



Parameter Estimation for a Gas Lifting Oil Well Model Using Bayes' Rule and the MetropolisHastings Algorithm

Zhe Ban¹ Ali Ghaderi² Nima Janatian¹ Carlos Pfeiffer¹

¹*Department of Electrical Engineering, IT and Cybernetics, University of South-Eastern Norway, 3918 Porsgrunn, Norway. E-mail: {zhe.ban, nima.janatianghadikolaei, carlos.pfeiffer}@usn.no*

²*Department of Mathematics and Science Education, University of South-Eastern Norway, 3918 Porsgrunn, Norway. E-mail: ali.ghaderi@usn.no*

Abstract

Oil well models are frequently used in the oil production process. Estimation of unknown parameters of these models has long been a question of great interest in the oil industry field. Data collected from an oil well system can be useful for identifying and characterizing the parameters in the corresponding model. In this article, we present a solution to estimate the parameters and uncertainty of a gas lifting oil well model by designing Bayesian inference and using the Metropolis-Hastings algorithm. To present and evaluate the estimation, the performance of the chains and the distributions of the parameters were shown, followed by posterior predictive distributions and sensitivity analysis. Compared with the conventional maximum likelihood estimation methods that tried to identify one optimum value for each parameter, more information of the parameters is obtained by using the proposed model. The insights gained from this study can benefit the optimization and advanced control for the oil well operation.

Keywords: Gas Lifting Oil Well, Parameter Estimation, Markov Chain Monte Carlo.

1 Introduction

Gas lifting is one of the dominant methods for oil extraction, especially for light oil. A simplified diagram for a gas lifting well is shown in Fig. 1. A subsurface reservoir is filled with hydrocarbons, and the well into the reservoir helps extract hydrocarbons from the reservoir. In a gas lifted process, a compressor injects gas to a distribution manifold connecting to the outer pipe of the tubing, or annulus. The gas from the annulus passes to tubing, mixing with the oil and water coming from the reservoir to produce a low-density fluid that flows out of the well.

The pressure difference between the reservoir and bottom hole is critical for oil production. The density reduction of the multi-phase fluid in the tubing

helps to increase the pressure difference. This is the main reason that gas is injected through the tubing. The injected gas also reduces the flow resistance Brown (1977), which also benefits pushing the oil from the reservoir towards the gathering manifold. The density of the fluid in the tubing plays an important role in the gas lifting model. After producing the fluid from the well, it flows to the separator where the gas, oil and water are separated. The final stage of the flow process is recirculating the separated gas to the compressor for the next production cycle.

Modelling is essential for describing the oil production process and benefits the oil production operations. With analysis of physical properties, mathematical representation is developed as the model of the system. An accurate model has a pivotal role in optimization

and advanced control, especially model-based control strategies such as Model Predictive Control. Parameter estimation has long been a question of great interest in the modelling fields. Apart from parameters of models, initial states and inputs are also common targets to be identified in the modelling step, which also demand similar methods. Therefore, the meaning of parameter estimation has been broadened to refer to evaluation of unknown constants, which embodies a multitude of inputs and initial states in systems. It is well established in a variety of studies, that those parameters, initial states and inputs are constants all the time or in a certain period of time.

Statistical methods are commonly used for estimating model parameters. Among these methods, Bayesian estimation [Lindley \(1961\)](#); [Jeffreys \(1935\)](#) has recently been applied to estimate parameters of models in different engineering applications, such as internal combustion engines [Khaliullin et al. \(2019\)](#), flow meters [Gyamfi et al. \(2018\)](#), power grids [Hu et al. \(2021\)](#) and so on. One of the main advantages of Bayesian estimation is that it offers an explicit way to include engineering knowledge of the parameters in the estimation process, and provides effective ways of addressing uncertainty. However, for complex models such like many oil well models, Bayes estimation requires intensive computations [Cumming and Goldstein \(2010\)](#).

Early research publications paid particular attention to simplifying the original Bayes analysis to decrease the computational cost. Bayes linear approach is one of the methods, where the estimation (mean) and its beliefs (variance) are linearly updated using the outputs from the simulator without producing the posterior distribution [Craig et al. \(1996\)](#). The update process is also called linear fitting. The acceptance of outputs is vital for the estimation. However, the criteria of outputs acceptance is difficult to design and requires empirical knowledge.

Based on the previous work, [Craig et al.](#) improved and identified a region of input values [Craig et al. \(1997\)](#). This input collection provided adequate belief and information of the uncertainty, compared with the traditional maximum likelihood estimation method which only tried to find one single value of an estimation. Resulting from the limitation of computation development, this method needs to compromise the time of iterations. Joint prior was chosen and only prior means, variances and covariances between all quantities of interest were used for beliefs update. Although this method provides a solution to balance computational cost and benefit of accurate calculation, the design of joint prior and the selection of inputs are complex. Lack of prior exploration also might lead to insufficient input collection.

Another research study [Busby et al. \(2007\)](#) designed emulators to evaluate the uncertainty of a reservoir. Combining fractional factorial designs and sequential designs, random processes with mean and variance of posterior were used to describe emulators, so that the number of simulations was required as small as possible. Numerical results from two-dimensional and high-dimensional input were presented. Because of a small number of simulations, this study might not be able to explore the whole input region. The results might be the local optimization, resulting from less exploration and comparison.

Based on multiscale computer experiments, an oil reservoir model evaluation was simplified using Bayes linear uncertainty analysis [Cumming and Goldstein \(2010\)](#). Similar to the Bayes linear approach, the probability in the conventional Bayes analysis is substituted by expectation in the Bayes linear uncertainty analysis. To begin the process, the evaluations of the coarse simulator were performed on different wells. Once the emulation of the coarse simulator is completed, coarse and accurate emulators were combined to update the prior. Meanwhile, the values of the best model input were addressed by the history matching process. Following the history matching, the study reduced the set of possible values of the model input and added time points for forecasting. The final stage of approximation provided a means of applying Bayesian analysis to a complex model, but it cannot identify and characterize accurate uncertainty for the individual parameter.

Nowadays the increase in computational speed and the development of more efficient algorithms make Bayesian technique accessible to solve large scale, complex problems, but it is still not well known by many engineers practitioners. There is a relatively small number of publications applying Bayesian statistical methods to the oil industry for improving the uncertainty analysis of well systems. One study by [Maraggi et al. \(2020\)](#) chose a Bayesian approach to estimate two parameters of tight oil reservoirs, which was simplified as a single dimensionless equation. Following an adaptative Markov Chain Monte Carlo (MCMC) algorithm, posterior predictive checks were adopted to examine the inferences. Uncertainty of the estimated ultimate recovery was addressed in the study. The convergence of the chains, acceptance ratio, posterior distributions, correlation between posterior parameters, the reliability plot and posterior predictive checks were presented to evaluate the approach.

Another recent study [Costa et al. \(2021\)](#) applied a similar method to an electric submersible pump system. With the inputs as prior, Bayesian inference and the MCMC methods were adopted for parameter estimation. After estimating five parameters, validation

and cross-validation were deployed using two sets of data to examine the model. Dynamic and steady-state uncertainty of the model were obtained by probability density function using uncertainty assessment JCGM (2008). Autocorrelation was used to evaluate the samples and sensitivity analysis was employed for capturing the region of convergence of the likelihood function. It would be better if the work provided the plots of multiple chains to proof the reliable parameter estimation.

