Proceedings of the 2015 Winter Simulation Conference
L. Yilmaz, W. K. V. Chan, I. Moon, T. M. K. Roeder, C. Macal, and M. D. Rossetti, eds.

ESTIMATION OF BOURGOYNE AND YOUNG MODEL COEFFICIENTS USING MARKOV CHAIN MONTE CARLO SIMULATION

Sanjay Formighieri
Paulo José de Freitas Filho

PerfomanceLab
Department of Computing and Statistics
Universidade Federal de Santa Catarina
PO Box 476 – 88040-900. Florianópolis, SC, Brazil

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 paper therefore describes a probabilistic model and Bayesian inference conducted using Markov Chain Monte Carlo. The results were satisfactory and the probability distributions obtained offer improved insight into the influence of different coefficients on the simulation results.

1 INTRODUCTION

Drilling oil wells involves a large number of risks. Optimization of the drilling process is normally achieved by increasing the rate of penetration (ROP) in an environment that is bounded by financial costs and physical limits. Achieving the optimum ROP involves understanding a series of operational parameters such as, for example, flow rate and pressure at the bottom of the well (which are related to well cleanliness and safety), the weight exerted on the drill bit (WOB) and the rotational speed of the bit. The greater the weight and the higher the rotation of the bit, the faster the ROP will be. However, increasing these parameters can also lead to excessive wear to the bit. Considering that the greatest costs are related to expenditure on rental of operational equipment, a bit change operation can be a very expensive procedure (Gandelman 2012).

In view of this, well drilling operations must be very carefully planned and executed to ensure that they run safely and within the time-frame predicted. Several data-driven models have been developed recently in attempts to deal with the current complexity of this problem and in response to advances in monitoring technologies. Some of these approaches employ neural networks as a black-box model of ROP and the operational variables (Edalatkhah et al. 2010; Rodrigues et al. 2014); while others employ Bayesian networks for decision support (Al-yami et al. 2012) or for prediction (Lima et al. 2014). However, there are also older mathematical models that are used for analytical support in parallel with the more modern models, primarily during the well planning phase. Of these, the model that has gained greatest acceptance is the Bourgoyne and Young Model (BYM) (Bourgoyne et al. 1986) because it takes the largest number of operational parameters into account (Edalatkhah et al. 2010) and still is widely employed (Moradi et al. 2010). The BYM is a system of ordinary differential equations (ODE) that models ROP as change in depth as a function of time, dD/dt, and includes a variable $h \in [0,1]$ that represents proportional bit wear:

$$\frac{dD}{dt} = f_1 \times f_2 \times f_3 \times f_4 \times f_5 \times f_6 \times f_7 \times f_8 \tag{1a}$$

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

with,

$$f_1 = e^{a_1} = K \qquad f_5 = \left[\frac{\frac{W}{d_b} - \left(\frac{W}{d_b}\right)_t}{4 - \left(\frac{W}{d_b}\right)_t}\right]^{a_5}$$

$$f_2 = e^{a_2\,(10000-D)} \qquad f_6 = \left(\frac{N}{60}\right)^{a_6}$$

$$f_3 = e^{a_3\,D^{0.69}(g_p-9)} \qquad f_7 = e^{-a_7\,h}$$

$$f_4 = e^{a_4\,D(g_p-\rho_c)} \qquad f_8 = \left(\frac{F_j}{1000}\right)^{a_8},$$

where the term f_i represents an influence on the determination of ROP weighted by a_i. Originally the coefficients a_1 through a_4 are multiplied by 2.303 to convert their functions into powers of base 10. In these equations,

a_1 to a_8 = coefficients that must be chosen on the basis of previous drilling experience;
D = true vertical well depth (ft);
g_p = pore pressure gradient (lbm/gal);
ρ_c = equivalent mud density (lbm/gal);
W = weight on bit (1000 lbf);
d_b = bit diameter (in);
$(W/d_b)_t$ = threshold of weight on bit at which the bit begins to drill;
N = rotary speed (rpm);
h = fractional tooth wear of the bit, for which $h = 0$ at zero wear;
F_j = jet impact force (lbf),

and,

H_1, H_2, $(W/d_b)_m$ = constants for physical specifications of bit;
τ_H = formation abrasiveness constant (hr).

In possession of the values for the constants relative to the bit, the operational parameters and those variables that can be observed during drilling, the values of the coefficients remain to be determined. When Bourgoyne and Young (1974) published their model, they suggested that the coefficients should be determined by multivariate regression of data from drilling of similar wells. However, it has been shown that this method can produce coefficient values that are negative or zero and, as such, physically meaningless (Moradi et al. 2010). For example, if a certain coefficient is zero, it would mean that increasing the weight on the bit would not affect ROP and if the coefficient were negative, increasing weight would reduce ROP. Bahari et al. (2008) and Hasan et al. (2011) attempted to determine coefficients that respect physical limits by estimating them with genetic algorithms.

