

Sanjay Formighieri

**Calibração Bayesiana do modelo de Bourgoyne e
Young via Monte Carlo em Cadeias de Markov**

**Florianópolis, SC
9 de Março de 2016**

Sanjay Formighieri

**Calibração Bayesiana do modelo de Bourgoyne e Young via
Monte Carlo em Cadeias de Markov**

Dissertação apresentada ao Programa de Pós Graduação em Ciência da Computação da Universidade Federal de Santa Catarina como requisito parcial para a obtenção do grau de Mestre em Ciência da Computação.

Universidade Federal de Santa Catarina

Orientador: Paulo José de Freitas Filho

Florianópolis, SC
9 de Março de 2016

Formighieri, Sanjay

Calibração Bayesiana do modelo de Bourgoyne e Young via Monte Carlo em Cadeias de Markov / Sanjay Formighieri; orientador, Paulo José de Freitas Filho - Florianópolis, SC, 2016.

82 p.

Dissertação (mestrado) - Universidade Federal de Santa Catarina, Centro Tecnológico. Programa de Pós-Graduação em Ciência da Computação.

Inclui Referências

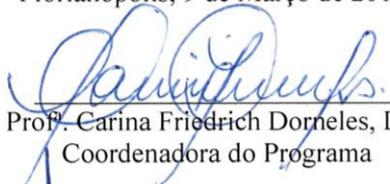
1.Ciência da Computação. 2.MCMC. 3.ROP. 4.Modelo de Borgoune e Young. I. Freitas Filho, Paulo José de. II. Universidade Federal de Santa Catarina. Programa de Pós-Graduação em Ciência da Computação. III. Calibração Bayesiana do modelo de Bourgoyne e Young via Monte Carlo em Cadeias de Markov.

Sanjay Formighieri

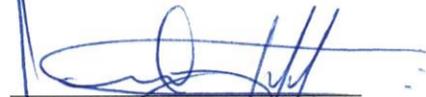
**CALIBRAÇÃO BAYESIANA DO MODELO DE BOURGOYNE E
YOUNG VIA MONTE CARLO EM CADEIAS DE MARKOV**

Esta dissertação foi julgada adequada para obtenção do título de mestre e aprovada em sua forma final pelo Programa de Pós-Graduação em Ciência da Computação.

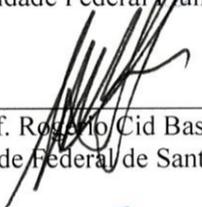
Florianópolis, 9 de Março de 2016.

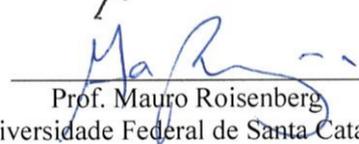

Prof.^a Carina Friedrich Dorneles, Dr.^a.
Coordenadora do Programa

Banca Examinadora:


Prof. Paulo José de Freitas Filho
Universidade Federal de Santa Catarina
Orientador


Prof. Carlos Magno Couto Jacinto
Universidade Federal Fluminense


Prof. Rosário Cid Bastos
Universidade Federal de Santa Catarina


Prof. Mauro Roisenberg
Universidade Federal de Santa Catarina

AGRADECIMENTOS

Agradeço à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior pelo recurso do Programa de Demanda Social (CAPES/DS) que fomentou esta pesquisa aliada ao Programa de Pós-Graduação em Ciência da Computação da UFSC; à todos os colegas do programa, professores, que compartilharam críticas e sugestões no decorrer desta pesquisa.

Agradeço o apoio e a paciência de meus familiares e amigos mais próximos; por compartilhar as comemorações da vida.

Agradeço à Adri, querida, por tudo o que ela já sabe e que não caberia no restante desta página.

RESUMO

O Modelo de Bourgoyne e Young (BYM) é utilizado para determinar a taxa de penetração no processo de perfuração de poços de petróleo. Para tanto é necessário que esteja parametrizado por coeficientes que devem ser estimados a partir de experiências anteriores. Por se tratar de um processo físico, a operação de perfuração pode apresentar medições ruidosas e o modelo em questão naturalmente pode não representá-la corretamente. Neste trabalho os coeficientes do BYM são determinados como distribuições de probabilidade, ao invés de valores fixos, propagando essas incertezas presentes nos dados e no próprio modelo. Para isso, é descrito um método que realiza inferência Bayesiana desses coeficientes através de Monte Carlo em Cadeias de Markov. Os resultados foram satisfatórios e as distribuições de probabilidade obtidas permitem uma melhor compreensão de como diferentes coeficientes atuam no resultado das simulações. Ao fim é feita uma análise sobre as possibilidades em aberto de melhorias para o método proposto.

Palavras-chave: Incerteza paramétrica. Modelo de Bourgoyne e Young. MCMC. Monte Carlo. ROP.

ABSTRACT

The Bourgoyne and Young Model (BYM) is used to determine the rate of penetration in oil well drilling processes. To achieve this the model must be parameterized with coefficients that are estimated on the basis of prior experience. Since drilling is a physical process, measurement data may include noise and the model may naturally fail to represent it correctly.

In this study the BYM coefficients are determined in the form of probability distributions, rather than fixed values, propagating the uncertainties present in the data and the model itself. This work therefore describes a method that performs a Bayesian inference conducted by a Monte Carlo Markov Chain algorithm. The results were satisfactory and the probability distributions obtained offer improved insight into the influence of different coefficients on the simulation results. At last, a discussion is carried of open possibilities for improving the proposed method.

Palavras-chave: Parameter uncertainty. Bourgoyne and Young Model. MCMC. Monte Carlo. ROP.

LISTA DE ILUSTRAÇÕES

Figura 1 – Funções de verossimilhança	38
Figura 2 – Passo-a-passo do GLUE	43
Figura 3 – Exemplo de resultado de uma simulação	45
Figura 4 – Exemplo de um parâmetro operacional e seu valor médio	47
Figura 5 – Exemplo de um parâmetro operacional ajustado por polinômio	48
Figura 6 – Aplicação da medida de verossimilhança dos dados ob- servados	50
Figura 7 – Efeito do desvio padrão na medida de verossimilhança .	51
Figura 8 – Grafo do modelo probabilístico	53
Figura 9 – Configuração inicial da função de transição	54
Figura 10 – Resumo do método	55
Figura 11 – Primeiras 1500 iterações para a Perfuração B	59
Figura 12 – Perfurações A e B: valores das variáveis observadas nos dados e os utilizados na simulação	60
Figura 13 – Perfurações C e D: valores das variáveis observadas nos dados e os utilizados na simulação	61
Figura 14 – Perfuração A: traços das cadeias.	62
Figura 15 – Perfuração B: traços das cadeias.	62
Figura 16 – Perfuração C: traços das cadeias.	63
Figura 17 – Perfuração D: traços das cadeias.	63
Figura 18 – Perfuração A: distribuições à posteriori dos parâmetros.	64
Figura 19 – Perfuração B: distribuições à posteriori dos parâmetros.	64
Figura 20 – Perfuração C: distribuições à posteriori dos parâmetros.	65
Figura 21 – Perfuração D: distribuições à posteriori dos parâmetros.	65
Figura 22 – Resultado das simulações do BYM	67

LISTA DE TABELAS

Tabela 1 – Valor padrão dos coeficientes	46
Tabela 2 – Intervalo dos parâmetros na distribuição à priori.	52
Tabela 3 – Parametrização das variáveis operacionais e observadas nas perfurações.	58
Tabela 4 – Medidas de erro	66
Tabela 5 – Intervalos de credibilidade inferidos para os coeficientes	68

LISTA DE ABREVIATURAS

ABC	–	Computação Bayesiana Aproximada	25
AG	–	Algoritmo Genético	21
BYM	–	Modelo de Bourgoyne e Young	22
GLUE	–	Estimação de incerteza de verossimilhança generalizada	24
HPD	–	Máxima Densidade à Posteriori	66
MAPE	–	Erro Médio Percentual Absoluto	66
MCMC	–	Monte Carlo em Cadeias de Markov	25
ODE	–	Equação Diferencial Ordinária	23
RMSE	–	Erro Quadrático Médio	66
RNA	–	Rede Neural Artificial	21
ROP	–	Taxa de Penetração	21
RPM	–	Rotações por Minuto	21
RWM	–	<i>Random Walk Metropolis</i>	52
SMC	–	Monte Carlo Sequencial	35
WOB	–	Peso Sobre a Broca	21

LISTA DE SÍMBOLOS

D	– Profundidade	22
d_b	– Diâmetro da broca	29
F_j	– Força de impacto do jato	30
g_p	– Gradiente de pressão dos poros	29
h	– Desgaste dos dentes da broca	30
N	– Rotação da broca	30
ρ_c	– Densidade equivalente de lama	29
τ_H	– Constante de abrasividade	31
W	– Peso sobre broca	29
$\left(\frac{W}{d_b}\right)_t$	– Limiar de peso de perfuração	29
$\pi(\theta)$	– Distribuição de probabilidade à priori	32
$L(\tilde{y} \theta)$	– Função de verossimilhança	32
$f(\theta \tilde{y})$	– Distribuição de probabilidade à posteriori	32
$q(\cdot \theta)$	– Função de transição	34

SUMÁRIO

1	INTRODUÇÃO	21
1.1	Justificativa e motivação	23
1.2	Objetivo geral	26
1.3	Objetivos específicos	26
1.4	Organização do texto	26
2	REVISÃO BIBLIOGRÁFICA	29
2.1	O modelo de Bourgoyne e Young	29
2.2	Inferência Bayesiana	32
2.2.1	Computação Bayesiana Aproximada (ABC)	32
2.2.2	Estimação de incerteza de verossimilhança generalizada (GLUE)	36
2.2.2.1	Funções de verossimilhança	37
2.2.2.2	Funcionamento do <i>framework</i>	42
3	DESENVOLVIMENTO DO MÉTODO	45
3.1	A solução do BYM	45
3.2	O modelo probabilístico	48
3.2.1	Monte Carlo em Cadeias de Markov (MCMC)	52
4	EXPERIMENTOS E RESULTADOS	57
5	CONCLUSÃO	69
5.1	Comentários e trabalhos futuros	69
	Bibliografia	75

1 INTRODUÇÃO

A perfuração de poços de petróleo é uma atividade que envolve muitos riscos. A otimização do processo de perfuração consiste normalmente em aumentar a taxa de penetração (ROP, *Rate of Penetration*) em um ambiente delimitado por custos financeiros e limites físicos.

Otimizar a ROP envolve a compreensão de uma série de parâmetros operacionais como por exemplo a vazão e pressão de fundo de poço (ligados à limpeza e segurança do poço), peso sobre a broca (WOB, *Weight On Bit*) e rotação da broca (RPM). Quanto maior o peso e a rotação da broca maior será a ROP. Entretanto, o aumento desses parâmetros pode também significar um desgaste excessivo da broca. Levando em conta que os maiores gastos estão associados ao aluguel dos equipamentos de operação, a manobra para a troca de uma broca pode ser considerada uma operação bastante dispendiosa. Para maiores detalhes sobre a segurança e as relações entre os parâmetros de perfuração, ver GANDELMAN (2012).

Assim, a atividade de perfuração de um poço deve ser bem planejada para que ocorra de forma segura e em um tempo ótimo. Para este efeito diversos modelos matemáticos e/ou computacionais (citados adiante) têm sido desenvolvidos. Alguns podem ser aplicados com a finalidade de estimar a ROP em atividades de análise e planejamento, outros com a finalidade na adequação dos parâmetros operacionais de perfuração em tempo real. Nesta linha pode-se citar a própria dissertação de GANDELMAN (2012) que emprega redes neurais artificiais (RNAs) para a estimação dos parâmetros WOB e RPM em tempo real.

A utilização de RNAs como ferramenta para construção de um modelo não-linear entre o ambiente de perfuração e a ROP é bastante comum há algum tempo. AREHART (1990), BILGESU; TETRICK et al. (1997), BILGESU; ALTMIS et al. (1998), DASHEVSKIY; DUBINSKY; MACPHERSON (1999) e GIDH; IBRAHIM (2011) e os trabalhos mais recentes as têm sofisticado ainda mais: EDALATKHAH; RASOUL; HASHEMI (2010) desenvolvem um sistema de controle da ROP baseado em RNAs e algoritmos genéticos (AGs), enquanto RO-

DRIGUES (2014) e RODRIGUES et al. (2014) propõem um novo modelo de rede neuro-fuzzy que permite manter a interpretabilidade da ROP calculada.

Em paralelo a isso existem ainda modelos matemáticos para predição da ROP destacados por BARRAGAN (1995) e GANDELMAN (2012). Dentre eles o mais aceito é o Modelo de Bourgoyne e Young (BYM, *Bourgoyne and Young Model*) (BOURGOYNE; YOUNG, 1974; BOURGOYNE; CHENEVERT et al., 1986) pois é o que considera a maior quantidade de parâmetros operacionais (FONSECA et al., 2006; EDALATKHAH; RASOUL; HASHEMI, 2010) e ainda é amplamente utilizado (MORADI et al., 2010).

BOURGOYNE; YOUNG (1974) observaram que a relação entre a ROP com os demais parâmetros operacionais se aproxima ao de uma curva exponencial. Propuseram, portanto, um modelo apresentado na equação (1.1) em que cada parâmetro operacional contribui com um certo “peso” para a formação desta curva.

$$\text{ROP} = \frac{dD}{dt} = \exp \left(a_1 + \sum_{i=2}^8 a_i x_i \right) \quad (1.1)$$

Estes “pesos”, (a_1, \dots, a_8) conhecidos como coeficientes do modelo de Bourgoyne e Young, são dependentes das características do solo e devem ser determinados a partir de experiências análogas de perfuração.

Junto com a divulgação do modelo, BOURGOYNE; YOUNG (1974) sugeriram que os coeficientes deveriam ser determinados através de regressão múltipla. Entretanto, este processo se mostrou passível de gerar valores negativos ou zero para os coeficientes e, portanto, fisicamente insignificantes (MORADI et al., 2010). Por exemplo, se um determinado coeficiente é zero, significa que o aumento do peso sobre a broca não afeta a ROP, ou até mesmo a diminuiria se o coeficiente fosse negativo. Em vista disso alguns trabalhos como BAHARI et al. (2008) e HASAN; ABOOZAR; HAMIDREZA (2011) têm estudado maneiras alternativas de os determinar.

Sob o ponto de vista de KENNEDY; O'HAGAN (2001), a determinação dos coeficientes do BYM pode ser vista como um problema de calibração do próprio modelo. O processo de calibração consiste, vagamente, em ajustar os parâmetros do modelo até as saídas se assemelharem aos dados. O trabalho de BAHARI et al. (2008) trata a calibração do BYM através de um processo de otimização por AGs. Já, HASAN; ABOOZAR; HAMIDREZA (2011), avançam na proposta e, identificando os coeficientes obtidos por AG, desmembram cada termo da somatória da equação (1.1) como entrada de uma RNA em que a saída é a ROP.

1.1 JUSTIFICATIVA E MOTIVAÇÃO

O que é comum nos trabalhos de BAHARI et al. (2008) e HASAN; ABOOZAR; HAMIDREZA (2011) é que eles tratam do problema de calibração como uma simples regressão, deixando para trás a dinâmica inerente do sistema. O BYM inclui uma equação que modela o desgaste da broca em função do tempo – fornecendo o valor de h no sétimo termo, $a_7(-h)$, da equação (1.1) – e que os trabalhos supracitados desconsideram. Por consequência, o conjunto de dados utilizados pelos mesmos se restringe àquele onde a broca ainda é considerada nova, como no início da perfuração ($h = 0$), ou constatada como totalmente desgastada ($h = 1$) ao retirá-la do poço.

Além disso, é sensato pensar num ambiente de perfuração onde os sensores produzem informações potencialmente ruidosas. Por consequência, o valor das variáveis informativas do sistema pode ser resultado de cálculos somente aproximados. Assim os valores de ROP observados incluem incertezas que são dificilmente identificadas através das variáveis.

O BYM é um sistema de equações diferenciais ordinárias (ODEs, *Ordinary Differential Equations*), portanto, determinístico. Entretanto o processo de perfuração, por se tratar de um processo físico, pode conter elementos inerentemente imprevisíveis pelo modelo. Essa imperfeição gera uma variabilidade residual no modelo (KENNEDY; O'HAGAN,

2001) que abre margem para uma equifinalidade (BEVEN; BINLEY, 1992) de modelos e variáveis. Isto significa que diferentes parametrizações devem ser consideradas ao invés de uma única.

