An adaptive robust optimization scheme for water-flooding optimization in oil reservoirs using residual analysis

Siraj, M. Mohsin; Van den Hof, Paul M.J.; Jansen, Jan Dirk

DOI
10.1016/j.ifacol.2017.08.1632

Publication date
2017

Document Version
Final published version

Published in
20th World Congress of the International Federation of Automatic Control (IFAC), 2017

Citation (APA)
Siraj, M. M., Van den Hof, P. M. J., & Jansen, J. D. (2017). An adaptive robust optimization scheme for water-flooding optimization in oil reservoirs using residual analysis. In D. Dochain, D. Henrion, & D. Peaucelle (Eds.), 20th World Congress of the International Federation of Automatic Control (IFAC), 2017 (1 ed., Vol. 50, pp. 11275-11280). (IFAC papers online; Vol. 50). Elsevier. https://doi.org/10.1016/j.ifacol.2017.08.1632

Important note
To cite this publication, please use the final published version (if applicable). Please check the document version above.

Copyright
Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

Takedown policy
Please contact us and provide details if you believe this document breaches copyrights. We will remove access to the work immediately and investigate your claim.
An adaptive robust optimization scheme for water-flooding optimization in oil reservoirs using residual analysis

M.Mohsin Siraj*, Paul M.J. Van den Hof* and Jan Dirk Jansen**

* Electrical Engineering, Control Systems group, Eindhoven University of Technology, The Netherlands. (e-mails: m.m.siraj, p.m.j.vandenhof@tue.nl).

** Department of Geoscience and Engineering, Delft University of Technology, The Netherlands. (e-mail: j.d.Jansen@tudelft.nl)

Abstract: Model-based dynamic optimization of the water-flooding process in oil reservoirs is a computationally complex problem and suffers from high levels of uncertainty. A traditional way of quantifying uncertainty in robust water-flooding optimization is by considering an ensemble of uncertain model realizations. These models are generally not validated with data and the resulting robust optimization strategies are mostly offline or open-loop. The main focus of this work is to develop an adaptive or online robust optimization scheme using residual analysis as a major ingredient. The models in an ensemble are confronted with data and an adapted ensemble is formed with only those models that are not invalidated. As a next step, the robust optimization is again performed (i.e., updated/adjusted) with this adapted ensemble. The adapted ensemble gives a less conservative description of uncertainty and also reduces the high computational cost involved in robust optimization. Simulation example shows that an increase in the objective function value with a reduction of uncertainty on these values is obtained with the developed adaptive robust scheme compared to an open-loop offline robust strategy with the full ensemble and an adaptive scheme using Ensemble Kalman Filter (EnKF), which is one of the most common parameter estimation methods in reservoir simulations.

© 2017, IFAC (International Federation of Automatic Control) Hosting by Elsevier Ltd. All rights reserved.

Keywords: Uncertainty handling, water-flooding optimization, residual analysis, online robust optimization

1. INTRODUCTION

Dynamic optimization of the water-flooding process has shown significant scope for improvement of the economic life-cycle performance of oil fields compared to a more conventional reactive strategy, see e.g., Brouwer and Jansen (2004); Jansen et al. (2008); Foss (2012); Van den Hof et al. (2012). Besides computational complexity, induced due to complex non-linear dynamics and hence non-convexity, one of the key challenges in this model-based dynamic optimization is the high levels of uncertainty arising from the modeling process of water flooding and from strongly varying economic conditions.

Various approaches to decision making under uncertainty can be implemented in two different ways, see e.g., Bertsimas and Thiele (2006). In open-loop or offline schemes, robust optimization is performed only once under a given description of uncertainty. Robust optimization can also be used in an adaptive or online fashion where the uncertainty is reduced with the information that is revealed over time. A general practice of quantifying uncertainty in water-flooding optimization is a scenario-based approach where an ensemble of uncertain parameters (e.g., reservoir models), see e.g., Van Essen et al. (2009); Capolei et al. (2013) is considered. These models are mostly either generated by using geostatistical tools, see e.g., Mariethoz and Caers (2014) or hand drawn, and are typically not (in)validated by the production data. Hence they may provide a (very) conservative description of uncertainty. An adaptive scheme, i.e., Closed-Loop Reservoir Management (CRLM) has been introduced in Jansen et al. (2005), where the reservoir model variables (states and/or parameters) are updated using data assimilation or Computer Assisted History Matching (CAHM) techniques, such as Ensemble Kalman Filter (EnKF), variational approaches, etc., see e.g., Evensen (2009); Aanonsen et al. (2009); Oliver and Chen (2011) and the optimization is adapted with updated model(s). In the robust settings, as robust optimization uses an ensemble of model realizations, posterior ensemble, e.g., estimated by EnKF, can be directly used in an adaptive fashion resulting in a robust CRLM, see e.g., Chen et al. (2009), Chen et al. (2010), Capolei et al. (2013).