It is common among the studies cited so far for the data employed to be acquired from observations of the drilling process dispersed in time. Such observations do not take the dynamics inherent in the system into consideration, restricting these studies to employing data for scenarios in which the bit is still considered new, at the start of drilling ($h = 0$), or when it is found to be completely worn ($h = 1$) when it is removed from the well. Additionally, it is sensible to allow for drilling scenarios in which the sensors provide data that is potentially noisy. As a result, the values for the system's informative variables can only be the result of approximate calculations. As such, observed ROP values include uncertainties that are unlikely to be identified through the variables.

Since drilling is a physical process, it can include elements that are inherently not predictable by the model. This imperfection generates a residual variability (Kennedy and O'Hagan 2001) that opens the way to an equifinality of models and variables (Beven and Binley 1992). This means that a range of parameterizations should be considered, rather than a single solution.

One method of dealing with this uncertainty is to treat the coefficients as probability distributions rather than fixed values (Vyshemirsky and Girolami 2008), with the result that determination of these coefficients becomes a process of inference of their distribution functions. This can be achieved by specifications within a Bayesian framework, which will be described in detail in Section 2. Section 3 will describe experiments conducted to determine the coefficients in this manner with an analysis of the results, and Section 4 closes with a summary of what has been achieved and reflections on how treatment of uncertainty allows for improved understanding of the values of these parameters.

2 BAYESIAN INFERENCE

The BYM is a system comprising two ODEs which, given the initial values $y_0 = [D_0, h_0]$ and the parameters $\theta = [a_1, \ldots, a_8]$, can be solved numerically (simulated) with the solution $y_t = \mathscr{S}(y_0, \theta, t)$ in a discrete-time vector $t = [t_1, \ldots, t_n]$. Here the initial value of h will always be zero, $h_0 = 0$, because only new bits are used, which reduces the observed state space to $y_t = D_t$.

When the depth observations are provided, $\tilde{y} = \{\tilde{D}_t \mid t = 1, \ldots, n\}$, interest moves on to inference of the posterior probability distribution $f(\theta \mid \tilde{y})$. According to Bayes' theorem (Gelman et al. 2009),

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

where the prior probability distribution $\pi(\theta)$, represents the initial knowledge available on parameters θ, and where $L(\tilde{y} \mid \theta)$ is the likelihood function for the parameters, that represents how likely is the observed data given the outputs of the model parameterized by θ. One way of making inferences about f is with Markov Chain Monte Carlo (MCMC) simulations, using an algorithm such as the Metropolis-Hastings (MH) method (Hastings 1970; Metropolis et al. 1953):

MH1 Given the current state with θ, propose a move to θ^*, according to a transition function $q(\cdot \mid \theta)$.

MH2 Calculate

$$\alpha = \min\left(1, \frac{L(\tilde{y} \mid \theta^*)\pi(\theta^*)q(\theta^* \mid \theta)}{L(\tilde{y} \mid \theta)\pi(\theta)q(\theta \mid \theta^*)}\right) \tag{2}$$

MH3 Go to state of θ^* with probability α, else remain at θ; go to MH1.

This algorithm generates a Markov chain which will begin to provide observations of $f(\theta \mid \tilde{y})$ as its stationary distribution after an uncertain number of iterations.

Estimation of the parameters for differential equation models using Bayesian inference is not an innovation in the literature (Kulhavý 2007, Girolami 2008, Golightly and Wilkinson 2011). However, one difficulty lies in defining the likelihood function that best describes the proximity between the simulated values and the corresponding real data (Vrugt and Sadegh 2013).

With Approximate Bayesian Computation (ABC) methods (Pritchard et al. 1999; Beaumont et al. 2002; Marjoram et al. 2003; Sisson et al. 2007) the likelihood function is ignored and proximity is formalized as a distance ρ between summary statistics, S, for the values. Inference are generally performed using methods based on sequential sampling (Beaumont et al. 2009; Del Moral et al. 2012; Lee 2012) and acceptable values of θ satisfy

$$\rho(S(y), S(\tilde{y})) < \varepsilon, \tag{3}$$

with $\varepsilon \to 0$. In this case, the posterior distribution inferred is $f(\theta \mid \rho(S(y), S(\tilde{y})) < \varepsilon)$.

Some ABC methods conduct sampling using an MCMC (Marjoram et al. 2003) in which a test of condition (3) is applied to the values of y simulated by the model parameterized by θ^*, before step MH2. Additionally, in these models the proportion at MH2 depends only on prior, $\pi(\cdot)$, and the transition function $q(\cdot \mid \cdot)$. In special cases, in which the transition function is symmetrical, $q(\theta^* \mid \theta) = q(\theta \mid \theta^*)$, and the prior is uniform, $\pi(\theta) = \pi(\theta^*)$, so $\alpha = 1$ in (2) and the condition for acceptance of samples is entirely determined by (3).