Uma maneira de lidar com a incerteza dos parâmetros é enxergá-los sob um prisma Bayesiano, em que seus valores, antes fixos θ , são agora uma função de densidade de probabilidade $f(\theta)$. A interpretação Bayesiana da probabilidade é epistêmica, o que significa que a densidade de probabilidade $f(\theta)$ representa o nosso conhecimento e incerteza sobre os valores θ . (JAYNES, 2003)

A calibração do BYM por uma abordagem Bayesiana consiste, portanto, na inferência dos valores de uma distribuição de probabilidade à posteriori $f(\theta | \tilde{y})$ dos parâmetros $\theta = (a_1, a_2, \dots, a_8)$ dada as observações de profundidade $\tilde{y} = \{\tilde{D}_t | t = 1, \dots, n\}$ em um vetor discreto de tempo t .

De acordo com o teorema de Bayes (GELMAN et al., 2009),

$$f(\theta | \tilde{y}) \propto L(\tilde{y} | \theta) \pi(\theta).$$

Aqui, $L(\tilde{y} | \theta)$ é conhecida como a verossimilhança dos parâmetros, enquanto $\pi(\theta)$ é a distribuição à priori dos mesmos.

Normalmente a função de verossimilhança não está disponível de forma fechada e o seu cálculo é computacionalmente inviável (FERNHEAD; PRANGLE, 2012). Porém quando o modelo pode ser trivialmente simulado (resolvido numericamente, no caso dos sistemas de ODEs), podemos então criar um modelo probabilístico sobre o erro entre os valores simulados, y , e os observados, \tilde{y} . Então a distribuição à posteriori dos parâmetros, $f(\theta | \tilde{y})$, pode ser inferida a partir das simulações que, com determinada parametrização, fornecem resultados “próximos” aos observados.

Este modelo de aproximação entre resultados observados e simulados pode ser definido por uma verossimilhança subjetiva que serve meramente como uma medida do grau de ajuste entre os dados. Neste caso, a teoria utilizada para a calibração do modelo simulado é conhecida como Estimção de incerteza de verossimilhança generalizada (GLUE, *Generalized likelihood uncertainty estimation*) (BEVEN; BIN-

LEY, 1992; BEVEN; FREER, 2001). Muitas críticas têm sido feitas por GLUE não utilizar valores de verossimilhança consistentes estatisticamente (MCMICHAEL; HOPE; LOAICIGA, 2006; MONTANARI, 2005), mas por outro lado, é um *framework* ainda bem aplicado com sucesso principalmente para avaliação de incerteza em modelos ambientais. (DELSMAN et al., 2013; SUN; HONG; HALL, 2014; ALAZZY; LÜ; ZHU, 2015; ZHANG; LI, 2015)

Existe também o caso em que nem mesmo uma verossimilhança subjetiva é utilizada; a verificação do erro de uma simulação é feita diretamente sobre os dados \tilde{y} e y . Essa ideia – de efetuar uma calibração “livre do cálculo da verossimilhança” – data de pelo menos RUBIN (1984) e foi aplicada inicialmente por PRITCHARD et al. (1999) como Computação Bayesiana Aproximada (ABC, *Approximate Bayesian Computation*).

Desde então, ABC tem evoluído e servido como um *framework* para a calibração de diversos modelos em genética (SIEGMUND; MARJORAM; SHIBATA, 2008; FOLL; BEAUMONT; GAGGIOTTI, 2008), epidemiologia (BLUM; FRANÇOIS, 2010; TANAKA et al., 2006), biologia populacional (RATMANN; JØRGENSEN et al., 2007; J. M. CORNUET et al., 2008), inferência de extremos (BORTOT; COLES; SISSON, 2007), também em sistemas dinâmicos (TONI et al., 2009), entre outros.

Recentemente, NOTT; MARSHALL; BROWN (2012) e VRUGT; SADEGH (2013) demonstraram a equivalência entre ABC e GLUE o que permitiu uma “ponte teórica” entre experiências realizadas em linhas de pesquisa distintas. Em ambos *frameworks*, os métodos de amostragem utilizados são os baseados em Monte Carlo, destacando-se o Monte Carlo em Cadeias de Markov (MCMC, *Monte Carlo Markov Chain*) (GELMAN et al., 2009) como o mais comum em GLUE, e os de Monte Carlo Sequencial e Populacional em ABC (BEAUMONT; J.-M. CORNUET et al., 2009).

1.2 OBJETIVO GERAL

Dispondo da capacidade de calibrar modelos computacionais através dos *frameworks* apresentados, o objetivo deste trabalho é especificar um método para estimar os coeficientes do Modelo de Bourgoyne e Young (BYM) através de Monte Carlo em Cadeias de Markov (MCMC), que permita a identificação e a avaliação das incertezas presentes no modelo.

1.3 OBJETIVOS ESPECÍFICOS

A potencialidade do BYM depende de como as constantes físicas e os parâmetros operacionais são fornecidos para o modelo. Portanto um dos objetivos específicos é:

- Determinar uma forma de representar os parâmetros operacionais obtidos nos dados de maneira com que eles possam ser usados durante a simulação do modelo.

Além disso, é necessário inferir os coeficientes e permitir a avaliação da incerteza do modelo. Assim, outros objetivos específicos consistem em:

- Definir um modelo para a medida de adequação dos dados simulados aos dados reais.
- Determinar uma parametrização para o algoritmo MCMC específico ao espaço amostral do problema.

1.4 ORGANIZAÇÃO DO TEXTO

O restante deste texto está organizado da seguinte forma: o Capítulo 2 faz um detalhamento do BYM com o conteúdo pertinente à este trabalho, depois segue fazendo um resgate das teorias comentadas aqui na introdução – durante este resgate são comentados os avanços mais recentes da área; o Capítulo 3 descreve amplamente o método proposto e detalha também como foram tratados os objetivos específicos; o

Capítulo 4 apresenta os experimentos e comenta de maneira sucinta os resultados enquanto o Capítulo 5 aprofunda na discussão dos detalhes que surgiram ao longo deste trabalho e sugere temas para trabalhos futuros.

2 REVISÃO BIBLIOGRÁFICA

2.1 O MODELO DE BOURGOYNE E YOUNG

A equação (1.1) foi posteriormente reescrita em BOURGOYNE; CHENEVERT et al. (1986) como um produto das exponenciações

$$\frac{dD}{dt} = f_1 \times f_2 \times f_3 \times f_4 \times f_5 \times f_6 \times f_7 \times f_8 \quad (2.1)$$

onde,

$$f_1 = e^{a_1} = K \quad (2.2a)$$

$$f_2 = e^{a_2 (10000-D)} \quad (2.2b)$$

$$f_3 = e^{a_3 D^{0.69} (g_p - 9)} \quad (2.2c)$$

$$f_4 = e^{a_4 D (g_p - \rho_c)} \quad (2.2d)$$

$$f_5 = \left[\frac{\frac{W}{d_b} - \left(\frac{W}{d_b}\right)_t}{4 - \left(\frac{W}{d_b}\right)_t} \right]^{a_5} \quad (2.2e)$$

$$f_6 = \left(\frac{N}{60}\right)^{a_6} \quad (2.2f)$$

$$f_7 = e^{-a_7 h} \quad (2.2g)$$

$$f_8 = \left(\frac{F_j}{1000}\right)^{a_8} . \quad (2.2h)$$

Originalmente os coeficientes de a_1 a a_4 são multiplicados por 2.303 para transformar seus respectivos termos em potências de base 10 ao invés de base e . Nessas equações,

a_1 a a_8 = os coeficientes que devem ser escolhidos de acordo com as perfurações anteriores;

D = profundidade vertical total – *Depth* (ft);

g_p = gradiente de pressão dos poros da rocha (lbm/gal);

ρ_c = densidade equivalente de lama (lbm/gal);

W = peso sobre a broca (1000 lbf);

d_b = diâmetro da broca (in);

$(W/d_b)_t$ = limiar de peso sobre a broca em que ela começa a perfurar;

N = velocidade de rotação (rpm);

h = desgaste fracionário dos dentes da broca, com $h = 0$ quando não há desgaste;

F_j = força de impacto do jato (lbf).

As funções f_2 e f_3 modelam o efeito da compactação da rocha na determinação da ROP. É importante perceber que a multiplicação dessas funções ($f_2 f_3$) é igual a 1.0 para quando o gradiente de pressão dos poros da rocha equivale à 9.0 lbm/gal numa profundidade de 10000 ft.

A função f_4 modela o *overbalance* – que é “um diferencial positivo de pressão entre o poço e as formações perfuradas” (GANDELMAN, 2012). Seu valor é igual a 1.0 quando a pressão da formação rochosa é igual a do fundo do poço.

As funções f_5 e f_6 modelam a influência que o peso sobre a broca e a rotação da mesma têm sobre a ROP. Como as funções anteriores, f_5 também pode assumir valor 1.0, quando (W/d_b) for 4000 lbf/in; f_6 assume valor 1.0 quando velocidade de rotação da broca for 60 rpm. Já o limiar de peso sobre a broca para perfuração pode ser considerado com valor 0 nas formações macias.

A função f_7 modela a influência que o desgaste da broca tem sobre a ROP. Este termo tem valor igual a 1.0 quando o desgaste, h , for nulo.

A função f_8 modela a influência da força de impacto do jato, F_j , na ROP, que quando assume valor de 1000 lbf normaliza este termo em 1.0.

A função que resta, f_1 , indica a resistência da formação rochosa. Este termo tem a mesma unidade da ROP e é conhecido como constante de perfurabilidade (K). É numericamente igual à ROP do sistema quando operado e observado nas condições que tornam todos os outros termos, de f_2 a f_8 , iguais a 1 e, em consequência, independente dos coeficientes. É um termo, então, que normaliza as funções restantes e deve ser calculado a partir de dados obtidos nas perfurações de poços similares.

BOURGOYNE; CHENEVERT et al. (1986) apresentaram também a seguinte ODE para modelar o desgaste da broca, h , na função f_7 :

$$\frac{dh}{dt} = \frac{1}{\tau_H} \left(\frac{N}{60} \right)^{H_1} \left[\frac{\left(\frac{W}{d_b} \right)_m - 4}{\left(\frac{W}{d_b} \right)_m - \left(\frac{W}{d_b} \right)} \right] \left(\frac{1 + H_2/2}{1 + H_2h} \right) \quad (2.3)$$

Nesta equação,

t = tempo (hr);

H_1, H_2 = constantes;

$(W/d_b)_m$ = constante. Relação do peso máximo suportado pela broca em seu diâmetro;

τ_H = constante de abrasividade da formação rochosa (hr).

Para parametrizar tal equação, BOURGOYNE; CHENEVERT et al. (1986, p. 218) fornecem uma tabela com os valores sugeridos das constantes H_1, H_2 e $(W/d_b)_m$ de acordo com cada tipo de broca.

A constante τ_H tem a mesma função de normalização que o coeficiente a_1 em (1.1). Significa o tempo em horas requerido para desgastar completamente uma determinada broca dada as condições de operação constante de $W/d_b = 4000$ lbf/in e $N = 60$ rpm. Para encontrá-la, os autores integraram (2.3) em relação a t , e obtiveram

$$\tau_H = \frac{t_b}{J_2 \left(h_f + H_2 h_f^2 / 2 \right)}$$

$$J_2 = \left(\frac{60}{N} \right)^{H_1} \left[\frac{\left(\frac{W}{d_b} \right)_m - \left(\frac{W}{d_b} \right)}{\left(\frac{W}{d_b} \right)_m - 4} \right] \left(\frac{1}{1 + H_2/2} \right) \quad (2.4)$$

onde,

t_b = tempo necessário para desgastar a broca (hr);

h_f = proporção de desgaste final observada, com $h_f = 1$ quando o desgaste é total.

2.2 INFERÊNCIA BAYESIANA

Retomando ao Teorema de Bayes (GELMAN et al., 2009),

$$f(\boldsymbol{\theta} | \tilde{y}) = \frac{L(\tilde{y} | \boldsymbol{\theta})\pi(\boldsymbol{\theta})}{L(\tilde{y})} \quad (2.5)$$

onde $L(\tilde{y}) = \int L(\tilde{y} | \boldsymbol{\theta})\pi(\boldsymbol{\theta})d\boldsymbol{\theta}$ é uma constante de normalização. A distribuição de probabilidade a priori $\pi(\boldsymbol{\theta})$, definida em Θ , representa o conhecimento inicial existente sobre o vetor de parâmetros $\boldsymbol{\theta}$; a função de verossimilhança $L(\tilde{y} | \boldsymbol{\theta})$ representa o quão provável são os dados observados dada as saídas de um modelo \mathcal{M} parametrizado por $\boldsymbol{\theta}$.

A maneira mais simples de gerar observações de $f(\boldsymbol{\theta} | \tilde{y})$ é através do método de rejeição:

REJ1 Gerar $\boldsymbol{\theta}$ a partir de $\pi(\cdot)$

REJ2 Aceitar $\boldsymbol{\theta}$ com probabilidade $\alpha = L(\tilde{y} | \boldsymbol{\theta})$; ir para REJ1.

Entretanto, o cálculo da verossimilhança $L(\tilde{y} | \boldsymbol{\theta})$ pode não ser trivial ou até mesmo inexistente. No caso da calibração de modelos numéricos, o passo REJ2 se desdobra em simular o modelo com os parâmetros obtidos em REJ1 e aceitá-los caso o resultado da simulação for suficientemente “bom” em relação aos dados observados.

2.2.1 Computação Bayesiana Aproximada (ABC)

A literatura caracteriza ABC como um *framework* de métodos “livres de verossimilhança”. Uma variação do algoritmo de rejeição, REJ, quando a verossimilhança não pode ser calculada explicitamente é a seguinte:

ABC-REJ1 Gerar $\boldsymbol{\theta}$ a partir de $\pi(\cdot)$

ABC-REJ2 Simular y a partir do modelo \mathcal{M} , parametrizado por $\boldsymbol{\theta}$

ABC-REJ3 Aceitar $\boldsymbol{\theta}$ se $y = \tilde{y}$; ir para ABC-REJ1.

Assume-se que o modelo \mathcal{M} seja fácil de se simular sob um conjunto de parâmetros. Este algoritmo não é formalmente conhecido

como um método ABC, no entanto foi utilizado por TAVARÉ et al. (1997) e produz uma estimativa imparcial (não tendenciosa) da distribuição à posteriori dos parâmetros $f(\boldsymbol{\theta} | \tilde{y})$; e quando o número de iterações cresce assintoticamente obtém-se, então, uma estimativa consistente.

Entretanto, como comentado na seção 1.1, a probabilidade de se obter valores simulados exatos aos fornecidos é praticamente nula na maioria dos modelos. Assim, a ideia que definiu o termo ABC, aplicada por PRITCHARD et al. (1999), foi a de substituir a condição de aceitação exata em ABC-REJ3 por

$$\rho(y, \tilde{y}) \leq \varepsilon, \quad (2.6)$$

sendo ρ uma função de distância e ε uma tolerância. A escolha de ε reflete a tensão entre computabilidade e precisão: se $\varepsilon \rightarrow \infty$, o teste de aceitação repassa os valores gerados da distribuição à priori $\pi(\boldsymbol{\theta})$; se $\varepsilon = 0$, só serão aceitos os valores $y = \tilde{y}$. Dependendo da escolha de d e ε , pode-se dizer que os valores aceitos são independentes e uniformemente distribuídos a partir de $f(\boldsymbol{\theta} | \rho(y, \tilde{y}) \leq \varepsilon)$, normalmente conhecida na literatura como distribuição à posteriori do ABC (*ABC posterior*) $f_{abc}(\boldsymbol{\theta} | \tilde{y})$. Se ε for suficientemente pequeno então $f_{abc}(\boldsymbol{\theta} | \tilde{y}) \approx f(\boldsymbol{\theta} | \tilde{y})$.

Quando \tilde{y} é multi-dimensional, por exemplo, pode ser difícil determinar um cálculo adequado da distância d . Neste caso, é aconselhável utilizar uma nova distância aplicada sobre uma estatística resumitiva $S(\tilde{y})$ e $S(y)$ dos dados – caso em que $f_{abc} \equiv f(\boldsymbol{\theta} | d(S(y), S(\tilde{y})) \leq \varepsilon)$. A escolha dessa função S também é uma tarefa delicada pois ela deve capturar as informações sobre $\boldsymbol{\theta}$. Muitos trabalhos têm analisado este problema desde BEAUMONT; ZHANG; BALDING (2002) até os recentes como FEARNHEAD; PRANGLE (2012), RATMANN; CAMACHO et al. (2013), RULI; SARTORI; VENTURA (2014) e CABRAS; NUEDA; RULI (2015).

