a priori* (linha cheia) e *a posteriori* (linha tracejada) dos parâmetros com reta vertical do valor real dos parâmetros – Estudo de Caso Sintético

Tabela 5-1 – Valores estimados e reais dos parâmetros para o cenário sintético pelo modelo Bayesiano Completo (BC)

Parâmetros	Bayesiano Completo				Valor Real
	2,5%	97,5%	Média	Maxpost	
a	58,7	104,8	83,3	95,3	90,0
b	99,4	100,2	99,8	100,03	100
c	1,55	1,74	1,65	1,60	1,67
γ_1	0	75	35	2,54	25
γ_2	0,04	0,18	0,11	0,17	0,10
$\varphi(\ln(\alpha))$	5,45	6,32	5,90	5,59	5,70
ξ	1135	1632	1418	1339	1500
κ	-0,44	0,11	-0,14	-0,19	-0,04

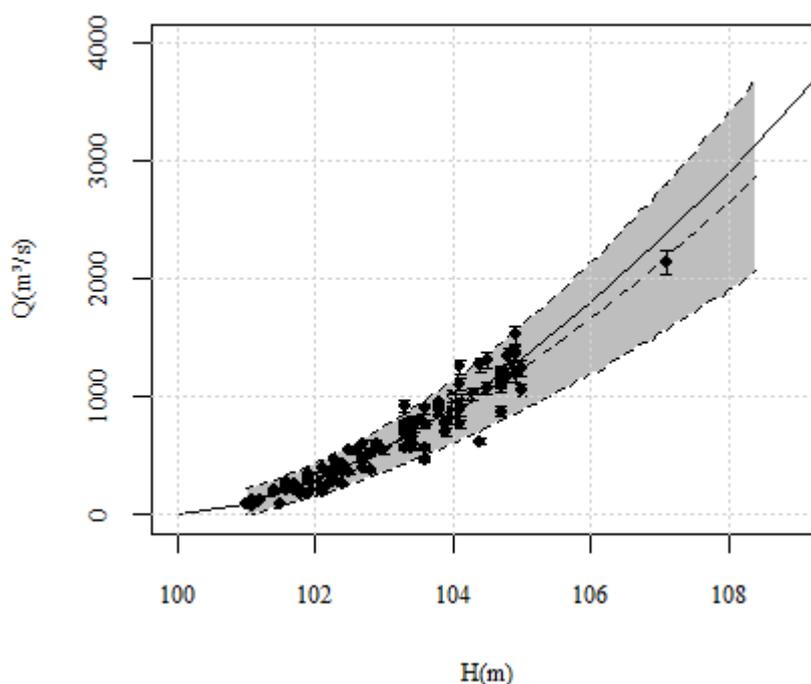


Figura 5-4 - Comparativos das curvas-chaves real (linha cheia) e média estimada (linha tracejada central) pelo método Bayesiano Completo (BC) com intervalo de credibilidade de 95% (hachura) e medições de descarga sintéticas com respectivos erros de medição (barra de erros) – Caso Sintético

É possível verificar, a partir dos gráficos da Figura 5-4, que o modelo gerou uma curva média muito próxima à curva sintética real e que 95 dos 100 pontos dos dados ficaram contidos no intervalo de credibilidade de 95%, corroborando as menções anteriores de que o modelo se comporta adequadamente para a estimativa dos parâmetros da curva-chave. Porém, ressalta-se que a curva-chave média estimada subestima o valor das vazões, o que torna fundamental que o resultado médio não seja analisado sem o intervalo de credibilidade a ele associado.

A exemplo do procedimento de geração da curva-chave a partir das distribuições dos parâmetros, podem ser estimadas as distribuições *a posteriori* de vazões máximas para a série de cotas máximas anuais. A vazão média e o intervalo de credibilidade dessa distribuição para cada valor de cota máxima anual podem ser observados na Figura 5-5 em conjunto com os valores das vazões máximas anuais gerados de forma sintética. Verifica-se na Figura 5-5 que os intervalos de credibilidade não são pequenos, demonstrando a importância da incorporação das incertezas da curva-chave na produção das séries de vazão e da melhor compreensão da limitação/restrição das informações utilizadas nos estudos que utilizam esses dados. Ressalta-se que esses intervalos representam os erros nas medições e no ajuste do modelo, além do desconhecimento *a priori* dos parâmetros. Além disso, é possível observar que, em função da curva-chave média estimada subestimar o valor das vazões, as vazões máximas reais são sistematicamente maiores que os valores médios estimados pelo modelo (linha cheia), embora estejam contidas no intervalo de credibilidade.

Finalmente, a partir da distribuição *a posteriori* dos parâmetros da GEV, são obtidas as distribuições dos quantis para diferentes tempos de retorno. Os gráficos da Figura 5-6 apresenta como exemplo as distribuições dos quantis de tempo de retorno de 10, 25, 50 e 100 anos. Com base nas distribuições dos diferentes quantis, obtêm-se a curva de frequência de cheias apresentada na Figura 5-7. A linha tracejada representa a média das distribuições dos quantis, e a hachura, seus respectivos intervalos de credibilidade. A linha cheia, do mesmo gráfico, representa a curva de frequência de cheias real e os pontos com barra de erros, as vazões máximas anuais e seus intervalos de credibilidade de 95% (ver Figura 5-5) plotados com sua probabilidade empírica de ocorrência.

Em relação a esse gráfico (Figura 5-7), devem ser apontados alguns pontos importantes. Primeiro, o resultado apresentado já demonstra o ganho da metodologia desenvolvida neste trabalho ao permitir a incorporação das incertezas da curva chave na estimativa da incerteza final dos quantis. Segundo, essas incertezas resultam em intervalos de credibilidade consideráveis, principalmente para períodos de retorno maiores, sendo insumo fundamental para análise de riscos. Terceiro, a curva de frequência estimada se distancia da curva real para tempos de retorno superiores a 10 anos. Entretanto, esse distanciamento demonstra uma propriedade importante da metodologia adotada de não consideração de que os eventos sejam equiprováveis. Isso significa que a curva se aproxima mais dos valores de maior ocorrência e menor incerteza, ou seja, vazões com períodos de retorno menores, e não necessariamente se ajusta às vazões maiores, menos frequentes e com maiores incertezas. Porém, como a análise proposta nesse trabalho não é um ajuste de único

de curva-chave com uma única curva de frequência de cheias, é possível concluir que a estimativa do modelo é adequada pelo fato do intervalo de credibilidade estimado incluir a curva real.

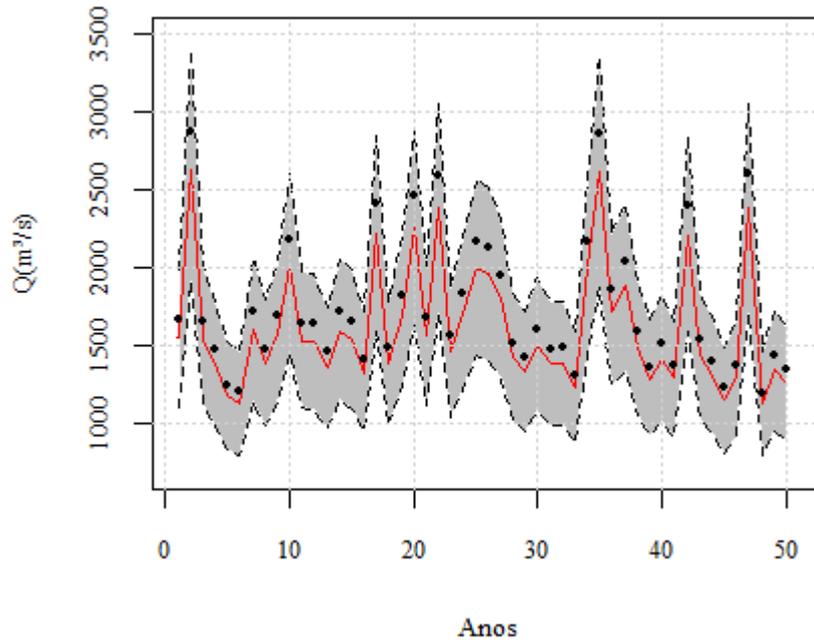


Figura 5-5- Série de vazões máximas anuais com valores médios (Linha cheia), intervalo de credibilidade (hachura) e dados de vazões máximas reais (pontos) – Caso Sintético

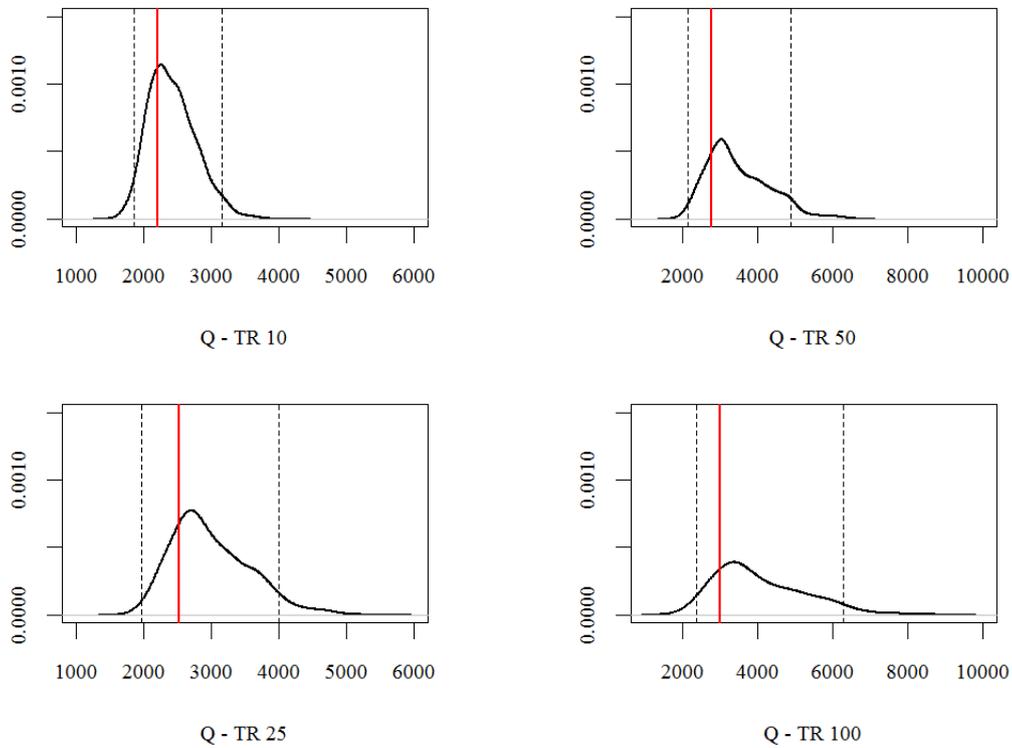


Figura 5-6 – Distribuições *a posteriori* dos quantis de cheia para os tempos de retorno de 10, 25, 50 e 100 anos com indicação do intervalo de credibilidade (linha tracejada) e valor real do quantil (linha cheia) – Caso Sintético

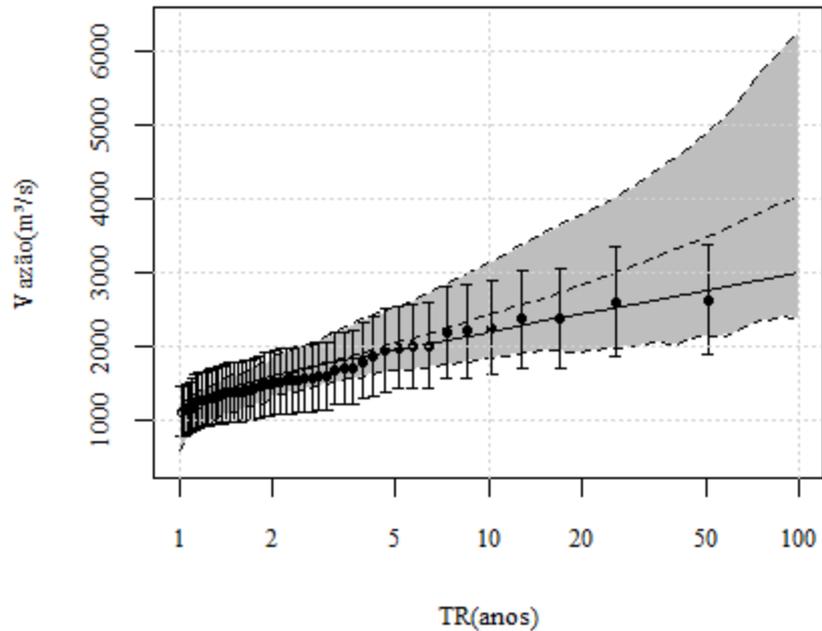


Figura 5-7 – Curva de frequência de cheias estimada (linha tracejada) e real (linha cheia), com intervalo de credibilidade de 95% (hachura), com os dados de vazão gerados e intervalo de credibilidade de 95% representado pela barra de erros – Caso Sintético

5.1.2 Resultados Gerais para o Modelo Bayesiano para GEV (BG)

O modelo BG estima apenas as distribuições *a posteriori* dos parâmetros φ , ξ e κ a partir de uma série única de vazões. Da mesma maneira que no modelo completo, é possível analisar os dados da cadeia e sua convergência, distribuições *a priori* e *a posteriori* e a curva de frequência de cheias estimada com seu intervalo de credibilidade.

Os resultados obtidos para o cenário PDH podem ser observados nas Figuras 5.8 a 5.12.

Observa-se pela análise das cadeias na Figura 5-8 e teste da Figura 5-9, que o modelo para GEV também converge. Além disso, por se tratar de um modelo de estimativa de menos parâmetros, são necessárias menos iterações para que essa convergência seja atingida. A Tabela 5-2 apresenta o resumo dos resultados obtidos nesse modelo com os valores reais.

Tabela 5-2 – Valores estimados e reais dos parâmetros para o cenário sintético pelo modelo Bayesiano da GEV (BG)

Parâmetros	Bayesiano GEV (BG)				Valor Real
	2,5%	97,5%	Media	Maxpost	
$\varphi(\ln(\alpha))$	5,38	5,86	5,61	5,55	5,70
ξ	1350	1509	1429	1428	1500
κ	-0,32	0,06	-0,13	-0,14	-0,04

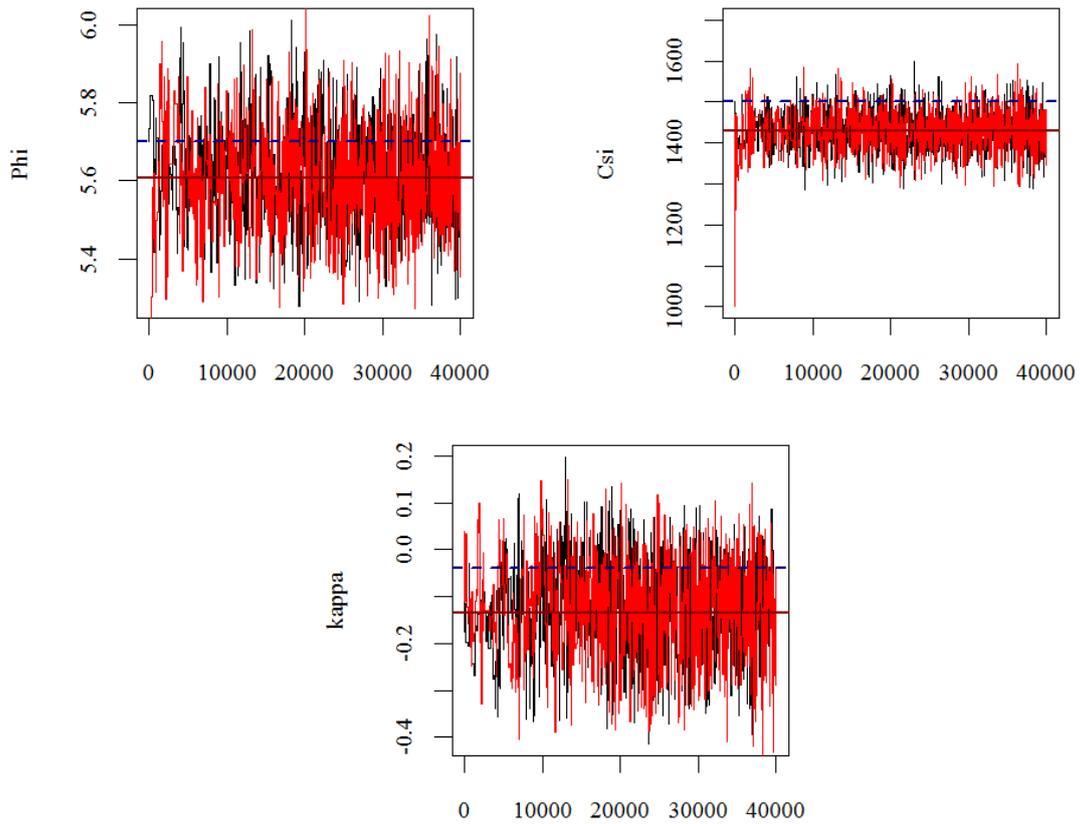


Figura 5-8 – Cadeias estimadas para 40.000 iterações pelo modelo Bayesiano da GEV (BG) para todos os parâmetros do cenário sintético PDH das cadeias 00 (pretas) e 01 (vermelhas) com o valor médio da cadeia (linhas cheias) e o valor real (tracejadas) – Caso Sintético

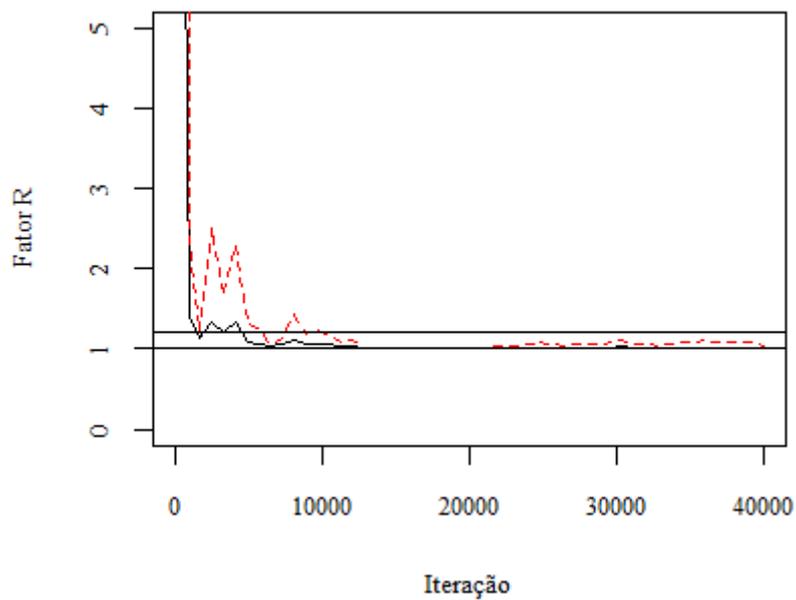


Figura 5-9 – Resultado do valor de R médio (linha cheia) e limite superior do intervalo de 95% de confiança (linha tracejada) obtido pelo teste de Convergência de Gelman e Rubin (1992) para o Cenário PDH com indicação dos limites de 1 e 1,2 – Caso Sintético

Como era esperado, pelo fato do modelo BG não considerar as incertezas das curvas-chaves, os resultados da Tabela 5-2 mostram que houve uma redução nos intervalos de credibilidade de todos os parâmetros, embora ainda contenham o valor real. Os gráficos da Figura 5-10 das distribuições *a priori* e *a posteriori* dos parâmetros da GEV estimadas pelo modelo BG também mostram esse resultado. A comparação direta dos resultados entre os diferentes modelos serão melhor detalhadas no cálculo da importância relativa no item 5.1.3.

Ainda com relação aos gráficos da Figura 5-10, é possível observar o reduzido conhecimento *a priori* que se tem em relação aos parâmetros da GEV e consequente redução desse desconhecimento na distribuição *a posteriori* após a utilização da técnica da inferência Bayesiana. É importante apontar nesses gráficos que a estimativa média difere do valor real, mas, conforme mencionado anteriormente, ele está contido no intervalo de credibilidade estimado.

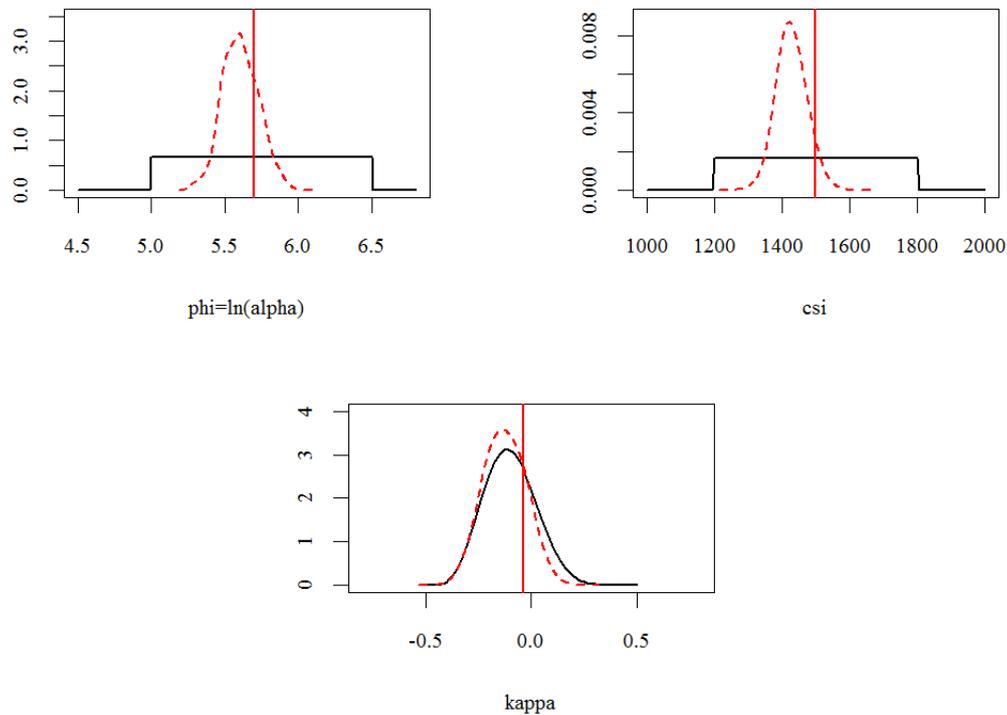


Figura 5-10 – Comparação entre as distribuições *a priori* (linha cheia) e *a posteriori* (linha tracejada) dos parâmetros com reta vertical do valor real dos parâmetros para o modelo BG – Caso Sintético

Com base nos parâmetros estimados, são calculados os quantis para os diferentes tempos de retorno. A Figura 5-11 exemplifica as distribuições calculadas para o modelo BG para os tempos de retorno de 10, 25, 50 e 100 anos.

A curva de frequência de cheias é, portanto, obtida pela média e intervalo de credibilidade das distribuições *a posteriori* dos quantis. Nesse caso, o intervalo de

credibilidade não incorpora a incerteza da curva-chave e representa apenas a incerteza amostral. O gráfico da Figura 5-12 mostra o comparativo entre a incerteza global e amostral para o cenário PDH na curva de frequência de cheias.

Nesse gráfico também são plotados os valores das vazões máximas e sua probabilidade empírica de ocorrência, ressaltando que para o modelo BG tem-se apenas uma série de vazões estimada com base em uma curva-chave única e não distribuições de valores de vazão, como ocorre no modelo BC.

Observa-se claramente na Figura 5-12 o quanto a incorporação da incerteza da curva-chave impacta no aumento do intervalo de credibilidade para todos os tempos de retorno. Essa observação corrobora ainda mais a necessidade da descrição clara dessas incertezas e sua incorporação nas análises de frequência de cheias.

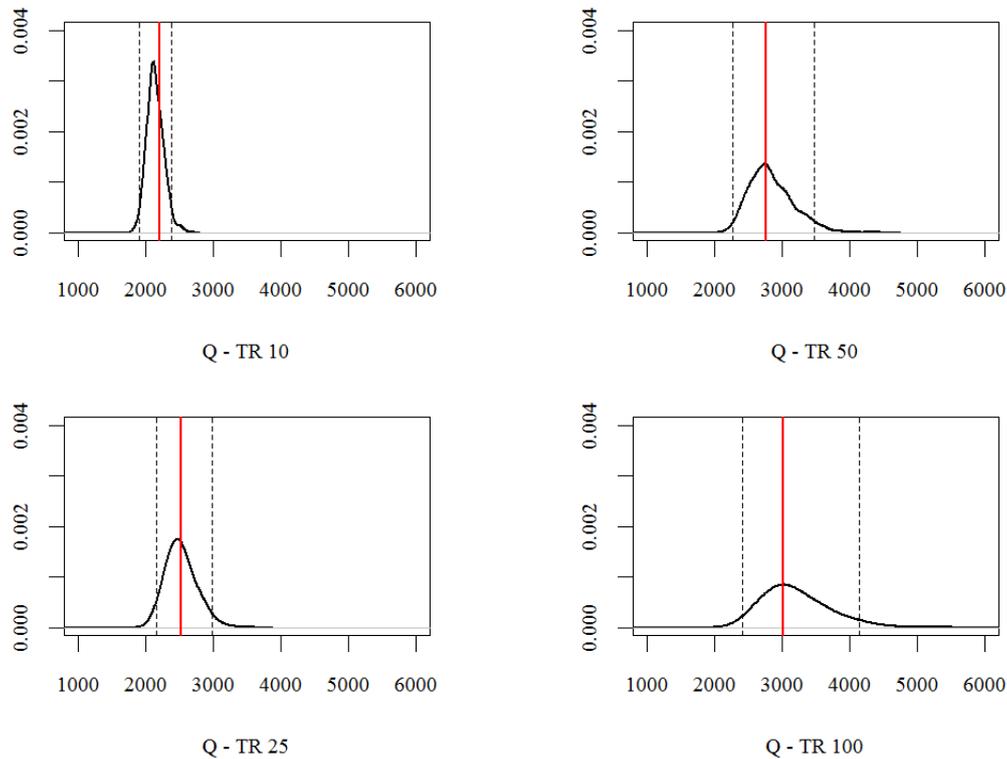


Figura 5-11 – Distribuições *a posteriori* dos quantis de cheia para os tempos de retorno de 10, 25, 50 e 100 anos com indicação do intervalo de credibilidade (linha tracejada) e valor real do quantil (linha cheia) para o modelo Bayesiano da GEV (BG) – Caso Sintético

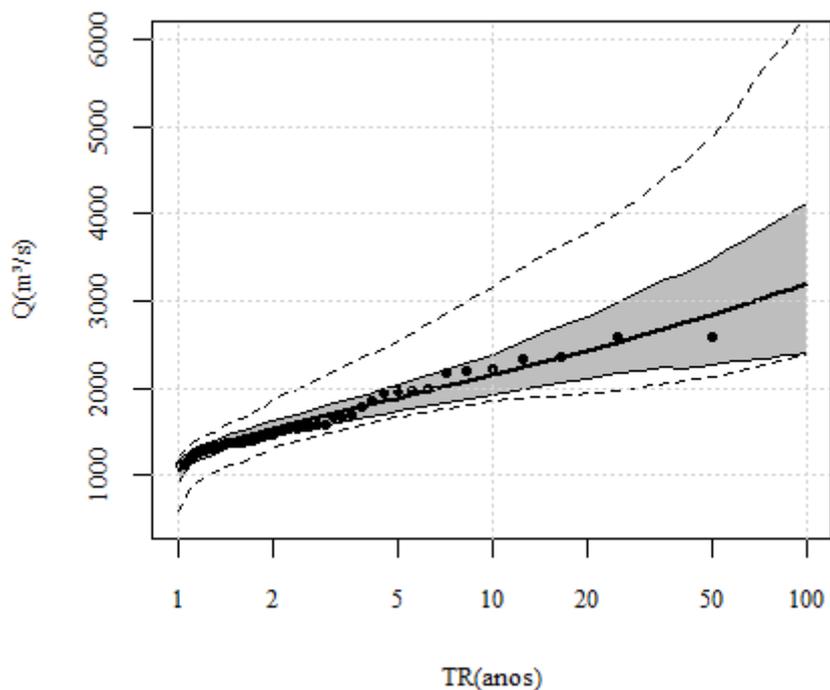


Figura 5-12 - Intervalos IG e IA das curvas de frequência de cheias para BG (hachurado) e BC (linhas tracejadas) para o cenário PDH com vazões máximas geradas pelos parâmetros da curva-chave de maior valor da *posteriori* (pontos) – Caso Sintético

5.1.3 Análise da Importância Relativa

Para análise da importância relativa da incerteza da curva chave em relação à incerteza global, os modelos BC e BG foram utilizados para estimativa da distribuição *a posteriori* dos parâmetros e dos quantis dos 8 cenários do estudo de caso sintético: PDH, P-DH, PD-H, PDH-, P-DH-, PD-H- e P-D-H-.

A partir das distribuições, foram calculadas as incertezas globais, amostrais e a importância relativa de cada um dos cenários para os tempos de retorno de 10, 25, 50 e 100 anos. O resultado obtido para as IR's pode ser observado no gráfico da Figura 5-13.

Verifica-se, pelos resultados apresentados, que a importância relativa da incerteza da curva-chave calculada para o caso sintético não é desprezível variando de 36% (P-DH-) a 65% (P-D-H).

Ainda a partir do gráfico da Figura 5-13, é possível, no exame individual de cada cenário, observar que há uma redução da IR à medida que se aumenta o tempo de retorno em todos os cenários. Esse resultado demonstra que, como era esperado, a incerteza da curva-chave se torna menos importante em relação à incerteza global para valores de vazão com menor probabilidade de ocorrência e consequentemente maiores incertezas amostrais.