In order to take full advantage of an MCMC sampler, it is necessary to define a likelihood function. Using the generalized likelihood uncertainty estimation (GLUE) methodology (Beven 2006), a synthetic likelihood function (Wood 2010) is defined to determine the extent to which the simulated values fit the observed data, taking into account possible modeling or measurement errors. Subject to the condition that errors are not correlated, normally distributed and with constant variance, σ_p^2, the likelihood function can be written as

$$L(\tilde{y} \mid \theta) = \prod_{t=1}^{n} \frac{1}{\sqrt{2\pi\hat{\sigma}_p^2}} \exp\left[-\frac{1}{2\hat{\sigma}_p^2}(\tilde{y}_t - y_t)^2\right], \tag{4}$$

where $\hat{\sigma}_p$ is an estimate of the standard deviation of error, which can be predetermined or inferred together with the values of θ (Vrugt et al. 2009). This measure of likelihood in the GLUE framework is similar to the use of squared errors in (3), $\rho(S(\tilde{y}), S(y)) = \sum_{t=1}^{n}(\tilde{y}_t - y_t)^2$, using ABC (Toni et al. 2009). Additionally, adding $\hat{\sigma}_p$ to the values to be inferred is similar to adding ε to the state space of the sampler, as proposed by Bortot et al. (2007).

Formal descriptions of the similarities between GLUE and ABC can be found in work by Nott et al. (2012), Sadegh and Vrugt (2013) and Vrugt and Sadegh (2013). Approximate Bayesian Computation methods have been used to calibrate models in the fields of genetics (Siegmund et al. 2008), epidemiology (Blum and François 2010) and population biology (Ratmann et al. 2007), among others; and GLUE has been applied to environmental problems (Delsman et al. 2013; Alazzy et al. 2015; Zhang and Li 2015).

2.1 The specific probabilistic model

The first element of note is that function f_1 in (1a) is a constant known as drillability. Its value will be numerically equal to the ROP when the system is operated and observed under conditions that result in all other terms, f_2 through f_8, being equal to 1 and, as a consequence, independent of the values of their coefficients. Its value is calculated by always using standard values (or values from similar wells) for all of the other coefficients, and isolating it in (1a), $f_1 = \frac{dD}{dt}/f_2 \times f_3 \times \ldots \times f_8$, populated with real data.

The physical constants relative to the bit, H_1, H_2 and $(W/d_b)_m$, are taken from a table of values suggested by Bourgoyne et al. (1986), p. 218. The value of τ_H is calculated by integrating (1b) with respect to t:

$$\tau_H = \frac{t_b}{J_2\left(h_f + H_2 h_f^2/2\right)}, \qquad J_2 = \left(\frac{60}{N}\right)^{H_1}\left[\frac{\left(\frac{W}{d_b}\right)_m - \left(\frac{W}{d_b}\right)}{\left(\frac{W}{d_b}\right)_m - 4}\right]\left(\frac{1}{1 + H_2/2}\right), \tag{5}$$

where h_f is the final tooth wear observed for the bit and t_b is the time taken to reach that state. Each inference run is conducted using the data for one full bit run, in which the bit was subject to normal wear and no defects occurred.

For the remaining variables $(N, W, F_j, \ldots)$ mean values are used if there is little or no variation, but when a variable, v, exhibits potentially non-negligible variability over time, a polynomial $P_v^*(t) = (c_n, \ldots, c_1, c_0)$

of n degree is defined with c_i coefficients, determined by least squares, to describe its tendency. While there are more sophisticated methods for fitting and selecting polynomials (Girolami 2008), in this study the degree of the polynomial was adjusted in such a way as to prevent overfitting. The first derivative is obtained from the polynomial P_v^*, defined for simplicity as $dP_v^*/dt = P_v$, with degree $n-1$, and added to the BYM state space with its respective initial value c_0.

The probabilistic model described here also allows for possible variation in the initial value of depth, D_0, since the observed values do not contain information at $t = 0$. The likelihood function described in (4) is used and its standard deviation is also added to the values to be inferred. A high standard deviation for the likelihood function will allow sub-optimum parametrization configurations to be considered in cases in which the BYM is unable to perfectly simulate the real data. However, this could introduce bias to the estimation since the models are compared with different levels of variance (Fearnhead and Prangle 2012; Andrieu et al. 2012). In order to minimize this bias it is necessary to verify the values of σ_p sampled and possibly only accept samples $\{\theta_i \mid \sigma_p < \sigma_T\}$ for a certain threshold, σ_T (Bortot et al. 2007).

As such, the parameter space is $\theta = (D_0, a_2, \ldots, a_8, \sigma_p)$, where all of the priors are non-informative, $\pi(\cdot) \sim Uniform$, with the limits shown in Table 1.

Table 1: Bounds of priors. The interval for initial value, D_0, is determined to a limit of 0.7% of the true value of initial depth, $\tilde{D}_1$. The maximum value of the likelihood function standard deviation is set at 1% of $\tilde{D}_0$.