Parameters in the real-life system are in certain ranges Narasimhan and Jordache (1999). Classic methods such as maximum likelihood and least squares, aim to find the best or "optimal" estimation of the parameters that describe the data, but they are difficult for assessing and quantifying the parameter uncertainty. An approximating distribution of parameters might benefit us with more insight into a system, such as quantification of the uncertainty of the estimated parameters.

In this work, we estimated the parameters of a gas lifting oil well dynamic model, considering that the main source of errors is Gaussian white noise on the measurements. The process noise is not included in this work. The purpose of this study is to obtain information from a first-principle model and data to provide estimated parameters and the related uncertainty. Bayes' Rule Lambert (2018) was used for inference and MCMC Kruschke (2014) was employed to approximate the posterior parameter distributions. The work can be the foundation of the case where oil fields contain multiple wells and it provides an important solution to advance the understanding of the uncertainty in a complex system.

The paper has been organised in the following way. The first section gives a brief overview of the recent history of Bayesian parameter estimation, especially its application in the oil industry. It will then go on to present more details of the gas lifting oil well model which is used in this work. The third section is concerned with the methodology adopted for this study. Section four briefly explain the process of the experiments. The remaining part of the paper analyses the results of the experiments and discuss the possible reasons for the issues in the experiments and the limitations of the current study.

2 Gas Lifting Oil Well Model

The gas lifting oil well model used in this work is based on a previous work Sharma et al. (2011). Including more details of an oil well model and states variation helps describe the dynamics of the gas lifting oil well. However, unnecessary details increase computational costs. Here we set out to improve the model considering the trade-off between simplicity and accuracy. Our

experiments focus on the short-term dynamics of a gas lifting oil well. We assume the reservoir pressure and compressibility factor are constant in our model, owing to the small variation in the data range and simplicity of the model. Compared with gravity, the influence of friction is small, so we ignore the pressure drop due to friction in this model. Some states are calculated and some of them are measured.

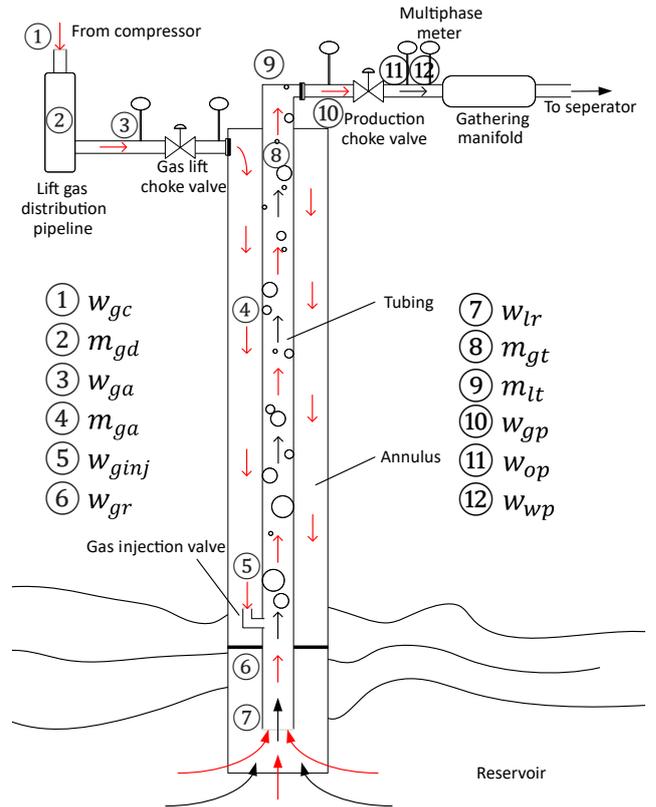


Figure 1: The components of a gas lifting oil well: the red arrows show the gas flow and black arrows present the liquid phase flow, which can contain oil and water. The mass and flow rate are depicted beside the corresponding components where these states occur in the gas lifted oil well.

A schematic diagram of a gas lifting oil well is shown in Fig. 1. In this process, valves are utilized to control the flow rates. The gas lift choke valve is located between the gas distribution manifold and annulus is used to control the amount of lift gas injected into the well. The production choke valve, which can be used to control the production, is the only valve which multi-phase fluid flows through and it connects gathering manifold and tubing. In the injection point, a gas injection valve is placed between the annulus and the tubing to connect the two parts. Since the gas lifting oil well here

is continuous in this study, the gas injection valve is fully opened and the lift gas is continuously injected from the annulus into the tubing. At the downstream of the production choke valve, a multiphase flow meter is installed to measure the flow rate of oil, gas and water. After separating the components by the separator, part of the gas is used as the lifting gas. Here we focus on the mass balance of the oil well and the gas distribution manifold.

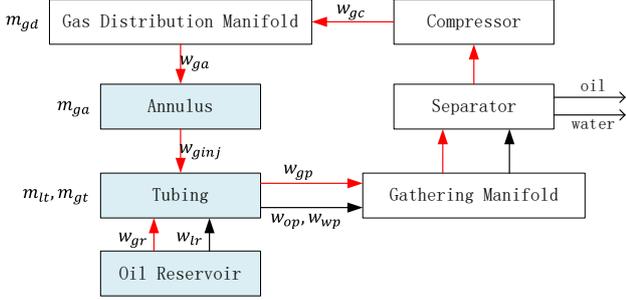


Figure 2: The flow chart of an oil field: the blue blocks are the components of an oil well. The red arrows show the gas flow and black arrows present the liquid phase flows, which can contain oil and water. The mass and flow rate are depicted beside the corresponding components where these states occur in the gas lifted oil well.

As the preceding explanation, reduction of bottom hole pressure due to gas injection typically increases oil production rate because gas injection lightens the fluid in tubing. However, injecting excessive gas increases the bottom hole pressure, which leads to a decrease in the oil production rate. This happened because excessive gas injection rate results in slippage, where the gas phase moves faster than the liquid, leaving the liquid phase behind. In this case, less amount of liquid will flow through the tubing. As the gas injection rate increases, the oil production rate rises to the maximum and decrease afterward. Model-based control contributes to the optimization of gas injection rate and then helps yield the maximum oil production rate but the control algorithm relies on the accurate model description. The first principle model is not enough to precisely describe a gas lifting oil well system. Data collected from the system is particularly useful in estimating parameters in the model.

The input of the model is the valve opening of the gas lift choke valve, u . The production choke valve is assumed fully opened during the whole experiment. As one of the parameters to be estimated, the water cut, WC, is the volume of water produced compared to the volume of total liquids produced from an oil

well. Other parameters in the model to be estimated are gas to oil ratio, GOR, and the productivity index, PI [$\frac{kg/hr}{bar}$]. GOR is the mass ratio of produced gas to produced liquid (oil and water). PI is a mathematical means of expressing the ability of a reservoir to deliver fluids to the wellbore. All these parameters are dimensionless.

