Managing Geological Uncertainty with Decision Analysis in Reservoir Management

Probabilistic History Matching, Robust Production Optimization, and Value-of-Information

by

Aojie Hong

Thesis submitted in fulfillment of
the requirements for degree of
PHILOSOPHIAE DOCTOR
(PhD)

Department of Petroleum Engineering
Faculty of Science and Technology

2017

University of Stavanger
N-4036 Stavanger
NORWAY
www.uis.no

ISBN: 978-82-7644-744-6

ISSN: 1890-1387

PhD Thesis UiS No. 371

Preface

This thesis has been written to partially fulfill the graduation requirements of the Philosophiae Doctor (PhD) degree at the Department of Petroleum Engineering, Faculty of Science and Technology, University of Stavanger, Norway. I was engaged in writing this thesis from May to October 2017.

This research was conducted under the main supervisor, Prof. Reidar B. Bratvold at the University of Stavanger and the co-supervisor, Dr. Geir Nævdal at the International Research Institute of Stavanger. It was funded by the National IOR Centre of Norway from October 2014 to October 2017. The work was conducted mainly at the University of Stavanger and partially at the University of Texas at Austin from February to July 2017.

This work intends to illustrate and discuss the implementation of decision analysis tools to manage geological and petrophysical uncertainties for better decision making in reservoir management contexts. I believe that this work will be of great interest to both managers and engineers in oil and gas companies and to scholars and researchers at academic and research institutes, who are engaged in improving the quality of oil and gas operation related decisions.

Aojie Hong

Stavanger, November 7, 2017

Abstract

Reservoir management (RM) is a decision-oriented activity where decision makers use their current knowledge to search for production strategies that maximize the value of hydrocarbon production from a reservoir. Two central components of RM are history matching (HM) and production optimization (PO). HM draws on information from data. The information is then used to support decisions on production strategies. The optimal production strategy is identified through PO.

Decisions will not be good unless they account for relevant and material uncertainties in a given decision context. Uncertainty is a result of not having perfect (i.e., complete) information. Although the oil and gas industry has long been aware of the importance of uncertainty understanding and management, decision-driven approaches that include consistent uncertainty quantifications are not commonly or comprehensively used. The intent of using decision-driven approaches is to manage uncertainties for good decision making.

This work intends to address three of the main challenges of using decision analysis (DA) tools for managing geological and petrophysical uncertainties in RM. The first challenge is in describing the geological and petrophysical uncertainties using probability distributions that result from HM. The second challenge is in the computational complexity of modeling flow behaviors and solving for the optimal production strategy when given a description of geological and petrophysical uncertainties. The third challenge is in the PO approach that can allow for learning over time. The ultimate goal of this work is to illustrate and discuss how these challenges can be overcome and to facilitate the application of decision quality in RM contexts.

The first challenge is addressed through illustrating and discussing the implementation of probabilistic HM approaches. Unlike a deterministic HM approach which results in a single combination of production model parameters that best matches the given production data, a probabilistic HM approach produces numerous combinations of production model parameters, each of which has a probability, to quantify the geological and petrophysical uncertainties. Modeling with these combinations of production model

parameters propagates the geological and petrophysical uncertainties to the uncertainty in future production.

The second challenge is addressed through illustrating and discussing the implementation of robust optimization (RO) algorithms and proxy production models. RO approaches identify the optimal production strategy that maximizes the expected value over many geological and petrophysical realizations. To speed up the process of solving an RO problem, we use a proxy production model to supplement a grid-based reservoir model. The proxy model captures only the most relevant physics and mechanisms affecting production prediction and thus is much more computationally attractive than the grid-based reservoir model.

The third challenge is addressed through illustrating and discussing the implementation of a fully structured reservoir management (FSRM) approach. The FSRM approach is based on the fully structured decision tree for a sequential decision-making problem in a RM context. It allows for learning over time by considering both the uncertainties associated with current available data and the uncertainties associated with future data. Therefore, the current decision does not depend only on the uncertainties that a decision maker has learned so far but on the uncertainties that the decision maker will learn in the future. The FSRM approach provides the optimal production strategy, whereas the state-of-the-art RM approach—closed loop reservoir management (CLRM)—might give a sub-optimal production strategy. Furthermore, we illustrate and discuss an a priori analysis on information valuation, known as value-of-information (VOI) analysis, in RM contexts. It evaluates the benefits of collecting additional information before one gathers the data and makes a decision. It can also be used to assess the value of accounting for learning over time in PO.

This work presents numerous examples to demonstrate the value of applying DA tools in RM. The main contribution of this work is the illustration and discussion of the implementation of decision-driven approaches for good decision making in RM contexts. To achieve this purpose, we:

1. show how to integrate model uncertainty in probabilistic decline curve analysis for unconventional oil production forecasting;

2. illustrate and discuss how to use a capacitance-resistance model as a proxy model to speed up the process of robust production optimization;
3. implement a fast analysis of the optimal improved-oil-recovery start time using a two-factor production model and the least-squares Monte Carlo algorithm;
4. illustrate and discuss how to implement a robust discretization of continuous probability distributions for VOI analysis;
5. show how to apply the VOI concept for production model parameter updating through HM.

Although other challenges remain in implementing some specific approaches in a real-world setting, we believe that this work is useful in conveying to RM professionals the benefits and implementation of DA techniques, and it can be used as guidance for future research.

Acknowledgements

This thesis would not have been possible without the help and support from many people, to whom I owe my deepest gratitude.

Firstly, I would like to express my sincere gratitude to my supervisor, Prof. Reidar B. Bratvold, for the continuous support of my PhD study and related research. His insightful supervision and guidance helped me throughout the research and writing of this thesis. Working under his supervision is one of the best things I have experienced in my life.

My grateful thanks also go to my co-supervisor, Dr. Geir Nævdal, for his support and guidance. I appreciate highly his help in my understanding of mathematical formulations and advice on coding and programming.

I am deeply indebted to Prof. Larry W. Lake for providing me the opportunity to visit his research group at the University of Texas at Austin. His tough questions spurred me to widen my research from various perspectives. His taking good care of me made me really enjoy the stay at Austin.

I gracefully acknowledge the financial support of the National IOR Centre of Norway, the Research Council of Norway and their industry partners (ConocoPhillips, Skandinavia AS, BP Norge AS, Det Norske Oljeselskap AS, Eni Norge AS, Maersk Oil Norway AS, DONG Energy A/S, Denmark, Statoil Petroleum AS, ENGIE E&P NORGE AS, Lundin Norway AS, Halliburton AS, Schlumberger Norge AS, and Wintershall Norge AS) for my entire PhD period. My sincere thanks also go to the University of Stavanger and the Petroleum Research School of Norway for offering me the scholarship to visit the University of Texas at Austin.

I would like to thank all my colleagues at the University of Stavanger, especially my office mate, Philip Thomas, for his stimulating discussion and daily chat.

Last but not least, I would like to thank my family—my parents, wife and baby boy—for their spiritual support throughout my PhD study and my life.

I would like to apologize that I could not name here all those who helped me during my PhD period or were important to the completion of this thesis and my PhD study.

Aojie Hong

List of Papers

Paper I: Integrating Model Uncertainty in Probabilistic Decline Curve Analysis for Unconventional Oil Production Forecasting
Hong, A.J., Bratvold, R.B., and Lake, L.W.
To be submitted for publication.

This work was presented at the SPE Workshop: Petroleum Economics—Optimisation versus Growth in Uncertain Times, 27–28 September 2017, Dubai, UAE.

Paper II: Robust Production Optimization with Capacitance-Resistance Model as Precursor
Hong, A.J., Bratvold, R.B., and Nævdal, G.
Published in *Computational Geosciences* (2017), **21** (5): 1423–1442. https://doi.org/10.1007/s10596-017-9666-8

This paper was also in the proceedings of the 15th European Conference on the Mathematics of Oil Recovery (ECMOR XV), 29 August–1 September 2016, Amsterdam, the Netherlands. https://doi.org/10.3997/2214-4609.201601840

Paper III: Fast Analysis of Optimal IOR Start Time Using a Two-Factor Production Model and Least-Squares Monte Carlo Algorithm
Hong, A.J., Bratvold, R.B., and Lake, L.W.
To be submitted for publication.

Paper IV: Robust Discretization of Continuous Probability Distributions for Value-of-Information Analysis
Bratvold, R.B., Thomas, P., Begg, S.H., and Hong, A.J.
Submitted to *Journal of Petroleum Science and Engineering*.

This paper was also in the proceedings of the International Petroleum Technology Conference, 10–12 December 2014, Kuala Lumpur, Malaysia. https://doi.org/10.2523/IPTC-17975-MS

Paper V: Value-of-Information for Model Parameter Updating through History Matching

Hong, A.J., Bratvold, R.B., Thomas, P., and Hanea, R.G.

Submitted to *Journal of Petroleum Science and Engineering*.

Similar content, titled "Value of Information from History Matching—How Much Information is Enough?" was also in the proceedings of the IOR NORWAY 2017—19th European Symposium on Improved Oil Recovery, 24–27 April 2017, Stavanger, Norway. https://doi.org/10.3997/2214-4609.201700327

Contents

Preface ..... **i**
Abstract ..... **ii**
Acknowledgements ..... **v**
List of Papers ..... **vii**
Contents ..... **ix**
List of Figures ..... **xi**
List of Tables ..... **xiii**
Abbreviations ..... **xiv**
Symbols ..... **xvi**
1 Introduction **1**
1.1 Motivation 1
1.2 Research Goals 3
1.3 Thesis Structure 4
2 Managing Geological Uncertainty in History Matching **6**
2.1 History Matching Approaches 6
2.1.1 Deterministic History Matching Approaches 6
2.1.2 Probabilistic History Matching Approaches 9
2.2 Integrating Model Uncertainty in Probabilistic History Matching . 18
2.2.1 Calculating Model Probabilities Using Bayes' Theorem 19
2.2.2 Example of Application—Probabilistic Decline Curve Analysis with Multiple Models 21
3 Managing Geological Uncertainty in Production Optimization **24**
3.1 Robust Production Optimization 24
3.2 Speeding up RO of Waterflooded Production Using CRM 30
3.2.1 Proxy Models vs. Rich Models 30

3.2.2 Proxy-model Workflow for RO 31
3.2.3 Value of Verisimilitude 32
3.2.4 CRM for Waterflooding 33
3.2.5 Example of Applying the Proxy-model Workflow 35
4 Managing Geological Uncertainty in Reservoir Management 42
4.1 Closed Loop Reservoir Management 42
4.2 Fully Structured Reservoir Management 44
4.3 Least-Squares Monte Carlo Algorithm for FSRM 47
4.3.1 LSM Algorithm 47
4.3.2 Example of Applying LSM 48
5 Value-of-Information in Reservoir Management 55
5.1 Definition of VOI 55
5.2 VOI Calculation for Continuous Probability Distributions 58
5.2.1 High Resolution Probability Tree Discretization Method 58
5.2.2 Accuracy of HRPT 59
5.2.3 Comparison of HRPT and MC-based Methods 61
5.3 VOI Calculation for Reservoir Management 62
5.3.1 Relationship between Terms in VOI Analysis and Reservoir Management 62
5.3.2 Decision-tree Example of VOI Calculation for Reservoir Management 64
5.3.3 VOI Calculation Using Ensemble-based Methods for Reservoir Management 67
6 Overview of Research Papers 75
7 Summary and Conclusions 78
8 Discussion and Future Research 81
References 83

List of Figures

Figure 2.1—Moving window approach for approximating measurement error SDs. 11
Figure 2.2—Reservoir simulation model representing the "truth." 15
Figure 2.3—The initial PermH field (in md) of an ensemble member. 16
Figure 2.4—The mean over the updated ensemble of the PermH field. 17
Figure 2.5—Simulation results of the initial ensemble and the updated ensemble. 17
Figure 2.6—Synthetic dataset to day 200. 22
Figure 2.7—Boxplots of cumulative oil production forecasted using solely one model and using the proposed approach given the synthetic data to day 200. 23
Figure 3.1—Illustration of the 2D model and well locations. 28
Figure 3.2—Three realizations of the permeability field for the 2D model. 28
Figure 3.3—Iterative ENPV and cumulative simulation runs. 28
Figure 3.4—Optimal injection scheme for the EnOpt example. 29
Figure 3.5—CDF of NPV under the optimal injection scheme for the EnOpt example. 29
Figure 3.6—Traditional and proxy-model workflows for RO. 31
Figure 3.7—Schematic of (a) CRMT, (b) CRMP, and (c) CRMIP. 34
Figure 3.8—Matching of total fluid production rate and water cut: (a) total fluid production rate from the 2D model, (b) water cut from the 2D model, (c) total fluid production rate from Coupled CRMP, and (d) water cut from Coupled CRMP. 36
Figure 3.9—Validation of total fluid production rate and water cut: (a) total fluid production rate from the 2D model, (b) water cut from the 2D model, (c) total fluid production rate from Coupled CRMP, and (d) water cut from Coupled CRMP. 36
Figure 3.10—CDFs of NPV for the proxy-model workflow example. 38
Figure 3.11—Optimal injection schemes for the proxy-model workflow example. 38
Figure 3.12—Decision tree with the optimal solution of the proxy-model workflow. 40
Figure 3.13—Decision tree with the optimal solution of the traditional workflow. 40

Figure 3.14—CDF of the VOV for the example of applying the proxy-model workflow. 41
Figure 3.15—Sensitivity of EVOV to valve cost. 41
Figure 4.1—Process flow of CLRM, adopted from Jansen et al. (2009)..... 43
Figure 4.2—Decision-tree representation for CLRM. 43
Figure 4.3—Decision-tree representation of FSRM. 44
Figure 4.4—Water cuts of R1 and R2..... 45
Figure 4.5—Decision tree for the example comparing FSRM and CLRM solutions. 46
Figure 4.6—Primary recovery as a function of time for three geological realizations. 50
Figure 4.7—Fully structured decision tree for the example of an optimal IOR start time problem. All NPV values are in million USD..... 51
Figure 5.1—Illustration of HRPT discretized outcomes and MC samples for a continuous probability distribution. 59
Figure 5.2—Average relative VOI error as a function of the number of degrees for the TALL-N problem. 60
Figure 5.3—Relative error of VOI estimates using HRDT, BMC, and EnKF. 62
Figure 5.4—Oil production rate profiles of three realizations. 65
Figure 5.5—Uncertainty trees in (a) assessed form, and (b) inferential form. 65
Figure 5.6—Decision tree for the case with information. 66
Figure 5.7—Decision tree for the case without information..... 66
Figure 5.8—Procedure of VOI calculation using the BvHJ approach. Adapted from Barros et al. (2016). 68
Figure 5.9—Schematic of VOI calculation using MC-based methods. 71
Figure 5.10—PDFs of VOI estimates of BvHJ, F1, and F2 approaches with ensemble size of 50. 73

List of Tables

Table 2.1—Loss function value of MLE for the Arps model, SEM, and Pan CRM for deterministic HM. 22
Table 2.2—Statistics for cumulative oil production forecasted by the Arps model. 22
Table 4.1—Likelihood matrix for the example comparing FSRM and CLRM solutions. 45
Table 4.2—Injection rates for decision nodes solved using FSRM and CLRM. 46
Table 4.3—Likelihood functions for the measured primary recovery efficiency. 50
Table 4.4—Path of measured data for LSM. 52
Table 4.5—NPVs for alternatives at time 2 for LSM. 52
Table 4.6—ENPVs for alternatives at time 2 for LSM. 52
Table 4.7—NPVs for alternatives at time 1 for LSM. 53
Table 4.8—ENPVs for alternatives at time 1 for LSM. 53
Table 4.9—NPVs for alternatives at time 0 for LSM. 54
Table 4.10—Table representation of optimal decision policy solved using LSM. 54
Table 5.1—Relationship between terms used in VOI analysis and those in RM. 63
Table 5.2—Likelihood function for the decision tree example. 65
Table 5.3—VOI estimates using BvHJ, F1, and F2 approaches with ensemble size of 10000. 73
Table 5.4—Statistics of VOI estimates of BvHJ, F1, and F2 approaches with ensemble size of 50. 73
Table 5.5—VOI estimates for the example with a reservoir simulation model. 74

Abbreviations

2D/3D	Two Dimentional / Three Dimensional
BHP	Bottom Hole Pressure
BMC	Bayes Monte Carlo
CDF	Cumulative Density Function
CI	Confidence Interval
CLRM	Closed Loop Reservoir Management
CRM	Capacitance-Resistance Model
CRMIP	Injector-Producer-pair-based CRM
CRMP	Producer-based CRM
CRMT	Single Tank CRM
DA	Decision Analysis
DM	Decision Maker
DWI	Decision with Additional Information
DWOI	Decision without Additional Information
EnKF	Ensemble Kalman Filter
EnOpt	Ensemble-based Optimization
ENPV	Expected NPV
EOR	Enhanced Oil Recovery
EV	Expected Value
EVOV	Expected VOV
EVWI	EV with Additional Information
EVWOI	EV without Additional Information
FOPR	Field Oil Production Rate
FSRM	Fully Structured Reservoir Management
FWCT	Field Water Cut
HM	History Matching
HRPT	High Resolution Probability Tree
HWS	Half Window Size
IOR	Improved Oil Recovery
LSE	Least Squares Estimation
LSM	Least-Squares Monte Carlo
MAP	Maximum a Posteriori Approach
MC	Monte Carlo
MCMC	Markov-Chain Monte Carlo
MCS	Monte Carlo Simulation
MLE	Maximum Likelihood Estimation
NPV	Net Present Value
O&G	Oil and Gas

OOIP	Original Oil in Place
PDF	Probability Density Function
PO	Production Optimization
RM	Reservoir Management
RO	Robust Optimization
SD	Standard Deviation
SEM	Stretched Exponential Model
USD	US Dollar
VOC	Value-of-Clairvoyance
VOI	Value-of-Information
VOPI	Value-of-Perfect-Information
VOV	Value of Verisimilitude
WBHP	Well Bottom Hole Pressure

Symbols

$\boldsymbol{a}$	Alternative
$\boldsymbol{A}$	Decision space
b	Discount factor
$\boldsymbol{C}$	Covariance matrix
D	Reference time for discounting
E	Recovery efficiency
f	Function
$\boldsymbol{g}$	Gradient matrix
I	Water injection rate
J	Objective function value
k	Index of time step
K	Kalman gain
$\boldsymbol{K}$	Kalman gain matrix
l	Index of iteration
L	Loss function value
m	Production model
n_e	Number of realizations (ensemble size)
o	Oil
p	Probability, or probability distribution
P	Price
q	Production rate
Q	Cumulative production
t	Time, or total fluid
T	Total number of time steps
$\boldsymbol{u}$	Control vector (vector of control variables)
v	Value
w	Water
$\boldsymbol{x}$	Vector of model parameters
y	Measured data
$\boldsymbol{y}$	Vector of measured data
$\tilde{y}$	Modeled data
α	Updating step length
ε	Measurement noise/error
$\boldsymbol{\varepsilon}$	Vector of measurement noises/errors
λ	Connectivity
σ	Standard deviation
τ	Time constant

Ω	Prior knowledge
Ω'	Posterior knowledge

1 Introduction

1.1 Motivation

The goal of an oil company is to maximize shareholder or stakeholder value by maximizing the net present value (NPV), which in turn can be maximized by minimizing capital investments and operating expenses while maximizing economic recovery of hydrocarbon from a reservoir through reservoir management (RM). Various definitions of "reservoir management" have been proposed (Haldorsen and Van Golf-Racht 1989, Robertson 1989, Wiggins and Startzman 1990, Satter et al. 1994). These definitions all emphasize that RM is a decision-oriented activity where a decision maker (DM) seeks a production strategy that maximizes the value (commonly quantified by NPV) of hydrocarbon production from a reservoir based on the DM's current knowledge.

A common practice is to use a production model to describe the flow behaviors in a reservoir. Such a model can be a decline curve model or a reservoir simulation model. It is assumed that once the values for the model parameters have been assessed, the model itself will correctly predict future production.[1] Model parameters are updated through history matching (HM),[2] and the optimal production strategy is determined through production optimization (PO) on the history-matched model. Therefore, HM and PO are the two central components of RM.

Begg et al. (2014) provided a formal definition of uncertainty— "Not knowing if a statement (or event), is true or false"—which we will use in this work. Numerous papers have shown the importance of uncertainty quantification in

[1] Nobody really believes that this is the case, but it is viewed as a reasonable assumption that will yield "good enough" results.

[2] In other modeling contexts, the "matching" of models to measured data is usually referred to as model calibration. We will use "history matching" and "model calibration" interchangeably. Originally, HM referred to the adjustment of production model parameters to reproduce the historical production data (rates and pressures) as closely as possible. Today, the term HM is often used in a broader context and includes model calibration using all relevant data and information (seismic data, log data, tracer behavior, etc.).

forecasting. Skov (1995) concluded that the quality of the production forecast can be improved if one can quantify the range of uncertainty and obtain feedback on the accuracy of the forecast. Jonkman et al. (2010) related a company's practices to its economic performance and found that the companies that rigorously took uncertainty into account and planned for multiple possibilities when planning for development, appeared to obtain better production forecasting and decision making, and consequently had higher economic performance. Wolff (2010) noted that it is more meaningful to generate multiple outcomes on a set of models based on uncertainties than to find a single ''true'' answer. Three comparative studies were done for the PUNQ-S3 problem (Floris et al. 2001, Barker et al. 2000, Hajizadeh et al. 2010). In these studies, uncertainty quantification methods were applied to a synthetic model. They illustrated the use of multiple history matched models for uncertainty quantification.

After fast growth in the understanding of uncertainty within the oil and gas (O&G) industry, many authors introduced the tools that can be used to capture both geological[3] and non-geological uncertainties. Clarkson and McGovern (2005) presented a coalbed methane prospecting tool that integrates reservoir simulators with a reservoir model, Monte Carlo simulation, and economic modules, and an infill-well locating tool that can evaluate the locations of an infill-well by combining simulations and economics. When time constraints or a lack of reservoir data make it infeasible to forecast by full-field reservoir simulation, these two tools will be very useful. Jannink and Bos (2005) suggested a fully probabilistic methodology that is able to model both discrete and continuous uncertainties. Their work integrated the pollutant discharge forecast uncertainties and decision making of asset investment, which means that the discharge risk can be converted to financial risk. A model introduced by Morgan (2005) can be used to predict a range of oil price in both the short term and the long term, one of the biggest non-geological uncertainties in O&G economics. His model utilized the oil price data from 1986 to 2003.

Numerous authors have emphasized that it is essential that relevant and material uncertainties (no matter geological or non-geological) need to be represented by multiple realizations or scenarios for decision making. This thesis focuses

[3] For brevity, this thesis uses "geological" to mean "geological and petrophysical."

on managing the geological uncertainty for making good decisions in RM contexts.[4] Non-geological uncertainty has been addressed in Thomas and Bratvold (2015), which accounted for the uncertainties in future oil and gas prices in a gas cap blowdown decision making context.

The O&G industry has long been aware of the importance of uncertainty management, but decision analysis (DA), which leverages consistent probabilistic approaches as a way to manage uncertainties for good decision making, is not commonplace in the industry because of technical and non-technical challenges. This thesis simply uses the phrase "uncertainty management" to refer to "decision-focused uncertainty management," meaning that the intent of uncertainty management is for making good decisions, as reflected in the statement by Bratvold et al. (2009) that quantifying uncertainty has no value in and of itself and value can be created only through our decisions.

Based on our communication with many industry insiders, an example of non-technical challenges is that many people are not comfortable working with probabilities and thus refuse to use a probabilistic approach. Non-technical challenges are outside the scope of this thesis. From the technical challenges arise the following questions, which form the research problems of this thesis:

1. How can geological uncertainty in HM be managed?
2. How can geological uncertainty in PO be managed?
3. How can DA be incorporated with RM?

1.2 Research Goals

Given that value can be created only through the implementation of high-quality decisions, which in turn rely on high-quality uncertainty assessments, we intend to use decision-driven approaches to manage uncertainties for good decision making. Through a literature review, we have identified three main challenges to making high-quality RM decisions:

[4] A good decision is "an action we take that is logically consistent with our objectives, the alternatives we perceive, the information we have, and the preference we feel" (Bratvold and Begg 2010).

1. describing geological uncertainty using probability distributions as a result of HM;
2. computational complexity of modeling flow behaviors and solving for the optimal production strategy when a description of geological uncertainty is given;
3. formulating a production optimization approach that can allow for learning over time.

The goal of this research is to illustrate and discuss how these challenges can be overcome and to facilitate the application of decision quality in RM contexts. To achieve this goal, we will address these challenges through

1. illustrating and discussing the implementation of probabilistic HM approaches, with a focus on how to integrate model uncertainty in probabilistic decline curve analysis for unconventional oil production forecasting;
2. illustrating and discussing the implementation of robust optimization (RO) algorithms and proxy production models, with a focus on how to use a proxy model to speed up the process of robust production optimization;
3. illustrating and discussing the implementation of the fully structured reservoir management (FSRM) approach, with focuses on how to use a two-factor production model and the least-squares Monte Carlo (LSM) algorithm for fast analysis of optimal improved-oil-recovery (IOR) start time and on how to apply the value-of-information (VOI) concept for production model parameter updating through HM.

We are concerned with providing *good approaches* to manage uncertainties for good decision making. A good approach must be both *useful* and *tractable*. By "useful," we mean "providing clear insight for decision making" and by "tractable," we mean that the analysis can be done with the time and resources available.

1.3 Thesis Structure

This thesis is structured in the form of paper collection, with two parts. The first part is an introduction summarizing the research and clarifying how the topics

of the included papers are interrelated. The second part consists of 5 papers which have been published or submitted, or will be submitted for journal publications.

The reminder of the thesis is divided into the following chapters:

- Chapter 2 reviews deterministic and probabilistic HM approaches and introduces an approach to integrate model uncertainty in probabilistic HM using the moving window approach and minimum likelihood estimation (MLE);
- Chapter 3 reviews RO with the ensemble-based optimization (EnOpt) algorithm and proposes a workflow of using a capacitance-resistance model (CRM) to speed up the RO process for waterflooded production;
- Chapter 4 reviews the state-of-the-art reservoir management tool—closed loop reservoir management (CLRM), discusses how it will improve decision making, and compares it with FSRM;
- Chapter 5 reviews the concept of VOI, which can be used to value the information in an RM context, and presents a workflow of assessing it using state-of-the-art HM and PO tools;
- Chapter 6 provides an overview of the research papers included in the second part of the thesis;
- Chapter 7 summarizes and concludes this research;
- Chapter 8 presents a discussion and suggestions for future research.

2 Managing Geological Uncertainty in History Matching

The purpose of a production model (a decline curve model, a reservoir simulation model, etc.) is to predict future production for our decisions on production strategy. Thus, the quality of a production model can affect our choice of the optimal production strategy. The quality of a production model means the similarity of the predicted production to the real production. This in turn relies on the similarity of the geological properties described by the model and its corresponding parameters to the real geological properties of a field. The higher similarity the model has, the higher quality the model has.

A production model is mathematically formulated based on our understanding of the underlying physical principles. Its parameters are based on our knowledge of the geological properties of a field. This knowledge comes from information provided by seismic data, core data, well log data, production data and so on. Our research focuses on the information provided by production data. The parameter settings of a production model are usually determined through an HM process where the model parameters are tuned such that it can closely reproduce historical production. The goal of HM is to correctly forecast production so that the forecast can be used for decision making.

However, the data we can access today represent only a part of the reservoir, and we can never know the subsurface in all its detail. Thus, geological uncertainties always exist in a production model, although HM can reduce the uncertainties. Because of the uncertainties, a more realistic goal of HM is to quantify our lack of knowledge (i.e., to quantify our uncertainty) about the future production, for the purpose of making good decisions.

2.1 History Matching Approaches

2.1.1 Deterministic History Matching Approaches

In a deterministic world, the goal is to identify the true nature of the relevant parameters of a given model. We do this by applying HM, which is an

optimization process where the objective function is a predefined loss function that we seek to minimize by tuning the model parameters. The model with its parameter settings which minimizes the loss function is referred as the best fit model.

2.1.1.1 Least Squares Estimation

An extensively used HM approach is least squares estimation (LSE) where the loss function is defined as the squared difference between model forecast and data:

$$L_{LSE}(\boldsymbol{x}) = \sum_{k=1}^{T} [\tilde{y}_k(\boldsymbol{x}) - y_k]^2, \tag{2.1}$$

where $L_{LSE}(\boldsymbol{x})$ is the loss function of LSE which is a function of the vector of model parameters $\boldsymbol{x}$, $\tilde{y}_k(\boldsymbol{x})$ is the model forecasted production at time step k, y_k is the measured production (i.e., the data) at time step k, and T is the total number of time steps of data.

2.1.1.2 Maximum Likelihood Estimation

Another approach is maximum likelihood estimation (MLE), which aims to maximize the likelihood function (i.e., the probability of observing the data given a set of model parameter settings). Assuming that the data measurements are independent, we can define the likelihood function as

$$p(y_1, y_2, \dots, y_T|\boldsymbol{x}) = \prod_{k=1}^{T} p(y_k|\boldsymbol{x}), \tag{2.2}$$

where p denotes probability, and $p(y_k|\boldsymbol{x})$ denotes the conditional probability of measured data y_k given parameters $\boldsymbol{x}$. Applying a Gaussian random error for the measurement with zero mean and standard deviation (SD) σ_k, we obtain

$$p(y_k|\boldsymbol{x}) = \frac{1}{\sqrt{2\pi\sigma_k^2}} \exp\left(-\frac{[\tilde{y}_k(\boldsymbol{x}) - y_k]^2}{2\sigma_k^2}\right). \tag{2.3}$$

Thus,

$$p(y_1, y_2, \dots, y_T|\boldsymbol{x}) = \frac{1}{\sqrt{(2\pi)^T \prod_{k=1}^{T} \sigma_k^2}} \exp\left(-\frac{1}{2}\sum_{k=1}^{T} \frac{[\tilde{y}_k(\boldsymbol{x}) - y_k]^2}{\sigma_k^2}\right). \quad (2.4)$$

Maximizing the likelihood function is equivalent to minimizing

$$L_{MLE}(\boldsymbol{x}) = \sum_{k=1}^{T} \frac{[\tilde{y}_k(\boldsymbol{x}) - y_k]^2}{\sigma_k^2}. \quad (2.5)$$

If $\sigma_1 = \sigma_2 = \cdots = \sigma_T$, Eq. 2.5 can be reduced to Eq. 2.1. Therefore, LSE is a special case of MLE. The advantage of using Eq. 2.5 instead of Eq. 2.1 is that σ_k acts as a weighting factor so that a more accurate data point (i.e., with smaller σ_k) will receive more weight than a less accurate data point (i.e., with larger σ_k).

2.1.1.3 Maximum A Posteriori

In the maximum a posteriori approach (MAP), we work on a posterior distribution. The posterior (i.e., the probability of a set of model parameter settings given the data) is derived using Bayesian inference that requires assessments of likelihood and prior:

$$p(\boldsymbol{x}|y_1, y_2, \dots, y_T) = \frac{p(y_1, y_2, \dots, y_T|\boldsymbol{x})p(\boldsymbol{x})}{p(y_1, y_2, \dots, y_T)}. \quad (2.6)$$

where $p(\boldsymbol{x})$ is the prior that describes our a priori knowledge about the model parameters and $p(y_1, y_2, \dots, y_T)$ is the preposterior.[5] Because the preposterior is merely a normalizing constant, it can be ignored for maximization. Thus, MAP maximizes $p(y_1, y_2, \dots, y_T|\boldsymbol{x})p(\boldsymbol{x})$. In contrast to MLE (maximizing $p(y_1, y_2, \dots, y_T|\boldsymbol{x})$), MAP includes $p(\boldsymbol{x})$ that can be regarded as a weighting factor for the likelihood. If we have an uninformative prior (i.e., a uniformly

[5] More detailed description of the terms "prior," "likelihood," "posterior," and "preposterior" can be found in **Paper V** and in Raiffa and Schlaifer (1961).

distributed prior), $p(\boldsymbol{x})$ is independent of the value of $\boldsymbol{x}$ and is constant. In such case, maximizing $p(y_1, y_2, \dots, y_T|\boldsymbol{x})p(\boldsymbol{x})$ is equivalent to maximizing $p(y_1, y_2, \dots, y_T|\boldsymbol{x})$. Therefore, MLE is a special case of MAP.

We have shown that LSE, MLE, and MAP are related to each other and all based on Bayes' theorem. Although probabilities are considered, these approaches give a single estimate rather than a distribution, which precludes uncertainties from being included in further analysis.

2.1.2 Probabilistic History Matching Approaches

In HM, uncertainties come from measurement (or observational) errors, non-uniqueness of inverse modeling, and the production model itself. By "measurement error," we mean the difference between a measured value of a quantity and its true value. It describes the inherent variability in the results of a measurement process. By "non-uniqueness of inverse modeling," we mean that multiple good fits to the data can be obtained with various combinations of model parameter settings. The third source of uncertainty (the production model) will be discussed in detail in Section 2.2. For now, the production model is a given. The mentioned uncertainties should be reflected by the model parameters with their distributions.

2.1.2.1 Non-uniqueness in Inverse Modeling

Tavassoli et al. (2004) presented the issue of non-uniqueness in inverse modeling (which they called the inherent uncertainty in HM) and demonstrated that different combinations of model parameter values, which give almost equally good HM results, can give different forecasts. To address this, Sayarpour et al. (2011) started with different sets of initial guesses of model parameters to history match data to generate numerous history matched solutions of model parameters. This approach provides a probability distribution that reflects the non-uniqueness issue, and it is applied in **Paper I**.