A desvantagem desse método de rejeição é que, mesmo bem escolhidas as métricas de distância e estatística resumitiva, a aceitação será baixa quando a distribuição à priori dos parâmetros for muito dife-

rente de à posteriori. Uma maneira de reduzir este problema é realizar a inferência das proposições através de MCMC.

O algoritmo clássico de Metropolis-Hastings (MH) (HASTINGS, 1970; METROPOLIS et al., 1953) consiste em gerar observações de $f(\theta | \tilde{y})$ a partir de uma Cadeia de Markov:

MH1 Dado o estado atual com θ , propôr uma mudança para θ^* de acordo com uma função de transição $q(\cdot | \theta)$.

MH2 Calcular

$$\alpha = \min \left(1, \frac{L(\tilde{y} | \theta^*)\pi(\theta^*)q(\theta^* | \theta)}{L(\tilde{y} | \theta)\pi(\theta)q(\theta | \theta^*)} \right) \quad (2.7)$$

MH3 Ir para estado em θ^* com probabilidade α , senão permanecer com θ ; ir para MH1.

O algoritmo constrói uma Cadeia de Markov que, após um número incerto de iterações, torna-se ergódica (ANDRIEU; FREITAS et al., 2003) e passa a oferecer, a partir de sua distribuição estacionária, observações da distribuição f .

É importante observar que a comparação feita no método de rejeição, em REJ2, é sempre em relação à $L(\tilde{y} | \theta)$ – uma comparação global. Desta forma os valores inferidos são estatisticamente independentes. Já em MH2, $L(\tilde{y} | \theta^*)$ é comparado com a verossimilhança no estado atual $L(\tilde{y} | \theta)$. Essa comparação local feita no MCMC garante maior aceitação das proposições porém gera resultados não-independentes.

Entretanto, o termo $L(\tilde{y} | \theta)$ ainda corresponde a uma medida de verossimilhança. No caso do *framework* GLUE, que será melhor detalhado adiante, uma das tarefas do especialista consiste em defini-la de maneira subjetiva. Tendo em mãos esta função, a aplicação do algoritmo MH é direta. Caso contrário, MARJORAM et al. (2003) propuseram uma versão que não necessita do cálculo da verossimilhança:

ABC-MCMC1 Dado o estado atual com θ , propôr uma mudança para θ^* de acordo com uma função de transição $q(\cdot | \theta)$.

ABC-MCMC2 Simular y a partir do modelo \mathcal{M} , parametrizado por θ .

ABC-MCMC₃ Se $y = \tilde{y}$ ir para ABC-MCMC₄, senão permanecer com θ e retornar para ABC-MCMC₁.

ABC-MCMC₄ Calcular

$$\alpha = \min\left(1, \frac{\pi(\theta^*)q(\theta^* | \theta)}{\pi(\theta)q(\theta | \theta^*)}\right)$$

ABC-MCMC₅ Ir para estado em θ^* com probabilidade α , senão permanecer com θ ; ir para ABC-MCMC₁.

Como discutido anteriormente, a condição de aceitação em ABC-MCMC₃ pode ser impraticável. Por isso é comum a substituição neste passo, por um teste de aproximação entre estatísticas aplicadas aos dados:

ABC-MCMC'₃ Se $d(S(y), S(\tilde{y})) \leq \varepsilon$ ir para ABC-MCMC₄, senão permanecer com θ e retornar para ABC-MCMC₁.

É importante perceber que quando a função de transição (também chamada de distribuição de proposição) é simétrica, $q(\theta | \theta^*) = q(\theta^* | \theta)$, então α dependerá somente da distribuição à priori. Além disso, se a distribuição à priori for uniforme, então $\alpha = 1$ e o algoritmo se torna o de rejeição simples (ABC-REJ).

Devido à perda da verificação de um valor de verossimilhança dos parâmetros, o algoritmo de MCMC aplicado em ABC acrescenta pouca informação para a escolha dos valores inferidos se o especialista tem pouco conhecimento à priori dos parâmetros. Também é uma tarefa delicada escolher uma distribuição de proposição que explore o espaço paramétrico, Θ , de maneira abrangente (tenha uma boa mixagem). Se $q(\cdot | \cdot)$ sugere grandes mudanças, há maior probabilidade de rejeição de suas proposições; já, em pequenos passos, as amostras são altamente correlacionadas e há maior possibilidade de prender-se em mínimos locais.

Em vista dessas dificuldades, outras técnicas de inferência têm ganhado popularidade em ABC. A mais notória é o uso de técnicas de Monte Carlo Sequencial (SMC, *Sequential Monte Carlo*) por DEL MORAL; DOUCET; JASRA (2006) e SISON; FAN; TANAKA (2007),

melhorado em TONI et al. (2009) e em BEAUMONT; J.-M. CORNUET et al. (2009) com Monte Carlo Populacional; e mais recentemente por DEL MORAL; DOUCET; JASRA (2012), LEE (2012) e ŁATUSZYŃSKI; LEE (2014).

Outra parte importante de pesquisa recai sobre a verificação do ajuste entre os valores observados e simulados, $d(S(y), S(\tilde{y})) \leq \varepsilon$. Como comentado anteriormente é preciso que boas estatísticas S sejam definidas. Porém, o dilema computabilidade vs precisão faz com que a escolha de um valor de ε também seja importante. Percebendo que o algoritmo ABC-MCMC poderia ter mais aceitações quanto maior o valor de ε , BORTOT; COLES; SISSON (2007) propuseram acrescentá-lo ao espaço paramétrico de $L(\tilde{y} | \theta)$. Isso resulta em uma Cadeia de Markov nos pares $(\theta, \varepsilon) \in \Theta$ onde valores de θ simulados com valores pequenos de ε têm uma distribuição próxima à posteriori $f(\theta | \tilde{y})$. Maiores valores de ε facilitam a transição para valores de θ que não seriam possíveis com ε fixo, permitindo assim uma melhor mixagem.

Ao permitir que o erro faça parte do modelo, o especialista assume que a calibração irá realizar o ajuste correto de parâmetros, mas para o modelo errado. Como sugeriram BORTOT; COLES; SISSON (2007), uma maneira de abrandar essas observações errôneas é determinar um limiar ε_τ e filtrar a série $\{(\theta_i, \varepsilon_i)\}$ como $\{\theta_i : \varepsilon_i < \varepsilon_\tau\}$. Essa modelagem do erro dentro do *framework* ABC é assunto criticado por alguns (FEARNHEAD; PRANGLE, 2012; ANDRIEU; BARTHELMÉ et al., 2012), mas outros como WILKINSON (2013) acreditam que, quando bem modelado, o erro pode representar a nossa crença em relação às discrepâncias entre os valores simulados e a realidade.

2.2.2 Estimação de incerteza de verossimilhança generalizada (GLUE)

Da mesma maneira que o ABC, o GLUE também se insere na Calibração Bayesiana de modelos. Cunhado no trabalho de BEVEN; BINLEY (1992, p. 281, tradução nossa), “GLUE reconhece a equivalência ou quase equivalência de diferentes conjuntos de parametrizações na calibração de modelos. É baseado na realização de uma grande quan-

tidade de execuções de um dado modelo, com diferentes conjuntos de valores de parâmetros escolhidos de maneira aleatória a partir de uma distribuição específica. Em se tratando de comparar respostas previstas e observadas, a cada conjunto de valores de parâmetros é associado uma verossimilhança (*likelihood*) de serem simuladores do sistema”.

A maneira como os parâmetros do modelo interagem no GLUE não chega a ser um problema pois, de qualquer forma, ela está refletida implicitamente nos valores de verossimilhança (BEVEN; BINLEY, 2014). Esta é generalizada – ou sintética (WOOD, 2010) – no sentido de permitir o uso de uma gama de medições, ou combinação de medições, o que inclui não só a multiplicação Bayesiana mas também operadores utilizados para definir grau de pertinência em lógica *fuzzy*.

Não se deve confundir aqui os conceitos de probabilidade e grau de pertinência *fuzzy*. Estas funções, algumas apresentadas a seguir, servem como uma definição matemática das medidas de distância que comparam os valores simulados.

2.2.2.1 Funções de verossimilhança

A função de verossimilhança (ou medida de verossimilhança) é um ponto chave do framework GLUE. Entretanto, ela difere aqui das usadas em outros métodos de calibração de modelos e de estimação de incertezas.

Diferentes tipos de medidas existem, e todas com qualidades específicas. Além disso, não há regras definitivas para a escolha de uma determinada função de verossimilhança.

Algumas medidas derivadas da estatística e da lógica *fuzzy*, que podem ser vistas na Figura 1, estão detalhadas à seguir:

Gaussiana

Figura 1a, a função de verossimilhança Gaussiana é definida como

$$L(\tilde{y} | \theta) = \frac{1}{\sqrt{2\pi\hat{\sigma}^2}} \exp \left[-\frac{1}{2\hat{\sigma}^2} (\tilde{y} - y)^2 \right], \quad (2.8)$$

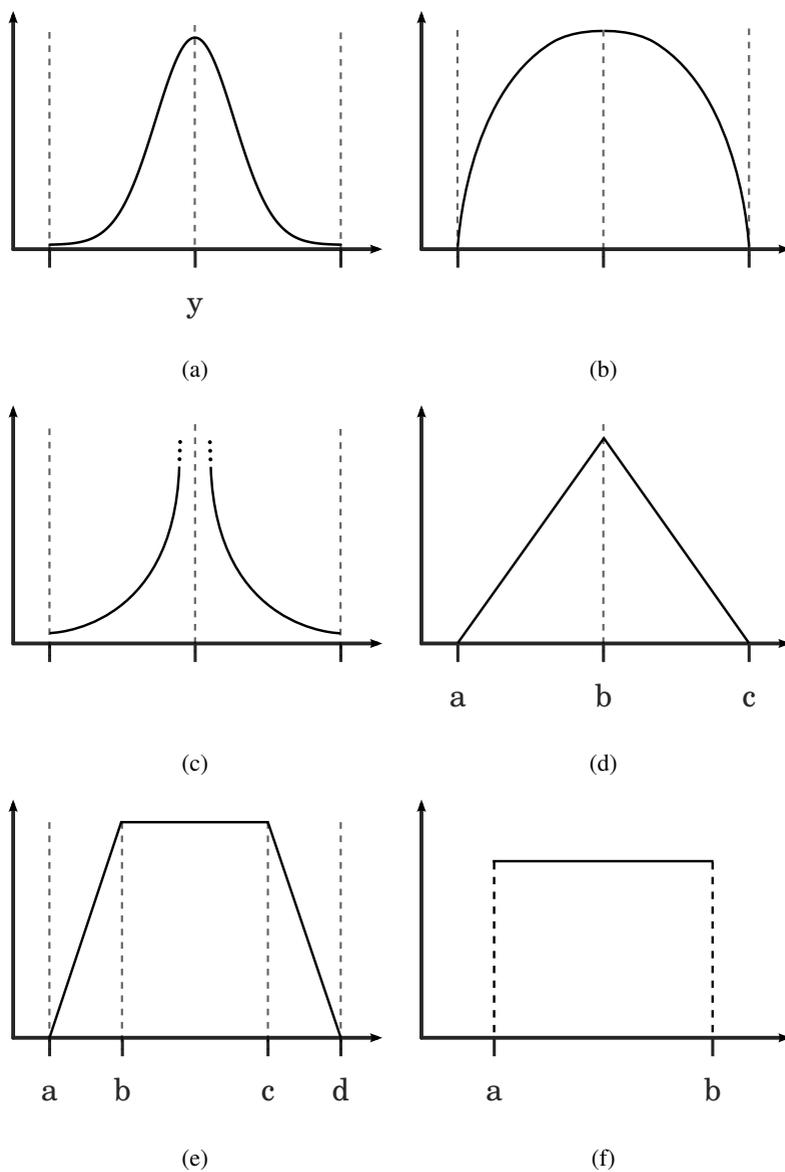


Figura 1 – Funções de verossimilhança: (a) Função Gaussiana, (b) Função de Eficiência do Modelo, (c) Função de Variância do Erro Invertida, (d) Função Triangular, (e) Função Trapezoidal, (f) Função Uniforme.

que é a mesma utilizada em outros *frameworks* estatísticos clássicos: Corresponde à função densidade de probabilidade.

Eficiência do Modelo (*model efficiency*)

Apresentada na Figura 1b e descrita por BEVEN; BINLEY (1992), é definida como

$$L(\tilde{y} | \theta) = \begin{cases} 1 - \sigma_{\varepsilon}^2 / \sigma_O^2 & \text{se } \sigma_O^2 < \sigma_{\varepsilon}^2 \\ 0 & \text{se } \sigma_{\varepsilon}^2 \geq \sigma_O^2 \end{cases} \quad (2.9)$$

onde,

$$\sigma_{\varepsilon}^2 = \frac{1}{n} (\tilde{y} - y)^T \mathbf{V} (\tilde{y} - y)$$

é a variância ponderada das diferenças entre os valores observados e simulados, σ_O^2 é a variância ponderada dos valores observados e \mathbf{V} , a matriz de pesos. Nesta medida a verossimilhança iguala a 1 (um) se não houver diferença entre valores observados e simulados, e a 0 (zero) se a variância ponderada das diferenças for maior que a variância ponderada das observações.

Variância do Erro Invertida (*Inverse error variance function*)

Ilustrada na Figura 1c, também sugerida por BEVEN; BINLEY (1992),

$$L(\tilde{y} | \theta) = (\sigma_{\varepsilon}^2)^{-N} \quad (2.10)$$

onde N é um fator de forma. Esta função concentra os pesos das melhores simulações ao passo que N aumenta de valor. Para $N \rightarrow \infty$, todo peso é dado unicamente à melhor simulação, enquanto que para os valores menores de N , todas as simulações tenderão a ter o mesmo peso.

Triangular

A Função Triangular, Figura 1d, definida como

$$L(\tilde{y} | \theta) = \begin{cases} 0 & \text{se } \tilde{y} < a \text{ ou } c < \tilde{y}, \\ \frac{\tilde{y}-a}{b-a} & \text{se } a \leq \tilde{y} < b, \\ \frac{c-\tilde{y}}{c-b} & \text{se } b \leq \tilde{y} \leq c, \end{cases} \quad (2.11)$$

deve ser parametrizada em torno do valor simulado, y , como por exemplo: $b = y$ e as variáveis restantes como uma proximidade relativa ao valor de y .

Trapezoidal

A Função Trapezoidal, Figura 1e, definida como

$$L(\tilde{y} | \theta) = \begin{cases} 0 & \text{se } \tilde{y} < a \text{ ou } d < \tilde{y}, \\ \frac{\tilde{y}-a}{b-a} & \text{se } a \leq \tilde{y} < b, \\ 1 & \text{se } b \leq \tilde{y} < c, \\ \frac{d-\tilde{y}}{d-c} & \text{se } c \leq \tilde{y} \leq d, \end{cases} \quad (2.12)$$

é também parametrizada com base no valor simulado como por exemplo y sendo a mediana da base do trapézio.

Uniforme

Por fim, a Função Uniforme, Figura 1f, é definida como

$$L(\tilde{y} | \theta) = \begin{cases} 1 & \text{se } a < \tilde{y} < b \\ 0 & \text{caso contrário.} \end{cases} \quad (2.13)$$

Alguns elementos devem ser considerados nestas medidas apresentadas. Dependendo da função, naturalmente existe um ponto de corte (ou rejeição) de onde em diante os valores de verossimilhança são zero. Em algumas funções, como por exemplo a Gaussiana, que é naturalmente definida em $(-\infty, +\infty)$, podem ser definidos limites artificiais, caso conveniente.

Além disso, as funções definidas pelas equações (2.9) e (2.10) retornam um único valor calculado sobre todos os valores simulados no modelo, enquanto as restantes são utilizadas para fornecer um único valor de verossimilhança para cada observação. Sendo assim, para efeito de comparação, é necessário que os valores de verossimilhança sejam agregados em um único valor global. Dentre as maneiras possíveis de agregação, podem ser destacados os operadores de agregação utilizados em lógica *fuzzy* dentre outros destacadas à seguir:

Multiplicação

$$L(\tilde{y} | \theta) = \prod_{t=1}^n L(\tilde{y}_t | \theta) \quad (2.14)$$

Este é um operador de agregação bastante restritivo se a função de verossimilhança utilizada tem um grande domínio que mapeia para o resultado zero. Neste caso, basta que somente um valor dos simulados não seja bom o suficiente (em relação ao observado correspondente) para que o resultado da função de verossimilhança agregada seja zero.

Mínimo

$$L(\tilde{y} | \theta) = \min_{t=1, \dots, n} L(\tilde{y}_t | \theta) \quad (2.15)$$

Também bastante restritivo no sentido de escolher o menor valor de verossimilhança calculado.