Parameter	Bounds of the priors $([a,b])$
D_0	$\left[\tilde{D}_1 \pm 0.007\tilde{D}_1\right]$
a_2	$\left[2.303 \times 10^{-6}, 0.012\right]$
a_3	$\left[2.303 \times 10^{-8}, 0.021\right]$
a_4	$\left[2.303 \times 10^{-6}, 2.303 \times 10^{-3}\right]$
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	$\left[10^{-4}, 0.01\tilde{D}_1\right]$

A graphical representation of the probabilistic model constructed is shown in Figure 1.

The inference of the posterior probability distribution, $f(\theta \mid \tilde{y})$, was performed by using the Metropolis random walk algorithm (Andrieu et al. 2003; Tierney 1994), which is a specific case of the MH algorithm in which the transition function, $q(\cdot \mid \theta) \sim Normal(\theta, \sigma_v)$, is symmetrical, so that $q(\theta^* \mid \theta)/q(\theta \mid \theta^*) = 1$. Additionally, since in this case the priors have constant probability values within the interval limits, the ratio $\pi(\theta^*)/\pi(\theta)$ becomes an indicative function: when q proposes a value within the limits, priors will be constant, otherwise the sample is rejected. Samples of q are independent for each variable v in Table 1, with standard deviation $\sigma_v = 1/6(b_v - a_v)$ and initial value $(b_v - a_v)/2$. Therefore, for the case starting from the centre of the interval $([a_v, b_v])$, the transition function will cover 99.73% of possible values.

3 EXPERIMENTS AND RESULTS

Data were acquired for three drill runs from the drilling of three deepwater wells, A, B and C. Analysis of these data showed that all but one of the operational variables remained constant. The exception was W, for which a polynomial was fitted in the hope that it would increase the model's information content. The values of the variables used in the BYM and the polynomials constructed for each well are shown in Table 2.

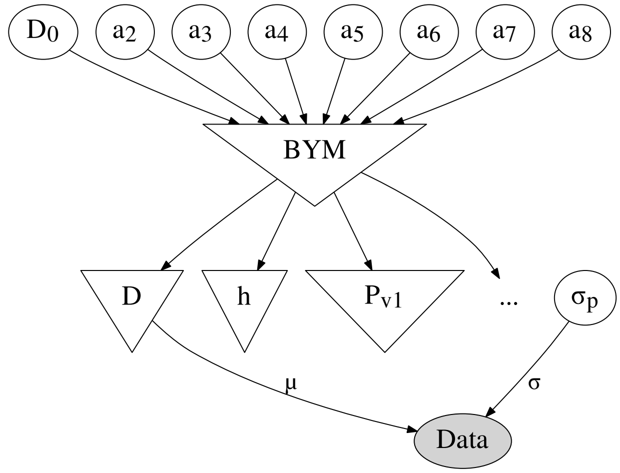

Figure 1: Graphical representation of the probabilistic model. The parameters θ to be inferred are represented by ellipses. The shaded ellipse is the likelihood, and the inverted triangles are components of the BYM.

Table 2: Parametrization of operational and observed variables for wells A, B and C. The variable W was represented by 4th degree polynomials for wells A and B, and by a 7th degree polynomial for the well C. For all wells, $a_1 = 2.4$.

Well	$H1$	$H2$	$(W/d_b)_m$	τ_H	d_b	$\tilde{D}_1$	g_p	ρ_c	W	N	F_j
A				581.87	12.25	16510	8.314	10.014	$P^4_{(A)}$	144.98	3431.82
B	1.5	2.0	10.0	256.79	8.5	16831	9.813	10.895	$P^4_{(B)}$	191.91	3491.19
C				129.57	9.0	16274	10.21	10.005	$P^7_{(C)}$	81.69	4161.76

$$P^4_{(A)} = \left[1.48\times10^{-7}, 1.1\times10^{-4}, -0.022, 1.239, 21.53\right]$$
$$P^4_{(B)} = \left[1.68\times10^{-6}, 2.5\times10^{-4}, -0.045, 1.839, 11.4\right]$$
$$P^7_{(C)} = \left[2.4\times10^{-9}, -6.4\times10^{-7}, 6.9\times10^{-5}, -0.0038, 0.1182, -1.891, 13.98, 4.873\right]$$

For each well, a total of 80000 iterations of the RWM algorithm were run for parameter inference. In all cases, the initial behavior observed was as illustrated in Figure 2. The initial value of σ_p is relatively high, allowing a variety of different parameterizations to be accepted. When the value of σ_p drops, condensing the scope off the likelihood function (4), parameterizations that previously had been accepted are more severely penalized. From this point, the values of θ migrate to the high probability areas that will be used for the analysis.

The first 1000 iterations were therefore discarded as burn-in and the resulting statistics for each parameter inferred are shown in Table 3.

The values chosen to illustrate the results are the median and the 25th and 75th quantiles. A graphical illustration of the posterior distribution values for well B is shown in Figure 3.

It can be observed that some of the parameters have multimodal distributions, as is the case of a_4 and a_2. This is because it is possible for them to be correlated in the BYM equation. The exponential nature of the system of equations (1) allows one term to compensate for another in the final result if its constituent variables are constants. Since the only variable term used in these examples was W, there is a greater chance that other coefficients, with constant bases, will be correlated.