2.1.2.2 Bootstrap Method

Cheng et al. (2010) and Jochen and Spivey (1996) used the bootstrap method to incorporate measurement errors in HM. The bootstrap method is a type of

Monte Carlo (MC) method for approximate Bayesian inference (Hastie et al. 2009) and assumes non-informative prior. The bootstrap method is applied through the following procedure:

1) Generate a sampled dataset by resampling data points from the original dataset with replacement;
2) Use LSE to history match the sampled dataset to obtain a set of model parameter settings;
3) Repeat Steps 1 and 2 numerous times to obtain numerous sets of model parameter values;
4) Use these sets of model parameter values as MC samples representing the distributions of the model parameters.

The bootstrap method does not require that explicit distributions be assigned to the measurement errors because they are approximated by the perturbed data (Hastie et al. 2009).

2.1.2.3 Probabilistic MLE Approach

Considering that a noisier data subset should be given less weight than a less noisy data subset, and that a data subset with sharp charges is less reliable and thus should be given less weight than a data subset with gradual changes, we propose an approach named the probabilistic MLE approach (P-MLE). A practical issue in using MLE is that distributions must be assigned to the measurement errors, which are seldom known a priori. To approximately estimate the measurement error distributions, we assume that measurement errors are independent and normally distributed, and we apply a moving window approach to estimate their SDs. **Figure 2.1** illustrates the moving window approach. For example, we want to estimate the measurement error SD at time k. If we use a half window size (HWS) of 2, we take the two previous data points (the data points at times $k-2$ and $k-1$), the two next data points (the data points at times $k+1$ and $k+2$), and the data at time k to form a data subset. The SD of this data subset is used as the measurement error SD at time k. A large number of sampled datasets can be generated by sampling from a normal distribution with the original data point as its mean, and the SD estimated using the moving window approach for each data point. MLE is then

used to history match each sampled dataset, yielding many sets of model parameter settings. This proposed approach is applied in **Paper I**.

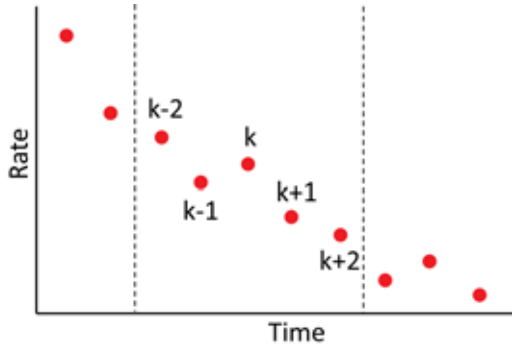


Figure 2.1—Moving window approach for approximating measurement error SDs.

2.1.2.4 Ensemble Kalman Filter

When a priori knowledge is considered, a fully Bayesian approach should be used. Approaches for Bayesian HM include the ensemble Kalman filter (EnKF) and Markov-chain Monte Carlo (MCMC). The present work applies EnKF because it is easily implemented with any production model, it is computationally attractive, and it can handle large datasets and a considerable number of model parameters. The first use of EnKF in petroleum engineering is probably that described by Lorentzen et al. (2001). Following its publication, this method attracted significant attention, and the number of papers applying and discussing EnKF increased rapidly.

The following briefly reviews EnKF. For a comprehensive introduction to EnKF, see Evensen (2009). Aanonsen et al. (2009) provided an extensive review of the application of EnKF in reservoir engineering.

2.1.2.4.1 EnKF Formulation

The basic idea behind EnKF is based on the Kalman (1960) filter. The Kalman filter is a solution to estimating the state[6] of a process in a recursive way and is

[6] Given a dynamic system, a state is an unobservable (in most cases) quantity that is required, and thus must be determined, to predict the system behavior (i.e., the future state). A state can be a scalar or a vector.

designed for linear filtering and prediction problems with discrete data. This method seeks an optimal weight between measured data and model forecast. Normal distribution is used to describe an uncertain quantity, say the production rate q, at a certain time: the forecasted q by a model is normally distributed $N(\bar{q}_{model}, \sigma_{model})$, where $\bar{q}_{model}$ is the model forecasted mean and σ_{model} is the SD of model forecast, and the measured q by a device is normally distributed $N(\bar{q}_{meas}, \sigma_{meas})$, where $\bar{q}_{meas}$ is the measured mean and σ_{meas} is the SD of measurement error. A linear combination of the model forecast and measured data can be constructed as

$$q_{update} = (1 - K) \cdot q_{model} + K \cdot q_{meas}, \tag{2.7}$$

where q_{update} is the updated estimate of q given measured data, and K (the Kalman gain) is a weighting factor for the model forecast versus the measured data. q_{update} is also normally distributed $N(\bar{q}_{update}, \sigma_{update})$, where $\bar{q}_{update}$ is the updated mean and σ_{update} is the updated SD. Given Eq. 2.7, $\bar{q}_{update}$ and σ_{update} can be calculated, respectively,

$$\bar{q}_{update} = (1 - K) \cdot \bar{q}_{model} + K \cdot \bar{q}_{meas}, \tag{2.8}$$

$$\sigma^2_{update} = (1 - K)^2 \cdot \sigma^2_{model} + K^2 \cdot \sigma^2_{meas}. \tag{2.9}$$

If we consider the best estimate as the one with the smallest SD, σ_{update} is minimized by adjusting K, yielding

$$\arg\min_K \sigma_{update} = K^* = \frac{\sigma^2_{model}}{\sigma^2_{model} + \sigma^2_{meas}}. \tag{2.10}$$

Replacing K in Eqs. 2.8 and 2.9 with K^*, we obtain the mean and SD of a normal distribution for the updated estimate.

Another way to obtain Eqs. 2.8, 2.9, and 2.10 is to start with Bayes' theorem:

$$\begin{aligned} p(q_{update}) &= p(q_{model}|q_{meas}) \\ &= \frac{p(q_{model})p(q_{meas}|q_{model})}{p(q_{meas})}. \end{aligned} \tag{2.11}$$

Using the properties of conjugate normal distributions, we can obtain

$$\bar{q}_{update} = \left(1 - \frac{\sigma^2_{model}}{\sigma^2_{model} + \sigma^2_{meas}}\right) \cdot \bar{q}_{model} + \frac{\sigma^2_{model} \cdot \bar{q}_{meas}}{\sigma^2_{model} + \sigma^2_{meas}} \quad (2.12)$$

and

$$\sigma^2_{update} = \frac{1}{\frac{1}{\sigma^2_{model}} + \frac{1}{\sigma^2_{meas}}}. \quad (2.13)$$

Eq. 2.12 and Eq. 2.13 can be rewritten as Eqs. 2.8 and 2.9, respectively. Using the Kalman filter, Bayesian inference is implicitly conducted.

Evensen (1994) introduced EnKF as a MC representation[7] of the Kalman filter. The prior distribution is represented by a set of realizations of model parameters (i.e., the prior or initial ensemble). The posterior distribution is represented by the updated realizations of model parameters (i.e., the posterior or updated ensemble), which are obtained by (Burgers et al. 1998)

$$\boldsymbol{Y}_a = \boldsymbol{Y}_f + \boldsymbol{K}(\boldsymbol{y}_p - \boldsymbol{H}\boldsymbol{Y}_f), \quad (2.14)$$

where matrix $\boldsymbol{Y}_a$ consists of the vectors containing the updated states, updated model parameters, and updated observations corresponding to each realization in the posterior ensemble; matrix $\boldsymbol{Y}_f$ contains the predicted states by forward modeling, model parameters, and predicted observations by forward modelling corresponding to each realization in the prior ensemble; $\boldsymbol{K}$ is the Kalman gain matrix, which weighs the influences of the prior predicted observations and the real-time observations (i.e., measured data); $\boldsymbol{y}_p$ is a matrix containing the perturbed observations;[8] and $\boldsymbol{H}$ is an operator that links $\boldsymbol{Y}_f$ to the predicted observations. As does Eq. 2.7, Eq. 2.14 describes a linear combination of the prior and the observations. The weighing factor $\boldsymbol{K}$ is calculated as

$$\boldsymbol{K} = \boldsymbol{C}_f\boldsymbol{H}^T(\boldsymbol{H}\boldsymbol{C}_f\boldsymbol{H}^T + \boldsymbol{C}_y)^{-1}, \quad (2.15)$$

[7] By "MC representation" we mean that a probability distribution is represented by its corresponding MC samples.

[8] When the EnKF is applied, an observation has to be perturbed with its corresponding statistics in order to avoid insufficient variance (Burgers et al. 1998).

where $\boldsymbol{C}_f$ is the covariance matrix of $\boldsymbol{Y}_f$ encoding the covariance matrix of the prior predicted observations, and $\boldsymbol{C}_y$ is the covariance matrix of the observations. As the measurements become noisier (i.e., the variance of an observation increases) or the variance of a prior predicted observation decreases, more weight is given to the prior; otherwise, more weight is given to the observations.

EnKF embodies the prior in $\boldsymbol{Y}_f$, the likelihood in $\boldsymbol{C}_y$ (when the model noise is ignored), and the posterior in $\boldsymbol{Y}_a$; and the preposterior is a normalizing constant of the posterior. Thus, Bayes' rule describing the relationship among the prior, the likelihood, the preposterior, and the posterior is no longer shown explicitly as in Eq. 2.11, but is implicit in Eqs. 2.14 and 2.15. Using EnKF in the context of HM requires the initial guess of the model parameters (i.e., the prior ensemble) together with a model that can predict both the production and observations given a production strategy, observations, and their associated statistics; and the result is the EnKF updated model parameters (i.e., the posterior ensemble).

To perform Bayesian inference using EnKF requires only a few matrix operations. Thus, it is very useful for updating a reservoir simulation model which usually has thousands of parameters for determination. EnKF is applied in **Paper V**.

2.1.2.4.2 Example of EnKF Application

Here is an example to illustrate the application of EnKF. A 2D, horizontal, two-phase (water and black oil) reservoir simulation model is built to represent the "truth." A horizontal injector penetrates along the left edge of the reservoir, and a horizontal producer along the right edge. For simplicity, the horizontal injector and the horizontal producer are represented by 10 vertical dummy injectors and 10 vertical dummy producers, respectively. The fluid injection rate is fixed at 15 sm^3/day in each dummy injector. The production rate in each dummy producer is constrained by an upper bound of 15 sm^3/day. Two highly permeable regions, with horizontal permeability of 1500 millidarcy (md), are used to mimic the water channels; and the horizontal permeability in the background is 250 md. The geometry of the model, the horizontal permeability field, and the dummy vertical wells are shown in **Figure 2.2**. The grid blocks

are cubic with length, width, and height equal to 10 m. The reservoir model is 450 m × 450 m × 10 m which scales to 45 blocks × 45 blocks × 1 block.

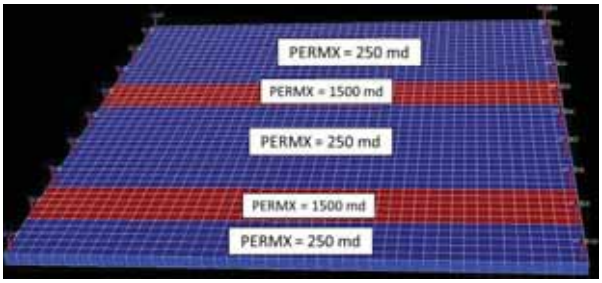


Figure 2.2—Reservoir simulation model representing the "truth."

The unknown is the horizontal permeability (PermH) at each grid block, which means 2025 parameters for the entire grid. The base 10 logarithms of the permeabilities are used as the model parameters in EnKF because grid block permeability is usually assumed to be lognormally distributed and the use of logarithm can prevent the updated permeability calculated by EnKF from being negative. The state variables are the grid block pressures and the grid block water saturations. The true observations are the well bottom hole pressures (WBHPs) and oil production rates in all the dummy wells, obtained by running simulation on the true model. The predicted observations (WBHPs and well oil production rates) are obtained by running simulations with the updated model and state variables. Observations are recorded every 50 simulated days over a 1000-day production period, totaling 20 data points. EnKF is implemented in MATLAB™ (2014), and the production model is an ECLIPSE™ (2014) simulation model, so an interface between MATLAB and ECLIPSE is constructed: the updated permeabilities calculated in MATLAB are written into the input files (the .DATA-files) of ECLIPSE, the forecasted and updated pressures and saturations are read from and written into the restart files (the X-files) of ECLIPSE, respectively, the predicted observations are read from the summary files (the S-files) of ECLIPSE, and the ECLIPSE simulator can be called to run in MATLAB. The standard EnKF algorithm (Eqs. 2.14 and 2.15) is used to calculate the updated model variables (PermH) and state variables (pressures and saturations).

Initial ensemble members of the PermH field are generated by sampling from a predefined multivariate normal distribution that describes our a priori

knowledge about the PermH field. One of the initial ensemble members is shown in **Figure 2.3**.

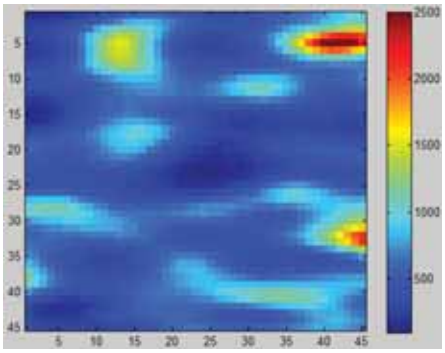


Figure 2.3—The initial PermH field (in md) of an ensemble member.

To avoid physically impossible PermH values, we truncate the permeability values so that they are limited to a range (from 100 to 2500 md in this example). We do the same for the updated saturations and pressures, such that the saturations are always between 0 and 1 and the pressures are always positive.

Observations are perturbed by adding an error term drawn from a predefined normal distribution that describes our a priori knowledge about the measurement errors. In this example, we assign relatively small errors (a SD of 1 bar for all the pressure measurements and a SD of 1 sm^3/day for all the oil rate measurements). In a more realistic case, the measurement errors can be much larger.

The mean over the updated ensemble of the PermH field is illustrated in **Figure 2.4**. Two high permeable channels similar to those illustrated in Figure 2.2 appear after applying the EnKF. The simulated results, WBHP for Producer 5 and Injector 5, field water cut (FWCT), and field oil production rate (FOPR), are shown in **Figure 2.5**. The grey lines represent the results generated using the initial ensemble. The cyan lines represent the results generated using the updated ensemble members. The black dashed line represents the result generated using the mean over the updated ensemble as shown in Figure 2.4. The red line represents the result generated using the true model. The measured WBHP from Producer 5, Injector 5, and measured FOPR are the red circles in Figure 2.5(a), (b), and (d), respectively. The results show that the use of EnKF not only reduces the spread of the initial ensemble (i.e., reduces the uncertainty)

but produces a good match to the measured data. A significant impact of using EnKF can be seen in Figure 2.5(c): none of the initial ensemble members predicts the water breakthrough time correctly, whereas the updated ensemble provides a nearly perfect match.

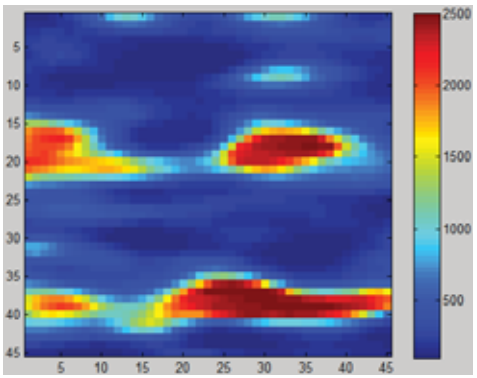
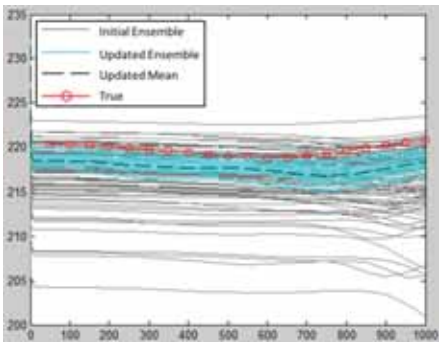
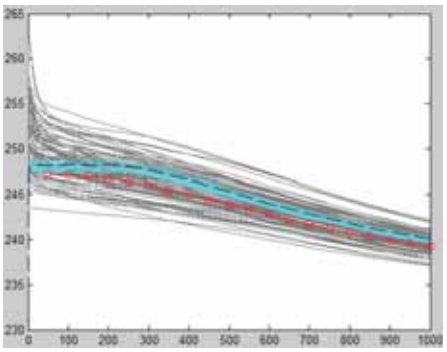


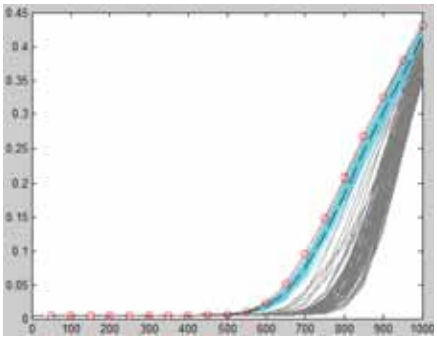
Figure 2.4—The mean over the updated ensemble of the PermH field.



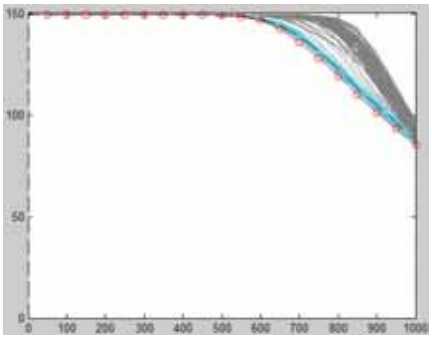
(a) Producer 5: WBHP [bars] vs. Time [days]



(b) Injector 5: WBHP [bars] vs. Time [days]



(c) FWCT [-] vs. Time [days]



(d) FOPR [sm^3/day] vs. Time [days]

Figure 2.5—Simulation results of the initial ensemble and the updated ensemble.

The updated results (the cyan lines) do not capture the truth (the red line). This does not mean that based on the updated results, the probability of the truth is zero. Recall that the EnKF uses MC samples to represent a probability distribution. The values on the cyan lines at a certain time are MC samples representing a Gaussian distribution. This underlying Gaussian distribution captures the truth.

2.2 Integrating Model Uncertainty in Probabilistic History Matching

The probabilistic HM approaches presented earlier incorporate measurement errors and non-uniqueness of inverse modeling and produce multiple realizations of model parameters to reflect these uncertainties. However, all of them assume a mathematical model that can "correctly" predict reservoir performance. The uncertainty is only in the model parameters and not in the model. This, of course, is never the case because "all models are wrong but some are useful" (Box 1979).

In this section, the context is the use of decline curve analysis for unconventional plays. Although numerical techniques for forecasting hydrocarbon production have developed rapidly over the past decades, the decline curve analysis technique is still used extensively in the O&G industry. A decline curve model is computationally attractive, and only production data that can be easily acquired is required for determining the model parameter values through the HM process. The history matched model is further used for forecasting hydrocarbon production and reserves.

One of the most extensively used decline curve models is the Arps (1945) model. However, the Arps model is often not ideal for unconventional plays (Joshi and Lee 2013). Several decline curve models have been developed in order to capture the flow behaviors in an unconventional well; for example, the power law exponential model (Ilk et al. 2008), the stretched exponential model (SEM) (Valko and Lee 2010), the Duong (2011) model, the logistic growth model (Clark et al. 2011), and the Pan (2016) CRM.

Given these models, a question arises that has not been discussed widely: which is the best model? This question seems trite because the meaning of "best" is

not well defined. In traditional practices, the model that can best fit the data in a least squares sense is regarded as the best model. However, this ignores two facts: the best fit model might not be the model that best describes the flow behaviors, and there might be several models that fit the data almost equally well.

2.2.1 Calculating Model Probabilities Using Bayes' Theorem

Instead of identifying the "best" decline curve model for unconventional plays, we propose to regard any model as potentially good, where goodness is characterized by a probability. If a model has higher probability than the others, we say that this model is more likely to be a good model. If several models have close probabilities, we say that the model uncertainty is high because it is difficult to tell which model is more likely to be a good model.

In this manner, the model uncertainty can be easily integrated in probabilistic decline curve analysis, using the probabilistic HM approaches presented in Section 2.1.2. However, the P-MLE approach (Section 2.1.2.3) will be used here for the following reasons:

- Data points should be weighted differently according to their noisiness, and distributions of measurement errors need to be assigned for calculating the model probabilities; thus, the Bootstrap method is not used.
- The parameters of different models have different physical meanings and are related to the reservoir property. It is not easy to assign distributions to the parameters of different models to represent the same reservoir property based on our a priori knowledge. To compare different models fairly, it would be better to assume an uninformative prior. Therefore, a fully Bayesian approach like the EnKF, which requires an informative prior, is not used.

Consider a given dataset $\boldsymbol{y} = [q_1, q_2, \dots, q_k, \dots, q_T]$ where q_k is the oil production rate measured at time k. The procedure to quantify the model uncertainty is

1) Estimate σ_k, which is the SD of the normal random error of the measurement of q_k, using the moving window approach.
2) Draw an MC sample from a normal distribution with mean q_k and SD σ_k, and repeat it for each data point to obtain a sampled dataset.
3) For a given sample dataset, use MLE to determine the parameters of each model under consideration.
4) Calculate the probability of each model with its MLE parameters, given the sampled dataset.
5) Repeat Steps 2–4 N times, where N is the total number of MC iterations (we use $N = 1000$).
6) Calculate the posterior probability of each model.

In Step 4, the probability of each model with its MLE parameters, given the sampled dataset, is calculated using Bayes' theorem:

$$p\left(\boldsymbol{x}_{ij}^{MLE}, m_i \middle| \boldsymbol{y}_j, \Omega'\right) = \frac{p\left(\boldsymbol{y}_j \middle| \boldsymbol{x}_{ij}^{MLE}, m_i\right) p\left(\boldsymbol{x}_{ij}^{MLE}, m_i \middle| \Omega\right)}{\sum_{i'} p\left(\boldsymbol{y}_j \middle| \boldsymbol{x}_{i'j}^{MLE}, m_{i'}\right) p\left(\boldsymbol{x}_{i'j}^{MLE}, m_{i'} \middle| \Omega\right)}, \quad (2.16)$$

where $\boldsymbol{x}_{ij}^{MLE}$ denotes the MLE parameters of model m_i given sampled dataset $\boldsymbol{y}_j$; Ω denotes a priori knowledge; and Ω' denotes a posteriori knowledge given $\boldsymbol{y}$. Applying a noninformative prior, we have $p\left(\boldsymbol{x}_{1j}^{MLE}, m_1 \middle| \Omega\right) = p\left(\boldsymbol{x}_{2j}^{MLE}, m_2 \middle| \Omega\right) = \cdots$, reducing Eq. 2.16 to

$$p\left(\boldsymbol{x}_{ij}^{MLE}, m_i \middle| \boldsymbol{y}_j, \Omega'\right) = \frac{p\left(\boldsymbol{y}_j \middle| \boldsymbol{x}_{ij}^{MLE}, m_i\right)}{\sum_{i'} p\left(\boldsymbol{y}_j \middle| \boldsymbol{x}_{i'j}^{MLE}, m_{i'}\right)}. \quad (2.17)$$

Inserting Eqs. 2.4 and 2.5 (MLE loss function) into Eq. 2.17, we obtain

$$\begin{aligned} p\left(\boldsymbol{x}_{ij}^{MLE}, m_i \middle| \boldsymbol{y}_j, \Omega'\right) &= \frac{\exp\left(-\frac{1}{2} L_{MLE}\left(\boldsymbol{x}_{ij}^{MLE}\right)\right)}{\sum_{i'} \exp\left(-\frac{1}{2} L_{MLE}\left(\boldsymbol{x}_{i'j}^{MLE}\right)\right)} \\ &= 1 \Big/ \sum_{i'} \exp\left(-\frac{1}{2}\left[L_{MLE}\left(\boldsymbol{x}_{i'j}^{MLE}\right) - L_{MLE}\left(\boldsymbol{x}_{ij}^{MLE}\right)\right]\right). \end{aligned} \quad (2.18)$$

Because $\boldsymbol{y}_j$ is a MC sample, $p\left(\boldsymbol{y}_j\right) = 1/N$ and

$$p(\boldsymbol{x}_{ij}^{MLE}, m_i, \boldsymbol{y}_j|\Omega') = p(\boldsymbol{x}_{ij}^{MLE}, m_i|\boldsymbol{y}_j, \Omega')p(\boldsymbol{y}_j)$$
$$= 1\Big/\left\{N\sum_{i'}\exp(-\frac{1}{2}[L_{MLE}\left(\boldsymbol{x}_{i'j}^{MLE}\right) - L_{MLE}(\boldsymbol{x}_{ij}^{MLE})])\right\}. \quad (2.19)$$

The probability calculated using Eq. 2.19 is the probability of the estimate of interest (e.g., reserves) forecasted by model m_i with its parameters $\boldsymbol{x}_{ij}^{MLE}$. By repeating this process over different models and different sampled datasets, we obtain the distribution of interest, which includes uncertainties in measurements, inverse modelling and models. The posterior probability of a model in Step 6 is calculated as

$$p(m_i|\Omega') = \sum_{j=1}^{N} p(\boldsymbol{x}_{ij}^{MLE}, m_i, \boldsymbol{y}_j|\Omega'). \quad (2.20)$$

2.2.2 Example of Application—Probabilistic Decline Curve Analysis with Multiple Models

We use an example to illustrate the impacts of using a deterministic approach or probabilistic approach without considering the model uncertainty on decline curve analysis, and to highlight the importance of integrating the model uncertainty. This example considers only three models—the Arps model, SEM, and Pan CRM—but other models can be easily added.

The "true" decline is generated using the Pan CRM. Random errors are added to the "true" decline to form the synthetic dataset. **Figure 2.6** illustrates the synthetic dataset and the "true" decline of oil production rate as well as the SDs of measurement errors assessed using the moving window approach with an HWS of 10 data points. Our interest is the cumulative oil production from day 200 to day 10950 (year 30). The "true" cumulative oil production given by the Pan CRM is 48.1 Mbbl. This value is used as a reference for the estimates.

We first use a deterministic approach to estimate the cumulative oil production. The resulting loss function value of MLE for each model is listed in **Table 2.1**. Because the Arps model best fits the data (minimum loss function value), it is used for forecasting, providing an estimated cumulative oil production of 132.6 Mbbl. This estimate is more than 2.5 times the "true" value.

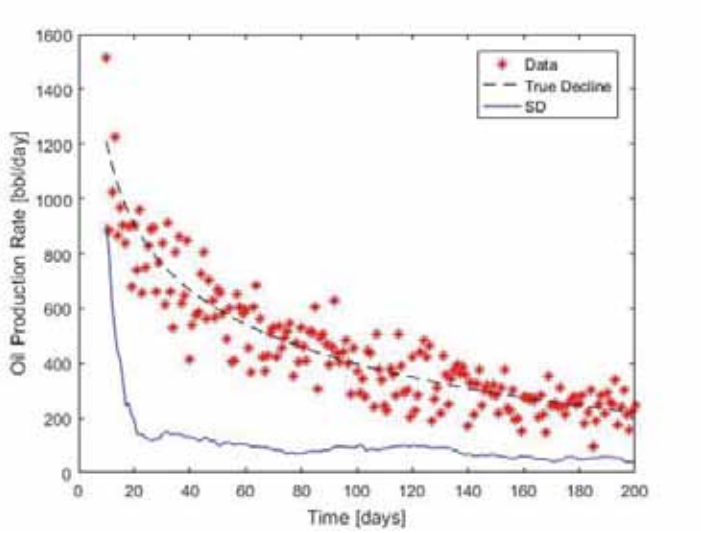


Figure 2.6—Synthetic dataset to day 200.

Arps Model	SEM	Pan CRM
171.65	171.75	172.29

Table 2.1—Loss function value of MLE for the Arps model, SEM, and Pan CRM for deterministic HM.

To produce a distribution of the cumulative oil production, we use the P-MLE approach with only the Arps model. The P10, P50, mean, and P90 values are listed in **Table 2.2**. The forecast of the Arps model is biased far from the "truth" (48.1 Mbbl), and the 80% confidence interval (CI), from P10 to P90, does not cover the "truth."

Statistic	P10	P50	Mean	P90
Cumulative Oil Production [Mbbl]	73.5	132.3	135.7	207.2

Table 2.2—Statistics for cumulative oil production forecasted by the Arps model.

We now integrate the model uncertainty in the P-MLE approach for this analysis. Table 2.1 shows that the minimized loss function values of the three models are indeed very close. Using Eq. 2.18 to convert these loss function values to probabilities, the posterior probabilities of Arps model, SEM and Pan CRM are 36.4%, 34.0%, and 29.6%, respectively. This suggests that the models are almost equally likely to be good, given the dataset. This result seems counter-intuitive as the correct model (the Pan CRM) is the least likely one.

This is because the dataset does not provide enough information to identify the correct model. Thus, the model uncertainty remains large even when the dataset is given. **Figure 2.7** shows the boxplots of cumulative oil production forecasted using solely the Arps model, SEM, or Pan CRM, and using our proposed approach. Among the three models, the Arps model has the largest uncertainty in inverse modelling, as it gives the largest 80% CI, whereas the Pan CRM has the smallest uncertainty in inverse modelling. Since the Pan CRM is the correct model, its estimate is the best. Using our proposed approach, the "truth" is covered in the 80% CI. This means that by integrating the model uncertainty in the analysis, we can reduce the risk of selecting a poor model for forecasting.

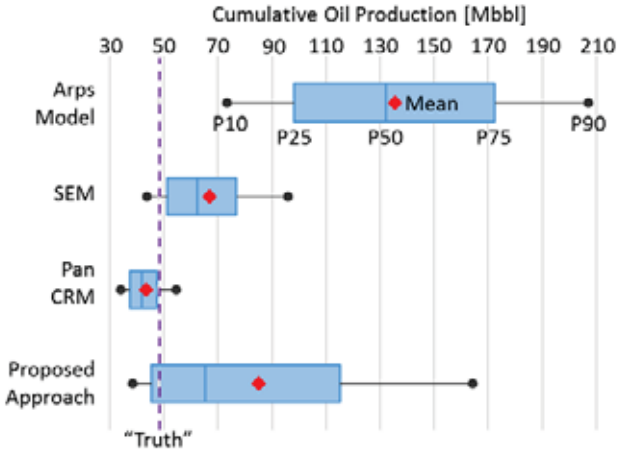


Figure 2.7—Boxplots of cumulative oil production forecasted using solely one model and using the proposed approach given the synthetic data to day 200.

The application of the proposed approaches for probabilistic decline curve analysis for unconventional wells in two fields is illustrated and discussed in **Paper I**.

3 Managing Geological Uncertainty in Production Optimization

PO is a powerful approach to design a production strategy. The goal of PO is to maximize hydrocarbon production by adjusting the control (or decision) variables (e.g., water injection rates and producer WBHPs). However, this can be expressed in different ways, such as to minimize the water cut of a producer, to optimize the WBHP of an injector, or to maximize the economic value over the life of a reservoir. This work defines the objective of PO as, to maximize NPV. The context in this work is waterflooding, and we assume that all revenues stem from oil production and that all costs are induced by water injection and water production. Thus, the objective function for a deterministic case, where a single realization is considered, is defined as

$$J(\boldsymbol{u}) = NPV(\boldsymbol{u}) = \sum_{k=1}^{n_T} \frac{[q_o^k(\boldsymbol{u})P_o - q_w^k(\boldsymbol{u})P_{wp} - I^k(\boldsymbol{u})P_{wi}]\Delta t_k}{(1+b)^{t_k/D}}, \quad (3.1)$$

where $\boldsymbol{u}$ is the control vector (i.e., a vector of control variables) defined as $\boldsymbol{u} = [\boldsymbol{u}_1, \boldsymbol{u}_1, \dots, \boldsymbol{u}_N]^{\mathrm{T}}$, where N is the number of control variables; q_o^k is the field oil production rate at time k; q_w^k is the field water production rate at time k; I^k is the field water injection rate at time k; P_o is the oil price; P_{wp} is the water production cost; and P_{wi} is the water injection cost; b is the discount factor; t_k is the cumulative time for discounting; and D is the reference time for discounting (D = 365 days if b is expressed as a fraction per year and the cash flow is discounted daily). q_o^k, q_w^k, and I^k are forecasted by a given production model.

3.1 Robust Production Optimization

RO of production is performed over an ensemble of realizations representing the geological uncertainty. For a risk-neutral decision maker, the objective of RO is to maximize the expected value (EV) over the ensemble,

$$\bar{J}(\boldsymbol{u}) = \frac{\sum_{r=1}^{n_e} J_r}{n_e}, \tag{3.2}$$

where $\bar{J}$ is the EV over all realizations (i.e., the objective function for a probabilistic case), J_r is the objective function (Eq. 3.1) for a deterministic case with a single realization r, and n_e is the number of realizations (i.e., ensemble size).