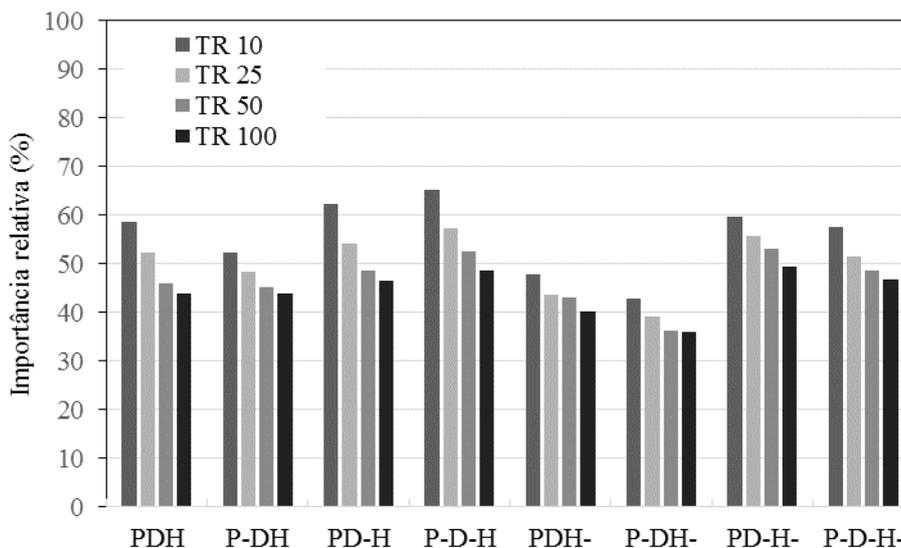


Figura 5-13 - Valor de IR para todos os cenários e tempos de retorno de 10, 25, 50 e 100 anos – Estudo de Caso Sintético

O gráfico da Figura 5-14 apresenta o resultado dos cenários de disponibilidade de dados de cotas máximas H, em função dos diferentes tempos de retorno. Nesse gráfico, fica mais clara a variação da IR para diferentes combinações de informação *a priori* (P e P-) e disponibilidade de dados de resumo de descarga (D e D-)

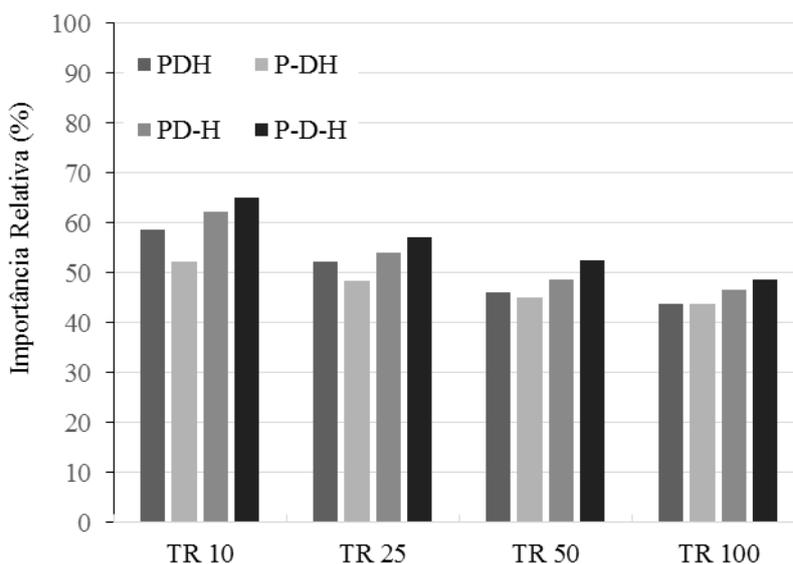


Figura 5-14 – Comparativo da importância relativa para PDH, P-DH, PD-H e P-D-H (Caso Sintético)

Observa-se na Figura 5-14 que, à medida que se reduz a disponibilidade de dados (D para D-), há um aumento no valor da IR. Esse resultado é razoável, tendo em vista que a expectativa é de que as incertezas na estimativa da curva-chave sejam de fato maiores quando existem poucos dados de medição disponíveis. Além disso, esse aumento

supracitado da IR pode ser explicado, nesse caso especificamente, pelo aumento da incerteza global (Modelo BC) das distribuições *a posteriori* dos parâmetros a , b , γ_1 e γ_2 , ξ e κ constatado a partir do exame das distribuições *a posteriori* dos parâmetros apresentadas na Figura 5-15.

A Figura 5-15 apresenta as distribuições *a priori*, *a posteriori* e o valor real de todos os parâmetros para ambos os cenários (PDH e PD-H). A partir desses gráficos, é possível analisar a forma das distribuições estimadas pelo modelo nos dois cenários em relação à informação que se tinha *a priori*, e o quanto essas distribuições se distanciam/aproximam do valor real. Comparando PDH e PD-H, verifica-se que além de impactar no aumento da incerteza global dos parâmetros, a redução da disponibilidade de dados de medição (D-) também afeta a estimativa média dos parâmetros que acaba se distanciando dos valores reais. Porém, embora a média esteja distante, os intervalos de credibilidade ainda cobrem os valores reais e carregam a informação desse aumento da incerteza que é propagado na estimativa dos quantis de cheia (Figura 5-17). É importante ressaltar que os gráficos das distribuições são suavizados para melhor apresentação dos resultados, portanto em alguns casos (ex: distribuição *a posteriori* de Γ) o que parece um erro da estimativa pelo fato da distribuição *a posteriori* possuir ocorrências em locais onde *a priori* é nula, é apenas um efeito dessa suavização e na realidade esses valores não acontecem.

A Tabela 5-3 demonstra as parcelas das incertezas amostral (IA) e global (IG) calculadas para os dois cenários e o valor final da importância relativa (IR).

Tabela 5-3 - IG, IA e IR para os cenários PDH e PD-H – Estudo de Caso Sintético

TR	PDH			PD-H		
	IG	IA	IR(%)	IG	IA	IR(%)
TR=10	0.53	0.22	59	0.67	0.25	62
TR=25	0.68	0.32	52	0.84	0.39	54
TR=50	0.79	0.42	46	0.98	0.50	49
TR=100	0.96	0.54	44	1.17	0.63	47

Analisando as métricas calculadas e apresentadas na Tabela 5-3 observa-se que:

- Em função da redução da disponibilidade dos dados, a incerteza global (IG) aumentou em torno de 20 a 25% de PDH para PD-H para todos os tempos de retorno;
- A incerteza amostral (IA) aumentou entre 15 a 20% aproximadamente. Como não houve variação da disponibilidade de H entre os cenários, entende-se que esse aumento se deu em função do uso de duas curvas-chaves diferentes para estimativa das séries de vazões máximas, não representando necessariamente um aumento na incerteza;

- O aumento da Importância Relativa (IR) foi relativamente pequeno (entre 4 e 5%), principalmente porque tanto IG como IA aumentaram, o que se compensou no cálculo da métrica de IR. Isso pode indicar que a metodologia adotada para se obter IA talvez não seja adequada;

- IG é sempre significativamente maior (85% a 170%) que IA indicando a pertinência e importância da incorporação das incertezas da curva-chave no cálculo dos quantis, foco principal deste trabalho.

Resultados semelhantes foram encontrados para os cenários de disponibilidade H- e eles podem ser vistos em maior detalhe no APÊNDICE II.

Partindo para análise dos cenários com diferentes níveis de informação *a priori* (P e P-), a expectativa era de que também houvesse um aumento da importância relativa nesses casos. Entretanto, pela análise do gráfico da Figura 5-14, o que se observa, é que há uma redução da IR de PDH para P-DH. O mesmo ocorre para PDH- e P-DH- (Figura 5-13). A Tabela 5-4 apresenta os resultados comparativos de IG, IA e IR para os cenários PDH para P-DH.

Tabela 5-4 – IG, IA e IR para os cenários PDH e P-DH – Estudo de Caso Sintético

TR	PDH			P-DH		
	IG	IA	IR(%)	IG	IA	IR(%)
TR=10	0,53	0,22	59	0,52	0,25	52
TR=25	0,68	0,32	52	0,70	0,36	48
TR=50	0,79	0,42	46	0,86	0,47	45
TR=100	0,96	0,54	44	1,05	0,59	44

Em relação aos resultados para esses cenários, o que se observa é que:

- IG aumentou em torno de 2 a 10% para tempos de retorno de 25, 50 e 100 anos, e reduziu para o tempo de retorno de 10 anos, aumento inferior ao observado anteriormente. Isso pode indicar que o impacto da redução da informação *a priori* é menos significativo que o impacto da redução da disponibilidade de dados de medição de cota de nível d'água e vazão;

- IA também aumentou entre 9 e 11% pelos mesmos motivos explicados no cenário comparativo anterior;

- Como IA aumentou mais do que IG, a importância relativa (IR) diminuiu para quase todos os tempos de retorno indicando que o cálculo dessa métrica deve ser melhor estudado.

O aumento de IG não tão significativo quanto dos cenários anteriores, pode ser explicado a partir da análise das distribuições *a posteriori* e *a priori* de PDH e P-DH. A Figura 5-16 apresenta o comparativo entre as distribuições.

Verifica-se que os intervalos de credibilidade das distribuições *a posteriori* aumentaram para os parâmetros a , b , γ_1 e κ . Entretanto, foram compensados pela redução dos demais. Nesse caso, isso pode ser explicado pela observação das distribuições comparativas do parâmetro c . Observa-se que, quando o desvio-padrão da distribuição *a priori* aumenta (P-), a estimativa de c tende para valores menores que 1,6. Isso demonstra que, por algum motivo, a amostra de dados de medição de descarga gerada é melhor explicada por esse valor de parâmetro, já que a verossimilhança está direcionando o resultado para esse extremo. Assim, a distribuição *a posteriori* de c para P-DH ficou com intervalo de credibilidade menor que a de PDH, ocasionando uma compensação entre os parâmetros e aumento menor de IG. A Figura 5-18 apresenta o exemplo da distribuição dos quantis e o intervalo de credibilidade global.

Resultado semelhante ocorre de PDH- para P-DH- e de PD-H para P-D-H. Os resultados comparativos completos para todos os cenários podem ser observados no APÊNDICE II.

Por fim, o último aspecto analisado em relação à importância relativa das incertezas da curva-chave, é a redução da disponibilidade de dados de cotas máximas anuais (H e H-). No gráfico da Figura 5-19, os cenários de mesmo fator P e D são plotados junto com a variação apenas de H (H e H-). A expectativa em relação a essa análise é que a incerteza amostral aumente de H para H-, reduzindo o valor da importância relativa. Isso fica confirmado na Figura 5-19 para os cenários PD, P-D (quatro primeiras barras de cada bloco) e P-D- (últimas duas barras dos blocos) em que se observa a redução de IR para todos os valores de TR.

Por outro lado, para os cenários de disponibilidade de dados de medição de descarga reduzida com *priori* informativa (PD-), o resultado foi menos intuitivo. Para o tempo de retorno de 10 anos, observa-se a redução esperada da IR de PD-H para PD-H-, enquanto que para os demais (TR= 25, 50 e 100) ocorre o inverso. A Tabela 5-5 permite analisar o porquê do aumento da IR nesses casos.

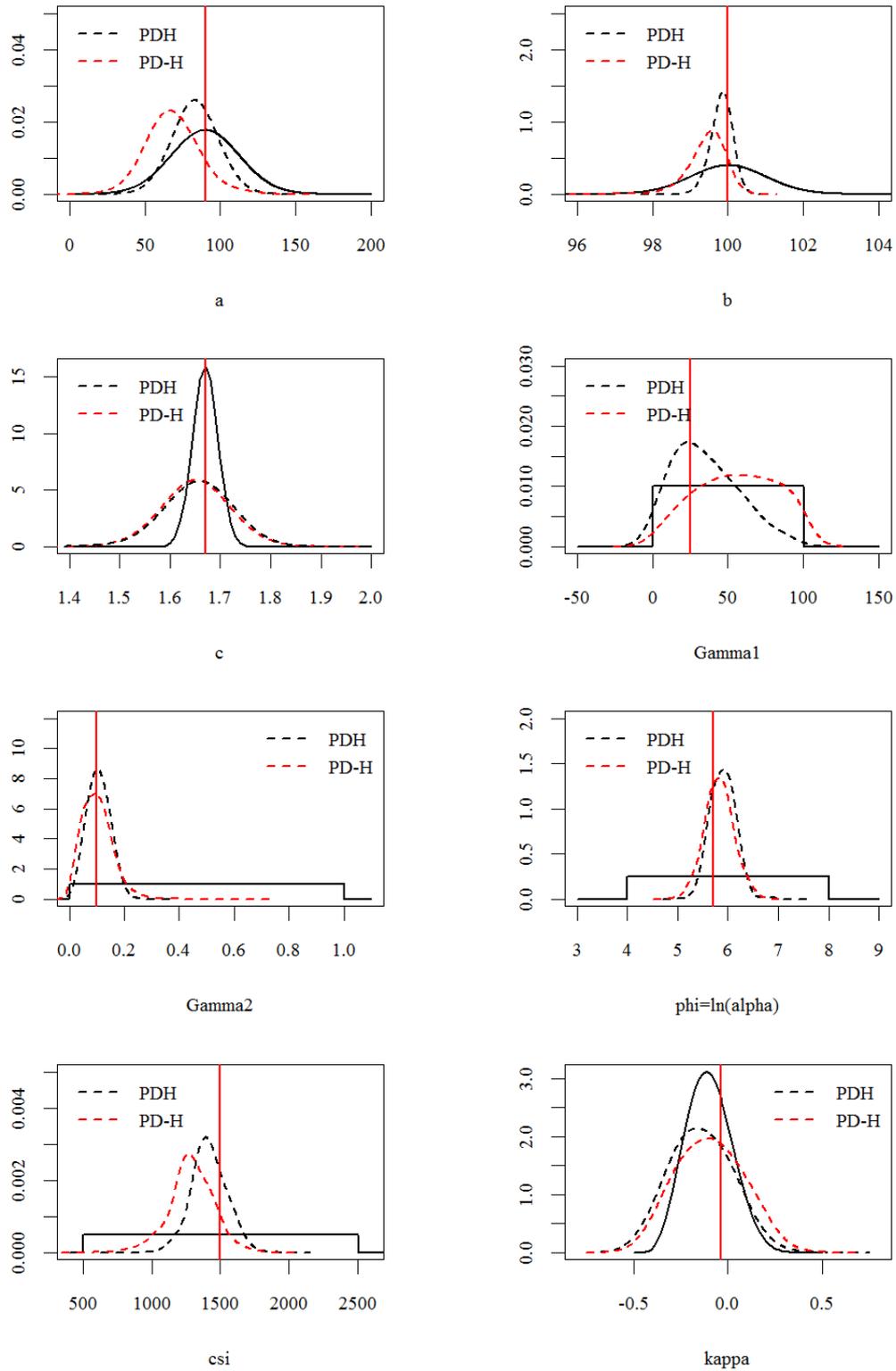


Figura 5-15 - Comparativo das distribuições *a priori* (linha cheia) e *a posteriori* (linha tracejada) de todos os parâmetros dos cenários PDH (preto tracejado) e PD-H (vermelho tracejado) com indicação do valor real (linha vertical) – Caso Sintético

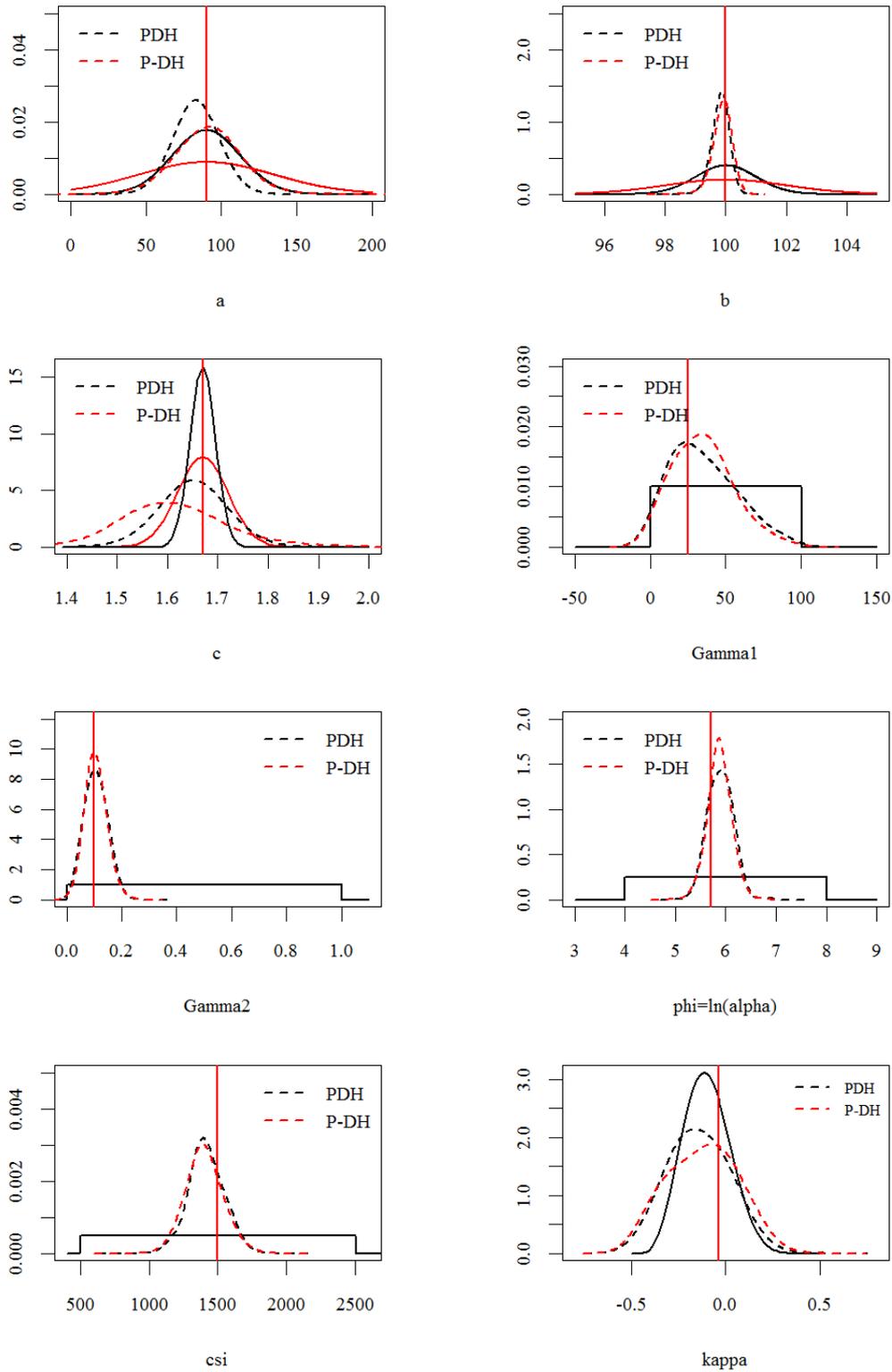


Figura 5-16 - Comparativo das distribuições *a priori* (linha cheia) e *a posteriori* (linha tracejada) de todos os parâmetros dos cenários PDH (preto tracejado) e P-DH (vermelho tracejado) com indicação do valor real (linha vertical) – Caso Sintético

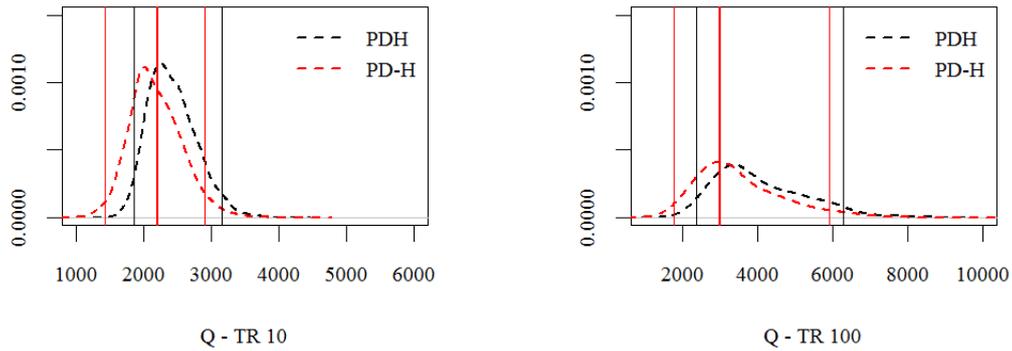


Figura 5-17 – Comparativo das distribuições *a posteriori* dos quantis de tempo de retorno de 10 e 100 anos para os cenários PDH e PD-H com a indicação dos respectivos intervalos de credibilidade (linhas cheias verticais) – Caso Sintético

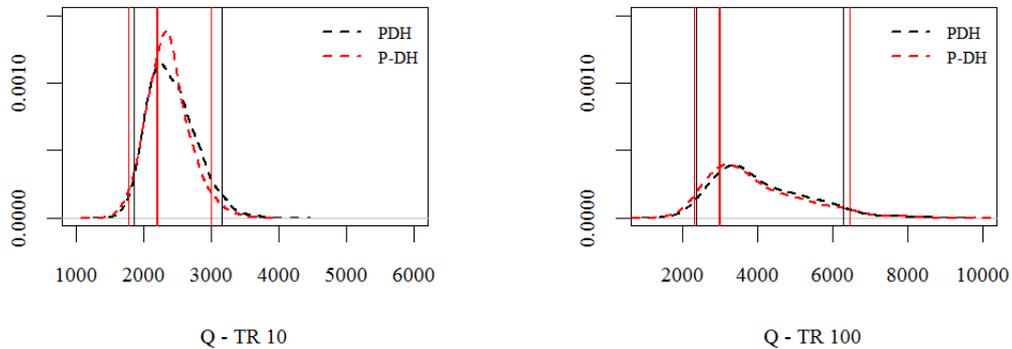


Figura 5-18 – Comparativo das distribuições *a posteriori* dos quantis de tempo de retorno de 10 e 100 anos para os cenários PDH e P-DH com a indicação dos respectivos intervalos de credibilidade (linhas cheias verticais) – Caso Sintético

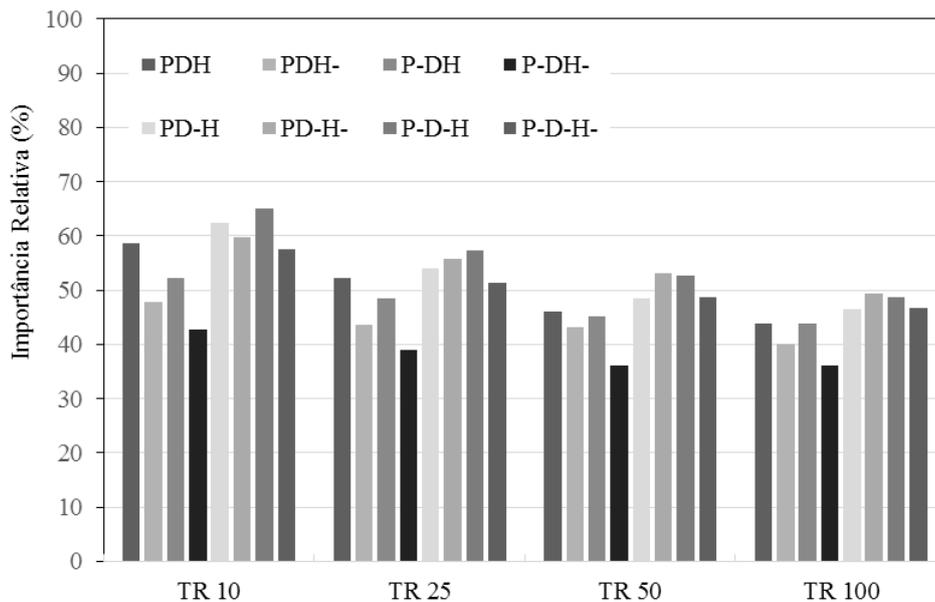


Figura 5-19 Comparativo da importância relativa para variações de H para H- (Caso Sintético)

Tabela 5-5 Cálculo da Importância Relativa (%) para os cenários PD-H e PD-H- - Caso Sintético

TR	PD-H			PD-H-		
	IG	IA	IR(%)	IG	IA	IR(%)
TR=10	0,67	0,25	62	0,73	0,29	60
TR=25	0,84	0,39	54	0,92	0,41	56
TR=50	0,98	0,50	49	1,07	0,50	53
TR=100	1,17	0,63	47	1,22	0,62	49

Verifica-se que tanto IG quanto IA aumentaram quando se reduziu a disponibilidade de dados de cotas máximas anuais. Esse resultado é esperado, tendo em vista que a redução da amostra impacta ambas as incertezas. Entretanto, o que ocorre é que, para TR maiores que 10 anos, o aumento observado em IA foi menor do que o aumento em IG, resultando no aumento da importância relativa. Isso demonstra que o impacto da variação na disponibilidade de dados de cotas máximas anuais pode não ser muito intuitivo pela métrica proposta e que mais estudos devem ser desenvolvidos para compreender a importância do aumento da disponibilidade de dados, levando em consideração também sua variabilidade.

5.2 Estudo de Caso – Cáceres

5.2.1 Resultados Gerais para o Modelo Bayesiano Completo (BC)

Para o caso real foram gerados os mesmos resultados discutidos no caso sintético. A grande diferença entre os dois, é que para o caso real os valores dos parâmetros são desconhecidos e as informações *a priori* se basearam nas informações das características hidráulicas de campo acerca da seção em estudo. A possibilidade de utilização dessas informações é uma das grandes vantagens do uso da inferência Bayesiana no método proposto. Estão apresentados nesse item os resultados encontrados para o cenário PDH.

As cadeias de Markov para todos os parâmetros podem ser observadas na Figura 5-20 e a verificação da convergência dessas cadeias encontra-se apresentada no gráfico da Figura 5-21. Observa-se que a partir de 40.000 iterações os valores do Fator R, tanto a média como o limite superior do intervalo de confiança de 95%, se aproximam da unidade e ficam inferiores ao limite de 1,2, indicando, de acordo com o teste, uma possível convergência. Em função disso, considerou-se também nesse caso um aquecimento de 50.000 iterações para o cálculo dos resultados. As distribuições *a priori* consideradas e *a posteriori* estimadas a partir das cadeias podem ser observadas na Figura 5-22.

Verifica-se, pela análise dessas distribuições, que a informação *a priori* dos parâmetros da curva-chave não parece condizer com os dados observados. Por exemplo, para o parâmetro a , a distribuição *a posteriori* com seu valor mais provável encontra-se na cauda inferior da distribuição *a priori*. Efeito semelhante ocorre para os parâmetros b e c .

Essas diferenças podem ser compreendidas a partir do exame da curva-chave estimada com seus intervalos de credibilidade apresentada na Figura 5-23. Verifica-se que as medições plotadas junto com a curva indicam uma possível inflexão a partir da cota de nível d'água 112. A não consideração dessa possível mudança, além de explicar a diferença nas estimativas da *posteriori* em relação ao conhecimento hidráulico *a priori*, implica na subestimação dos valores da vazão para cotas superiores a esse ponto e no aumento da incerteza do ajuste da curva-chave como um todo. Entretanto, como nesse estudo adotamos a premissa de que a curva possui apenas um segmento, isso não será considerado e deverá ser melhor aprofundado em estudos futuros. Ressalta-se, porém, que para considerar esse segundo segmento, é importante conhecer o comportamento hidráulico da seção para entender o motivo da inflexão e aplicar o modelo proposto neste trabalho para os dois segmentos também de forma simultânea.

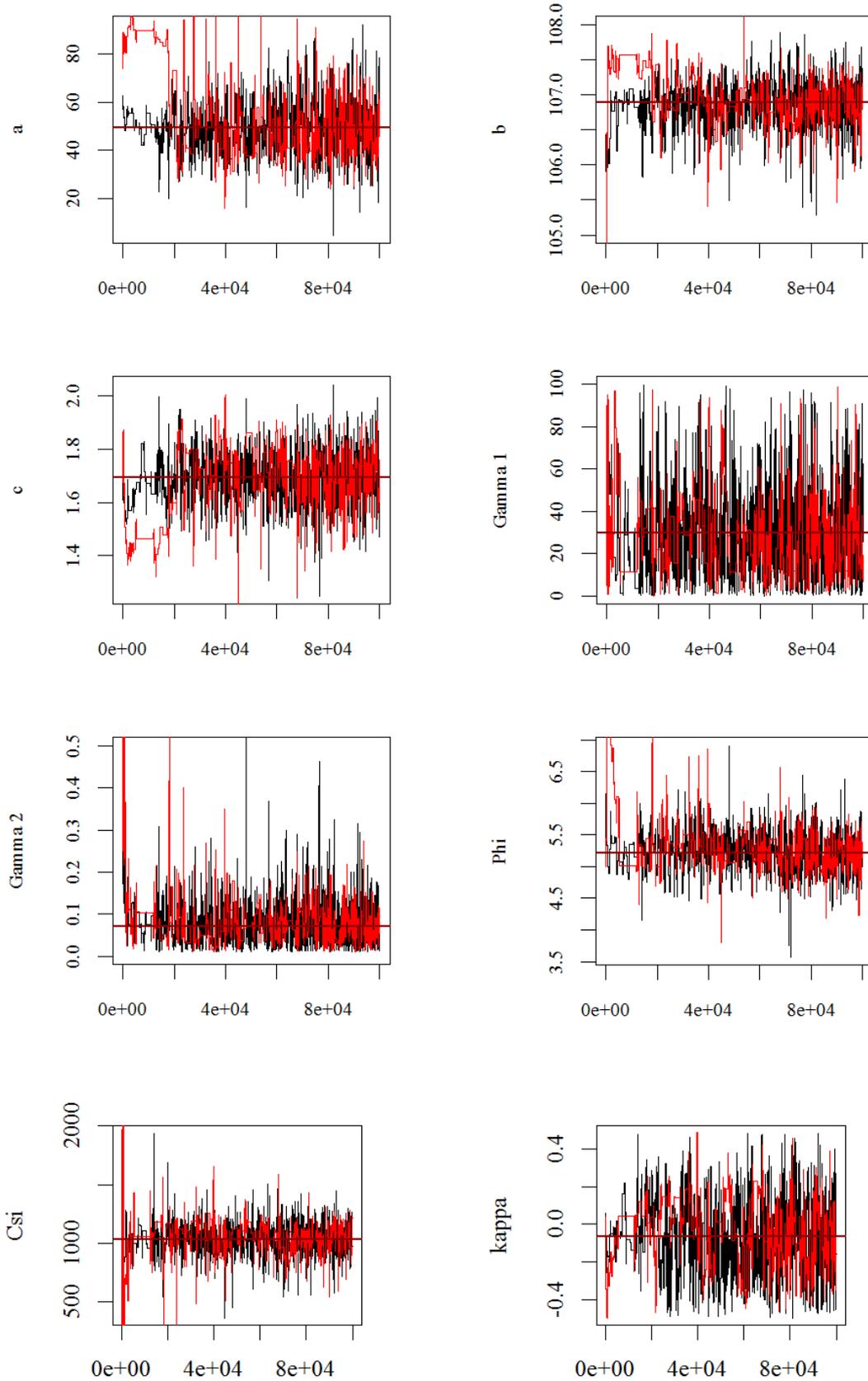


Figura 5-20 – Cadeias de Markov para 100.000 iterações pelo modelo Bayesiano Completo para todos os parâmetros do cenário sintético PDH das cadeias 00 (preto) e 01 (vermelho), valor médio (linha cheia) e o valor real (linha tracejada) para a Estação de Cáceres.

