Parameter and State Estimation for an Oil Production Model using Julia

Zhe Ban, Carlos Pfeiffer, Bernt Lie

Department of Electrical Engineering, IT and Cybernetics, University of South-Eastern Norway, 3918 Porsgrunn, Norway
{zhe.ban, carlos.pfeiffer, Bernt.Lie}@usn.no

Abstract

Dynamic models of industrial processes play an instrumental role in the operation of such processes from smart sensors, data reconciliation, to advanced control. For good performance, a precise model is normally required. The issue of improving models has received considerable critical attention. In this work, we consider the estimation of model parameters and initial states of a gas lifting oil well model, followed by filtering of its states. By utilizing information from both first-principle model and data, the results are presented to show the estimated values and their uncertainties. Julia is the main programming language used in this study. This research study provided an opportunity to advance the understanding of the optimization and estimation for the oil well operation.

1. Introduction

As a common statistic method, Bayesian inference has been used in many scientific fields and industries to provide estimates for unknown quantities considering uncertainty. The application of the Bayesian approach can be classified as batch and sequential according to the property of the data set. In the case of the estimation related to batch datasets, all data are collected at once before processing. Markov chain Monte Carlo (MCMC) algorithms, as stochastic simulation methods, are commonly used for the batch case to solve model uncertainty problems. In the sequential application, the data arrive sequentially, so the estimation or the data evolution is required in real-time. The relevant algorithms for Bayesian sequential updating of probability distributions include the Kalman Filter (KF) [1], the Particle Filter (PF) [2], the Ensemble Kalman Filter (EnKF) [3].

With the information of prior and likelihood, the posterior distribution can be calculated, which helps gain more insight into the estimated quantities. In some cases, the posteriors are simple and can be presented or approximated as tractable and common distributions. However, the posterior distribution may be complex for sophisticated models, for example, the distribution is multimodal or high dimensional. The MCMC algorithm provides a solution where the distribution of the accepted samples converges to the true posterior distribution in the long run [4]. Compared with many deterministic approximation methods, the MCMC algorithm is conceptually easy to adopt for complex systems.

Although the MCMC algorithm has less requirement for the system, the algorithm is computationally slow. Therefore, it was not commonly applied to oil well systems in early research. Benefiting from the development of computer technology, a parallel computing scheme can be used for the MCMC algorithm.

For the dynamically evolving datasets, KF and PF have been used in the past to solve data reconciliation and data assimilation problems. Derived KF approaches, algorithms such as the Extended Kalman Filter (EKF), the Unscented Kalman Filter (UKF), and the EnKF, solve the problem where the model is nonlinear. Models and algorithms are formed in these methods to recursively update and estimate quantities.

By linearizing the model using differentiation, the covariance matrices are propagated assuming Gaussian distribution in EKF. However, the analytical computation is not feasible for non-Gaussian models or models which can not be succinctly linearized. In these cases, numerical strategies can be applied to estimate the system states.

Another approach to solving nonlinear problems is PF. As a nonparametric approach, PF estimates the belief by sampling the model output and assigning weights to these samples. Resampling schemes are used to estimate the posterior distribution in each iteration. The high weighted particles are used to design copies and low weighted particles are rejected [5, 6].

In the EnKF method, stochastic models are used to calculate the probability distribution and time evolution of the states. During the initialization, particles are generated around the initial states and these particles are designed according to the prior probability distribution. The term ‘ensemble members’ will be used in this paper to refer to these particles in the EnKF algorithm. All ensemble members are propagated and updated during the iterations. The uncertainty is predicted by calculating the error covariances. Because only a few ensemble members play a vital role eventually in the integration and the weights of most ensemble members become very small, the accuracy of estimation is related to the ensemble size [7]. Typically, the number of ensemble members should be in the range of 50-200 for computing mean and covariance, and considerably larger for computing higher order statistics. However, the large ensemble size also causes high computational cost. Compared with EKF, EnKF is easy to be implemented as it does not demand differentiation. For the sequential filters, EnKF requires substantially less computational cost than PF.