Máximo

$$L(\tilde{y} | \theta) = \max_{t=1, \dots, n} L(\tilde{y}_t | \theta) \quad (2.16)$$

Menos restritivo pois considera o valor de maior verossimilhança dentre os simulados e observados.

Média Aritmética Ponderada

$$L(\tilde{y} | \theta) = \frac{1}{n} \sum_{t=1}^n \omega_t L(\tilde{y}_t | \theta) \quad (2.17)$$

onde ω_t é o peso correspondente à observação \tilde{y}_t .

Média Geométrica

$$L(\tilde{y} | \theta) = \sqrt[n]{\prod_{t=1}^n L(\tilde{y}_t | \theta)} \quad (2.18)$$

Pode ser tão restritivo quanto como a Multiplicação, mas em compensação produz uma superfície de resposta menos abrupta ao ponderar os valores multiplicados. Esta qualidade torna também este agregador menos sensível a grandes quantidades de valores comparados.

2.2.2.2 Funcionamento do *framework*

Depois de definida a função de verossimilhança, $L(\tilde{y} | \theta)$, o algoritmo REJ, definido na seção 2.2, pode ser aplicado de maneira direta para gerar observações da distribuição à posteriori $f(\theta | \tilde{y})$. Entretanto, de maneira extremamente ineficiente.

A amostragem por hipercubo latino pode ser utilizado para reduzir a quantidade de simulações necessárias quando a distribuição à priori dos parâmetros, $\pi(\theta)$, é fortemente informativa (BEVEN; BINLEY, 2014). Entretanto, para cenários mais complexos, a estratégia de amostragem é semelhante a descrita para o *framework* ABC na seção 2.2.1: utilizar métodos como o MCMC. Mais especificamente no âmbito do GLUE, há abordagens como amostragem por MCMC adaptativo em BLASONE et al. (2008) e em VRUGT; BRAAK; GUPTA et al. (2009).

Em linhas gerais, a Figura 2 apresenta a metodologia de um algoritmo GLUE que calibra um modelo \mathcal{M} parametrizado por θ . Nela, os blocos cinzas representam a atuação da função de verossimilhança no *framework*.

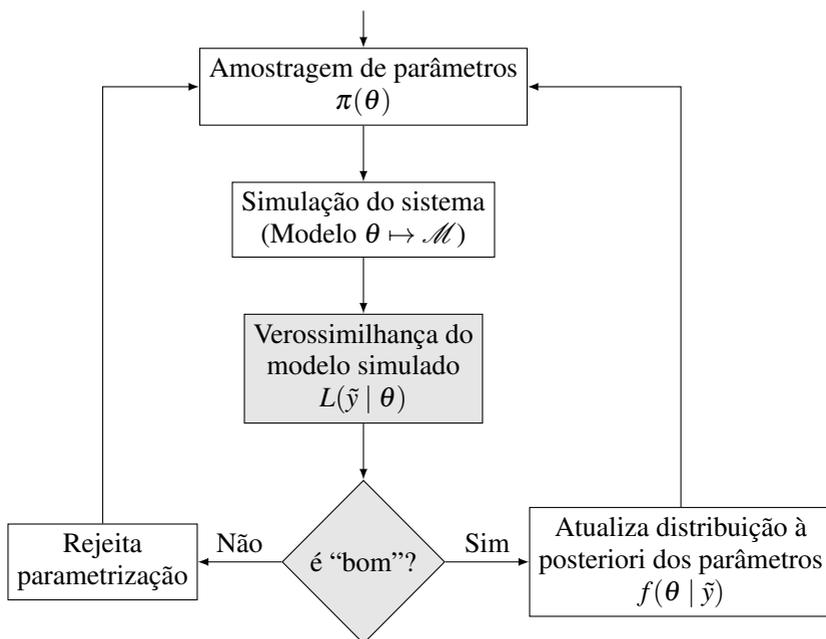


Figura 2 – Passo-a-passo do GLUE

3 DESENVOLVIMENTO DO MÉTODO

Este capítulo trata da descrição do método proposto como objetivo deste trabalho. Apesar de seu escopo claro, o propósito aqui é descrever de maneira vasta as formalizações necessárias e os detalhes que construíram os passos deste método.

Será exposto inicialmente maiores detalhes sobre a parametrização inicial do BYM: como ela é realizada a partir dos dados. Na sequência, é definido o modelo probabilístico e a maneira como é feita a inferência sobre os parâmetros restantes.

3.1 A SOLUÇÃO DO BYM

Com o intuito de formalizar descrições matemáticas, o BYM é um sistema de duas ODEs que, dados os valores iniciais $y_0 = [D_0, h_0]$ e os parâmetros $\theta = [a_1, \dots, a_8]$, pode ser resolvido numericamente (simulado) com solução $y_t = \mathcal{S}(y_0, \theta, t)$ em um vetor discreto de tempo $t = [t_1, \dots, t_n]$. Aqui, o valor inicial de h será sempre zero, $h_0 = 0$, pois são utilizadas somente brocas novas, o que reduz o espaço de estados observados para $y_t = D_t$.

A Figura 3 mostra o exemplo do resultado de uma simulação com $n = 72000$, produzindo o aspecto de uma curva contínua.

O termo f_1 em (2.1) é um termo livre, fixo, que inicialmente deve

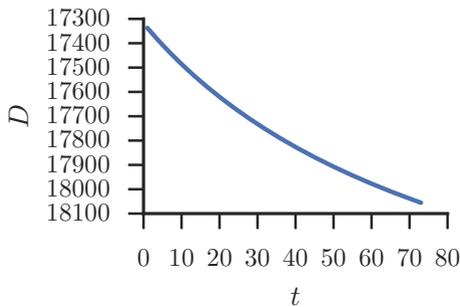


Figura 3 – Exemplo de resultado de uma simulação. A variável t está em horas e D em pés.

ser calculado isolando-o na equação,

$$f_1 = \frac{dD}{dt} / f_2 \times f_3 \times f_4 \times f_5 \times f_6 \times f_7 \times f_8,$$

a partir do conhecimento de poços de perfuração similar. Quando não há valores de coeficientes pré-determinados de dados de perfuração anteriores, utilizam-se os indicados na Tabela 1. Para os parâmetros operacionais restantes, detalhados em (2.2), este cálculo é realizado com os respectivos valores médios observados.

Tabela 1 – Valor padrão dos coeficientes, utilizados na ausência de cálculos anteriores.

Coef.	Valor Padrão
a_2	2.45161×10^{-5}
a_3	9.67742×10^{-6}
a_4	2.31472×10^{-5}
a_5	0.503226
a_6	0.988155
a_7	1.49828
a_8	0.6

Ainda na definição de parametrização fixa, os valores de H_1 , H_2 e $(W/d_b)_m$ – constantes relativas à broca – são escolhidos conforme a tabela fornecida em BOURGOYNE; CHENEVERT et al. (1986, p. 218). A constante restante, τ_H , é obtida na equação (2.4) onde o desgaste final observado na broca, h_f , e o tempo necessário para tal, t_b , são colhidos nos dados de perfuração.

As demais variáveis, $\{N, W, F_j, g_p, \rho_c\}$, correspondem à variáveis operacionais ou de ambiente também coletadas e realizadas durante uma perfuração. Dependendo de como são coletadas e apresentadas, essas variáveis podem conter simplificações ou até mesmo apresentar valores espúrios (*outliers*). Normalmente são adotados valores médios para tais variáveis principalmente quando existe pouca variabilidade nos seus valores. Em um exemplo apresentado na Figura 4, a escolha

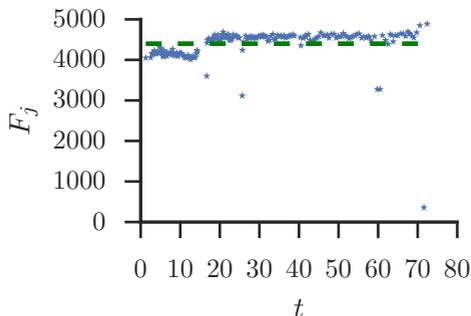


Figura 4 – Exemplo de um parâmetro operacional observado com pouca variabilidade, e seu valor médio no intervalo (linha tracejada).

do valor médio para tal apresenta-se como uma opção sensata além de desconsiderar alguns valores espúrios.

Entretanto, quando há certa variabilidade, esta deve ser sem dúvida considerada. Como a solução do BYM se trata de uma integração numérica, não é possível inserir variabilidade em termos que supostamente são constantes em uma ODE sem impactar fortemente na eficiência da solução – o processo de integração leva em conta o cálculo do derivativo dos estados das equações em cada passo de tempo e a alteração relativamente livre de um parâmetro gera singularidades na função sendo integrada.

A maneira encontrada aqui para compreender os diferentes valores de uma variável, v , na solução do BYM é definir um polinômio $P_v^*(t) = (c_n, \dots, c_1, c_0)$ de grau n e coeficientes c_i , que descreva a sua tendência. Esta ideia, apresentada graficamente na Figura 5, pode também lidar, de certa forma, com valores espúrios.

O método utilizado neste trabalho para definir esses polinômios é o de mínimos quadrados. Apesar de existirem diversas formas mais sofisticadas de ajuste e escolha de polinômios (GIROLAMI, 2008), este não é o foco aqui. Tendo esses polinômios em mãos, é possível adicioná-los ao espaço de estados do BYM – inicialmente constituído por D e h – agregando-os, assim, ao processo de integração numérica. Para isso é obtida a primeira derivada do polinômio P_v^* , por simplicidade

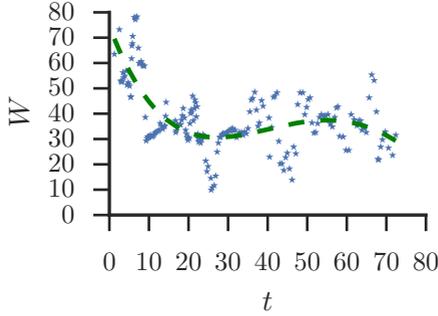


Figura 5 – Exemplo de um parâmetro operacional observado com variabilidade substantiva, e o polinômio de quarto grau (linha tracejada) ajustado aos dados.

denotada $dP_v^*/dt \equiv P_v$, com grau $n - 1$, que é adicionada ao BYM com o respectivo valor inicial c_0 .

Assim o sistema de ODEs do BYM resultante tem a seguinte estrutura:

$$\begin{aligned} \frac{dD}{dt} &= f_1 \times f_2 \times f_3 \times f_4 \times f_5 \times f_6 \times f_7 \times f_8 \\ \frac{dh}{dt} &= \frac{1}{\tau_H} \left(\frac{N}{60} \right)^{H_1} \left[\frac{\left(\frac{W}{d_b} \right)_m - 4}{\left(\frac{W}{d_b} \right)_m - \left(\frac{W}{d_b} \right)} \right] \left(\frac{1 + H_2/2}{1 + H_2h} \right) \end{aligned} \quad (3.1)$$

$$P_v(t) = \dots$$

⋮

onde cada variável definida por um polinômio tem o seu valor atualizado nos passos de integração numérica – processo inverso da obtenção da derivada – resultando nos valores definidos pelo polinômio ajustado por mínimos quadrados.

3.2 O MODELO PROBABILÍSTICO

Quando fornecidas as observações de profundidade, $\tilde{y} = \{\tilde{D}_t \mid t = 1, \dots, n\}$, o interesse passa a ser a inferência da distribuição de probabilidade à posteriori, $f(\theta \mid \tilde{y})$. Para inferi-la através do teorema de Bayes,

definido na equação (2.5), pode-se utilizar tanto o ABC (seção 2.2.1) quanto o GLUE (seção 2.2.2) dependendo da amostragem utilizada e da maneira de comparar os valores simulados e observados.

Neste trabalho, optou-se por utilizar uma função de verossimilhança Gaussiana, nos moldes do *framework* GLUE e apresentada na equação (2.8), como medida do erro (diferença entre os valores simulados e observados). Devido a essa escolha se supõe, portanto, que os erros são não correlacionados, com variância constante (homoscedástico), e também, distribuídos normalmente.

É verdade que para tal sistema físico esta suposição acaba não sendo realística. Entretanto, aprofundar na modelagem do erro – pelo menos em termos estatísticos – acabou não sendo tema deste trabalho. Pensar estruturas de erros correlacionados e heteroscedásticos é o foco de trabalhos como o de SCHOUPS; VRUGT (2010) na área de funções de verossimilhança formais, e BLUM; FRANÇOIS (2010) nos métodos de inferência livre de verossimilhança. Promover estas adequações aqui é tema de trabalhos futuros.

Voltando na definição utilizada neste trabalho, a Figura 6 mostra uma representação de como é realizada essa medida para o erro. De fato essa medição é realizada para cada ponto do tempo onde há valor de profundidade observado.

Com o intuito de comparação, todas essas medidas são agregadas utilizando-se o operador multiplicativo, definido na equação (2.14), fazendo com que a função de verossimilhança resultante se torne

$$L(\tilde{y} | \theta) = \prod_{t=1}^n \frac{1}{\sqrt{2\pi\hat{\sigma}_p^2}} \exp \left[-\frac{1}{2\hat{\sigma}_p^2} (\tilde{y}_t - y_t)^2 \right]. \quad (3.2)$$

Como a função Gaussiana está definida em $[-\infty, +\infty]$, este operador acaba não tornando a multiplicação uma agregação restritiva. Entretanto é possível que resulte valores muito pequenos, dependendo do caso, como será discutido posteriormente na apresentação dos experimentos.

Aqui, $\hat{\sigma}_p^2$ é uma estimativa do desvio padrão do erro que pode ser pré-determinada ou inferida em conjunto com os valores de θ (VRUGT;

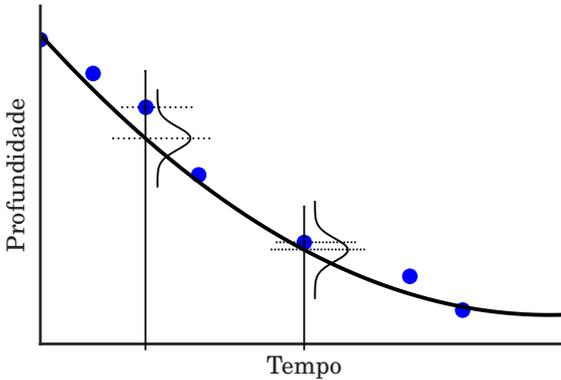


Figura 6 – Aplicação da medida de verossimilhança dos dados observados representados pelos círculos.

BRAAK; GUPTA et al., 2009; SADEGH; VRUGT, 2013). Esta medida de verossimilhança no framework GLUE equivale à utilização dos erros ao quadrado em (2.6), $\rho(\tilde{y}, y) = \sum_{t=1}^n (\tilde{y}_t - y_t)^2$, no ABC (TONI et al., 2009). E, adicionar $\hat{\sigma}_p$ aos valores que devem ser inferidos é semelhante ao que BORTOT; COLES; SISSON (2007) propuseram ao adicionar ε ao estado de espaços da amostragem.

Escolher um valor fixo para $\hat{\sigma}_p$ pode não ser uma tarefa trivial. Um alto valor permite que configurações de parametrização não-ótimas possam ser consideradas caso o BYM não consiga simular perfeitamente os dados reais; um valor pequeno de desvio padrão pode não permitir a exploração de possíveis boas configurações do espaço paramétrico ao criar entraves de mínimos locais. Portanto, escolheu-se aqui adicioná-lo aos valores a serem inferidos, o que permite variação do seu valor durante a amostragem. A tendência natural é de seu valor reduzir pois, como mostrado na Figura 7, para um mesmo conjunto de valores $\{\tilde{y}, y\}$, com y sendo o resultado de uma boa simulação, o valor da verossimilhança será maior quando o desvio padrão $\hat{\sigma}_p$ for menor.

Outra variável importante na solução do BYM é o valor inicial de profundidade, D_0 , pois determina o ponto inicial do processo de integração numérica. É possível que não exista nos dados de perfuração

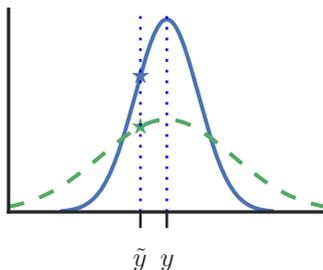


Figura 7 – Efeito do desvio padrão na medida de verossimilhança: a função Gaussiana em tracejado tem maior desvio padrão.

registro de operação exatamente na profundidade inicial. Neste caso, o valor de D_0 deve ser inferido juntamente com as outras variáveis, o que permite aliás um maior grau de flexibilidade na calibração do modelo.