Maximizing Eq. 3.2 can be done iteratively using the steepest ascent method. For each iteration, the control vector is updated as

$$\widehat{\boldsymbol{u}}_{l+1} = \widehat{\boldsymbol{u}}_l + \alpha_l \frac{\boldsymbol{g}_l}{\|\boldsymbol{g}_l\|_\infty}, \tag{3.3}$$

where the subscript l denotes the iteration number, $\widehat{\boldsymbol{u}}_{l+1}$ is the currently updated control vector, $\widehat{\boldsymbol{u}}_l$ is the previously updated control vector, α is the step length for updating, and $\boldsymbol{g}$ is the gradient for updating the control variables. Several methods have been proposed to estimate $\boldsymbol{g}$ for each iteration. Nævdal et al. (2006) used an adjoint method to calculate $\boldsymbol{g}$. However, deriving the adjoint equation requires access to the mathematical formulation of a production model, which is impossible for commercial reservoir simulators, such as ECLIPSE that is used in this work. Chen et al. (2009) proposed the ensemble-based optimization method (EnOpt), to calculate $\boldsymbol{g}$. The EnOpt treats a production model as a black box (i.e., no requirement for its mathematical formulation) and thus can be easily implemented with any type of production model. Besides, EnOpt is specially designed for the case where an ensemble of realizations is involved, so it can significantly reduce the number of simulations required for gradient calculation. Thus, we use EnOpt for this work, as discussed in **Papers II** and **V**.

In the original EnOpt proposed by Chen et al. (2009), the ensemble of values for control vector $\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_M$, where M is the ensemble size,[9] is generated from a multivariate normal distribution with predefined mean $\widehat{\boldsymbol{u}}$ and predefined covariance matrix $\boldsymbol{C}_{\boldsymbol{u}}$, which is used to specify the temporal correlation of the

[9] $M = n_e$ if each realization is coupled with one sample of values for the control vector (i.e., a 1:1 ratio is applied), as in this work.

controls to limit the frequency of changes to the controls. The predefined mean $\hat{\boldsymbol{u}}$ can be approximated by its sample mean $\bar{\boldsymbol{u}}$, i.e.,

$$\hat{\boldsymbol{u}} \approx \bar{\boldsymbol{u}} = \frac{\sum_{r=1}^{M} \boldsymbol{u}_r}{M}, \tag{3.4}$$

and the average objective value $\bar{J}(\hat{\boldsymbol{u}})$ can be approximated by

$$\bar{J}(\hat{\boldsymbol{u}}) = \frac{\sum_{r=1}^{M} J_r(\hat{\boldsymbol{u}})}{M} \approx \frac{\sum_{r=1}^{M} J_r(\boldsymbol{u}_r)}{M}. \tag{3.5}$$

Chen et al. (2009) used the approximation in Eq. 3.4 to calculate the mean-shifted ensemble matrix $\bar{\boldsymbol{U}}$ and the approximation in Eq. 3.5 to calculate the mean-shifted objective function vector $\boldsymbol{j}$. However, Do and Reynolds (2013) did not find any advantage, theoretical or practical, in approximating these two terms. Hence, they calculate the mean-shifted ensemble matrix directly using the predefined mean $\hat{\boldsymbol{u}}$ by

$$\bar{\boldsymbol{U}} = [\boldsymbol{u}_1 - \hat{\boldsymbol{u}}, \quad \boldsymbol{u}_2 - \hat{\boldsymbol{u}}, \quad \dots, \quad \boldsymbol{u}_M - \hat{\boldsymbol{u}}]^{\mathbf{T}}. \tag{3.6}$$

Moreover, Fonseca et al. (2014) suggested calculating the mean-shifted objective function vector $\boldsymbol{j}$ with respect to the objective value of the predefined mean for each individual realization instead of the average objective value (Eq. 3.5), i.e.,

$$\begin{aligned} \boldsymbol{j} \\ = [J_1(\boldsymbol{u}_1) - J_1(\hat{\boldsymbol{u}}), \quad J_2(\boldsymbol{u}_2) - J_2(\hat{\boldsymbol{u}}), \quad \dots, \quad J_M(\boldsymbol{u}_M) - J_M(\hat{\boldsymbol{u}})]^{\mathbf{T}}. \end{aligned} \tag{3.7}$$

The cross-covariance matrix is then

$$\boldsymbol{C}_{\boldsymbol{uJ}} = \frac{1}{M}\left(\bar{\boldsymbol{U}}^{\mathrm{T}}\boldsymbol{j}\right). \tag{3.8}$$

Chen et al. (2009) approximated the gradient by

$$\boldsymbol{g} \approx \boldsymbol{C}_{\boldsymbol{uJ}}. \tag{3.9}$$

Eqs. 3.6–3.9 and 3.3 constitute the modified EnOpt formulation. The original EnOpt formulation can be obtained by replacing $\hat{\boldsymbol{u}}$ with $\bar{\boldsymbol{u}}$ in Eq. 3.6 and $J_r(\hat{\boldsymbol{u}})$ with the approximated $\bar{J}(\hat{\boldsymbol{u}})$ in Eq. 3.7. Fonseca et al. (2014) showed that the

modified EnOpt converges to a higher objective value and more quickly than does the original EnOpt. Therefore, we use the modified EnOpt.

The step length α significantly affects the rate of convergence. A naive line search procedure is to simply reduce α by half if $\bar{J}(\hat{\boldsymbol{u}}_{l+1}) < \bar{J}(\hat{\boldsymbol{u}}_l)$. The present work uses an interpolation-based line search procedure to find a relatively large value of α that satisfies the Armijo condition (Wright and Nocedal 1999). The line search procedure halts when $\bar{J}(\hat{\boldsymbol{u}}_{l+1}) > \bar{J}(\hat{\boldsymbol{u}}_l)$ or a maximum number of iterations is reached, whichever comes first.

The stopping criterion of the iteration of the steepest ascent method is usually that the increase in the objective function value is smaller than a user-defined threshold value or a maximum number of iterations is reached, whichever comes first. In our use of EnOpt, an optimal solution is usually found during the first 5–15 iterations. In order to observe clear evidence of convergence, our stopping criterion is that the number of iterations has reached 30, which is about twice as many as have been required based on our experience.

We use a synthetic 2D reservoir model (the 2D model) to illustrate the application of the EnOpt. The 2D Model is an isotropic heterogeneous model with four injectors and one producer in a five-spot pattern. The heterogeneity pertains to permeability only. **Figure 3.1** shows the geometry of the 2D model and well locations. The color map in Figure 3.1 shows a realization of the permeability field in md. The producer bottom hole pressures (BHPs) are fixed at 200 bars. The well injection rate is controllable and constrained from 0 to 200 m^3/day. The control variables are injection rates for 50 time intervals of 30 days each (i.e., a life-cycle of 1500 days). For simplicity, the injection rates of all injectors for a given day are assumed to be identical. Thus, there are 50 control variables. To incorporate permeability uncertainty, we generate 100 realizations of the permeability field from a multivariate normal distribution. Three of these realizations are shown in **Figure 3.2**. NPV calculation uses the oil price 315 \$/m^3, water production cost 47.5 \$/m^3, water injection cost 12.5 \$/m^3, and discount rate 8%.

For robust production optimization, we use the modified EnOpt to maximize the expected NPV (ENPV) over all the realizations under a single injection

scheme. The starting point (the base case) uses an injection rate of 100 m^3/day for all the producers and time steps.

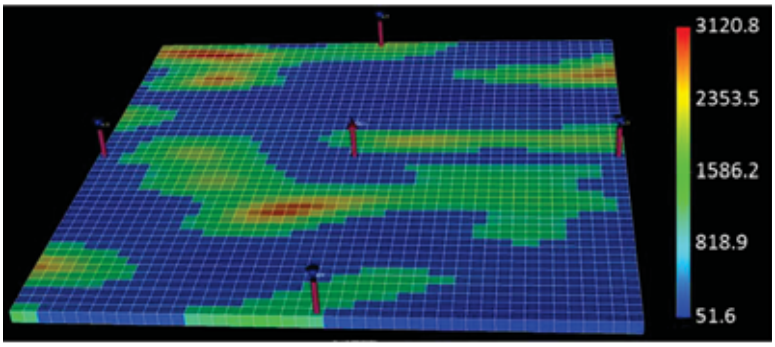


Figure 3.1—Illustration of the 2D model and well locations.

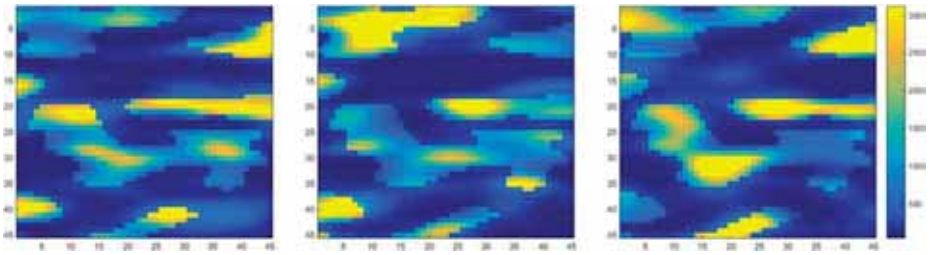


Figure 3.2—Three realizations of the permeability field for the 2D model.

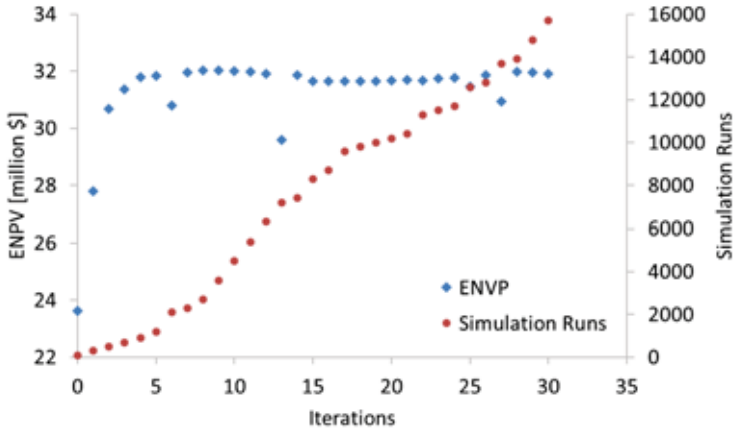


Figure 3.3—Iterative ENPV and cumulative simulation runs.

The ENPV and cumulative simulation runs of each EnOpt iteration are shown in **Figure 3.3**. The ENPV increases sharply at the beginning and starts to

converge after 7 iterations. The ENPV decreases at some iterations because the maximum number of line search iterations (set to 7 in this example) was reached. The optimal ENPV is $32.03 million, corresponding to an improvement of 35.59% over the base case ENPV of $23.62 million. The total number of simulation runs is 15700 for 30 iterations.

Figure 3.4 illustrates the optimal injection scheme. Its major trend is that the injection rate decreases. Because of the discounting nature of NPV calculation, the more oil produced at the early times, the better. At late times, the water cut is high, so water injection rate is reduced to a sufficient rate for continuing production.

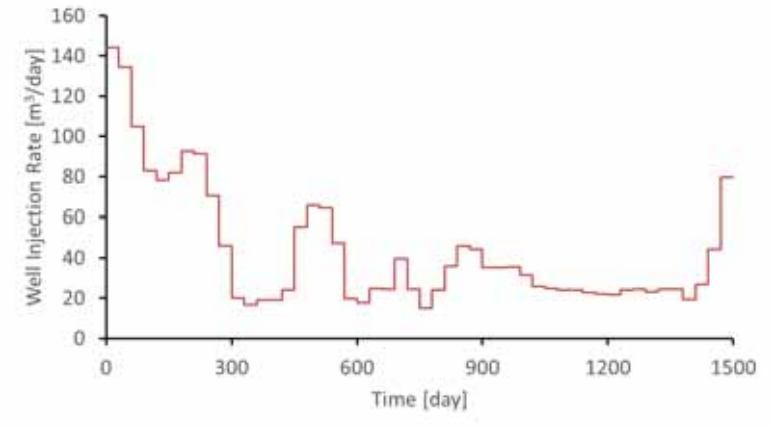


Figure 3.4—Optimal injection scheme for the EnOpt example.

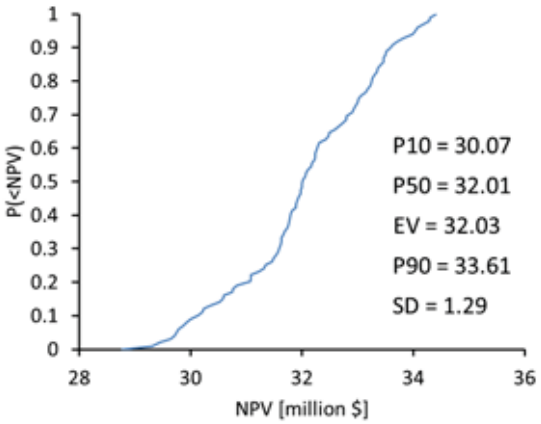


Figure 3.5—CDF of NPV under the optimal injection scheme for the EnOpt example.

The cumulative density function (CDF) of NPVs corresponding to the optimal injection scheme on the 2D model with the 100 realizations is illustrated in **Figure 3.5**. The 80% CI, from P10 to P90, is from $30.07 million to $33.61 million. The SD is $1.29 million, an order of magnitude smaller than the EV. The distribution is quite symmetric.

3.2 Speeding up RO of Waterflooded Production Using CRM

Although EnOpt has significantly reduced the number of simulation runs required, RO can still be computationally intensive when based on grid-based reservoir models using hundreds of realizations. This section presents an RO workflow that embeds proxy models for further reduction of the computational cost for RO.

3.2.1 Proxy Models vs. Rich Models

Whether using rich[10] models or proxy models, one should keep in mind the counsel of Box (1979), "All models are wrong but some are useful."

Engineering uses models to support decision making. Good decision models are both useful and tractable. By "useful," we mean that the model must be relevant and generate insight to resolve the decisions at hand. Usefulness also requires the model to be credible and transparent—will the DMs believe the result of the analysis and can the approach be clearly explained and understood? By "tractable," we mean that the required analysis can be done within the time and resources available.

Both grid-based models and proxy models can be useful. However, a reservoir simulator such as ECLIPSE, with its voluminous code, is so feature rich and detail oriented that it is not transparent to most users. Furthermore, a rich and computationally intensive grid-based model is often not tractable—particularly when the underlying problem is uncertain.

[10] We use "rich" interchangeably with "complex" and "verisimilar" to indicate a high level of detail built into a model.

Decision-making contexts require cogent (compelling) models. Companies tend to build too much detail into their decision-making models from the start and focus too much energy on specific cases or inputs that do not influence the decisions at hand. As stated by Bickel and Bratvold (2008), "This level of modeling detail is really a shirking of responsibility on the part of the decision analyst who will and can build a model that includes only the most salient factors. Building in detail is easy. Building in incisiveness is hard work."

3.2.2 Proxy-model Workflow for RO

When geological uncertainty is considered and is represented by hundreds of realizations, the EnOpt usually requires thousands of production prediction runs. In the traditional workflow, production is predicted by running grid-based reservoir simulation. The computational intensiveness of a grid-based reservoir model can be reduced by having a useful and tractable proxy model serve as a precursor. **Figure 3.6** depicts the traditional and proxy-model workflows for RO.

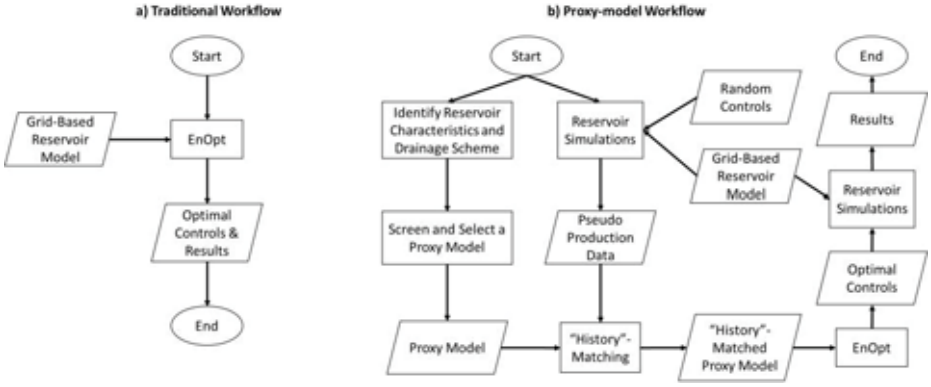


Figure 3.6—Traditional and proxy-model workflows for RO.

The proxy-model workflow starts with identifying reservoir characteristics and drainage scheme for selecting potential proxy models. The potential proxy models are screened by testing their qualities of matching and prediction. "Pseudo production data"[11] are generated by reservoir simulation that includes

[11] We use the term "pseudo production data" because the data are simulated rather than measured.

a random injection scheme. An HM process is used to tune the parameters of a proxy model to where the production predicted by the proxy model fits the pseudo production data. The production predicted by the history-matched proxy model is then compared to that of the grid-based reservoir model, using another random injection scheme for validation. If the difference is unacceptably large either at the HM step or at the validation step, we test another potential proxy model. If the difference is small at both steps, we use the proxy model for PO. After using the optimization algorithm with the proxy model, we run simulation again on the grid-based reservoir model with the optimal control vector of the proxy model to get the final results (e.g., NPV). This last step is for evaluating the optimal control vectors of the reservoir simulation and proxy models on the same basis. This will be discussed later.

3.2.3 Value of Verisimilitude

Similar to the statement that "one cannot value information outside of a particular decision context" (Bratvold et al. 2009), the value of using a more verisimilar or richer model cannot be assessed outside of a particular decision context. Using a more verisimilar model has no value in and of itself. It can add value only if it holds the possibility of changing our decisions. We introduce the value of verisimilitude (VOV) to quantify the economic impact of any difference in optimal solutions obtained by the proxy-model and traditional workflows.

VOV is the benefit that the verisimilitude can offer, and calculated as

$$VOV = v(\boldsymbol{u}_{rich}) - v\big(\boldsymbol{u}_{proxy}\big), \tag{3.10}$$

where $v(\boldsymbol{u})$ is the true value of the decision given a control vector $\boldsymbol{u}$, $\boldsymbol{u}_{rich}$ is the control vector found by using the traditional workflow, and $\boldsymbol{u}_{proxy}$ is that found by the proxy-model workflow. However, being that a given field can support only one production strategy and cannot yield the true value until the end of production, the rich model is assumed to represent the truth. For the case with multiple realizations, Eq. 3.10 is used to calculate the VOV for each realization. This results in a distribution of VOV.

If VOV is zero, using a more verisimilar model does not create any value. The VOV can be compared with the costs saved (converted to monetary terms) by using a proxy model. The costs include the computational cost itself and the costs induced by longer waiting time and later decisions [12] when a more verisimilar model is used. If the costs saved exceeds VOV, then the proxy model is preferable.

Bratvold and Begg (2009) noted, "Companies tend to build too much detail into their decision-making models from the start." This pitfall might be avoided by calculating the VOV. Unfortunately, calculating the VOV is, in itself, computationally expensive, as it can be assessed only by running both the rich and the proxy models. Ideally, the VOV should be obtained without running the rich model. One possible solution is to build a database of VOV assessments, which can eventually be used to estimate the value of using a richer model for similar cases. Another possibility is to formulate the VOV as a sensitivity analysis problem, with the purpose of assessing what the value with a rich model, $v(\boldsymbol{u}_{rich})$, must be in order for the VOV to be greater than the costs. This minimum $v(\boldsymbol{u}_{rich})$ can then be evaluated by the subject matter experts, for determining whether it is realistic for a given decision context.

3.2.4 CRM for Waterflooding

The choice and usefulness of a proxy model is a function of the reservoir characteristics, drainage scheme, and decision context. For RO of waterflooded production, a CRM is a potential candidate.

A CRM is based on material balance and derived from the total fluid continuity equation. It contains considerably fewer parameters and needs significantly less computation time than does a grid-based reservoir model. Required input data for a CRM are production rates, injection rates, and producers' BHPs. The two main parameters of a CRM are connectivity and time constant. For an oil-water system, connectivity is the proportion of injected water in an injector that

[12] An example of the costs induced by longer waiting time and later decisions is: Suppose we need to decide the location of a new well. The waiting time of using a rich model is 7 days and that of a proxy is only 1 day. Then, there will be costs associated with having the drilling team wait for 6 days (the difference between the waiting times of using the rich model and the proxy model) and delaying the production for 6 days.

contributes to the total fluid production in a producer. The time constant indicates how long a pressure wave from an injector takes to reach a producer.

Based on control volume, CRM can be divided into three categories: single-tank CRM (CRMT), producer-based CRM (CRMP), and injector-producer-pair-based CRM (CRMIP) (Sayarpour et al. 2009). As illustrated in **Figure 3.7**, the control volume in a CRMT is the entire drainage volume of a reservoir covering a single pseudo-producer (encompassing all physical producers) and a single pseudo-injector (encompassing all physical injectors), enabling the entire reservoir to be treated as a single tank with one inlet and one outlet. A CRMP has producer-based control volumes, each of which covers all the injectors influencing its corresponding producer. A CRMIP has one control volume for each injector-producer pair.

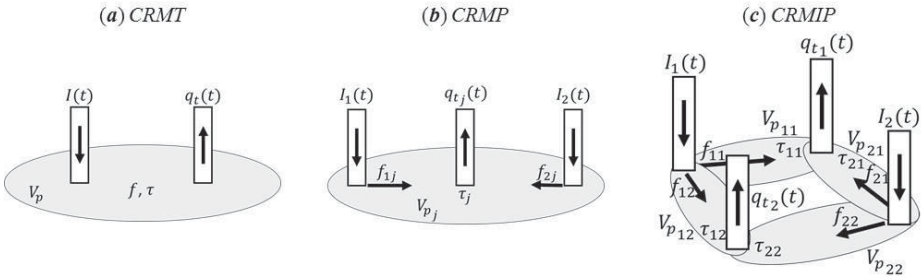


Figure 3.7—Schematic of (a) CRMT, (b) CRMP, and (c) CRMIP.

This work focuses on CRMP. The fundamental equation of CRMP is the total fluid production equation. For constant producer BHPs, the equation is (Sayarpour 2008)

$$q_{t_j}^{k} = q_{t_j}^{k-1} \exp\left(-\frac{\Delta t_k}{\tau_j}\right) + \left[1 - \exp\left(-\frac{\Delta t_k}{\tau_j}\right)\right]\left(\sum_{i=1}^{n_{inj}} \lambda_{ij} I_i^k\right), \quad (3.11)$$

where sub/superscripts i, j, and k are the indices of injector, producer, and time step, respectively; $q_{t_j}^{k}$ is the total fluid production rate of producer j at time k; Δt_k is the time step length between times k and $k - 1$; τ_j is the time constant for producer j; n_{inj} is the number of injectors; λ_{ij} is the connectivity between injector i and producer j; and I_i^k is the water injection rate of injector i during the period Δt_k. Both λ and τ are assumed to be constant with respect to time.

Because Eq. 3.11 can predict only total fluid production, combining it with a fractional flow model is necessary in order to separate oil production from total production. Sayarpour et al. (2011) combined CRMP with a Buckley-Leverett-based fractional flow model, Gentil (2005) with an empirical fractional flow equation, and Cao et al. (2015) with the Koval (1963) model. Cao (2014) demonstrated that the Koval model might not yield a good match for mature waterfloods because it might approach an abrupt end of 100% water cut. She also noted that the Gentil model might not work well for immature waterfloods because the relationship between the natural logs of water-oil ratio and cumulative water injection might be non-linear at that stage. These concerns prompted the development of a fully coupled two-phase-flow-based CRMP (Coupled CRMP), which is applicable in all stages of maturity (Cao 2014, Cao et al. 2015).

The parameters of CRM and fractional flow model are determined by history-matching the production data.

A detailed review of CRMP combined with the Koval model and Coupled CRMP is provided in **Paper II**.

3.2.5 Example of Applying the Proxy-model Workflow

To illustrate the application of the proxy-model workflow with Coupled CRMP and the concept of VOV, we use the 2D model with 100 realizations that was presented in Section 3.1. **Paper II** applies this approach to a more realistic case with a synthetic 3D reservoir simulation model.

It yields 100 sets of the parameters of Coupled CRMP (i.e., an ensemble of Coupled CRMP) to repeat the procedure of matching Coupled CRMP to the pseudo production data simulated by every realization of the 2D model. The qualities of matching and prediction are shown in **Figure 3.8** and **Figure 3.9**, respectively. The matching quality is good: the minimal R^2 (the worst matching) for total production rate and water cut are 0.9950 and 0.9929, respectively. The prediction quality is less good, but still satisfactory, with a minimal R^2 of 0.9946 for total production rate and 0.9546 for water cut.

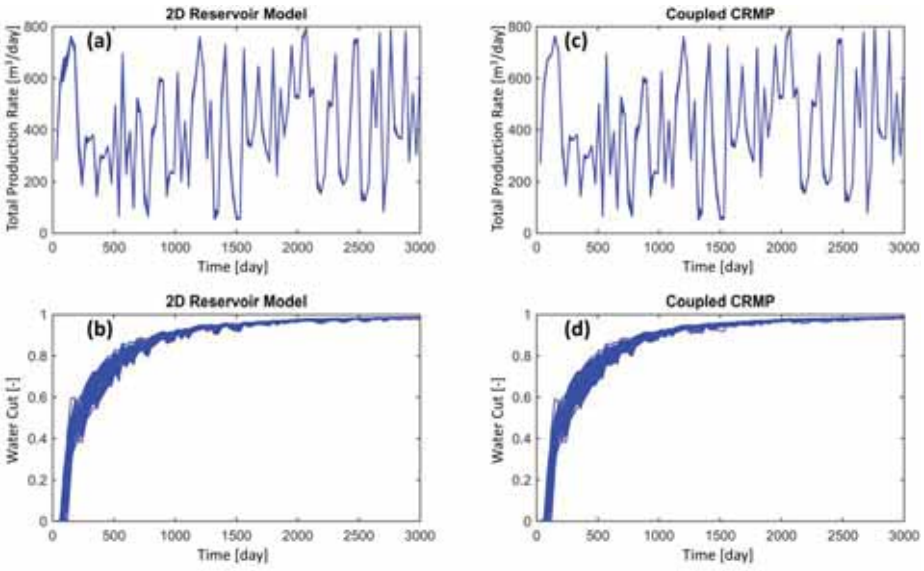


Figure 3.8—Matching of total fluid production rate and water cut: (a) total fluid production rate from the 2D model, (b) water cut from the 2D model, (c) total fluid production rate from Coupled CRMP, and (d) water cut from Coupled CRMP.

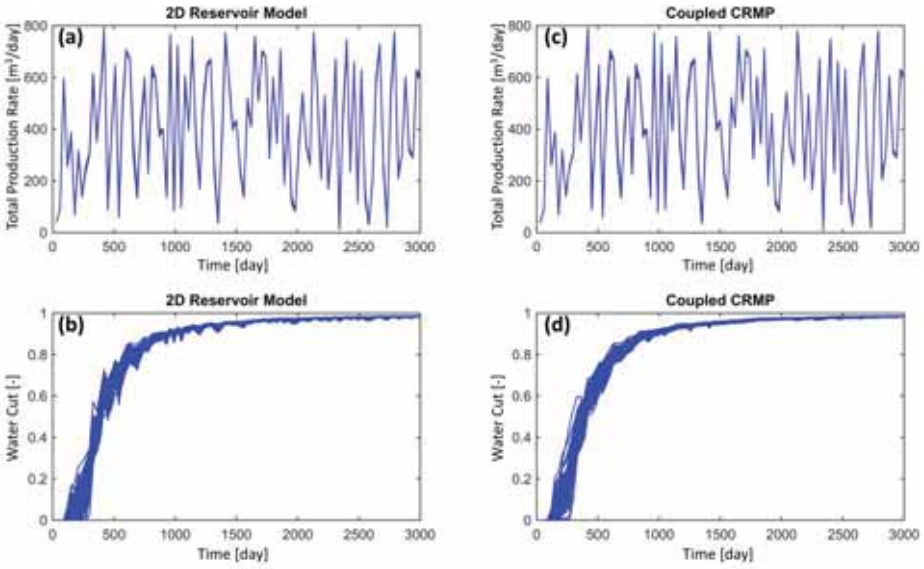


Figure 3.9—Validation of total fluid production rate and water cut: (a) total fluid production rate from the 2D model, (b) water cut from the 2D model, (c) total fluid production rate from Coupled CRMP, and (d) water cut from Coupled CRMP.

For robust production optimization, we use the modified EnOpt to maximize the expected NPV over all the realizations, under a single injection scheme. The optimal ENPV of the 2D model is $32.03 million and of Coupled CRMP is $31.12 million. The expected NPV of the proxy-model workflow is $31.84 million, which is close to that of the traditional workflow. **Figure 3.10** illustrates the CDFs of NPV corresponding to the optimal solutions of the 2D model and of Coupled CRMP, and to the 2D model under the optimal injection scheme of Coupled CRMP. The optimal injection schemes are shown in **Figure 3.11**.

The proxy-model and traditional workflows lead to similar optimal expected NPVs and similar NPV distributions (the solid red curve vs. the solid blue curve in Figure 3.10), although they provide different optimal injection schemes. This indicates that these two injection schemes are almost equally good if NPV is the only criterion. To determine the better injection scheme, we could add relevant criteria; for example, we might want to minimize the frequency of significant changes in the injection rates and thus the optimal injection scheme of Coupled CRMP is better than that of the 2D model.

Coupled CRMP underestimates NPV (the dashed red curve vs. the solid red curve in Figure 3.10). We aim to investigate the difference in NPVs due to the difference in decisions made based on the rich and proxy models. The difference in NPVs should not include that due to the calculation of NPVs based on different models. Therefore, we should evaluate the optimal control vectors of the rich and proxy models on the same basis (i.e., the rich model). This is why the proxy-model workflow runs the simulation again on the grid-based reservoir model with the optimal control vector of the proxy model to get the final NPV.

The total number of production prediction runs during EnOpt was 21800 for Coupled CRMP and 15600 for the 2D model. The run times of Coupled CRMP and the 2D model for a prediction of 3000 days were 0.17 and 4.1 seconds, respectively, and the computational time of one HM was 1.1 seconds. The 1500-day run time of a life-cycle prediction is one half of the run time of a 3000-day prediction. The total computational time of the traditional workflow is 31980 seconds, and of the proxy-model workflow is 2578 seconds. The computation time required by the proxy-model workflow is therefore less than

1/10 that of the traditional workflow. The distinction will become more pronounced if the grid-based reservoir model has a finer grid block size or higher geological complexity. Longer computational time or waiting time caused by using a richer model can potentially come with significant cost of decision delays (see the discussion related to Footnote 12).

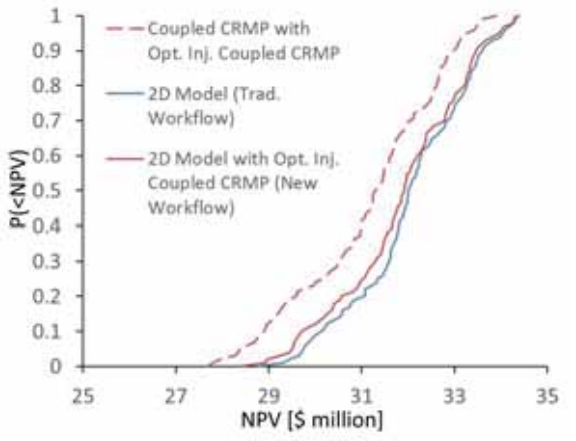


Figure 3.10—CDFs of NPV for the proxy-model workflow example.

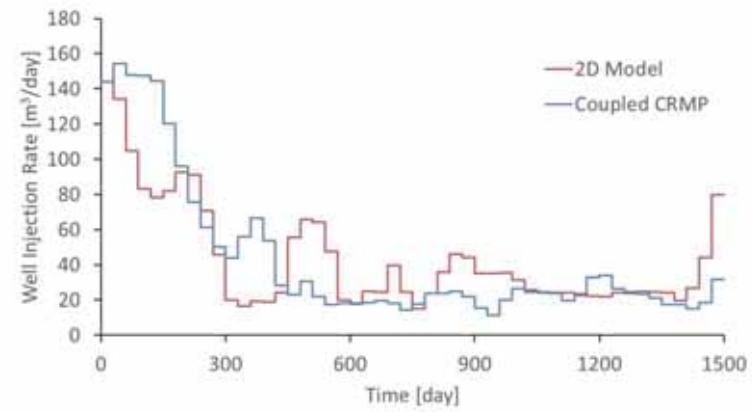


Figure 3.11—Optimal injection schemes for the proxy-model workflow example.

To calculate the VOV, we consider the following decision-making context that involves three decisions in series. The first decision has two options: to continue developing the field at a cost of $10 million or to walk away. The second decision is whether to install valves on the injectors at a cost of $2

million, given that water would be injected over the entire production life-cycle at a constant rate of 100 m^3/day without valves or at any rate from 0 to 200 m^3/day with valves. The third decision is to choose a water injection scheme if valves are installed. All the decisions will be based on our current knowledge about the field. We assume a risk-neutral case, where the objective is to optimize the EV.

The decision trees are illustrated in **Figure 3.12** and **Figure 3.13**, where a square, a circle, and a triangle represent a decision node, an uncertainty node, and a payoff node, respectively. (Bratvold and Begg (2010) includes a detailed description of using decision trees to solve hydrocarbon-production-related decision-making problems.) The P10, P50, and P90 values are shown in the decision trees. The distributions of the NPV are expected to differ because the proxy-model and traditional workflows lead to different optimal injection schemes. Given our objective of maximizing the ENPV of a particular decision-making problem, the solution of the proxy-model workflow gives an ENPV very close to that of the traditional workflow.