In the oil well area, MCMC and EnKF have been used in some research work. One study by Maraggi et al. [8] chose a Bayesian approach to estimate two parameters of oil reservoirs, which was simplified as a single dimensionless equation. Following an adaptative 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.

Kang et al. [9] applied a similar method to an electric submersible pump system. With given 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 via probability density functions using uncertainty assessment [10]. Autocorrelation was used to evaluate the samples, and sensitivity analysis was employed for capturing the region of convergence of the likelihood function. However, the research study did not take into account the samples process in much detail. The convergence of multiple chains was not clearly shown to provide reliable parameter estimation.

In terms of the application of the EnKF to the oil well field, a recent study [11] proposed to use principal component analysis for selecting the initial models for EnKF, so that fewer ensemble members were needed to predict production performances of the channel reservoirs. Meanwhile, the accuracy of the prediction is better than the prediction using one model when the ensemble size is the same. With a smaller ensemble size, computational time decreased significantly. Compared with the original EnKF, this work with model selection scheme provided a solution for filtering data of a complex model with less time. However, the selection of representative initial models demands empirical information of the system, for example, permeability distribution in this work. The selection is the part of the work which belongs to prior design.

With coupled machine learning and EnKF, a data-driven method was proposed to estimate the properties of the reservoir using pressure transient data [12]. Prior to commencing parameter estimation, the random forest method was adopted to classify the model using the discrete linear segment slopes in transition parts. After deciding which model should be used, the grid search method was conducted to estimate three hyper-parameters. With partitioned subspace, a decision tree was designed to address the optimal partitions. The accuracy of the optimization was validated by using cross validation. Once the model and parameters were decided, EnKF was applied to predict the pressure transients of the water drive gas reservoir well.

The purpose of this investigation is to estimate parameters and initial states in the process of a gas lifting oil well model using a first-principle model and batch data, and then sequentially estimate the states online. The dominant noise is Gaussian white noise on the measurements. MCMC was employed to approximate the posterior distributions of parameters and initial states in the first stage. Once the estimates and uncertainty were extracted, EnKF was carried out to estimate states at each sampling time. This study provides a solution to advance the understanding of the uncertainty in a complex system in real-time.

The overall structure of the study is as follows. Section 2 presents the methodology which is adopted in this work. The third section introduces the gas lifting oil well model which is used in this work and explains the procedure of the simulation. Section 4 validates the performance of the method, Section 5 provides a discussion of the results, while conclusions are given in Section 6.

2. Parameter Estimation and States Filtering

In this work, the estimation include two parts: parameters and initial state estimation, and state estimation. We used the Metropolis-Hastings (MH) Algorithm to estimate the distribution of parameters and initial states based on Bayesian analysis. Once we have found the distribution of the parameters and initial states, we draw samples from part of the accepted list for the initialization of the EnKF.

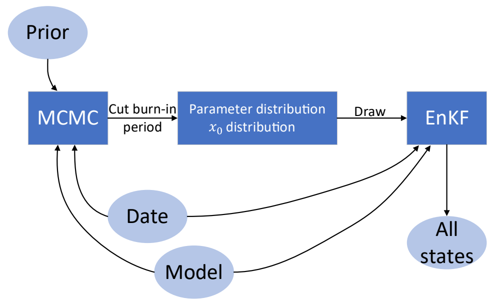

Figure 1: Outline of the estimation process.

2.1. Parameter and Initial State Estimation

Among the MCMC algorithms, the MH algorithm is one of the algorithms which has least requirement of the posterior and is simple to apply [13]. We use MH here to estimate the parameters and initial states by accepting or rejecting proposed samples.

A strong relationship between the accuracy of the MCMC estimation and the chain length has been reported in the literature [14]. Similar to most MCMC algorithms, the step length of MH is a pivotal parameter in the algorithm and impacts the efficiency of the algorithm. A series of experiments were run and step length for each quantity was selected after comparing the results, so that proposed samples can reach large parameter space and can also converge to an optimised value.