The purpose of this work is to devise an adaptive robust scheme that can be updated with the given production data. The main focus is to address the question: how the available information (data) with time can be used to shrink the uncertainty space by selecting fewer number of
representative models in an ensemble (reduce the ensemble size) or in other words how uncertainty is propagated and can be reduced in an adaptive online setting? The concept of residual analysis is used, where the models in an ensemble are confronted with data and are invalidated if they do not sufficiently agree with the observed data. An adapted ensemble is formed with only those models that are not invalidated thus providing a less conservative description of uncertainty with a reduced number of models in an ensemble. The adapted ensemble is used in a robust optimization in an online fashion. Residual analysis follows an ‘exclusion approach’ to uncertainty modeling, see e.g., Tarantola (2006); Caers (2011), which focuses on starting from all possibilities (models) and then excluding those possibilities (models) that can be ‘rejected’ by any information available to us, e.g., with rejection sampling which is a probabilistic approach of rejecting models. EnKF is also implemented and the posterior ensemble is used in robust optimization. The results of the EnKF in terms of uncertainty reduction and performance improvement are compared with the developed approach.

The paper is organized as follows: in the next section, model-based water-flooding optimization and uncertainty quantification are discussed. The concept of residual analysis is presented in section 3. In section 4, the adaptive robust scheme is introduced. Simulation example with the introduced scheme is given in section 5 followed by the conclusions in section 6.

2. UNCERTAINTY IN WATER-FLOODING OPTIMIZATION

Water-flooding involves the injection of water in an oil reservoir to increase oil production. Net Present Value (NPV), as an objective for the dynamic optimization of the water-flooding process, can be mathematically represented in the usual fashion as:

\[ J = \sum_{k=1}^{K} \left[ \frac{r_o \cdot q_{o,k} - r_w \cdot q_{w,k} - r_{inj} \cdot q_{inj,k}}{(1 + b)^{\frac{\Delta t_k}{b}}} \right] \]

where \( r_o, r_w \) and \( r_{inj} \) are the oil price, the water production cost and the water injection cost in \( \frac{m^2}{day} \) respectively. \( K \) represents the production life-cycle i.e., the total number of time steps \( k \) and \( \Delta t_k \) the time interval of time step \( k \) in days. The term \( b \) is the discount rate for a certain reference time \( \tau \). The terms \( q_{o,k}, q_{w,k} \) and \( q_{inj,k} \) represent the total flow rate of produced oil, produced water and injected water at time step \( k \) in \( \frac{m^3}{day} \).

Model uncertainty is the prime source of uncertainty in model-based optimization of the water-flooding process. Traditionally an ensemble of uncertain dynamic model realizations is considered to quantify the uncertainty space \( \Theta \), i.e., \( \{ M(\theta_1), M(\theta_2), \ldots, M(\theta_{N_{geo}}) \} \), where \( M \) is a dynamic model with \( \theta_i \in \Theta, i = 1, 2, \ldots, N_{geo} \) a realization of a vector of uncertain parameters. This ensemble-based uncertainty set can be used with various robust schemes. One of the simplest offline or open-loop scenario-based robust approaches is to maximize the average of NPV over the model uncertainty ensemble, as introduced in Van Essen et al. (2009). Robust optimization (or Mean Optimization (MO)) can be formulated as:

\[ J_{MO} = \frac{1}{N_{geo}} \sum_{i=1}^{N_{geo}} J(u, \theta_i) \]

where \( J \) is NPV and \( u \) is the input decision variable. Other offline (open-loop) robust approaches in water-flooding optimization, e.g., mean-variance and mean-CVaR have been introduced in e.g., Capolei et al. (2015); Siraj et al. (2015). In these offline strategies, the optimal solution is devised for the complete production life of the reservoir under a given uncertainty set and the optimization is not updated/adapted to the available information with time. These offline (open-loop) approaches mainly aim only at minimizing the negative effect of uncertainty on the achieved NPV and do not deal with shrinking the uncertainty space \( \Theta \).