The optimal choice is the same for the first and second decision nodes. However, the two workflows make the third decision diverge, resulting in an expected revenue of \$19.84 million for the proxy-model workflow (applying the blue line in Figure 3.11) and of \$20.03 million for the traditional workflow (applying the red line in Figure 3.11). The expected VOV (EVOV) is \$20.03 million - \$19.84 million = \$0.19 million. Using a more verisimilar production prediction model has no value in and of itself, as it must be material to potentially change our decisions. Thus, the 2D model (a richer model) adds no value to the first and second decisions in this case. It adds value only to the third decision. The EVOV is positive because the 2D model can capture more minor flow behaviors than can Coupled CRMP, giving a more accurate production prediction under our assumption that the 2D model represents the truth. However, the EVOV (\$0.19 million) is more than an order of magnitude less than the ENPV improvement (\$19.84 million − \$13.62 million = \$6.22 million) by installing the valves and conducting the optimal injection scheme found by performing EnOpt on Coupled CRMP. This is because Coupled CRMP is already a good approximation to the 2D model. The CDF of the VOV is illustrated in **Figure 3.14** which shows that the chance of a negative VOV is about 5.4%. The minimal VOV is −\$0.096 million, and the maximal is \$0.783

million. This indicates that using a richer model for RO cannot guarantee a better control than using a proxy model for each individual realization.

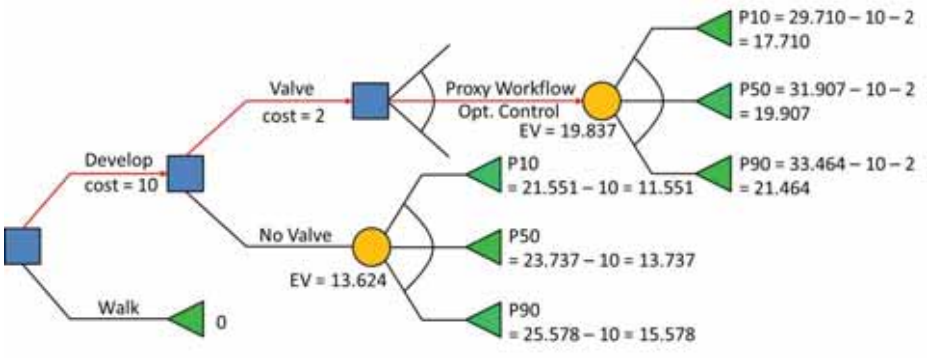


Figure 3.12—Decision tree with the optimal solution of the proxy-model workflow.

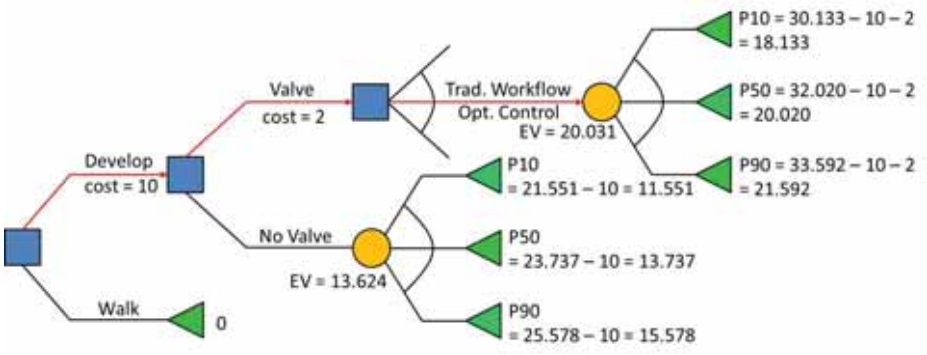


Figure 3.13—Decision tree with the optimal solution of the traditional workflow.

We investigate the sensitivity of the EVOV to the valve cost by varying the latter. The results are plotted in **Figure 3.15**, which can be divided into three regions: valve cost lower than $8.21 million (Region 1), from $8.21 million to $8.41 million (Region 2), and greater than $8.41 million (Region 3). In Region 1, the EVOV is a constant $0.194 million because only the third decision is affected by whether the proxy-model or traditional workflow is chosen. In Region 2, installing valves is called for by the traditional workflow but not by the proxy-model workflow. The linear decrease in the EVOV in Region 2 is because the value with valves installed is a linear function of the valve cost and the value with valves not installed is a constant. In Region 3, the EVOV drops to 0 because neither the proxy-model nor the traditional workflow calls for

installing valves. This illustrates that the value of using a richer model is decision-dependent.

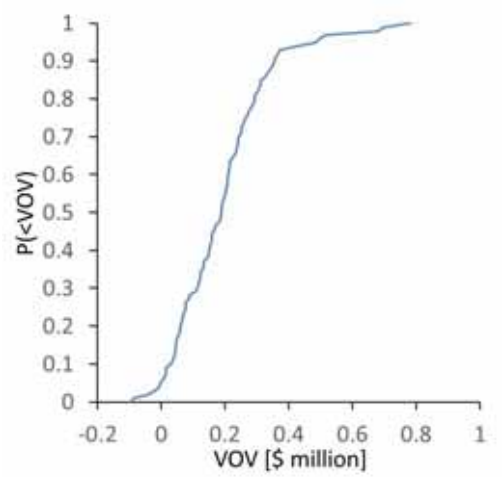


Figure 3.14—CDF of the VOV for the example of applying the proxy-model workflow.

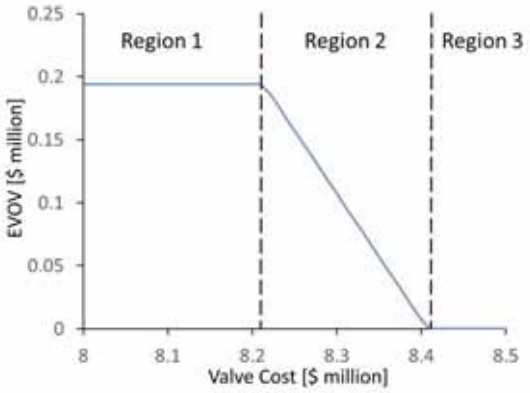


Figure 3.15—Sensitivity of EVOV to valve cost.

4 Managing Geological Uncertainty in Reservoir Management

Previous chapters have presented approaches to managing geological uncertainty in HM and PO. This chapter will present iterative approaches to managing geological uncertainty that combine HM and PO for making sequential RM decisions. This iteration is central to RM.

4.1 Closed Loop Reservoir Management

In CLRM (closed loop reservoir management—also known as "real time reservoir management," "smart reservoir management," or "closed loop optimization"), the loop of HM and PO is closed by continuously updating a production model with several realizations and performing life-cycle optimization whenever new data become available (Brouwer et al. 2004, Nævdal et al. 2006, Chen et al. 2009, Wang et al. 2009, Jansen et al. 2009). **Figure 4.1** illustrates the process flow of CLRM. An initial ensemble of a production model with numerous realizations is built based on prior knowledge on the reservoir, and an initial production strategy for the whole life-cycle is determined by performing RO on the initial ensemble. The initial production strategy is applied to the real field until new data becomes available. The new data are used to update the production model, and a new production strategy for the remainder of the life-cycle is determined by performing RO on the updated ensemble of the production model. The new production strategy is applied to the real field until new data become available. Repeating the process keeps the RM up to date.

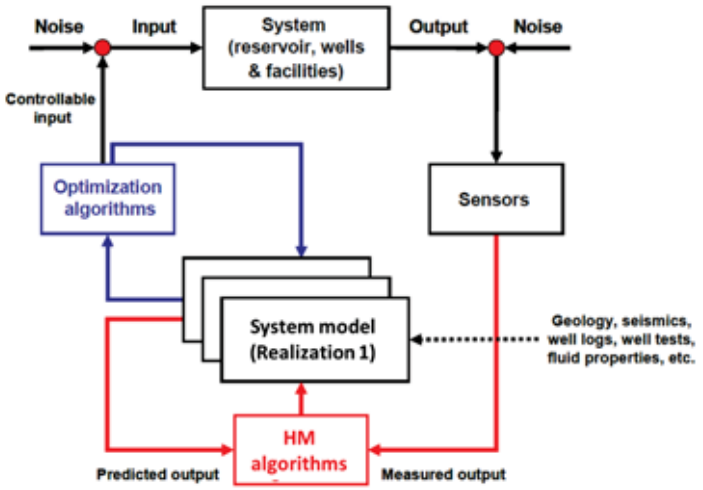


Figure 4.1—Process flow of CLRM, adopted from Jansen et al. (2009).

CLRM provides a myopic or naive decision policy in that the uncertainties and decisions associated with currently available data are considered but the uncertainties and decisions associated with future data are not (Kullawan 2016, Thomas 2016). A decision-tree representation for this approach is illustrated in **Figure 4.2**, where D_k denotes decisions made at t_k and U_k denotes the uncertainties associated with current available data until t_k. U_k is commonly represented by a production model with numerous realizations. The production strategy $D_k, D_{k+1}, \ldots, D_{end}$ is determined with the consideration of only the immediate relevant uncertainty U_k.

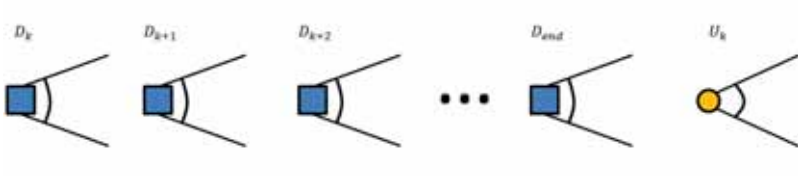


Figure 4.2—Decision-tree representation for CLRM.

CLRM greatly simplifies the structure of a RM decision problem and consequently incurs less computational cost for solving the problem. However, its failure to reflect the full structure of an RM decision problem might lead to a suboptimal production strategy.

4.2 Fully Structured Reservoir Management

The decision-tree representation of fully structured reservoir management (FSRM) is illustrated in **Figure 4.3**.[13] The problem structure required for using FSRM is sequential decision making. If there is only one decision point (i.e., only D_k and U_k are included in the decision tree), then FSRM and CLRM will yield the same solutions.

FSRM explicitly considers both the uncertainties associated with current available data (U_k) and those with future data (U_{k+1}, …, U_{end}). Therefore, the current decision (D_k) does not depend only on the uncertainties that a DM has learned so far, but on the uncertainties that the DM will learn in the future.[14] FSRM includes the DM's learning (D_{k+1}, …, D_{end}) from the revelation of uncertainties associated with future data (U_{k+1}, …, U_{end}).

Solving the decision tree in Figure 4.3 gives the optimal production strategy. However, it can be computationally intensive or even prohibitive because of the dependencies among the tree elements[15] or the computational intensiveness of a production model.

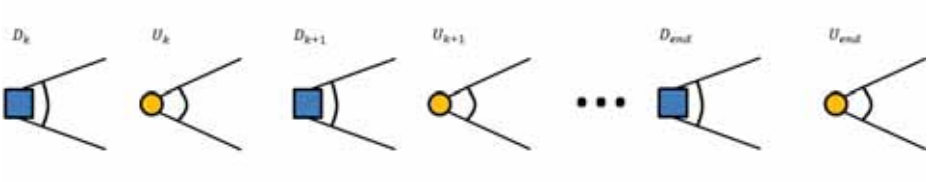


Figure 4.3—Decision-tree representation of FSRM.

An example illustrates the differences between FSRM and CLRM solutions. Consider two geological realizations: Realization 1 (R1) has a high permeability channel between the injector and producer, and Realization 2 (R2)

[13] We call this approach "fully structured" because we intend to (approximately) solve the fully structured decision tree (Figure 4.3), which reflects the real structure of a sequential decision-making problem, rather than the "simplified" decision tree (Figure 4.2).

[14] The uncertain quantities under consideration won't change over time, but the quantification of their uncertainties will change with new knowledge (learning) over time.

[15] For example, a value function depends on reservoir properties (uncertainties) and the chosen production strategy (decisions).

doesn't. The prior probabilities of these two realizations are 0.5 and 0.5, respectively. The water cut behavior for these two realizations is shown in **Figure 4.4**. Because of the channel, R1 (red curve) will have earlier water breakthrough than R2 (blue curve) under the same production strategy. Consider a life-cycle of 120 months. The decision variables are the injection rates for 3 time intervals (Time Interval 1 (TI1): Month 1–Month 40, Time Interval 2 (TI2): Month 41–Month 80, and Time Interval 3 (TI3): Month 81–Month 120). The injection rate is bounded at 0 to 150 m^3/day. The decisions will be made at the end of Month 0, Month 40, and Month 80. Before making each decision, information from production data will be used to update our belief on the existence of the channel. Assuming that the reliability of the information is 0.9, we have the likelihood matrix listed in **Table 4.1**. The decision tree for this context is illustrated in **Figure 4.5**.

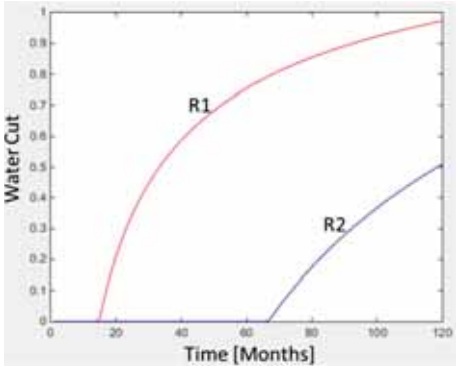


Figure 4.4—Water cuts of R1 and R2.

		Given the truth is	
		R1	**R2**
Information says	**"R1"**	0.9	0.1
	"R2"	0.1	0.9

Table 4.1—Likelihood matrix for the example comparing FSRM and CLRM solutions.

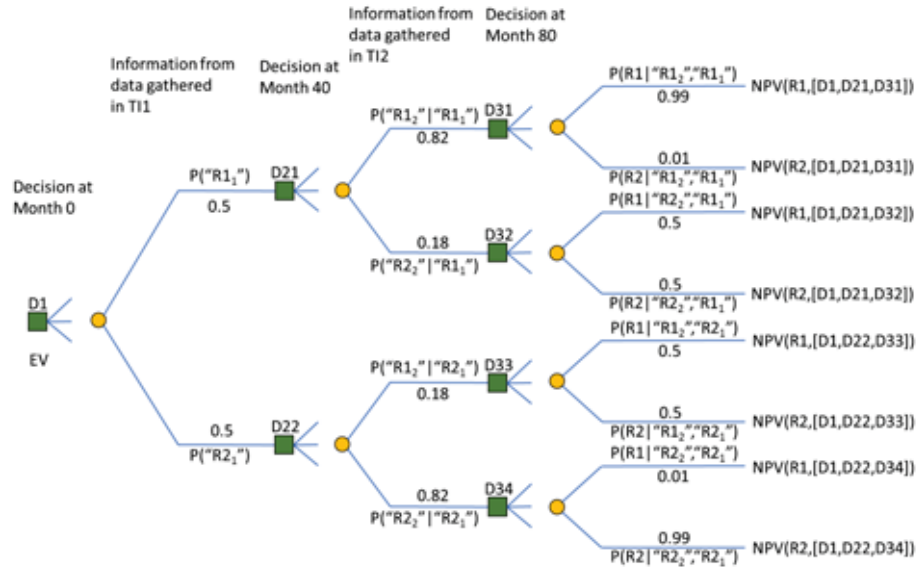


Figure 4.5—Decision tree for the example comparing FSRM and CLRM solutions.

The FSRM and CLRM solutions are listed in **Table 4.2**. The FSRM solution leads to the optimal EV of $45.963 million, and the CLRM solution gives a lower EV of $45.927 million.

Decision Node	D1	D21	D22	D31	D32	D33	D34
FSRM	150.0	56.4	150.0	0.0	150.0	82.6	136.3
CLRM	150.0	76.9	150.0	0.0	150.0	82.6	136.3

Table 4.2—Injection rates for decision nodes solved using FSRM and CLRM.

The only difference between the two solutions is that FSRM assigns a lower injection rate to D21. Given that the first information indicates channel existence ("$R1_1$"), the probability of channel existence (R1) is 0.9 and that of no channel (R2) is 0.1; thus, CLRM assigns a moderate injection rate of 76.9 m^3/day to D21. Indeed, given that the first information indicates channel existence ("$R1_1$"), we can "foresee" that there is a high chance of 0.82 that the second information will indicate channel existence, resulting in a probability of channel existence (R1) of 0.99; thus, FSRM further reduces injection rate to 56.4 m^3/day for D21. This example illustrates that FSRM and CLRM can lead to different decisions and the CLRM solution can be suboptimal.

The difference between the FSRM and CLRM solutions ($45.963 million vs. $45.927 million) is very small for this example. The improvement of using FSRM instead of CLRM depends on the characteristics of a problem (uncertainties, decisions, and values). **Paper III** presents an example in the context of determining the optimal IOR start time. That example shows a 6.7% improvement in ENPV (from $2641.2 million to $2818.0 million) when FSRM is used instead of CLRM.

4.3 Least-Squares Monte Carlo Algorithm for FSRM

The LSM algorithm is an approximate dynamic programming algorithm, which approximates the conditional future value of an alternative by regression. Thus, it does not suffer from the "curse of dimensionality" induced by the uncertainties. However, its computational time increases linearly with the computational time of a production model and exponentially with the number of alternatives. It is promising for solving an FSRM problem with a computationally attractive production model and small number of alternatives.

4.3.1 LSM Algorithm

The LSM algorithm was proposed by Longstaff and Schwartz (2001) for American option problems that involve a yes-no decision: at any time, an option holder can decide whether to immediately exercise the option at the current stock price or to continue holding the option for exercising it at a future stock price. The stock price is an uncertain quantity in an American option problem, whose uncertainty can be modeled as a Markov process, and thus an alternative's future value conditions only to the current stock price.

However, geological uncertainty depends not only on the currently measured data but on the previously measured data, and consequently an alternative's future value conditions to currently and previously measured data. To address geological uncertainty, we slightly modify the LSM algorithm to include the dependency of the future value on the currently and previously measured data. The two central steps of our modified LSM algorithm are

(1) Monte Carlo Simulation (MCS) Step:

N independent samples of model parameters representing geological uncertainty are generated using MCS. For one sample of model parameters, forward modeling is performed to provide modeled production data from t_0 (time 0) to t_{end} (end time); and then, random noises generated based on the statistics of the measurement errors are added to the modeled production data to provide a sample of measured data. Because this sample of measured data consists of a series of data points in time, it is also called a path of measured data. Repeating this procedure for each of the N sampled sets of model parameters, we obtain N paths of measured data, i.e., $[\boldsymbol{y}_1, \boldsymbol{y}_2, \dots, \boldsymbol{y}_N]^{\mathrm{T}}$.

(2) Least Squares Step:
For the ith sample of model parameters, the NPV of alternative a is calculated, giving $NPV_i(a)$. This is repeated for the N sampled sets of model parameters, resulting in $[NPV_1(a), NPV_2(a), \dots, NPV_N(a)]^{\mathrm{T}}$. To estimate the ENPV with alternative a conditioned on the measured data, $ENPV(a)|\boldsymbol{y}$, we regress $[NPV_1(a), NPV_2(a), \dots, NPV_N(a)]^{\mathrm{T}}$ on $[\boldsymbol{y}_1, \boldsymbol{y}_2, \dots, \boldsymbol{y}_N]^{\mathrm{T}}$. This procedure is repeated for each of the alternatives.

The modified LSM algorithm is applied in **Paper III**.

4.3.2 Example of Applying LSM

An example illustrates the detailed steps of applying our modified LSM algorithm to solve an FSRM problem.

A field has a life-cycle of 15 years. The decision-problem setting is relevant to the optimal IOR start time. We consider a yes-no decision at the end of Years 0, 5, and 10: whether primary recovery should be shifted to secondary recovery. The production is modeled using a two-factor model (Parra-Sanchez 2010) that includes a factor for the recovery increment due to the change in recovery mechanism and another factor for the time constant of a recovery mechanism. For primary recovery, the model is formulated as

$$E_{R1}(t) = E_{R1}^{\infty}\left(1 - \exp\left(-\frac{t}{\tau_1}\right)\right), \tag{4.1}$$

and for secondary recovery,

$$E_{R2}(t) = E_{R1}(t_{LT1}) + \Delta E_{R2}^{\infty}\left(1 - \exp\left(-\frac{t - t_{LT1}}{\tau_2}\right)\right), \qquad (4.2)$$

where E_{R1} is the primary oil recovery efficiency, E_{R1}^{∞} is the theoretical ultimate recovery efficiency for primary recovery, τ_1 is the primary recovery time constant for production, E_{R2} is the secondary oil recovery efficiency, t_{LT1} is the life time of primary recovery, ΔE_{R2}^{∞} is the theoretical ultimate recovery increment for secondary recovery, and τ_2 is the secondary recovery time constant for production. Then, the cumulative oil production is calculated as

$$Q_o(t) = OOIP \cdot E_R(t), \qquad (4.3)$$

where $E_R = E_{R1}$ or E_{R2}.

The geological uncertainty is represented by three realizations (R1, R2, and R3) of the production model parameters. The realizations are a priori equi-probable. **Figure 4.6** illustrates the primary recovery as a function of time for the realizations.

The measured primary recovery efficiency is used to inform our decisions. The likelihood functions are listed in **Table 4.3**. For example, given that the truth is R1, the probability of measurement at t_1 saying "high recovery" (i.e., a measured recovery efficiency of 0.3) is 1/4, and the probability of measurement at t_2 saying "high recovery" (i.e. a measured recovery efficiency of 0.33) is 3/4.

Our objective is to solve for a decision policy that maximizes the value of the flexibility of shifting to secondary recovery at various times. This can be achieved using a decision tree. The fully structured decision tree for this problem setting is illustrated in **Figure 4.7**, where the optimal decision policy is indicated by the red branches. The optimal ENPV is $2454.45 million.

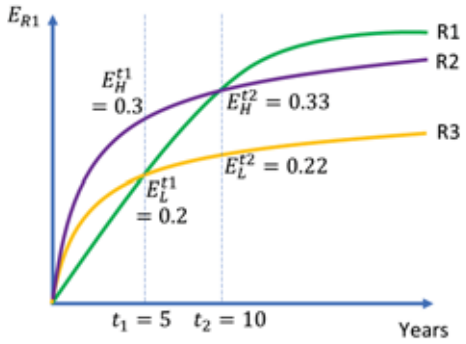


Figure 4.6—Primary recovery as a function of time for three geological realizations.

		Given the truth is		
		R1	R2	R3
Measurement at t_1 says	"High Recovery" ("E_H^{t1}")	1/4	3/4	1/4
	"Low Recovery" ("E_L^{t1}")	3/4	1/4	3/4
Measurement at t_2 says	"High Recovery" ("E_H^{t2}")	3/4	3/4	1/4
	"Low Recovery" ("E_L^{t2}")	1/4	1/4	3/4

Table 4.3—Likelihood functions for the measured primary recovery efficiency.

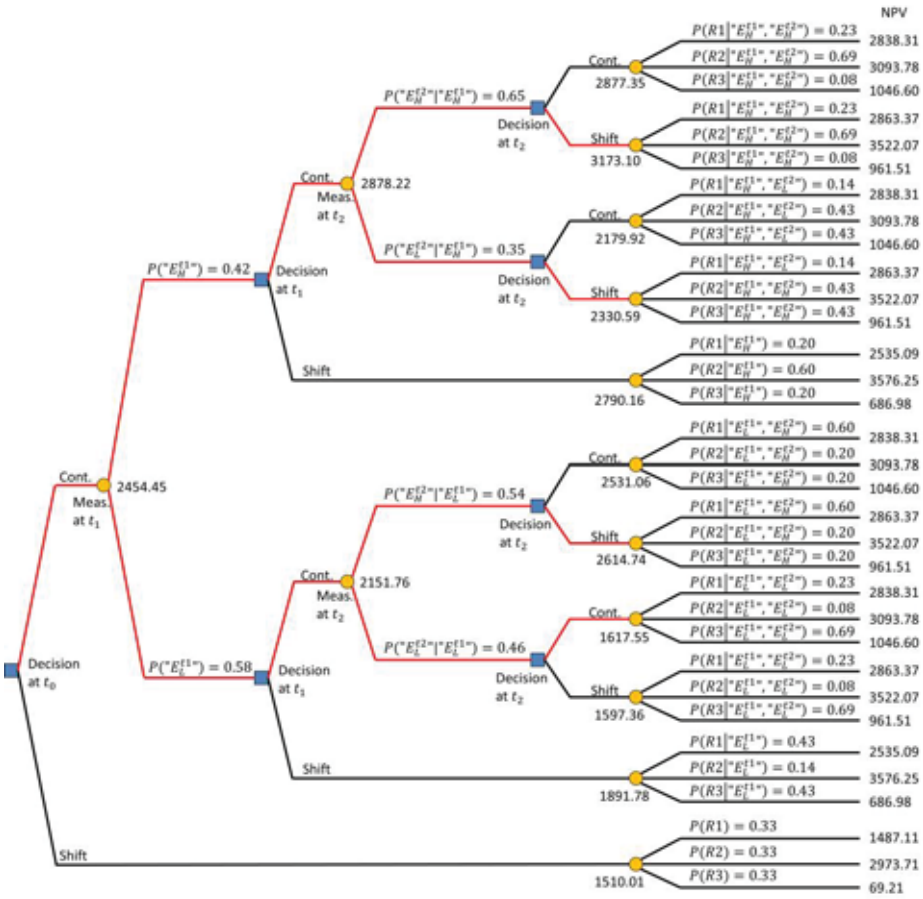


Figure 4.7—Fully structured decision tree for the example of an optimal IOR start time problem. All NPV values are in million USD.

We use the modified LSM to solve the same problem. For illustration, we use only 5 MC samples. The number of samples will be increased later. We first sample 5 paths of the measured data as listed in **Table 4.4**. Then, we start the LSM from the last decision point, as we solve the decision tree backwards. **Table 4.5** lists the NPVs corresponding to alternatives of "continue (with primary recovery)" and "shift (to secondary recovery)" at t_2 for each path.

Path	Geological Realization	Data at	
		t_1 ("E_{R1}^{t1}")	t_2 ("E_{R1}^{t2}")
1	R2	0.30	0.22
2	R3	0.20	0.22
3	R3	0.20	0.33
4	R1	0.20	0.33
5	R3	0.30	0.22

Table 4.4—Path of measured data for LSM.

Path	Geological Realization	NPV [million USD]	
		Continue	Shift
1	R2	3093.78	3522.07
2	R3	1046.60	961.51
3	R3	1046.60	961.51
4	R1	2838.31	2863.37
5	R3	1046.60	961.51

Table 4.5—NPVs for alternatives at time 2 for LSM.

When we are making the decision at t_2, we have obtained the data at both t_1 and t_2. Therefore, the ENPV is conditioned on the data at both t_1 and t_2. We estimate the ENPV by regressing NPV on the data. Using $ENPV|data = a_0 + a_1"E_{R1}^{t1}" + a_2"E_{R1}^{t2}"$, the resulting conditional expectation function for "continue" is $ENPV|data = -2792.3 + 10235.9 \cdot "E_{R1}^{t1}" + 8144.1 \cdot "E_{R1}^{t2}"$. The same is done for "shift". **Table 4.6** lists the resulting ENPV corresponding to alternatives of "continue" and "shift" at t_2 for each path. The optimal decision can now be made by taking the alternative that gives the highest ENPV for each path. Each optimal decision is colored in red in Table 4.6.

Path	ENPV [million USD]	
	Continue	Shift
1	2070.19	2241.79
2	1046.60	961.51
3	1942.46	1912.44
4	1942.46	1912.44
5	2070.19	2241.79

Table 4.6—ENPVs for alternatives at time 2 for LSM.

We move one time step backwards to t_1. **Table 4.7** lists the NPVs corresponding to alternatives of "continue" and "shift" at t_1 for each path. The

NPVs for "continue" in Table 4.7 come from the NPVs in Table 4.5 for the optimal decisions shown in Table 4.6.

Path	Geological Realization	NPV [million USD]	
		Continue	Shift
1	R2	3522.07	3576.25
2	R3	1046.60	686.98
3	R3	1046.60	686.98
4	R1	2838.31	2535.09
5	R3	961.51	686.98

Table 4.7—NPVs for alternatives at time 1 for LSM.

When we are making the decision at t_1, only the data at t_1 are available but not at t_2. Therefore, the ENPV is conditioned on only the data at t_1. Using $ENPV|data = b_0 + b_1"E_{R1}^{t1}"$, the resulting conditional expectation function for "continue" is $ENPV|data = 447.9 + 5979.5 \cdot "E_{R1}^{t1}"$. The same is done for "shift". **Table 4.8** lists the resulting ENPV corresponding to alternatives of "continue" and "shift" at t_1 for each path, and the optimal decisions are in red.

Path	ENPV [million USD]	
	Continue	Shift
1	2241.79	2131.62
2	1643.84	1303.02
3	1643.84	1303.02
4	1643.84	1303.02
5	2241.79	2131.62

Table 4.8—ENPVs for alternatives at time 1 for LSM.

We move to t_0. **Table 4.9** lists the NPVs corresponding to alternatives of "continue" and "shift" at t_0 for each path. The NPVs for "continue" in Table 4.9 come from the NPVs in Table 4.7 for the optimal decisions shown in Table 4.8.

Because no data are available at t_0, we calculate the ENPVs for "continue" and "shift" by calculating the mean of the values in their respective columns in Table 4.9. The resulting ENPV for "continue" is \$1883.02 million, and that for "shift" is \$933.69 million. Thus, the optimal decision at t_0 is "continue", and the optimal ENPV is \$1883.02.

Path	Geological Realization	NPV [million USD]	
		Continue	Shift
1	R2	3522.07	2973.71
2	R3	1046.60	69.21
3	R3	1046.60	69.21
4	R1	2838.31	1487.11
5	R3	961.51	69.21

Table 4.9—NPVs for alternatives at time 0 for LSM.

The optimal decision policy can be determined by looking back at Table 4.8 and Table 4.6, and it is represented by **Table 4.10**. This indicates that we should continue with primary recovery up to t_2 irrespective of the measurement at t_1, and shift to secondary recovery at t_2 if the measurement says "E_H^{t1}" at t_1 and "E_L^{t2}" at t_2.

Path	Data at		Optimal Decision at		
	t_1 ("E_{R1}^{t1}")	t_2 ("E_{R1}^{t2}")	t_0	t_1	t_2
1	0.30	0.22	Continue	Continue	Shift
2	0.20	0.22	Continue	Continue	Continue
3	0.20	0.33	Continue	Continue	Continue
4	0.20	0.33	Continue	Continue	Continue
5	0.30	0.22	Continue	Continue	Shift

Table 4.10—Table representation of optimal decision policy solved using LSM.

The optimal ENPV and optimal decision policy solved using the modified LSM are different from the analytical solution from the decision tree (Figure 4.7) because we used only 5 paths. The accuracy can be improved by increasing the number of paths. We increase to 500000 paths and repeat it 1000 times. This gives a mean estimated optimal ENPV of \$2454.48 million (which is almost the same as the analytical solution \$2454.45 million) and an SD of \$1.49 million (which is relatively small).

5 Value-of-Information in Reservoir Management

The concept of VOI was first introduced in the O&G industry for drilling decisions that are related to RM (Grayson 1960). More than 50 years later, this concept was adopted to assess the value of information from HM (Barros et al. 2015, Barros et al. 2016, He et al. 2017). The information from HM is used to support decisions on production strategy.

5.1 Definition of VOI

The concept of VOI[16] originates from the DA community. Schlaifer (1959) was the first to define VOI in the context of business decisions. Other early references to VOI analysis can be found in Raiffa and Schlaifer (1961) and Howard (1966). Its use has been documented in broad areas of real-world application, from nuclear waste storage assessment (Eppel and von Winterfeldt 2008) to biosurveillance (Willis and Moore 2013). Grayson (1960) introduced this concept in the O&G industry. Bratvold et al. (2009) presented the definition of VOI as well as an overview of its use in the O&G industry. Eidsvik et al. (2015) provided an exposition of VOI in the earth sciences.

VOI analysis is concerned with two fundamental uncertainties: (1) the uncertainties we hope to learn about but cannot directly observe, which we call the distinctions (or events) of interest; and (2) the test results, referred to as the observable distinctions (Bratvold et al. 2009). In the following, we use $\boldsymbol{x}$ to denote the distinctions of interest and $\boldsymbol{y}$ to denote the observable distinctions.

VOI is defined as the most that the DM should pay for additional information on the distinctions of interest. It is calculated before the additional information is actually acquired. If the DM is risk neutral, then[17]

[16] To separate "value-of-information (VOI)" from other definitions, we use the term as it is defined in the DA community.
[17] This is not the general definition of VOI. This formulation is true only if the DM is risk-neutral or risk-averse with an exponential utility function (Bratvold et al. 2009).

$$VOI = \begin{bmatrix}\text{Expected value } with \\ \text{additional information}\end{bmatrix} - \begin{bmatrix}\text{Expected value } without \\ \text{additional information}\end{bmatrix}. \quad (5.1)$$

If the additional information is perfect (i.e., the information reveals the truth), we refer to the VOI as the value-of-perfect-information (VOPI) or the value-of-clairvoyance (VOC). No test, no matter how sophisticated, can be worth more than the VOPI (Bratvold et al. 2009).

In mathematical form,

$$VOI = \max[0, \Delta], \quad (5.2)$$
$$\Delta = EVWI - EVWOI, \quad (5.3)$$

where $EVWOI$ is the EV without additional information and $EVWI$ is the EV with additional information. The lower bound of VOI is always 0 because if Δ is negative when $EVWOI > EVWI$, one can always choose to not gather the information. VOI is an indicator of the maximal buying price or cost of an information-gathering activity. If the VOI is greater than the cost, the DM should gather the information; otherwise, the DM should not do so.