In compensation, as shown in Figure 2, the values that remain after exclusion of the burn-in period are restricted to a relatively narrow band, in comparison with the interval occupied by the priors. This allows

Table 3: Belief intervals inferred for the BYM coefficients, $[a_2, \dots, a_8]$, for the standard deviation of the likelihood function, σ_p, and for the initial value, D_0, for the three wells.

Variable	Quantil	Well		
		A	B	C
a_2	25th	$5.17e-05$	$5.76e-05$	0.00103
	50th	$5.28e-05$	$6.32e-05$	0.00103
	75th	$6.02e-05$	$7.29e-05$	0.00103
a_3	25th	$2.11e-06$	0.00202	0.00162
	50th	$5.09e-06$	0.00206	0.00172
	75th	$1.02e-05$	0.00208	0.00193
a_4	25th	$8.65e-05$	0.000195	0.000834
	50th	$8.71e-05$	0.000196	0.000838
	75th	$8.83e-05$	0.000199	0.000845
a_5	25th	0.30	0.31	0.40
	50th	0.31	0.32	0.51
	75th	0.31	0.34	0.65
a_6	25th	1.49	1.45	0.40
	50th	1.50	1.48	0.73
	75th	1.50	1.49	1.36
a_7	25th	0.39	1.01	0.11
	50th	0.63	1.18	0.11
	75th	0.71	1.28	0.13
a_8	25th	0.89	0.85	0.49
	50th	0.90	0.88	0.53
	75th	0.90	0.89	0.56
σ_p	25th	9.24	16.82	6.44
	50th	9.82	17.74	6.78
	75th	11.03	18.77	7.18
D_0	25th	16527.4	16858.4	16270.4
	50th	16529.5	16861.5	16271.7
	75th	16536.8	16866.1	16273.0

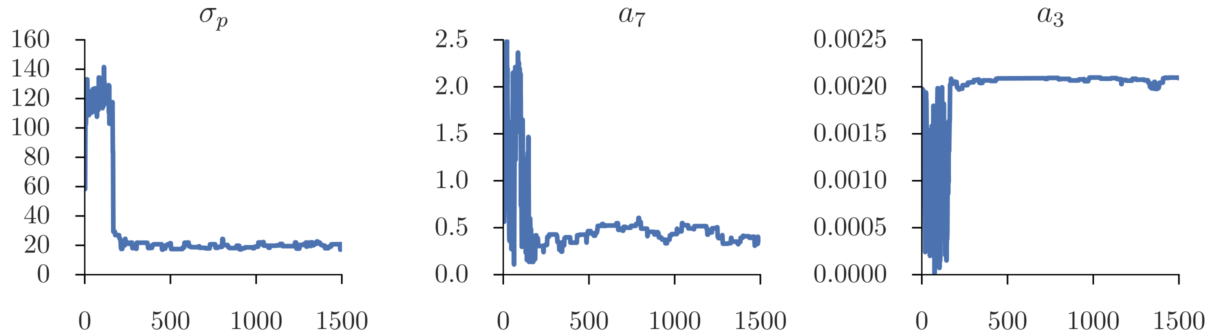

Figure 2: First 1500 iterations for well B. This figure illustrates an example of how the chain mixes when σ_p attains "lower" values (≈ 20): all variables migrate to their high probability areas.