The uncertainty reduction can be achieved with the help of available information. The estimation of physical parameters with the available production data is one of the ways to reduce uncertainty. Data assimilation or CAHM algorithms such as the Ensemble Kalman Filter (EnKF), variational methods are typically used in reservoir simulation offering a joint state and parameter estimation. This parameter estimation problem, due to a large number of to-be-estimated parameters, is ill-posed, i.e., many combination of parameter values will result in the same minimum value of the cost function. Therefore, CAHM typically uses a Bayesian framework with a prior distribution of the parameters reflected by a prior ensemble. Hence the estimation of physical parameters is highly influenced by the selection of this prior ensemble of parameter. These estimated parameters and the resulting posterior (adapted) ensemble with CAHM can be used in an adaptive fashion to update the robust strategies as presented, e.g., in Chen et al. (2009); Capolei et al. (2013), which is also referred to as closed-loop approaches in the petroleum engineering literature. The application of CAHM with nominal optimization has been presented, e.g., in Jansen et al. (2005); Sarma et al. (2005); Jansen et al. (2009). As the number of realizations in the adapted ensemble with CAHM are not reduced, the computational complexity of these adaptive robust optimization is not decreased. Another way to adapt the ensemble is by using clustering techniques. The number of realizations in an ensemble is reduced by clustering the models with the similar (static or dynamic) behavior, see e.g., Sarma et al. (2013). The clustering techniques are not data driven and the ensemble size reduction may not correspond to uncertainty reduction, hence it will only reduce the complexity of the robust optimization problems.

In the next sections, residual analysis is introduced which offers minimization of uncertainty space and a reduction of computational complexity of robust optimization.

3. RESIDUAL ANALYSIS

Model validation is usually performed by confronting the model with available information, e.g., production data, a priori knowledge. For a given ensemble of models, residual analysis follows an ‘exclusion approach’ to uncertainty, which focuses on rejecting models by available production data. Therefore, it does not only aim at minimizing the uncertainty space \( \Theta \) but it also reduces the size of the ensemble, leading to complexity reduction of subsequent
robust optimization. A principle difference between residual analysis and other data assimilation techniques is that residual analysis is performed on the model space which is smaller in size compared to the assimilation methods which focus on the large parameter space. Hence residual analysis does not suffer from the problem of ill-posedness and the effect of the selection of a poor prior ensemble.  

Fig. 1 illustrates a detailed overview of techniques which can be used to update the ensemble of realizations. CAHM techniques, such as variational methods follow a gradient-based approach while EnKF uses ensemble-based approach for joint state and parameter estimation.  

Regression problem, first order statistics provide a complete characterization of the validation problem, e.g., in the correlation analysis, the residual should be asymptotically uncorrelated with past input samples. For a nonlinear regression problem, the first order moments are not sufficient to draw any conclusions about the validity of the models. As the reservoir models are strongly nonlinear in nature, we use a deterministic metric, i.e., Best-fit-ratio (BFR) to define an invalidation test. The available production data is used for invalidation.  

The residual $\epsilon$ is defined as the difference between the observed (measured) output $y$ and the simulation output $\hat{y}$.  

$$\epsilon = y - \hat{y}.$$  

The Best-Fit Ratio (BFR) or the fit ratio is defined as:  

$$\text{BFR} = 100\% \times \max \left( 1 - \frac{||\epsilon||_2}{||y - \hat{y}||_2} , 0 \right)$$  

where $\hat{y}$ is the average of the measured output $y$. The BFR percentage is a relative measure often used in system identification, and a low value of BFR indicates a poor fit to data, see Ljung (1999). BFR is a unit-less quantity and gives an indication of fit in a percentage. Generally, as the reservoir models contain multiple outputs, an average BFR over each individual output channel is considered. The selection of BFR is not unique and other metrics, e.g., Mean Squared Error (MSE) or Variance Accounted For (VAF), see Ljung (1999) can also be used for defining the invalidation test. VAF measures how much variation in data (variance) is captured by the model output and disregards the mismatch (bias) of data with model output while the MSE measure is dependent on the unit of physical quantity being measured, therefore BFR is used for residual analysis.  