For each MH iteration, differential equations in the oil well model need to be solved. Therefore, it takes a long time to run more than 20 chains with a large number of iterations, for example, 20 chain with 1000 iterations for the outputs which contain 3600 samples takes more than six hours with the processor Intel(R) Core(TM) i7-10750H CPU @ 2.60GHz 2.59 GHz. To increase the computational efficiency, the number of chains and the number of samples for each chain were chosen carefully by trial and error.

After running the MH algorithm, we have a list of accepted samples. These samples form a chain. We can check how the chain explores the parameter space by plotting these samples. The distribution of these samples can be used to approximate the distribution of the posterior of the parameters.

Algorithm 1: Ensemble Kalman Filter algorithm

Initialization:

Draw samples for initial states, $x^i_{0|0}$, from Q, $i \in \{1, ..., n\}$

Calculate the mean of the samples $\hat{x}_{0|0} = \frac{1}{n}\sum_{i=1}^{n} x^i_{0|0}$

Give uncertainty

$X_{0|0} = \frac{1}{n-1}\sum_{i=1}^{n}(x^i_{0|0} - \hat{x}_{0|0})(x^i_{0|0} - \hat{x}_{0|0})^T$

for k = *1,2,...,*N **do**

Propagation:

$x^i_{k|k-1} = f(x^i_{k-1|k-1}, u_{k-1}, e^i_{k-1})$

$\hat{x}_{k|k-1} = \frac{1}{n}\sum_{i=1}^{n} x^i_{k|k-1}$

Update:

$y^i_{k|k-1} = h(x^i_{k|k-1}, u_{k-1}, v^i_{k-1})$

$\hat{y}_{k|k-1} = \frac{1}{n}\sum_{i=1}^{n} \hat{y}^i_{k|k-1}$

$Z_{k|k-1} =$

$\frac{1}{n-1}\sum_{i=1}^{n}(x^i_{k|k-1} - \hat{x}_{k|k-1})(y^i_{k|k-1} - \hat{y}_{k|k-1})^T$

$\varepsilon_{k|k-1} = \frac{1}{n-1}\sum_{i=1}^{n}(y^i_{k|k-1} - \hat{y}_{k|k-1})(y^i_{k|k-1} - \hat{y}_{k|k-1})^T$

$K_k = Z_{k|k-1}\varepsilon^{-1}_{k|k-1}$

$x^i_{k|k} = x^i_{k|k-1} + K_k(y_k - y^i_{k|k-1})$

$\hat{x}_{k|k} = \frac{1}{n}\sum_{i=1}^{n} x^i_{k|k}$

$X_{k|k} = \frac{1}{n-1}\sum_{i=1}^{n}(x^i_{k|k} - \hat{x}_{k|k})(x^i_{k|k} - \hat{x}_{k|k})^T$

end

Finally, a subset Q is prepared to include the estimation information for the initialization of the estimation of the state. According to the posterior distribution, Q is designed to cover the most likely estimation by ignoring the burn-in, which is the beginning part of the chain before the chain converges to a certain value.

2.2. State Estimation

The aim of this section is to estimate the model states, given a first-principle model and a set of measurements with unknown uncertainties. The data assimilation method, EnKF, was applied as in Algorithm 1 [7].
The distribution of states at each sampling time were approximated via a number of ensemble members x^i_k. The superscript shows the number of the ensemble member and the subscript presents the sampling time. Initial ensemble members are drawn from the subset Q to include the prior knowledge. In the EnKF algorithm, previous studies evaluating the algorithm results observed that the ensemble size is important [7]. The ensemble size that we used in the algorithm is n. The covariance of the states at sampling time k is X_k. The experiment contains N time steps.
In the propagation stage, the states are forecasted based on the last states and the dynamic system model. During the update stage, the estimated outputs are expressed based on the dynamic system model. e and v are the process noise and measurement noise with zero means. y_k is the measurement vector which contains noise. The means and covariance of the states are updated at the end of each time step. The states of the oil well model can be filtered in real time.