There are seven measurements: w_{ga} , w_{gp} , w_{op} , w_{wp} , P_{wf} , P_{wh} , P_a in this model, which are output 1 to 7. Flow meters measure the flow rate of the gas in different places, including w_{ga} and w_{gp} . Also, the flow rate of the liquid, including w_{op} and w_{wp} . Flow measurements are impacted by bubbles and are not reliable as the data such as temperature and pressure. Pressure transmitters are used to measure the bottom hole pressure or well flow pressure (P_{wf}), in the tubing upstream the production choke valve (P_{wh}), and in the annulus downstream the lift gas choke valve (P_a). We assume the pressure of the gas in the gas distribution pipeline (P_c) is a constant of 200 [bar]. In the production process, temperature sensors detect the temperature in the gas distribution pipeline, the annulus and the tubing. Because of the small difference between these temperatures, we assumed the temperature is constant everywhere and all these temperatures are assumed to be equal to T [K]. The corresponding notation table is shown in Appendix 1.

Considering the principle of the gas lifted oil well, its model is designed based on the mass balance of three states: the mass of gas in the annulus m_{ga} , in the tubing above injection point m_{gt} and the mass of liquid in the tubing above injection m_{lt} , which is mainly present as Eq. (1):

$$\begin{aligned} \dot{m}_{ga} &= w_{ga} - w_{ginj}, \\ \dot{m}_{gt} &= w_{ginj} + w_{gr} - w_{gp}, \\ \dot{m}_{lt} &= w_{lr} - w_{lp}. \end{aligned} \quad (1)$$

, where w_{ga} is the flow rate of the gas through the gas lift choke valve which is injected into the annulus. The flow rates of the lift gas from the annulus and reservoir to the tubing are w_{ginj} and w_{gr} respectively. The flow rate of produced gas through the production choke valve is presented as w_{gp} . w_{lr} and w_{lp} are the liquid phase flow from the reservoir into the well and through the production choke valve, respectively. These flows are marked in the corresponding position in the flow chart Fig. 2. Auxiliary equations are shown in Appendix 2.

3 Methodology

Fig. 3 describes the process where parameters are estimated using Bayes' rule and MCMC algorithm. To be-

gin this process, Bayesian inference needs to be carried out as the first step. Once the proposed parameters are generated according to the distribution of prior, the posterior can be calculated based on the inference. After comparing the current posterior with the previous one, MCMC decides to accept or reject the parameters. On the long run, the distribution of the accepted samples converges to the true posterior distribution. The final stage of the study is checking the posterior predictive distribution to predict the real output.

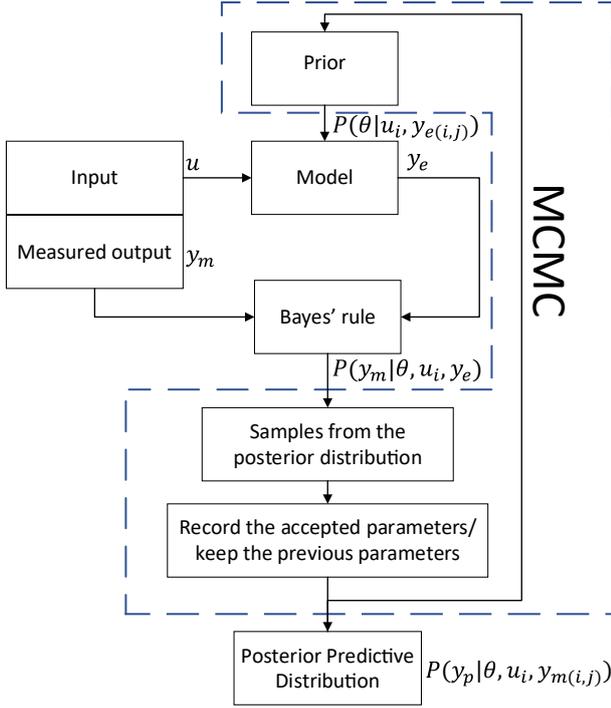


Figure 3: Outline of the estimation process: according to Bayes' rule, posterior is calculated using the proposed parameters as prior, the output of model and input. The blue box contains the MCMC algorithm, which consists of two steps: proposal and acceptance/rejections of the samples from the parameter space.

3.1 Bayes' Rule

In this gas lifting oil well model, there is one input and seven measurements. Here are the notations that are used in Bayes' Rule.

θ : the model parameters, which is presented as a 1×3 array: [PI, GRO, WC].

u_i : there are n inputs in the system. We used i to count them. The input vector can be presented as a $1 \times n$ array.

$y_{m(i,j)}$: measurements of the outputs. In this system, there are 7 outputs: w_{ga} , w_{gp} , w_{op} , w_{wp} , P_{wf} , P_{wh} , P_a , which are denoted by $j = 1, 2, 3 \dots 7$. Each of these outputs contains n samples. Therefore, the measurements can be presented as a $7 \times n$ matrix.

$y_{e(i,j)}$: the estimated outputs, namely the outputs of the model given the input and the parameters. Similarly to the measured output, each estimated output contains the number of n samples. The estimated output is a $7 \times n$ matrix. The model that we used for the estimation process does not contain noise.

3.1.1 Prior

The GOR and WC parameters represent ratios, so they are between 0 and 1. The order of magnitude of the parameter PI is also known. Therefore, the following prior distributions for these parameters were chosen:

$$P(\theta | u_i, y_{e(i,j)}) = \begin{cases} B & PI \in [10^4, 10^5], GRO \in [0, 1], WC \in [0, 1] \\ 0 & \text{otherwise} \end{cases} \quad (2)$$

3.1.2 Likelihood

We assume the measurement contains white noise which is normally distributed with mean 0 and it can be present as Eq. (3).

$$\begin{aligned} y_{m(i,j)} &= y_{e(i,j)} + \varepsilon_{i,j}, \\ \varepsilon_{i,j} &\sim N(0, \sigma_{i,j}), \\ y_{m(i,j)} &\sim N(y_{e(i,j)}, \sigma_{i,j}). \end{aligned} \quad (3)$$

The errors for each sample $y_{m(i,-)}$ are independent and the standard variation of the errors for the samples of a given output is considered to be the same, namely

$$\sigma_{i,j} = \sigma_j \quad i = 1, 2, 3, \dots n \quad (4)$$

In the case where the standard deviation is known (for example from the specifications of the sensors), the likelihood is the product of the probability of the samples of each output:

$$P(y_m | \theta, u_i, y_e) = \prod_{j=1}^7 \left(\frac{1}{\sigma_j \sqrt{2\pi}} \right)^n \exp\left(-\frac{1}{2} \sum_{i=1}^n \left(\frac{y_{m(i,j)} - y_{e(i,j)}}{\sigma_j} \right)^2 \right) \quad (5)$$

When the standard deviation σ_j is unknown, the likelihood should be averaged over all possible values of the model parameters. This process is known as marginalization and the marginal likelihood is given

by:

$$\begin{aligned}
 P(y_{m(i,j)} | \theta, u_i, y_{e_j}) &= \int P(\sigma_j, y_{m(i,j)} | \theta, u_i, y_{e_j}) d\sigma_j \\
 &= \int P(y_{m(i,j)} | \sigma_j, \theta, u_i, y_{e_j}) P(\sigma_j | \theta, u_i, y_{e_j}) d\sigma_j \\
 &= \int P(y_{m(i,j)} | \sigma_j, y_{e_j}(\theta, u_i)) P(\sigma_j | y_{e_j}(\theta, u_i)) d\sigma_j
 \end{aligned} \quad (6)$$

The noise is not impacted by the model, so $P(\sigma_j | y_{e_j}(\theta, u_i)) = P(\sigma_j) \propto \frac{1}{\sigma_j}$, where $\frac{1}{\sigma_j}$ is the Jeffreys prior for the standard deviation when $\sigma > 0$.