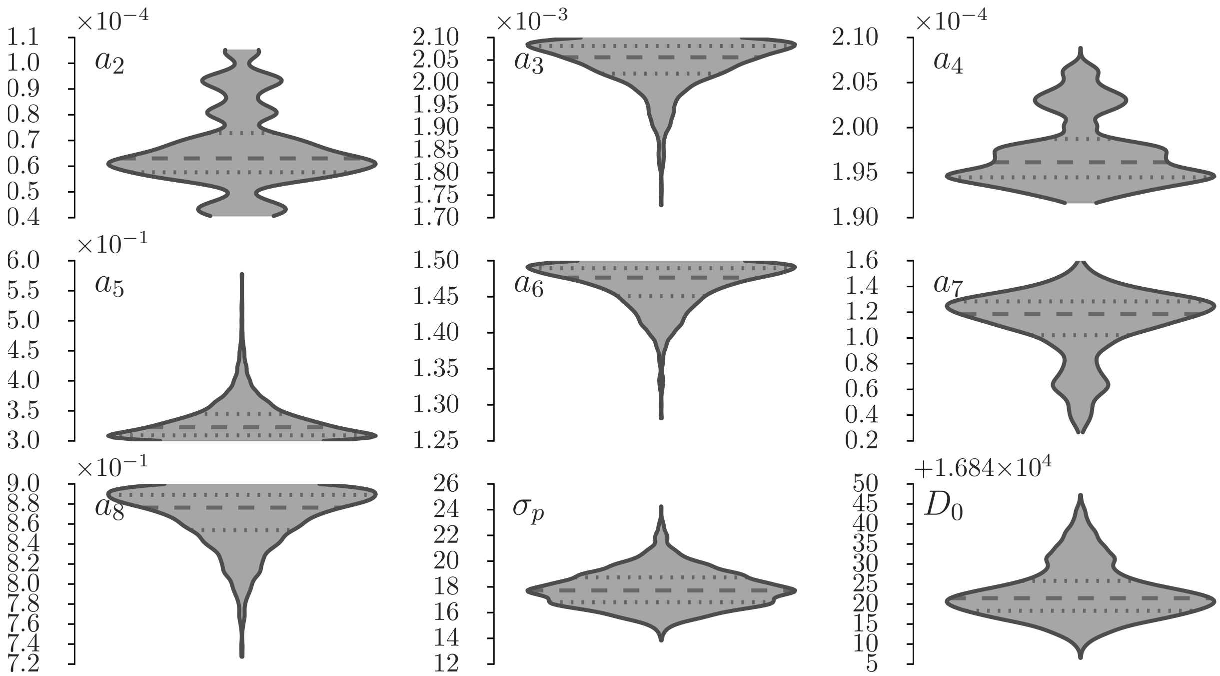

Figure 3: Violin plots of the posterior distributions for well B. The dashed line illustrates the median and the dotted lines represent the 25th and 75th quantiles.

the medians of potentially multimodal parameter distributions to be inferred, obtaining BYM simulation results with little variation. The root mean squared error and the mean absolute percentage between the median parameterized simulation results and the drilling data, as illustrated in Figure 4, are shown in Table 4.

The RMSE values correspond to standard deviations between observed and simulated data and as such, because of the likelihood function employed in this case, they are the same as the values for the posterior probability distributions of σ_p. This value illustrates the extent to which the BYM simulations for each well approach the true values.

Table 4: Errors between the median parameterized simulation results and the drilling data.

Well	RMSE	MAPE
A	9.73661	0.0454%
B	17.2434	0.0881%
C	6.6368	0.0325%

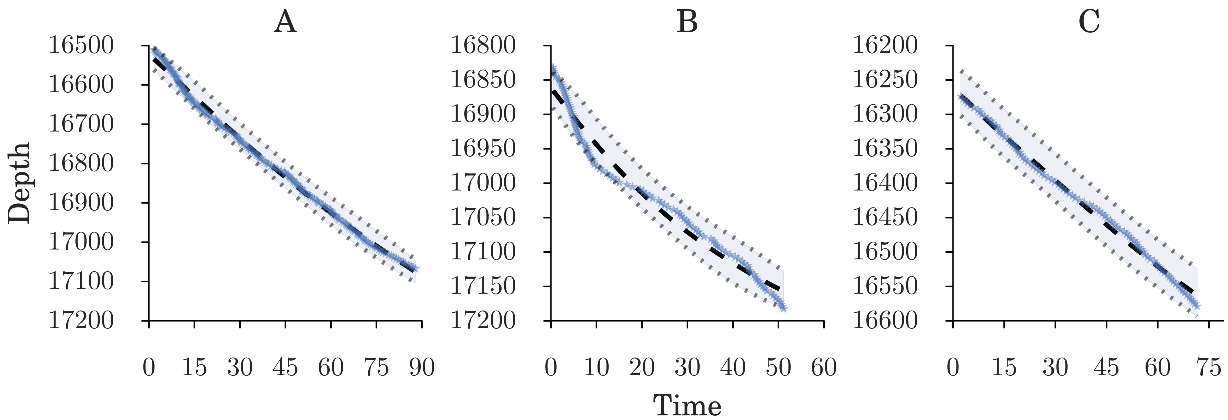

Figure 4: Results of BYM simulations. The blue lines illustrate the real drilling data. The dashed line illustrates the median simulation posteriors and the dotted lines represent quantiles 2.5 and 97.5. The shaded area is the 95% Highest Posterior Density (HPD) interval.

4 CONCLUSION

This study estimated the BYM coefficients using Bayesian inference by MCMC. Simulations were run to provide estimates for three drill runs from three different wells in deep water regions.

It can be observed that the true values shown in Figure 4 do not strictly adhere to the curve plotted by the BYM. This could indicate measurement limitations or a range of different types of errors linked to the model itself (Kennedy and O'Hagan 2001). These errors signal the uncertainties that must be considered in the process of determination of the model's coefficients. Analysis of the parameters' posterior probability distributions enabled inference of an interval of values that produced acceptable simulation results. This means that the BYM is capable of predicting, for a given interval of time, a range of depths that the bit could possibly reach, within which the observed depth data is included.

As such, the treatment of uncertainty used to determine the BYM coefficients in this study offers a better understanding of the model than earlier research that resulted in fixed values for these parameters.

ACKNOWLEDGMENTS

We gratefully acknowledge the financial support of the National Council for the Improvement of Higher Education through their Social Demand Scholarship Program (CAPES/DS), and the Postgraduate Program at Universidade Federal de Santa Catarina for the opportunity.