3. Simulation Study

3.1. Gas Lifting Oil Well Model

The gas lifting oil well model used in this work is based on previous work [15, 16]. The input of the model is the valve opening of the gas lift choke valve, u. 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 [$\frac{\text{kg/hr}}{\text{bar}}$]. PI is a mathematical means of expressing the ability of a reservoir to deliver fluids to the wellbore. All these parameters are dimensionless.

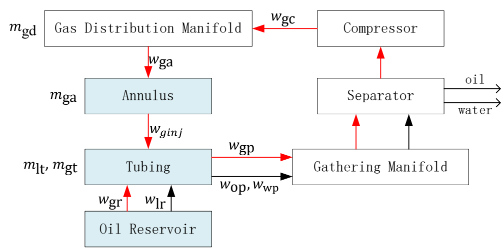

Figure 2: Flow chart of an oil field: the blue blocks are the components of an oil well. The red arrows show the gas flow and black arrows present the liquid phase flows, which can contain oil and water. The mass and flow rate are depicted beside the corresponding components where these states occur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

The flow chart in Fig. 2 shows the change of flows and masses in the gas lifting oil well. m denotes accumulated mass, and w denotes mass flow rate. A dot on top of a mass implies the time derivative of mass. A subscript indicates the phase (g: gas, o: oil) and location of the various variables. The output vector is $y = \{w_{\text{ga}}, w_{\text{gp}}, w_{\text{op}}, w_{\text{wp}}, P_{\text{wf}}, P_{\text{wh}}, P_{\text{a}}\}$. Flow measurements are impacted by bubbles and are not as reliable as other data such as temperature and pressure. Pressure transmitters are used to measure the bottom hole pressure and well flow pressure (P_{wf}), the pressure in the tubing upstream the production choke valve (P_{wh}), and the pressure 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 as 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 assume the temperature is constant everywhere and all these temperatures are assumed to be equal to T [K].

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}, the mass in the tubing above injection point m_{gt}, and the mass of liquid in the tubing above injection m_{lt}. After time derivative, the mass balance is mainly presented as Eq. (1):

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

, 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. The state vector is $x = \{m_{\mathrm{ga}}, m_{gt}, m_{lt}\}$.

3.2. Simulation Setup

There has been an increased interest in the programming language Julia [17] in recent years. As a free language, it has shown rapid advances in the field of numerical computing. In this work, we adopted Julia as the main language, especially for the state estimation part. The parameter and initial state estimation part is based on the previous work [16]. The Julia packages $Plots$, $StatsPlots$, $Distributions$, $DifferentialEquations$, $Noise$, $Random$ were used in the this work.

The work began by generating data from the oil well model with true parameters and initial states, and then Gaussian noise with zero means was added to the outputs. The oil well model does not contain process noise. The noise in the measurements only include measurement noise. The true parameters are $WC = 0.18, PI = 2.4 \times 10^4, GOR = 0.15$. The true initial states of $m_{\mathrm{ga}}, m_{\mathrm{gt}}$ and m_{lt} are $m_1 = 8650, m_2 = 3306, m_3 = 18250$ respectively. The true values were only revealed after the estimation for evaluation. The information of noise is not accessible during the experiments. The simulation will then go on to estimate the parameters and the initial states, as well as the uncertainty of the model. Once the distribution of these quantities are found, the states will be estimated.

The prior of the parameters and initial states were set as uniform distribution in Eq. (2). The GOR and WC parameters represent ratios, so they are between 0 and 1.

$$P(\theta, m_0 \mid u_i, y_{e(i,j)}) = \begin{cases} B & PI \in [10^4, 10^5], GOR \in [0,1], WC \in [0,1], \\ & m_1 \in [4000, 10000], m_2 \in [1000, 5000], \\ & m_3 \in [1 \times 10^4, 2.5 \times 10^4] \\ 0 & \text{otherwise} \end{cases} \quad (2)$$