$$\begin{aligned}
 P(y_{m(i,j)} | \theta, u_i, y_{e_j}) &= \\
 &\int N(y_{m(i,j)} - y_{e_j}(u_i, \theta), \sigma_j) P(\sigma_j) d\sigma_j \quad (7)
 \end{aligned}$$

, namely the likelihood for the j output is calculated as:

$$\begin{aligned}
 P(y_{m_j} | \theta, u_i, y_{e_j}) &\propto \\
 &\int \frac{1}{\sigma_j} \exp\left(-\frac{1}{2\sigma_j^2} \sum_{i=1}^n (y_{m(i,j)} - y_{e_j}(u_i, \theta))^2\right) d\sigma_j \quad (8)
 \end{aligned}$$

3.1.3 Posterior

According to the Bayes' Rule, the posterior is calculated by prior and likelihood:

$$P(\theta | y_m, u_i, y_e) \propto P(y_m | \theta, u_i, y_e) P(\theta | u_i, y_e) \quad (9)$$

The logarithms of likelihood is used here to avoid underflow during the computation. It is difficult to graph the likelihood because it is often a tiny number. With the logarithms of likelihood, the large products become sums and it is easier to plot.

When the standard deviation is known, the log of the posterior is

$$\begin{aligned}
 &\ln P(y_m | \theta, u_i, y_e) \\
 &= n \sum_{j=1}^7 \ln\left(\frac{1}{\sigma_j \sqrt{2\pi}}\right) - \frac{1}{2} \sum_{j=1}^7 \sum_{i=1}^n \left(\frac{y_{m(i,j)} - y_{e_j}(u_i, \theta)}{\sigma_j}\right)^2
 \end{aligned} \quad (10)$$

In terms of the cases where the standard deviation is unknown, since the likelihood and prior are followed to Gaussian distribution and uniform distribution individually, the posterior is obeyed to student-t distribution [Sivia and Skilling \(2006\)](#). Considering the various inputs and multiply outputs, the distribution of the posterior is

$$\begin{aligned}
 P(\theta | y_m, u_i, y_e) &\propto \\
 &\prod_{j=1}^7 \prod_{i=1}^n \left(1 + \left(\frac{y_{m(i,j)} - y_{e_j}(u_i, \theta)}{S_j}\right)^2\right)^{-\frac{n}{2}} \quad (11)
 \end{aligned}$$

, where S_j is the standard deviation of corresponding measured output. With the measurement during steady state, the standard deviation can be calculated as Eq.(12).

$$S_j^2 = \frac{1}{n} \sum_{i=1}^n (y_{m(i,j)} - \bar{y}_{m(i,j)})^2 \quad (12)$$

The log of posterior can be calculated as:

$$\begin{aligned}
 \ln P(\theta | y_m, u_i, y_e) &\propto \\
 &-\sum_{j=1}^7 \left(\frac{n}{2} \sum_{i=1}^n \ln\left(1 + \left(\frac{y_{m(i,j)} - y_{e_j}(u_i, \theta)}{S_j}\right)^2\right)\right) \quad (13)
 \end{aligned}$$

3.2 Markov Chain Monte Carlo

The distribution of the posterior provides vital information for helping tackle the optimal parameters. In this study, the posterior is a three-dimensional space which is related to the joint distribution of combinations of PI, GRO, WC. Considering the complexity of the posterior, the distribution can not be adequately represented by a simple distribution, such as Binomial Distribution, Normal distribution, Cauchy distribution and so on.

Algorithm 1: MetropolisHastings algorithm

```

Initialisation of  $\theta_0; N_i; q$ 
Calculate
 $\ln P(y_m^t | \theta^t, u_i, y_{e_j}) = \ln P(y_0 | \theta_0, u_i, y_{e_j})$ 
for iteration = 1, 2, ...,  $N_i$  do
    Generate samples
     $\theta_n^* \sim \text{Beta}(q * \theta_n^t, q * (1 - \tilde{\theta}_n^t));$ 
    finn Generate a random constant
     $a \sim U(0, 1);$ 
    Calculate  $\ln P(y_m^* | \theta_n^*, u_i, y_{e_j});$ 
    if  $\ln a <$ 
     $\ln \frac{P(y_m^* | \theta_n^*, u_i, y_{e_j})}{P(y_m^t | \theta_n^t, u_i, y_{e_j})} + \sum_{j=1}^3 \ln \frac{P(\theta_{n,j}^t | \theta_{n,j}^*)}{P(\theta_{n,j}^* | \theta_{n,j}^t)}$ 
    then
        if  $\theta_n^* \neq 0$  then
             $\theta_n^t = \theta_n^*;$ 
        end
        update  $\ln P(y_m^t | \theta_n^t, u_i, y_{e_j});$ 
    end
    Save  $\theta_n^t;$ 
end
    
```

Random walk Metropolis-Hastings algorithm [Metropolis et al. \(1953\)](#); [Hastings \(1970\)](#) is a well-known MCMC method for indirect sampling from a distribution. The MCMC algorithm used in this work is shown in Algorithm 1. Considering the prior

is in a certain range, the samples were randomly drew from these ranges as initialization. During the iteration, we utilized proposed beta distributions to generate a sequence of random samples. Since the three parameters are independent, three beta distributions, $\text{Beta}(q * \tilde{\theta}_n^t, q * (1 - \tilde{\theta}_n^t))$, were used and tuned individually. q is tuned to adjust the distribution. The benefit of the modification is that all samples satisfy the prior constraints by adjusting from the beta distribution range $\tilde{\theta}_n^t \in [0, 1]$ to prior θ_n^t range. Compared with normal distribution and gamma distribution, beta distribution generates more workable samples.

During random walk, Beta distribution is adopted here as the proposal distribution, which is recentered at θ_n^{t-1} during each step. Some of the Beta distributions used here are skew. The mean value and the mode of the Beta distribution do not coincide. Therefore, the random walk does not satisfy the symmetry requirement. Owing to the asymmetry of the proposal distribution, the correction term $\sum_{j=1}^3 \ln \frac{P(\theta_{n,j}^t | \theta_{n,j}^*)}{P(\theta_{n,j}^* | \theta_{n,j}^t)}$ was needed to add [Holder et al. \(2005\)](#).

In the algorithm, the criteria of accepting the proposed parameter is that the posterior of the proposed parameter is larger than the posterior of the previous parameter. To avoid being stuck in a local optimum, we drew a random number between 0 and 1 and compared with the division between previous posterior and the current posterior. In the case where the standard deviation is known, the calculation can be simplified. Only the part: $\sum_{j=1}^7 \sum_{i=1}^n (\frac{y_{m(i,j)} - y_{e(i,j)}(u_i, \theta)}{\sigma_j})^2$ is used for the condition, as the constant part $n \sum_{j=1}^7 \ln(\frac{1}{\sigma_j \sqrt{2\pi}})$ are canceled out after using quotient of natural logarithm.

After the number of N_i iterations, we get a list of accepted samples which form a chain. We can check how the chain explore the parameter space by plotting these samples. The distribution of these samples can be used to approximate the distribution of the marginal posterior distribution of the parameters.