Assim, o espaço paramétrico final, $\theta = (D_0, a_2, a_3, \dots, a_8, \hat{\sigma}_p)$, é consolidado através das variáveis de desvio padrão do erro, de valor inicial da perfuração e dos próprios coeficientes do BYM.

Para tais parâmetros é necessário definir a distribuição de probabilidade à priori, $\pi(\theta)$, abreviadamente também referida como *prior*. Devido ao fato de não existir conhecimento inicial sobre as possibilidades dos diferentes valores de cada parâmetro, os *priors* aqui são definidos como não-informativos. Ou seja, para cada parâmetro v que compõe θ é definido um intervalo $[a_v, b_v]$, disposto na Tabela 2, com distribuição de probabilidade uniforme: $\pi(\cdot) \sim U(a, b)$.

O intervalo para os coeficientes (a_2, \dots, a_8) segue o sugerido por BOURGOYNE; CHENEVERT et al. (1986) com um pequeno relaxamento nos limites superiores. Para a profundidade inicial, D_0 , é definido um intervalo relativo à primeira observação de profundidade, \tilde{D}_1 , com 0.7% do seu valor para menos e para mais. Por fim, o desvio padrão da função de verossimilhança, σ_p , vai de um valor relativamente pequeno, 10^{-4} , até o valor de um 1% de \tilde{D}_1 – relação de grandeza necessária uma vez que a função de verossimilhança estará centrada (terá seu valor de média, μ) nos valores de profundidade.

A Figura 8 apresenta graficamente o modelo probabilístico defi-

Tabela 2 – Intervalo dos parâmetros na distribuição à priori.

Parâmetro	Intervalo dos <i>priors</i> ($[a, b]$)
D_0	$[\tilde{D}_1 \pm 0.007\tilde{D}_1]$
a_2	$[2.303 \times 10^{-6}, 0.012]$
a_3	$[2.303 \times 10^{-8}, 0.021]$
a_4	$[2.303 \times 10^{-6}, 2.303 \times 10^{-3}]$
a_5	$[0.3, 2.5]$
a_6	$[0.2, 1.5]$
a_7	$[0.1, 2.5]$
a_8	$[0.1, 0.9]$
σ_p	$[10^{-4}, 0.01\tilde{D}_1]$

nido até aqui. A partir de então, resta definir uma maneira de realizar a inferência sobre a distribuição $f(\boldsymbol{\theta} | \tilde{y})$.

3.2.1 Monte Carlo em Cadeias de Markov (MCMC)

O fato de escolher uma função de verossimilhança como nos preceitos do GLUE permite a utilização direta de um método baseado em MCMC para realizar a inferência da distribuição à posteriori, $f(\boldsymbol{\theta} | \tilde{y})$.

Um método bastante comum, e escolhido aqui, é o *Random Walk Metropolis* (RWM) (ANDRIEU; FREITAS et al., 2003; TIERNEY, 1994), ou somente Metropolis, que consiste em um caso específico do algoritmo Metropolis-Hastings (MH) definido na seção 2.2.1.

A diferença significativa do RWM reside em uma simplificação da razão de aceitação, α , definida na equação (2.7). Nela, a função de transição é simétrica, portanto $q(\boldsymbol{\theta}^* | \boldsymbol{\theta})/q(\boldsymbol{\theta} | \boldsymbol{\theta}^*) = 1$. Além disso, como neste caso os *priors*, $\boldsymbol{\pi}(\boldsymbol{\theta})$, têm valores constantes de probabilidade dentro dos limites intervalares, a razão $\boldsymbol{\pi}(\boldsymbol{\theta}^*)/\boldsymbol{\pi}(\boldsymbol{\theta})$ passa a agir como uma função indicadora: quando q propõe um valor dentro dos limites, os *priors* serão constantes, caso contrário a amostra é rejeitada.

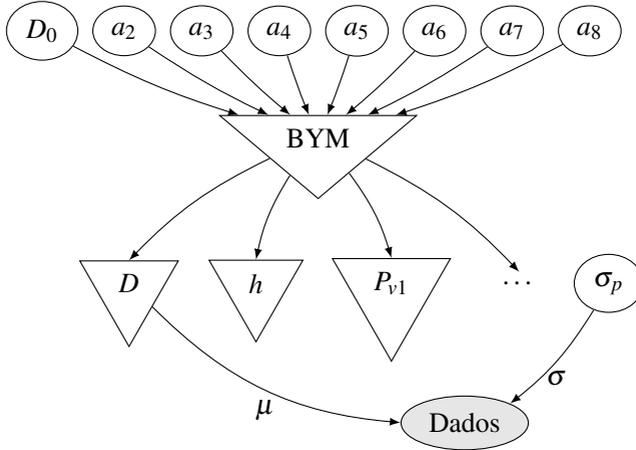


Figura 8 – Grafo do modelo probabilístico: os parâmetros a serem inferidos, θ , são representados por elipses; a verossimilhança é a elipse sombreada, e os triângulos invertidos são componentes do BYM.

Assim, a razão de aceitação simplificada resulta em

$$\alpha = \min \left(1, \frac{L(\tilde{y} | \theta^*)}{L(\tilde{y} | \theta)} \mathbb{1}_{\pi(\theta^*) > 0} \right). \quad (3.3)$$

A função de transição simétrica é definida como

$$q(\cdot | \theta) \sim \text{Normal}(\theta, \sigma_v),$$

onde o próximo valor, θ^* , parte da amostragem de uma distribuição normal centrada no atual valor, θ . Cada parâmetro do conjunto de *priors*, θ_v , é amostrado de maneira independente. O valor inicial respectivo, que parametriza a média da função de transição, é o centro do intervalo, $(b_v - a_v)/2$, e o desvio padrão para cada um é $\sigma_v = 1/6(b_v - a_v)$. Assim, já na primeira proposição $q(\theta^* | \theta)$, 99.73% dos valores possíveis cobrem o intervalo conforme ilustra a Figura 9. Esta proporção do passo de transição mostrou-se grande o suficiente para garantir uma boa mixagem – ou seja, toda a gama de valores de parâmetros pode ser sugerida em poucos passos – e pequena o suficiente para não permitir

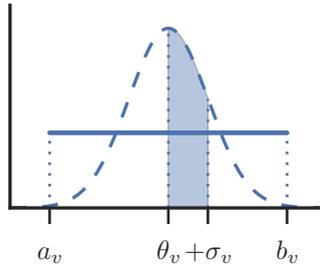


Figura 9 – Configuração inicial da função de transição, em tracejado, no intervalo do *prior* não-informativo.

demasiadas propostas rejeitadas por se encontrarem além do intervalo definido, onde $\pi(\theta^*) = 0$.

Esta maneira de amostragem – sugerindo uma transição partindo do valor anterior – gera valores não-independentes estatisticamente, segundo comentado na seção 2.2.1, e é uma característica dos métodos MCMC. Observações da distribuição à posteriori, f , só devem ser consideradas a partir de um número incerto de iterações, quando a Cadeia de Markov gerada pelo algoritmo torna-se ergódica, ou em termos mais práticos, quando se identifica o estabelecimento de uma distribuição estacionária. Para se verificar esta condição, deve-se conferir o traço dos diferentes valores aceitos de parametrização e desconsiderá-los em uma quantidade inicial.

Especialmente neste trabalho foi verificada uma condição especial de mixagem da cadeia por permitir que o elemento de desvio padrão da função de verossimilhança, σ_p , participasse do espaço de parâmetros inferidos. Permitir variação no seu valor é permitir variação na própria definição do erro. Na prática, isso induz um viés (*bias*) na calibração dos parâmetros uma vez que os modelos são comparados com diferentes níveis de variância (FEARNHEAD; PRANGLE, 2012). Como será mostrado e discutido nos capítulos subsequentes, para reduzir este viés é necessário verificar os valores de σ_p amostrados e possivelmente considerar somente as amostras $\{\theta_i \mid \sigma_p < \sigma_T\}$ para algum valor de limiar σ_T . Trata-se de uma tipificação do sugerido por BORTOT; CO-

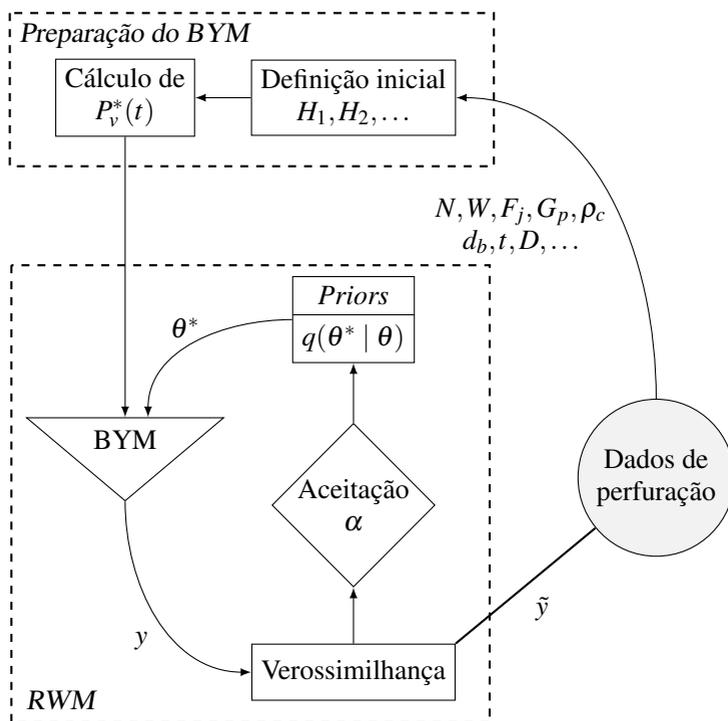


Figura 10 – Resumo do método. Depois de definido os parâmetros iniciais necessários para o BYM, o RWM começa sua execução pela função de transição, q .

LES; SISSON (2007), comentado na seção 2.2.1.

Um resumo visual do método definido neste capítulo é apresentado na Figura 10. Perceba-se que a região do algoritmo RWM é uma especificação da metodologia GLUE ilustrada no fluxograma da Figura 2 e cuja relação entre as variáveis já foi exposta no grafo da Figura 8.

4 EXPERIMENTOS E RESULTADOS

Este capítulo apresenta, de maneira extensiva, os dados utilizados e o resultado de cada passo do processo de calibração do BYM. São feitas também pequenas pontuações técnicas sobre detalhes do método. Já uma análise crítica é apresentada somente no capítulo 5.

Como experimentação, quatro modelos foram calibrados baseados em dados de diferentes perfurações. Cada perfuração, A, B, C e D, consiste nos dados de execução de uma broca, respectivamente, em quatro diferentes poços do pré-sal brasileiro. Os resultados obtidos são confrontados com os próprios dados reais de perfuração no intuito de averiguar a calibração dos modelos.

Os valores das variáveis definidas e calculadas para serem utilizadas no BYM, em cada uma das perfurações, pode ser visto na Tabela 3.

Coincidentemente os quatro modelos tiveram a variável W modelada por um polinômio. Dos dados das Perfurações A,B e D foi modelado um polinômio de quarto grau para W (peso sobre a broca); da Perfuração C foi modelado um polinômio de sétimo grau. Os coeficientes que definem estes polinômios respectivamente estão descritos na equação (4.1):

$$\begin{aligned}
 P_{W(A)}^4 &= [1.48 \times 10^{-7}, 1.1 \times 10^{-4}, -0.022, 1.239, 21.53] \\
 P_{W(B)}^4 &= [1.68 \times 10^{-6}, 2.5 \times 10^{-4}, -0.045, 1.839, 11.4] \\
 P_{W(C)}^7 &= [2.4 \times 10^{-9}, -6.4 \times 10^{-7}, 6.9 \times 10^{-5}, -0.0038, \\
 &\quad 0.1182, -1.891, 13.98, 4.973] \\
 P_{W(D)}^4 &= [5.19 \times 10^{-6}, -1.49 \times 10^{-3}, 0.129, -4.12, 7.45].
 \end{aligned} \tag{4.1}$$

Para fins de completude e análise posterior, as Figuras 12 e 13 justapõem esses valores utilizados na preparação dos modelos com os dados reais de cada variável.

Para cada modelo de perfuração foram executadas 80000 iterações do algoritmo RWM com a finalidade da inferência dos parâmetros, θ . Em todos os casos foi percebido um comportamento inicial

Tabela 3 – Parametrização das variáveis operacionais e observadas nas perfurações.

Parâm.	Perfuração			
	A	B	C	D
H_1		1.5		
H_2		2.0		
$(W/d_b)_m$		10.0		
a_1		2.4		
\tilde{D}_1	16510	16832	16274	17356
τ_H	581.87	256.80	129.57	379.37
d_b	12.25	8.50	9.00	12.25
g_p	8.31	9.81	10.22	9.88
ρ_c	10.01	10.90	10.01	10.01
W	$P_{W(A)}^4$	$P_{W(B)}^4$	$P_{W(C)}^7$	$P_{W(D)}^4$
N	144.99	191.91	81.69	65.05
F_j	3431.82	3491.19	4161.76	4400.37

exemplificado na Figura 11: o valor inicial de σ_p é relativamente alto, permitindo que diversas parametrizações sejam aceitas, até o momento em que sua queda (para ≈ 20 neste caso) condensa a abrangência da função de verossimilhança, o que penaliza mais severamente as parametrizações anteriormente aceitas. Então os valores de θ caminham para regiões de maior probabilidade.

Com efeito, essas simulações iniciais são prontamente descartadas. Para a construção das distribuições de probabilidade à posteriori dos parâmetros (também referenciada aqui como *posteriors*) é descartado de fato todos os 10,000 primeiros valores sugeridos pelo RWM. Esta quantidade inicial é uma larga garantia que representa o momento de “aquecimento” (*burn-in*) do algoritmo – enquanto a Cadeia de Markov se torna ergódica. Além disso, devido ao autocorrelacionamento intrínseco dos valores gerados, é realizado um desbaste (*thinning*) de ordem dez. Ou seja, são considerados somente valores intervalados por

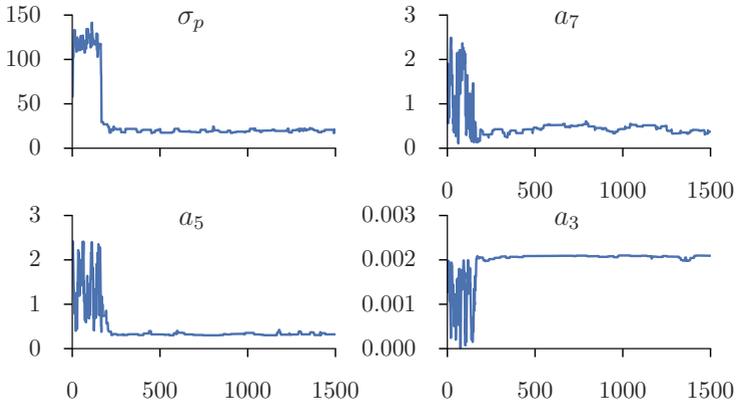


Figura 11 – Primeiras 1500 iterações nos dados da Perfuração B: exemplo da influência dos valores de σ_p no procedimento de calibração.

outros dez valores. As Figuras 14, 15, 16 e 17 apresentam os traços dos valores para todas as cadeias. Perceba-se que de todas as iterações são consideradas ao final somente 7000 valores.

Ao agrupá-los, obtêm-se os gráficos de distribuição à posteriori dos parâmetros, $f(\theta | \tilde{y})$, que podem ser vistos nas Figuras 18, 19, 20 e 21. Nelas, a linha tracejada corresponde à mediana e as linhas pontilhadas aos 25^o e 75^o percentis, o que permite se recriar visualmente um *box-plot*, caso conveniente.

É possível perceber que para alguns parâmetros foi inferida uma distribuição multimodal. O caso mais extremo é entre a_3 e a_6 na Perfuração C. Isso se deve à possibilidade que estes têm de se correlacionarem na equação do BYM. Devido à natureza exponencial do sistema de equações (2.1), um termo pode acabar compensando o outro no resultado final caso suas variáveis integrantes sejam constantes. Como nestes exemplos somente W foi incluída como termo variável, há uma maior chance de outros coeficientes com bases constantes se correlacionarem.