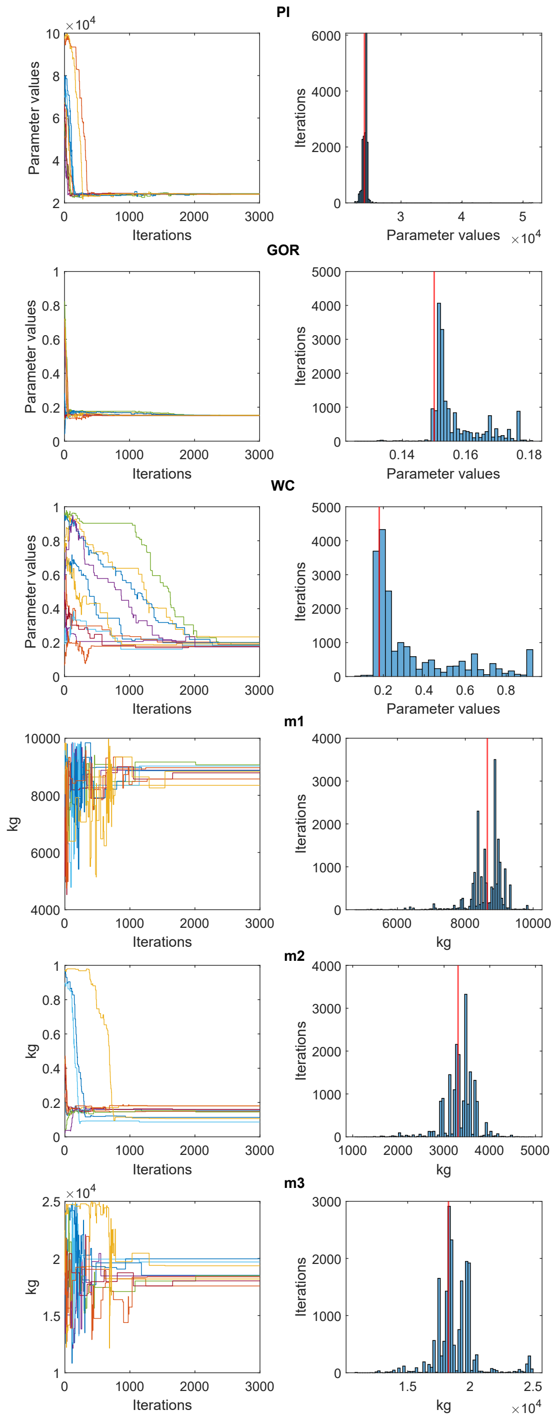

Figure 3: Estimation of the parameters and the initial states. The top three rows of plots show the parameter estimates and the remaining plots present the initial value estimates. The red vertical lines in the right columns of plots show the true values of the parameters and initial states.

In this work, we run 10 chains, and each chain contains 3000 iterations to identify the parameters and initial states using data with 40 samples. We assume the measurements are independent and contain white noises which are normally distributed with zero means, but with different variances. For every individual measurement, the variance is assumed to be constant during the experiments. The process error, e^i_{k-1}, used at the propagation stage in EnKF was tuned by trial and error. The measurement noise, v^i_{k-1}, used at the update stage was calculated from the data.

4. Estimate validation

The main aim of this section is to present and analyse the results of experiments in various ways. To gain insights into every parameter in the estimation process, we present plots of individual quantities distributions and chains. The interactions between every pair parameters are shown in scatter plots and contour plots. The estimated states are compared with the true states.

4.1. The Distribution of Quantities and Chains

In order to test the influence of the random initial value of the MCMC algorithm and check the explored range of the estimated distribution, we present the plots of the chains and the distribution for each parameter and initial value, shown in Fig.3. The plots include the estimations of ten chains of the MCMC algorithm with 3000 iterations.

The left column of plots show trace plots with various initial values, which provides information on the converging speed of the chains. The vertical axis is the estimated value, and the horizontal axis is the iteration number. The burn-in periods of these chains are generally less than 2000. Compared with other chains, WC and m_1 took longer time to converge to a certain range. We cut the burn-in periods and used the last 1000 samples as the subset Q.