The test for invalidating models is chosen as:  

$\mathcal{M}(\theta, u)$ is not invalidated if $E[\text{BFR}(\mathcal{M}(\theta, u))] > 30\%$, for a given $u$  

where $E[\cdot]$ is the expected value and calculated as the average over all outputs. It implies that all those models with an average BFR of above 30% are retained in the adapted ensemble. The selection of 30% is chosen in an ad-hoc way. One of the risks with this selection criterion is that all the models in an ensemble can be rejected. An alternative choice for the invalidation test is by considering e.g., 10% models who score the highest average BFR within the ensemble. This criterion is not used in this work.  

A flow chart as shown in Fig. 2 explains residual analysis. At a given time step, the concept of residual analysis from generating prior models to forming an adaptive ensemble is performed and later, as a final step of the developed adaptive strategy as discussed in the next section, the robust optimization is conducted with the adaptive ensemble. At the next time step, the adaptive ensemble in the previous step is considered as a prior and the procedure is repeated as shown in Fig. 2.
adapted ensemble as shown at the left side of the figure. Throughout this work, the MO approach of Van Essen et al. (2009) is used for robust optimization. Other robust measures such as mean-variance, mean-CVaR can also be used. An implementation of the adaptive robust strategy is presented in the next section to investigate if model invalidation by residual analysis can be an appropriate tool for reducing uncertainty and improving robust optimization of economic performance of oil reservoir in an online setting.

5. SIMULATION EXAMPLE

5.1 Ensemble of reservoir models

An ensemble of 100 geological realizations of the standard egg model is considered. Each model in the ensemble is a three-dimensional realization of a channelized reservoir produced under water flooding conditions with eight water injectors and four producers based on the original Egg model proposed in Van Essen et al. (2009). The life-cycle of each reservoir model is 3600 days. The complete list of parameters and the details about the model set are presented in Jansen et al. (2014). The true permeability field is considered to be the unknown parameter. The number of 100 realizations is assumed to be large enough to be a good representative of this parametric uncertainty space. The absolute-permeability field of the first realization in the set is shown in Fig. 4. Fig. 5 shows the permeability fields of six randomly chosen realizations of the standard egg model in an ensemble of 100 realizations. Each realization in the set is considered equiprobable.

5.2 An offline (open-loop) MO approach with complete ensemble

All the simulation experiments in this work are performed using MATLAB Reservoir Simulation Toolbox (MRST), see Lie et al. (2012).

Economic data for NPV An un-discounted NPV is used. Other economic parameters, i.e., oil price $r_o$, water injection $r_{inj}$ and production cost $r_p$ are chosen as $126 \frac{8}{m^3}$, $6 \frac{8}{m^3}$ and $19 \frac{8}{m^3}$ respectively.
Other economic parameters, i.e., oil price using MATLAB Reservoir Simulation Toolbox (MRST), ensemble.

5.1 Ensemble of reservoir models
An ensemble of 100 geological realizations of the standard egg model is considered. Each model in the ensemble is considered to be the unknown parameter. The number of realizations is set as 100.

5.2 An offline (open-loop) MO approach with complete set is shown in Fig. 4. Fig. 5 shows the permeability fields absolute-permeability field of the first realization in the ensemble to collect simulation data ˆy. Residual analysis with synthetic truth is performed with the adapted ensemble by residual analysis and EnKF. The economic parameters are the total production rate from each production well, i.e., y,

MO with the adapted ensemble: Robust optimization is updated/adjusted with information revealed, i.e., MO is again performed with the adapted ensembles generated by residual analysis and EnKF. The economic parameters are kept the same. The initial input value for the optimization is u_0. As the first sample of u_0, i.e., from the time 0 day to 360 days is already applied to the system, the remaining part of the input from time 360 days to 3600 days (end of simulation time) is used. The time horizon for MO is reduced to 3600 − 360 = 3240 days. Optimal inputs are obtained as a result of robust optimization over the different ensembles for both residual analysis u_on,RA and EnKF u_on,EnKF and they are applied to the synthetic truth. Fig. 8 shows the comparison of time evolutions of NPV with the offline approach (open-loop with complete ensemble) and the adaptive approaches (online with adapted ensembles). We note that NPV is often defined as the cumulative discounted cash flow (CDCF) over the entire project life. Here we have chosen to use the term NPV (specifically NPV time-evolution) also for intermediate values of the CDCF. The introduced adaptive approach using residual analysis gives an increase of 0.62% in the NPV value compared to the offline (open-loop) approach and hence provides better optimization results. MO with the posterior ensemble with EnKF gives poor results and a decrease of 1.2% in NPV value compared to offline (open-loop) approach is observed. From the computational complexity viewpoint, as the number of model realizations in posterior ensemble by EnKF are not reduced compared to the adapted ensemble by residual analysis with only 22 members, the computational complexity of robust optimization using residual analysis is drastically reduced.