Em compensação, como visto anteriormente, os valores considerados após o *burn-in* estão localizados em uma faixa relativamente

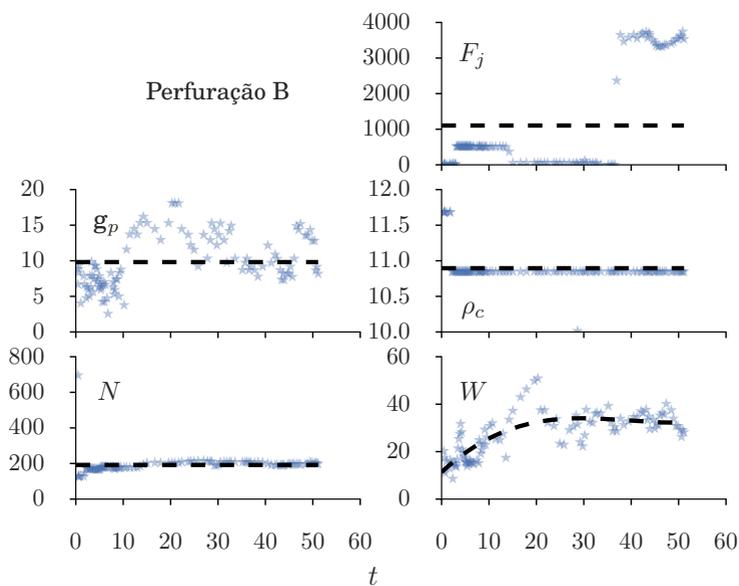
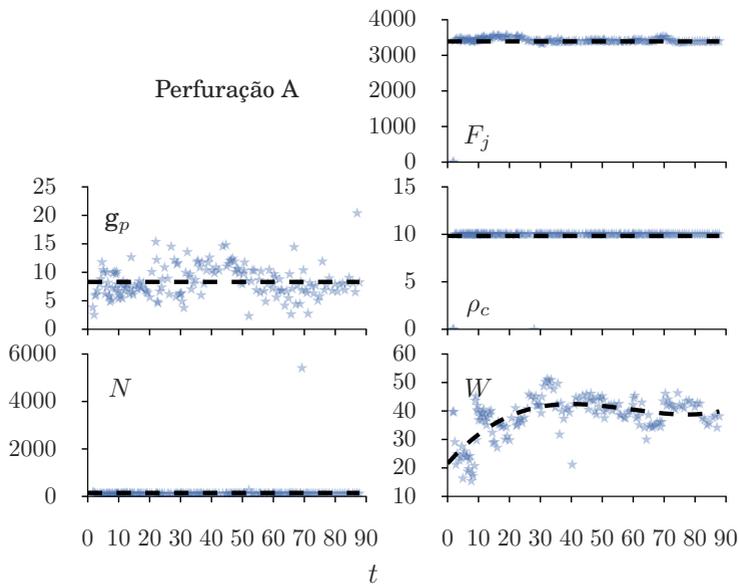


Figura 12 – Perfurações A e B: valores das variáveis observadas nos dados, em marcas, e os utilizados no modelo simulado, apresentado nas linhas pontilhadas.

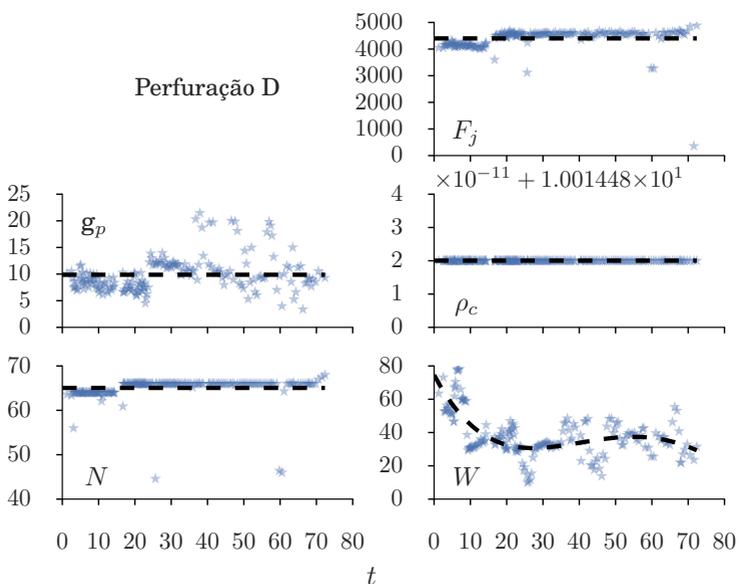
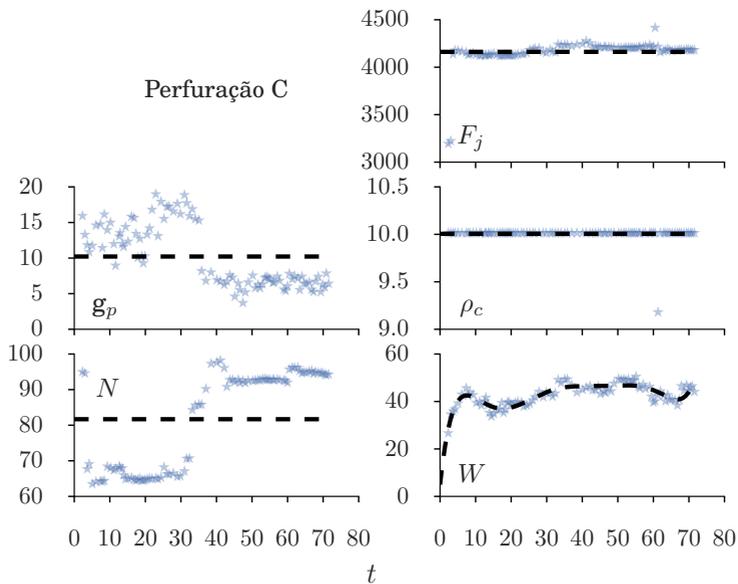


Figura 13 – Perfurações C e D: valores das variáveis observadas nos dados, em marcas, e os utilizados no modelo simulado, apresentado nas linhas pontilhadas.

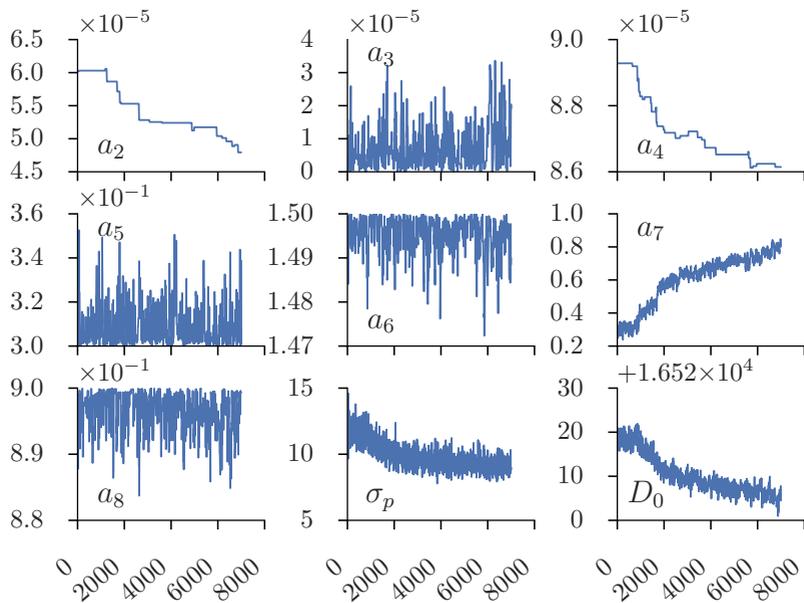


Figura 14 – Perforação A: traços das cadeias.

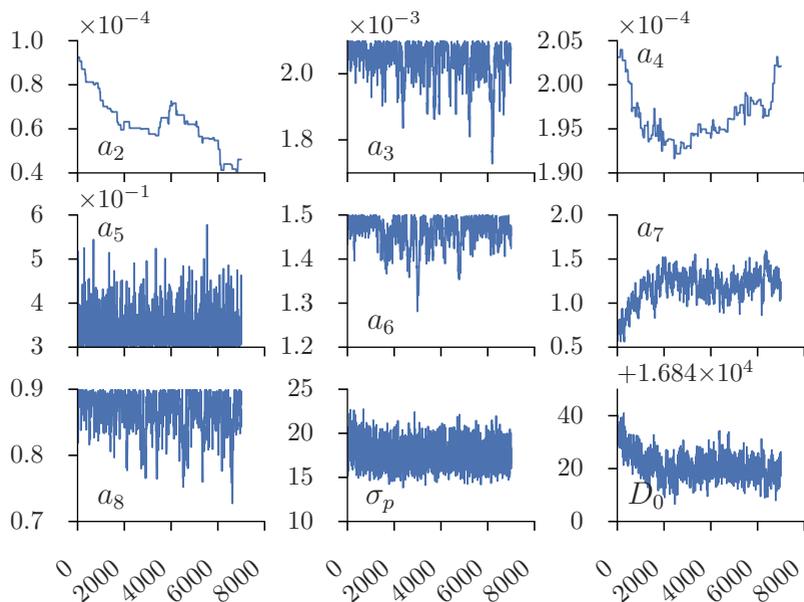


Figura 15 – Perforação B: traços das cadeias.

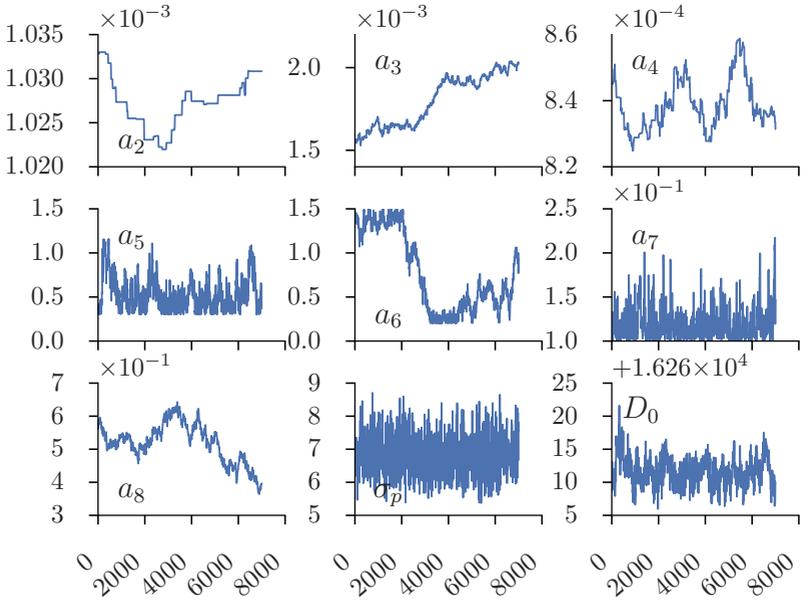


Figura 16 – Perfuração C: traços das cadeias.

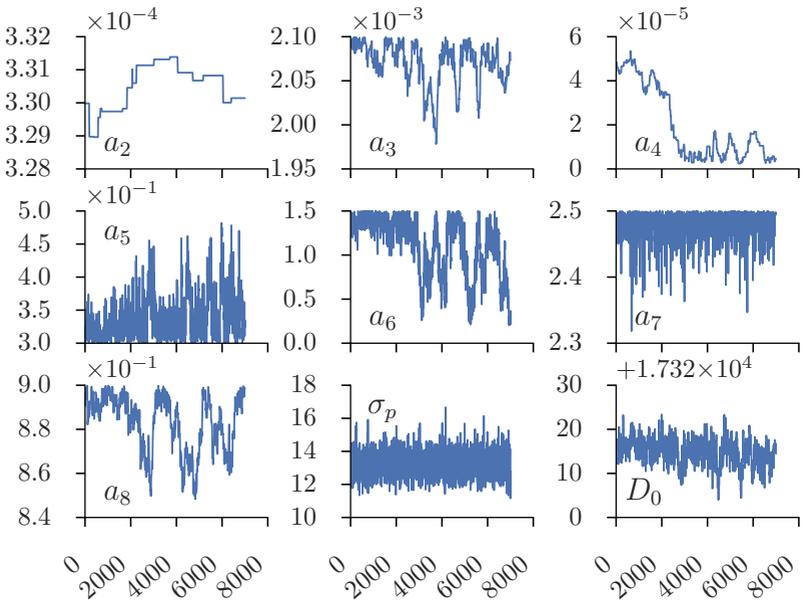


Figura 17 – Perfuração D: traços das cadeias.

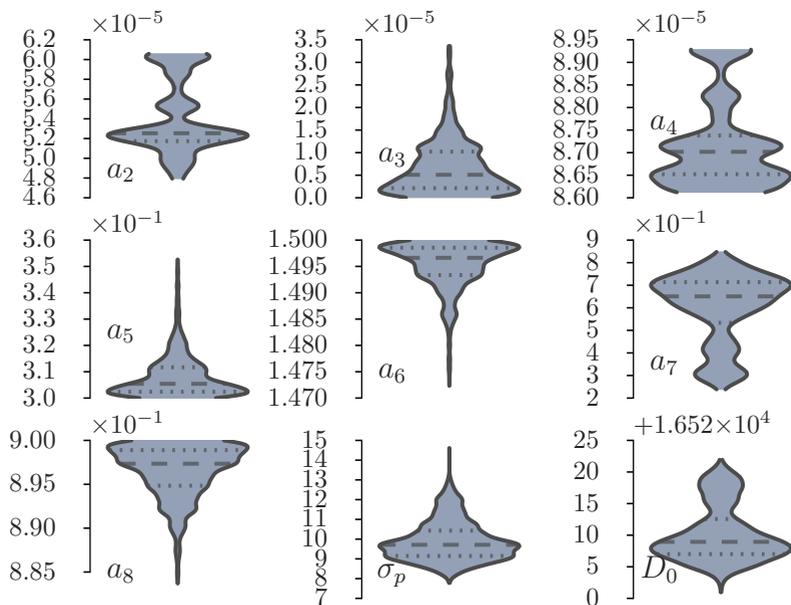


Figura 18 – Perfuração A: distribuições à posteriori dos parâmetros.

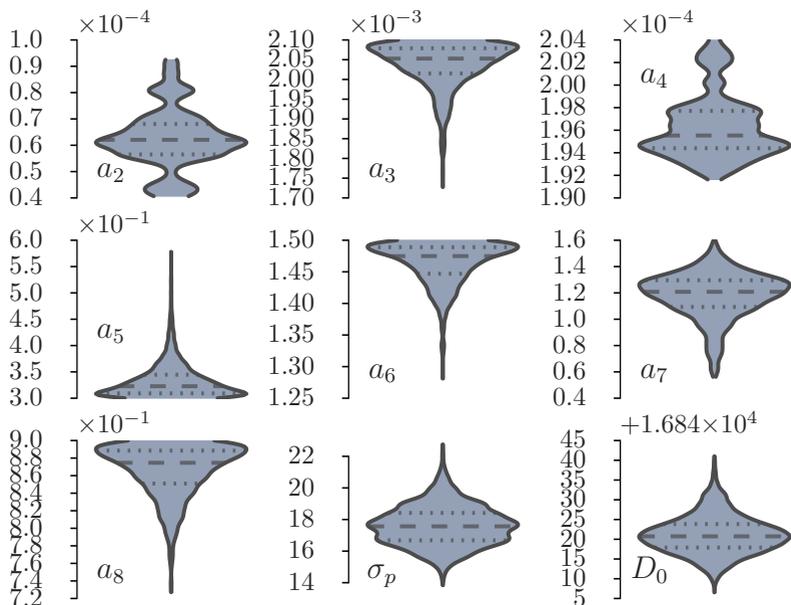


Figura 19 – Perfuração B: distribuições à posteriori dos parâmetros.

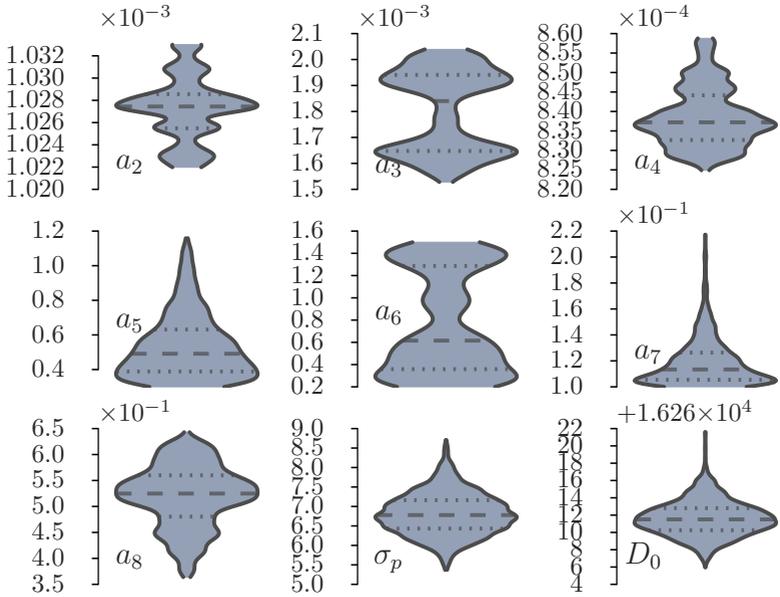


Figura 20 – Perfuração C: distribuições à posteriori dos parâmetros.