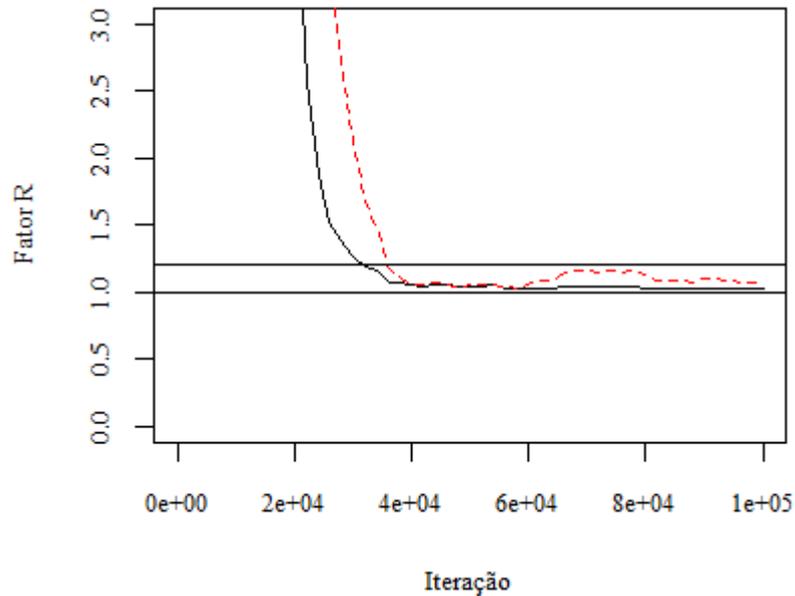


Figura 5-21 - Resultado do valor de R médio (linha cheia) e limite superior do intervalo de 95% de confiança (linha tracejada) obtido pelo teste de Convergência de Gelman e Rubin (1992) para o Cenário PDH com indicação dos limites de 1 e 1,2 – Caso Real - Cáceres

A Figura 5-24 explicita as incertezas nos valores de vazões máximas anuais calculados com base na série de cotas máximas anuais e nas cadeias dos parâmetros da curva-chave e do erro do modelo. O exame dessa série de vazões permite concluir que as incertezas nas séries de vazões máximas anuais não são desprezíveis, colocando em dúvida o procedimento usualmente adotado de consideração desses dados como sendo perfeitos.

Por fim, os gráficos da Figura 5-25 e Figura 5-26 mostram a distribuição *a posteriori* dos quantis e a curva de frequência de cheias calculada a partir da média e intervalo de credibilidade de 95% dessas distribuições. Esses gráficos são a representação mais importante do presente trabalho, pois demonstram as incertezas da curva-chave, derivadas do conhecimento das características hidráulicas do local e dos erros de medição, e amostrais no resultado final da distribuição dos quantis de cheia. A partir desse resultado, é possível avaliar com mais clareza os eventos e suas probabilidades de ocorrência com seu respectivo nível de incerteza, objetivo principal deste estudo.

Com relação ao resultado específico da estação de Cáceres, o que se observa é que as incertezas em relação aos parâmetros e a existência de um segundo segmento da curva não considerado, resultaram em incertezas bastante elevadas. Para vazão com tempo de retorno de 100 anos, por exemplo, o intervalo de credibilidade é de quase 2.500 m³/s. A recomendação nesse caso, é investigar melhor as características hidráulicas e tentar reduzir as incertezas aplicando o modelo para dois segmentos.

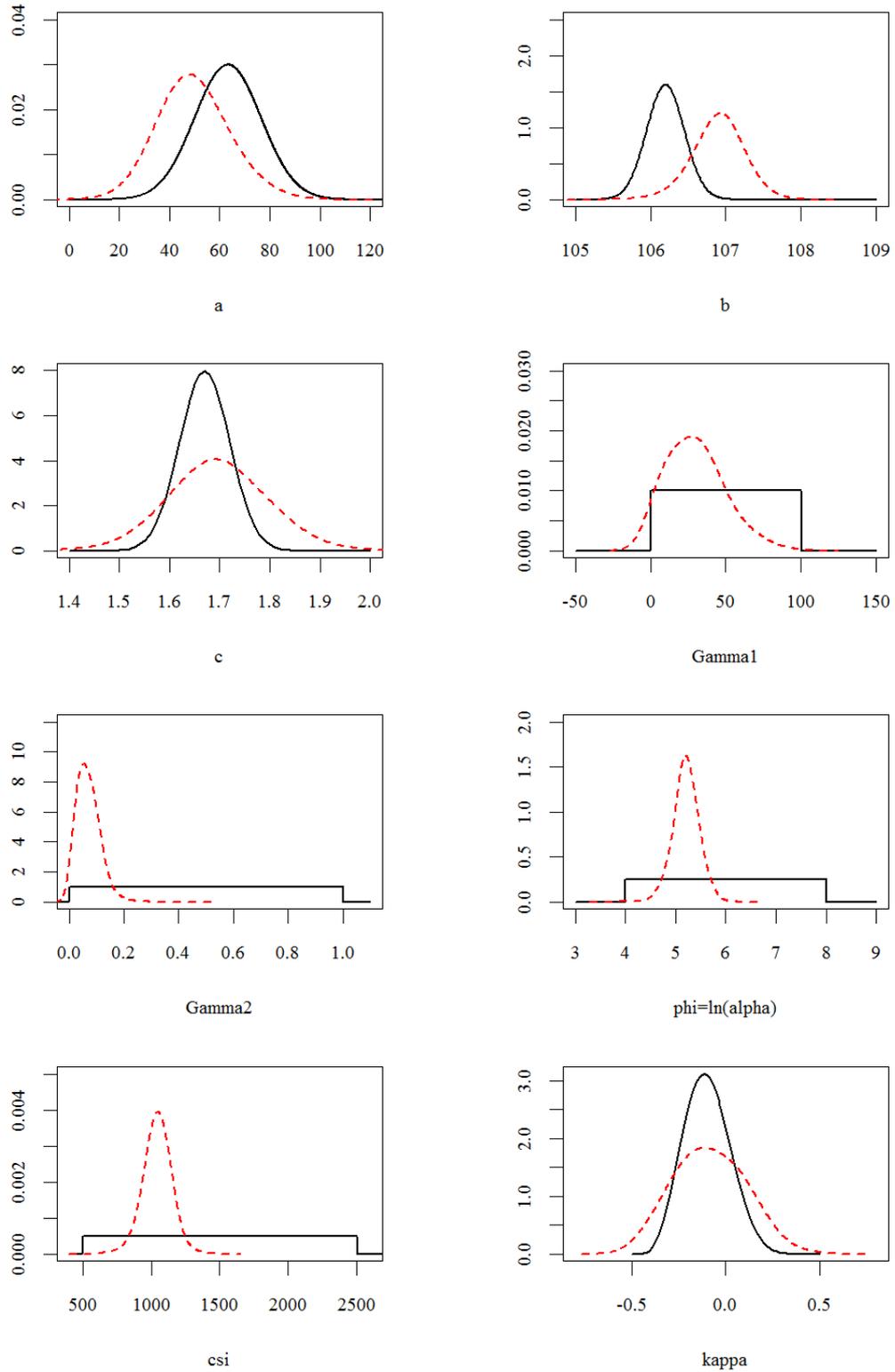


Figura 5-22 – Comparação entre as distribuições *a priori* (linha cheia) e *a posteriori* (linha tracejada) dos parâmetros para o modelo BC – Estudo de caso real em Cáceres

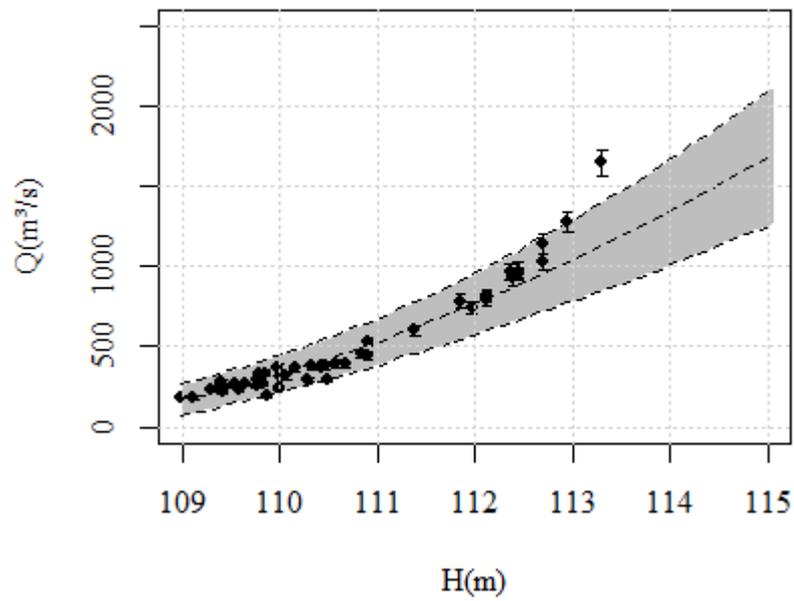


Figura 5-23 –Comparativos das curvas-chaves real (linha cheia) e estimada (linha tracejada central) pelo método Bayesiano Completo (BC) com intervalo de credibilidade de 95% (hachura) e medições de descarga observadas com erro de medição (barra de erros) – Caso Real - Cáceres

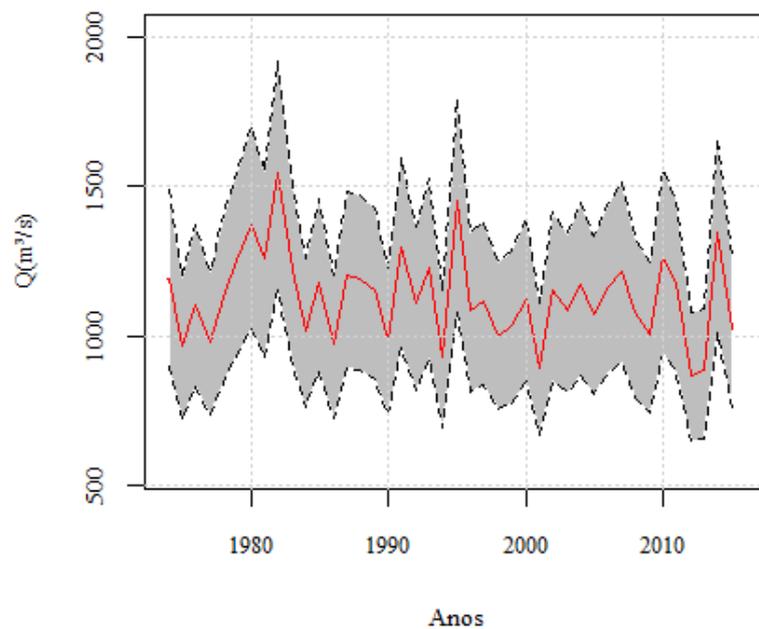


Figura 5-24 Série de vazões máximas anuais. Linha cheia corresponde aos valores médios e hachura com linha tracejada o intervalo de credibilidade de 95 % - Caso Real - Cáceres

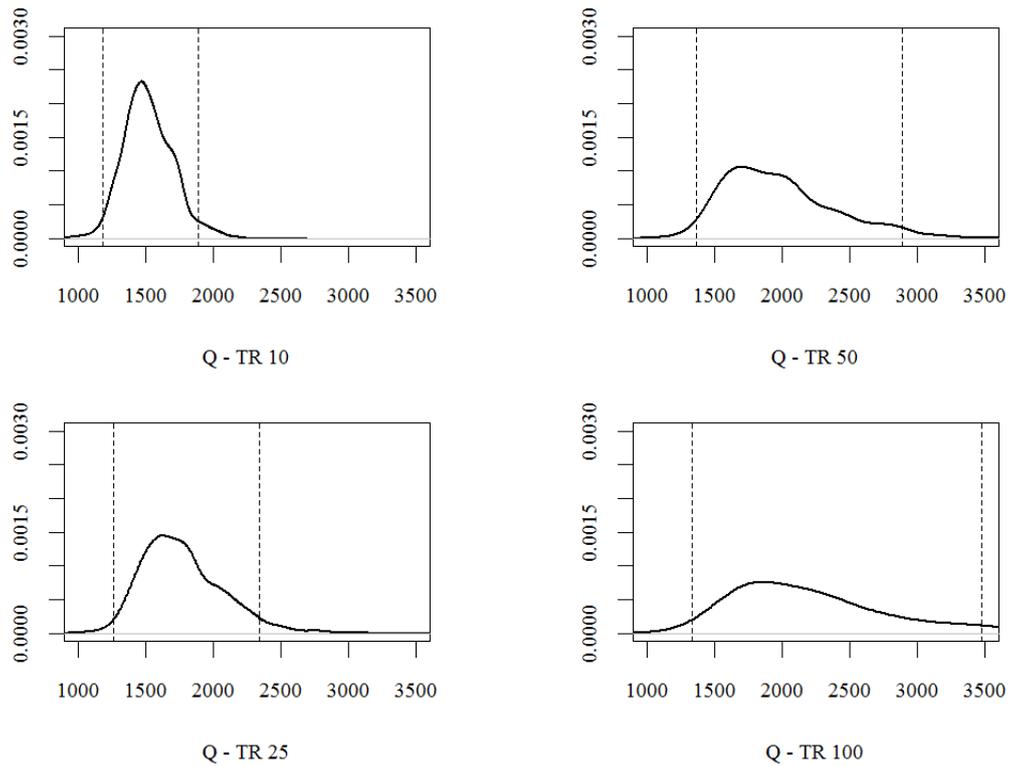


Figura 5-25 – Distribuições *a posteriori* dos quantis de cheia para os tempos de retorno de 10, 25, 50 e 100 anos com indicação do intervalo de credibilidade (linha tracejada) – Modelo BC – Caso Real

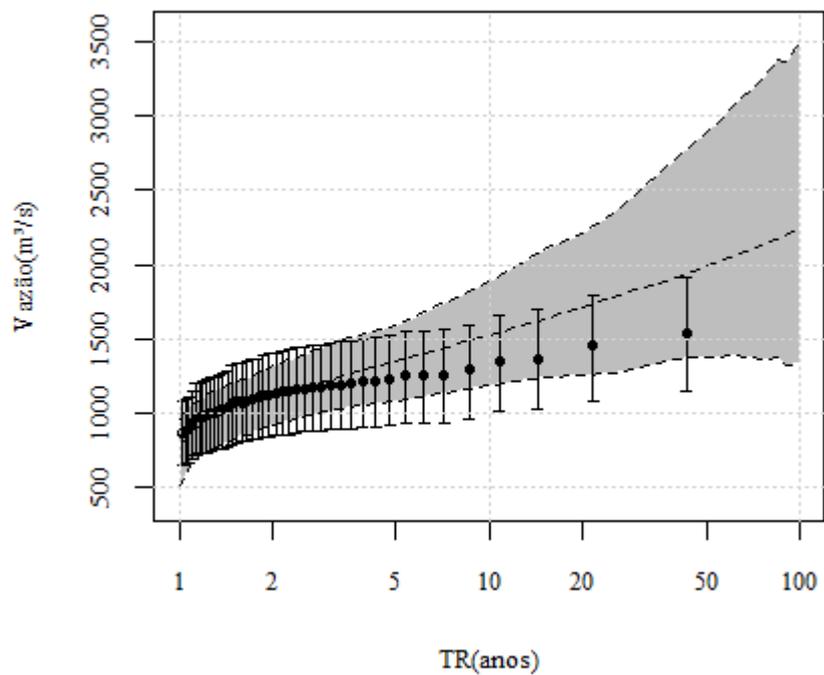


Figura 5-26 – Curva de frequência de cheias estimada pelo método bayesiano completo (BC) com intervalo de credibilidade de 95% para os dados da Estação de Cáceres

5.2.2 Resultados Gerais para o Modelo Bayesiano para GEV (BG)

Nesse item são apresentados os resultados do modelo Bayesiano para estimativa dos parâmetros da GEV (BG), cujo resultado será utilizado para calcular a incerteza amostral.

As Figuras 5.27, a 5.31 exemplificam os resultados encontrados para um dos cenários simulados, no caso, o cenário PDH.

O gráfico da Figura 5-27 mostra as cadeias de Markov dos 3 parâmetros da GEV. A análise visual dessas cadeias permite inferir que os resultados convergem para a distribuição *a posteriori*. O teste de convergência apresentada na Figura 5-28 corrobora essa constatação demonstrando uma possível convergência da cadeia a partir da iteração 10.000.

As distribuições *a posteriori* dos parâmetros plotadas nos gráficos da Figura 5-29 permitem verificar a redução do intervalo de credibilidade quando comparadas com as distribuições do modelo BC (Figura 5-22). Como as incertezas da curva-chave não são consideradas nesse modelo, era esperado que essa redução ocorresse.

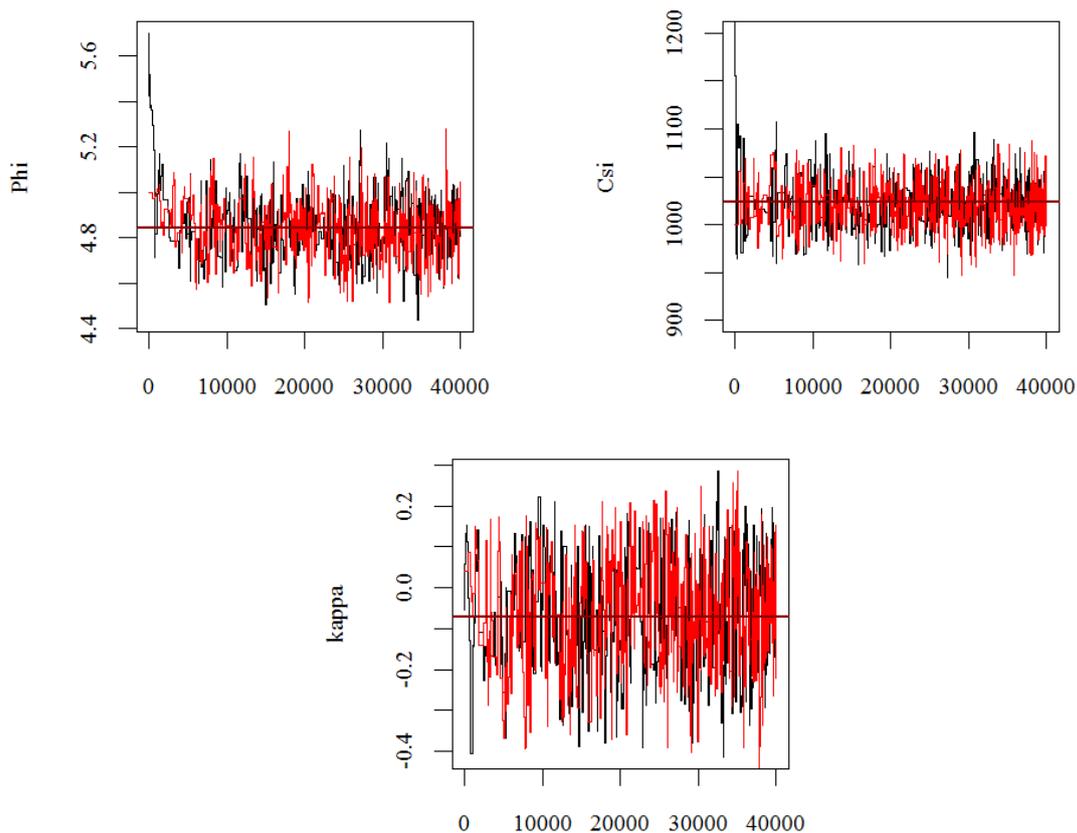


Figura 5-27 – Cadeias de Markov com 40.000 iterações pelo modelo Bayesiano da GEV (BG) para todos os parâmetros do cenário real PDH das cadeias 00 (preto) e 01 (vermelho), valor médio da cadeia sem o aquecimento (linha cheia) - Estudo de Caso – Cáceres

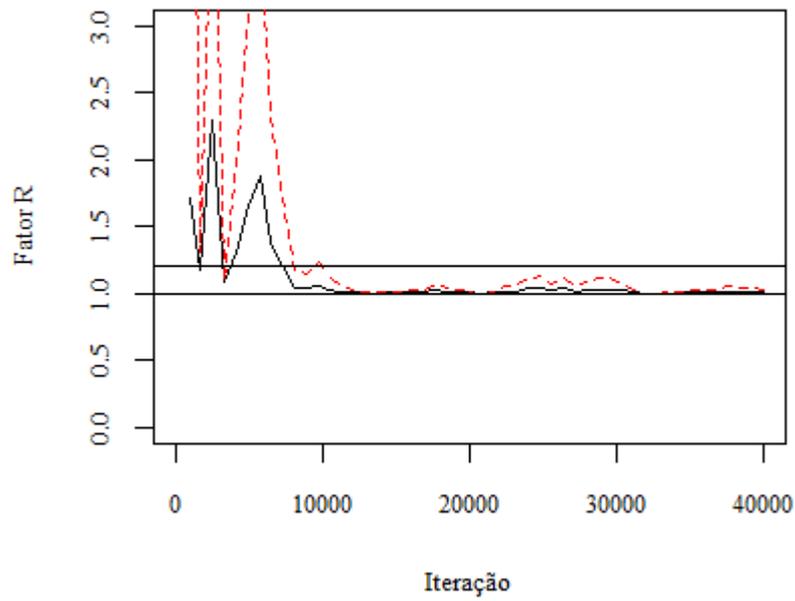


Figura 5-28 – Resultado do Teste de Convergência de Gelman e Rubin(1992) para o Modelo BG– Cenário PDH com indicação dos limites de 1 e 1,2 – Estudo de Caso – Cáceres

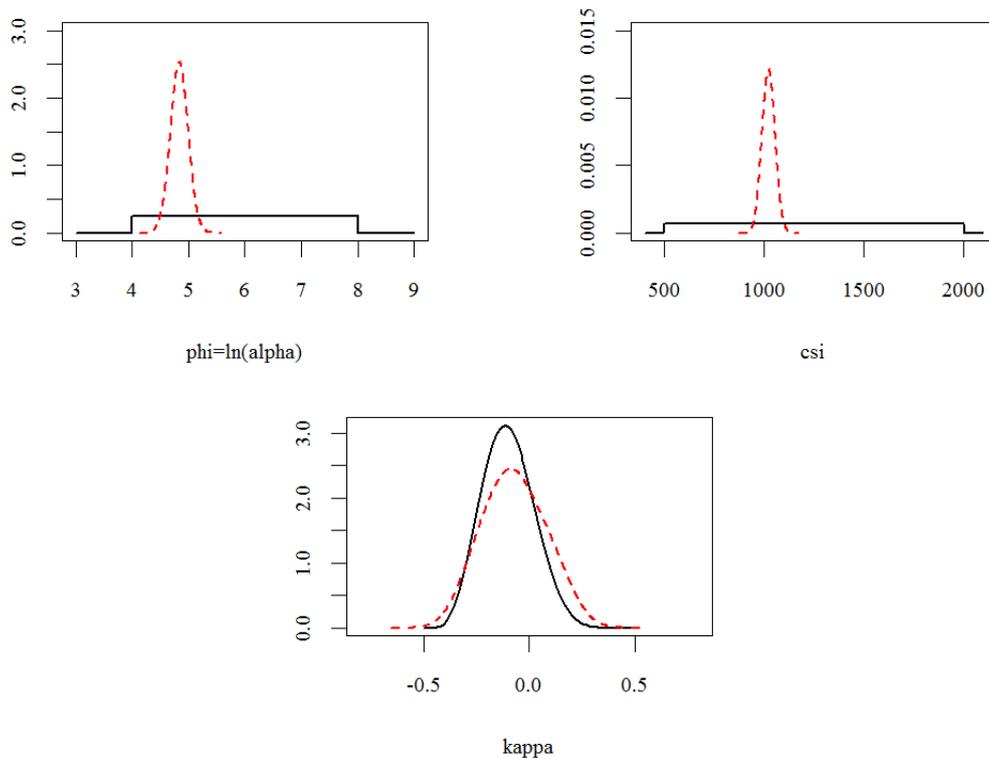


Figura 5-29 – Comparação entre as distribuições *a priori* (linha cheia) e *a posteriori* (linha tracejada) dos parâmetros para o modelo BC – Estudo de caso - Cáceres

Os gráficos da Figura 5-30 apresentam a distribuição *a posteriori* dos quantis de cheia sem a incorporação das incertezas. Da mesma forma como foi observado nas distribuições dos parâmetros, é possível notar a redução dos intervalos de credibilidade em relação aos resultados da aplicação do modelo Bayesiano completo (Figura 5-25).

A diferença fica ainda mais clara quando se comparam as curvas de frequência de cheias estimadas pelos dois modelos (Figura 5-31). Observa-se que os limites inferiores dos intervalos de credibilidade ficam muito próximos. Os limites superiores, por sua vez, são bastante diferentes, sendo o resultado do modelo BG inferior àquele obtido para BC, o que acarreta uma subestimação no valor das vazões.

Por outro lado, a curva do modelo BG (linha cheia) parece se adequar melhor aos valores de vazão estimados. Isso provavelmente acontece devido ao fato de que para esse modelo não são consideradas as incertezas nos valores das vazões. Assim, a estimativa dos parâmetros da GEV é feita apenas para se ajustar aos valores de vazão indicados com pontos no gráfico, enquanto no modelo BC essa estimativa leva em consideração uma incerteza significativa, principalmente nas vazões de maior magnitude, demonstrando mais uma vez a necessidade de consideração dessas incertezas nos dados de vazão.

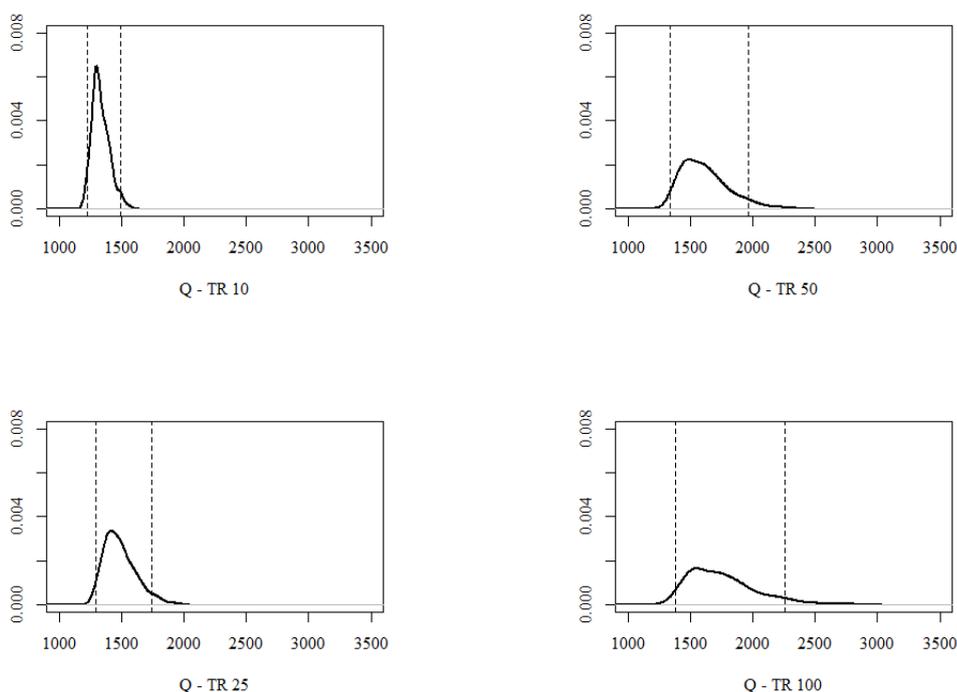


Figura 5-30 – Distribuições *a posteriori* dos quantis de cheia para os tempos de retorno de 10, 25, 50 e 100 anos com indicação do intervalo de credibilidade (linha tracejada) – Modelo BG – Caso Real

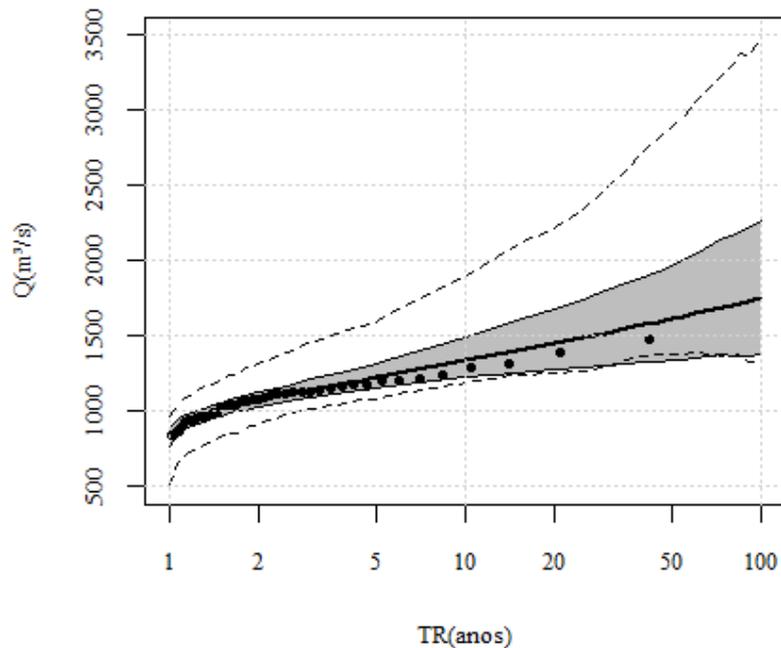


Figura 5-31 – Intervalos IG e IA das curvas de frequência de cheias para BG (hachurado) e BC (linhas tracejadas) para o cenário PDH com vazões máximas geradas pelos parâmetros de valor de máxima *posteriori* (pontos) – Caso Real

5.2.3 Análise da Importância Relativa

Os resultados para estação de Cáceres, apresentados nos itens 5.2.1 e 5.2.2, foram gerados para os 8 cenários propostos: PDH, P-DH, PD-H, PDH-, P-DH-, PD-H- e P-D-H-, cujas características foram descritas no item 4.5.2.

A partir desses resultados, calculou-se a importância relativa da incerteza da curva-chave. O gráfico da Figura 5-32 apresenta o resultado geral das importâncias relativas calculadas.

De forma geral, verifica-se, pelos resultados apresentados, que a importância relativa da incerteza da curva-chave no caso real também é considerável, variando de 47% (P-DH-) a 73% (P-D-H) e superior aos resultados encontrados no caso sintético.

Seguindo os mesmos passos de análise do caso sintético, procedeu-se a avaliação do impacto da importância relativa à medida que se reduz: o tempo de retorno (1), o nível de conhecimento *a priori* (2), a disponibilidade de medições de descarga (3) e de cotas máximas anuais (4).

Com relação à variação do tempo de retorno (1), chega-se a mesma conclusão obtida para o estudo de caso sintético. Analisando os blocos da Figura 5-32 individualmente, verifica-se que a importância relativa reduz com a redução da probabilidade de excedência do evento, como era esperado.

Para análise dos pontos (2) e (3), plotou-se o resultado comparativo para os cenários de disponibilidade H em função do tempo de retorno (Figura 5-33).