REFERENCES

Al-yami, A. S., J. Schubert, and A. Texas. 2012. "Using Bayesian Network to Model Drilling Fluids Practices in Saudi Arabia". In *SPE International Production and Operations Conference and Exhibition*, Number May, 14–16. Doha, Qatar.

Alazzy, A. A., H. Lü, and Y. Zhu. 2015, February. “Assessing the Uncertainty of the Xinanjiang Rainfall-Runoff Model: Effect of the Likelihood Function Choice on the GLUE Method”. *Journal of Hydrologic Engineering*.

Andrieu, C., S. Barthelmé, N. Chopin, J. Cornebise, A. Doucet, M. A. Girolami, I. Kosmidis, A. Jasra, A. Lee, J.-M. Marin, P. Pudlo, C. P. Robert, M. Sedki, and S. S. Singh. 2012. “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* Jan:10.

Andrieu, C., N. de Freitas, A. Doucet, and M. I. Jordan. 2003, September. “An introduction to MCMC for machine learning”. *Machine learning* 50 (1-2): 5–43.

Bahari, M. H., A. Bahari, F. N. Moharrami, and N. Sistani. 2008. “Determining Bourgoyne and Young Model Coefficients Using Genetic Algorithm to Predict Drilling Rate”. *Journal of Applied Sciences* 8 (17): 3050–3054.

Beaumont, M. A., J.-M. Cornuet, J.-M. Marin, and C. P. Robert. 2009. “Adaptive approximate Bayesian computation”. *Biometrika* 96 (4): 983–990.

Beaumont, M. A., W. Zhang, and D. J. Balding. 2002. “Approximate Bayesian computation in population genetics”. *Genetics* 162 (December): 2025–2035.

Beven, K. J. 2006. “A manifesto for the equifinality thesis”. *Journal of Hydrology* 320:18–36.

Beven, K. J., and A. Binley. 1992. “The future of distributed models: model calibration and uncertainty prediction.”. *Hydrological Processes* 6 (3): 279–298.

Blum, M. G. B., and O. François. 2010. “Non-linear regression models for Approximate Bayesian Computation”. *Statistics and Computing* 20:63–73.

Bortot, P., S. G. Coles, and S. A. Sisson. 2007. “Inference for Stereological Extremes”. *Journal of the American Statistical Association* 102:84–92.

Bourgoyne, A. T., M. E. Chenevert, K. K. Millheim, and F. S. Young. 1986. *Applied drilling engineering*. Richardson, TX: Society of Petroleum Engineers.

Bourgoyne, A. T., and F. S. Young. 1974. “A Multiple Regression Approach to Optimal Drilling and Abnormal Pressure Detection”. *SPE Journal* 14 (4): 371–384.

Del Moral, P., A. Doucet, and A. Jasra. 2012. “An adaptive sequential Monte Carlo method for approximate Bayesian computation”. *Statistics and Computing* 22 (1): 1009–1020.

Delsman, J. R., G. H. P. Oude Essink, K. J. Beven, and P. J. Stuyfzand. 2013. “Uncertainty estimation of end-member mixing using generalized likelihood uncertainty estimation (GLUE), applied in a lowland catchment”. *Water Resources Research* 49:4792–4806.

Edalatkhah, S., R. Rasoul, and A. Hashemi. 2010, October. “Bit Selection Optimization Using Artificial Intelligence Systems”. *Petroleum Science and Technology* 28 (18): 1946–1956.

Fearnhead, P., and D. Prangle. 2012. “Constructing summary statistics for approximate Bayesian computation: Semi-automatic approximate Bayesian computation”. *Journal of the Royal Statistical Society. Series B: Statistical Methodology* 74:419–474.

Gandelman, R. A. 2012. *Predição da ROP e otimização em tempo real de parâmetros operacionais na perfuração de poços de petróleo offshore*. Ph. D. thesis, UFRJ.

Gelman, A., J. B. Carlin, H. S. Stern, and D. B. Rubin. 2009. *Bayesian data analysis*. 2nd ed. Washington, D.C.: Taylor & Francis e-Library.

Girolami, M. A. 2008. “Bayesian inference for differential equations”. *Theoretical Computer Science* 408 (1): 4–16.

Golightly, A., and D. J. Wilkinson. 2011. “Bayesian parameter inference for stochastic biochemical network models using particle Markov chain Monte Carlo”. *Interface Focus* 1 (September): 807–820.

Hasan, B. M., B. Aboozar, and M. Hamidreza. 2011. “Intelligent drilling rate predictor”. *International Journal of Innovative Computing, Information and Control* 7 (4): 1511–1519.

Hastings, W. K. 1970. “Monte carlo sampling methods using Markov chains and their applications”. *Biometrika* 57:97–109.

Kennedy, M. C., and A. O'Hagan. 2001. "Bayesian calibration of computer models". *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 63 (3): 425–464.

Kulhavý, R. 2007. "Bayesian Analysis of Stochastic System Dynamics". In *25th International Conference of the System Dynamics*, edited by J. Sterman, R. Oliva, R. S. Langer, J. I. Rowe, and J. M. Yanni. Boston, Massachusetts.