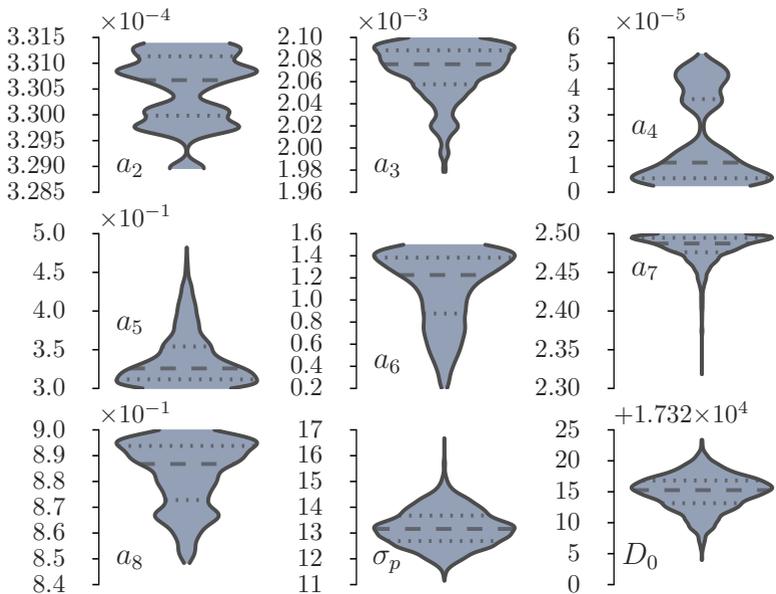


Figura 21 – Perfuração D: distribuições à posteriori dos parâmetros.

restrita de valores em comparação ao intervalo inicialmente definido aos *priors* (Tabela 2). Isso permite inferir a mediana das distribuições multimodais dos parâmetros e obter pequena variação nas simulações do BYM.

Realmente, como mostra a Figura 22, os modelos parametrizados com os valores medianos dos *posteriors* simulam valores muito próximos aos dados reais. O Erro Quadrático Médio (RMSE, *Root Mean Squared Error*), definido como

$$RMSE = \sqrt{\frac{\sum_{t=1}^n (\tilde{y} - y)^2}{n}}, \quad (4.2)$$

e o Erro Médio Percentual Absoluto (MAPE, *Mean Absolute Percentage Error*),

$$MAPE = \frac{1}{n} \sum_{t=1}^n \left| \frac{\tilde{y} - y}{\tilde{y}} \right|, \quad (4.3)$$

dessas simulações estão expostos na Tabela 4. Ainda Figura 22, as áreas preenchidas em cinza, correspondem ao intervalo de 95% de Máxima Densidade à Posteriori (HPD, *Highest Posterior Density*). Este intervalo, portanto, abrange o resultado de todas as simulações realizadas pelos modelos parametrizados por todos os valores possíveis de parâmetros entre o percentis 2.5º e 97.5º das distribuições à posteriori. Perceba-se que, em todas as Perfurações, os dados reais estão contidos neste intervalo de simulações possíveis.

Tabela 4 – Medidas de erro entre os valores observados e simulados pelo modelo parametrizado com a mediana das distribuições à posteriori.

Perf.	RMSE	MAPE
A	9.61761	0.04497%
B	17.17884	0.08835%
C	6.63449	0.03249%
D	13.04604	0.06202%

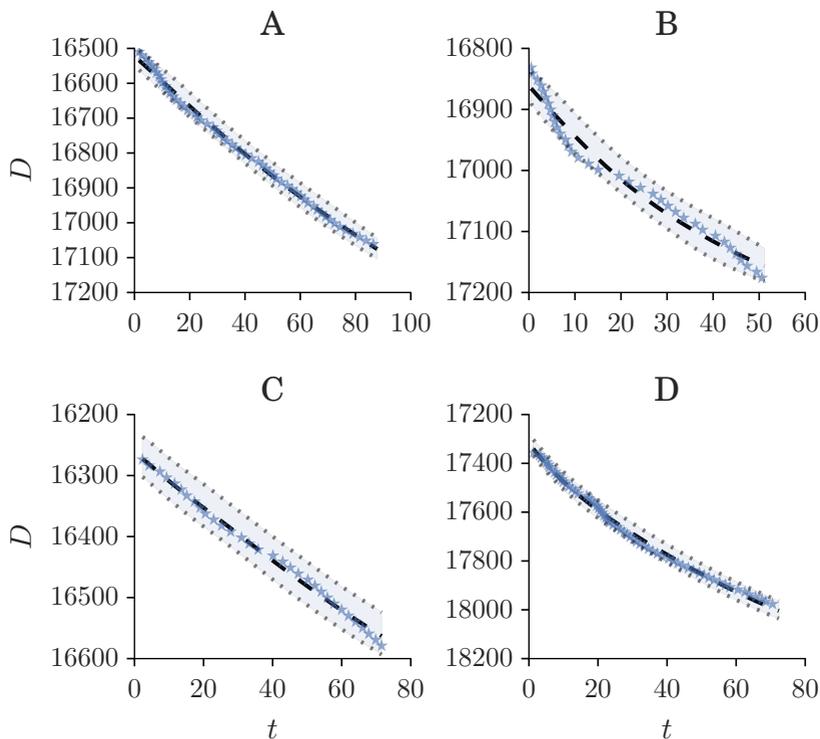


Figura 22 – Resultado das simulações do BYM. Em estrela estão os dados reais da perfuração. No tracejado encontra-se o resultado da simulação com a mediana dos *posteriors*; e os pontilhados restantes delimitam a área de 95% HPD.

Por fim, a Tabela 5 apresenta os valores que definem as distribuições à posteriori dos parâmetros ilustrados anteriormente. Os valores informados de percentis correspondem aos tracejados e pontilhados das Figuras 18, 19, 20 e 21.

Tabela 5 – Intervalos de credibilidade inferidos para os coeficientes do BYM, para o valor de profundidade inicial, D_0 , e para σ_p .

Parâm.	Percentil	Perfuração			
		A	B	C	D
a_2	25°	$5.17e-05$	$5.76e-05$	0.00103	0.00033
	50°	$5.28e-05$	$6.32e-05$	0.00103	0.00033
	75°	$6.02e-05$	$7.29e-05$	0.00103	0.000331
a_3	25°	$2.11e-06$	0.00202	0.00162	0.00206
	50°	$5.09e-06$	0.00206	0.00172	0.00208
	75°	$1.02e-05$	0.00208	0.00193	0.00209
a_4	25°	$8.65e-05$	0.000195	0.000834	$5.87e-06$
	50°	$8.71e-05$	0.000196	0.000838	$1.4e-05$
	75°	$8.83e-05$	0.000199	0.000845	$4.38e-05$
a_5	25°	0.30	0.31	0.40	0.31
	50°	0.31	0.32	0.51	0.32
	75°	0.31	0.34	0.65	0.35
a_6	25°	1.49	1.45	0.40	0.93
	50°	1.50	1.48	0.73	1.27
	75°	1.50	1.49	1.36	1.41
a_7	25°	0.39	1.01	0.11	2.47
	50°	0.63	1.18	0.11	2.49
	75°	0.71	1.28	0.13	2.49
a_8	25°	0.89	0.85	0.49	0.87
	50°	0.90	0.88	0.53	0.89
	75°	0.90	0.89	0.56	0.90
σ_p	25°	9.24	16.82	6.44	12.71
	50°	9.82	17.74	6.78	13.20
	75°	11.03	18.77	7.18	13.73
D_0	25°	16527.4	16858.4	16270.4	17333.6
	50°	16529.5	16861.5	16271.7	17335.6
	75°	16536.8	16866.1	16273.0	17337.3

5 CONCLUSÃO

Detalhes específicos da formulação do método desprenderam-se dos dados utilizados e da natureza particular do sistema de equações do BYM. Aumentar o espaço de estados com polinômios que representam as variáveis operacionais ou de ambiente – apresentado na seção 3.1 – foi a solução desenvolvida para lidar com o primeiro objetivo específico deste trabalho.

Dentre os demais objetivos específicos, definidos na seção 1.3, a utilização da Gaussiana como função de verossimilhança e, em especial, a possibilidade de variar sua abrangência sob certas condições, permitiu uma boa medição da adequação entre os dados simulados e os reais.

Já, como contribuição ainda mais específica, a parametrização do algoritmo RWM na seção 3.2.1 mostrou-se satisfatória na exploração do espaço amostral paramétrico do BYM.

Além disso, a decisão de tomar dados de perfuração de brocas completas permitiu que a equação que modela o desgaste da broca, (2.3), fosse considerada no processo de calibração do modelo – uma diferenciação deste trabalho e também uma das motivações para sua realização.

Um apanhado de comentários críticos e detalhes observados no decorrer de cada passo do método são apresentados à seguir como uma análise dos pontos fortes ou fracos e das respectivas melhorias possíveis do método desenvolvido neste trabalho.

5.1 COMENTÁRIOS E TRABALHOS FUTUROS

Inicialmente, durante o tratamento das variáveis $\{N, W, F_j, g_p, \rho_c\}$, ao se determinar o polinômio ou o valor médio para definição de cada uma – Figuras 12 e 13 – é possível que uma quantidade notável de informação esteja sendo desprezada. Como observado anteriormente, é sensato desconsiderar valores espúrios existentes nos dados à exemplo do ocorrido nas variáveis N e ρ_c , da Perfuração B, ou F_j nas Perfurações C e D. Entretanto, há o caso inverso, como nas variáveis F_j da

Perfuração B e N da Perfuração C que apresentam clara variabilidade em momentos específicos mas são consideradas constantes no processo de calibração.

Devido à natureza exponencial do BYM, um aumento intenso no valor de uma base de exponenciação – neste caso as definidas nas equações (2.2h) e (2.2f) – causa grandes valores de diferenciação no processo de integração, o que resulta em valores impraticáveis de profundidade ou até mesmo em transbordamento numérico (*overflow*).

O tratamento específico dos valores coletados nos poços de perfuração não é mérito deste trabalho. A decisão de fixar valores de variáveis, como o ocorrido neste caso, só agrega mais incertezas ao processo de calibração. Estas incertezas já vêm se somando desde as primeiras decisões de simplificação ou interpolação dos dados brutos, entre outras questões técnicas de operação dos próprios sistemas de engenharia.

Além dessas “incertezas paramétricas”, o fato do BYM produzir variações suaves no resultado das simulações pode acrescentar incertezas relativas à “inadequação do modelo”. Ademais, o processo de perfuração pode conter elementos naturalmente aleatórios e inverossímeis ao construtor do modelo, o que se apresenta como “variabilidade residual”, ou até mesmo as próprias observações podem ser errôneas (KENNEDY; O’HAGAN, 2001). Ao final, a discretização de todas estas “incertezas epistêmicas” (BEVEN, 2015) pode ser até mesmo impossível.

Lidar especificamente com a inadequação do modelo e com a variabilidade residual do processo é tarefa dos engenheiros de petróleo e foge do escopo deste trabalho. Aqui, todas as incertezas foram representadas indistintamente pela função de verossimilhança. Essas informações descartadas ao se definir um valor / polinômio para variáveis operacionais poderiam servir como clara fonte de incerteza a ser considerada pela função de verossimilhança. Para isto seria necessário reescrevê-la tendo em conta uma medida de erro ou dispersão resultantes deste processo de “preparação do BYM”.

São bons temas para trabalhos futuros várias outras alterações na definição de tal função: inicialmente poderia verificar-se o comportamento do processo de calibração ao se utilizar outras fórmulas, como

as sugeridas na Figura 1, e também com outros operadores de agregação. Enquanto o BYM apresenta resultados exponenciais sem grande variação intermediária, poder-se-ia utilizar, por exemplo, uma agregação ponderada dos valores de verossimilhança onde as profundidades finais têm maior peso comparativo que as iniciais.

Além da utilização de outras fórmulas e agregadores, é possível cogitar uma função que considere mais de um modelo de simulação no cálculo de verossimilhança de uma parametrização. Por exemplo, para cada valor de parametrização proposta, seriam simulados os modelos da Perfuração B e C simultaneamente e o cálculo da verossimilhança poderia ser uma espécie de média das diferenças entre os valores simulados e os dados reais em cada modelo. Como resultado se obteria uma parametrização θ baseada nos dados de mais de uma perfuração, o que permitiria eliminar as condições que fizeram com que os *posteriors* a_3 e a_6 na Figura 20 resultassem em distribuições multimodais – além dos motivos já comentados na apresentação dos resultados.

Esta flexibilidade de se experimentar diversas maneiras de aferir o “erro” entre dados reais e simulados exige um cuidado maior do modelador. Neste exemplo em que mais de uma Perfuração é utilizada simultaneamente, já não seria possível permitir a inferência do valor inicial de profundidade, D_0 , uma vez que eles não se corresponderiam. No caso em que a função de verossimilhança construída resulte em uma maneira muito complexa de análise de valores, ou se muitos valores do seu domínio mapearem para zero, $L(\tilde{y} | \theta) \mapsto \{0\}$, cabe então utilizar o paradigma do ABC: representá-la como na equação (2.6) e realizar a amostragem através de Monte Carlo Sequencial a fim de se explorar adequadamente o espaço paramétrico.

Em se tratando agora desta exploração, a escolha neste trabalho da função de proposição, $q(\theta^* | \theta)$, descrita na seção 3.2.1 e criticamente definida com $\sigma_v = 1/6(b_v - a_v)$, foi o resultado de tentativas e erros. Apesar de uma técnica aparentemente grotesca, tentativa e erro é a única solução para o melhor dimensionamento da exploração no algoritmo RWM, o que lhe rende vastas críticas na literatura.

Percebe-se, por exemplo, que na Perfuração A, Figura 14, o es-

paço das variáveis a_3 , a_5 , a_6 e a_8 , foi bem explorado enquanto que em a_7 e D_0 os passos foram muito pequenos e em a_2 e a_4 os passos foram grandes – gerando diversas proposições rejeitadas o que mantinha como aceito o valor anterior. Cada caso é especificado dependendo do intervalo de alta probabilidade resultante da estabilização do valor de σ_v , exemplificado na Figura 11. Este cenário é o que dificultou a utilização de uma função multivariada para a proposição conjunta dos parâmetros. Ao se fixar neste caso a matriz de covariância, teríamos a probabilidade de rejeitar toda a amostra no caso de somente um dos parâmetros ter sido amostrado fora do seu intervalo de distribuição à priori.

Uma solução para esta situação exploratória é monitorar a taxa de aceitação das proposições q , adaptando o valor de σ_v para reduzir sua abrangência quando esta taxa for muito baixa e aumentando caso contrário. Outra opção seria utilizar algum algoritmo MCMC naturalmente adaptativo como o *Adaptive Metropolis* (HAARIO; SAKSMAN; TAMMINEN, 2001), o *Differential Evolution Adaptive Metropolis* (VRUGT; BRAAK; DIKS et al., 2009) ou o *No-U-Turn Sampler* (HOFFMAN; GELMAN, 2014).

Em contrapartida, nos resultados deste trabalho, os parâmetros com alta taxa de rejeição estão situados num limite intervalar muito pequeno o que praticamente fixa o seu valor nas diversas simulações. Isto demonstra uma redução de sua influência nos valores simulados mas ainda produz bons resultados finais.

A utilização de algoritmos MCMC sempre traz à tona o questionamento sobre a eficiência da solução. A execução dos experimentos em um núcleo Intel® Core™i3-4330, com Python 2.7.6, Ubuntu 14.04 e com as últimas versões do SciPy (JONES; OLIPHANT; PETERSON et al., 2001–) e do PyMC (PATIL; HUARD; FONNESBECK, 2010) resultou em cerca de 1800 iterações por minuto. Portanto as 80 mil iterações realizadas para cada Perfuração necessitaram de aproximadamente 45 minutos para serem completadas. Este não é um tempo proporcionalmente longo se levarmos em conta que o BYM é praticamente utilizado somente na fase de planejamento da perfuração de

poços. De qualquer maneira, poder-se-ia realizar somente 45 mil iterações, continuar descartando as 10 mil iniciais como período de aquecimento e desbastá-las (*thinning*) a cada 5, ao invés de 10. Ao final resultariam mesmos 7 mil valores cuja distribuição de probabilidade seria semelhante às ilustradas nos resultados.