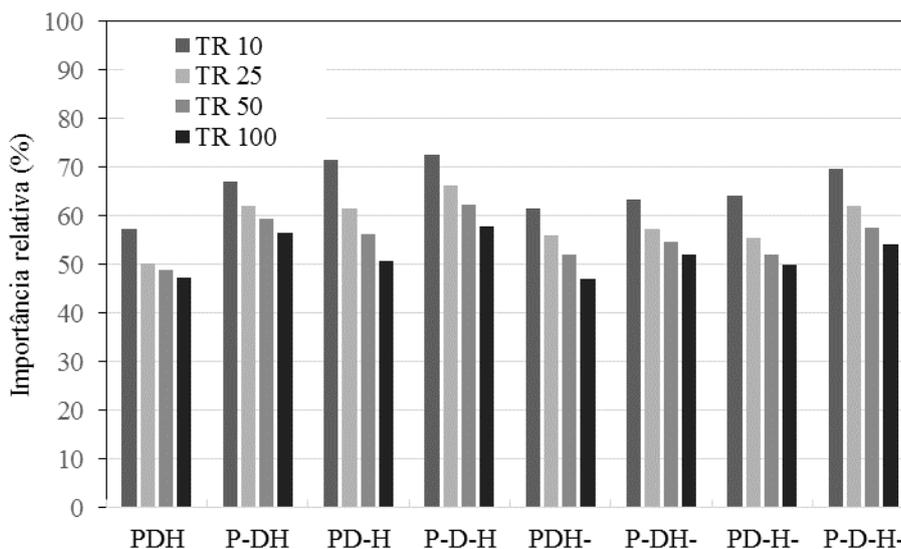


Figura 5-32 – Valor de IR para todos os cenários e tempos de retorno de 10, 25, 50 e 100 anos – Estudo de Caso - Cáceres.

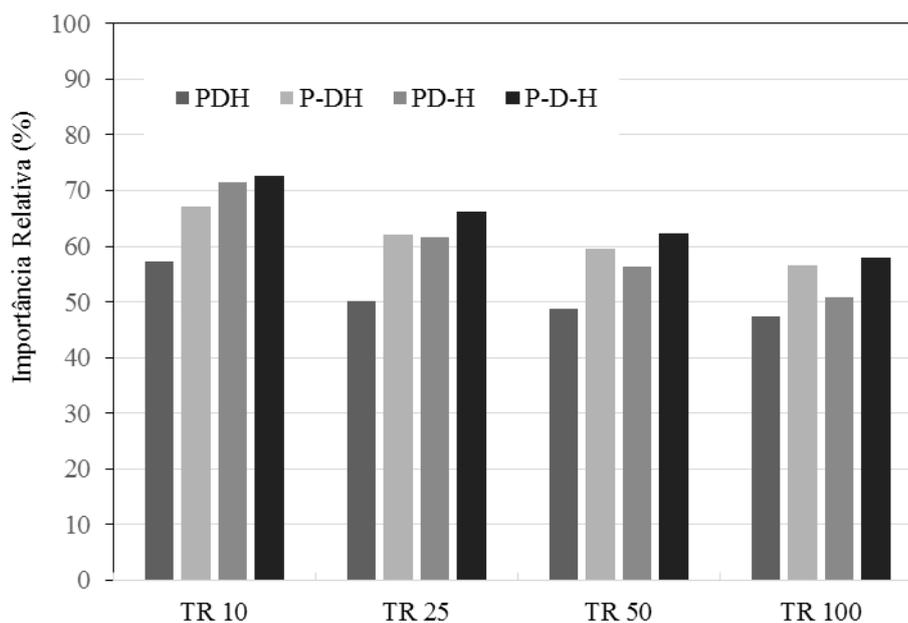


Figura 5-33 – Comparativo da importância relativa para PDH, P-DH, PD-H e P-D-H – Estudo de Caso - Cáceres

A partir da observação do gráfico da Figura 5-33, nota-se que à medida que se reduz o nível de conhecimento *a priori* (P para P-) e a disponibilidade de dados (D para D-) há um aumento no valor da IR em relação ao cenário PDH (primeira coluna de todos os blocos) para todos os tempos de retorno, o que era esperado tendo em vista o aumento da incerteza da curva-chave. Para os tempos de retorno de 10 anos, a redução da disponibilidade dos

dados resultou em um aumento maior da IR do que a redução do conhecimento *a priori* (1º bloco). Por outro lado, para os tempos de retorno maiores, ocorreu o inverso. Isso pode ser explicado pelo fato de que a maior parte das medições de descarga representam vazões com tempo de retorno inferior a 10 anos, já que não é usual existirem medições de cheias extremas, impactando, portanto, na incerteza do trecho mais baixo da curva de frequência. Por outro lado, para períodos de retorno maiores, essa redução tem pouca influência na incerteza final, por essas vazões fazerem parte de segmento cujos dados de medição de descarga normalmente já são restritos.

As constatações em relação ao aumento da IR podem ser explicadas pela análise das distribuições *a posteriori* dos parâmetros para os diferentes cenários em relação aos cenários de referência (PDH e PDH-) apresentadas no APÊNDICE II. Para exemplificar, são apresentados na Tabela 5-6 e Figura 5-34 a comparação entre o cenário PDH e P-DH.

Tabela 5-6 Cálculo da Importância Relativa (%) para os cenários PDH e P-DH – Caso Real

TR	PDH			P-DH		
	IG	IA	IR(%)	IG	IA	IR(%)
TR=10	0.46	0.20	57	0.52	0.17	67
TR=25	0.60	0.30	50	0.70	0.26	62
TR=50	0.76	0.39	49	0.86	0.35	59
TR=100	0.96	0.50	47	1.05	0.46	56

Fica claro, a partir da Tabela 5-6, que há um aumento da IG e redução da IA para todos os casos, o que explica o aumento observado e esperado para IR. A redução da IA só acontece em função da utilização de parâmetros de curvas-chaves diferentes em cada cenário, sendo essa redução apenas uma coincidência não relacionada com a diminuição da incerteza amostral de fato.

Outra constatação importante em relação ao exame da Tabela 5-6, conforme já apontado no caso sintético, é de que a incerteza global é bastante superior a incerteza amostral. Para TR de 10 anos, IG chega a ser 200% maior que IA. Esse resultado demonstra a importância da consideração de todas as incertezas inerentes ao processo de produção das séries de vazões e propagação das mesmas nas estimativas dos quantis.

Comparando os resultados das distribuições *a posteriori* dos cenários PDH e P-DH apresentados na Figura 5-34, a conclusão em relação ao aumento da incerteza de um cenário para outro é menos intuitiva. Na maioria dos parâmetros não houve aumento significativo do intervalo de credibilidade de um cenário para o outro (b , γ_2 , ξ , φ , e κ) sendo que em

alguns casos, ainda é possível se observar uma pequena redução desse intervalo (a, γ_1). No caso do parâmetro a , por exemplo, é possível verificar que com a mudança da distribuição *a priori*, ela passou a explicar melhor os dados observados, e a distribuição *a posteriori* tende para valores menores que 50 com intervalo de credibilidade ligeiramente menores. Fato semelhante foi visualizado no parâmetro c no estudo de caso sintético.

Entretanto, apesar dessa diferença, a combinação das pequenas variações dos outros parâmetros levou ao aumento da incerteza global dos quantis observada. Para visualização mais clara dessa diferença, a distribuição dos quantis com seu intervalo de credibilidade de maior densidade de probabilidade (HPD) de 95% foi plotado no gráfico da Figura 5-36. A partir das linhas é possível ver esse aumento, que varia de 9 a 12%, da incerteza global de um cenário para o outro para os tempos de retorno de 10 e 100 anos.

Resultados semelhantes são observados também do cenário PDH- para P-DH- (APÊNDICE II).

É importante apontar que, devido à suavização na hora de representar os gráficos, a distribuição de γ_1 e γ_2 indicam a existência de valores negativos desses parâmetros, porém isso não ocorre na realidade sendo apenas um efeito da plotagem.

Para exemplificar o que ocorre quando se reduz a disponibilidade dos dados de medição mantendo a mesma condição distribuição *a priori* (P) são apresentados na Tabela 5-7 os resultados detalhados de IG, IA e IR para os cenários PDH e PD-H.

Tabela 5-7 - Cálculo da Importância Relativa (%) para os cenários PDH e PD-H – Caso Real

TR	PDH			PD-H		
	IG	IA	IR(%)	IG	IA	IR(%)
TR=10	0.46	0.20	57	0.59	0.17	72
TR=25	0.60	0.30	50	0.73	0.28	62
TR=50	0.76	0.39	49	0.86	0.38	56
TR=100	0.96	0.50	47	1.05	0.52	51

Conforme já constatado anteriormente, a redução dos dados de D para D- resultou em um aumento da IR. O que se observa no exame da Tabela 5-7 é que, da mesma maneira como ocorre em P para P-, houve um aumento da IG acompanhado da redução de IA, mais uma vez com diferenças significativas entre essas incertezas.

Os gráficos das distribuições comparativas dos cenários apresentados na Figura 5-35

Figura 5-35 permitem analisar o que acontece com os parâmetros nesse caso. Verifica-se, bem mais claramente nesses cenários, que todos os parâmetros tiveram suas

incertezas/desvio padrão aumentados quando da redução da disponibilidade de dados de medição de descarga. Além disso, o valor médio de γ_1 e γ_2 referentes ao erro do ajuste do modelos também aumentaram, o que era esperado já que eles representam o erro do ajuste do modelo. A combinação de todos os fatores resultou portanto, no aumento da incerteza global dos quantis estimado. A distribuição dos quantis de tempo de retorno de 10 e 100 anos podem ser observados na Figura 5-37.

Por fim, analisa-se o último ponto (3), referente ao impacto da redução da disponibilidade de dados de cotas máximas. Assim a exemplo do que foi feito no Estudo de Caso sintético, plotam-se os cenários de mesmo fator P e D juntamente com a variação apenas de H. O gráfico pode ser observado na Figura 5-38.

O gráfico da Figura 5-38 mostra para cada bloco de TR a comparação do cenário de disponibilidade H ao lado do seu respectivo cenário H-. Assim, analisando os pares para cada TR, observa-se que para os cenários P-D, PD- e P-D-, a redução da disponibilidade dos dados de cotas máximas resultou em uma redução da IR para todos os tempos de retorno. Esse resultado vai de encontro à expectativa de que à medida que se reduz a disponibilidade dos dados, a incerteza amostral passa a ser mais significativa, reduzindo assim a importância da incerteza da curva-chave e conseqüentemente a importância relativa (IR). Entretanto, no cenário PD isso não ocorre.

A Tabela 5-8 apresenta as incertezas globais e amostrais para ambos os casos que permitem entender o que aconteceu nesse cenário especificamente.

Tabela 5-8 - Cálculo da Importância Relativa (%) para os cenários PDH e PDH- (Caso Real)

TR	PDH			PDH-		
	IG	IA	IR(%)	IG	IA	IR(%)
TR=10	0.46	0.20	57	0.59	0.17	72
TR=25	0.60	0.30	50	0.73	0.28	62
TR=50	0.76	0.39	49	0.86	0.38	56
TR=100	0.96	0.50	47	1.05	0.52	51

Observa-se pela Tabela 5-8 que, conforme esperado, os valores de IG e IA aumentaram de H para H-. Porém, o aumento de IG (9 a 26%) foi superior ao aumento de IA, resultando no aumento da IR. Além disso, observa-se que a redução da disponibilidade de dados de cotas máximas (H para H-), nesse caso, quase não impactou nos resultados das incertezas amostrais, podendo a diferença observada entre os cenários estar relacionada apenas à diferença nas curvas-chaves adotadas. Para se tirar conclusões melhores quanto ao

impacto na redução dos dados de cotas máximas anuais seriam necessários estudos mais aprofundados com diferentes cenários diferentes de H.

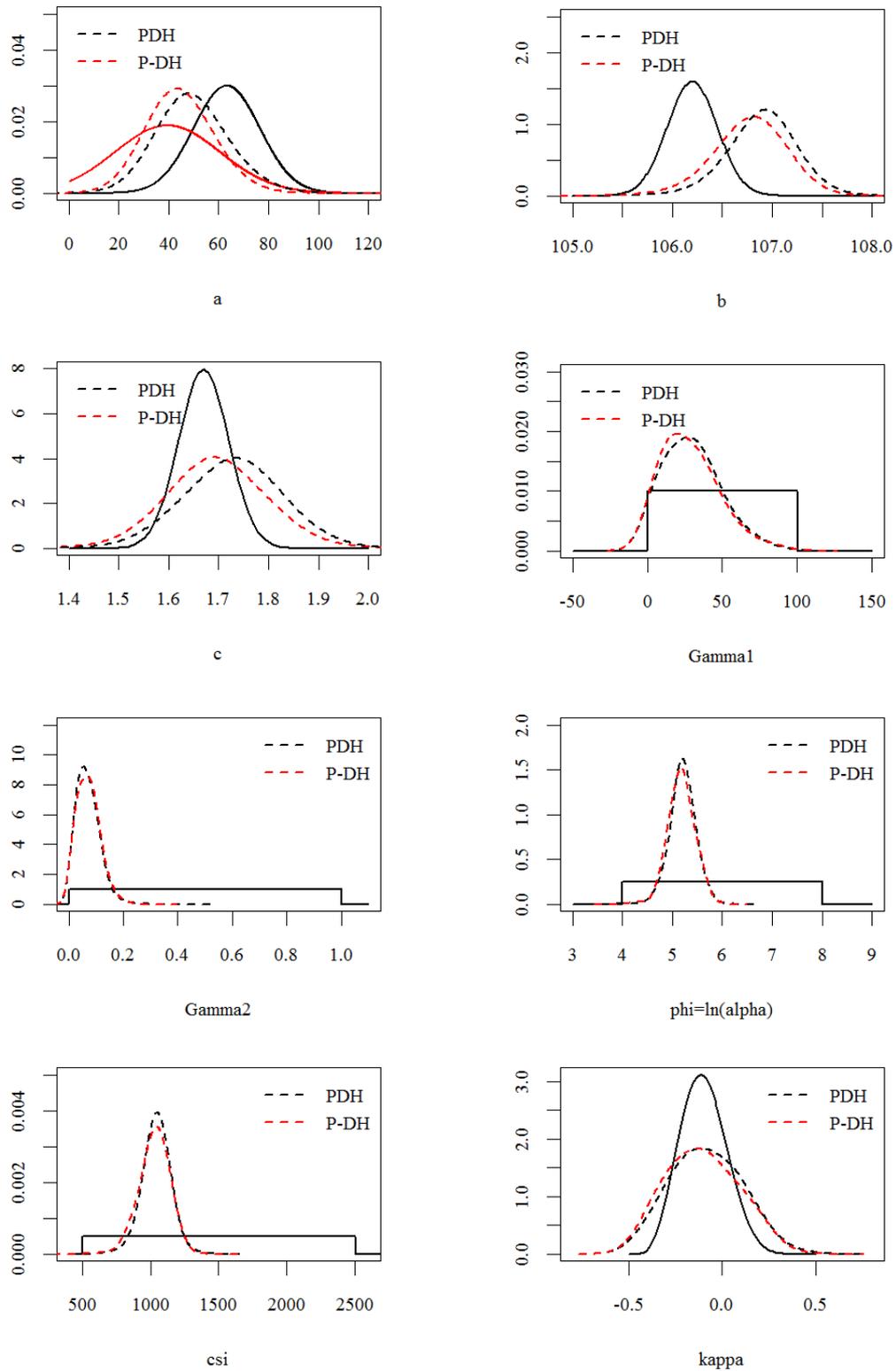


Figura 5-34 - Comparativo das distribuições *a priori* (linha cheia) e *a posteriori* (linha tracejada) dos cenários PDH (preto) e P-DH (vermelho, preto quando *a priori* for a mesma) com indicação do intervalo HPD nas distribuições dos quantis (linhas verticais) – Caso Real

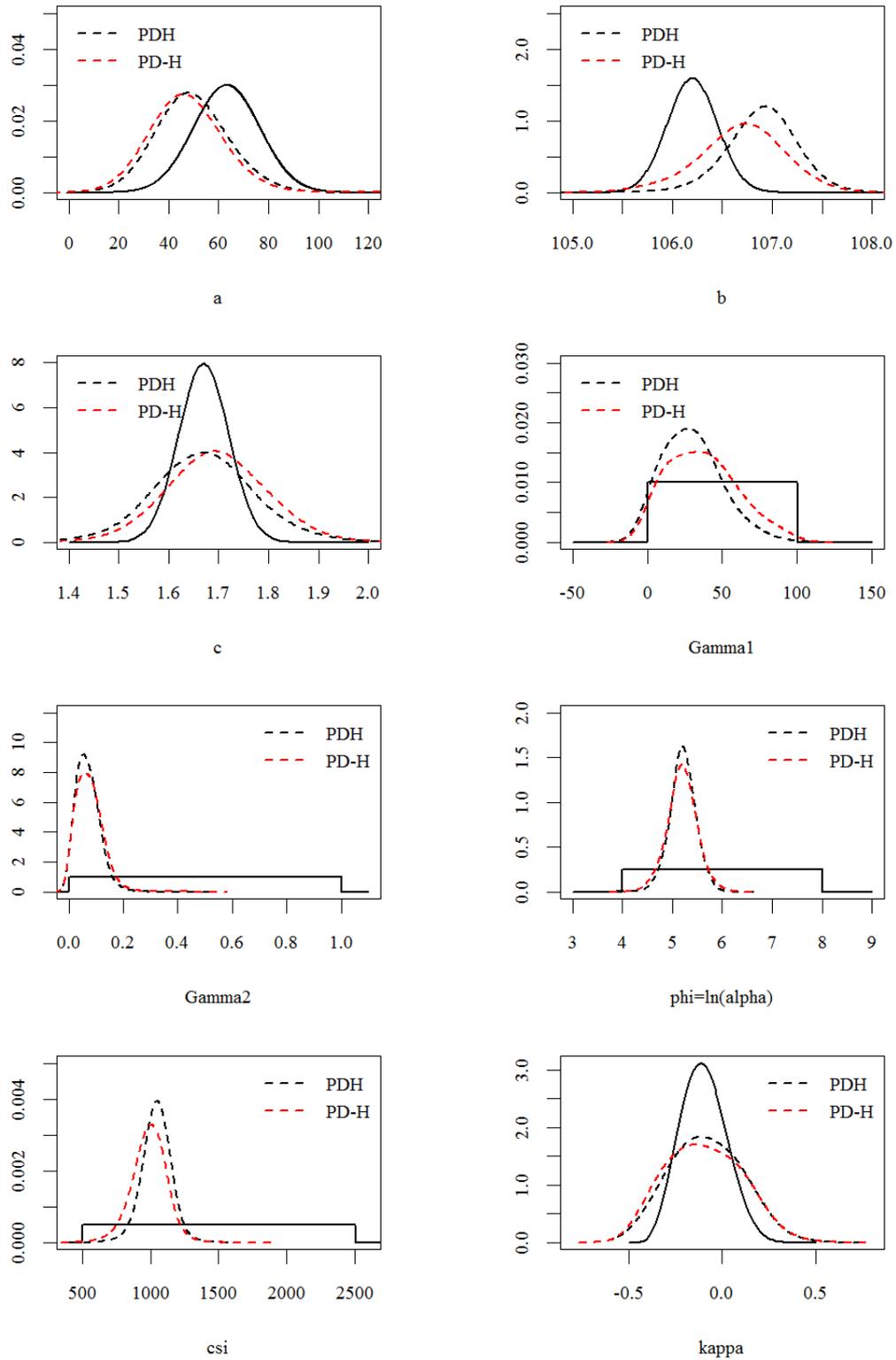


Figura 5-35 Comparativo das distribuições *a priori* (linhas cheias) e *a posteriori* (linhas tracejadas) dos cenários PDH (preto) e PD-H (vermelho) com indicação do intervalo HPD nas distribuições dos quantis (linhas verticais) – Caso Real

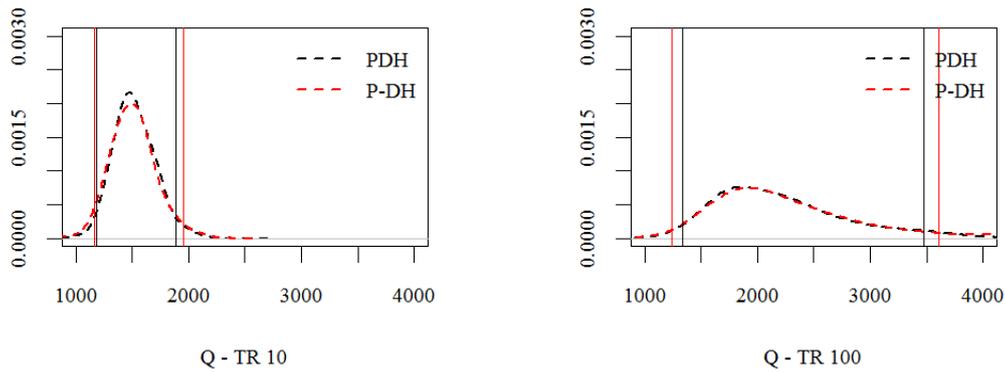


Figura 5-36 - Comparativo das distribuições *a posteriori* dos quantis de tempo de retorno de 10 e 100 anos para os cenários PDH e P-DH com a indicação dos respectivos intervalos de credibilidade (linhas cheias verticais) – Caso Real

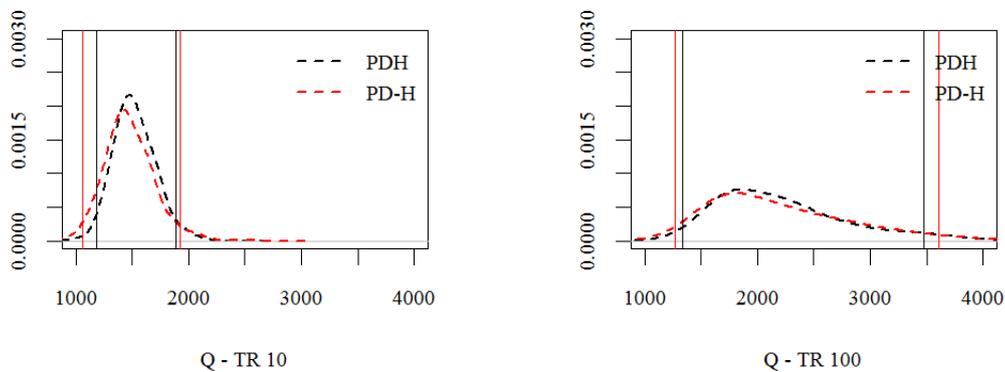


Figura 5-37 - Comparativo das distribuições *a posteriori* dos quantis de tempo de retorno de 10 e 100 anos para os cenários PDH e PD-H com a indicação dos respectivos intervalos de credibilidade (linhas cheias verticais) – Caso Real

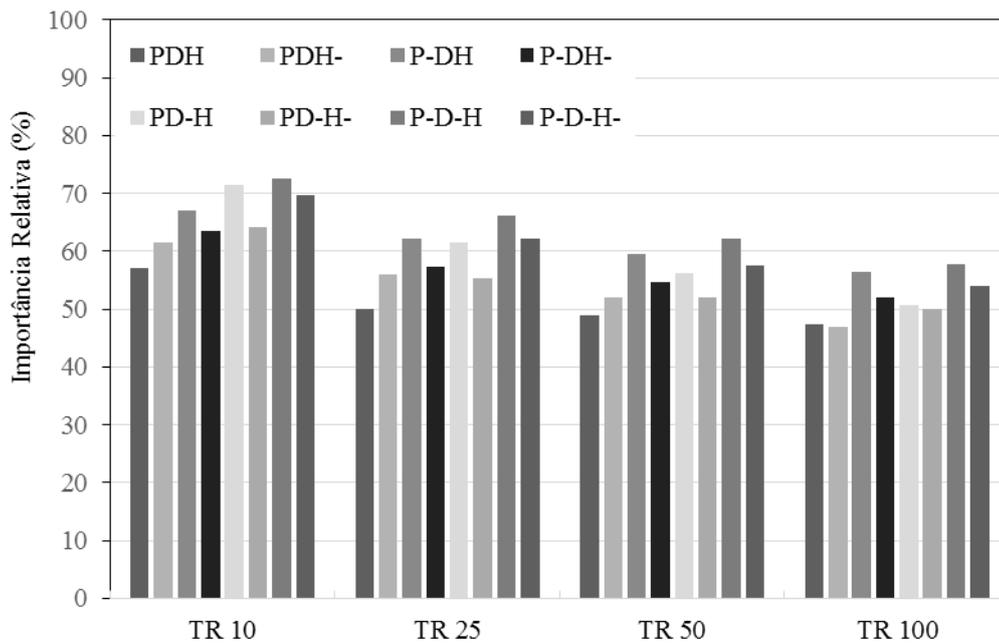


Figura 5-38 Comparativo da importância relativa para variações de H para H- (Caso Real)

6 CONCLUSÕES E RECOMENDAÇÕES

Este trabalho propõe um modelo Bayesiano completo para modelagem e incorporação das incertezas associadas à estimativa dos parâmetros da curva-chave nos estudos de frequência de cheias, permitindo empregar, quando disponível, informações sobre as características hidráulicas do trecho do rio em questão. O modelo proposto, além de proporcionar uma estimativa simultânea dos parâmetros da curva-chave e da distribuição teórica de probabilidades, que resultam na estimativa dos quantis de cheia e suas respectivas incertezas, possibilita também empregar explicitamente o conhecimento, ainda que incerto, do comportamento hidráulico do rio, de forma a incorporar informação relevante à análise na tentativa de reduzir as incertezas envolvidas no processo.

Além do desenvolvimento desse modelo Bayesiano para a análise de frequência de cheias (Objetivo específico 1), buscou-se entender, por meio da aplicação do mesmo, a contribuição das incertezas da curva-chave na incerteza global dos quantis de cheia (Objetivo específico 2), denominada aqui de importância relativa (IR), e analisar como o nível de conhecimento sobre as características hidráulicas do trecho do rio, a quantidade de medições de descarga e o tamanho da série histórica de cotas máximas anuais afetam a IR das incertezas da curva-chave nas incertezas totais dos quantis estimados (Objetivo específico 3).

As estimativas das séries de vazões máximas anuais, obtidas pelo modelo proposto nos dois casos estudados, demonstraram que as incertezas nessas séries, normalmente consideradas como sendo perfeitas em estudos de engenharia, podem ser elevadas, com intervalos de credibilidade (95%) com largura de aproximadamente 50% do valor médio estimado. Esse intervalo descreve as incertezas contidas nas séries de vazão, oriundas das imperfeições no ajuste da curva-chave e dos erros de medição de descarga. O grau de incerteza nas séries de vazões máximas anuais obtido nos dois casos indica, portanto, a necessidade de se considerar essas incertezas em estudos dessa natureza, sob pena de subestimar as incertezas nos quantis de cheia estimados.

Em ambos os casos tratados nesta dissertação, a IR das incertezas oriundas do ajuste da curva-chave em relação às incertezas totais nas estimativas dos quantis de cheia variou entre 40 e 70%, dependendo das características dos cenários avaliados. Esses valores são considerados altos, apesar da possibilidade de ter havido uma superestimativa desses valores, devido ao fato do modelo desenvolvido empregar apenas um segmento de curva-chave, o que empobrece o ajuste da curva aos dados, resultando num aumento das incertezas. Mas

ainda assim, entende-se que os resultados obtidos aqui mostram que a propagação dos erros da curva-chave nas séries de vazões máximas anuais pode impactar de forma significativa a precisão das estimativas dos quantis.

O estudo de diferentes cenários com diferentes graus de disponibilidade de dados de medição de descarga e nível de conhecimento *a priori* sobre as características hidráulicas do trecho do rio permitiram avaliar a importância dos dados na redução das incertezas dos quantis. Os cenários criados variavam o nível de informação *a priori* (P e P), disponibilidade de dados de resumo de descarga (D e D-) e disponibilidade de dados de cotas máximas anuais (H e H-). O que se pôde observar foi que a importância relativa calculada foi considerável, principalmente para a estação de Cáceres que chegou a mais de 70% nos casos com menos informação (P- e D-). Além disso, verificou-se o aumento da importância relativa com o aumento das incertezas (P para P-, e D para D-), tendo-se observado também para os casos específicos que a redução da disponibilidade de dados se mostrou mais importante que a redução do nível de informação *a priori*. Além disso, demonstrou-se que a redução da disponibilidade de dados de cotas máximas anuais resulta em uma redução da importância relativa, tendo em vista que a incerteza amostral aumenta com a redução do tamanho da amostra para tempos de retorno maiores.

Por restrição de tempo, o modelo Bayesiano completo desenvolvido nesta dissertação se limitou a curvas-chave de apenas um segmento. Do ponto de vista prático, essa é uma limitação importante porque muitas estações fluviométricas apresentam variações geométricas importantes na seção transversal à medida que o nível d'água aumenta, caracterizando a necessidade de múltiplos segmentos de curva para um bom ajuste da relação cota-descarga. A primeira sugestão para trabalhos futuros é estender o modelo desenvolvido aqui para propiciar o ajuste a vários segmentos de curva-chave, proporcionando uma maior aplicabilidade da metodologia desenvolvida nesta dissertação. Ainda em relação ao ajuste da curva-chave, é necessário investigar outras possibilidades de parametrização da equação matemática empregada. A que foi utilizada aqui na dissertação é a mais popular na engenharia, porém ela é adequada apenas em regime de canal, e somente quando a largura da seção transversal é muito superior à sua profundidade. Em situações onde essa aproximação não é adequada, a equação tem dificuldade em representar a verdadeira relação entre a cota do nível d'água e a vazão que escoar pelo rio. Sugere-se então que outras formas de parametrização da equação, publicadas na literatura científica na área de hidráulica, sejam avaliadas na tentativa de encontrar um equacionamento mais apropriado aos diferentes tipos de situação encontrados na natureza.

As conclusões obtidas neste trabalho em relação ao papel das incertezas da curva-chave na análise de frequências de cheia devem ser encaradas apenas como um primeiro passo num longo caminho a ser percorrido. Por exemplo, o caso sintético empregado neste estudo utilizou um valor relativamente conservador para o parâmetro de forma κ da distribuição GEV, $\kappa = -0.10$. Sabe-se que valores mais negativos de κ resultam não apenas em quantis de cheia mais extremos, mas também com maiores incertezas amostrais. Isso significa que a IR obtida no caso sintético não deve ser extrapolada para outras situações, já que deve variar com o valor de κ empregado. Sendo assim, sugere-se para trabalhos futuros, avaliar como a IR das incertezas da curva-chave em relação às incertezas globais dos quantis de cheia varia à medida que o parâmetro κ se torna mais negativo. Essa avaliação proporcionaria uma visão mais completa do papel das incertezas da curva-chave na análise de frequência de cheias.

Além disso, estudos realizados no passado mostram que empregar uma distribuição *a priori* mais informativa para o parâmetro κ pode reduzir significativamente as incertezas amostrais nos quantis de cheia estimados. Essa distribuições *a priori* são obtidas com base em alguma técnica de regionalização hidrológica, geralmente baseada em estudos de regressão. Portanto, seria importante também avaliar como a IR tratada nesta dissertação iria variar em situações onde a análise de frequência empregasse informação regional na estimativa de κ , e como essa informação regional seria obtida levando em consideração as incertezas na curva-chave em cada uma das estações utilizadas no estudo de regressão.