To analyze uncertainty reduction, the optimal solutions, i.e., u_0, u_on,RA and u_on,EnKF are applied to the complete and the adapted/posterior ensembles respectively. NPV points are collected and the corresponding PDFs are obtained by approximating a non-parametric KDE with MATLAB routine 'ksdensity' on these NPV data values.

The effect of uncertainty on the time evolutions of NPV is shown in Fig. 9. The maximum and the minimum values of the time evolutions of NPV from both ensembles will form a band. The width of the band shows the variability of the strategy over the ensemble of the model realizations. The smaller width of resulting from the adaptive strategy with residual analysis ensemble shows the effect of uncertainty reduction compared to adaptive strategy with EnKF and offline approach.
Chen, Y., Oliver, D.S., et al. (2010). Ensemble-based closed-loop optimization applied to brugge field. *SPE Reservoir Evaluation & Engineering*, 13(01), 56–71.

Evensen, G. (2009). *Data Assimilation: the Ensemble Kalman Filter*. Springer Science & Business Media.

Foss, B. (2012). Process control in conventional oil and gas fields challenges and opportunities. *Control Engineering Practice*, 20(10), 1058–1064.

Jansen, J.D., Brouwer, D.R., Naevdal, G., and Van Kruijsdijk, C.P.J.W. (2005). Closed-loop reservoir management. *First Break*, 23(1), 43–48.

Jansen, J.D., Douma, S.G., Brouwer, D.R., Van den Hof, P.M.J., Bosgra, O.H., and Heemink, A.W. (2009). Closed-loop reservoir management. *Paper SPE 119098, SPE Reservoir Simulation Symposium, The Woodlands, Texas, USA.*

Jansen, J.D., Fonseca, R.M., Kahrroaei, S., Siraj, M.M., Van Essen, G.M., and Van den Hof, P.M.J. (2014). The egg model—a geological ensemble for reservoir simulation. *Geoscience Data Journal*, 1(2), 192–195.

Jansen, J.D., Bosgra, O.H., and Van den Hof, P.M.J. (2008). Model-based control of multiphase flow in subsurface oil reservoirs. *Journal of Process Control*, 18(9), 846–855, DOI: 10.1016/J.JPROCONT.2008.06.011.

Lie, K.A., Krogstad, S., Ligaarden, I.S., Natvig, J.R., Nilsen, H.M., and Skadestad, B. (2012). Open-source MATLAB implementation of consistent discretisations on complex grids. *Computational Geosciences*, 16(2), 297–322.

Ljung, L. (1999). *System Identification – Theory for the User*. Prentice-Hall.

Mariethoz, G. and Caers, J. (2014). *Multiple-point geostatistics: stochastic modeling with training images*. John Wiley & Sons.

Oliver, D.S. and Chen, Y. (2011). Recent progress on reservoir history matching: a review. *Computational Geosciences*, 15(1), 185–221.

Sarma, P., Chen, W.H., Xie, J., et al. (2013). Selecting representative models from a large set of models. In *SPE Reservoir Simulation Symposium, 18-20 February, The Woodlands, Texas, USA*. Society of Petroleum Engineers.

Sarma, P., Durlofsky, L.J., and Aziz, K. (2005). Efficient closed-loop production optimization under uncertainty (spe94241). In *67th EAGE Conference & Exhibition*.

Siraj, M.M., Van den Hof, P.M.J., and Jansen, J.D. (2015). Risk management in oil reservoir water-flooding under economic uncertainty. In *Proc. of at 54th IEEE Conference on Decision and Control, 15-18 December, Osaka, Japan*, 7542–7547.

Tarantola, A. (2006). *Popper, bayes and the inverse problem*. *Nature physics*, 2(8), 492–494.

Van den Hof, P.M.J., Jansen, J.D., and Heemink, A. (2012). Recent developments in model-based optimization and control of subsurface flow in oil reservoirs. In *Proc. of the 2012 IFAC Workshop on Automatic Control in Offshore Oil and Gas Production*, 189–200.

Van Essen, G., Zandvliet, M., Van den Hof, P.M.J., Bosgra, O., and Jansen, J.D. (2009). Robust waterflood-optimization of multiple geological scenarios. *SPE Journal*, 14(01), 202–210, DOI: 10.2118/102913–PA.