Lee, A. 2012. "On the choice of MCMC kernels for approximate Bayesian computation with SMC samplers". In *Proceedings of the 2012 Winter Simulation Conference*, edited by C. Laroque, J. Himmelspach, and R. Pasupathy, 304–315. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Lima, M. D. C., S. M. Nassar, P. J. d. Freitas Filho, and C. M. C. Jacinto. 2014. "Heuristic discretization method for bayesian networks". *Journal of Computer Science* 10 (5): 869–878.

Marjoram, P., J. Molitor, V. Plagnol, and S. Tavaré. 2003. "Markov chain Monte Carlo without likelihoods". *Proc Natl Acad Sci U.S A* 100 (26): 15324–15328.

Metropolis, N., A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller. 1953. "Equation of State Calculations by Fast Computing Machines". *The Journal of Chemical Physics* 21:1087–1092.

Moradi, H., M. H. Bahari, M. Bagher, and N. Sistani. 2010. "Drilling rate prediction using an innovative soft computing approach". *Scientific Research and Essays* 5 (13): 1583–1588.

Nott, D. J., L. Marshall, and J. Brown. 2012. "Generalized likelihood uncertainty estimation (GLUE) and approximate Bayesian computation: What's the connection?". *Water Resources Research* 48 (November): 1–7.

Pritchard, J. K., M. T. Seielstad, A. Perez-Lezaun, and M. W. Feldman. 1999. "Population growth of human Y chromosomes: a study of Y chromosome microsatellites.". *Molecular Biology and Evolution* 16 (12): 1791–1798.

Ratmann, O., O. Jø rgensen, T. Hinkley, M. Stumpf, S. Richardson, and C. Wiuf. 2007. "Using likelihood-free inference to compare evolutionary dynamics of the protein networks of H. pylori and P. falciparum". *PLoS Computational Biology* 3 (11): 2266–2278.

Rodrigues, D. G., G. Moura, C. M. C. Jacinto, P. J. d. Freitas Filho, and M. Roisenberg. 2014. "Generating Interpretable Mamdani-type fuzzy rules using a Neuro-Fuzzy System based on Radial Basis Functions". *IEEE International Conference on Fuzzy Systems*:1352–1359.

Sadegh, M., and J. A. Vrugt. 2013. "Bridging the gap between GLUE and formal statistical approaches: Approximate Bayesian computation". *Hydrology and Earth System Sciences* 17:4831–4850.

Siegmund, K. D., P. Marjoram, and D. Shibata. 2008. "Modeling DNA methylation in a population of cancer cells.". *Statistical applications in genetics and molecular biology* 7:Article 18.

Sisson, S. A., Y. Fan, and M. M. Tanaka. 2007. "Sequential Monte Carlo without likelihoods.". *Proceedings of the National Academy of Sciences of the United States of America* 104 (6): 1760–1765.

Tierney, L. 1994, December. "Markov Chains for Exploring Posterior Distributions". *The Annals of Statistics* 22 (4): 1701–1728.

Toni, T., D. Welch, N. Strelkowa, A. Ipsen, and M. P. H. Stumpf. 2009. "Approximate Bayesian computation scheme for parameter inference and model selection in dynamical systems.". *Journal of the Royal Society, Interface* 6 (July 2008): 187–202.

Vrugt, J. A., and M. Sadegh. 2013. "Toward diagnostic model calibration and evaluation: Approximate Bayesian computation". *Water Resources Research* 49:4335–4345.

Vrugt, J. A., C. J. F. ter Braak, H. V. Gupta, and B. A. Robinson. 2009. "Equifinality of formal (DREAM) and informal (GLUE) Bayesian approaches in hydrologic modeling?". *Stochastic Environmental Research and Risk Assessment* 23:1011–1026.

Vyshemirsky, V., and M. A. Girolami. 2008. "Bayesian ranking of biochemical system models". *Bioinformatics* 24 (6): 833–839.

Wood, S. N. 2010. "Statistical inference for noisy nonlinear ecological dynamic systems.". *Nature* 466 (7310): 1102–1104.

Zhang, W., and T. Li. 2015. "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* 29:2059–2072.

AUTHOR BIOGRAPHIES

SANJAY FORMIGHIERI is a Masters student at the Department of Computing and Statistics at the Universidade Federal de Santa Catarina (UFSC), Brazil. His research interests lie in the areas of machine learning and probabilistic methods. His email address is sanjay.f@posgrad.ufsc.br.

PAULO FREITAS is an associate professor at the Department of Computing and Statistics at the Universidade Federal de Santa Catarina (UFSC), Brazil. He received a doctoral degree in production engineering from UFSC in 1994. His research interests include simulation of computer systems for performance improvement, Monte Carlo methods, risk modeling and simulation, analysis for input modeling, and output analysis. He is a member of the Society for Computer Simulation (SCS) and the Brazilian Computer Society (SBC). His email address is freitas.filho@ufsc.br.