REFERÊNCIAS BIBLIOGRÁFICAS

- Beven, K., Hall, J. (2014). *Applied Uncertainty Analysis for Flood Risk Management*. London, Imperial College Press, 500p.
- Boiten, W. (2003). *Hydrometry*. Swets & Zeitlinger B. V., Lisse, Holanda, 248p.
- Chow, V. T. (1959). *Open-channel hydraulics*. Nova Iorque, McGraw-Hill, 680 p.
- Chow, V. T., Maidment, D. R., Mays, L. W. (1988). *Applied Hydrology*. New York, McGraw-Hill, 572 p.
- Clarke, R. T. (1999). Uncertainty in the estimation of mean annual flood due to rating-curve indefiniton. *Journal of Hydrology*, 222(1-4), 185–190.
- Clarke, R. T., Mendiondo, E. M., Brusa, L. C. (2000). Uncertainties in mean discharges from two large South American rivers due to rating curve variability. *Hydrological Sciences Journal*, 45(2), 221–236.
- Coles, S. G., Tawn, J. A. (1996). Modelling Extremes of the areal rainfall process. *Journal Royal Statistical Society, B*, 58.
- Coles, S. (2001). *An Introduction to Statistical Modeling of Extreme Values*. Londres, Springer-Verlag, 208 p.
- Coxon, G. Freer, J. Westerberg, K. Wagener, T. Woods, R. Smith, P. J. (2015). A novel framework for discharge uncertainty quantification applied to 500 UK gauging stations. *Water Resources, Res.* 51, 5331-5546.
- Despax, A., Perre, C., Garçon, R., Hauet, A., Belleville, A., Le Coz, J., Favre, A. C. (2016). Considering Sampling strategy and cross section complexity for estimating the uncertainty of discharge measurements using the velocity-area method. *Journal of Hydrology*, 533, pg. 128-140.
- Di Baldassarre, G., Montanari, A. (2009). Uncertainty in river discharge observations: a quantitative analysis. *Hydrology and Earth System Sciences*, 13(6), 913–921.
- Di Baldassarre, G., Claps, P. (2011). A hydraulic study on the applicability of flood rating curves. *Hydrology Research*, 42(1), 10-19.
- Di Baldassarre, G., Laio, F., Montanari, A. (2012). Effect of observation errors on the uncertainty of design floods. *Physics and Chemistry of the Earth*. 42-44, 82-90.
- Domeneghetti, A., Castellarin, A., Brath, A. (2012). Assessing rating-curve uncertainty and its effects on hydraulic calibration. *Hydrology and Earth System Sciences*, 26, 1191-1202.

- Dymond, J. R., Christian, R. (1982). Accuracy of discharge determined from a rating-curve. *Hydrological Sciences Journal*, 27(4), 493-504.
- Gelman, A., Rubin, D. (1992). Inference from Iterative Simulation Using Multiple Sequences. *Statistical Science*, 7, 457-511.
- Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., Rubin, D. B. (2014). *Bayesian Data Analysis*. Nova Iorque, Taylor & Francis, 668 p.
- Hersch, R. (1999). *Hydrometry. Principles and Practices*. Australia, Wiley, 511 p.
- Haario, H., Eero, S., Tamminen, J. (2001). An adaptive Metropolis algorithm. *Bernoulli*, 7, 223-242.
- Jacon, G., Cudo, K. J. (1989). *Curva-chave: Análise e Traçado*. DNAEE, Brasília, 273 p.
- Kuczera, G. (1996). Correlated rating curve error in flood frequency inference. *Water Resources Research*, 32(7), 2119–2127.
- Kuczera, G. (1999). Comprehensive at-site flood frequency analysis using Monte Carlo Bayesian Inference, *Water Resources Research*, 35(5), 1551-1557.
- Jalbert, J., Mathevet, T., Favre, A. (2011). Temporal Uncertainty estimation of discharges from rating curves using a variographic analysis. *Journal of Hydrology*. 397 (1-2), 83-92.
- Lang, M., Pobanz, K., Renard, B., Renouf, E., Sauquet, E. (2010). Extrapolation of rating curves by hydraulic modelling, with application to flood frequency analysis. *Hydrological Sciences Journal*, 55(6), 883–898.
- Le Coz, J. (2012). A literature review of methods for estimating the uncertainty associated with stage-discharge relations. Cemagref, *Hydrology-Hydraulics WMO*, 1-21.
- Le Coz, J., Renard, B., Bonnifait, L., Branger, F., Le Boursicaud, R. (2014). Combining hydraulic knowledge and uncertain gaugings in the estimation of hydrometric rating curves: A bayesian approach. *Journal of Hydrology*, 509, 573-587.
- Le Coz, J., Renard, B., Bonnifait, L., Branger, F., Le Boursicaud, R. (2014a). *Guide Pratique de la méthode Baratin pour l'analyse des courbes de tarage et de leurs incertitudes*. IRSTEA, Unidade de Pesquisa em Hidrologia e Hidráulica de Lyon.
- Mcmillan, H., Freer, J., Pappenberger, F., Krueger, T., Clark, M. (2010). Impacts of uncertain river flow data on rainfall-runoff model calibration and discharge prediction. *Hydrological Processes*, 24 (10), 1270-1284.
- Mansanarez, V., Le Coz, J., Renard, B., Lang, M., Vauchel, P. (2016). Bayesian Analysis of stage fall discharge rating curves and their uncertainties. *Water Resources Research*, 52, doi:10.1002/2016WR018916

- Marshall, L. (2005). Bayesian analysis of rainfall-runoff models: Insights to parameter estimation, model comparison and hierarchical model development., Tese de Doutorado, Universidade de New South Wales, Australia.
- Martinez, W. L., Martinez, A. R. (2002). Computational Statistics Handbook with MATLAB. Chapman & Hall/CRC.
- Martins, E. S., Stedinger, J. R. (2000). Generalized Maximum Likelihood generalized extreme-value quantile estimators for hydrologic data. *Water Resources Research*, 36(3), 737-744.
- Martins, E.S., Reis, D. S., Alexanre, A. M. (2008). Uso de Informação Regional na estimativa de quantis de cheia GEV. *RBRH*, Volume 13 n. 2, 127-139
- Mason, R., Kiang, J. Cohn, T. (2016) Rating curve uncertainty: A comparison of estimation Methods. *River Flow*, 533, 128-140.
- Moyeed, R. A., Clarke, R. T. (2005). The use of Bayesian methods for fitting rating curves, with case studies. *Advances in Water Resources*, 28, 807-818.
- Montanari, A. Di Baldassarre, G. (2013). Data errors and hydrological modelling: The role of model structure to propagate observation uncertainty. *Advances in Water Resources*, 51, 498-504.
- Morlot, T. Perret, C. Favre, A. Jalbert, J. (2014). Dynamic Rating Curve Assessmente for hydrometric stations and computation of the associated uncertainties: quality and station management. *Journal of Hydrology*, 517, 173-186.
- Naghetini, M., Pinto, E. J. A. (2007). *Hidrologia Estatística*. Belo Horizonte, Serviços Geológico do Brasil – CPRM, 552 p.
- Neppel, L., Renard, B., Lang, M., Ayrat, P. A., Coeur, D., Gaume, E., Jacob, N., Pobanz, K., Vinet, F. (2010). Flood Frequency analysis using historical data: accounting for random and systematic errors. *Hydrological Sciences Journal*, 55 (2), 192-208.
- O'Connell, D. R. H. (2005). Nonparametric flood frequency estimation. *Journal of Hydrology*, 313 (2005), 79-96.
- Pappenberger, F., Matgen, P., Beven, K. J., Henry, J. B., Pfister, L., Fraipont, P. (2006). Influence of uncertain boundary conditions and model structure on flood inundation predictions. *Advanced Water Resources*, 29, 1430-1449.
- Petersen-Overleir, A. (2004). Accounting for heteroscedasticity in rating curve estimates. *Journal of Hydrology*, 292 (1-4), 173-181.

- Petersen-Overleir, A. (2006). Modelling stage-discharge relationships affected by hysteresis using the Jones formula and nonlinear regression. *Hydrological Sciences Journal*, 51 (3), 365-388.
- Petersen-Overleir, A. (2008). Fitting depth-discharge relationships in rivers with floodplains. *Hydrology Research*, 39 (5-6), 369-384.
- Petersen-Overleir, A., Reitan, T. (2005). Objective segmentation in compound rating curves. *Journal of Hydrology*, 311 (1-4), 188-201.
- Petersen-Overleir, A., Reitan, T. (2009). Bayesian analysis of stage-fall-discharge models for gauging stations affected by variable backwater. *Hydrological Processes*, 23 (21), 3057-3074.
- Petersen-Overleir, A., Reitan, T. (2009a). Accounting for rating curve imprecision in flood frequency analysis using likelihood-based methods. *Journal of Hydrology*, 366 (2009), 89-100.
- Rampinelli, C. G. (2016) Modelagem hidrológica sob uma abordagem Bayesiana: comparação de algoritmos MCMC e análise da influência da função verossimilhança na estimativa dos parâmetros e descrição das incertezas. Dissertação de Mestrado, Universidade de Brasília, Brasil.
- Reis, D. S., Stedinger, J. R. (2005). Bayesian MCMC flood frequency analysis with historical information. *Journal of Hydrology*, 313, 97-116.
- Reitan, T., Petersen-Overleir, A. (2008). Bayesian power-law regression with a location parameter, with applications for construction of discharge rating curves. *Stochastic Environmental Research and Risk Assessment*, 22 (3), 351-365.
- Reitan, T., Petersen-Overleir, A. (2009). Bayesian methods for estimating multi-segment discharge rating curves. *Stochastic Environmental Research and Risk Assessment*, 23 (5), 627-642.
- Schmidt, A. R. (2002). Analysis of stage-discharge relations for open-channel flows and their associated uncertainties. Tese de Doutorado, Universidade de Illinois, Urbana-Champaign.
- Sefione, A. L. (2002). Estudo Comparativo de Métodos de Extrapolação superior de Curvas-chaves. Dissertação de Mestrado. Universidade Federal do Rio Grande do Sul, Instituto de Pesquisas Hidráulicas.
- Singh, V. P., Jain S. K., Tyagi, A. (2007). Risk and reliability analysis: a handbook for civil and environmental engineers. ASCE

- Stedinger, J.R., Vogel, R. M., Foufoula-Georgiou, E. (1993). Frequency analysis of extreme events. Handbook of Hydrology, New York, McGraw-Hill Book Co.
- Steinbakk, G. H., Thorarinsdottir, T. L., Reitan, T., Schlichting, L., Holleland, S., Engeland, K. (2016). Propagation of rating curve uncertainty in design flood estimation. Water Resources Research, 52, doi:10.1002/2015WR018516
- Tucci, C. E. M., Silveira, A. L. L. (1993). Hidrologia, Ciência e Aplicação. Porto Alegre, UFRGS/ABRH, 943p.
- Tucci, C. E., Silveira, R. L. (1985). Análise de Consistência de Dados Fluviométricos. Porto Alegre: Instituto de Pesquisas Hidráulicas, UFRGS, paginação irregular.
- UFPR/ITTI. (2015) Estudo de Viabilidade Técnica, Econômica e Ambiental da Hidrovia do Paraguai.
- Venetis, C. (1970). A note on the estimation of the parameters in logarithmic stage-discharge relationships with estimate of their error. Bulletin of the International Association of Scientific Hydrology, XV (2), 105-111.

APÊNDICE I – ROTINAS DOS MODELOS BC E BG

Código – Bayesiano Completo - BC

```
##### CÓDIGO BAYESIANO COMPLETO
```

```
#####VERSÃO: DEZEMBRO DE 2016
```

```
##### DESCRIÇÃO: CÓDIGO PARA ESTIMATIVA SIMULTÂNEA DOS  
PARÂMETROS DA CURVA CHAVE E DA  
##### GEV POR INFERÊNCIA BAYESIANA PARA CÁLCULO DA INCERTEZA  
NOS QUANTIS.
```

```
##### VERSÃO PARAMETRIZADA
```

```
#####
```

```
##### ALGORITMO ADAPTIVE METROPOLIS
```

```
#####
```

```
rm(list=ls())
```

```
library(pracma)
```

```
#####
```

```
##PARTE INICIAL: DEFINE AS CONFIGURACOES GERAIS
```

```
args = commandArgs(trailingOnly=TRUE)
```

```
diretoribase <- args[1]
```

```
codigo <- args[2]
```

```
fatorP <- args[3]
```

```
fatorRD <- args[4]
```

```
fatorH <- args[5]
```

```
cenario1 <- strcat(c(args[3], args[4], args[5]))
```

```
simulacao <- args[6]
```

```
estudo <- args[7]
```

```
pastadosdeentrada <- strcat(c(diretoribase,"Dados de Entrada"))
```

```
pastadosdesaida <-strcat(c(diretoribase,cenario1,"/",codigo))
```

```
arquivodeentradaRD <- strcat(c(fatorRD,".csv"))
```

```
arquivodentradaHmax <- strcat(c(fatorH,".csv"))
```

```
arquivodentradaPrioris <- strcat(c(fatorP,".csv"))
```

```
arquivoconfig <- strcat(c('config_',simulacao,'.csv'))
```

```
# Nome do Cenario para salvar no arquivo
```

```
cenario = strcat(c(codigo,estudo,cenario1,'_',simulacao))
```

```
#####DEFiNE AS CONFIGURACOES DAS SIMULACOES
```

```
### Erro de Medicao
```

```
erromed <- 0.05
```

```
### Prioris
```

```
print(pastadosdeentrada)
```

```

setwd(pastadadosdeentrada)
prioris <- read.csv(arquivodentradaPrioris,head=FALSE)

### Ponto de Partida e número de iterações
setwd(pastadadosdeentrada)
config <- read.csv(arquivoconfig,head=FALSE)

startvalue <-
c(config[1,1],config[1,2],config[1,3],config[1,4],config[1,5],config[1,6],config[1,7],config
[1,8])
nIter=config[2,1]
iterations <- nIter

#####
##PARTE 1: DADOS INICIAIS
#####
# vai para pasta de dados de entrada
setwd(pastadadosdeentrada)
#####
##### ALGORITMO ADAPTIVE METROPOLIS
#####

#####
##PARTE 2: DADOS,FUNÇÕES PRIORIS E VEROSSIMILHANÇA
#####
#Dados
#Carrega os dados. Deve ser um arquivo .csv em que a primeira coluna representa os
valores de H observados e a segunda os valores de Q
# DADOS RIO PARAGUAI
dados<-read.csv(arquivodeentradaRD,sep=" ",head=FALSE)
dadoscotasmximas<-read.csv(arquivodentradaHmax,sep=" ",head=FALSE)

#Passo 3: aloca os dados carregados nos vetores coluna H e Q e calcula o desvio padrão do
erro de medição (uqi) como sendo 5% do valor da vazão
##### DADOS #####
H<-dados[,1]
Q<-dados[,2]
#Desvio padrao do erro de 5%, pode alterar
Erro.med<-erromed*Q
hmax<-dadoscotasmximas

##Função Posterior
##Ln do produto da verossimilhança pela priori

posterior <- function(parametros,prioris,Q,H,Erro.med,hmax){

```

```

a<-parametros[1]
b<-parametros[2]
c<-parametros[3]
gamma1 <- parametros[4]
gamma2<-parametros[5]
## parametros do estudo de frequencia de cheias
### alteracao : introduz phi = log(alpha)
#alpha<-parametros[6]
phi <- parametros[6]
qui<-parametros[7]
kapa<-parametros[8]
##Qtde de dados de vazão
nflows <-nrow(hmax)

#Verossimilhança 1 -L1
bref <- min(H)

##o erro é proporcional ao valor estimado e não observado, corrija esse ponto
if (
(a>0.01)&&(b>0.01)&&(b<bref)&&(c>0.01)&&(gamma1>0.01)&&(gamma2>0.01)) {
  Fh<-a*(H-b)^c
  sigf<-gamma1 + gamma2*Fh
  DP<-((sigf)^2 + Erro.med^2)^.5
  Fhref <- min(Fh)
  DPref <- min(DP)
  teste1 <- Fhref - DPref
  verif <- 0
  if (teste1<=0) {
    verif <- 1
  }
  }else{
  verif<- 1
  }
}

if (verif==0) {
  L1 <- sum(dnorm(Q,Fh,DP,log=TRUE))
}else{
  L1 <- -Inf
}
##Verossimilhança 2 - L2

if (b<bref){
  Qmaxsemerro<-a*(hmax-b)^c
  sigmafqmax<- gamma1+gamma2*Qmaxsemerro
}else{
  Qmaxsemerro <- a*(hmax-b)#para gerar valores negativos e dar indicio de erro
  sigmafqmax <- Qmaxsemerro
}

```

```

}

vazaoref <- min(Qmaxsemerro)
sigmafqmaxref <- min(sigmafqmax)
teste <- vazaoref - sigmafqmaxref
#if (teste <= 0) {
# indicadordefalha <-1
#}
#corrigindo o erro de geraçao de NaN

if (min(sigmafqmax) < 0||teste<=0){
  indicadordefalha <- 1
} else {
  indicadordefalha<-0
  erro <- rnorm(nrow(hmax),0,sigmafqmax)
  Qmaxcomerro<-Qmaxsemerro + erro
}

## usando as formulações de martins e stedinger (2000)

if (indicadordefalha == 1)
{
  L2<- -Inf
}

if (indicadordefalha == 0)
{
  verifGEV = qui + exp(phi)/kapa
  minimo = min(Qmaxcomerro)
  maximo = max(Qmaxcomerro)

  if (kapa>0) {

    if (verifGEV > maximo) #a distribuição existe nesse intervalo
    {
      y <- (1-kapa/exp(phi))*(Qmaxcomerro-qui)
      term2<- ((1/kapa -1)*log(y)-(y)^(1/kapa))
      L2 <- -nflows*phi+sum(term2)
    } else{
      L2 <- -Inf
    }
  }

  }else if (kapa>0){

    if (verifGEV<minimo) {

```

```

y <- (1-kapa/exp(phi))*(Qmaxcomerro-qui)
term2<- ((1/kapa -1)*log(y)-(y)^(1/kapa))
L2 <- -nflows*phi+sum(term2)

} else {
  L2 <- -Inf
}

} else {
  term2 <- -(Qmaxcomerro-qui)/exp(phi) - exp(-(Qmaxcomerro-qui)/exp(phi))
  L2 <- -nflows*phi +sum(term2)
}

}

#Prioris

### priori a (normal)
mua<-prioris[1,1]
siga<-prioris[1,2]
priora <- dnorm(a,mua,siga,log=TRUE)

### priori b (normal)
mub<-prioris[2,1]
sigb<-prioris[2,2]
priorib <- dnorm(b,mub,sigb,log=TRUE)

### priori c (normal)
muc<-prioris[3,1]
sigc<-prioris[3,2]
prioric = dnorm(c,muc,sigc,log=TRUE)

### priori gamma1 (uniforme)

priorigamma1 <- dunif(gamma1, min = prioris[4,1], max = prioris[4,2], log = TRUE)

### priori gamma2 (uniforme 0-10.000)
priorigamma2 <- dunif(gamma2, min = prioris[5,1], max = prioris[5,2], log = TRUE)

#alterei a priori de alpha para uniforme de 0 a 1000
###priori de alpha (uniforme de 0 a 1000)
prioriphi <- dunif(phi,min=prioris[6,1],max=prioris[6,2],log=TRUE)

###priori de qui (uniforme de -10.000 a 10.000)
prioriqui <- dunif(qui,min=prioris[7,1],max=prioris[7,2],log=TRUE)

###priori de kapa beta(media = -0,10, var = 0,122^2, p=6, e q=9)
if (kapa>0)

```

```

kapamod <- kapa
if (kapa<0)
  kapamod<- -kapa
p=6
q=9
if (kapamod<=0.5) {
  bpq <- gamma(p)*gamma(q)/gamma(p+q)
  logdistbeta <- log(((0.5+kapa)^(p-1))*((0.5-kapa)^(q-1))/bpq)
  priorikapa <- logdistbeta
else{
  priorikapa <- -Inf
}
}

#Posteriori

if ((L1 == "NaN")||(L2 == "NaN")||(L1==-Inf)||(L2==-Inf)) {
  posterior<- -Inf
} else {
  posterior<- L1 + L2+
priora+priorib+prioric+priorigamma1+priorigamma2+prioriphi+priorikapa+prioriqui
}
}

```

#5.3- Algoritmo Adaptive - Metropolis
#####

```

run_AdaptiveMetropolis_MCMC <-
function(startvalue,prioris,iterations,H=H,Q=Q,Erro.med=Erro.med,hmax=hmax){

```

Define a cadeia formada por colunas correspondentes ao número de parametros e linhas
#correspondente ao número de iterações + 1 (chute inicial)

```
chain = array(dim = c(iterations+1,8))
```

#1a linha da cadeia, valor do chute inicial

```
chain[1,] = startvalue
```

#Define matriz nula de diagonal correspondente às variâncias dos parâmetros a priori

```

siga<-prioris[1,2]
sigb<-prioris[2,2]
sigc<-prioris[3,2]
siggamma1 <-prioris[4,2]
siggamma2 <-prioris[5,2]
sigphi <- prioris[6,2]
sigcsi <- prioris[7,2]

```

```

vara = (siga^2)*0.1
varb = sigb^2
varc = sigc^2
vargamma1 = (siggamma1-0)^2/2
vargamma2 = (siggamma2-0)^2/2
varphi = (sigphi-0)^2/2*0.001
varqui = (sigcsi-(-sigcsi))^2/2*0.01
varkappa = 0.3^2/2*0.1

C=matrix(0,nrow=8,ncol=8)
#C<-C
diag(C)=c(vara,varb,varc,vargamma1,vargamma2,varphi,varqui,varkappa)

##ID

Id=diag(x=1,nrow=8,ncol=8)

sd=(2.4)^2/8

E=0.001
naceito=0
nrejeitado=0

#matriz da proposta normal multivariada (inicia ela nula com n° de parametros colunas)

p=matrix(0,nrow=iterations,ncol=8)
posteriori = matrix(0,nrow=iterations,ncol=1)
guardarazao = matrix(0,nrow=iterations,ncol=1)

#para as primeiras 600 amostras aplica as variancias iniciais e a média corresponde
#ao valor no passo anterior

for(i in 1:600)
{
  proposal=mvrnorm(1,chain[i,],C)

  print(c(i,proposal))

  p[i,]=proposal
  posteriori[i,] = posterior(proposal,prioris,H=H,Q=Q,Erro.med=Erro.med,hmax=hmax)

  #testa o valor da razao entre o valor proposto e o valor no passo anterior

  razao=posterior(proposal,prioris,H=H,Q=Q,Erro.med=Erro.med,hmax=hmax)-
posterior(chain[i,],prioris,H=H,Q=Q,Erro.med=Erro.med,hmax=hmax)
  guardarazao[i,]<-razao
}

```

```

if (exp(razao) == "NaN"){

  chain[i+1,]=chain[i,]
  nrejeitado=nrejeitado+1
}

else if(runif(1,0,1)<min(1,exp(razao))){

  chain[i+1,]=proposal
  naceito=naceito+1
}

else{

  chain[i+1,] = chain[i,]
  nrejeitado=nrejeitado+1

}
if (mod(i,1000)==0) {
  print(c(i,proposal,posteriori[i,]))
}

}
##A partir da iteração 1000 a matriz de variancias começa a ser atualizada

for (i in 600:iterations){

chainMed1=c(mean(chain[1:i,1]),mean(chain[1:i,2]),mean(chain[1:i,3]),mean(chain[1:i,4]
),mean(chain[1:i,5]),mean(chain[1:i,6]),mean(chain[1:i,7]),mean(chain[1:i,8]))
  chainMed2=c(mean(chain[1:(i-1),1]),mean(chain[1:(i-1),2]),mean(chain[1:(i-
1),3]),mean(chain[1:(i-1),4]),mean(chain[1:(i-1),5]),mean(chain[1:(i-
1),6]),mean(chain[1:(i-1),7]),mean(chain[1:(i-1),8]))
  C=((i-1)/i)*C+(sd/i)*(i*chainMed2%*%t(chainMed2)-
(i+1)*chainMed1%*%t(chainMed1)+chain[i,]%*%t(chain[i,])+E*Id)

  proposal=mvnorm(1,chain[i,],C)

  print(c(i,proposal))

  p[i,]=proposal

  posteriori[i,] = posterior(proposal,prioris,H=H,Q=Q,Erro.med=Erro.med,hmax=hmax)

  razao=posterior(proposal,prioris,H=H,Q=Q,Erro.med=Erro.med,hmax=hmax)-
posterior(chain[i,],prioris,H=H,Q=Q,Erro.med=Erro.med,hmax=hmax)
  guardarazao[i,]<-razao
}

```

```

if (exp(razao) == "NaN"){

  chain[i+1,]=chain[i,]
  nrejeitado=nrejeitado+1
}

else if(runif(1,0,1)<min(1,exp(razao))){

  chain[i+1,]=proposal
  naceito=naceito+1
}

else{

  chain[i+1,] = chain[i,]
  nrejeitado=nrejeitado+1

}

if (mod(i,1000)==0) {
  print(c(i,proposal,posteriori[i,]))
}

}

lista=list(chain,naceito,nrejeitado,p,C,posteriori,guardarazao)

return(lista)
}

#####
##INICIO SIMULACAO

chain=0

library(fExtremes)
library(MASS)
date()
resultadoAM<-
run_AdaptiveMetropolis_MCMC(startvalue,prioris,iterations,H=H,Q=Q,Erro.med=Erro.m
ed,hmax=hmax)
date()

#####
##PARTE 5: GUARDA OS RESULTADOS EM ARQUIVOS .CSV PARA PLOTAGEM
#####

#### seta a pasta para o lançamento dos resultados

```

```
setwd(pastadadosdesaida)
```

```
# a) Salva arquivo dos vetores das propostas da cadeia
```

```
propostas<-resultadoAM[[4]]
```

```
filepropostas <- strcat("propostas",cenario)
```

```
write.csv(propostas,file=filepropostas)
```

```
# c) Salva arquivo das cadeias de todos os parâmetros
```

```
chainAM=resultadoAM[[1]]
```

```
chainAM=matrix(unlist(chainAM),ncol=8,byrow=F)
```

```
filecadeia <- strcat("cadeia_cenario",cenario)
```

```
write.csv(chainAM,file=filecadeia)
```

```
alpha <- exp(chainAM[,6])
```

```
filealpha <- strcat("cadeiaalpha_cenario",cenario)
```

```
write.csv(alpha,file=filealpha)
```

```
# d) Guarda a Taxa de rejeição/aceitação
```

```
taxa <- resultadoAM[[2]]/resultadoAM[[3]]
```

```
aceitarejeita <- c(resultadoAM[[2]],resultadoAM[[3]],taxa)
```

```
fileaceitarejeita <-strcat("aceitarejeitacenario",cenario)
```

```
write.csv(aceitarejeita,file=fileaceitarejeita)
```

```
# e) Guarda o Valor da posteriori para cada vetor
```

```
posteriori <- resultadoAM[[6]]
```

```
filepost <- strcat("posterioricenario",cenario)
```

```
write.csv(posteriori, file = filepost)
```

```
# f) Guarda o resultado final de todos os parâmetros
```

```

burninfinal <- 50000
library(TeachingDemos)

hpda<-emp.hpd(chainAM[burninfinal:100000,1],conf=0.95)
maxposta = mean(chainAM[burninfinal:100000,1])
hpdb<-emp.hpd(chainAM[burninfinal:100000,2],conf=0.95)
maxpostb = mean(chainAM[burninfinal:100000,2])
hpdc<-emp.hpd(chainAM[burninfinal:100000,3],conf=0.95)
maxpostc = mean(chainAM[burninfinal:100000,3])
hpdgamma1<-emp.hpd(chainAM[burninfinal:100000,4],conf=0.95)
maxpostgama1 = mean(chainAM[burninfinal:100000,4])
hpdgamma2<-emp.hpd(chainAM[burninfinal:100000,5],conf=0.95)
maxpostgama2 = mean(chainAM[burninfinal:100000,5])
hpdalpha<-emp.hpd(chainAM[burninfinal:100000,6],conf=0.95)
maxpostalpha = mean(chainAM[burninfinal:100000,6])
hpdcsi<-emp.hpd(chainAM[burninfinal:100000,7],conf=0.95)
maxpostcsi = mean(chainAM[burninfinal:100000,7])
hpdkapa<-emp.hpd(chainAM[burninfinal:100000,8],conf=0.95)
maxpostkapa = mean(chainAM[burninfinal:100000,8])
resultssummary1 <-
rbind(hpda,hpdb,hpdc,hpdgamma1,hpdgamma2,hpdalpha,hpdcsi,hpdkapa)
resultssummary2<-
rbind(maxposta,maxpostb,maxpostc,maxpostgama1,maxpostgama2,maxpostalpha,maxpost
csi,maxpostkapa)
summary <- cbind(resultssummary1,resultssummary2)

filessummary <- strcat("resultssummary_cenario",cenario)
write.csv(summary,file=filessummary)

#####
##PARTE 6: CALCULA A DISTRIBUIÇÃO À POSTERIORI DOS QUANTIS
#####

#a) Calcula a distribuição à posteriori dos diferentes tempos de retorno
# Burn in definido com base na análise da cadeia

alphachain <-alpha[burninfinal:iterations]

kapachain <-chainAM[burninfinal:iterations,8]

csichain <-chainAM[burninfinal:iterations,7]

# b) Gera Matriz em que cada coluna é um TR e para cada TR calcula o valor da Vazão
máxima provável Xp

tempoderetorno <-
c(1.01,1.1,1.2,1.3,1.4,1.5,1.6,1.7,1.8,1.9,2,2.5,3,4,4.5,5,5.5,10,15,20,25,30,35,40,45,50,55,
60,75,80,85,90,100)

```

```

tempoderetorno1 <-matrix(tempoderetorno,nrow=33,ncol=1)

tempoderetorno<-matrix(tempoderetorno,nrow=1,ncol=33)

p = 1/tempoderetorno

n = 33

estfreq<-matrix(0,nrow=(iterations-burninfinal+1),ncol=n)

total <- iterations - burninfinal +1

for (i in 1:n){
  prob <- p[i]
  yprob <- -log(1-prob)

  for (k in 1:total) {

    if ((kapachain[k] < 0.01)&&(kapachain[k]> -0.01)) {
      estfreq[k,i]<- csichain[k] -alphachain[k]*(log(yprob))

    } else
      estfreq[k,i] <- csichain[k] + alphachain[k]/kapachain[k]*(1-(yprob)^kapachain[k])

  }
}

filefreq <- strcat("estfreq_cenario",cenario)
write.csv(estfreq,file=filefreq)

#####
##PARTE 7: CALCULA A CURVA DE FREQUÊNCIA DE CHEIAS + IC
#####

# a) Intervalo de confiança fornecido pelo usuário
IC<- 0.95

#primeira coluna o mais provável , segunda coluna a parte inferior, segunda coluna a parte
superior

library(TeachingDemos)

curvadefrequenciadecheias <- matrix(0,nrow=n,ncol=3)

for (i in 1:n){
  curvadefrequenciadecheias[i,] <- c(mean(estfreq[,i]),emp.hpd(estfreq[,i],conf=IC))
}