The decision without information (DWOI) is the alternative that optimizes EV over the prior, and the EVWOI is the optimal EV over the prior, i.e.,

$$\boldsymbol{a}_{pri}^{opt} = \underset{\boldsymbol{a}\in\boldsymbol{A}}{\operatorname{argmax}}\{EV[v(\boldsymbol{x},\boldsymbol{a})]\} = \underset{\boldsymbol{a}\in\boldsymbol{A}}{\operatorname{argmax}}\left\{\int_{\boldsymbol{x}} v(\boldsymbol{x},\boldsymbol{a})p(\boldsymbol{x})d\boldsymbol{x}\right\}, \quad (5.4)$$
$$EVWOI = EV\left[v\left(\boldsymbol{x},\boldsymbol{a}_{pri}^{opt}\right)\right] = \int_{\boldsymbol{x}} v\left(\boldsymbol{x},\boldsymbol{a}_{pri}^{opt}\right)p(\boldsymbol{x})d\boldsymbol{x}, \quad (5.5)$$

where $\boldsymbol{a}$ is an alternative from the decision space $\boldsymbol{A}$, $\boldsymbol{a}_{pri}^{opt}$ is the DWOI, $v(\boldsymbol{x},\boldsymbol{a})$ is the value function that assigns a value to each alternative-outcome pair for a given $\boldsymbol{x}$, and $p(\boldsymbol{x})$ is the prior probability distribution of $\boldsymbol{x}$. Similarly, for given observations, the decision with information (DWI) is the alternative that optimizes EV over the posterior, i.e.,

$$\begin{aligned}\boldsymbol{a}_{pos}^{opt}(\boldsymbol{y}) &= \underset{\boldsymbol{a}\in\boldsymbol{A}}{\operatorname{argmax}}\{EV[v(\boldsymbol{x},\boldsymbol{a})|\boldsymbol{y}]\} \\ &= \underset{\boldsymbol{a}\in\boldsymbol{A}}{\operatorname{argmax}}\left\{\int_{\boldsymbol{x}} v(\boldsymbol{x},\boldsymbol{a})p(\boldsymbol{x}|\boldsymbol{y})d\boldsymbol{x}\right\},\end{aligned} \tag{5.6}$$

$$\begin{aligned}EV_{pos}^{opt}(\boldsymbol{y}) &= EV\left[v\left(\boldsymbol{x},\boldsymbol{a}_{pos}^{opt}(\boldsymbol{y})\right)|\boldsymbol{y}\right] \\ &= \int_{\boldsymbol{x}} v\left(\boldsymbol{x},\boldsymbol{a}_{pos}^{opt}(\boldsymbol{y})\right)p(\boldsymbol{x}|\boldsymbol{y})d\boldsymbol{x},\end{aligned} \tag{5.7}$$

where $\boldsymbol{a}_{pos}^{opt}$ denotes the optimal decision policy, $\boldsymbol{a}_{pos}^{opt}(\boldsymbol{y})$ is the DWI given observations $\boldsymbol{y}$, $p(\boldsymbol{x}|\boldsymbol{y})$ is the posterior probability distribution, and EV_{pos}^{opt} is the optimal EV over the posterior. The posterior probability distribution is assessed by using Bayes' theorem.

Before a DM actually acquires information, the observations $\boldsymbol{y}$ is an uncertain quantity. Thus, the EVWI is defined as the expectation over the observations,

$$\begin{aligned}EVWI &= \int_{\boldsymbol{y}} EV_{pos}^{opt}(\boldsymbol{y})p(\boldsymbol{y})d\boldsymbol{y} \\ &= \int_{\boldsymbol{y}}\int_{\boldsymbol{x}} v\big(\boldsymbol{x},\boldsymbol{a}_{pos}^{opt}(\boldsymbol{y})\big)p(\boldsymbol{x}|\boldsymbol{y})d\boldsymbol{x}\,p(\boldsymbol{y})d\boldsymbol{y},\end{aligned} \tag{5.8}$$

or , equivalently

$$EVWI = \int_{(\boldsymbol{x},\boldsymbol{y})} v\big(\boldsymbol{x},\boldsymbol{a}_{pos}^{opt}(\boldsymbol{y})\big)p(\boldsymbol{x},\boldsymbol{y})d(\boldsymbol{x},\boldsymbol{y}), \tag{5.9}$$

where $(\boldsymbol{x},\boldsymbol{y})$ denotes a joint space of $\boldsymbol{x}$ and $\boldsymbol{y}$, and $\int_{(\boldsymbol{x},\boldsymbol{y})}\cdot\, d(\boldsymbol{x},\boldsymbol{y})$ stands for the integration over all combinations of $\boldsymbol{x}$ and $\boldsymbol{y}$. The rest of this thesis refers to Eq. 5.8 as Formulation 1 (F1) and Eq. 5.9 as Formulation 2 (F2). F1 is the standard formulation used in VOI's definition and is commonly used for VOI calculation because its order of integrals is consistent with the steps of solving a decision tree for the case with information. We define F2 in addition because it is relevant to the MC approach that we will discuss in Section 5.3.3.

Eqs. 5.6–5.9 do not show explicitly time-dependent components. Each component of the decision vector $\boldsymbol{a}_{pos}^{opt}(\boldsymbol{y})$ is a function of all the observations available at the time of evaluating the decision. The Bellman equation contains

a complete expression of Eq. 5.8 for a time-dependent system (Eidsvik et al. 2015, Bellman 1957).

5.2 VOI Calculation for Continuous Probability Distributions

An analytical solution for Eq. 5.8 does not exist for many cases involving continuous probability distributions. In these contexts, VOI calculation can be performed only using numerical approximation. Such methods include MCS or discretization rules (e.g., extended Pearson-Tukey, McNamee-Celona shortcut, or extended Swanson-Megill). Bickel et al. (2011) compared these approximation methods for cases where both the prior and likelihood functions were normal distributions, and they concluded that all the above-mentioned discretization rules result in significant errors in the VOI estimate and that numerous MC samples are required for an accurate VOI estimate.

The accuracy of VOI estimate depends on the decision characteristics resulting from the probability distributions and how well these are preserved using a heuristic discretization method or MCS.

5.2.1 High Resolution Probability Tree Discretization Method

We introduce a discretization method called the high resolution probability tree (HRPT) approach, which approximates a probability distribution by discretizing the space of an uncertain quantity into numerous outcomes. The following briefly presents this approach and its advantage. More details on the development and discussion of the HRPT approach can be found in **Paper IV**.

Figure 5.1 illustrates the discretization of a probability distribution $p(x)$ with 9 degrees[18] and its comparison to MC sampling with 9 samples. The continuous probability distribution is represented by the blue curve, the HRPT discretized outcomes by the orange dots, and the MC samples by the black rings. The shape of the continuous probability distribution is better preserved by the HRPT

[18] The number of degrees is referred to as the number of discrete outcomes, which is also the number of branches for an uncertainty node in a decision tree.

discretized outcomes than the MC samples because the HRPT discretized outcomes distribute evenly throughout the range of x, whereas most of the MC samples cluster around the mode. Better preservation of the shape of the continuous probability distribution can better preserve the decisions characteristics and in turn improve the accuracy of VOI estimate.

If the number of degrees for HRPT equals the number of MC samples, the computational time for VOI estimation is the same using HRPT as using MCS. Thus, using HRPT won't introduce additional computational time for VOI estimation, compared to MC method.

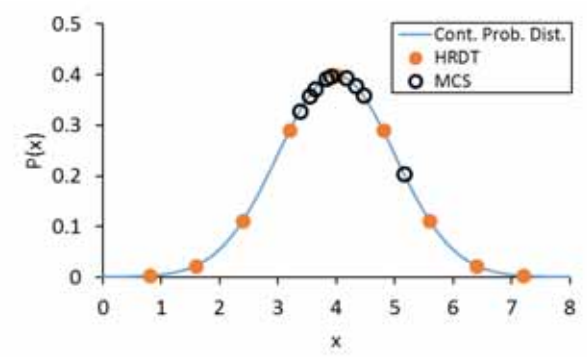


Figure 5.1—Illustration of HRPT discretized outcomes and MC samples for a continuous probability distribution.

5.2.2 Accuracy of HRPT

We investigate a two action linear-loss normal (TALL-N) problem to explore the accuracy of HRPT. The reason for using a TALL-N problem is that its VOI can be calculated analytically using the formula derived by Bickel (2008).

The TALL-N problem is as follows. A risk-neutral DM is considering whether or not to drill a well in an undeveloped area. The DM is uncertain about the NPV for the well and a priori characterizes it by a normal distribution with a mean of \$10 million and an SD of \$20 million. If the well is not drilled, the NPV is a sure \$0. To improve this decision, a seismic survey can be conducted at a cost. The DM uses a positive correlation coefficient ρ to describe the correlation between the true value of the well and the seismic results. We consider the value of ρ being any of 0.70, 0.71, 0.72, …, 1.00 because the VOI

error is significant when $\rho \geq 0.70$ for this example.[19] The VOI for the seismic survey needs to be calculated to support the decision of whether the DM should pay for conducting the seismic survey. For this TALL-N problem, the VOI has a closed-form solution, as presented by Bickel (2008). In order to assess how well HRPT approximates the analytical solution for different ρ values, we calculate the VOIs analytically and using HRPT with various degrees, and then calculate the error of the HRPT estimate relative to the analytical solution. The average relative VOI error is calculated by averaging over all the ρ values. **Figure 5.2** shows the average relative VOI error as a function of the number of degrees.

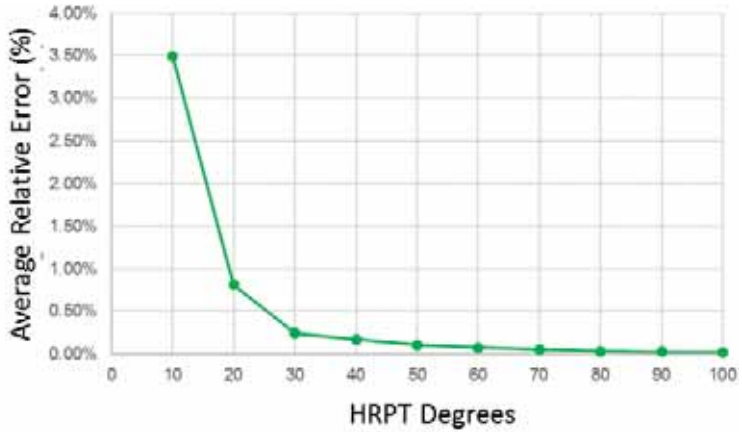


Figure 5.2—Average relative VOI error as a function of the number of degrees for the TALL-N problem.

In this figure, the average relative VOI error decreases (i.e., the accuracy of HRPT increases) as the number of degrees increases. The average relative VOI error is as small as 3.5% with only 10 degrees, and it is further reduced to 0.25% with only 30 degrees.

Paper IV investigates the accuracy of applying HRPT for VOI estimation to two other structures of prior distribution—PERT and lognormal distributions. These distributions are used extensively in the O&G industry.

[19] The value of ρ is allowed to be any from 0 to 1.

5.2.3 Comparison of HRPT and MC-based Methods

We compare the accuracy of using HRPT, Bayes Monte Carlo (BMC), and EnKF in estimating VOI for a TALL-N problem. Both of the latter two methods use MC sampling, but in different manners (see the appendices in **Paper IV** for detailed description of these two methods). This comparison provides an indication of how many MC samples are required to provide an equally good estimate as using HRPT.

The TALL-N problem for this comparison is as follows. A risk-neutral DM is considering whether to conduct a project: if the expected revenue of the project is positive, it will be conducted; otherwise, not conducted. The DM is uncertain about the revenue of the project and a priori characterizes it by a normal distribution with a mean of \$3 million and an SD of \$10 million. Information on the revenue can be obtained at a cost. The information has a Gaussian distributed error with zero mean and an SD of \$5 million. The VOI for this information needs to be calculated to support the decision of whether the DM should pay for this information.

The comparison is shown in **Figure 5.3**. To achieve a relative VOI error of 0.75%, HRPT requires 20 degrees, EnKF requires 80 ensemble members, and BMC requires more than 18000 MC samples. To achieve a relative VOI error of 0.5%, the number of degrees for HRPT would need to be increased only to 30, but the necessary ensemble size for EnKF would exceed 10000 and the required number of MC samples for BMC would far exceed 18000.[20] This indicates that HRPT is significantly more efficient than the other two methods in calculating the VOI for a TALL-N problem.

Although a closed-form solution exists for the TALL-N problem, this type of problem is constrained with Gaussian distributions, two actions, and linear value function. Thus, there is a need for more general solutions that would allow the VOI principle to be applied in real-world settings having any distributions,

[20] We did not calculate the relative VOI errors from using EnKF with more than 10000 samples or the BMC with more than 18000 samples because our computer did not have sufficient memory.

any number of actions, and any value function. A possible solution will be illustrated and discussed in the next section.

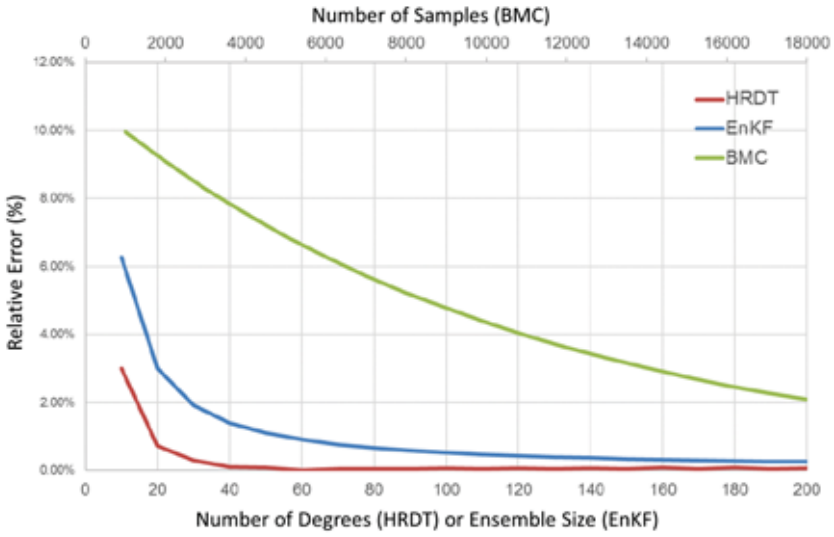


Figure 5.3—Relative error of VOI estimates using HRDT, BMC, and EnKF.

5.3 VOI Calculation for Reservoir Management

In a given RM context, there might be hundreds of relevant uncertain quantities. If we use HRPT with 30 degrees for each uncertain quantity, then having 100 uncertain quantities would mean 30^{100} combinations of discretized outcomes, once again raising the "curse of dimensionality." Thus, a different method is required for calculating VOI for many RM decision problems.

5.3.1 Relationship between Terms in VOI Analysis and Reservoir Management

Decision analysts and reservoir engineers use different terms in VOI analysis. Thus, it can be difficult for decision analysts and reservoir engineers to understand each other and communicate efficiently on decision problems. To alleviate this difficulty, **Table 5.1** gives the two disciplines' corresponding terms for common VOI concepts.

Notation	VOI Analysis	History Matching / Optimization	Ensemble-based Methods (e.g., EnKF and RO)
$\boldsymbol{x}$	Distinctions of Interest	Model Parameters	Model Parameters
$\tilde{\boldsymbol{y}}$	Observable Distinctions (without measurement noise)	Predicted Observations	Predicted Observations
$\boldsymbol{y}$	Observable Distinctions (with measurement noise)	Observations (with measurement noise) / Measurement Data	Observations (with measurement noise)
$p(\boldsymbol{x})$	Prior	Probability Distribution of Model Parameters	Prior Ensemble—a set of realizations from the prior
$p(\tilde{\boldsymbol{y}}\|\boldsymbol{x})$	Reliability of Model*	Probability Distribution of Predicted Observations Given Model Parameters	Model Noise (higher noise level implies lower reliability)
$p(\boldsymbol{y}\|\tilde{\boldsymbol{y}})$	Reliability of Measurements*	Probability Distribution of Observations Given Predicted Observations	Measurement Noise (higher noise level implies lower reliability)
$p(\boldsymbol{y}\|\boldsymbol{x})$	Likelihood	Probability Distribution of Observations Given Model Parameters	Assessed implicitly by assessing model and measurement noise
$p(\boldsymbol{y})$	Preposterior	Probability Distribution of Observations	Realizations generated by forward modelings plus realizations of model and measurement noise
$p(\boldsymbol{x}\|\boldsymbol{y})$	Posterior	Probability Distribution of Model Parameters Given Observations	Updated/Posterior Ensemble—a set of realizations from the posterior
	Bayesian Inference	Bayesian Calibration	Using e.g., EnKF
$\boldsymbol{a}$	Alternative	Control Variables	Control Variables
	Decision Making	Optimization	Using e.g., RO

*The term is not used in common VOI analysis. It is introduced here for the HM context.

Table 5.1—Relationship between terms used in VOI analysis and those in RM.

Assume a production model (e.g., an ECLIPSE reservoir simulation model) that formulates the dynamics of reservoirs is already given. Examples of the distinctions of interest in a RM context are the permeabilities in every grid block cell, and examples of the observable distinctions are the oil rates measured at different times. The priors are the probability distributions of the permeabilities based on current knowledge; the likelihoods are the probability distributions over the measured oil rates given the permeabilities; the preposteriors are the probability distributions over the measured oil rates; and the posteriors are the probability distributions over the permeabilities given the measured oil rates.

5.3.2 Decision-tree Example of VOI Calculation for Reservoir Management

We use a decision-tree example to illustrate the standard procedure of VOI analysis in an RM context.

We assume that a production model $f(\cdot)$ is given but that its parameters $\boldsymbol{x}$ (say, the permeability distribution) are uncertain and that once we have determined the values for these parameters, the model itself will correctly "predict" future production. We consider three realizations of the permeability distribution, denoted $\boldsymbol{x}_1$, $\boldsymbol{x}_2$ and $\boldsymbol{x}_3$. These realizations are equi-probable a priori. We are considering whether to measure and match the oil production rate at time t. The information provided by the measured data is used to support the decision on choosing one of three production strategies $\boldsymbol{a}_1$, $\boldsymbol{a}_2$, and $\boldsymbol{a}_3$ for time t forward. We want to estimate the value of the data measurement at time t.

The predicted oil production rates $\tilde{q}_o$ by the three realizations are shown in **Figure 5.4**, where $\tilde{q}_o^{0:T}$ is the oil production rate predicted for the period from time 0 to T, which is obtained by running the production model $f(\cdot)$ with parameters $\boldsymbol{x}_i$ and a production strategy $\boldsymbol{a}$. If the measurement is perfect, we will definitely measure a high rate $q_{o,H}^t$ given the permeability distribution of $\boldsymbol{x}_3$. Because of measurement errors, the probability of measuring a high rate $q_{o,H}^t$ is 4/5 and of measuring a low rate $q_{o,L}^t$ is 1/5 given $\boldsymbol{x}_3$. That is, the likelihood $p\left(q_{o,H}^t \middle| \boldsymbol{x}_3\right) = 4/5$ and $p\left(q_{o,L}^t \middle| \boldsymbol{x}_3\right) = 1/5$. The likelihood function is listed in **Table 5.2**.

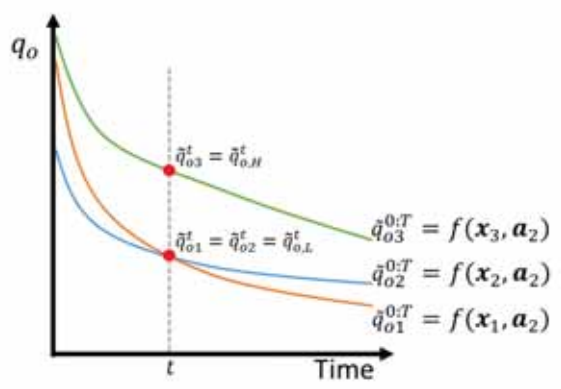


Figure 5.4—Oil production rate profiles of three realizations.

Likelihood Function p(**says**|**is**)		Given the set of model parameters **is** … $\boldsymbol{x}_1$	$\boldsymbol{x}_2$	$\boldsymbol{x}_3$
… the measurement	"Low" ($q^t_{o,L}$)	4/5	4/5	1/5
at time t **says**	"High" ($q^t_{o,H}$)	1/5	1/5	4/5
	Total	1	1	1

Table 5.2—Likelihood function for the decision tree example.

Figure 5.5a shows the uncertainty tree with the prior and likelihood probabilities in assessed form. Its corresponding inferential form (flipped tree) is shown in Figure 5.5b, where the preposterior and posterior probabilities are shown. These probabilities are calculated using Bayes' theorem, from the given prior and likelihood probabilities.

(a) Assessed Form: Information we have (b) Inferential Form: Information we need

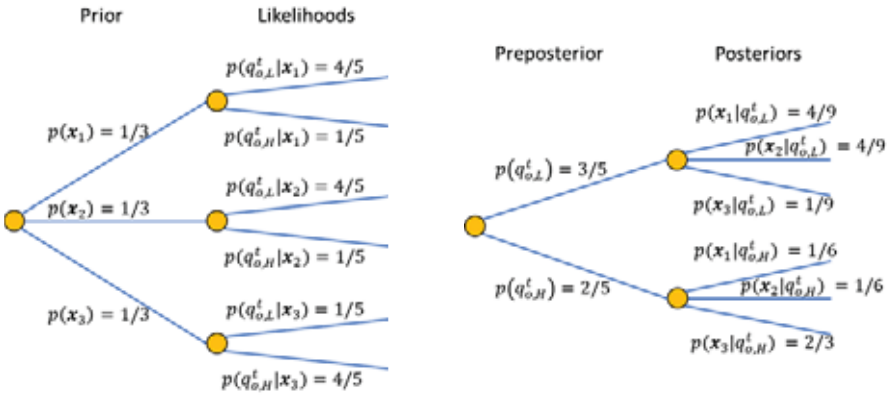


Figure 5.5—Uncertainty trees in (a) assessed form, and (b) inferential form.

The decision tree for the case with information, as illustrated in **Figure 5.6** where all monetary values are in million USD, is constructed by including decision and payoff nodes in the flipped uncertainty tree (Figure 5.5b). Solving the decision tree yields the optimal decision policy that if we have measured a low rate $q_{o,L}^t$, the optimal production strategy is $\boldsymbol{a}_1$; if we have measured a high rate $q_{o,H}^t$, the optimal production strategy is $\boldsymbol{a}_3$. This results in the optimal EV (i.e., EVWI) of \$70.8 million.

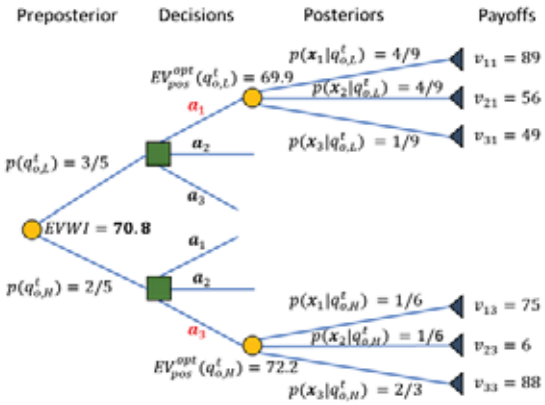


Figure 5.6—Decision tree for the case with information.

Figure 5.7 illustrates the decision tree for the case without information. The DWOI is $\boldsymbol{a}_2$ and EVWOI is \$67.3 million. The VOI is thus \$70.8 million – \$67.3 million = \$3.5 million. The information creates a value of \$3.5 million because it holds the possibility to change the decision when it becomes available. If the cost of acquiring the information is lower than \$3.5 million, we should conduct the measurement; otherwise, we should not.

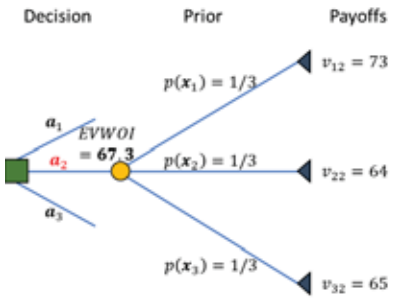


Figure 5.7—Decision tree for the case without information.

5.3.3 VOI Calculation Using Ensemble-based Methods for Reservoir Management

MC sampling is an easily accessible and powerful approach to handle continuous probability distributions and numerous uncertain quantities. MC-based methods, such as EnKF for Bayesian inference and EnOpt for RO (i.e., identifying the optimal alternative), can also be used to calculate VOI, as will be discussed below.

5.3.3.1 The BvHJ Approach

An approach of using EnKF and EnOpt for VOI calculation was proposed by Barros et al. (2016). We refer to this as the BvHJ approach.

The BvHJ approach is derived using common terminology in reservoir engineering. It is based on a twin experiment where the DWI and DWOI are valued on a synthetic truth and where the difference between the values corresponding to the DWI and DWOI is calculated. A synthetic truth in the BvHJ approach is a realization drawn from the prior probability distribution. n_e realizations are first drawn from the prior probability distribution, one of which is then chosen as a synthetic truth, and the rest form the prior ensemble with $n_e - 1$ realizations (Barros et al. 2016). The procedure of conducting a twin experiment is repeated for all n_e synthetic truths. For each synthetic truth, there is a corresponding difference between its values corresponding to the DWI and DWOI. The BvHJ approach then calculates the VOI as the EV over all these differences associated with synthetic truths. **Figure 5.8** shows the procedure of VOI calculation using the BvHJ approach.

In the BvHJ approach, the concept of synthetic truth is central to the analysis. However, "synthetic truth" has no meaning in VOI analysis. In reservoir engineering, a synthetically true model is used to mimic the actual state of a reservoir, by which reservoir engineers can investigate whether the actual outcome falls within the range that was predicted by updated models. This is not the case in VOI analysis, where the importance is to inform decisions, which in turn requires relevant and material uncertainties be quantified and analysis of how new information will update the distribution over the distinctions of interest and thus potentially affect the decisions. Moreover, VOI as defined by

Barros et al. (2016) implies a VOI distribution. However, as shown by Bratvold et al. (2009), the VOI is not a distribution but the difference between two expected values. Thus, the BvHJ approach is inconsistent with the definition of VOI analysis. However, two small modifications to the BvHJ approach can address this, as we will show later.

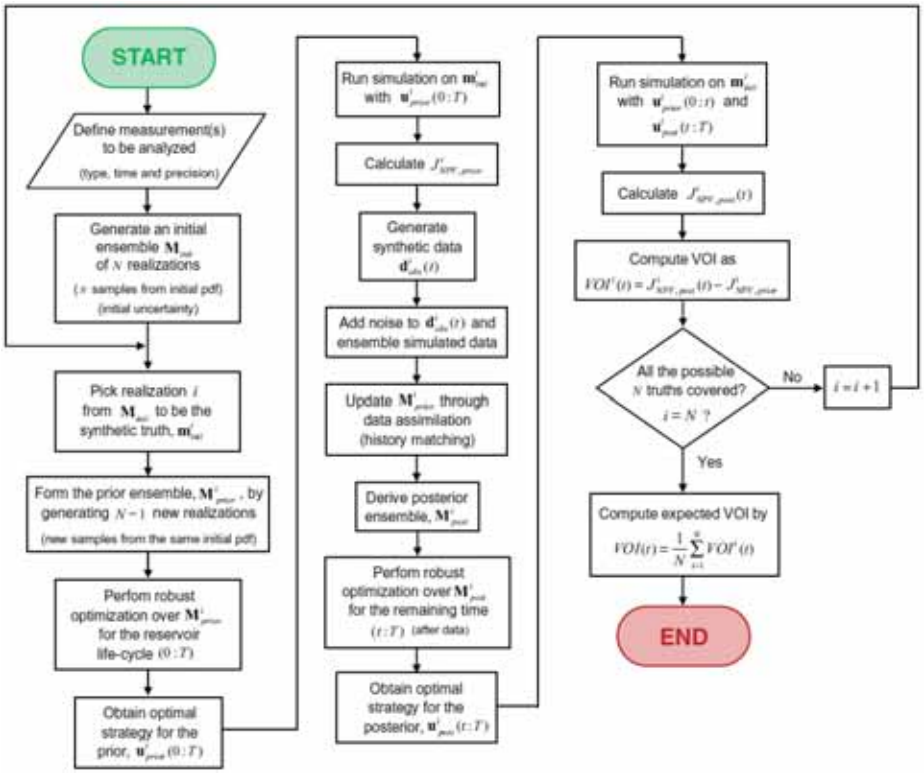


Figure 5.8—Procedure of VOI calculation using the BvHJ approach. Adapted from Barros et al. (2016).

5.3.3.2 Derivation of an Approach Consistent with VOI's Definition

Barros et al. (2016) did not elaborate on the connection between their approach and VOI analysis as defined in DA (Bratvold et al. 2009). We seek to accomplish this by deriving an approach using DA terminology and definitions.

Prior. n_e MC samples are drawn from the prior probability distribution $p(\boldsymbol{x})$ to form the prior ensemble [$\boldsymbol{x}_1$, $\boldsymbol{x}_2$, …, $\boldsymbol{x}_{n_e}$]. The probability of each sample $p(\boldsymbol{x}_i)$ is $1/n_e$. The prior ensemble is an MC representation of the prior probability distribution.

Likelihood. The likelihood function $p(\boldsymbol{y}|\boldsymbol{x})$ is assessed through a measurement noise distribution. For a dynamic system, we have the following mathematical relationship between the model parameters and measured data at time t_k:

$$\tilde{\boldsymbol{y}}_k = f(\boldsymbol{x}, \boldsymbol{a}_{0:k}), \quad (5.10)$$

$$\boldsymbol{y}_k = \tilde{\boldsymbol{y}}_k + \boldsymbol{\varepsilon}_{y,k}, \quad (5.11)$$

where the subscript k is the index of time, $f(\cdot)$ is a production model, and $\boldsymbol{\varepsilon}_{y,k}$ is a vector of the measurement noise at time t_k. When using EnKF together with a reservoir simulation model, the standard practice is to assign a multivariate Gaussian distribution to $\boldsymbol{\varepsilon}_{y,k}$ with zero mean and a diagonal covariance matrix $\boldsymbol{C}_{y,k}$ (i.e., $p(\boldsymbol{\varepsilon}_{y,k}|\tilde{\boldsymbol{y}}_k) \sim N(\boldsymbol{0}, \boldsymbol{C}_{y,k})$); this yields $p(\boldsymbol{\varepsilon}_{y,k}|\boldsymbol{x}) \sim N(f(\boldsymbol{x}, \boldsymbol{a}_{0:k}), \boldsymbol{C}_{y,k})$. Therefore, assessing a probability distribution for the measurement noise is equivalent to assessing the likelihood function.

Preposterior. The MC samples representing the preposterior probability distribution $p(\boldsymbol{y})$ can be generated as follows: (1) draw an MC sample $\boldsymbol{x}_i$ from $p(\boldsymbol{x})$ (this is done when the prior ensemble has been generated), (2) run the model $f(\boldsymbol{x}_i, \boldsymbol{a}_{0:k})$ to time t_k to obtain a realization of the predicted observations $\tilde{\boldsymbol{y}}_{k,i}$, (3) draw an MC sample $\boldsymbol{\varepsilon}_{y,k,i}$ from $p(\boldsymbol{\varepsilon}_{y,k}|\boldsymbol{x}_i)$, (4) add $\boldsymbol{\varepsilon}_{y,k,i}$ to $\tilde{\boldsymbol{y}}_{k,i}$ to obtain a sample of the observations with noise $\boldsymbol{y}_{k,i}$, and (5) repeat steps (1)–(3) to generate a set of $\boldsymbol{y}_k$, [$\boldsymbol{y}_{k,1}$, $\boldsymbol{y}_{k,2}$, …, $\boldsymbol{y}_{k,n_e}$], which is the preposterior ensemble representing the preposterior probability distribution. Moreover, the samples of ($\boldsymbol{x}$, $\boldsymbol{y}_k$)-pairs [($\boldsymbol{x}_1$, $\boldsymbol{y}_{k,1}$), ($\boldsymbol{x}_2$, $\boldsymbol{y}_{k,2}$), …, ($\boldsymbol{x}_{n_e}$, $\boldsymbol{y}_{k,n_e}$)] represent the joint probability distribution $p(\boldsymbol{x}, \boldsymbol{y}_k)$.

Posterior. The posterior ensemble [$\hat{\boldsymbol{x}}_1|\boldsymbol{y}_{k,i}$, $\hat{\boldsymbol{x}}_2|\boldsymbol{y}_{k,i}$, …,$\hat{\boldsymbol{x}}_{n_e}|\boldsymbol{y}_{k,i}$], representing the posterior probability distribution $p(\boldsymbol{x}|\boldsymbol{y}_{k,i})$, is generated by using EnKF to update the prior ensemble with the given observations $\boldsymbol{y}_{k,i}$ and assessed measurement noise $\boldsymbol{\varepsilon}_{y,k}$.