3.3 Posterior Predictive Distribution

Once the parameter estimation is completed, we can use the oil well model with the estimated parameters to estimate other outputs of the system. After running the MCMC algorithm, we get a group of samples from the accept list, which approximately shows the distribution of parameters $\theta \sim P(\theta | y_m, y_e, u_i)$. Following cut out of the burn-in period of the chain, outputs of the model y_p with estimated parameters and inputs are simulated, namely we get $P(y_p | \theta, u_i, y_e, y_m)$. Though the iterations of the MCMC algorithm are limited, the

distribution of these outputs in each sampling time can generally present where the output locates. Then we can calculate the posterior predictive distribution as:

$$\begin{aligned} P(y_p | u_i, y_e, y_m) &= \int P(y_p, \theta | u_i, y_e, y_m) d\theta \\ &= \int P(y_p | \theta, u_i, y_e, y_m) P(\theta | y_m, y_e, u_i) d\theta \\ &= \int P(y_p | y_e(\theta, u_i), y_m) P(\theta | y_m, u_i, y_e) d\theta \quad (14) \end{aligned}$$

We assumed the noise in the measurement follows a normal distribution. Therefore, posterior predictive distribution follows the student-t distribution with $(n-1)$ degree of freedom.

$$P(y_{pj} | u_i, y_{m(i,j)}) \propto (1 + (\frac{y_{pj} - y_{ej}(u, \theta)}{S_j})^2)^{-\frac{n}{2}} \quad (15)$$

By calculating a large number of y_{pj} , we get the posterior predictive distribution. The accuracy and efficiency of the parameter estimation process can be evaluated by comparing posterior predictive distribution and measurements.

4 Experiments

In this project, we used MATLAB as our experimental platform. The work began by generating data from a simulator with preset parameters, and then Gaussian noise was added to the outputs. The preset parameters are $WC = 0.18$, $PI = 2.4 \times 10^4$, $GOR = 0.15$. The preset parameters were only unveiled after the experiments for evaluating the estimation. The information of noise is not accessible during the experiments. It will then go on to estimate the parameters and the uncertainty of the model.

The first step in the MCMC algorithm was to generate samples. Here we use beta distribution to propose samples from the parameter space, as shown in [Section 3.2](#). The value of GOR and WC can be used for shape parameters directly, since they are between 0 and 1. PI is between 10^4 to 10^5 , so PI needs to be transferred to $[0, 1]$ before being used as the shape parameter. To find proper q in beta distribution. A series of experiments were run with $q = 1, 10, 100, 1000, 10000, 300$ was selected after comparing the results, so that proposed samples can reach relatively far area and can also converge to an optimised value.

Because the proposed parameters are impacted by current samples and the shape parameter in beta distribution should be positive, a zero value would lead to a nonsense beta distribution in the next iteration. Therefore a conditional statement is developed, which helps keep the shape parameter as the current value.

With proposed parameters and inputs, the estimated outputs are simulated by the model. Then, the corresponding posterior can be calculated. Once the calculation is completed, the MCMC algorithm moves to the decision stage by comparing the posterior corresponding to proposed parameters and current parameters.

For each MCMC iteration, differential equations in the oil well model need to be solved. Therefore, it takes a long time to run a couple of chains with a large number of iterations, for example, a chain with 1000 iterations for the outputs which contain 3600 samples takes more than 20 minutes with the processor Intel(R) Core(TM) i7-10750H CPU @ 2.60GHz 2.59 GHz. To increase the computation efficiency, the number of measurement data and iterations of the chains need to be chosen exquisitely.

To balance the quality and quantity of the samples, we choose more samples when the inputs change and fewer samples during the steady-state to gain more information from the data. In this work, we run 10 chains and each chain contains 2000 iterations to identify the parameters using data with 20 samples. The number of data samples is not large and the production of normal distribution and t-distribution is similar. Therefore, the results for the case with known variation and unknown variation are approximately similar. The MCMC algorithm can estimate parameters in both cases and chains converge at a similar speed. We will present one of the results in the next section. All codes and resources in this study have been uploaded to Github¹.

5 Result

In this section, we will present and analyse the results of experiments in various ways. Reject rate and autocorrelation offer an effective way of checking the efficiency of the chains. To gain insights into every parameter in the estimation process, we drew figures of individual parameter distribution and Markov Chains. The interaction between every two parameters is shown in scatter plots and contour plots. Posterior predictive distribution plots provide the results of the accuracy of the estimation. The impacts of parameters are compared by sensitivity analysis.

5.1 Autocorrelation

Autocorrelation offers further in-depth information on the efficiency of the chains, which benefits us to select the step length of MCMC. Furthermore, the estimations of parameters are impacted by the correlation

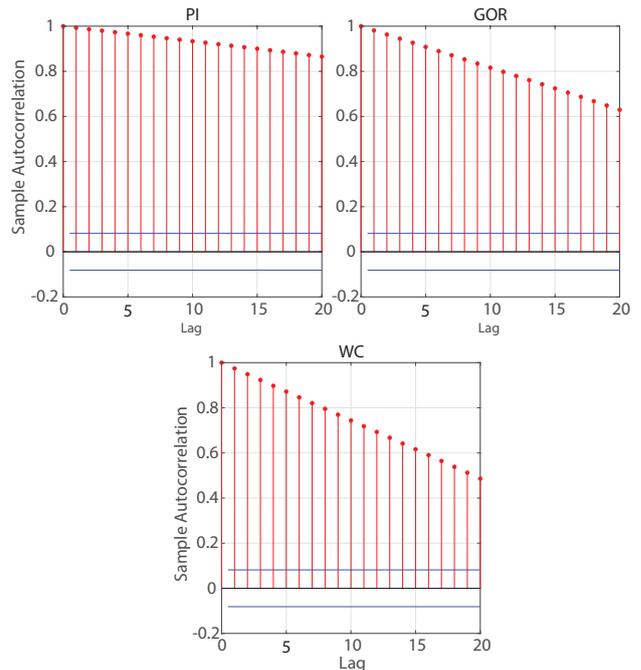


Figure 4: Autocorrelation of chains for three parameters without burn-in period. This plot shows the autocorrelation across a spectrum of candidate lags. The autocorrelation are plotted as bars at every integer lag. The blue lines show the upper and lower confidence bounds.

between the samples. We expect independent sampling in our experiments or at least sampling with low autocorrelation. However, according to the Metropolis-Hastings algorithm, every proposed value is based on the previous value. It is impossible to make every step independent from others, while proper step length contributes to an effective sample size.

The autocorrelation plot displays the dependence structure of the chain. Fig. 4 illustrates the chain autocorrelation after cutting off the first 1000 samples where the chain has not converged. The correlogram shows the correlations decreases as the lag increases.

5.2 Parameter Distribution and Chains

In order to test the influence of the random initial value of the chain and check the explored range of the estimated parameter distribution, we present the plots of the chains and the distribution for each parameter, shown in Fig.5.