```

```

filecurva <- strcat("curvadefrequenciadecheias_cenario",cenario)

curvadefrequenciadecheiasfinal <- cbind(tempoderetorno1,curvadefrequenciadecheias)

write.csv(curvadefrequenciadecheiasfinal,file = filecurva)

#####
##PARTE 8: CURVA CHAVE PARA PLOTAR
#####
# a) Usuário fornece o valor mínimo, máximo e o intervalo dos valores de h

hminimo <- min(H)
hmaximo <- max(hmax)+0.5
intervalo <- 0.1

hparaplotar <- seq(hminimo,hmaximo,intervalo)

# b) Processo de geração das niter valores de vazao para cada valor de h

n = length(hparaplotar)

parametrosFB <- chainAM

resumodedescarga <- dados

burnin <- 50000

total <- length(parametrosFB[,1])

parametrosseburnin <- parametrosFB[burnin:total,]

n2 = length(parametrosseburnin[,1])

library(TeachingDemos)
### Gerar as infinitas curvas chaves
vetorzoesparaplotar <- matrix(0,nrow=n,ncol=3)
for (i in 1:n) {
  vetorzoes <- matrix(0,nrow=n2,ncol =1)
  for (k in 1:n2) {
    Q1 <- parametrosseburnin[k,1]*(hparaplotar[i]-
parametrosseburnin[k,2])^(parametrosseburnin[k,3])
    desvioerromodelo <- parametrosseburnin[k,4] + parametrosseburnin[k,5]*Q1
    erro = rnorm(1,0,desvioerromodelo)
    vetorzoes[k] <- Q1 + erro
  }
}

```

```

hpdvazao<-emp.hpd(vetorvazoes,conf=0.95)
maxpostvazao = mean(vetorvazoes)
vetorvazoesparaplotar[i,1] <- maxpostvazao
vetorvazoesparaplotar[i,2] <- hpdvazao[1]
vetorvazoesparaplotar[i,3] <- hpdvazao[2]
}

filehparaplotar <- strcat('hparaplotarfbcenario',cenario)
write.csv(hparaplotar,file=filehparaplotar)
filevetorvazoesparaplotar <- strcat('vazoescurvachaveparaplotarcenario',cenario)
vetorvazoesparaplotarfinal <- cbind(hparaplotar,vetorvazoesparaplotar)
write.csv(vetorvazoesparaplotarfinal,file = filevetorvazoesparaplotar)

##### MAX POST

cadeia<-chainAM
post <- posteriori

## Calcula o max post

n <- nrow(post)

for (i in 2:n) {
  if (i==2) {
    posttemp <-post[i-1,1]
    maxpost <- cadeia[i,]
  }
  else {
    if (post[i,1]>posttemp) {
      posttemp <- post[i-1,1]
      maxpost <- cadeia[i,]
    }
  }
}

filemaxpost <- strcat('maxpost',cenario)
write.csv(maxpost,file = filemaxpost)

```

Código – Bayesiano para GEV – BG

```
##### CÓDIGO BAYESIANO PARA ESTIMATIVA DOS PARÂMETROS PARA  
CURVA DE FREQUÊNCIA DE CHEIAS
```

```
library(pracma)
```

```
#####VERSÃO: NOVEMBRO DE 2016
```

```
##### DESCRIÇÃO: CÓDIGO PARA ESTIMATIVA DOS PARÂMETROS DA  
CURVA DE FREQUÊNCIA DE CHEIAS POR INFERÊNCIA BAYESIANA
```

```
rm(list=ls())
```

```
#####
```

```
#####
```

```
##PARTE INICIAL: DEFINE AS CONFIGURACOES GERAIS
```

```
args = commandArgs(trailingOnly=TRUE)
```

```
diretoribase <- args[1]
```

```
codigo <- args[2]
```

```
fatorP <- args[3]
```

```
fatorRD <- args[4]
```

```
fatorH <- args[5]
```

```
cenario1 <- strcat(c(args[3], args[4], args[5]))
```

```
simulacao <- args[6]
```

```
estudo <- args[7]
```

```
pastadosdeentrada <- strcat(c(diretoribase, "Dados de Entrada"))
```

```
pastadosdeentrada2 <- strcat(c(diretoribase, cenario1, "/BC"))
```

```
pastadosdesaida <- strcat(c(diretoribase, cenario1, "/", codigo))
```

```
arquivodeentradaRD <- strcat(c(fatorRD, ".csv"))
```

```
arquivodentradaHmax <- strcat(c(fatorH, ".csv"))
```

```
arquivodentradaPrioris <- strcat(c(fatorP, ".csv"))
```

```
arquivoconfig <- strcat(c('config_', simulacao, '.csv'))
```

```
### Ponto de Partida e número de iterações
```

```
setwd(pastadosdeentrada)
```

```
config <- read.csv(arquivoconfig, head=FALSE)
```

```
# Nome do Cenário para salvar no arquivo
```

```
cenario = strcat(c(codigo, estudo, cenario1, '_', simulacao))
```

```
cenarioref = strcat(c("BC", estudo, cenario1, '_', "00"))
```

```
startvalue <- c(config[1,6], config[1,7], config[1,8])
```

```
#nIter=config[2,1]/2
```

```
#iterations <- nIter
```

```

##### Dados de entrada

hmax<-read.csv(arquivodentradaHmax,sep="," ,head=FALSE)

##### Funções

posterior <- function(parametros,Qmax){

phi <- parametros[1]
qui<-parametros[2]
kapa<-parametros[3]
##Qtde de dados de vazão
nflows <-nrow(Qmax)

##Verossimilhança 2 - L2

verifGEV = qui + exp(phi)/kapa
minimo = min(Qmax)
maximo = max(Qmax)

if (kapa>0) {

if (verifGEV > maximo) #a distribuição existe nesse intervalo
{
y <- (1-kapa/exp(phi))*(Qmax-qui)
term2<- ((1/kapa -1)*log(y)-(y)^(1/kapa))
L2 <- -nflows*phi+sum(term2)
} else
L2 <- -Inf

}else if (kapa>0){

if (verifGEV<minimo) {

y <- (1-kapa/exp(phi))*(Qmax-qui)
term2<- ((1/kapa -1)*log(y)-(y)^(1/kapa))
L2 <- -nflows*phi+sum(term2)

} else
L2 <- -Inf

} else {
term2 <- -(Qmax-qui)/exp(phi) - exp(-(Qmax-qui)/exp(phi))
L2 <- -nflows*phi +sum(term2)
}
}

```

```

}

#Prioris

#alterei a priori de alpha para uniforme de 0 a 1000
###priori de alpha (uniforme de 0 a 1000)
prioriphi <- dunif(phi,min=1,max=10,log=TRUE)

###priori de qui (uniforme de -10.000 a 10.000)
prioriqui <- dunif(qui,min=-10000,max=10000,log=TRUE)

###priori de kapa beta(media = -0,10, var = 0,122^2, p=6, e q=9)
if (kapa>0)
  kapamod <- kapa
if (kapa<0)
  kapamod<- -kapa
p=6
q=9
if (kapamod<=0.5) {
  bpq <- gamma(p)*gamma(q)/gamma(p+q)
  logdistbeta <- log(((0.5+kapa)^(p-1))*((0.5-kapa)^(q-1))/bpq)
  priorikapa <- logdistbeta
}else
  priorikapa <- -Inf

#Posteriori

if ((L2 == "NaN")||(L2== -Inf)) {
  posterior<- -Inf
} else {
  posterior<- L2+prioriphi+priorikapa+prioriqui
}
}

#####
##PARTE 3: ALGORITMO ADAPTIVE METROPOLIS (NÃO DEVE SER
ALTERADO)
#####

run_AdaptiveMetropolis_MCMC_GEV <- function(startvalue, iterations,Qmax=Qmax){

# Define a cadeia formada por colunas correspondentes ao número de parametros e linhas
#correspondente ao número de iterações + 1 (chute inicial)

```

```

chain = array(dim = c(iterations+1,3))

#1a linha da cadeia, valor do chute inicial
chain[1,] = startvalue

#Define matriz nula de diagonal correspondente às variâncias dos parâmetros a priori

varphi = (7-0)^2/2*0.001
varqui = (10000-(-10000))^2/2*0.1
varkappa = 0.3^2/2*0.1

C=matrix(0,nrow=3,ncol=3)
#C<-C
diag(C)=c(varphi,varqui,varkappa)

##ID

Id=diag(x=1,nrow=3,ncol=3)

sd=(2.4)^2/3

E=0.001
naceito=0
nrejeitado=0

#matriz da proposta normal multivariada (inicia ela nula com nº de parametros colunas)

p=matrix(0,nrow=iterations,ncol=3)
posteriori = matrix(0,nrow=iterations,ncol=1)
guardarazao = matrix(0,nrow=iterations,ncol=1)

#para as primeiras 600 amostras aplica as variancias iniciais e a média corresponde
#ao valor no passo anterior

for(i in 1:600)
{

proposal=mvrnorm(1,chain[i,],C)
p[i,]=proposal
posteriori[i,] = posterior(proposal,Qmax=Qmax)
print(i)
print(proposal)

#testa o valor da razao entre o valor proposto e o valor no passo anterior

razao=posterior(proposal,Qmax=Qmax)-posterior(chain[i,],Qmax=Qmax)
guardarazao[i,]<-razao

if (exp(razao) == "NaN"){

```

```

    chain[i+1,]=chain[i,]
    nrejeitado=nrejeitado+1
}

else

    if(runif(1,0,1)<min(1,exp(razao))){

        chain[i+1,]=proposal
        naceito=naceito+1
    }

else{

    chain[i+1,] = chain[i,]
    nrejeitado=nrejeitado+1

}

}

##A partir da iteração 1000 a matriz de variancias começa a ser atualizada

for (i in 600:iterations){

    print(i)
    chainMed1=c(mean(chain[1:i,1]),mean(chain[1:i,2]),mean(chain[1:i,3]))
    chainMed2=c(mean(chain[1:(i-1),1]),mean(chain[1:(i-1),2]),mean(chain[1:(i-1),3]))
    C=((i-1)/i)*C+(sd/i)*(i*chainMed2%*%t(chainMed2)-
(i+1)*chainMed1%*%t(chainMed1)+chain[i,]%*%t(chain[i,])+E*Id)

    proposal=mvrnorm(1,chain[i,],C)
    print(proposal)
    p[i,]=proposal

    posterior[i,] = posterior(proposal,Qmax=Qmax)

    razao=posterior(proposal,Qmax=Qmax)-posterior(chain[i,],Qmax=Qmax)
    guardarazao[i,]<-razao

    if (exp(razao) == "NaN"){

        chain[i+1,]=chain[i,]
        nrejeitado=nrejeitado+1
    }

else

```

```

if(runif(1,0,1)<min(1,exp(razao))){
  chain[i+1,]=proposal
  naceito=naceito+1
}
else{
  chain[i+1,] = chain[i,]
  nrejeitado=nrejeitado+1
}
}
}

lista=list(chain,naceito,nrejeitado,p,C,posteriori,guardarazao)

return(lista)
}

```

```

#####
##PARTE 4: RODA O ALGORITMO (NÃO DEVE SER ALTERADO)
#####

```

```

##### CENÁRIO PDH#####

```

```

setwd(pastadosdeentrada2)

```

```

#b)Carrega os parâmetros maxpost

```

```

arquivoref <- strcat('maxpost',cenarioref)
print(arquivoref)
parametroscurvachave <- read.csv(arquivoref,sep=" ", head=TRUE)
#c) Calcula as vazoes maximas com base nos valores dos parametros max post
a<- parametroscurvachave[1,2]
b<-parametroscurvachave[2,2]
c<-parametroscurvachave[3,2]

```

```

n <- nrow(hmax)
Q <- matrix(0,nrow=n,ncol=1)
Qmax <- matrix(0,nrow=n,ncol=1)

```

```

for (i in 1:n) {
  Qmax[i,1] <- a*(hmax[i,1]-b)^(c)

}

#b) Número de Iterações
nIter=40000
iterations=nIter

#c) Zera a cadeia e carrega os pacotes para rodar o algoritmo

chain=0

library(fExtremes)
library(MASS)
date()
resultadoAM<-run_AdaptiveMetropolis_MCMC_GEV(startvalue,iterations,Qmax=Qmax)
date()

#####
##PARTE 5: GUARDA OS RESULTADOS EM ARQUIVOS .CSV PARA PLOTAGEM
#####
burninfinal = 6000

setwd(pastadadosdesaida)

#### Salva arquivo das cadeias de todos os parâmetros

chainAM=resultadoAM[[1]]

chainAM=matrix(unlist(chainAM),ncol=3,byrow=F)

filecadeia <- strcat("cadeia_cenario",cenario)

write.csv(chainAM,file=filecadeia)

alpha <- exp(chainAM[,1])

filealpha <- strcat("cadeiaalpha_cenario",cenario)

write.csv(alpha,file=filealpha)

# d) Guarda a Taxa de rejeição/aceitação

taxa <- resultadoAM[[2]]/resultadoAM[[3]]

```

```
aceitarejeita <- c(resultadoAM[[2]],resultadoAM[[3]],taxa)
```

```
fileaceitarejeita <-strcat("aceitarejeitacenario",cenario)
```

```
write.csv(aceitarejeita,file=fileaceitarejeita)
```

e) Guarda o Valor da posteriori para cada vetor

```
posteriori <- resultadoAM[[6]]
```

```
filepost <- strcat("posterioricenario",cenario)
```

```
write.csv(posteriori, file = filepost)
```

f) Guarda o resultado final de todos os parâmetros

```
library(TeachingDemos)
```

```
hpdalpha<-emp.hpd(chainAM[burninfinal:iterations,1],conf=0.95)
```

```
maxpostalpha = mean(chainAM[burninfinal:iterations,1])
```

```
hpdcsi<-emp.hpd(chainAM[burninfinal:iterations,2],conf=0.95)
```

```
maxpostcsi = mean(chainAM[burninfinal:iterations,2])
```

```
hpdkapa<-emp.hpd(chainAM[burninfinal:iterations,3],conf=0.95)
```

```
maxpostkapa = mean(chainAM[burninfinal:iterations,3])
```

```
resultssummary1 <- rbind(hpdalpha,hpdcsi,hpdkapa)
```

```
resultssummary2<- rbind(maxpostalpha,maxpostcsi,maxpostkapa)
```

```
summary <- cbind(resultssummary1,resultssummary2)
```

```
filesummary <- strcat("resultssummary_cenario",cenario)
```

```
write.csv(summary,file=filesummary)
```

#a) Calcula a distribuição à posteriori dos diferentes tempos de retorno

Burn in definido com base na análise da cadeia

```
alphachain <-alpha[burninfinal:iterations]
```

```
kapachain <-chainAM[burninfinal:iterations,3]
```

```
csichain <-chainAM[burninfinal:iterations,2]
```

b) Gera Matriz em que cada coluna é um TR e para cada TR calcula o valor da Vazão máxima provável X_p

```
temperetorno <-
```

```
c(1.01,1.1,1.2,1.3,1.4,1.5,1.6,1.7,1.8,1.9,2,2.5,3,4,4.5,5,5.5,10,15,20,25,30,35,40,45,50,55,60,75,80,85,90,100)
```

```

tempoderetorno<-matrix(tempoderetorno,nrow=1,ncol=33)

p = 1/tempoderetorno

n = ncol(tempoderetorno)

estfreq<-matrix(0,nrow=(iterations-burninfinal+1),ncol=n)

total <- iterations - burninfinal + 1

for (i in 1:n){
  prob <- p[i]
  yprob <- -log(1-prob)

  for (k in 1:total) {

    if ((kapachain[k] < 0.01)&&(kapachain[k]> -0.01)) {
      estfreq[k,i]<- csichain[k] -alphachain[k]*(log(yprob))

    } else
      estfreq[k,i] <- csichain[k] + alphachain[k]/kapachain[k]*(1-(yprob)^kapachain[k])

  }
}

estfreqfinal <- rbind(tempoderetorno,estfreq)
filefreq <- strcat("estfreq_cenario",cenario)
write.csv(estfreqfinal,file=filefreq)

#####
##PARTE 7: CALCULA A CURVA DE FREQUÊNCIA DE CHEIAS + IC
#####

# a) Intervalo de confiança fornecido pelo usuário
IC<- 0.95

#primeira coluna o mais provável , segunda coluna a parte inferior, segunda coluna a parte superior

library(TeachingDemos)

curvadefrequenciadecheias <- matrix(0,nrow=n,ncol=3)

for (i in 1:n){
  curvadefrequenciadecheias[i,] <- c(mean(estfreq[,i]),emp.hpd(estfreq[,i],conf=IC))
}

```

```

tempoderetorno1 <-matrix(tempoderetorno,nrow=33,ncol=1)

filecurva <- strcat("curvadefrequenciadecheias_cenario",cenario)

curvadefrequenciadecheiasfinal <- cbind(tempoderetorno1,curvadefrequenciadecheias)

write.csv(curvadefrequenciadecheiasfinal,file = filecurva)

#####Gráficos

cadeiadosvaloresdephi <- strcat("Cadeia Valores de Phi Cenario ")
plot(chainAM[1:iterations,1],type="l",xlab="Iteração",ylab="Phi",main=cadeiadosvalores
dephi)
abline(h = mean(chainAM[1:iterations,1]),lwd=2,col="darkred")
#abline(h = 5.70 ,lwd=2,lty=2,col="blue")
legend("bottomright",lty=c(1,2),lwd=c(2,2),bty="n",cex=0.8,col=c("darkred","blue"),
legend=c("Média Cadeia -AM","Valor Estimado MT-Alpha"))

cadeiadosvaloresdecsi <- strcat("Cadeia Csi Cenario ")
plot(chainAM[1:iterations,2],type="l",xlab="Iteração",ylab="Csi",main=cadeiadosvalores
decsi)
abline(h = mean(chainAM[1:iterations,2]),lwd=2,col="darkred")
#abline(h = 1500 ,lwd=2,lty=2,col="blue")
legend("bottomright",lty=c(1,2),lwd=c(2,2),bty="n",cex=0.8,col=c("darkred","blue"),
legend=c("Média Cadeia -AM","Valor Estimado MT-Csi"))

cadeiadosvaloresdekapa <- strcat("Cadeia Kapa Cenario ")
plot(chainAM[1:iterations,3],type="l",xlab="Iteração",ylab="Kapa",main=cadeiadosvalore
sdekapa)
abline(h = mean(chainAM[1:iterations,3]),lwd=2,col="darkred")
#abline(h = -0.04 ,lwd=2,lty=2,col="blue")
legend("bottomright",lty=c(1,2),lwd=c(2,2),bty="n",cex=0.8,col=c("darkred","blue"),
legend=c("Média Cadeia -AM","Valor Estimado MT-Kapa"))

##### MAX POST

cadeia<-chainAM
post <- posteriori

## Calcula o max post

n <- nrow(post)

for (i in 2:n) {
  if (i==2) {

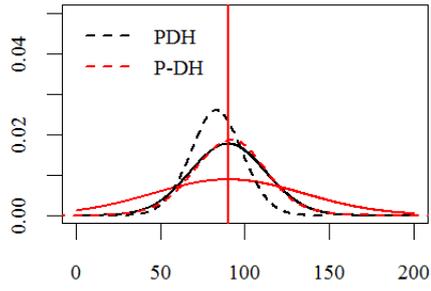
```

```
posttemp <- post[i-1,1]
maxpost <- cadeia[i,]
}
else {
  if (post[i,1]>posttemp) {
    posttemp <- post[i-1,1]
    maxpost <- cadeia[i,]

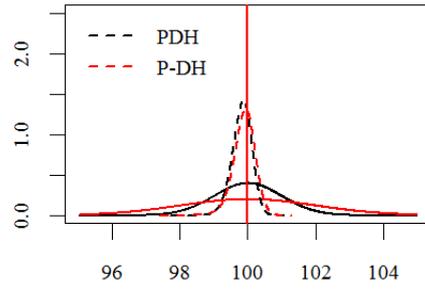
  }
}
}
```

```
filemaxpost <- strcat('maxpost',cenario)
write.csv(maxpost,file = filemaxpost)
```

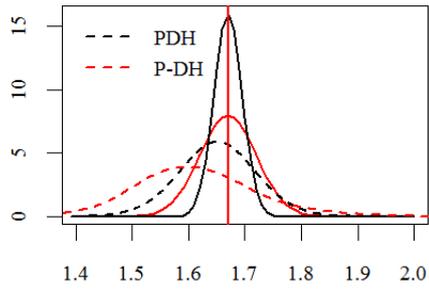
**APÊNDICE II – DISTRIBUIÇÕES A *POSTERIORI* COMPARATIVA
DOS CENÁRIOS**



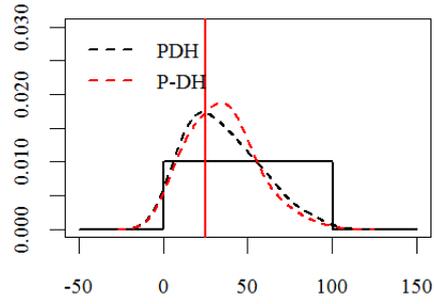
a



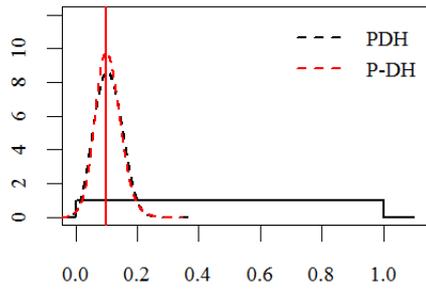
b



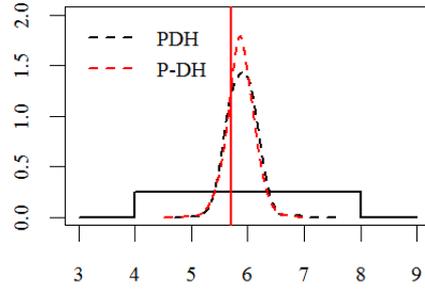
c



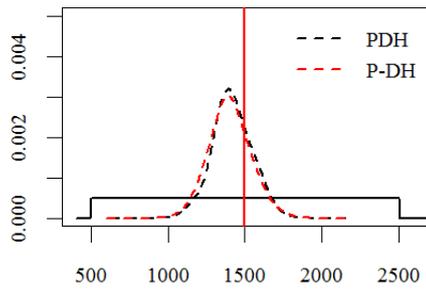
Gamma1



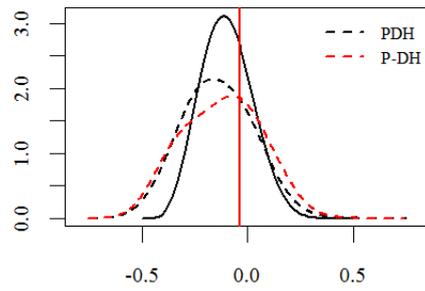
Gamma2



phi=ln(alpha)

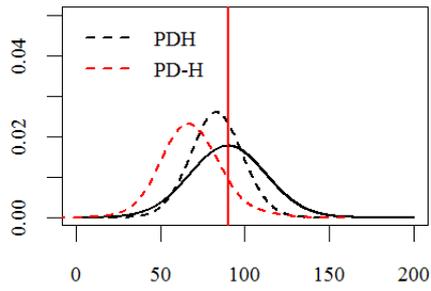


csi

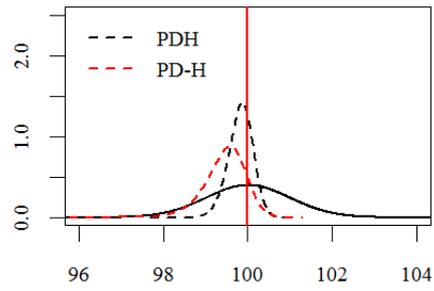


kappa

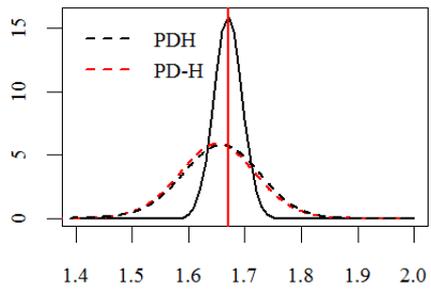
ESTUDO DE CASO SINTÉTICO - Comparativo das distribuições *a priori* (linha cheia) e *a posteriori* (linha tracejada) de todos os parâmetros dos cenários PDH (preto tracejado) e P-DH (vermelho tracejado) com indicação do valor real (linha vertical)



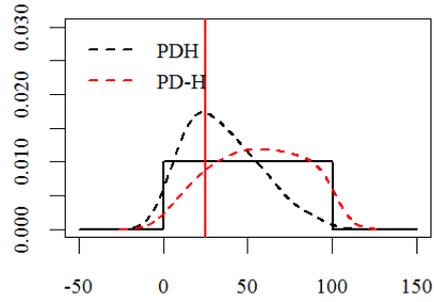
a



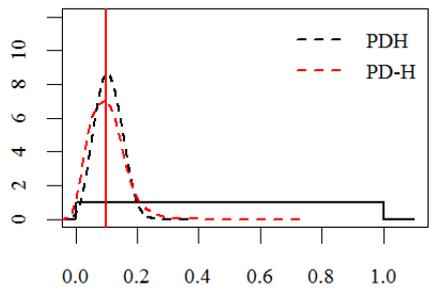
b



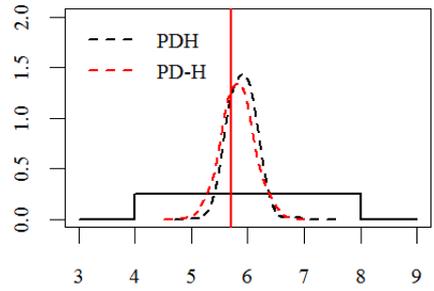
c



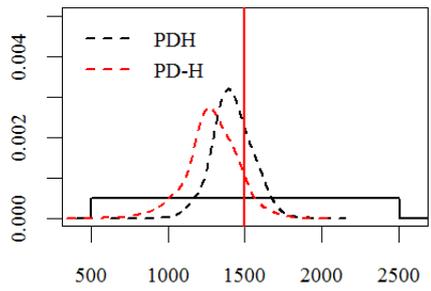
Gamma1



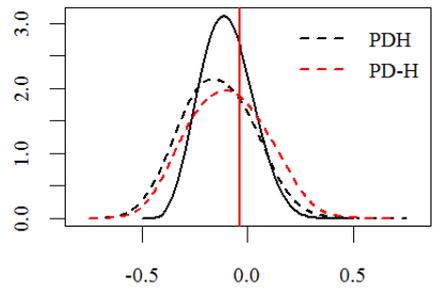
Gamma2



phi=ln(alpha)

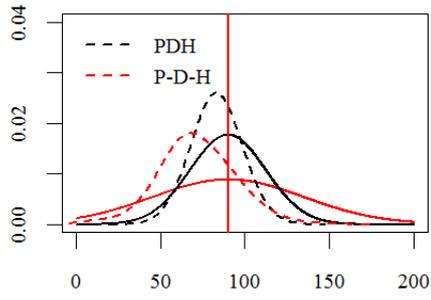


csi

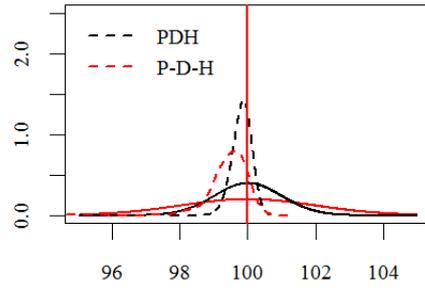


kappa

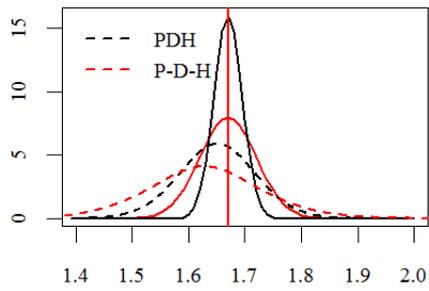
ESTUDO DE CASO SINTÉTICO - Comparativo das distribuições *a priori* (linha cheia) e *a posteriori* (linha tracejada) de todos os parâmetros dos cenários PDH (preto tracejado) e PD-H (vermelho tracejado) com indicação do valor real (linha vertical)



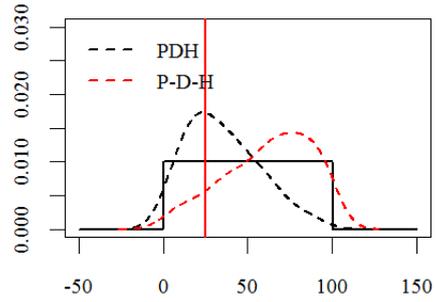
a



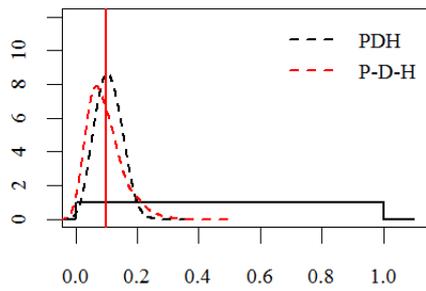
b



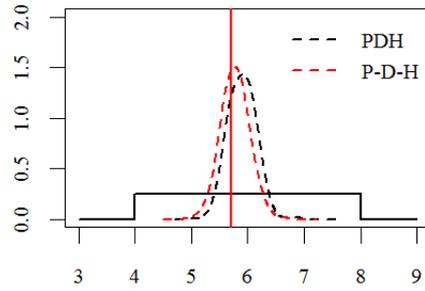
c



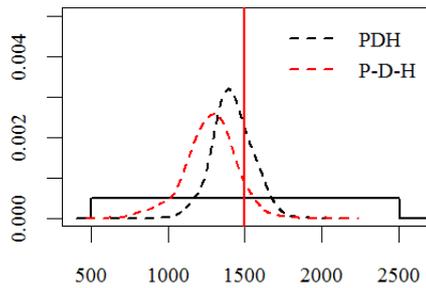
Gamma1



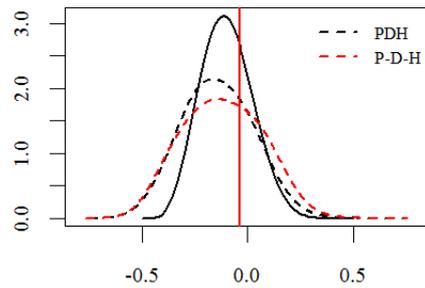
Gamma2



phi=ln(alpha)