According to these plots, all chains converge to around the true values after exploring the whole prior range. The end of the chains overlap each other and fairly smoothly drift around the optimal estimated value, namely the estimate uncertainty decreased as the iterations progressed. All the chains for the same parameters converge to a similar region and mix well, which indicates convergence is achieved.

The right column of the figures illustrates the distribution of each individual estimate. The histograms of the quantities are drawn based on the accepted lists of all chains. According to the distribution, the most likely estimates of these quantities are around the true values. The MCMC algorithm was able to identify the true values of the quantities.

4.2. Pair plots of the Distribution of Quantities

Scatter plots and contour plots are shown in Fig.4 to present the relationship between each pair of the parameters and initial states. According to the top two rows of plots, the main distribution of PI is between 2.35×10^4 to 2.50×10^4. The first and third rows of plots show that GOR distributes around [0.14, 0.20]. The samples of WC spread between 0.2 to 0.8. The initial state estimates are not as converged as the parameter estimates. Most of m_1 samples are between 8000 and 10000. m_2 is estimated within [3000, 4000], and the value of m_3 is locate in $[1.6, 2]\times10^4$. These plots confirm the estimates in Fig.3. The contour plots in the second column demonstrate more details in the converging area and show that some of the posterior distributions are multimodal.

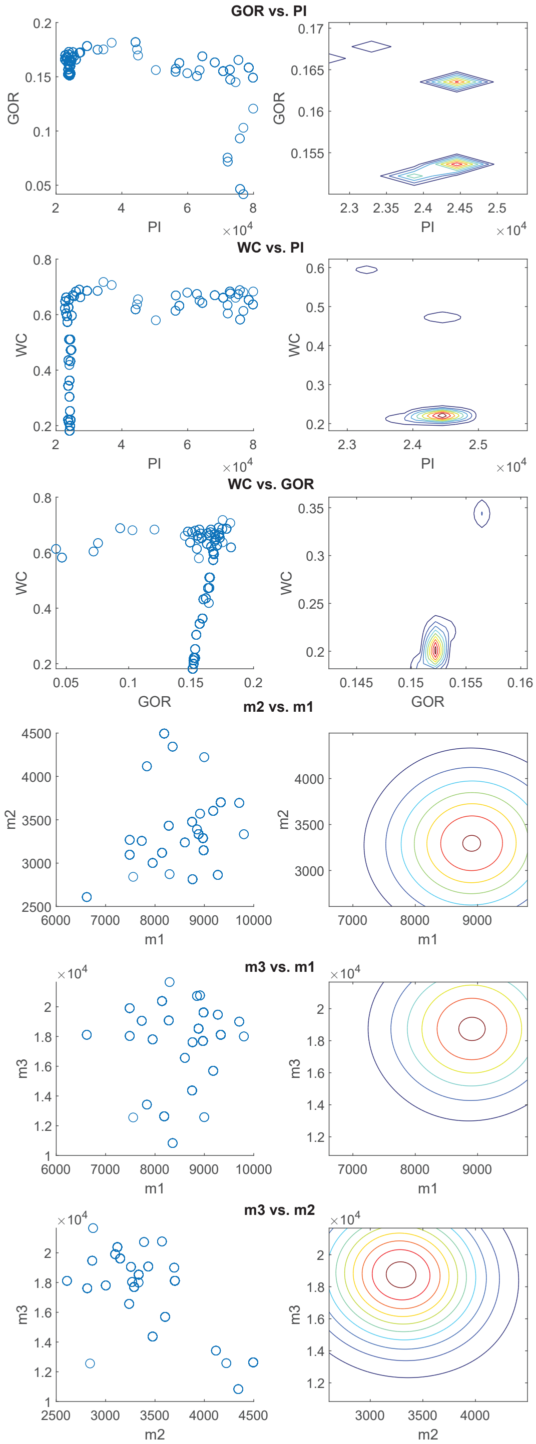

Figure 4: Estimation of parameters and initial states with 2D plots in scatter plots and contour plots. The top three rows of plots show the parameter estimates and the remaining plots present the initial value estimates. The contour plots are drawn in the range with minimum and maximum values. The colors of the contour plots show the density of estimated values. The red curves shows the estimation at high density, while the blue curves correspond to the low density contour. For the estimation of parameters, the converge parts were zoomed in to present the contour clearly.