The first row provides information on the converging speed of the chains. The burn-in period of these chains is less than 1000 for both *GOR* and *PI*, while the chain of *WC* took a longer time to converge to a certain range. The end of the chains overlap each other and

¹<https://github.com/Zhe-Jessica/Parameter-estimation-MCMC>

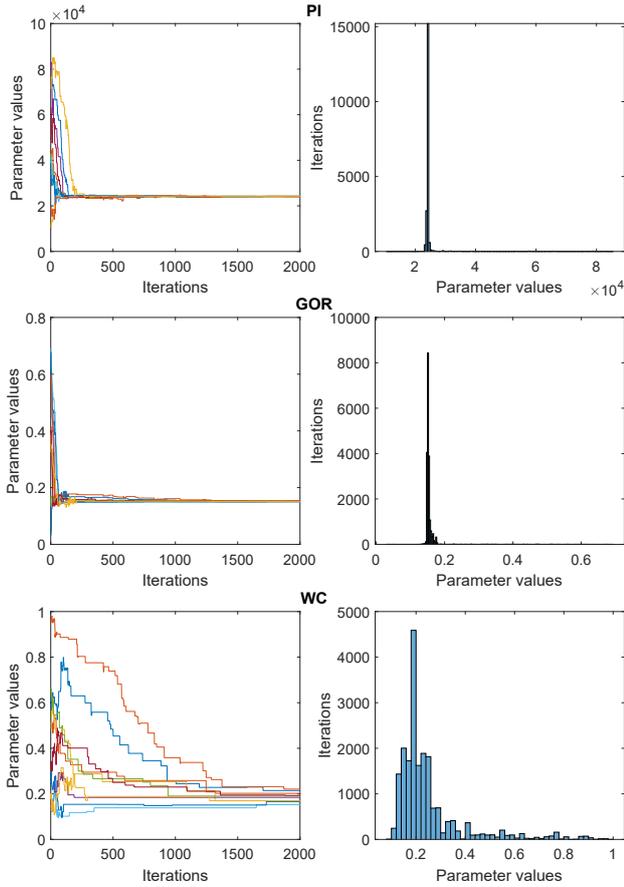


Figure 5: The first column of plots show trace plots with various initial values. The vertical axis is the parameter value, and the horizontal axis is the step in the chain, namely the iterations. The second column in the figures illustrate the distribution of each individual parameter, using samples of all chains.

fairly smoothly meander around the optimal estimated value.

The second column illustrates histograms of the parameter from the accepted lists of all chains. According to these plots, all chains converge to around the preset values after exploring the whole prior range. All the chains for the same parameters converge to a similar region and mix well, which indicates the convergence is achieved.

5.3 Reject Rate

When running MCMC experiments, we recorded the number of reject decisions during iterations. Following divided by the number of iterations, the reject rate is calculated. The reject rate shows how many proposed samples were rejected when the chain moved. The re-

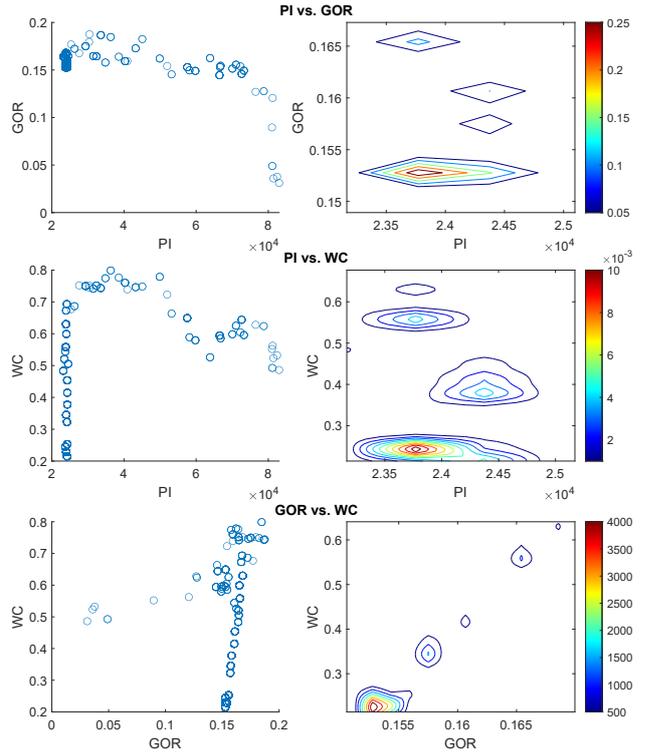


Figure 6: Estimation of parameters with 2D plots in scatter plots and contour plots. The contour plots were drawn in the range with minimum and maximum values. The converge parts were zoomed in to present the contour clearly.

ject rate in our experiments is around 94.8%.

5.4 2D Parameters Distributions

Three scatter plots are shown in Fig.6 to present the distribution of each two parameters. These plots confirm the conclusion in Fig.5. The contour plots in the second column demonstrate more details in the converging area.

5.5 Posterior Predictive Distributions

According to Section 3.2, we calculated y_p . There are 1000 samples of y_p for each sampling time. The quantiles of these samples were chosen and presented with the measurement and the output of the model with preset parameters in Fig.7.

Most posterior predictive distributions are around the real outputs, especially the first dynamic stage. For output 1 and 7, the posterior predictive distributions tackled the first dynamic change and parted from the real output after the second input change. Compared

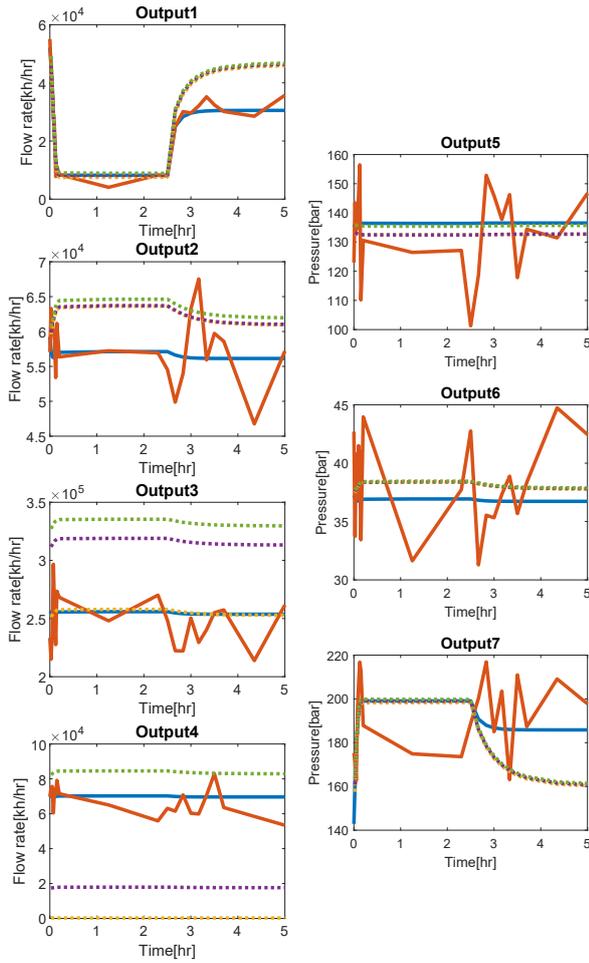


Figure 7: Posterior predictive distribution: In each plot, dash lines are quantiles of the posterior predictive and the purple dash line is the median. The solid orange lines are the data that contains noise. To compare with the real outputs, we also drew the outputs of the system with preset parameters. The real outputs do not contain noise, which are presented as blue solid lines.

with other posterior predictive distributions, the one for output 4 is wider and the measurement stays within the quantile range. For output 3, the posterior predictive distributions are slightly above the measurement.

5.6 Sensitivity Analysis

To identify the impact of each parameter on the outputs, we increased and decreased the parameters within ranges and checked the outputs. As shown in Fig. 8, WC does not impact on output 1,2,5,6,7 a lot. If PI changes tremendously, it will have an effect on all outputs, apart from output 4 compared with two other

parameters. GOR also contributes to the changes of outputs, especially output 2 and output 6, compared with the other parameters.