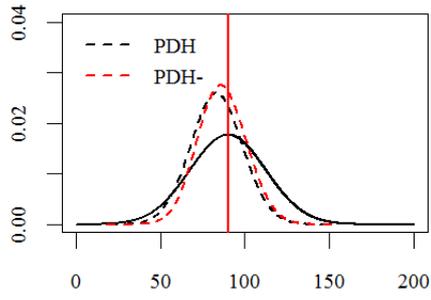


csi

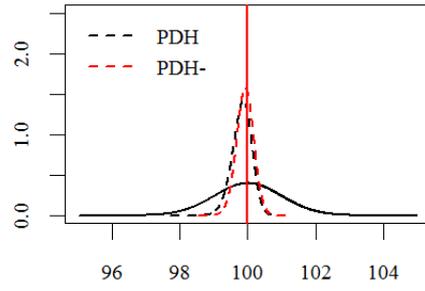


kappa

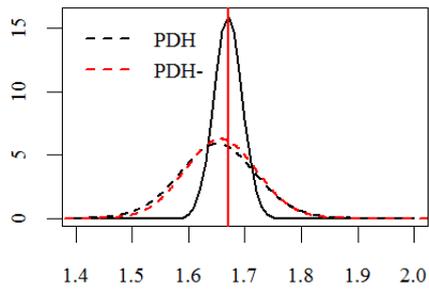
ESTUDO DE CASO SINTÉTICO - Comparativo das distribuições *a priori* (linha cheia) e *a posteriori* (linha tracejada) de todos os parâmetros dos cenários PDH (preto tracejado) e P-D-H (vermelho tracejado) com indicação do valor real (linha vertical)



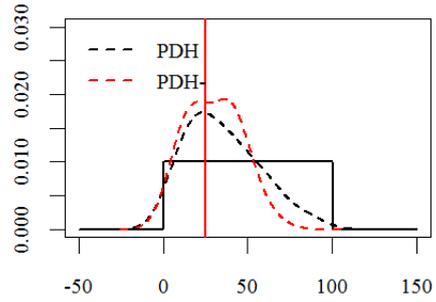
a



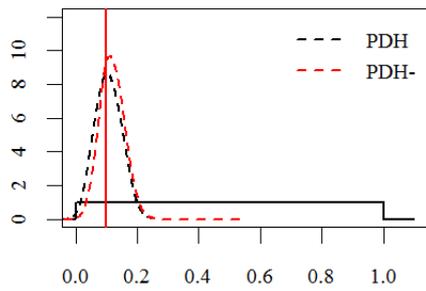
b



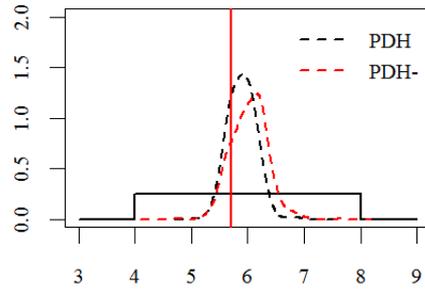
c



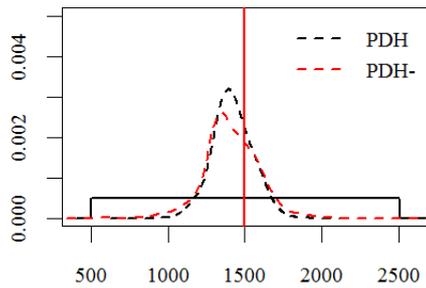
Gamma1



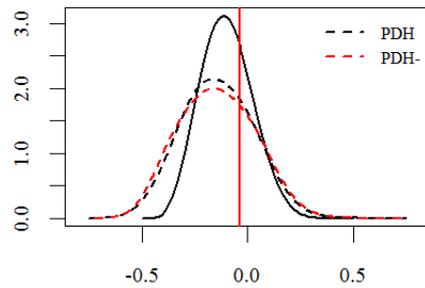
Gamma2



phi=ln(alpha)

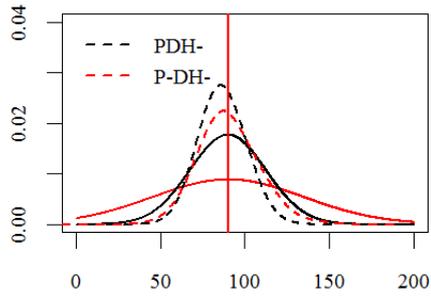


csi

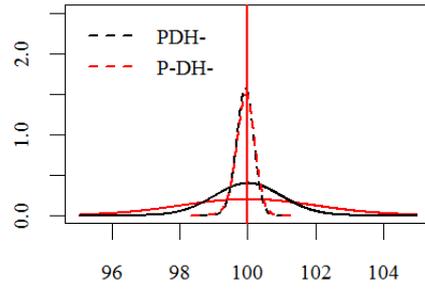


kappa

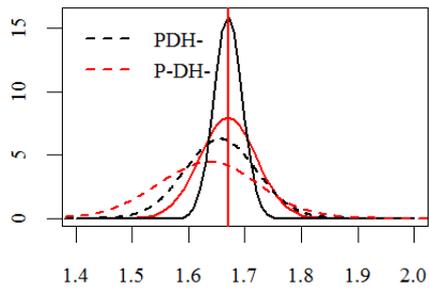
ESTUDO DE CASO SINTÉTICO - Comparativo das distribuições *a priori* (linha cheia) e *a posteriori* (linha tracejada) de todos os parâmetros dos cenários PDH (preto tracejado) e PDH- (vermelho tracejado) com indicação do valor real (linha vertical)



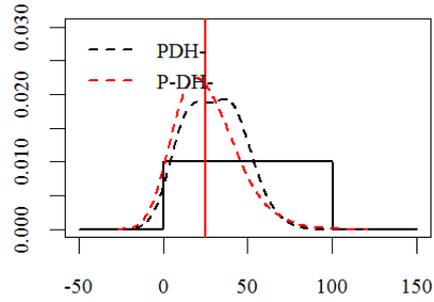
a



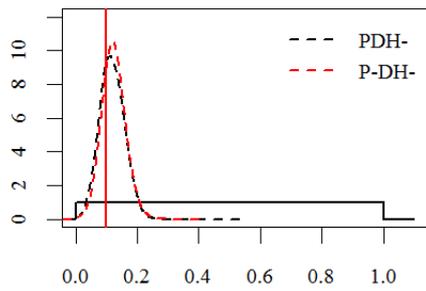
b



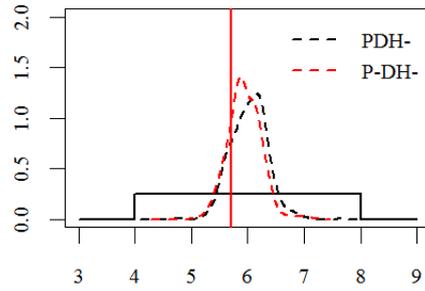
c



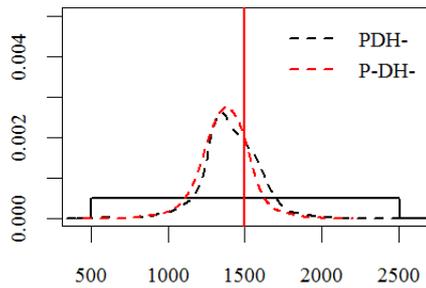
Gamma1



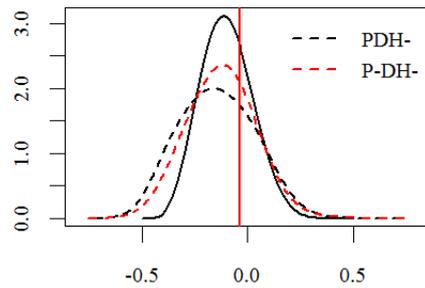
Gamma2



phi=ln(alpha)

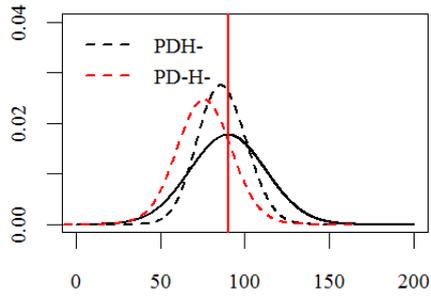


csi

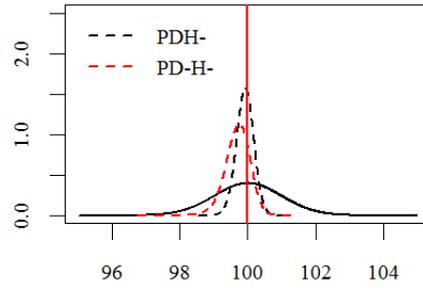


kappa

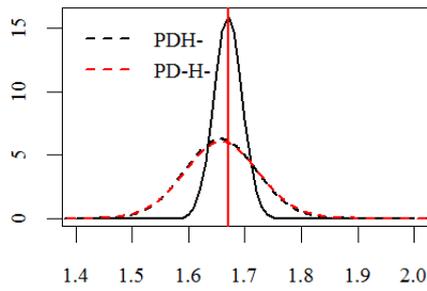
ESTUDO DE CASO SINTÉTICO - Comparativo das distribuições *a priori* (linha cheia) e *a posteriori* (linha tracejada) de todos os parâmetros dos cenários PDH- (preto tracejado) e P-DH- (vermelho tracejado) com indicação do valor real (linha vertical)



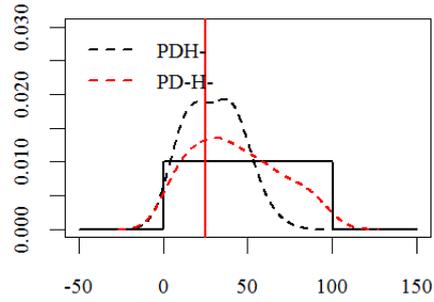
a



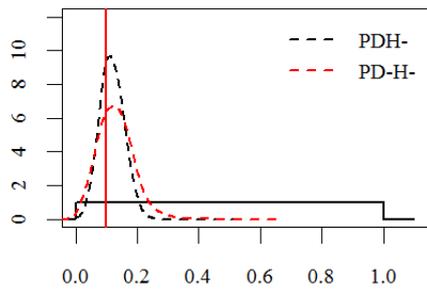
b



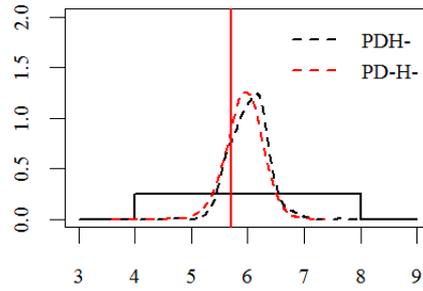
c



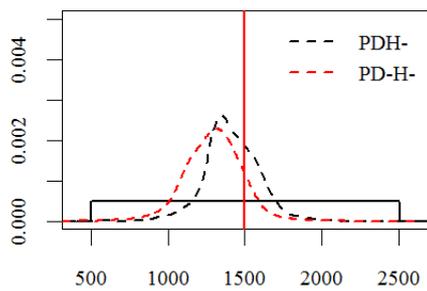
Gamma1



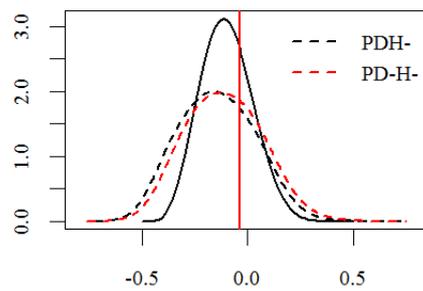
Gamma2



phi=ln(alpha)

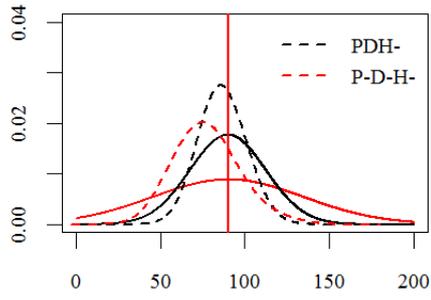


csi

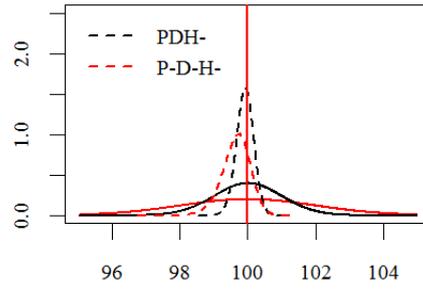


kappa

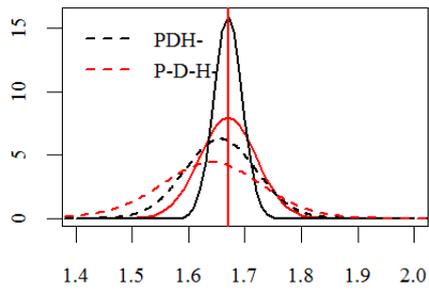
ESTUDO DE CASO SINTÉTICO - Comparativo das distribuições *a priori* (linha cheia) e *a posteriori* (linha tracejada) de todos os parâmetros dos cenários PDH- (preto tracejado) e PD-H- (vermelho tracejado) com indicação do valor real (linha vertical)



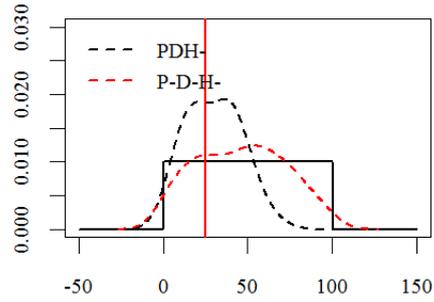
a



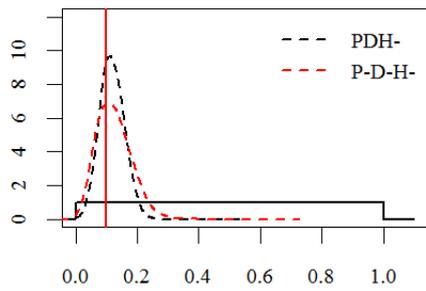
b



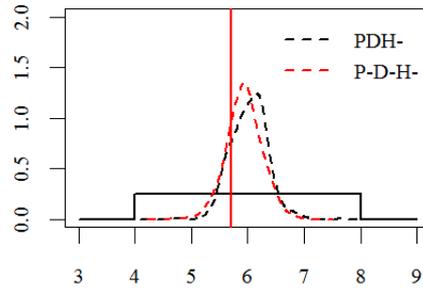
c



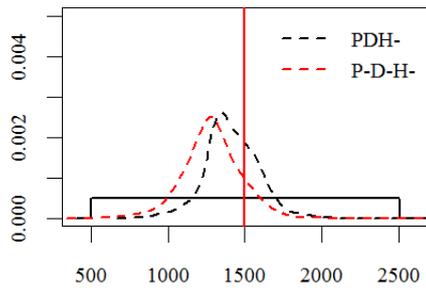
Gamma1



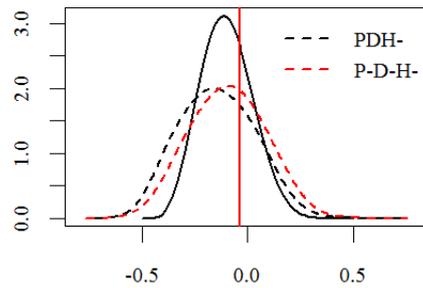
Gamma2



phi=ln(alpha)

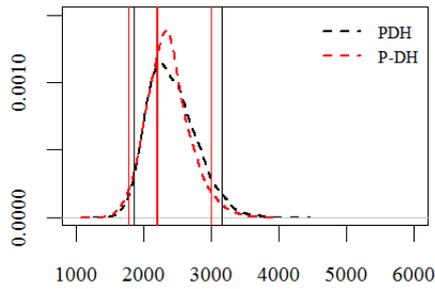


csi

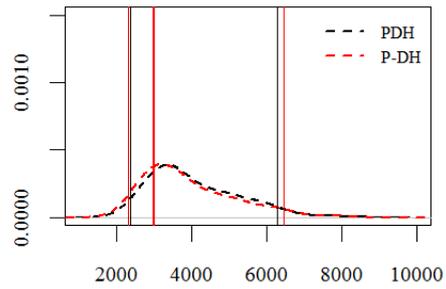


kappa

ESTUDO DE CASO SINTÉTICO - Comparativo das distribuições *a priori* (linha cheia) e *a posteriori* (linha tracejada) de todos os parâmetros dos cenários PDH- (preto tracejado) e P-D-H- (vermelho tracejado) com indicação do valor real (linha vertical)

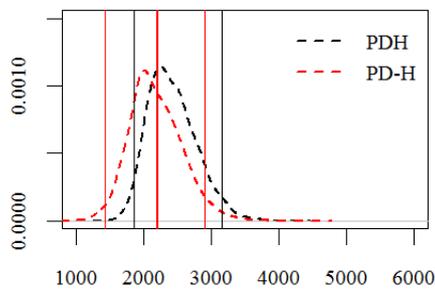


Q - TR 10

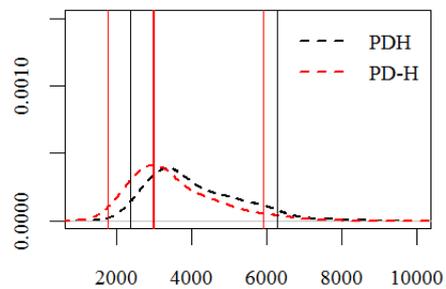


Q - TR 100

ESTUDO DE CASO SINTÉTICO -Comparativo das distribuições *a posteriori* dos quantis de tempo de retorno de 10 e 100 anos para os cenários PDH e P-DH com a indicação dos respectivos intervalos de credibilidade (linhas cheias verticais).

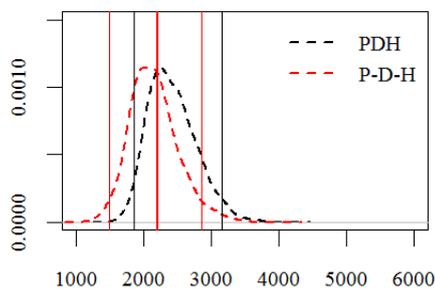


Q - TR 10

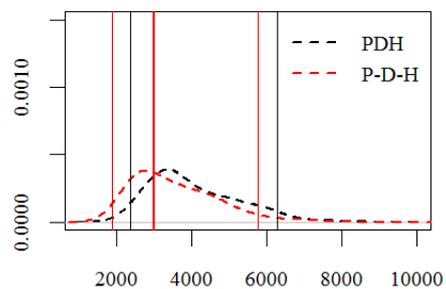


Q - TR 100

ESTUDO DE CASO SINTÉTICO -Comparativo das distribuições *a posteriori* dos quantis de tempo de retorno de 10 e 100 anos para os cenários PDH e PD-H com a indicação dos respectivos intervalos de credibilidade (linhas cheias verticais).

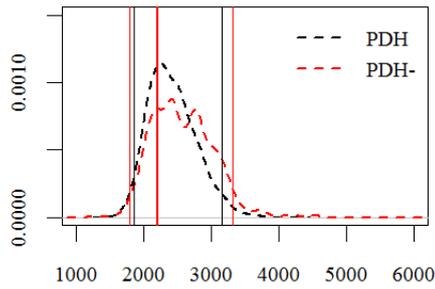


Q - TR 10

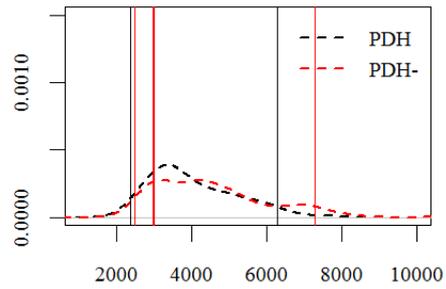


Q - TR 100

ESTUDO DE CASO SINTÉTICO -Comparativo das distribuições *a posteriori* dos quantis de tempo de retorno de 10 e 100 anos para os cenários PDH e P-D-H com a indicação dos respectivos intervalos de credibilidade (linhas cheias verticais).

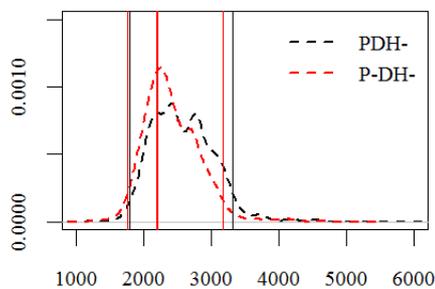


Q - TR 10

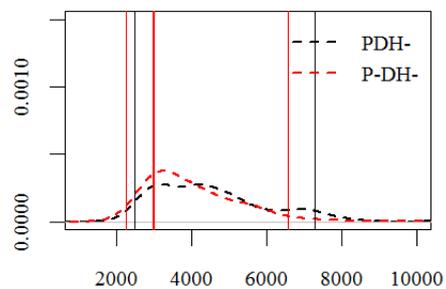


Q - TR 100

ESTUDO DE CASO SINTÉTICO -Comparativo das distribuições *a posteriori* dos quantis de tempo de retorno de 10 e 100 anos para os cenários PDH e PDH- com a indicação dos respectivos intervalos de credibilidade (linhas cheias verticais).

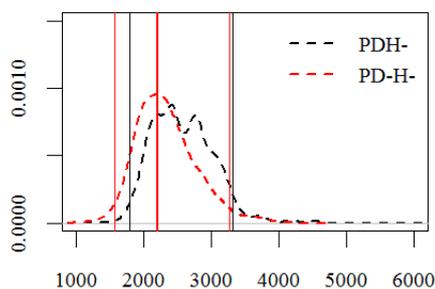


Q - TR 10

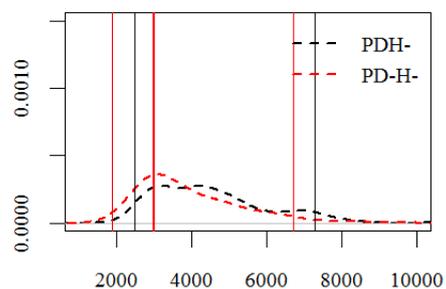


Q - TR 100

ESTUDO DE CASO SINTÉTICO -Comparativo das distribuições *a posteriori* dos quantis de tempo de retorno de 10 e 100 anos para os cenários PDH- e P-DH- com a indicação dos respectivos intervalos de credibilidade (linhas cheias verticais).

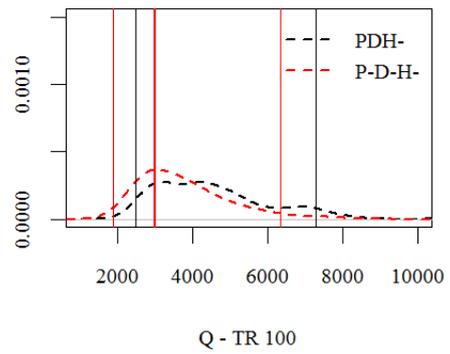
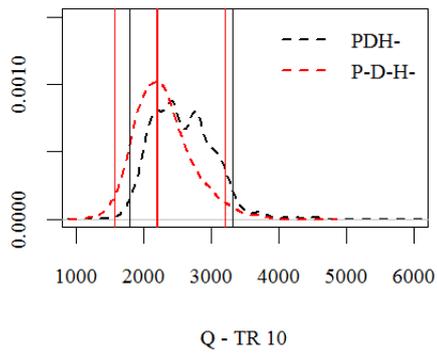


Q - TR 10

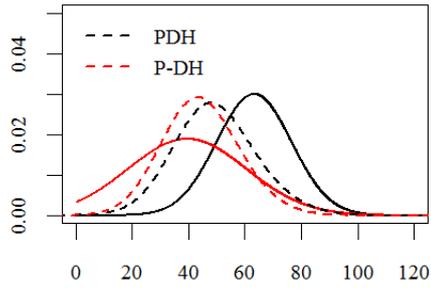


Q - TR 100

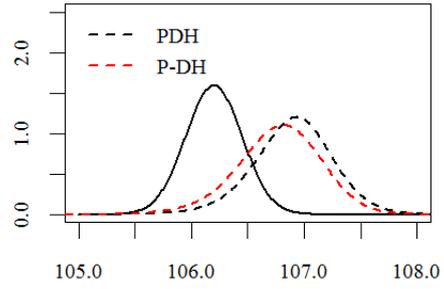
ESTUDO DE CASO SINTÉTICO -Comparativo das distribuições *a posteriori* dos quantis de tempo de retorno de 10 e 100 anos para os cenários PDH- e PD-H- com a indicação dos respectivos intervalos de credibilidade (linhas cheias verticais).



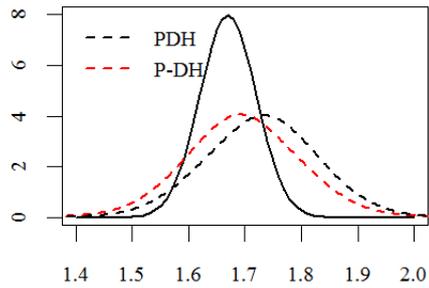
ESTUDO DE CASO SINTÉTICO -Comparativo das distribuições *a posteriori* dos quantis de tempo de retorno de 10 e 100 anos para os cenários PDH- e P-D-H- com a indicação dos respectivos intervalos de credibilidade (linhas cheias verticais).



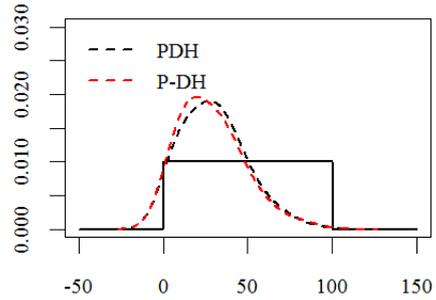
a



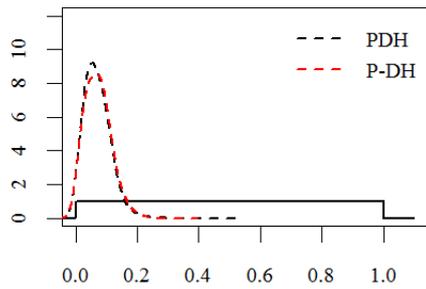
b



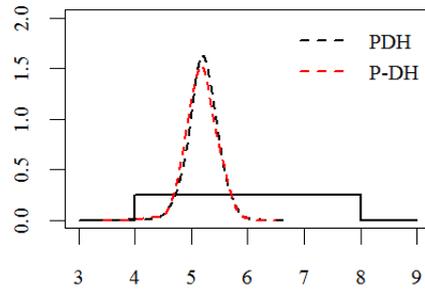
c



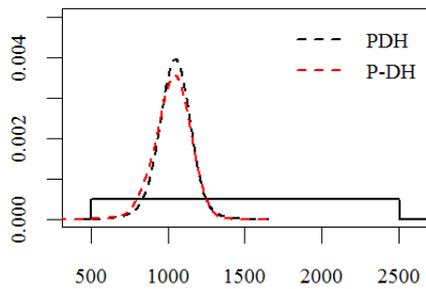
Gamma1



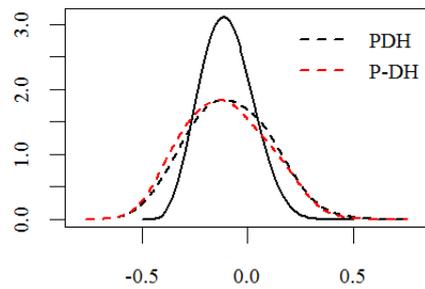
Gamma2



phi=ln(alpha)

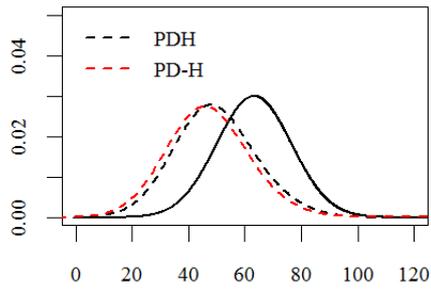


csi

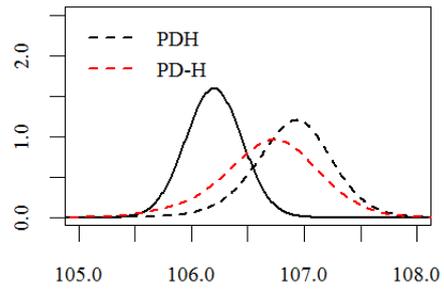


kappa

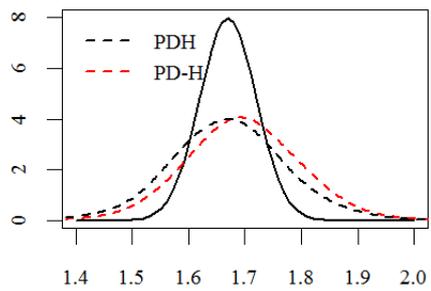
ESTUDO DE CASO EM CÁCERES - Comparativo das distribuições *a priori* (linha cheia) e *a posteriori* (linha tracejada) de todos os parâmetros dos cenários PDH (preto tracejado) e P-DH (vermelho tracejado)



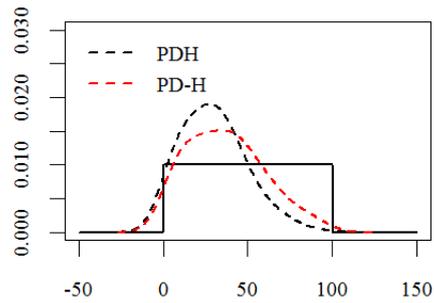
a



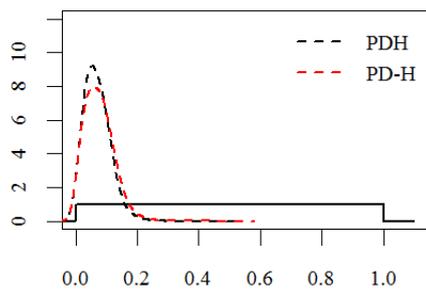
b



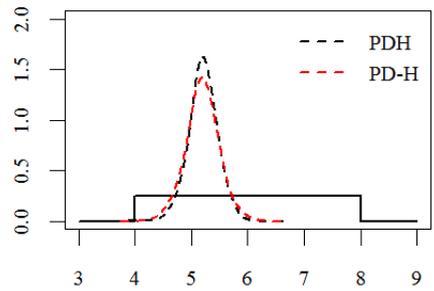
c



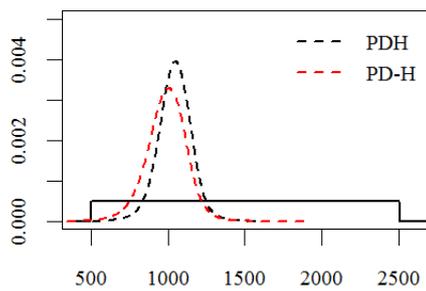
Gamma1



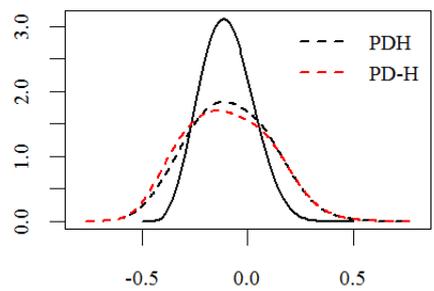
Gamma2



phi=ln(alpha)