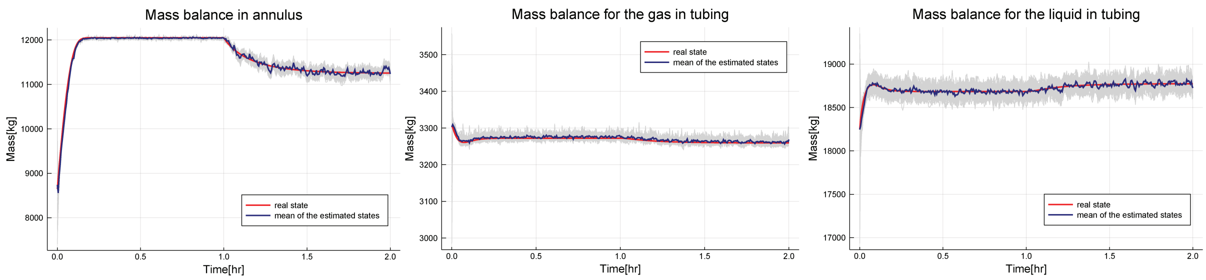

Figure 5: Estimation of states. The grey shadows are the estimated states which include 100 estimates for each sampling time. The dark blue lines are the mean of these ensemble values. The true values of these states are presented as red lines.

4.3. State estimation

We compare the estimated states and the true states in this subsection. Figure 5 shows the estimation result. With 1000 samples from the accepted list Q, the EnKF algorithm is initialised with 100 ensemble members. The time evolution of these ensemble members are shown as grey shadows in the plots. The starting points of these grey shadows indicate the initialisation range of ensemble. For example, the initial ensemble members of mass in the annulus is distributed within the range (7000, 10000) and the initial ensemble members of mass balance for gas and liquid in tubing are in the ranges (2950, 3550) and (16800,19400), respectively. The true states were collected from model simulation without any noise.

According to the result, the mean of all estimated states are close to the true states. All the ensemble values converge quickly towards the true states, though some of the initial values are distributed with a considerable deviation from the true initial states. Because the estimation of the states is based on the model and the measurements with noise, the estimates are less smooth than the true state.

The covariance matrix shows the uncertainty of the estimation. The uncertainty can also be checked through the width of the grey shadow. The uncertainty of the estimates of the mass for the gas and liquid in the tubing does not change significantly after the input changing at $t = 1$, while the uncertainty of the estimates of the mass for the gas in tubing increases significantly. The increase of uncertainty results from the propagation of the error between the estimated output and the measurement. Compared with the last two states, the first state changes more noticeable after the input changes.

5. Disscussion

Due to the lack of real data, it is not possible to exam the method here to a real gas lifting oil well system. In real life, the model mismatch might lead to undesirable estimation. Besides, in real oil well systems, most states and measurements are coupling. We simplified the problem and assumed the noises of these quantities are independent, so that the measurement error in the EnKF algorithm could be easily found by calculating the variance of the measurement during steady state. The assumption might lead to some mismatch in the estimation.

During the data filtering stage, the initial ensemble contains the prior information. However, it is not possible to add constraints to the time evolution of these ensemble. The estimated states are calculated based on the measurement and the last estimated states. For example, the mass should be positive all the time according to physics knowledge, but the EnKF algorithm can not guarantee that the estimates are always positive. The shortage of the EnKF algorithm might lead to some improper estimation.

Another potential problem is that this research study does not engage with the circumstance where the parameters change over time. In reality, models are always imperfect representations of a system, and it may be necessary to allow for parameters to change over time to achieve the best possible model fit.