Soluções que utilizam *priors* conjugados (MACKAY, 2003, p. 319), como a de GIROLAMI (2008), resultam em uma fórmula fechada para a inferência direta da distribuição à posteriori. Assim, por não se tratarem de métodos iterativos, sua eficiência é máxima. Entretanto, a utilização de métodos baseados em amostragem, como o apresentado neste trabalho, permite a liberdade de se experimentar diferentes relações entre conhecimentos e incertezas sobre o domínio modelado.

Em especial sobre o objetivo geral deste trabalho, o método proposto permite propagar as incertezas existentes nos dados e inerentes ao próprio modelo. A inferência de distribuições de probabilidade para representar seus coeficientes proporciona um entendimento ampliado das diferentes parametrizações, o que permite ao modelador uma melhor escolha para valor dos coeficientes – um avanço perante os métodos anteriores que calculavam sempre valores únicos ou fixos.

BIBLIOGRAFIA

- ALAZZY, A. A.; LÜ, H.; ZHU, Y. Assessing the Uncertainty of the Xinanjiang Rainfall-Runoff Model: Effect of the Likelihood Function Choice on the GLUE Method. *Journal of Hydrologic Engineering*, p. 04015016, 2015.
- ANDRIEU, C.; BARTHELMÉ, S. et al. Some discussions of D. Fearnhead and D. Prangle's Read Paper "Constructing summary statistics for approximate Bayesian computation: Semi-automatic approximate Bayesian computation". *ArXiv e-prints*, v. Jan, p. 10, 2012. arXiv: 1201.1314v1.
- ANDRIEU, C.; FREITAS, N. de et al. An introduction to MCMC for machine learning. *Machine learning*, v. 50, n. 1-2, p. 5–43, set. 2003.
- AREHART, R. A. Drill-Bit Diagnosis With Neural Networks. *SPE Computer Applications*, v. 2, n. 4, p. 24–28, 1990.
- BAHARI, M. H. et al. Determining Bourgoyne and Young Model Coefficients Using Genetic Algorithm to Predict Drilling Rate. *Journal of Applied Sciences*, v. 8, n. 17, p. 3050–3054, 2008.
- BARRAGAN, R. V. Otimização dos parametros mecanicos nas brocas para obter o custo minimo de uma fase de um poço. Tese de doutorado. UNICAMP, 1995, p. 105.
- BEAUMONT, M. A.; CORNUET, J.-M. et al. Adaptive approximate Bayesian computation. *Biometrika*, v. 96, n. 4, p. 983–990, 2009. arXiv: 0805.2256.
- BEAUMONT, M. A.; ZHANG, W.; BALDING, D. J. Approximate Bayesian computation in population genetics. *Genetics*, v. 162, n. December, p. 2025–2035, 2002.
- BEVEN, K. J. Facets of uncertainty: epistemic uncertainty, non-stationarity, likelihood, hypothesis testing, and communication. *Hydrological Sciences Journal*, n. June, p. 150409130216006, 2015.

- BEVEN, K. J.; BINLEY, A. GLUE: 20 years on. *Hydrological Processes*, v. 28, n. 24, p. 5897–5918, 2014.
- BEVEN, K. J.; BINLEY, A. The future of distributed models: model calibration and uncertainty prediction. *Hydrological Processes*, v. 6, n. 3, p. 279–298, 1992.
- BEVEN, K. J.; FREER, J. E. Equifinality, data assimilation, and uncertainty estimation in mechanistic modelling of complex environmental systems using the GLUE methodology. *Journal of Hydrology*, v. 249, p. 11–29, 2001.
- BILGESU, H.; ALTMIS, U. et al. A New Approach to Predict Bit Life Based on Tooth or Bearing Failures. Em: *SPE Eastern Regional Meeting*. Pittsburgh, Pennsylvania: Society of Petroleum Engineers, 1998, p. 253–257.
- BILGESU, H.; TETRICK, L. et al. A New Approach for the Prediction of Rate of Penetration (ROP) Values. Em: *SPE Eastern Regional Meeting*. Lexington, Kentucky: Society of Petroleum Engineers, out. 1997.
- BLASONE, R. S. et al. Generalized likelihood uncertainty estimation (GLUE) using adaptive Markov Chain Monte Carlo sampling. *Advances in Water Resources*, v. 31, n. 4, p. 630–648, 2008.
- BLUM, M. G. B.; FRANÇOIS, O. Non-linear regression models for Approximate Bayesian Computation. *Statistics and Computing*, v. 20, p. 63–73, 2010.
- BORTOT, P.; COLES, S. G.; SISSON, S. A. Inference for Stereological Extremes. *Journal of the American Statistical Association*, v. 102, p. 84–92, 2007.
- BOURGOYNE, A. T.; CHENEVERT, M. E. et al. *Applied drilling engineering*. Richardson, TX: Society of Petroleum Engineers, 1986, p. 502.
- BOURGOYNE, A. T.; YOUNG, F. S. A Multiple Regression Approach to Optimal Drilling and Abnormal Pressure Detection. *SPE Journal*, v. 14, n. 4, p. 371–384, 1974.

- CABRAS, S.; NUEDA, M. E. C.; RULI, E. Approximate Bayesian Computation by Modelling Summary Statistics in a Quasi-likelihood Framework. *Bayesian Analysis*, p. 1–29, 2015.
- CORNUET, J. M. et al. Inferring population history with DIY ABC: A user-friendly approach to approximate Bayesian computation. *Bioinformatics*, v. 24, n. 23, p. 2713–2719, 2008. arXiv: arXiv:0804.4372v2.
- DASHEVSKIY, D.; DUBINSKY, V.; MACPHERSON, J. D. Application of Neural Networks for Predictive Control in Drilling Dynamics. Em: *SPE Annual Technical Conference and Exhibition*. Houston Texas: Society of Petroleum Engineers, 1999, p. 1–9.
- DEL MORAL, P.; DOUCET, A.; JASRA, A. An adaptive sequential Monte Carlo method for approximate Bayesian computation. *Statistics and Computing*, v. 22, n. 1, p. 1009–1020, 2012.
- DEL MORAL, P.; DOUCET, A.; JASRA, A. Sequential Monte Carlo samplers. *Journal of the Royal Statistical Society. Series B: Statistical Methodology*, v. 68, p. 411–436, 2006. arXiv: 0212648v1 [cond-mat].
- DELSMAN, J. R. et al. Uncertainty estimation of end-member mixing using generalized likelihood uncertainty estimation (GLUE), applied in a lowland catchment. *Water Resources Research*, v. 49, p. 4792–4806, 2013.
- EDALATKHAH, S.; RASOUL, R.; HASHEMI, A. Bit Selection Optimization Using Artificial Intelligence Systems. *Petroleum Science and Technology*, v. 28, n. 18, p. 1946–1956, out. 2010.
- FEARNHEAD, P.; PRANGLE, D. Constructing summary statistics for approximate Bayesian computation: Semi-automatic approximate Bayesian computation. *Journal of the Royal Statistical Society. Series B: Statistical Methodology*, v. 74, p. 419–474, 2012. arXiv: arXiv:1201.1314v1.
- FOLL, M.; BEAUMONT, M. A.; GAGGIOTTI, O. An approximate bayesian computation approach to overcome biases that arise when using

- amplified fragment length polymorphism markers to study population structure. *Genetics*, v. 179, n. June, p. 927–939, 2008.
- FONSECA, T. C. et al. A Genetic Neuro-Model Reference Adaptive Controller for Petroleum Wells Drilling Operations. Em: *International Conference on Computational Intelligence for Modelling Control and Automation and International Conference on Intelligent Agents Web Technologies and International Commerce*. Sydney, NSW, 2006.
- GANDELMAN, R. A. Predição da ROP e otimização em tempo real de parâmetros operacionais na perfuração de poços de petróleo offshore. Diss. de mestrado. UFRJ, 2012, p. 195.
- GELMAN, A. et al. *Bayesian data analysis*. 2nd. Washington, D.C.: Taylor & Francis e-Library, 2009, p. 689.
- GIDH, Y.; IBRAHIM, H. Real-Time Drilling Parameter Optimization System Increases ROP by Predicting/Managing Bit Wear. Em: *SPE Digital Energy Conference and Exhibition*. 2011.
- GIROLAMI, M. A. Bayesian inference for differential equations. *Theoretical Computer Science*, v. 408, n. 1, p. 4–16, 2008.
- HAARIO, H.; SAKSMAN, E.; TAMMINEN, J. An adaptive Metropolis algorithm. *Bernoulli*, v. 7, n. 2, p. 223–242, 2001.
- HASAN, B. M.; ABOOZAR, B.; HAMIDREZA, M. Intelligent drilling rate predictor. *International Journal of Innovative Computing, Information and Control*, v. 7, n. 4, p. 1511–1519, 2011.
- HASTINGS, W. K. Monte carlo sampling methods using Markov chains and their applications. *Biometrika*, v. 57, p. 97–109, 1970.
- HOFFMAN, M. D.; GELMAN, A. The No-U-Turn Sampler: Adaptively Setting Path Lengths in Hamiltonian Monte Carlo. *Journal of Machine Learning Research*, v. 15, n. 2008, p. 30, 2014. arXiv: 1111.4246.
- JAYNES, E. T. *Probability theory: the logic of science*. Ed. por BRETTHORST, G. L. 1ª ed. V. 27. Cambridge: Cambridge University Press, 2003, p. 727.

- JONES, E.; OLIPHANT, T.; PETERSON, P. et al. *SciPy: Open source scientific tools for Python*. 2001–. URL: <http://www.scipy.org/> (acesso em 25/01/2014).
- KENNEDY, M. C.; O'HAGAN, A. Bayesian calibration of computer models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, v. 63, n. 3, p. 425–464, 2001.
- ŁATUSZYŃSKI, K.; LEE, A. Variance bounding and geometric ergodicity of Markov chain Monte Carlo kernels for approximate Bayesian computation. *Biometrika*, v. 101, n. 3, p. 655–671, 2014. arXiv: arXiv:1210.6703v2.
- LEE, A. On the choice of MCMC kernels for approximate Bayesian computation with SMC samplers. Em: *Proceedings - Winter Simulation Conference*. 2012, p. 304–315.
- MACKAY, D. J. *Information theory, inference and learning algorithms*. 7^a ed. Cambridge University Press, 2003, p. 640.
- MARJORAM, P. et al. Markov chain Monte Carlo without likelihoods. *Proc Natl Acad Sci U.S.A.*, v. 100, n. 26, p. 15324–15328, 2003.
- MCMICHAEL, C. E.; HOPE, A. S.; LOAICIGA, H. A. Distributed hydrological modelling in California semi-arid shrublands: MIKE SHE model calibration and uncertainty estimation. *Journal of Hydrology*, v. 317, p. 307–324, 2006.
- METROPOLIS, N. et al. Equation of State Calculations by Fast Computing Machines. *The Journal of Chemical Physics*, v. 21, p. 1087–1092, 1953. arXiv: 5744249209.
- MONTANARI, A. Large sample behaviors of the generalized likelihood uncertainty estimation (GLUE) in assessing the uncertainty of rainfall-runoff simulations. *Water Resources Research*, v. 41, p. 1–13, 2005.
- MORADI, H. et al. Drilling rate prediction using an innovative soft computing approach. *Scientific Research and Essays*, v. 5, n. 13, p. 1583–1588, 2010.

- NOTT, D. J.; MARSHALL, L.; BROWN, J. Generalized likelihood uncertainty estimation (GLUE) and approximate Bayesian computation: What's the connection? *Water Resources Research*, v. 48, n. November, p. 1–7, 2012.
- PATIL, A.; HUARD, D.; FONNESBECK, C. J. PyMC: Bayesian Stochastic Modelling in Python. *Journal of statistical software*, v. 35, n. 4, p. 1–81, 2010.
- PRITCHARD, J. K. et al. Population growth of human Y chromosomes: a study of Y chromosome microsatellites. *Molecular Biology and Evolution*, v. 16, n. 12, p. 1791–1798, 1999.
- RATMANN, O.; CAMACHO, A. et al. Statistical modelling of summary values leads to accurate Approximate Bayesian Computations. *ArXiv e-prints*, n. 1, p. 35, 2013. arXiv: 1305.4283.
- RATMANN, O.; JØRGENSEN, O. et al. Using likelihood-free inference to compare evolutionary dynamics of the protein networks of *H. pylori* and *P. falciparum*. *PLoS Computational Biology*, v. 3, n. 11, p. 2266–2278, 2007.
- RODRIGUES, D. G. Um modelo de rede neuro-fuzzy baseada em funções de base radial capaz de inferir regras do tipo Mamdani. Diss. de mestrado. UFSC, 2014, p. 63.
- RODRIGUES, D. G. et al. Generating Interpretable Mamdani-type fuzzy rules using a Neuro-Fuzzy System based on Radial Basis Functions. *IEEE International Conference on Fuzzy Systems*, p. 1352–1359, 2014.
- RUBIN, D. B. *Bayesianly Justifiable and Relevant Frequency Calculations for the Applied Statistician*. 1984.
- RULI, E.; SARTORI, N.; VENTURA, L. Approximate Bayesian Computation with composite score functions. *ArXiv e-prints*, p. 1–27, 2014. arXiv: 1311.7286v2.
- SADEGH, M.; VRUGT, J. A. Bridging the gap between GLUE and formal statistical approaches: Approximate Bayesian computation. *Hydrology and Earth System Sciences*, v. 17, p. 4831–4850, 2013.

- SCHOUPS, G.; VRUGT, J. A. A formal likelihood function for parameter and predictive inference of hydrologic models with correlated, heteroscedastic, and non-Gaussian errors. *Water Resources Research*, v. 46, p. 1–17, 2010.
- SIEGMUND, K. D.; MARJORAM, P.; SHIBATA, D. Modeling DNA methylation in a population of cancer cells. *Statistical applications in genetics and molecular biology*, v. 7, Article 18, 2008.
- SISSON, S. A.; FAN, Y.; TANAKA, M. M. Sequential Monte Carlo without likelihoods. *Proceedings of the National Academy of Sciences of the United States of America*, v. 104, n. 6, p. 1760–1765, 2007.
- SUN, N.; HONG, B.; HALL, M. Assessment of the SWMM model uncertainties within the generalized likelihood uncertainty estimation (GLUE) framework for a high-resolution urban sewershed. *Hydrological Processes*, v. 28, p. 3018–3034, 2014.
- TANAKA, M. M. et al. Using approximate bayesian computation to estimate tuberculosis transmission parameters from genotype data. *Genetics*, v. 173, n. July, p. 1511–1520, 2006.
- TAVARÉ, S. et al. Inferring coalescence times from DNA sequence data. *Genetics*, v. 145, p. 505–518, 1997.
- TIERNEY, L. Markov Chains for Exploring Posterior Distributions. *The Annals of Statistics*, v. 22, n. 4, p. 1701–1728, dez. 1994.
- TONI, T. et al. Approximate Bayesian computation scheme for parameter inference and model selection in dynamical systems. *Journal of the Royal Society, Interface*, v. 6, n. July 2008, p. 187–202, 2009. arXiv: 0901.1925.
- VRUGT, J. A.; BRAAK, C. J. F. ter; DIKS, C. G. H. et al. Accelerating Markov Chain Monte Carlo Simulation by Differential Evolution with Self-Adaptive Randomized Subspace Sampling. *International Journal of Nonlinear Sciences and Numerical Simulation*, v. 10, n. March, p. 271–288, 2009.

- VRUGT, J. A.; BRAAK, C. J. F. ter; GUPTA, H. V. et al. Equifinality of formal (DREAM) and informal (GLUE) Bayesian approaches in hydrologic modeling? *Stochastic Environmental Research and Risk Assessment*, v. 23, p. 1011–1026, 2009.
- VRUGT, J. A.; SADEGH, M. Toward diagnostic model calibration and evaluation: Approximate Bayesian computation. *Water Resources Research*, v. 49, p. 4335–4345, 2013.
- WILKINSON, R. D. Approximate Bayesian computation (ABC) gives exact results under the assumption of model error. *Statistical Applications in Genetics and Molecular Biology*, v. 12, p. 129–141, 2013. arXiv: 0811.3355.
- WOOD, S. N. Statistical inference for noisy nonlinear ecological dynamic systems. *Nature*, v. 466, n. 7310, p. 1102–1104, 2010.
- ZHANG, W.; LI, T. The Influence of Objective Function and Acceptability Threshold on Uncertainty Assessment of an Urban Drainage Hydraulic Model with Generalized Likelihood Uncertainty Estimation Methodology. *Water Resources Management*, 2015.