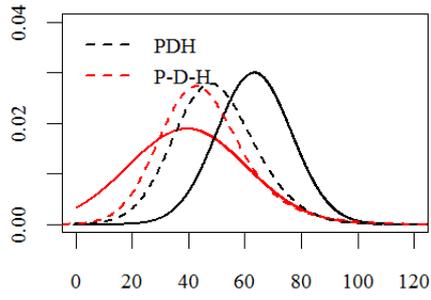


csi

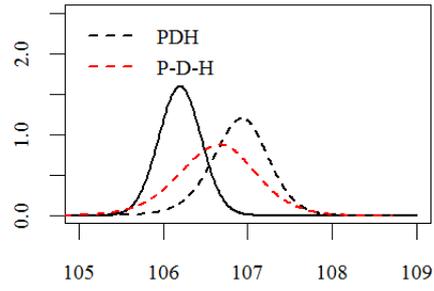


kappa

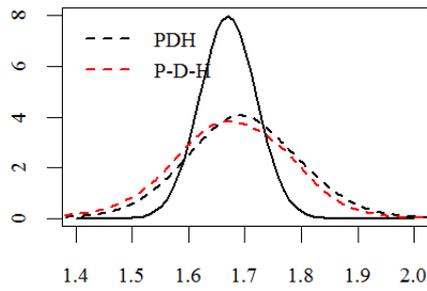
ESTUDO DE CASO EM CÁCERES - Comparativo das distribuições *a priori* (linha cheia) e *a posteriori* (linha tracejada) de todos os parâmetros dos cenários PDH (preto tracejado) e PD-H (vermelho tracejado)



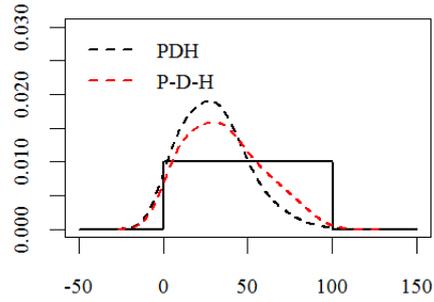
a



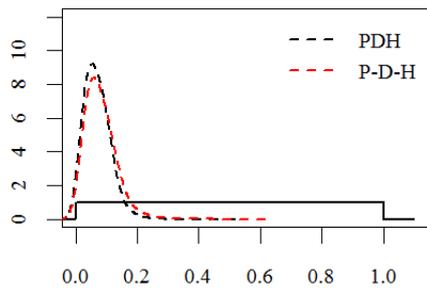
b



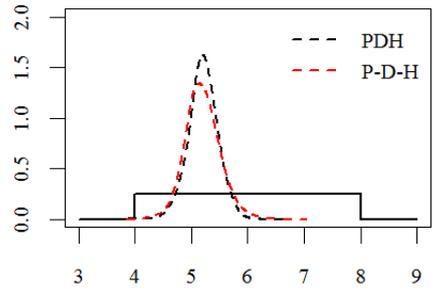
c



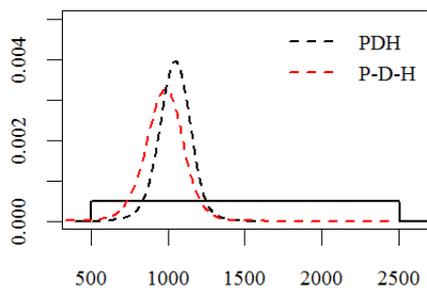
Gamma1



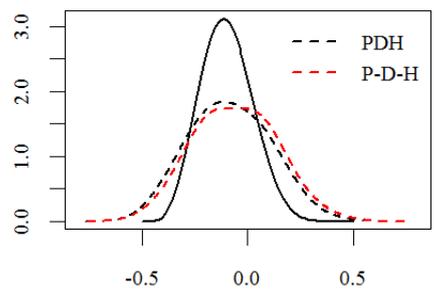
Gamma2



phi=ln(alpha)

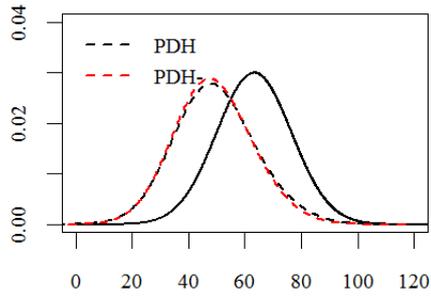


csi

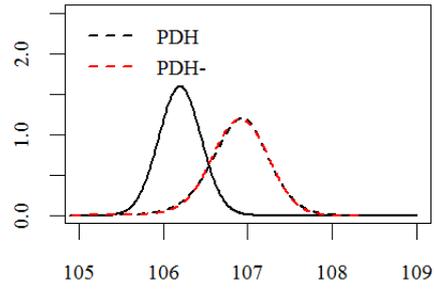


kappa

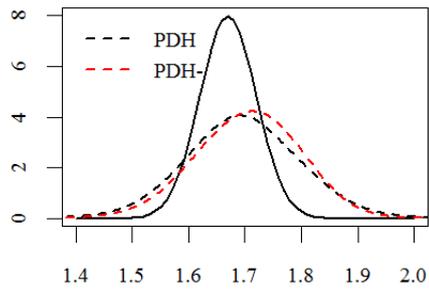
ESTUDO DE CASO EM CÁCERES - Comparativo das distribuições *a priori* (linha cheia) e *a posteriori* (linha tracejada) de todos os parâmetros dos cenários PDH (preto tracejado) e P-D-H (vermelho tracejado)



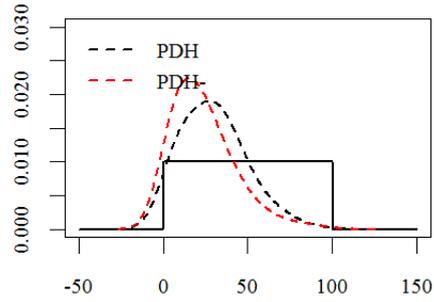
a



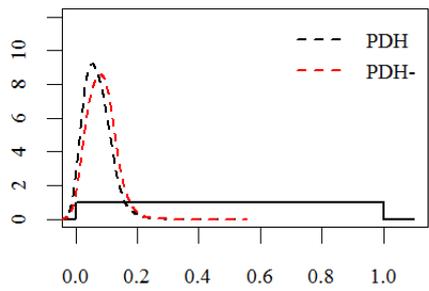
b



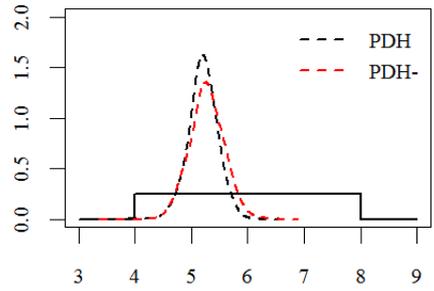
c



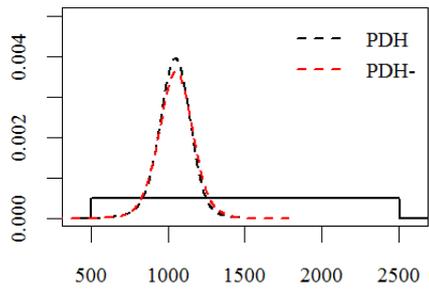
Gamma1



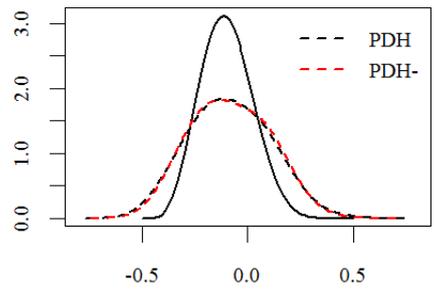
Gamma2



phi=ln(alpha)

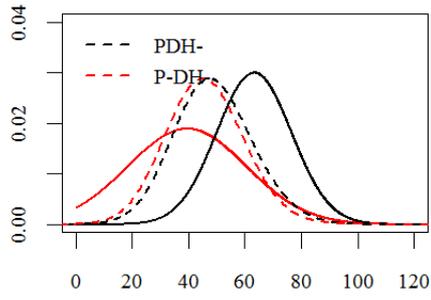


csi

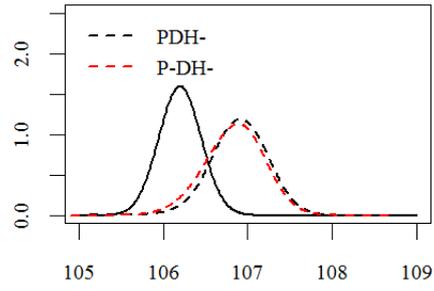


kappa

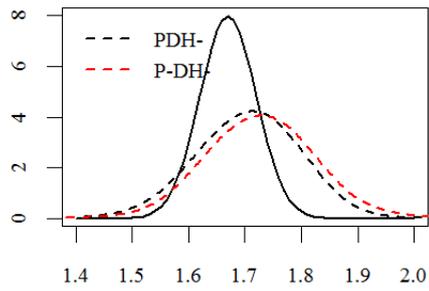
ESTUDO DE CASO EM CÁCERES - Comparativo das distribuições *a priori* (linha cheia) e *a posteriori* (linha tracejada) de todos os parâmetros dos cenários PDH (preto tracejado) e PDH- (vermelho tracejado)



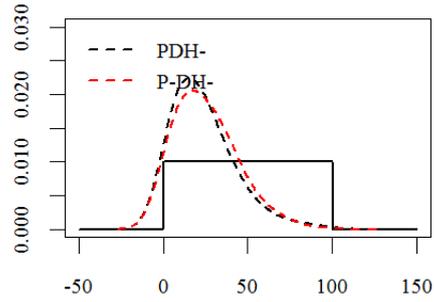
a



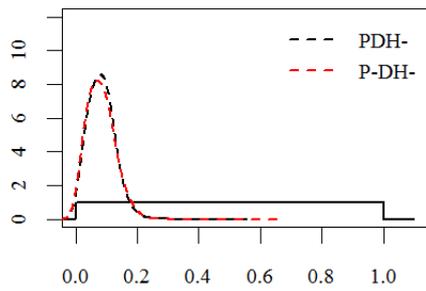
b



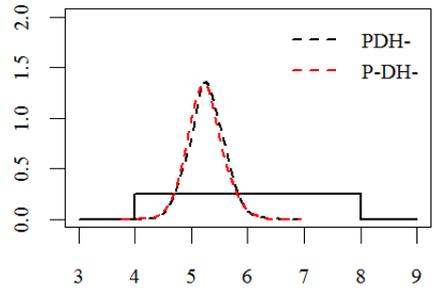
c



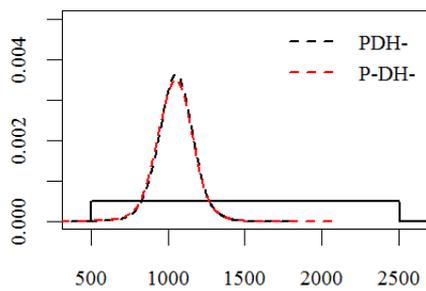
Gamma1



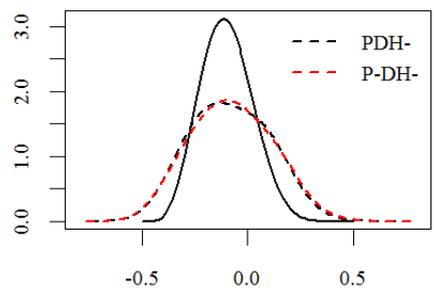
Gamma2



phi=ln(alpha)

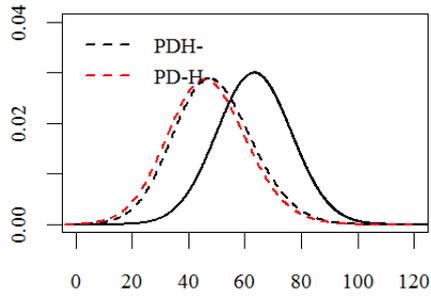


csi

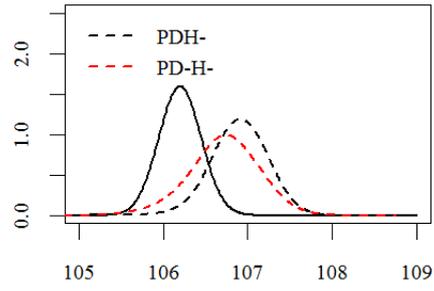


kappa

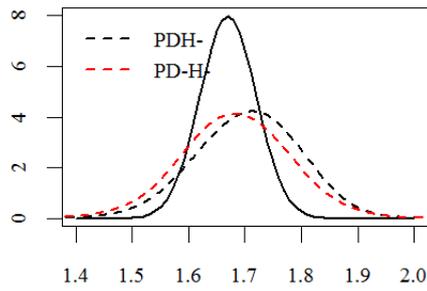
ESTUDO DE CASO EM CÁCERES - Comparativo das distribuições *a priori* (linha cheia) e *a posteriori* (linha tracejada) de todos os parâmetros dos cenários PDH- (preto tracejado) e P-DH- (vermelho tracejado)



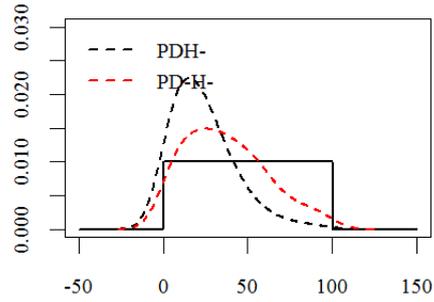
a



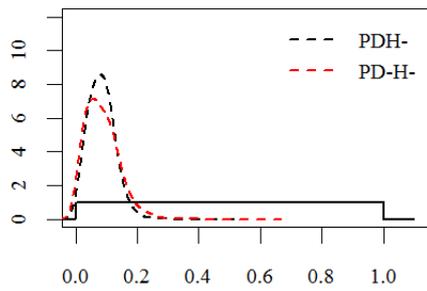
b



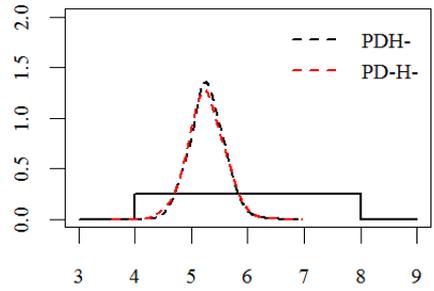
c



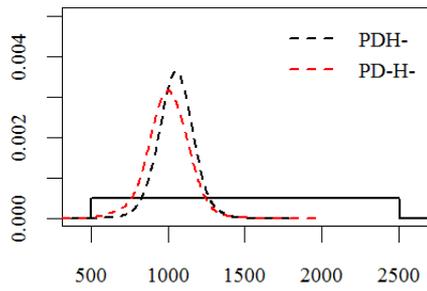
Gamma1



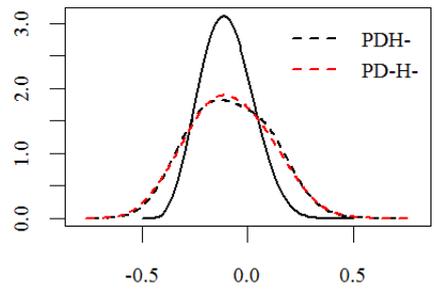
Gamma2



phi=ln(alpha)

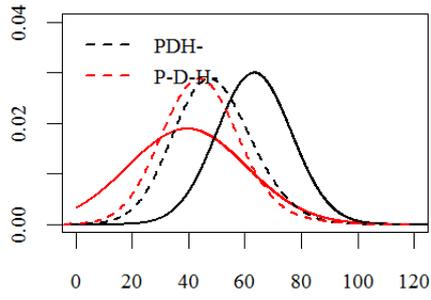


csi

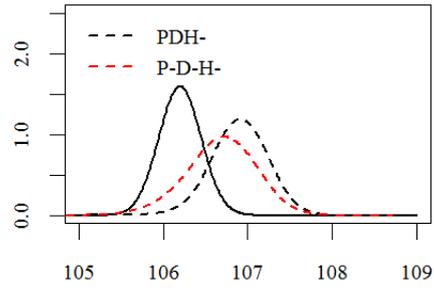


kappa

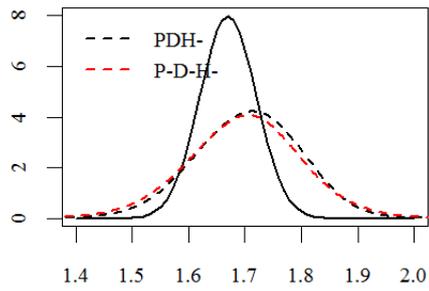
ESTUDO DE CASO EM CÁ CERES - Comparativo das distribuições *a priori* (linha cheia) e *a posteriori* (linha tracejada) de todos os parâmetros dos cenários PDH- (preto tracejado) e PD-H- (vermelho tracejado)



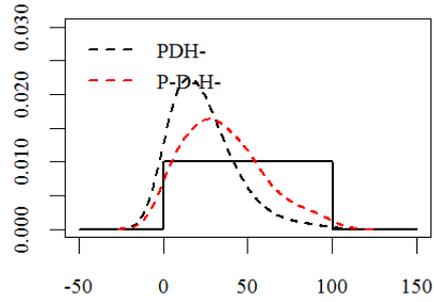
a



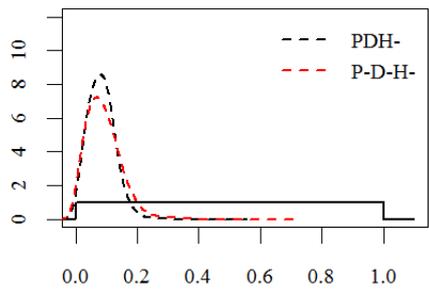
b



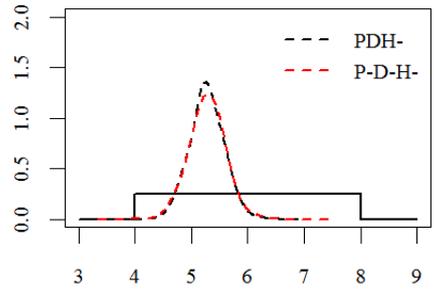
c



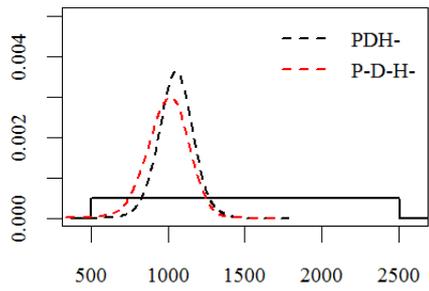
Gamma1



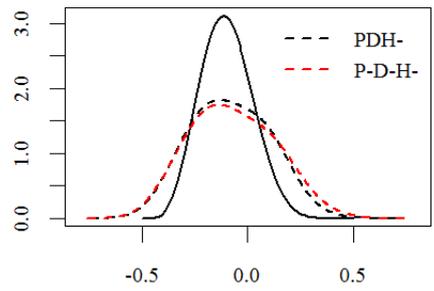
Gamma2



phi=ln(alpha)

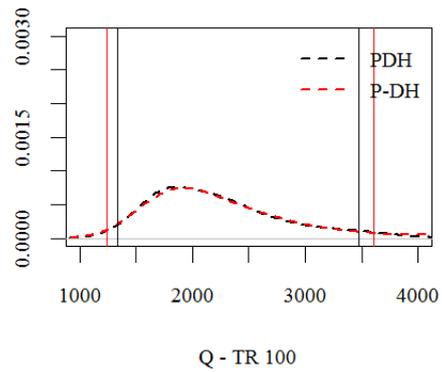
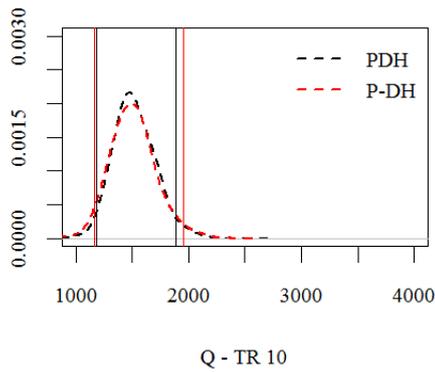


csi

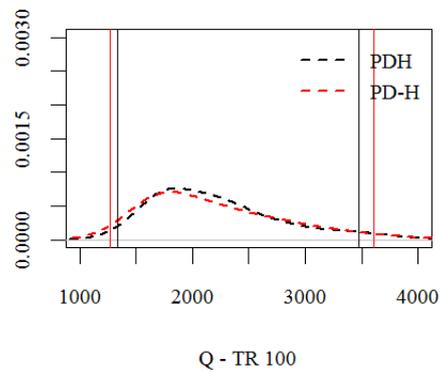
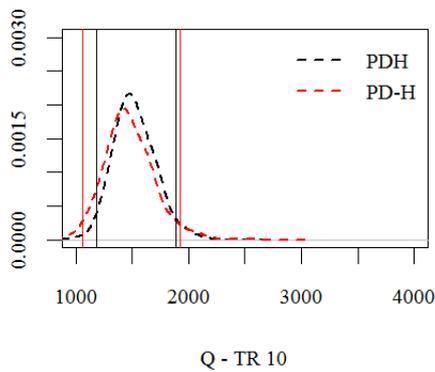


kappa

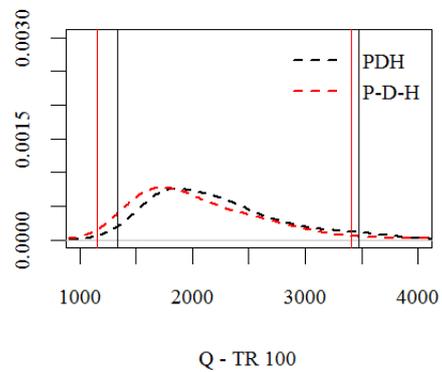
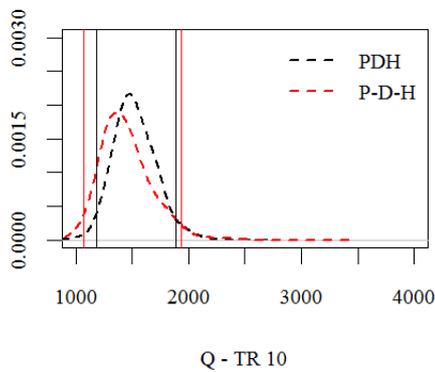
ESTUDO DE CASO EM CÁ CERES - Comparativo das distribuições *a priori* (linha cheia) e *a posteriori* (linha tracejada) de todos os parâmetros dos cenários PDH- (preto tracejado) e P-D-H- (vermelho tracejado)



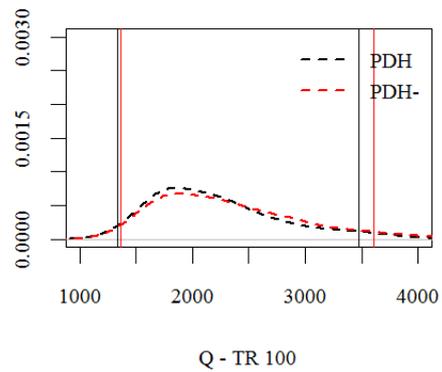
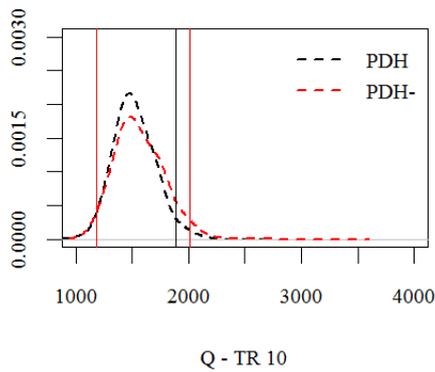
ESTUDO DE CASO EM CÁCERES -Comparativo das distribuições *a posteriori* dos quantis de tempo de retorno de 10 e 100 anos para os cenários PDH e P-DH com a indicação dos respectivos intervalos de credibilidade (linhas cheias verticais).



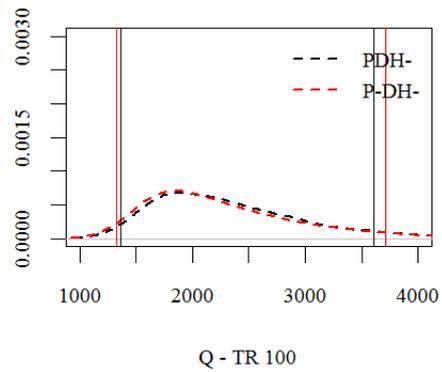
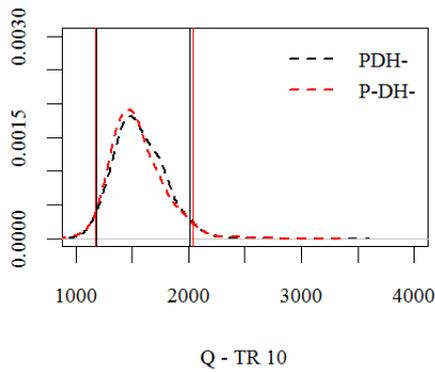
ESTUDO DE CASO EM CÁCERES - Comparativo das distribuições *a posteriori* dos quantis de tempo de retorno de 10 e 100 anos para os cenários PDH e PD-H com a indicação dos respectivos intervalos de credibilidade (linhas cheias verticais).



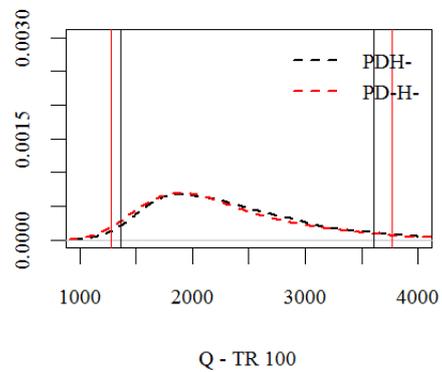
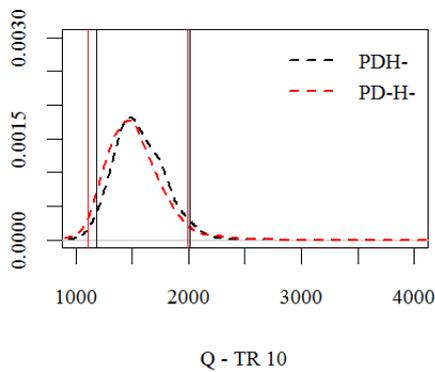
ESTUDO DE CASO EM CÁCERES -Comparativo das distribuições *a posteriori* dos quantis de tempo de retorno de 10 e 100 anos para os cenários PDH e P-D-H com a indicação dos respectivos intervalos de credibilidade (linhas cheias verticais).



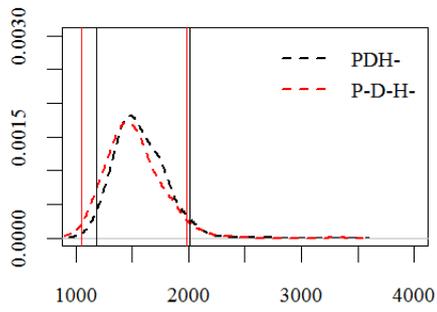
ESTUDO DE CASO EM CÁCERES -Comparativo das distribuições *a posteriori* dos quantis de tempo de retorno de 10 e 100 anos para os cenários PDH e PDH- com a indicação dos respectivos intervalos de credibilidade (linhas cheias verticais).



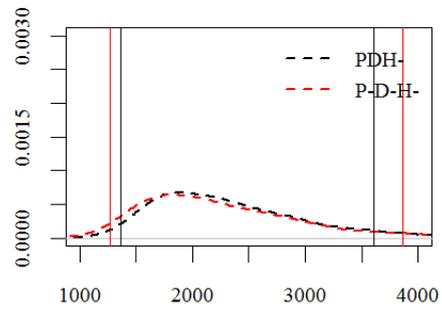
ESTUDO DE CASO EM CÁCERES -Comparativo das distribuições *a posteriori* dos quantis de tempo de retorno de 10 e 100 anos para os cenários PDH- e P-DH- com a indicação dos respectivos intervalos de credibilidade (linhas cheias verticais).



ESTUDO DE CASO EM CÁCERES -Comparativo das distribuições *a posteriori* dos quantis de tempo de retorno de 10 e 100 anos para os cenários PDH- e PD-H- com a indicação dos respectivos intervalos de credibilidade (linhas cheias verticais).



Q - TR 10



Q - TR 100

ESTUDO DE CASO EM CÁCERES -Comparativo das distribuições *a posteriori* dos quantis de tempo de retorno de 10 e 100 anos para os cenários PDH- e P-D-H- com a indicação dos respectivos intervalos de credibilidade (linhas cheias verticais).

TR	PDH			P-DH			PDH-			P-DH-		
	IG(%)	IA(%)	IR	IG(%)	IA(%)	IR	IG(%)	IA(%)	IR	IG(%)	IA(%)	IR
TR=10	53	22	59	52	25	52	59	31	48	58	33	43
TR=25	68	32	52	70	36	48	75	42	44	74	45	39
TR=50	79	42	46	86	47	45	92	52	43	88	56	36
TR=100	96	54	44	105	59	44	106	64	40	106	68	36
TR	PD-H			P-D-H			PD-H-			P-D-H-		
	IG(%)	IA(%)	IR	IG(%)	IA(%)	IR	IG(%)	IA(%)	IR	IG(%)	IA(%)	IR
TR=10	67	25	62	63	22	65	73	29	60	71	30	58
TR=25	84	39	54	80	34	57	92	41	56	85	41	51
TR=50	98	50	49	94	44	53	107	50	53	100	51	49
TR=100	117	63	47	109	56	49	122	62	49	117	62	47

Tabela comparativa com todos os resultados dos 8 cenários para o ESTUDO DE CASO SINTÉTICO

TR	PDH			P-DH			PDH-			P-DH -		
	IG(%)	IA(%)	IR	IG(%)	IA(%)	IR	IG(%)	IA(%)	IR	IG(%)	IA(%)	IR
10	46	20	57	52	17	67	53	20	61	56	20	63
25	60	30	50	70	26	62	68	30	56	71	30	57
50	76	39	49	86	35	59	81	39	52	87	40	55
100	96	50	47	105	46	56	97	51	47	105	50	52
TR	PD-H			P-D-H			PD-H-			P-D-H-		
	IG(%)	IA(%)	IR	IG(%)	IA(%)	IR	IG(%)	IA(%)	IR	IG(%)	IA(%)	IR
10	59	17	72	60	16	73	58	21	64	62	19	70
25	73	28	62	78	26	66	73	33	55	79	30	62
50	86	38	56	93	35	62	89	43	52	95	40	58
100	105	52	51	107	45	58	109	54	50	114	52	54

Tabela comparativa com todos os resultados dos 8 cenários para o ESTUDO DE CASO EM CÁCERES