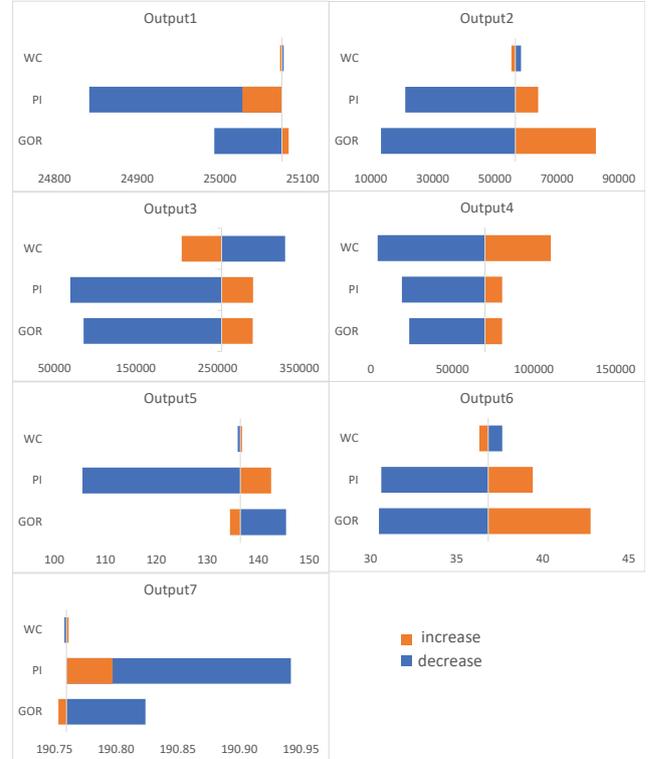


Figure 8: Sensitivity analysis of three parameters for seven outputs using tornado diagram. The baseline values are set with the preset values in the above experiments.

6 Discussion

- Proper step length plays a vital role in the Metropolis-Hastings algorithm [Kruschke \(2014\)](#). We used beta distribution in MCMC to propose new parameters. The parameters in beta distribution impact the step length of MCMC. For example, the first shape parameter in the beta distribution is related to the previous accept value. However, the mode of the beta distribution is not exactly the same as the first shape parameter. When q is large, the shape of the beta distribution is sharp and closer to the previous accept value. However, the large shape parameter leads to small walking steps, which causes a longer time for the chain to explore a wide area. Therefore, tuning the first shape parameter is a trade-off between the relevance of the previous step and the walking speed.

- According to the result of autocorrelation, the effective sample size still can be improved by tuning the step length of MCMC. Furthermore, the model here is a gas lifting oil well model only with boundaries of parameters. For a specific system, the empirical engineering knowledge might make a significant contribution to the prior, which further benefit the efficiency of the chains.

7 Conclusion

This study set out to estimate three parameters and their uncertainty of a gas lifting oil well model. With Bayesian interference and Metropolis-Hastings algorithm, the proposed solution was able to identify the real values of the parameters and present their distributions. Based on the estimation, we demonstrated posterior predictive distributions and sensitivity analysis, which has gone some way towards enhancing our understanding of the true value of the outputs of the system and the impact of the parameters. The model with estimated parameters can benefit optimization and advanced control in oil production. The work here can be considered as the first step of the improvement of a real-life gas lifting oil well model. A further study could apply this solution to an oil well system with real data.

Acknowledgments

This work was supported by The Research Council of Norway and Equinor ASA through Research Council project 308817 - Digital wells for optimal production and drainage (DigiWell).

References

- Brown, K. E. Technology of artificial lift methods. volume 1. inflow performance, multiphase flow in pipes, the flowing well. 1977.
- Busby, D., Farmer, C. L., and Iske, A. Uncertainty evaluation in reservoir forecasting by bayes linear methodology. In *Algorithms for Approximation*, pages 187–196. Springer, 2007. doi:[10.1007/978-3-540-46551-5_14](https://doi.org/10.1007/978-3-540-46551-5_14).
- Costa, E., de Abreu, O., Silva, T. d. O., Ribeiro, M., and Schnitman, L. A bayesian approach to the dynamic modeling of esp-lifted oil well systems: An experimental validation on an esp prototype. *Journal of Petroleum Science and Engineering*, 2021. 205:108880. doi:[10.1016/j.petrol.2021.108880](https://doi.org/10.1016/j.petrol.2021.108880).
- Craig, P., Goldstein, M., Seheult, A., and Smith, J. Bayes linear strategies for matching hydrocarbon reservoir history. *Bayesian statistics*, 1996. 5:69–95.
- Craig, P. S., Goldstein, M., Seheult, A. H., and Smith, J. A. Pressure matching for hydrocarbon reservoirs: a case study in the use of bayes linear strategies for large computer experiments. In *Case studies in Bayesian statistics*, pages 37–93. Springer, 1997. doi:[10.1007/978-1-4612-2290-3_2](https://doi.org/10.1007/978-1-4612-2290-3_2).
- Cumming, J. A. and Goldstein, M. Bayes linear uncertainty analysis for oil reservoirs based on multiscale computer experiments. *The Oxford handbook of applied Bayesian analysis*, 2010. pages 241–270.
- Gyamfi, K. S., Brusey, J., Hunt, A., and Gaura, E. Linear dimensionality reduction for classification via a sequential bayes error minimisation with an application to flow meter diagnostics. *Expert Systems with Applications*, 2018. 91:252–262. doi:[10.1016/j.eswa.2017.09.010](https://doi.org/10.1016/j.eswa.2017.09.010).
- Hastings, W. K. Monte carlo sampling methods using markov chains and their applications. 1970. doi:[10.1093/biomet/57.1.97](https://doi.org/10.1093/biomet/57.1.97).
- Holder, M. T., Lewis, P. O., Swofford, D. L., and Larget, B. Hastings ratio of the local proposal used in bayesian phylogenetics. *Systematic biology*, 2005. 54(6):961–965. doi:[10.1080/10635150500354670](https://doi.org/10.1080/10635150500354670).
- Hu, C., Zhang, C., Zhang, Z., and Xie, S. Comparative study on defects and faults detection of main transformer based on logistic regression and naive bayes algorithm. In *Journal of Physics: Conference Series*, volume 1732. IOP Publishing, page 012075, 2021. doi:[10.1088/1742-6596/1732/1/012075](https://doi.org/10.1088/1742-6596/1732/1/012075).
- JCGM. Evaluation of measurement dataguide to the expression of uncertainty in measurement. *JCGM*, 2008. 100(2008):1–116.
- Jeffreys, H. Some tests of significance, treated by the theory of probability. In *Mathematical Proceedings of the Cambridge Philosophical Society*, volume 31. Cambridge University Press, pages 203–222, 1935. doi:[10.1017/S030500410001330X](https://doi.org/10.1017/S030500410001330X).
- Khaliullin, F. K., Matyashin, A. V., Akhmetzyanov, R. R., Medvedev, V. M., and Lushnov, M. A. Prospects for using the bayes algorithm for assessing the technical condition of internal combustion engines. In *IOP Conference Series: Materials Science and Engineering*, volume 635. IOP Publishing, page 012016, 2019. doi:[10.1088/1757-899X/635/1/012016](https://doi.org/10.1088/1757-899X/635/1/012016).