DWOI and EVWOI. We identity the DWOI and calculate the EVWOI using Eq. 5.12 (the MC form of Eq. 5.4) and Eq. 5.13 (the MC form of Eq. 5.5), respectively,

$$\boldsymbol{a}_{pri}^{opt} = \underset{\boldsymbol{a}\in\boldsymbol{A}}{\text{argmax}}\left\{\frac{1}{n_e}\sum_{i=1}^{n_e} v(\boldsymbol{x}_i, \boldsymbol{a})\right\}, \tag{5.12}$$

$$EVWOI = \frac{1}{n_e}\sum_{i=1}^{n_e} v\left(\boldsymbol{x}_i, \boldsymbol{a}_{pri}^{opt}\right). \tag{5.13}$$

DWI. We identify the DWI for the given observations using Eq. 5.14 (the MC form of Eq. 5.6),

$$\boldsymbol{a}_{pos}^{opt}(\boldsymbol{y}) = \underset{\boldsymbol{a}\in\boldsymbol{A}}{\text{argmax}}\left\{\frac{1}{n_e}\sum_{i=1}^{n_e} v(\hat{\boldsymbol{x}}_i|\boldsymbol{y}, \boldsymbol{a})\right\}. \tag{5.14}$$

EVWI. The EVWI can be calculated using either Eq. 5.15 (the MC form of F1 (Eq. 5.8)),

$$EVWI = \frac{1}{n_e}\sum_{j=1}^{n_e}\left[\frac{1}{n_e}\sum_{i=1}^{n_e} v\left(\hat{\boldsymbol{x}}_i|\boldsymbol{y}_j, \boldsymbol{a}_{pos}^{opt}(\boldsymbol{y}_j)\right)\right], \tag{5.15}$$

or Eq. 5.16 (the MC form of F2 (Eq. 5.9)),

$$EVWI = \frac{1}{n_e}\sum_{i=1}^{n_e} v\left(\boldsymbol{x}_i, \boldsymbol{a}_{pos}^{opt}(\boldsymbol{y}_i)\right). \tag{5.16}$$

VOI. The VOI is calculated using Eqs. 5.2 and 5.3.

A schematic of VOI calculation using MC-based methods is shown in **Figure 5.9**.

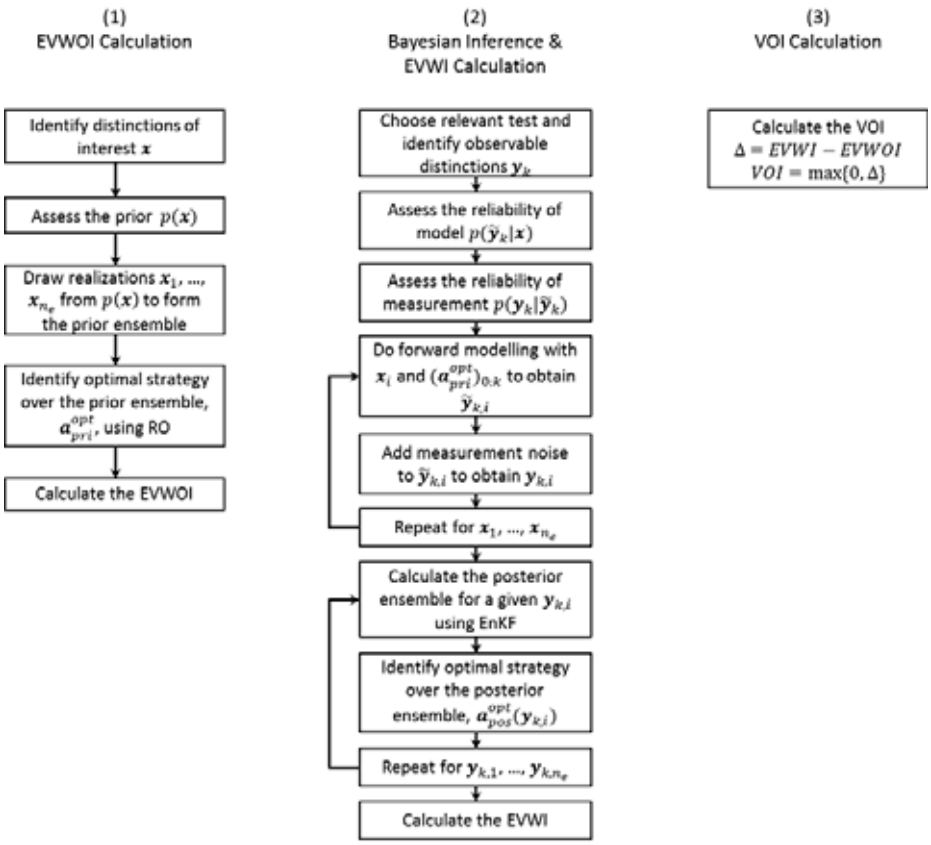


Figure 5.9—Schematic of VOI calculation using MC-based methods.

5.3.3.3 Comparison of the Approaches for VOI Calculation

The F2 approach is compared first with the BvHJ approach and then with the F1 approach.

F2 Approach vs. BvHJ Approach. There are two differences between these approaches. The first is that in the F2 approach, no realization is excluded from the prior ensemble; whereas in the BvHJ appraoch, a "synthetic truth" is excluded from the prior ensemble and the remaining ensemble members form a new prior ensemble. This difference leads to different VOI estimates as will be discussed later. The second difference is that the F2 approach calculates VOI as the difference between two EVs, whereas the BvHJ approach calculates VOI as an EV of a set of differences. Although the second difference does not affect

the VOI estimate, it constitutes a conceptual difference. The BvHJ approach with two minor modifications with respect to these two differences is identical to the F2 approach that is consistent with the VOI definition.

F2 Approach vs. F1 Approach. Based on the law of large numbers, these two approaches give the same limit when the ensemble size n_e approaches infinity. However, the convergence rate can be different, resulting in different estimates when the ensemble size is small because these two approaches introduce different levels of sampling error.

Accuracies of the Approaches. We use an example to investigate the accuracy of the VOI estimates of the three approaches.

Consider that the distinction of interest x is the revenue of a project and that the observable distinction y is the information on the revenue. The relationship of x, y, and the error of the information ε is $y = x + \varepsilon$. A priori, x is normally distributed with mean $\bar{x}$ and standard deviation σ_x, i.e., the prior $p(x) \sim N(\bar{x}, \sigma_x)$. ε is normally distributed with zero mean and standard deviation σ_ε, i.e., the likelihood function $p(y|x) \sim N(x, \sigma_\varepsilon)$. The decision a to be made is whether the project will be conducted; $a = 1$ and $a = 0$ correspond to "conduct" and "not conduct," respectively. If the expected revenue is positive, the project should be conducted; otherwise, it should not. Thus, the value function is $v(x, a) = ax$. With $\bar{x} = 3000$, $\sigma_x = 10000$ and $\sigma_\varepsilon = 5000$ (the monetary unit can be arbitrary here), the VOI for this problem setting can be calculated analytically, which is 2267.

We first test the three approaches using a large ensemble size: 10000. The VOI estimates are listed in **Table 5.3**. The BvHJ approach gives a result almost identical to that of the F2 approach because the impact of excluding a "synthetic truth" (i.e., an x_i) from the prior ensemble is very small when the ensemble size is large. The VOIs calculated using the F1 and F2 approaches differ slightly. All three methods estimate the VOI with an error smaller than 1%.

Approach	Analytical	BvHJ	F1	F2
VOI	2267	2283	2246	2283
Error	-	0.71%	0.93%	0.71%

Table 5.3—VOI estimates using BvHJ, F1, and F2 approaches with ensemble size of 10000.

Then, a small ensemble size, 50, is used. The VOI calculation is repeated 10000 times for each approach. The statistics of the VOI estimates are listed in **Table 5.4**, and the probability density functions (PDFs) are plotted in **Figure 5.10**.[21] Among the three approaches, the F2 approach is the most accurate, as it gives the best VOI estimate on average and has the smallest SD in its estimates. More detail on why the F2 approach is better is given in **Paper V**.

Approach	BvHJ	F1	F2
Average	2323	2257	2259
Error	2.47%	0.44%	0.35%
SD	933	776	695
[Min, Max]	[0, 10421]	[0, 5497]	[0, 5279]

Table 5.4—Statistics of VOI estimates of BvHJ, F1, and F2 approaches with ensemble size of 50.

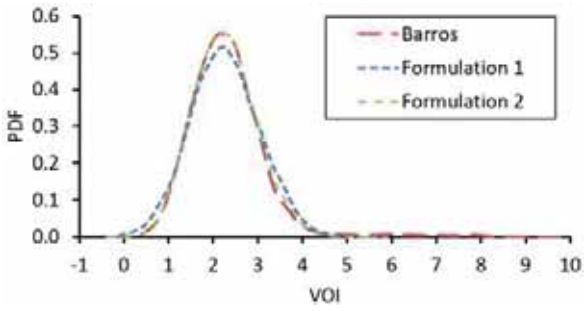


Figure 5.10—PDFs of VOI estimates of BvHJ, F1, and F2 approaches with ensemble size of 50.

This example has demonstrated that the three approaches give almost equivalent results for large ensembles. However, they can lead to quite different results when the ensemble size is small. As there are no significant

[21] Here, the uncertainty in the VOI estimate is due to the sampling error associated with MC methods. As discussed earlier, the VOI itself is not a distribution but the difference between two expected values.

computational increases resulting from using the F2 approach for VOI calculation with MC methods, this would be our approach of choice.

5.3.3.4 Example of Applying the F2 Approach with a Reservoir Simulation Model

We briefly show an example of applying the F2 approach to estimate VOI for a more realistic case, where a reservoir simulation model is involved. For more details on the model and problem setting, please refer to **Paper V**.

VOI analysis is used to analyze whether devices should be placed for acquiring saturation data, in addition to oil and water production rate data. The measured data are used to support the decision on polymer concentration in the enhanced-oil-recovery (EOR) phase.

The VOI is calculated using the F2 approach combined with EnKF and EnOpt. The resulting VOIs for rate data only, saturation data only, and both rate and saturation data, and VOPI are listed in **Table 5.5**. The values of rate data and saturation data are very close to each other. Although the saturation data has a value of $\$3.36\cdot10^5$ when it is considered individually, it adds a value of only $\$1.01\cdot10^5$ ($= \$4.11\cdot10^5 - \$3.10\cdot10^5$) to that of the rate data. This indicates that VOI is not additive (Samson et al. 1989). The VOI analysis informs us that we should not gather saturation data in addition to rate data if the cost of placing saturation measuring devices is greater than $\$1.01\cdot10^5$. The VOPI is $\$6.12\cdot10^5$, which is $\$6.12\cdot10^5 - \$3.10\cdot10^5 = \$3.02\cdot10^5$ more than the value of rate data. Therefore, we should decline any other data gathering activity that costs more than $\$3.02\cdot10^5$ given that the rate data will be gathered.

Data Type	Rate Only	Saturation Only	Rate and Saturation	Perfect Information
VOI [10^5 USD]	3.10	3.36	4.11	6.12

Table 5.5—VOI estimates for the example with a reservoir simulation model.

6 Overview of Research Papers

Paper I: Integrating Model Uncertainty in Probabilistic Decline Curve Analysis for Unconventional Oil Production Forecasting

This paper introduces an approach to integrating model uncertainty in probabilistic decline curve analysis. The aim is to assess the probabilities of different decline curve models and use these probabilities for further analysis instead of identifying and using a "best" model for decline curve analysis. We focus on unconventional oil production because the widely used Arps decline curve model might not be ideal for unconventional play; thus, several novel decline curve models have been proposed. Examples are used to illustrate the impact of the integration of model uncertainty. The proposed approach is applied to real data. We conclude that using our proposed approach can reduce the risk of using a best fit but very wrong model for forecasting and that our proposed approach performs well in propagating the model uncertainty to the uncertainty in forecast.

Paper II: Robust Production Optimization with Capacitance-Resistance Model as Precursor

This paper presents a proxy-model workflow where a grid-based model is supplemented by a useful yet tractable proxy model for speeding up the process of RO. Specifically, we review CRMs as potential proxy models for waterflooding systems. A selected CRM is embedded into the proxy-model workflow. We illustrate the use of a CRM and investigate its pros and cons using synthetic 2D and 3D models. The results obtained from the proxy-model and traditional workflows are compared. The impact of any differences is assessed, and the value of using a proxy model is quantified by considering a PO-relevant decision-making context. We discuss the desiderata of proxy models. We conclude that CRMs have high potential to serve as a cogent proxy model for waterflooding-related decision-making contexts and that the proxy-model workflow, leveraging a faster but relevant production model, significantly speeds up the optimization yet gives robust results that lead to a near-optimal solution.

Paper III: Fast Analysis of Optimal IOR Start Time Using a Two-Factor Production Model and Least-Squares Monte Carlo Algorithm

This paper presents an approach of using a two-factor production model together with the LSM algorithm for fast analysis of optimal IOR start time. We review the two-factor production model introduced by Parra-Sanchez (2010). This production model is computationally more attractive than a more complex model, such as a reservoir simulation model. Thus, using it can significantly speed up the uncertainty analysis where a large number of production forecast runs is required. Identifying the optimal IOR start time using LSM allows for learning over time. Unlike the myopic decision policy, which considers only the current information, LSM considers both the current information and the information that will be obtained in the future. Therefore, the LSM solves approximately for the global optimum of an FSRM problem. The impact of allowing learning over time in decision making is quantified using the VOI framework. An example illustrates the application of the proposed approach. We conclude that combining the two-factor model with LSM facilitates a fast analysis of optimal IOR start time and that it can create significant value to include learning over time in decision making.

Paper IV: Robust Discretization of Continuous Probability Distributions for Value-of-Information Analysis

This paper introduces a practical, flexible, efficient, and very general discretization method for continuous probability distributions involved in VOI analysis. This discretization method is validated by comparing its VOI estimate to the analytical solutions. Commonly used discretization methods for VOI calculation are reviewed and compared to the proposed method. In addition, we investigate two MC-based methods for VOI calculation. We conclude that the proposed discretization method estimates VOI very accurately (with a relative error of around 0.25% using a degree of discretization of only 30), that the commonly used discretization methods generally give poor VOI estimates, and that more than thousands of MC samples are required for the MC-based methods to provide accurate VOI estimates.

Paper V: Value-of-Information for Model Parameter Updating through History Matching

This paper presents a consistent, DA-based VOI analysis framework to assess the value of HM. It is consistent and DA-based because it is derived from VOI's definition originally developed in the DA community. In order to make the VOI framework understandable and accessible to both the RM and DA communities, we bridge the nomenclature and terminology used in VOI calculations and that used in state-of-the-art HM and PO methods. In particular, our framework uses EnKF for Bayesian inference and EnOpt for RO. We discuss two formulations (F1 and F2) of VOI calculation. In addition, we investigate the similarities and differences between the F2 approach and the BvHJ approach. Several examples are used to illustrate and discuss the use of this framework in an HM context. We conclude that the BvHJ approach with two minor modifications will be identical to the F2 approach, and that the F2 approach gives the most accurate VOI estimate among the three approaches investigated.

7 Summary and Conclusions

RM involves a process of gathering data and information to inform and support decision making. HM is used to extract information from measured data. The information is then used to support the decisions on production strategies through an optimization process. Thus, RM is a decision-oriented activity, for which a DA framework will add value.

The major goal of DA is to make good decisions. This requires the quantification of relevant and material uncertainties, a consistent way of updating uncertainties when new information arrives, and a means to assess the value of such information. We referred to this combination of activities as "decision-focused uncertainty management." [22] This work illustrated and discussed a decision-focused uncertainty-management framework for RM using DA tools, with a focus on geological uncertainty.

We first investigated how geological uncertainty should be managed for the goal of making good decisions in HM and PO separately, and then investigated how it should be managed in RM as a loop of HM and PO. We finally investigated how to incorporate a powerful decision analysis tool—VOI analysis—into RM.

To manage geological uncertainty in HM, HM should be conducted probabilistically rather than deterministically. Probabilistic HM approaches use probability distributions to quantify uncertainties. Probabilistic HM can be achieved by simply repeating deterministic HM in an MC manner, such as the Bootstrap method (Section 2.1.2.2) or the P-MLE method (Section 2.1.2.3). When normal distributions are used to represent a DM's uncertainties, EnKF is a proven method for Bayesian HM. It can efficiently update a production model with numerous uncertain parameters. To integrate model uncertainty in decline curve analysis, we proposed an approach to assess the probabilities of various decline curve models by repeating the P-MLE method for each decline curve

[22] For brevity, the phrase "decision-focused uncertainty management" was simply referred to as "uncertainty management" throughout this thesis.

model, and concluded that considering model uncertainty gives a more complete estimate of uncertainty in a production forecast.

Geological uncertainty in PO can be managed by optimizing the EV over geological realizations if the DM is risk-neutral. This is known as RO. However, the challenge is that when a traditional optimization algorithm is used, to optimize the EV over numerous geological realizations is much more computationally intensive than to optimize the value of a single realization. EnOpt is designed to address this challenge. It is based on the traditional steepest ascent method, but it uses a more efficient approach to approximate the gradient. Although EnOpt has significantly sped up RO, the computational time can still be long when a reservoir simulation model is involved. Thus, we proposed a proxy-model workflow where a CRM is used to supplement a reservoir simulation model for RO. The results showed that the proxy-model workflow gives a near-optimal solution with significantly reduced computational time.

CLRM is a state-of-the-art approach to managing geological uncertainty in RM. It closes the loop of HM and PO by continuously updating a production model and performing life-cycle optimization whenever new data become available. However, it is based on a myopic decision policy, in that it does not account for future uncertainty revelations and their impacts on future decisions. Thus, the CLRM solution can be suboptimal. Since RM is a sequential decision-making problem, it should consider whether geological uncertainty informed by current data or by future information. We demonstrated the full structure of a dynamic programming approach to RM decision making using decision trees. FSRM has a more complex structure and requires longer computational time than does CLRM. However, the optimal solution of FSRM is a global optimum. The LSM algorithm is very efficient for solving an FSRM problem with relatively small numbers of decisions and of decision alternatives. It does not suffer from the "curse of dimensionality" as the number of uncertainties increases. We illustrated this by an example of using LSM combined with a two-factor model to identify the optimal IOR start time.

VOI analysis assesses the value of additional information before it is obtained. The assessed value is used to support the decision on whether to buy the additional information. We proposed the HRPT method to discretize

continuous probability distributions used in VOI calculation. It was compared with other common discretization methods and MC-based methods. The comparison showed that HRPT provides significantly more accurate VOI estimates than do the other methods.

Although the concept of VOI has been illustrated and discussed in an RM context by other research, there is a gap between the terms used in the DA and RM communities for VOI analysis. This may make the communication between these two communities difficult. To reduce the gap, we related the RM terminology to the VOI terminology commonly used in DA and derived a general VOI analysis framework for RM, based on VOI's definition. We compared the VOIs estimated using different formulations. We found that although these formulations give the same analytical results, the results from their numerical implementations can be different when MC methods are used. The results showed that the F2 approach provides a more accurate VOI calculation.

This thesis used several examples to illustrate insights and guidance for managing geological uncertainty with DA tools in RM. Challenges remain in practice that might prevent the development of a broad and deep understanding and routine application of DA. Nevertheless, we are optimistic about the value of applying DA in RM and the entire O&G industry.

8 Discussion and Future Research

Our work has shown the value of implementing DA to manage geological uncertainty in RM. We propose the following areas for future research.

As engineers and geoscientists, we have been taught a modeling philosophy that emphasizes detailed deterministic realism. However, as the main purpose of a model used in an RM context is to inform and support decisions, we need to evaluate whether any model is useful and tractable for that purpose. Building in a lot of details into our production models (e.g., building reservoir simulation models) might not serve this purpose. Less detailed models that allow us to consistently capture our uncertainties will often do a better job in informing and supporting decision making. Details can then be added to the model if they are deemed to be relevant and material for the decisions at hand.

Although the current state-of-the-art production models are suitable for some types of RM decision, they often include detailed features that are not relevant and material for other decisions. This makes decision analysis for RM computationally intensive or even prohibitive. More focus should be on production models that are both useful and tractable for decision making. CRM is a cogent model for the context of deciding water injection rates for waterflooding. It should be extended to include more features (e.g., adding new wells and chemical flooding) so that other RM-relevant decision contexts can benefit from its speed.

CLRM, as the state-of-the-art RM tool, is based on a myopic decision policy and thus can result in suboptimal solutions. Because the computational requirement for solving a CLRM problem is significantly less than that for an FSRM problem, CLRM might be regarded as a simplification of FSRM. However, it is never a substitute for FSRM. Future research should focus on investigating the differences between CLRM and FSRM solutions and on more efficient methods for solving an FSRM problem.

Since data are gathered for informing decisions on production strategies, HM should be decision-driven rather than data-driven. “Data-driven” means that we first gather data and then consider what decision the data can inform. In contrast, “decision-driven” means we first identify the decisions we are going

to make and then consider what data should be gathered to inform these decisions. The advantage of decision-driven HM is that we can avoid spending money on valueless data gathering activities. Thus, future research should focus on decision-driven HM using a VOI analysis framework.

Our work considers only one decision point in VOI analysis for RM. Although Barros et al. (2015) extended the BvHJ approach to account for sequential decisions, they calculate the exclusive VOI for CLRM rather than the general VOI for FSRM. Because CLRM might result in suboptimal solutions as discussed earlier, the VOI for CLRM may be smaller than the VOI for FSRM (the VOI defined in the DA community). Focus should be on efficient methods to calculate the VOI for FSRM.

The present work focused on geological uncertainty. Future research should also consider economic uncertainty (e.g., the uncertainty in oil price) in RM.

References

Aanonsen, S.I., Nævdal, G., Oliver, D.S. et al. 2009. The Ensemble Kalman Filter in Reservoir Engineering—a Review. *SPE Journal* **14** (3): 393-412. SPE-117274-PA. https://doi.org/10.2118/117274-PA.

Arps, J.J. 1945. Analysis of Decline Curves. *Transactions of the AIME* **160** (1): 228-247. SPE-945228-G. https://doi.org/10.2118/945228-G.

Barker, J.W., Cuypers, M., Holden, L. 2000. Quantifying Uncertainty in Production Forecasts: Another Look at the PUNQ-S3 Problem. Presented at the SPE Annual Technical Conference and Exhibition, Dallas, USA, 1-4 October. SPE-62925-MS. https://doi.org/10.2118/62925-MS.

Barros, E., Leeuwenburgh, O., Van den Hof, P. et al. 2015. Value of Multiple Production Measurements and Water Front Tracking in Closed-Loop Reservoir Management. Presented at the SPE Reservoir Characterisation and Simulation Conference and Exhibition, Abu Dhabi, UAE, 14-16 September. SPE-175608-MS. https://doi.org/10.2118/175608-MS.

Barros, E., Van den Hof, P., Jansen, J. 2016. Value of Information in Closed-Loop Reservoir Management. *Computational Geosciences* **20** (3): 737-749. https://doi.org/10.1007/s10596-015-9509-4.

Begg, S.H., Welsh, M.B., Bratvold, R.B. 2014. Uncertainty vs. Variability: What's the Difference and Why is it Important? Presented at the SPE Hydrocarbon Economics and Evaluation Symposium, Houston, USA, 19-20 May. SPE-169850-MS. https://doi.org/10.2118/169850-MS.

Bellman, R.E. 1957. *Dynamic Programming*. New Jersey, USA: Princeton University Press.

Bickel, J.E. 2008. The Relationship Between Perfect and Imperfect Information in a Two-Action Risk-Sensitive Problem. *Decision Analysis* **5** (3): 116-128.

Bickel, J.E., Bratvold, R.B. 2008. From Uncertainty Quantification to Decision Making in the Oil and Gas Industry. *Energy Exploration & Exploitation* **26** (5): 311-325. https://doi.org/10.1260/014459808787945344.

Bickel, J.E., Lake, L.W., Lehman, J. 2011. Discretization, Simulation, and Swanson's (Inaccurate) Mean. *SPE Economics & Management* **3** (3): 2150-1173. SPE-148542-PA. https://doi.org/10.2118/148542-PA.

Box, G.E. 1979. Robustness in the Strategy of Scientific Model Building. In *Robustness in Statistics*, 201-236. New York, USA: Academic Press.

Bratvold, R.B., Begg, S. 2009. Would You Know a Good Decision if You Saw One? *The Way Ahead* **5** (2): 21-23. SPE-0209-021-TWA. https://doi.org/10.2118/0209-021-TWA.

Bratvold, R.B., Begg, S. 2010. *Making Good Decisions*. Texas, USA: Society of Petroleum Engineers.

Bratvold, R.B., Bickel, J.E., Lohne, H.P. 2009. Value of Information in the Oil and Gas Industry: Past, Present, and Future. *SPE Reservoir Evaluation & Engineering* **12** (4): 630-638. SPE-110378-PA. https://doi.org/10.2118/110378-PA.

Brouwer, D.R., Nævdal, G., Jansen, J.D. et al. 2004. Improved Reservoir Management Through Optimal Control and Continuous Model Updating. Presented at the SPE Annual Technical Conference and Exhibition, Houston, Texas, 26-29 September. SPE-90149-MS. https://doi.org/10.2118/90149-MS.

Burgers, G., Jan van Leeuwen, P., Evensen, G. 1998. Analysis Scheme in the Ensemble Kalman Filter. *Monthly Weather Review* **126** (6): 1719-1724. https://doi.org/10.1175/1520-0493(1998)126<1719:ASITEK>2.0.CO;2.

Cao, F. 2014. Development of a Two-Phase Flow Coupled Capacitance Resistance Model. PhD Thesis, University of Texas at Austin, Austin, USA.

Cao, F., Luo, H., Lake, L.W. 2015. Oil-Rate Forecast by Inferring Fractional-Flow Models From Field Data With Koval Method Combined With the Capacitance-Resistance Model. *SPE Reservoir Evaluation & Engineering* **18** (4): 534-553. SPE-173315-PA. https://doi.org/10.2118/173315-PA.

Chen, Y., Oliver, D.S., Zhang, D. 2009. Efficient Ensemble-Based Closed-Loop Production Optimization. *SPE Journal* **14** (4): 634-645. SPE-112873-PA. https://doi.org/10.2118/112873-PA.

Cheng, Y., Wang, Y., McVay, D. et al. 2010. Practical Application of a Probabilistic Approach to Estimate Reserves Using Production Decline Data. *SPE Economics & Management* **2** (1): 19-31. SPE-95974-PA. https://doi.org/10.2118/95974-PA.

Clark, A.J., Lake, L.W., Patzek, T.W. 2011. Production Forecasting with Logistic Growth Models. Presented at the SPE Annual Technical Conference and Exhibition, Denver, USA, 30 October-2 November. SPE-144790-MS. https://doi.org/10.2118/144790-MS.

Clarkson, C.R., McGovern, J.M. 2005. Optimization of CBM Reservoir Exploration and Development Strategies through Integration of Simulation and Economics. *SPE Reservoir Evaluation & Engineering* **8** (6): 502-519. SPE-88843-PA. https://doi.org/10.2118/88843-PA.

Do, S.T., Reynolds, A.C. 2013. Theoretical Connections between Optimization Algorithms based on an Approximate Gradient. *Computational Geosciences* **17** (6): 959-973. https://doi.org/10.1007/s10596-013-9368-9.

Duong, A.N. 2011. Rate-Decline Analysis for Fracture-Dominated Shale Reservoirs. *SPE Reservoir Evaluation & Engineering* **14** (3): 377-387. SPE-137748-PA. https://doi.org/10.2118/137748-PA.

ECLIPSE. 2016. http://www.software.slb.com/products/eclipse (accessed 12 October 2016).

Eidsvik, J., Mukerji, T., Bhattacharjya, D. 2015. *Value of Information in the Earth Sciences: Integrating Spatial Modeling and Decision Analysis*. Cambridge, UK: Cambridge University Press.

Eppel, T., von Winterfeldt, D. 2008. Value-of-Information Analysis for Nuclear Waste Storage Tanks. *Decision Analysis* **5** (3): 157-167.

Evensen, G. 1994. Sequential Data Assimilation with a Nonlinear Quasi-Geostrophic Model Using Monte Carlo Methods to Forecast Error Statistics. *Journal of Geophysical Research: Oceans* **99** (C5): 10143-10162. https://doi.org/10.1029/94JC00572.

Evensen, G. 2009. *Data Assimilation: The Ensemble Kalman Filter*. Berlin, Germany: Springer Science & Business Media.

Floris, F.J.T., Bush, M.D., Cuypers, M. et al. 2001. Methods for Quantifying the Uncertainty of Production Forecasts: a Comparative Study. *Petroleum Geoscience* **7**: S87-S96. https://doi.org/10.1144/petgeo.7.S.S87.

Fonseca, R.M., Stordal, A.S., Leeuwenburgh, O. et al. 2014. Robust Ensemble-based Multi-Objective Optimization. Presented at the ECMOR XIV-14th European Conference on the Mathematics of Oil Recovery Catania, Italy, 8-11 September. https://doi.org/10.3997/2214-4609.20141895

Gentil, P.H. 2005. The Use of Multilinear Regression Models in Patterned Waterfloods: Physical Meaning of the Regression Coefficients. Master's Thesis, University of Texas at Austin, Austin, USA.

Grayson, C.J. 1960. *Decisions under Uncertainty: Drilling Decisions by Oil and Gas Operators*. Cambridge, USA: Graduate School of Business, Harvard University.

Hajizadeh, Y., Christie, M.A., Demyanov, V. 2010. Comparative Study of Novel Population-based optimization Algorithms for History Matching and Uncertainty Quantification: PUNQ-S3 Revisited. Presented at the Abu Dhabi International Petroleum Exhibition and Conference, Abu Dhabi, UAE, 1-4 November. SPE-136861-MS. https://doi.org/10.2118/136861-MS.

Haldorsen, H.H., Van Golf-Racht, T. 1989. Reservoir Management into the Next Century. Presented at the New Mexico Tech Centennial Symposium, Socorro, USA, 16-19 October.

Hastie, T., Tibshirani, R., Friedman, J. 2009. *The Elements of Statistical Learning : Data Mining, Inference, and Prediction*. New York, USA: Springer.

He, J., Sarma, P., Bhark, E. et al. 2017. Quantifying Value of Information Using Ensemble Variance Analysis. Presented at the SPE Reservoir Simulation Conference, Montgomery, USA, 20-22 February. SPE-182609-MS. https://doi.org/10.2118/182609-MS.

Howard, R.A. 1966. Information Value Theory. *IEEE Transactions on Systems Science and Cybernetics* **2** (1): 22-26. https://doi.org/10.1109/TSSC.1966.300074.

Ilk, D., Rushing, J.A., Perego, A.D. et al. 2008. Exponential vs. Hyperbolic Decline in Tight Gas Sands: Understanding the Origin and Implications for Reserve Estimates Using Arps' Decline Curves. Presented at the SPE Annual Technical Conference and Exhibition, Denver, USA, 21-24 September. SPE-116731-MS. https://doi.org/10.2118/116731-MS.

Jannink, J.W., Bos, C.F.M. 2005. Probabilistic Discharge Forecasting for Improved Asset Investment Decision Support (SPE94116). Presented at the 67th EAGE Conference & Exhibition, Madrid, Spain, 13-16 June.

Jansen, J.-D., Brouwer, R., Douma, S.G. 2009. Closed Loop Reservoir Management. Presented at the SPE Reservoir Simulation Symposium, The Woodlands, Texas, 2-4 February. SPE-119098-MS. https://doi.org/10.2118/119098-MS.

Jochen, V.A., Spivey, J.P. 1996. Probabilistic Reserves Estimation Using Decline Curve Analysis with the Bootstrap Method. Presented at the SPE Annual Technical Conference and Exhibition, Denver, USA, 6-9 October. SPE-36633-MS. https://doi.org/10.2118/36633-MS.

Jonkman, R.M., Bos, C.F.M., Morgan, D.T.K. et al. 2010. Best Practices and Methods in Hydrocarbon Resource Estimation, Production and Emissions Forecasting, Uncertainty Evaluation, and Decision Making. *SPE Reservoir Evaluation & Engineering* **5** (2): 146-153. SPE-77280-PA. https://doi.org/10.2118/77280-PA.

Joshi, K., Lee, W.J. 2013. Comparison of Various Deterministic Forecasting Techniques in Shale Gas Reservoirs. Presented at the SPE Hydraulic Fracturing Technology Conference, The Woodlands, USA, 4-6 February. SPE-163870-MS. https://doi.org/10.2118/163870-MS.

Kalman, R.E. 1960. A New Approach to Linear Filtering and Prediction Problems. *Journal of Basic Engineering* **82** (1): 35-45. https://doi.org/10.1115/1.3662552.

Koval, E.J. 1963. A Method for Predicting the Performance of Unstable Miscible Displacement in Heterogeneous Media. *Society of Petroleum Engineers Journal* **3** (2): 145-154. SPE-450-PA. https://doi.org/10.2118/450-PA.

Kullawan, K. 2016. A Bayesian Framework for Real-Time Optimization of Well Placement. PhD Thesis, University of Stavanger, Stavanger, Norway.

Longstaff, F.A., Schwartz, E.S. 2001. Valuing American Options by Simulation: A Simple Least-Squares Approach. *The Review of Financial Studies* **14** (1): 113-147. https://doi.org/10.1093/rfs/14.1.113.

Lorentzen, R.J., Fjelde, K.K., Frøyen, J. et al. 2001. Underbalanced and Low-head Drilling Operations: Real Time Interpretation of Measured Data and Operational Support. Presented at the SPE Annual Technical Conference and Exhibition, New Orleans, USA, 30 September-3 October. SPE-71384-MS. https://doi.org/10.2118/71384-MS.

MATLAB. 2016. http://se.mathworks.com/products/matlab (accessed 12 October 2016).

Morgan, M.D. 2005. A Model of Canadian Oil and Gas Price Fluctuations. *Journal of Canadian Petroleum Technology* **44** (7). PETSOC-05-07-03. https://doi.org/10.2118/05-07-03.