6. Conclusions

The purpose of this study was to investigate the feasibility of estimating parameters and states of a gas lifting oil field. In the first stage, the posterior distributions of the parameters and the initial states were identified using a MCMC algorithm. Then, we drew samples from a subset of the posterior distribution for the initialization of the EnKF and estimated the states. The analysis of experimental results undertaken here has confirmed that the proposed solution was able to find the true values of the unknown quantities with uncertainty during estimation. The study contributes to our understanding of the uncertainty in the estimation of the quantities of a system in real-time. This understanding helps to improve the gas lifting oil well model and provide estimations of the process states in the system, why might benefit the operation and advanced control in real life. Further research could be conducted to improve the algorithm for the estimation of changing parameters.

Acknowledgment

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

[1] R. E. Kalman, “A new approach to linear filtering and prediction problems,” 1960.

[2] A. Doucet, S. Godsill, and C. Andrieu, “On sequential monte carlo sampling methods for bayesian filtering,” *Statistics and computing*, vol. 10, no. 3, pp. 197–208, 2000.

[3] G. Evensen *et al.*, *Data assimilation: the ensemble Kalman filter*, vol. 2. Springer, 2009.

[4] S. Brooks *et al.*, “Handbook of markov chain monte carlo. boca raton: Crc,” 2011.

[5] W.-s. Chen, B. R. Bakshi, P. K. Goel, and S. Ungarala, “Bayesian estimation via sequential monte carlo sampling: unconstrained nonlinear dynamic systems,” *Industrial & engineering chemistry research*, vol. 43, no. 14, pp. 4012–4025, 2004.

[6] A. Doucet, N. De Freitas, N. J. Gordon, *et al.*, *Sequential Monte Carlo methods in practice*, vol. 1. Springer, 2001.

[7] G. Evensen and P. J. Van Leeuwen, “Assimilation of geosat altimeter data for the agulhas current using the ensemble kalman filter with a quasigeostrophic model,” *Monthly Weather Review*, vol. 124, no. 1, pp. 85–96, 1996.

[8] L. M. R. Maraggi, L. W. Lake, and M. P. Walsh, "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.

[9] E. Costa, O. de Abreu, T. d. O. Silva, M. Ribeiro, and L. Schnitman, "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*, vol. 205, p. 108880, 2021.

[10] JCGM, "Evaluation of measurement data—guide to the expression of uncertainty in measurement," *JCGM*, vol. 100, no. 2008, pp. 1–116, 2008.

[11] B. Kang, H. Jung, H. Jeong, and J. Choe, "Characterization of three-dimensional channel reservoirs using ensemble kalman filter assisted by principal component analysis," *Petroleum Science*, vol. 17, no. 1, pp. 182–195, 2020.

[12] L. Xue, S. Gu, L. Mi, L. Zhao, Y. Liu, and Q. Liao, "An automated data-driven pressure transient analysis of water-drive gas reservoir through the coupled machine learning and ensemble kalman filter method," *Journal of Petroleum Science and Engineering*, vol. 208, p. 109492, 2022.

[13] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, "Equation of state calculations by fast computing machines," *The journal of chemical physics*, vol. 21, no. 6, pp. 1087–1092, 1953.

[14] J. Kruschke, "Doing bayesian data analysis: A tutorial with r, jags, and stan," 2014.

[15] R. Sharma, K. Fjalestad, and B. Glemmestad, "Modeling and control of gas lifted oil field with five oil wells," in *52nd International Conference of Scandinavian Simulation Society, SIMS*, pp. 29–30, 2011.

[16] Z. Ban, A. Ghaderi, N. Janatian, and C. F. Pfeiffer, "Parameter Estimation for a Gas Lifting Oil Well Model Using Bayes' Rule and the Metropolis–Hastings Algorithm," *Modeling, Identification and Control*, vol. 43, no. 2, pp. 39–53, 2022.

[17] J. Bezanson, A. Edelman, S. Karpinski, and V. B. Shah, "Julia: A fresh approach to numerical computing," *SIAM review*, vol. 59, no. 1, pp. 65–98, 2017.