- Kruschke, J. Doing bayesian data analysis: A tutorial with r, jags, and stan. 2014.
- Lambert, B. *A students guide to Bayesian statistics*. Sage, 2018.
- Lindley, D. V. The use of prior probability distributions in statistical inference and decision. In *Proc. 4th Berkeley Symp. on Math. Stat. and Prob.* pages 453–468, 1961.
- Maraggi, L. M. R., Lake, L. W., and Walsh, M. P. A bayesian framework for addressing the uncertainty in production forecasts of tight oil reservoirs using a physics-based two-phase flow model. In *SPE/AAPG/SEG Latin America Unconventional Resources Technology Conference*. OnePetro, 2020. doi:[10.15530/urtec-2020-10480](https://doi.org/10.15530/urtec-2020-10480).
- Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., and Teller, E. Equation of state calculations by fast computing machines. *The journal of chemical physics*, 1953. 21(6):1087–1092. doi:[10.1063/1.1699114](https://doi.org/10.1063/1.1699114).
- Narasimhan, S. and Jordache, C. *Data reconciliation and gross error detection: An intelligent use of process data*. Elsevier, 1999.
- Sharma, R., Fjalestad, K., and Glemmestad, B. Modeling and control of gas lifted oil field with five oil wells. In *52nd International Conference of Scandinavian Simulation Society, SIMS*. pages 29–30, 2011.
- Sivia, D. and Skilling, J. *Data analysis: a Bayesian tutorial*. OUP Oxford, 2006.

Appendix 1

Name	Notation	Unit	Classification
valve opening of the gas lift choke valve	u	-	input
flow rate of oil production	w_{op}	kg/hr	output
gas flow rate into annulus	w_{ga}	kg/hr	output
gas flow rate through the production choke valve	w_{gp}	kg/hr	output
water flow rate through production choke valve	w_{wp}	kg/hr	output
pressure of the gas in the annulus downstream the lift gas choke valve	P_a	bar	output
the bottom hole pressure or well flow pressure	P_{wf}	bar	output
pressure in the tubing upstream the production choke valve	P_{wh}	bar	output
pressure of gas in the gas distribution pipeline upstream the lift gas	P_c	bar	constant
temperature in the gas lifting oil well system	T	K	constant
water cut	WC	-	parameter
gas to oil ratio	GOR	-	parameter
productivity index	PI	$\frac{kg/hr}{bar}$	parameter

Appendix 2

Valve characteristic as a function of its opening for the lift gas choke valve is shown in Eq (1). u is the input of the model, the valve opening and C_v is the valve characteristic as a function of valve opening. Both notations are dimensionless. $u \in [0, 100]$. 0 and 100 are for entirely closed and opened respectively.

$$C_v(u) = \begin{cases} 0 & u < 5 \\ 0.111u - 0.556 & 5 < u < 50 \\ 0.5u - 20 & 50 < u \end{cases} \quad (1)$$

Density of gas in the gas distribution pipeline:

$$\rho_{gp} = \frac{m_{gd}}{A_p L_{p,tl}} \quad (2)$$

Density of gas in the annulus:

$$\rho_{ga} = \frac{m_{ga}}{A_a L_{a,tl}} \quad (3)$$

We assume the mixture in the tubing is homogeneous, and the density is constant everywhere. m_{gt} and m_{lt} are the mass of gas and liquid respectively above the injection point. The density of the mixture of liquid and gas in the tubing above the injection point:

$$\rho_m = \frac{m_{gt} + m_{lt}}{A_t L_{t,tl}} \quad (4)$$

Pressure of gas in the annulus downstream the lift gas choke valve is presented in Eq. (6), where Z_{P_a} is a compressibility factor. We regard $Z_{P_a} = 0.6816$ as a constant at 170 bar.

$$P_a = \frac{Z_{P_a} m_{ga} RT}{M A_a L_{a,tl}} \quad (5)$$

Pressure of gas in the annulus upstream the gas injection valve:

$$P_{ainj} = P_a + \rho_{ga} g L_{a,vl} \quad (6)$$

Gas expansion factor in the gas lift choke valve:

$$Y_1 = 1 - \alpha_Y \left(\frac{P_c - P_a}{\max(P_c, P_c^{min})} \right) \quad (7)$$

Density of the liquid based on the water cut:

$$\rho_l = \rho_w WC + \rho_o (1 - WC) \quad (8)$$

The volume of gas present in the tubing above the gas injection point:

$$V_G = A_t L_{t,tl} - \frac{m_{lt}}{\rho_l} \quad (9)$$

Pressure in the tubing downstream the gas injection valve is depicted in Eq. (10), where Z_{P_G} is a compressibility factor. We regard $Z_{P_G} = 0.6741$ as a constant at 150 bar.

$$P_{tinj} = \frac{Z_{P_G} m_{gt} RT}{M V_G} + \frac{1}{2} \rho_m g L_{t,vl} \quad (10)$$

Pressure in the tubing upstream the production choke valve:

$$P_{wh} = \frac{Z_{P_G} m_{gt} RT}{M V_G} - \frac{1}{2} \rho_m g L_{t,vl} \quad (11)$$

The bottom hole pressure or well flow pressure:

$$P_{wfl} = P_{tinj} + \rho_r g L_{r,vl} \quad (12)$$

Gas expansion factor in the gas injection valve:

$$Y_2 = 1 - \alpha_Y \left(\frac{P_{ainj} - P_{tinj}}{\max(P_{ainj} - P_{ainj}^{min})} \right) \quad (13)$$

Gas expansion factor in the production choke valve:

$$Y_3 = 1 - \alpha_Y \left(\frac{P_{wh} - P_s}{\max(P_{wh} - P_{wh}^{min})} \right) \quad (14)$$

The mass flow rate of the gas through the gas lift choke valve is shown in Eq. (15), where N_6 is valve constant and used to calculate the flow through the valve based on the pressure gradient:

$$w_{ga} = N_6 C_{v1} (u_1) Y_1 \sqrt{\rho_{gp} \max(P_c - P_a, 0)} \quad (15)$$

The mass flow rate of the gas injected into the tubing from the annulus:

$$w_{ginj} = k Y_2 \sqrt{\rho_{ga} \max(P_{ainj} - P_{tinj}, 0)} \quad (16)$$

The mass flow rate of the liquid from the reservoir:

$$w_{lr} = PI \max(P_r - P_{wf}, 0) \quad (17)$$

The mass flow rate of gas from the reservoir:

$$w_{gr} = GOR w_{lr} \quad (18)$$

The mass flow rate of the mixture of gas and liquid from the production choke valve:

$$w_{gop} = \bar{N}_6 C_v (u_2) Y_3 \sqrt{\rho_m \max(P_{wh} - P_s, 0)} \quad (19)$$

Mass flow rate of gas through the production choke valve:

$$w_{gp} = \frac{m_{gt}}{m_{gt} + m_{lt}} w_{gop} \quad (20)$$

Mass flow rate of liquid through the production choke valve:

$$w_{lp} = \frac{m_{lt}}{m_{gt} + m_{lt}} w_{gop} \quad (21)$$

Oil compartment mass flow rate from liquid product considering water cut:

$$w_{op} = \frac{\rho_o}{\rho_l} (1 - WC) w_{lp} \quad (22)$$

Water compartment mass flow rate from liquid product considering water cut:

$$w_{wp} = w_{lp} - w_{op} \quad (23)$$