Nævdal, G., Brouwer, D.R., Jansen, J.-D. 2006. Waterflooding Using Closed-Loop Control. *Computational Geosciences* **10** (1): 37-60. https://doi.org/10.1007/s10596-005-9010-6.

Pan, Z. 2016. Revised Productivity Index Equation to Improve Transient History Match for the Capacitance Resistance Model. Master's Thesis, University of Texas at Austin, Austin, USA.

Parra-Sanchez, C. 2010. A Life Cycle Optimization Approach to Hydrocarbon Recovery. Master Thesis, University of Texas at Austin, Austin, USA.

Raiffa, H., Schlaifer, R. 1961. *Applied Statistical Decision Theory*. New York, USA: Wiley.

Robertson, J.D. 1989. Reservoir Management Using 3D Seismic Data. *Journal of Petroleum Technology* **41** (7): 663-667. SPE-19887-PA. https://doi.org/10.2118/19887-PA.

Satter, A., Varnon, J.E., Hoang, M.T. 1994. Integrated Reservoir Management. *Journal of Petroleum Technology* **46** (12): 1057-1064. SPE-22350-PA. https://doi.org/10.2118/22350-PA.

Sayarpour, M. 2008. Development and Application of Capacitance-Resistive Models to Water/CO2 Floods. PhD Thesis, University of Texas at Austin, Austin, USA.

Sayarpour, M., Kabir, C., Sepehrnoori, K. et al. 2011. Probabilistic History Matching with the Capacitance–Resistance Model in Waterfloods: a Precursor to Numerical Modeling. *Journal of Petroleum Science and Engineering* **78** (1): 96-108. https://doi.org/10.1016/j.petrol.2011.05.005.

Sayarpour, M., Zuluaga, E., Kabir, C.S. et al. 2009. The Use of Capacitance–Resistance Models for Rapid Estimation of Waterflood Performance and Optimization. *Journal of Petroleum Science and Engineering* **69** (3): 227-238. https://doi.org/10.1016/j.petrol.2009.09.006.

Schlaifer, R. 1959. *Probability and Statistics for Business Decisions*. New York, USA: McGraw-Hill.

Skov, A.M. 1995. An Analysis of Forecasts of Energy Supply, Demand, and Oil Prices. Presented at the SPE Hydrocarbon Economics and Evaluation Symposium, Dallas, USA, 26-28 March. SPE-30058-MS. https://doi.org/10.2118/30058-MS.

Tavassoli, Z., Carter, J.N., King, P.R. 2004. Errors in History Matching. *SPE Journal* **9** (3): 352-361. SPE-86883-PA. https://doi.org/10.2118/86883-PA.

Thomas, P. 2016. Managing Uncertainty with Decision Analysis in The Oil and Gas Industry. PhD Thesis, University of Stavanger, Stavanger, Norway.

Thomas, P., Bratvold, R.B. 2015. A Real Options Approach to the Gas Blowdown Decision. Presented at the SPE Annual Technical Conference and Exhibition, Houston, Texas, 28-30 September. SPE-174868-MS. https://doi.org/10.2118/174868-MS.

Valko, P.P., Lee, W.J. 2010. A Better Way to Forecast Production from Unconventional Gas Wells. Presented at the SPE Annual Technical Conference and Exhibition, Florence, Italy, 19-22 September. SPE-134231-MS. https://doi.org/10.2118/134231-MS.

Wang, C., Li, G., Reynolds, A.C. 2009. Production Optimization in Closed-Loop Reservoir Management. *SPE Journal* **14** (3): 506-523. SPE-109805-PA. https://doi.org/10.2118/109805-PA.

Wiggins, M., Startzman, R. 1990. An Approach to Reservoir Management Presented at the SPE Annual Technical Conference and Exhibition, New Orleans, USA, 23-26 September. SPE-20747-MS. https://doi.org/10.2118/20747-MS.

Willis, H.H., Moore, M. 2013. Improving the Value of Analysis for Biosurveillance. *Decision Analysis* **11** (1): 63-81. https://doi.org/10.1287/deca.2013.0283.

Wolff, M. 2010. Probabilistic Subsurface Forecasting-What Do We Really Know? *Journal of Petroleum Technology* **62** (5): 86-92. SPE-118550-JPT. https://doi.org/10.2118/118550-JPT.

Wright, S.J., Nocedal, J. 1999. *Numerical optimization*. Berlin, Germany: Springer.

Paper I

Integrating Model Uncertainty in Probabilistic Decline Curve Analysis for Unconventional Oil Production Forecasting

Hong, A.J., Bratvold, R.B., and Lake, L.W.

Manuscript

Integrating Model Uncertainty in Probabilistic Decline Curve Analysis for Unconventional Oil Production Forecasting

Aojie Hong and **Reidar B. Bratvold**, University of Stavanger and The National IOR Centre of Norway; and **Larry W. Lake**, University of Texas at Austin

Abstract

Decline curve analysis is an industry-accepted and extensively used method in the oil and gas industry. As a result of the realization that the long-used Arps model may not be ideal for describing flow from unconventional plays, and may lack physical parameters, several models have been developed to capture the characteristics of different flow regimes.

We use a Bayesian method and actual field data to identify the "best" (most probable) model given data. Each model has a unique extrapolation and uncertainty. No one model is best in all circumstances, so instead of identifying a single "best" model for unconventional production, we propose to regard any model as a potentially good model whose goodness is described by a probability representation. These probabilities of the models are further used to weight the model forecasts.

The main contributions of this work are (1) using probability to describe the goodness of a model, (2) an approach to integrate the model uncertainty in probabilistic decline curve analysis, (3) illustration and discussion of the impact of the model uncertainty, and (4) to illustrate our proposed approach in a real case.

We demonstrate and conclude that using our proposed approach can reduce the risk of using a best fit but very wrong model for forecasting. Our proposed approach performs well in propagating the model uncertainty to the uncertainty in forecast.

Introduction

Although numerical techniques for forecasting hydrocarbon production have developed rapidly over the past decades, decline curve analysis (DCA) remains an industry-accepted method that is used extensively in the oil and gas industry. Decline curve models are very computational attractive because only production data, which can be easily acquired, is required for determining the parameter values of a decline curve model through history matching. The history-matched model is further used for forecasting the hydrocarbon production and reserves.

The Arps (1945) model[1] has been used extensively for both conventional and unconventional plays. However, the Arps model may not be ideal for unconventional plays because unconventional wells usually completed with hydraulic fractures so that several flow regimes (formation linear flow, apparent boundary dominated flow by fracture interference, linear flow in unstimulated matrix and true boundary dominated flow) may appear during the life of a well (Joshi and Lee 2013). Therefore, several alternative decline curve models have been developed to capture the characteristics of different flow regimes. Examples are the power law exponential model (Ilk et al. 2008), the stretched exponential model (Valko 2009), the Duong (2011) model, the logistic growth model (Clark et al. 2011) and the Pan (2016) capacitance-resistance model (CRM).

Given these models, a question, that arises naturally, but has not been discussed widely, is: which is the best model? This question is subtler than it appears because the meaning of "best" is not well defined. In many practices, the model that can best fit the data in a least squares sense is regarded as the best model. However, this ignores two facts: the best-fit model may not be the model that best describes the flow behavior and there may be several models which fit the data almost equally well.

In deterministic DCA, a single model with best-fit parameter values is used for forecasting. Deterministic analysis alone does not quantify the uncertainty in a forecast and thus it often leads to a "precisely wrong" result. A

[1] Commonly, the Arps model is categorized into three types based on the value of the decline exponent b—exponential ($b = 0$), hyperbolic ($0 < b < 1$) and harmonic ($b = 1$). In the following, we use "the Arps model" to refer to all these three types by allowing $0 \leq b \leq 1$.

"vaguely right" forecast in a probabilistic/uncertain sense is more useful and essential for decision making (Read 1920, Bratvold and Begg 2008).

Although probabilistic DCA has not been used extensively in the industry, it is gathering more focus. Jochen and Spivey (1996), and Cheng et al. (2010) used a bootstrap method with the Arps model to quantify the uncertainty in forecast for conventional plays. Gong et al. (2014) used Markov-chain Monte Carlo (MCMC) method with the Arps model to quantify the uncertainty in forecast for unconventional plays. These authors considered measurement errors in decline curve analysis.[2] However, they assumed that a model for forecasting flow behavior is already given. Gonzalez et al. (2012) used MCMC with different individual models to investigate their performances in quantifying the uncertainty in forecast for unconventional plays; their objective was to identify a single best model for probabilistic DCA.

Instead of identifying the "best" model for unconventional plays, we propose to regard any model as a potentially good model and its goodness is described by a probability representation. If a model has larger probability of being better than the others, we say that this model is more likely to be a good model. If several models have close probabilities, we say that the uncertainty in the models is large because it is difficult to tell which model is more likely to be a good model. In this manner, the model uncertainty[3] can be easily integrated in probabilistic decline curve analysis.

An analogy to our proposal is the aggregation of forecasts from diverse experts. Such a forecast is a probability distribution of an uncertain quantity given subjectively or modeled by an expert. Thus, a decline curve model can be analogized to an expert. The topic of aggregating forecasts has been furthered significantly over the past decades in other fields (e.g., risk analysis and operational research). More than 40 years ago, Bunn (1975) reasoned as follows: "The principal motivation for combining forecasts has been to avoid the a priori choice of which forecasting method to use by attempting to aggregate together all the information which each forecasting model embodies. In selecting the 'best' mode, the forecaster is often discarding useful independent evidence in those models which are rejected. Hence, the methodology of combining forecasts is founded upon the axiom of maximal information usage."

Clemen and Winkler (1999) provided a comprehensive review on the aggregation approaches ranging from simple linear aggregation approach (e.g., Stone 1961) to complex copula approach (e.g., Jouini and Clemen 1996) which uses a copula function to include the dependences among the experts' forecasts. Some of their conclusions, that gives support to our work, are (1) "it is worthwhile to consult multiple experts and combine their probabilities;" (2) "in general, simple combination approaches perform quite well;" (3) "heterogeneity among experts is highly desirable" and it is suggested to use three to five experts; and (4) rather than finding a single, all-purpose aggregation approach, an approach should be designed based on the details of each individual situation.

Bunn (1975) quantified the weights in a linear aggregation approach using the probability of a model outperforming the other models. Although his interpretation of probability is different from ours, he used Bayesian theorem to update probabilities based on historical data, which is similar to what we do. Dillon et al. (2002) presented a Monte Carlo based approach to aggregate forecasts. They assigned each individual forecast the same weight. We use the same Monte Carlo based approach but assign each individual forecast different weight which is quantified by the probability.

In the remainder of the paper, we first briefly review four decline curve models. We then introduce an approach to assess the model probabilities which can be used for probabilistic decline curve analysis. An example with synthetic data is used to illustrate the importance of considering the model uncertainty, followed by an application of our proposed approach with real data from two unconventional fields. Finally, we present our conclusions.

The main contributions of this work are (1) using probability to describe the goodness of a model, (2) an approach to integrate the model uncertainty in probabilistic decline curve analysis, (3) illustration and discussion of the impact of the model uncertainty, and (4) illustration of the use of our proposed approach in a real case.

Decline Curve Models

One of the most popular models for decline curve analysis is the Arps model. Over its more than 60 years' application, it has been realized that it may not be ideal for unconventional plays. Thus, several models were developed for analysis of unconventional plays. For example, the power law exponential model (Ilk et al. 2008), the

[2] In decline curve analysis, the data fluctuations caused by changes in operating conditions are treated as caused by measurement errors (Jochen and Spivey 1996).

[3] We use "a model" to refer to a mathematical formulation for describing the physics of a phenomena. For a given model, its parameter values can be uncertain. This is referred to as "the uncertainty in the model parameters". On the other hand, we can be uncertain about which mathematical formulation (i.e. which model) should be used to describe the physics of a phenomena. This is referred to as "the model uncertainty" or "the uncertainty in the models." When only one single model is considered, many authors have used "the model uncertainty" to refer to the uncertainty in the model parameters.

streched exponential model (Valko 2009), Duong (2011) model, the logistic growth model (Clark et al. 2011) and Pan (2016) capacitance-resistance model (CRM). In this work, we focus on four of them: the Arps model, the stretched exponential model (SEM), the logistic growth model (LGM) and the Pan CRM.

Arps Model. It is an empirical model formulated as

$$q_t = \begin{cases} q_0(1 + bD_i t)^{-1/b}, & 0 < b \leq 1 \\ q_0 e^{-D_i t}, & b = 0 \end{cases} \quad (1)$$

where q_t is the rate at time t, q_0 is the initial rate at $t = 0$, D_i is the initial decline rate and b is the decline exponent. With the assumption of boundary-dominated flow, the upper bound of b is 1. For the transient flow regime of unconventional production, b is often greater than 1 (Valko and Lee, 2010) and as a result the cumulative production is unbounded. To avoid unbounded cumulative production, we use the range from 0 to 1 for b.

Stretched Exponential Model. Based on the analysis of the Barnett shale wells, Valko (2009) presented the SEM to formulate an empirical time-rate relation

$$q_t = q_0 e^{-(\frac{t}{\tau})^n} \quad (2)$$

where τ is the characteristic time parameter and n is the exponent parameter. Valko and Lee (2010) interpreted the SEM as: the actual production decline is an integrated effect of multiple contributing volumes each in its individual exponential decay with a specific distribution of characteristic time constants. The distribution can be determined by τ and n—τ is the median of the characteristic time constants and n describes the fatness of the tail of the distribution.

Logistic Growth Model. The LGM is an empirical model originally developed for modeling population growth. It describes the limit of a biological population growth. Hubbert (1956) adopted this model to model production for entire fields or producing regions. Recently, Clark et al. (2011) applied it to model production in a single unconventional well. The specific formulation of the LGM used by them is

$$q_t = \frac{aK\eta t^{\eta-1}}{(a + t^{\eta})^2} \quad (3)$$

where a is a constant, K is the carrying capacity and η is the hyperbolic exponent parameter. The value of the carrying capacity K represents the total amount of hydrocarbon that can be recovered under the primary recovery mechanism in a producer. The hyperbolic exponent parameter η controls the decline behavior—the larger the parameter value is, the slower the decline is. The constant a and the hyperbolic exponent parameter η together determine the time when half of the carrying capacity K has been produced, i.e., the cumulative production at $t = \sqrt[\eta]{a}$, is $Q_t = K/2$.

Pan CRM. Pan (2016) proposed a model to capture the productivity index behavior over both linear transient and boundary-dominated flow,

$$J = \frac{\beta}{\sqrt{t}} + J_\infty \quad (4)$$

where J is the productivity, β is the linear transient flow parameter and J_∞ is the constant productivity index that a well will eventually reach. β is related to the permeability in the analytical solution of linear flow into fractured wells presented by Wattenbarger et al. (1998). By combining Eq. 4 and a tank material balance equation, Pan (2016) derived the analytical solution of rate over time

$$q_t = \Delta P(\frac{\beta}{\sqrt{t}} + J_\infty) e^{-(2\beta\sqrt{t} + J_\infty t)/(c_t V_p)} \quad (5)$$

where ΔP is the difference between the initial reservoir pressure and the assumed constant flowing bottom hole pressure, c_t is the total compressibility and V_p is the drainage pore volume. The Pan CRM may give unrealistically large rate for small t, as q_t approaches infinity when t approaches 0. To deal with this issue, we use this model for

production after 10 days (i.e. $t \geq 10$). The Pan CRM is the only one of the four that is derived analytically and has physical parameters in it.

Determination of Model Parameters

For a given model, its parameters are determined through history matching with the goal to minimize a predefined loss function by adjusting the model parameters.

Deterministic Approach. In a deterministic approach, a model and its parameter values that best fit the data is found and used for forecasting. The best-fit model gives a single estimate of our interest (e.g., future production). A widely used approach is the least squares estimation (LSE) with the aim to minimize the difference between model forecast and data. Thus, the loss function of LSE is defined as

$$L_{LSE}(\boldsymbol{x}) = \sum_{k=1}^{T} [q_k(\boldsymbol{x}) - \hat{q}_k]^2 \tag{6}$$

where $L_{LSE}(\boldsymbol{x})$ is the loss function of LSE which is a function of the vector of model parameters $\boldsymbol{x}$, $q_k(\boldsymbol{x})$ is the model forecasted rate at time step k, $\hat{q}_k$ is the measured rate (i.e., data) at time step k and T is the total number of time steps of data.

Another approach is the maximum likelihood estimation (MLE). It aims to maximize the likelihood function (i.e., the probability of observing the data given a model). Assuming the data measurements are independent, the likelihood function is

$$P(\hat{q}_1, \hat{q}_2, \dots, \hat{q}_T|\boldsymbol{x}) = \prod_{k=1}^{T} P(\hat{q}_k|\boldsymbol{x}) \tag{7}$$

where P denotes probability. Further assuming the measurement of $\hat{q}_k$ has a Gaussian random error with zero mean and standard deviation σ_k, we have

$$P(\hat{q}_k|\boldsymbol{x}) = \frac{1}{\sqrt{2\pi\sigma_k^2}} e^{-[q_k(\boldsymbol{x})-\hat{q}_k]^2/(2\sigma_k^2)}. \tag{8}$$

Thus,

$$P(\hat{q}_1, \hat{q}_2, \dots, \hat{q}_T|\boldsymbol{x}) = \frac{1}{\sqrt{(2\pi)^T \prod_{k=1}^{T} \sigma_k^2}} e^{-\frac{1}{2}\sum_{k=1}^{T} \frac{[q_k(\boldsymbol{x})-\hat{q}_k]^2}{\sigma_k^2}}. \tag{9}$$

Maximizing the likelihood function (Eq. 9) is equivalent to minimizing

$$L_{MLE}(\boldsymbol{x}) = \sum_{k=1}^{T} \frac{[q_k(\boldsymbol{x}) - \hat{q}_k]^2}{\sigma_k^2}. \tag{10}$$

It can be seen that if $\sigma_1 = \sigma_2 = \cdots = \sigma_T$, Eq. 10 can be reduced to Eq. 6. Therefore, LSE is a special case of MLE. The advantage of using Eq. 10 instead of Eq. 6 is that σ_k acts as a weighting factor so that a more accurate data point (i.e., with small σ_k) will have more weight than a less accurate data point (i.e., with large σ_k).

A practical issue of using MLE is that we must assess σ_k, which we seldom know a priori. To approximately estimate σ_k, we suggest using the moving window approach. For example, if we want to assess σ_k, we first assign a value to the half window width w, then take a subset of data $\{\hat{q}_{k-w}, \dots, \hat{q}_{k-1}, \hat{q}_k, \hat{q}_{k+1}, \dots, \hat{q}_{k+w}\}$, and then calculate the sample standard deviation of this subset of data which is treated as σ_k. In this way, a data point in a period with sharp changes or large fluctuations will have large σ_k. This makes sense for DCA because sharp changes (usually sharp drops) in rate normally happens in the early time where the data is less reliable and when a subset of data with large fluctuations is less important for history matching and should be assigned less weight. To our knowledge, there is no strict rule for the determination of the optimal window width. Many empirical rules relate the window width to

the number of data pints. However, to find a rule for the optimal window width is outside the scope of this paper. Based on our practice, we use a half window width of 10 data points.

Probabilistic Approach. Instead of estimating a single value of our interest, a probabilistic approach focuses on assessing a distribution or a range (e.g., the 80% confidence interval) of our interest with the consideration of uncertainties in measurements, inverse modelling[4] and model type. The latter will be detailed in the next section.

Jochen and Spivey (1996) illustrated the use of the bootstrap method combined with DCA to give a probabilistic reserves estimation. The bootstrap method assumes a noninformative prior, so it does not require a priori knowledge on the distribution of model parameters. Because of unknown measurement errors, the bootstrap method includes uncertainties in measurements by using the Monte Carlo method to resample datasets from the original dataset with replacement. For each resampled dataset, the model parameters are determined using LSE. As it repeats for numerous resampled datasets, the posterior distributions[5] of model parameters can be obtained.

Tavassoli et al. (2004) presented the issue of non-uniqueness in inverse modelling (inherent uncertainty in history matching) and showed that different combinations of model parameters may give almost equally good history matching result but give different forecasts. Sayarpour et al. (2011) started with different sets of initial guesses of model parameters to history match data to generate numerous history matched solutions of model parameters.

To allow for the uncertainties in both measurements and inverse modeling, we start with different sets of initial guesses of model parameters when history matching each resampled dataset.

Integrating Model Uncertainty in Probabilistic Decline Curve Analysis

The approaches presented in the previous section assume that a single model, which predicts reservoir performance, has been proposed. None of the reviewed publications discussed a systematic approach of determining which decline curve model is the best for unconventional plays. Moreover, in some cases, it is not easy to select a best model because satisfactory history matching resutls can be produced by several different models. Instead of determining a best model, we propose to include the model uncertainty in the probabilistic approach by assessing the probability of each model considered.

This section illustrates our proposed approach to integrate model uncertainty in probabilistic DCA. We name this approach the multiple-model probabilistic MLE (MM-P-MLE) approach because it considers multiple models and applies probabilistic DCA with MLE for data matching. More particularly, in the following, we use single-model probabilistic MLE (SM-P-MLE) to refer to the circumstance where probabilistic DCA with MLE is applied for a single "best" model, and single-model deterministic MLE (SM-D-MLE) to the circumstance where deterministic DCA with MLE is applied for a single "best" model.

The procedure of performing MM-P-MLE is described as follows. We assume the measurement of rate $\hat{q}_k$ has a Gaussian random error with zero mean and standard deviation σ_k, and use the moving window approach to approximately estimate σ_k. Similar to the bootstrap method, numerous datasets are sampled. However, since the standard deviation of measurement error has been assessed, we can use the Monte Carlo method to sample from a Gaussian distribution with mean $\hat{q}_k$ and standard deviation σ_k for each data point. For a given sampled dataset, we use MLE to determine the parameters of each model considered. Then, we use Bayes' theorem to calculate the probability of each model with its parameters of MLE solution given the sampled dataset:

$$P\left(\boldsymbol{x}_{ij}^{MLE}, m_i \middle| \boldsymbol{d}_j, \Omega'\right) = \frac{P\left(\boldsymbol{d}_j \middle| \boldsymbol{x}_{ij}^{MLE}, m_i\right) P\left(\boldsymbol{x}_{ij}^{MLE}, m_i \middle| \Omega\right)}{\sum_{i'} P\left(\boldsymbol{d}_j \middle| \boldsymbol{x}_{i'j}^{MLE}, m_{i'}\right) P\left(\boldsymbol{x}_{i'j}^{MLE}, m_{i'} \middle| \Omega\right)} \tag{11}$$

where $\boldsymbol{x}_{ij}^{MLE}$ is the parameters of MLE solution given model m_i and sampled dataset $\boldsymbol{d}_j$; Ω denotes a priori knowledge; and Ω' denotes a posteriori knowledge. Using a noninformative prior distribution, we have $P(\boldsymbol{x}_{m1}^{MLE}, m_1|\Omega) = P(\boldsymbol{x}_{m2}^{MLE}, m_2|\Omega) = \cdots$, so Eq. 9 is reduced to

$$P\left(\boldsymbol{x}_{ij}^{MLE}, m_i \middle| \boldsymbol{d}_j, \Omega'\right) = \frac{P\left(\boldsymbol{d}_j \middle| \boldsymbol{x}_{ij}^{MLE}, m_i\right)}{\sum_{i'} P\left(\boldsymbol{d}_j \middle| \boldsymbol{x}_{i'j}^{MLE}, m_{i'}\right)}. \tag{12}$$

[4] We use "uncertainty in inverse modeling" interchangeably with "non-uniqueness in inverse modeling" to describe the fact that multiple combinations of model parameter values may give equally good match to data.

[5] In the context of history matching, "posterior distribution" is referred to the distribution of our interested uncertain quantity given additional information.

Insert Eqs. 9 and 10 into Eq. 12, we get

$$P\left(\boldsymbol{x}_{ij}^{MLE}, m_i \middle| \boldsymbol{d}_j, \Omega'\right) = \frac{e^{-\frac{1}{2}L_{MLE}\left(\boldsymbol{x}_{ij}^{MLE}\right)}}{\sum_{i'} e^{-\frac{1}{2}L_{MLE}\left(\boldsymbol{x}_{i'j}^{MLE}\right)}} = 1 \Big/ \sum_{i'} e^{-\frac{1}{2}\left|L_{MLE}\left(\boldsymbol{x}_{i'j}^{MLE}\right) - L_{MLE}\left(\boldsymbol{x}_{ij}^{MLE}\right)\right|}. \tag{13}$$

The probability calculated using Eq. 13 is the probability of the estimate of our interest forecasted by model m_i with its parameters $\boldsymbol{x}_{ij}^{MLE}$ given a sampled dataset $\boldsymbol{d}_j$. Because $\boldsymbol{d}_j$ is an independent Monte Carlo sample, $P(\boldsymbol{d}_j) = 1/N$ and

$$P\left(\boldsymbol{x}_{ij}^{MLE}, m_i, \boldsymbol{d}_j \middle| \Omega'\right) = P\left(\boldsymbol{x}_{ij}^{MLE}, m_i \middle| \boldsymbol{d}_j, \Omega'\right) P(\boldsymbol{d}_j) = 1 \Big/ \left\{ N \sum_{i'} e^{-\frac{1}{2}\left|L_{MLE}\left(\boldsymbol{x}_{i'j}^{MLE}\right) - L_{MLE}\left(\boldsymbol{x}_{ij}^{MLE}\right)\right|} \right\} \tag{14}$$

where N is the total number of Monte Carlo samples ($N = 1000$ in the following applications). The probability calculated using Eq. 14 is regarded as the weight given to the estimate forecasted by model m_i with its parameters $\boldsymbol{x}_{ij}^{MLE}$. By repeating this process over the models and the sampled datasets, we obtain a set of weighted samples of the estimate as shown in **Table 1** where M is the total number of models considered, f_{ij} is the estimate forecasted using m_i and $\boldsymbol{x}_{ij}^{MLE}$, and P_{ij} denots $P\left(\boldsymbol{x}_{ij}^{MLE}, m_i, \boldsymbol{d}_j \middle| \Omega'\right)$. This set of weighted samples represents the distribution of the estimate with the integration of model uncertainty. The marginal posterior probability of a model regardless of any sampled dataset is calculated as

$$P(m_i|\Omega') = \sum_{j=1}^{N} P\left(\boldsymbol{x}_{ij}^{MLE}, m_i, \boldsymbol{d}_j \middle| \Omega'\right). \tag{15}$$

We interpret this probability as a measure of the relative truthfulness of model i to the other models.

Estimate forecasted using m_i and $\boldsymbol{x}_{ij}^{MLE}$	f_{11}	…	f_{M1}	f_{12}	…	f_{M2}	…	f_{1N}	…	f_{MN}
Weight assigned to the estimate	P_{11}	…	P_{M1}	P_{12}	…	P_{M2}	…	P_{1N}	…	P_{MN}

Table 1—Set of weighted samples of estimate.

Illustrative Example with Synthetic Data

This section uses a synthetic dataset to illustrate the impacts of using the deterministic and probabilistic approaches without considering the model uncertainty (i.e. SM-D-MLE and SM-P-MLE) on cumulative oil production estimation, and to highlight the importance of using our proposed approach (MM-P-MLE). We consider only the Arps model, SEM and Pan CRM for this example, but more models can be easily included in our approach as we will include the LGM for the analyses with actual field data.

The "true" decline is generated using the Pan CRM. Random errors, drawn from Gaussian distributions with zero mean and standard deviation equal to 20% of true rate, are added to the "true" decline to form the synthetic dataset. **Fig. 1** illustrates the synthetic dataset and the "true" decline of oil production rate as well as the standard deviation of measurement error assessed using the moving window approach with a half window size of 10. Our interest is the cumulative oil production from the time of the last data point to day 10950 (year 30). The "true" cumulative oil production given by the Pan CRM is 48.1 Mbbl. This value is used as a reference for subsequent estimates.

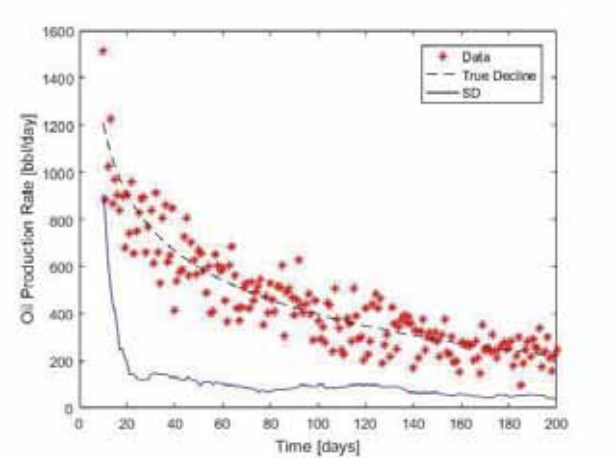


Fig. 1—Synthetic dataset to day 200.

We first use SM-D-MLE to estimate the cumulative oil production. The resulting best-fit parameters and corresponding loss function value of MLE for each model is listed in **Table 2**. We see that Arps model with $q_0 = 1121.6$ bbl/day, $b = 0.75$ and $D_i = 0.016$ day^{-1} best fits the data (minimal loss function). Therefore, it is used for forecasting, providing an estimated cumulative oil production of 132.6 Mbbl. This estimate is more than 2.5 times the "true" value.

Arps Model		**SEM**		**Pan CRM**	
q_0 [bbl/day]	1121.6	q_0 [bbl/day]	1519.8	ΔP [psi]	507.8
b [-]	0.75	τ [day]	56.5	J_∞ [bbl/day/psi]	1.16
D_i [day^{-1}]	0.016	n [-]	0.53	$c_t V_p$ [bbl/psi]	283.7
				β [bbl/day$^{1/2}$/psi]	4.2
L_{MLE}	171.65	L_{MLE}	171.75	L_{MLE}	172.29

Table 2—Best-fit parameters for the Arps, SEM and Pan CRM models for the synthetic dataset to day 200.

To take the uncertainties in measurements and inverse modeling into consideration, we use SM-P-MLE with the Arps model. This produces a distribution of the cumulative oil production, whose P10, P50, mean and P90 are listed in **Table 3**. The forecast of the Arps model is biased, far from the "truth" (48.1 Mbbl), and the 80% confidence interval (from P10 to P90) does not contain it.

Statistics	**P10**	**P50**	**Mean**	**P90**
Cumulative Oil Production [Mbbl]	73.5	132.3	135.7	207.2

Table 3—Statistics of the cumulative oil production forecasted by the Arps model given the synthetic dataset to day 200.

We use MM-P-MLE to integrate the model uncertainty in this analysis. Indeed, the minimized loss function values of the three models are very close to each other, which means no model is superior to the others. The marginal posterior probabilities calculated using Eq. 15 are 36.4%, 34.0% and 29.6% for the Arps, SEM and Pan CRM models, respectively. The models are almost equally likely to be the correct model given the synthetic dataset. This result seems counter-intuitive as the "correct" model (the Pan CRM) is the least likely one because the dataset does not show an obvious linear transient flow behavior that the Arps model may not capture. Thus, the model uncertainty remains large even when the dataset is given.

Figure 2 shows Box plots of cumulative oil production forecasted using SM-P-MLE solely with the Arps, SEM or Pan CRM models, and using MM-P-MLE. Among the three models, the Arps model has the largest uncertainty in inverse modeling as it gives the largest 80% confidence interval, while the Pan CRM model has the smallest uncertainty in inverse modeling. Since the Pan CRM is the one used to generate the data, the "correct" model, its estimate is the best. Using the proposed approach MM-P-MLE, the "truth" is contained in the 80% confidence interval. This means that by integrating the model uncertainty in the analysis, we can reduce the risk of selecting a wrong model for forecasting when the correct model is unknown.

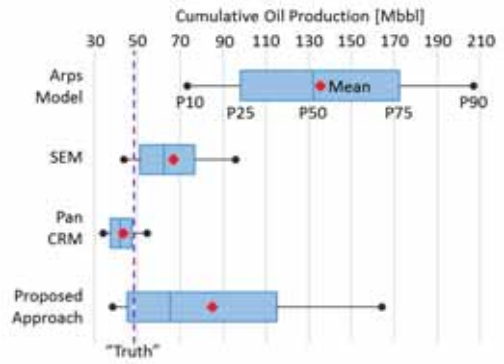


Fig. 2—Box plots of cumulative oil production forecasted using a single model and compared to that of the proposed approach given the synthetic data to day 200.

We extend the dataset to day 400 (**Fig. 3**) and use MM-P-MLE to estimate the cumulative oil production from day 400 to day 10950 (year 30). The dataset is in Fig. 3. The resulting marginal posterior probabilities are 14.2%, 38.3% and 47.5% for the Arps, SEM and Pan CRM models, respectively. Given the additional data from day 200 to day 400, the Pan CRM becomes the most likely, whereas the Arps model is the least likely. This ranking given by the marginal posterior probabilities is consistent with the ranking given by the matching quality shown in **Fig. 4**, where each decline curve represents a MLE solution for a sampled dataset and a single model. **Fig. 5** shows the Box plots of cumulative oil production forecasted using SM-P-MLE solely with the Arps, SEM or Pan CRM model, and using MM-P-MLE. Because the SEM and Pan CRM now have much larger probability than the Arps model, the distribution forecasted using MM-P-MLE is skewed toward the distributions forecasted solely by the SEM and Pan CRM with a tail toward that of the Arps model. When considering the Arps or SEM model individually, none of these models provides an 80% confidence interval containing the truth. However, when all these three models are considered so that the "correct" model is included, the estimate is improved as the 80% confidence interval obtained using MM-P-MLE contains the "truth".

This example indicates that to include multiple models for forecasting reduces the risk of rejecting a good model by selecting a single "best" model.

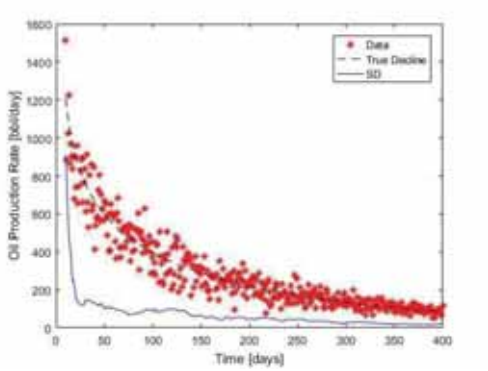


Fig. 3—Synthetic dataset to day 400.

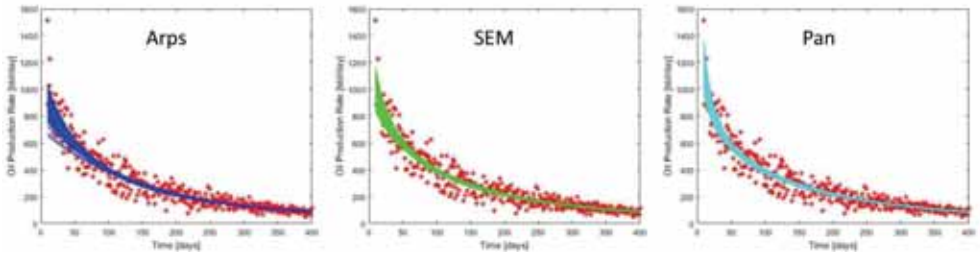


Fig. 4—Plots of the history matching results for the Arps, SEM and Pan CRM models.

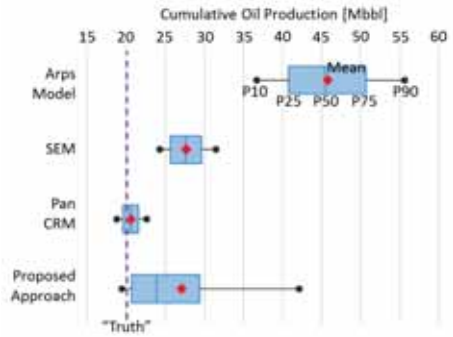


Fig. 5—Boxplots of cumulative oil production forecasted using solely one model and using the proposed approach given the synthetic data to day 400.

Application to the Bakken Field Data

Our proposed approach is applied to daily oil production rate data from the Bakken field. The Bakken wells that we study in this work are completed in a fractured shale reservoir. All the wells we select for this study are completed in the same formation. First, we conduct a hindcast test where the first part of a dataset is used for history matching and the second part is used for comparison with the model forecast. Finally, we use our proposed approach to forecast the cumulative oil production for selected wells.

Hindcast Test. The oil production rate data is from Bakken Well UT-ID 220 (**Fig. 5**). There are 660 data points in total, covering the period from day 10 to day 695. The rates are measured daily. We use the data from day 10 to day 100 for history matching, and then use the history-matched model for forecasting. In this test, we consider the Arps, SEM and Pan CRM models. The resulting marginal posterior probabilities are 14.9%, 37.4% and 47.7% for the Arps, SEM and Pan CRM models, respectively. Although the Pan CRM is the most likely model, we are not confident enough to say that Pan CRM is superior to the other two because its probability is only slightly larger than that of the others. Therefore, the possibilities of the Arps and SEM models should be included in the forecast using MM-P-MLE. The forecast from day 100 to day 695 compared with the corresponding data is shown in **Fig. 6**. We see that most of the data falls in the 80% confidence interval.

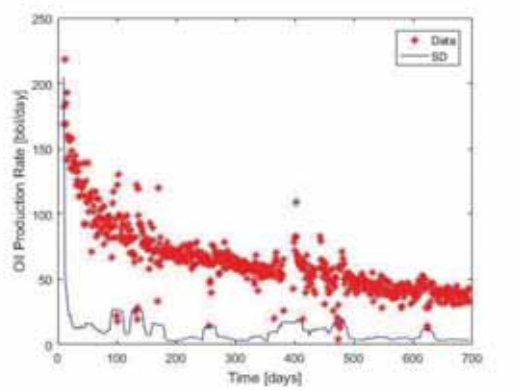


Fig. 5—Data from Bakken Well UT-ID 220 from day 10 to day 695.

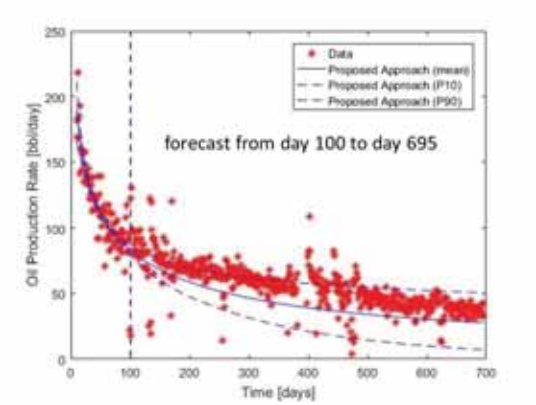


Fig. 6—The forecast from day 100 to day 695 compared with the corresponding data from Bakken Well UT-ID 220.

We use more data points to investigate how this will impact the forecast. Given data from day 10 to day 200, the probability of the Pan CRM increases from 47.7% (given data from day 10 to day 100) to 79.3%, the probability of the SEM drops to 20.7% and the probability of the Arps model drops to 0%. The forecast of oil production rate from day 200 to day 695 is in **Fig. 7**. The forecast matches the data better and its uncertainty is reduced.

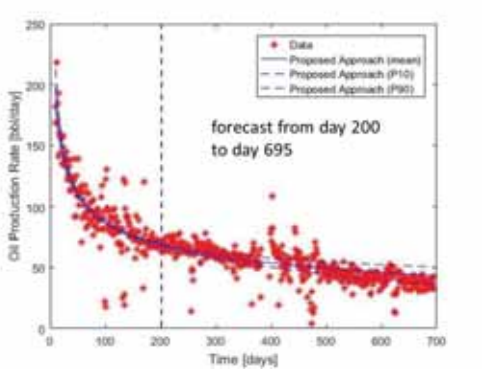


Fig. 7—The forecast from day 200 to day 695 compared with the corresponding data from Bakken Well UT-ID 220.

Model Probabilities and Cumulative Oil Production Estimates for Bakken Wells. We select 28 Bakken wells with relatively less noisy data. The model probabilities and cumulative oil production are assessed for these wells. The wells have different operation times; some have a short period of data while some are long. The data length of each well is listed in **Table 4**.

Well UT-ID	**222**	**208**	**401**	**195**	**69**	**194**	**413**	**187**	**216**	**804**	**223**	**78**	**80**	**232**
Data Length [days]	213	223	233	260	292	309	336	337	372	377	392	434	499	515

Well UT-ID	**265**	**209**	**385**	**227**	**220**	**391**	**386**	**228**	**67**	**4**	**198**	**197**	**81**	**82**
Data Length [days]	531	533	546	707	740	769	886	993	1136	1241	1306	1377	1620	1642

Table 4—Data length of selected Bakken wells.

This analysis considers the Arps, SEM, LGM and Pan CRM models. The marginal posterior probabilities of the decline curve models are illustrated in **Fig. 8**. Among the 28 selected wells, the Pan CRM is most likely for 13 wells, the SEM for 7 wells, and each of the LGM and Arps models for only 4 wells. This confirms that the Arps model may not be ideal for unconventional plays, and indicates that the Pan CRM, as the only analytical model, is more likely to better describe the unconventional flow behavior than the empirical models.

For the wells with short data length (smaller than 400 days), the model uncertainty is large and the probability of the Arps model is comparable to the other three models. As the data length increase, one of the models becomes dominating with a probability higher than 90%. However, for a few of the well with long data length, the uncertainty remains in the SEM, LGM and Pan CRM (for example, Wells UT-ID 228 and 197). Thus, using solely one of them for further analysis may underestimate the uncertainty even when there is a lot of data.

We see two wells (Wells UT-ID 187 and 227) with the Arps model dominating. The reason may be that the data is so noisy that it masks the decline features of this unconventional oil or that the fracture flow is indeed not dominating in that well. Additional and less noisy data is required for these wells.

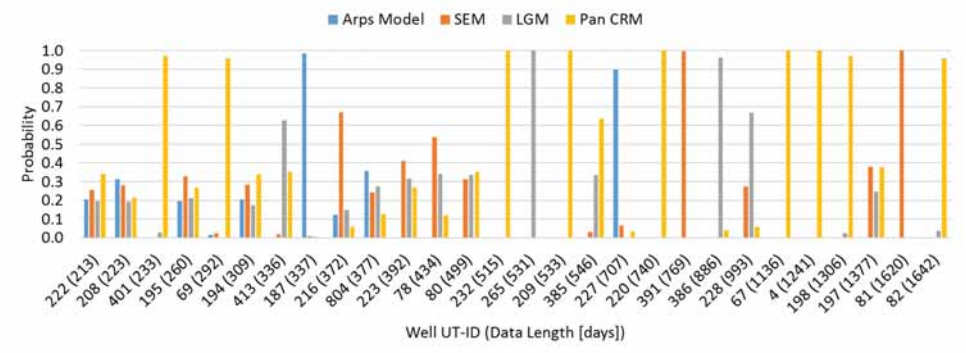


Fig. 8—The marginal posterior probabilities of the decline curve models for the 28 selected Bakken wells.

Considering a well's life cycle of 30 years, we are interested in how much oil is left in each well (i.e., the cumulative oil production from the last day with data to day 10950). The mean value or expected value of the cumulative oil production for each of the selected Bakken wells is in **Fig. 9**. The figure includes the mean values estimated using solely one of the models and using our proposed approach. The Box plot of the estimated cumulative oil production using MM-P-MLE is in **Fig. 10**. In the Box plot, all the values are normalized by the mean value estimated using MM-P-MLE for each well.

From Fig. 9, we see that the LGM and Arps models gives a higher estimate of the mean value of the cumulative oil production than the other two models for most of the wells with short data length. This means that using the LGM or Aprs models solely with short data length may give an optimistic estimate in cumulative oil production. When the proposed approach is used, 15 wells' mean values are between the highest and lowest mean values estimated using the models individually, 8 wells hits the highest mean values estimated using the models individually, and 5 wells hits the lowest mean values estimated using the models individually. Thus, using the proposed approach to take the model certainty into account will give a result that is neither too optimistic nor too pessimistic.

In Fig. 10, for most of the wells with short data length, the uncertainty in estimated cumulative oil production is large with wide 80% confidence interval, whereas for most of the wells with long data length, the uncertainty is significantly smaller. It is because the uncertainty in the estimated cumulative oil production is large as the model uncertainty is large. Well UT-ID 197 has wide 80% confidence interval even it has long data length because of the uncertainty in the SEM, LGM and Pan CRM models. Well UT-ID 385 has an extremely skewed distribution. This is because the Pan CRM forecasts low cumulative oil production with small SD while the LGM forecasts much larger value with large SD; when 63.4% weight is given to the Pan CRM and 33.6% to the LGM, the resulting distribution skews towards the Pan CRM's forecast with a long tail toward the LGM's forecast.

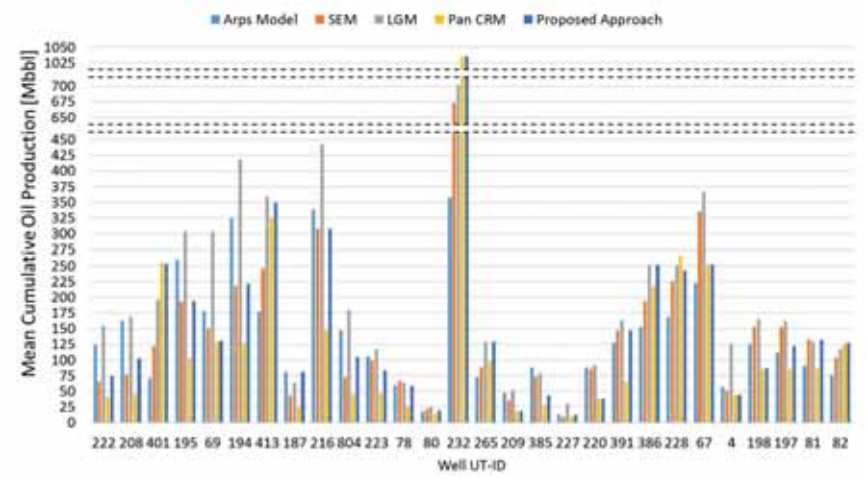


Fig. 9—Mean values of estimated cumulative oil production for the 28 selected Bakken wells.

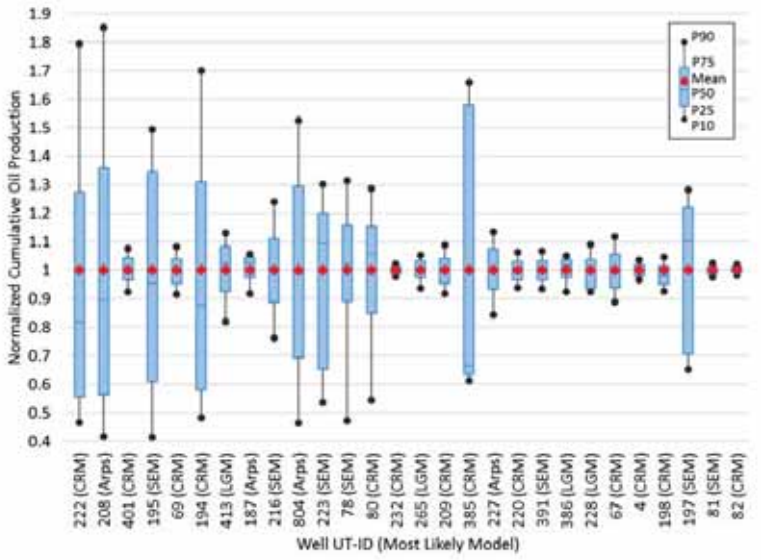


Fig. 10—Box plots of estimated cumulative oil production using MM-P-MLE for the 28 selected Bakken wells.

Application to the Midland Field Data

We perform the same analysis for the Midland field, as we did for the Bakken field. The Midland wells that we study in this work are completed in a fractured reservoir. Unlike the daily oil production data from the Bakken field, the Midland field data is monthly. Thus, the Midland field data is smoother than the Bakken field data. We select 31 Midland wells with relatively less noisy data. Each well has been operated under different lengths of time. **Table 5** lists the length of data for each selected well.

Well ID	28	24	47	9	22	11	14	48	25	15	29	32	42	27	39	45
Data Length [months]	62	68	71	72	75	76	76	76	77	78	78	78	78	80	81	82

Well ID	43	49	26	13	31	33	34	40	41	44	17	18	19	30	16
Data Length [months]	83	83	84	85	86	86	87	89	93	94	97	97	97	97	104

Table 5—Data length of selected Midland wells.

Figure 11 illustrates the marginal posterior probabilities, calculated using MM-P-MLE, for the Arps, SEM, LGM and Pan CRM models. Among the 31 selected Midland wells, the Pan CRM is most likely for 18 wells, the Arps model for 8 wells, the SEM for only 3 wells, and the LGM for only 2 wells. This indicates that, in general, the Pan CRM has higher chance to be a good model for describing the Midland field data than the other three models. Nevertheless, no one model is best for all wells. For some wells, the Pan CRM can be the least likely model. Moreover, even if the Pan CRM is the most likely one, some other models may have probabilities close to that of the Pan CRM; for example, the marginal posterior probabilities for Well 14 are 35.2% for the Pan CRM, 27.7% for the SEM and 25.4% for the LGM.

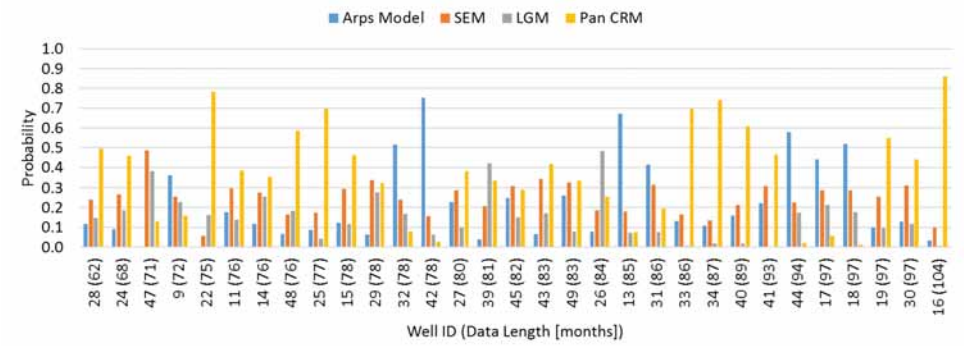


Fig. 11—The marginal posterior probabilities of the decline curve models for the 31 selected Midland wells.

Figure 12 illustrates the mean values of estimated cumulative oil production from the last data point to year 30 using SM-P-MLE with the Arps, SEM, LGM or Pan CRM model solely, and MM-P-MLE with the consideration of all these four models. When we consider these models individually, the LGM tends to give the highest estimate while the Pan CRM tends to give the lowest estimate for most of the wells. Although the Arps model does not tend to give an estimate as high as the LGM, it tends to give an estimate higher than the SEM and Pan CRM. Our proposed approach MM-P-MLE weights each individual model and gives a moderate estimate.

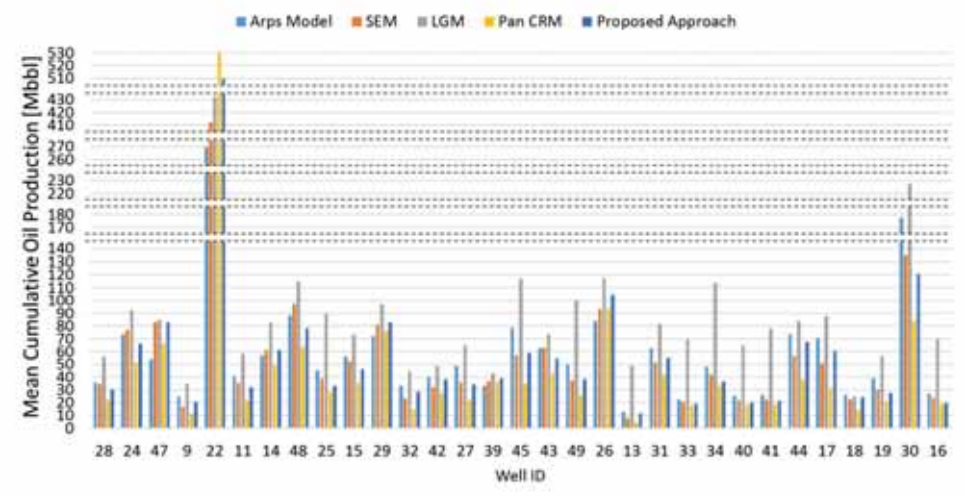


Fig. 12—Mean values of estimated cumulative oil production for the 31 selected Bakken wells.

Figure 13 illustrates the Box plots of estimated cumulative oil production using MM-P-MLE for the 31 selected Midland wells. All the values are normalized by the estimated mean for each well. For most of the wells, the 80% confidence interval is between 0.5 and 1.5 times the mean value. Compared to the Bakken wells, the uncertainties in the forecasts of Midland wells are relatively large (wider 80% confidence intervals). This is because Bakken wells' daily production data contains more data points than Midland wells' monthly production data. As the number of data points increases, less uncertainty in measurement propagates to the uncertainty in forecast.

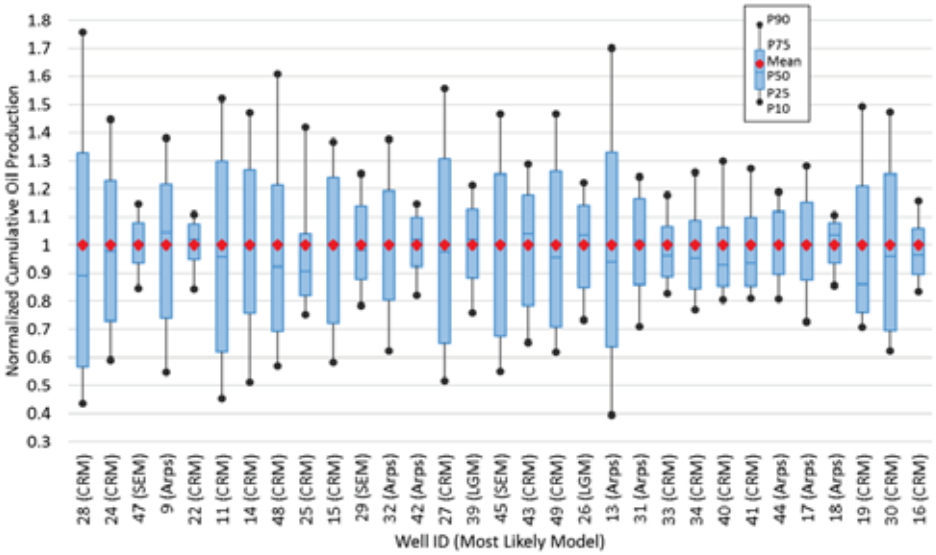


Fig. 13—Box plots of estimated cumulative oil production using MM-P-MLE for the 31 selected Midland wells.

Discussion

Ambiguous choice of a single model for DCA can be caused by cognitive biases[6]. Overconfidence is one of the most common cognitive biases in the oil and gas industry (Welsh et al., 2005). When a single model is selected for DCA, one gives a 100% weight to this selected model (i.e., trust this model 100%). However, it is obvious that a single model should not be trust 100%. Thus, he/she is overconfident in the selected model. When diverse models are considered, the impact of overconfidence is reduced. Our proposed approach MM-P-MLE provides a consistent and systematic framework for including multiple models in probabilities DCA and updating the model probabilities/weights based on given data. Therefore, the risk of over/underestimate given by a single model is reduced and MM-P-MLE gives a moderate estimate.

Traditional measures of the goodness of fit—(weighted) sum of squares and (weighted) R^2—can only be used to rank different models but they do not have any useful interpretation for probabilistic analysis. It is more useful for probability analysis to convert them into corresponding probability representation. Moreover, probability has a more intuitive interpretation to many people than sum of squares and R^2.

The Pan CRM is the most likely model for most of the unconventional wells selected for this work. It tends to give the lowest estimate on cumulative oil production over long time among the four models studies in this work. This means that it is very likely that the true cumulative oil production is lower than the forecast by the other models (for example, the widely used Arps model). This might explain that why an oil company's estimate on future oil production has often turned out to be too optimistic.

However, no one model is best in all circumstances, as for some wells, the Arps, SEM or LGM models may performs better than the others. MM-P-MLE avoids the a priori choice of which model to use and the rejection of a possible good model. Thus, we are open-minded to other models and weights them according to their quality of data match. Any model holds the possibility to be the "best" when the data provides enough evidence as more data becomes available.

When the number of data points is small, the model uncertainty remains large, because these data points do not clearly reflect the characteristics of different flow regimes of unconventional production. As the number of data points increases, the characteristics of different flow regimes is shown more clearly, and the model that cannot well capture the physics of all the flow regimes will be gradually eliminated (i.e., its probability approaches 0).

The Pan CRM is the only analytical model among the models we consider in this work. It is designed to capture the major flow regimes—transient flow regime and semi-steady state flow regime—involved in an unconventional well. Therefore, it is not surprising that the data from two unconventional fields show that the Pan CRM is the most likely one to well describe the flow behaviors, among the four models. Some wells show that one of the other three models is the most likely for some possible reasons: the measure errors masks the characteristics of different flow regimes, different flow regimes show up simultaneously, minor flow regimes appears, fracture flow is not dominating, or unclear flow behaviors appear. Besides, unlike that the parameters of an empirical model have no physical meaning, all the parameters of the Pan CRM are physically defined, and the fit parameter values reveal important reservoir/well properties. For example, the V_p in the Pan CRM reveals the effective drainage volume of a producer.

Our investigation is limited to the Arps, SEM, LGM and Pan CRM models. More models should be included in this analysis. Other models (e.g., a dual-porosity reservoir simulation model) might be superior to the Pan CRM. However, we argue that using a more complex model with more parameters can make probabilistic analysis too computationally intensive and may not create much additional value (see the discussion in Bratvold and Begg (2009) and Hong et al. (2017)). One of the advantage of DCA is its speed. Thus, we prefer models with small number of parameters for DCA.

Conclusions

We propose an approach, MM-P-MLE, that integrates model uncertainty in probabilistic DCA for unconventional oil production forecasting. Different from the approach that uses only a single best-fit model for further analysis, the proposed approach interprets the goodness of fit of a model with a probability representation that can be carried to uncertainty analysis. The uncertainty in the Arps ($0 \leq b \leq 1$), SEM and Pan CRM models for analyzing unconventional plays was investigated.

An example illustrated that the best-fit model may not be the model that best fits the flow behavior. Using the proposed approach can reduce the risk of using a best-fit but very wrong model for forecasting.

The proposed approach was applied with real oil production data from the Bakken and Midland fields. The hindcast test showed that the model uncertainty was reduced and the forecast was improved as more data points

[6] A cognitive bias refers to the unconscious deviation from rationality in judgment (Haselton et al., 2005).

became available. The model probabilities were assessed and the cumulative oil productions were estimated using our proposed approach for 28 selected Bakken wells and 31 selected Midland wells. The proposed approach performed well in propagating the model uncertainty to the uncertainty in forecast. The results showed that:

- It is confirmed that the Arps model ($0 \leq b \leq 1$) may not be ideal for unconventional plays.
- The Pan CRM, as the only analytical model, is more likely to better describe the unconventional flow behavior than the other three empirical models.
- No one model is the most likely for all wells.
- The model uncertainty can remain large even when the data length is long.
- The Arps and LGM models tends to give a larger estimate of the expected value of the cumulative oil production than the other two models for the wells with small number of data points. Using the Aprs or LGM model with small number of data points may result in a too optimistic estimate in cumulative oil production.
- By weighting the models, the proposed approach gives a moderate estimate of the cumulative oil production; neither too optimistic or too pessimistic.

Acknowledgements

The authors acknowledge the support of the Research Council of Norway and industry partners: ConocoPhillips, Skandinavia AS, BP Norge AS, Det Norske Oljeselskap AS, Eni Norge AS, Maersk Oil Norway AS, DONG Energy A/S, Denmark, Statoil Petroleum AS, ENGIE E&P NORGE AS, Lundin Norway AS, Halliburton AS, Schlumberger Norge AS, and Wintershall Norge AS of The National IOR Centre of Norway.

Aojie Hong would like to thank the staffs and colleagues at the University of Texas at Austin for their support during his visiting. Larry W. Lake holds the Shahid and Sharon Ullah Chair at the University of Texas.

References

Arps, J. J. 1945. Analysis of Decline Curves. *Transactions of the AIME* **160** (1): 228–247. https://doi.org/10.2118/945228-G.

Bunn, D. W. 1975. A Bayesian Approach to the Linear Combination of Forecasts. *Journal of the Operational Research Society* **26** (2): 325–329. http://dx.doi.org/10.1057/jors.1975.67.

Bratvold, R. B., and Begg, S. H. 2008. I Would Rather be Vaguely Right than Precisely Wrong: a New Approach to Decision Making in the Petroleum Exploration and Production Industry. *AAPG bulletin* **92** (10): 1373–1392. http://dx.doi.org/10.1306/06040808070.

Bratvold, R. B., and Begg, S. H. 2009. Would You Know a Good Decision if You Saw One? *The Way Ahead* **5** (2): 21–23. SPE-0209-021-TWA. https://doi.org/10.2118/0209-021-TWA.

Cheng, Y., Wang, Y., McVay, D., and Lee, W. J. 2010. Practical Application of a Probabilistic Approach to Estimate Reserves Using Production Decline Data. *SPE Economics & Management* **2** (1): 19–31. SPE-95974-PA. http://dx.doi.org/10.2118/95974-PA.

Clark, A. J., Lake, L. W., & Patzek, T. W. 2011. Production Forecasting with Logistic Growth Models. Society of Petroleum Engineers. Presented at the SPE Annual Technical Conference and Exhibition, Denver, USA, 30 October–2 November. SPE-144790-MS. http://dx.doi.org/10.2118/144790-MS.

Clemen, R. T., and Winkler, R. L. 1999. Combining Probability Distributions from Experts in Risk Analysis. *Risk Analysis* **19** (2): 187–203. http://dx.doi.org/ 10.1111/j.1539-6924.1999.tb00399.x.

Duong, A. N. 2011. Rate-Decline Analysis for Fracture-Dominated Shale Reservoirs. Society of Petroleum Engineers. *SPE Reservoir Evaluation & Engineering* **14** (3): 377–387. SPE-137748-PA. http://dx.doi.org/10.2118/137748-PA.

Dillon, R. L., John, R., and von Winterfeldt, D. 2002. Assessment of Cost Uncertainties for Large Technology Projects: a Methodology and an Application. *Interfaces* **32** (4): 52–66.

Gong, X., Gonzalez, R., McVay, D. A., and Hart, J. D. 2014. Bayesian Probabilistic Decline-Curve Analysis Reliably Quantifies Uncertainty in Shale-Well-Production Forecasts. *SPE Journal* **19** (6): 1047–1057. SPE-147588-PA. http://dx.doi.org/10.2118/147588-PA.

Gonzalez, R. A., Gong, X., and McVay, D. A. 2012. Probabilistic Decline Curve Analysis Reliably Quantifies Uncertainty in Shale Gas Reserves Regardless of Stage of Depletion. Presented at the SPE Eastern Regional Meeting, Lexington, USA, 3–5 October. SPE-161300-MS. http://dx.doi.org/10.2118/161300-MS.

Hong, A. J., Bratvold, R. B., and Nævdal, G. 2017. Robust Production Optimization with Capacitance-Resistance Model as Precursor. *Computational Geosciences*. https://doi.org/10.1007/s10596-017-9666-8.

Haselton, M. G., Nettle, D., and Murray, D. R. 2005. The Evolution of Cognitive Bias. *The Handbook of*

Evolutionary Psychology: 724–746. https://doi.org/10.1002/9781119125563.evpsych241.

Ilk, D., Rushing, J. A., Perego, A. D., and Blasingame, T. A. 2008. Exponential vs. Hyperbolic Decline in Tight Gas Sands: Understanding the Origin and Implications for Reserve Estimates Using Arps' Decline Curves. Presented at the SPE Annual Technical Conference and Exhibition, Denver, USA, 21–24 September. SPE-116731-MS. http://dx.doi.org/10.2118/116731-MS.

Jouini, M. N., and Clemen, R. T. 1996. Copula Models for Aggregating Expert Opinions. *Operations Research* **44** (3): 444–457. http://dx.doi.org/10.1287/opre.44.3.444

Joshi, K., and Lee, W. J. 2013. Comparison of Various Deterministic Forecasting Techniques in Shale Gas Reservoirs. Presented at the SPE Hydraulic Fracturing Technology Conference, The Woodlands, USA, 4–6 February. SPE-163870-MS. http://dx.doi.org/10.2118/163870-MS.

Jochen, V. A., and Spivey, J. P. 1996. Probabilistic Reserves Estimation Using Decline Curve Analysis with the Bootstrap Method. Presented at the SPE Annual Technical Conference and Exhibition, Denver, USA, 6–9 October. SPE-36633-MS. http://dx.doi.org/10.2118/36633-MS.

Pan, Z. 2016. *Revised Productivity Index Equation to Improve Transient History Match for the Capacitance Resistance Model*. Master Thesis. The University of Texas at Austin, USA.

Read, C. 1920. *Logic: Deductive and Inductive*, reprinted. London, UK: Simkin.

Stone, M. 1961. The Opinion Pool. *The Annals of Mathematical Statistics* **32** (4), 1339–1342. http://dx.doi.org/10.1214/aoms/1177704873.

Sayarpour, M., Kabir, C. S., Sepehrnoori, K., and Lake, L. W. 2011. Probabilistic History Matching with the Capacitance–Resistance Model in Waterfloods: a Precursor to Numerical Modeling. *Journal of Petroleum Science and Engineering* **78** (1), 96–108. https://doi.org/10.1016/j.petrol.2011.05.005.

Tavassoli, Z., Carter, J. N., and King, P. R. 2004. Errors in History Matching. *SPE Journal* **9** (3): 352–361. SPE-86883-PA. http://dx.doi.org/10.2118/86883-PA

Valko, P. P. 2009. Assigning Value to Stimulation in the Barnett Shale: a Simultaneous Analysis of 7000 plus Production Histories and Well Completion Records. Presented at the SPE Hydraulic Fracturing Technology Conference, The Woodlands, USA, 19–21 January. SPE-119369-MS. http://dx.doi.org/10.2118/119369-MS.

Valko, P. P., and Lee, W. J. 2010. A Better Way to Forecast Production from Unconventional Gas Wells. Presented at the SPE Annual Technical Conference and Exhibition, Florence, Italy, 19–22 September. SPE-134231-MS. http://dx.doi.org/10.2118/134231-MS

Welsh, M. B., Bratvold, R. B., and Begg, S. H. 2005. Cognitive Biases in the Petroleum Industry: Impact and Remediation. Presented at the SPE Annual Technical Conference and Exhibition, Dallas, USA, 9–12 October. SPE-96423-MS. https://doi.org/10.2118/96423-MS.

Wattenbarger, R. A., El-Banbi, A. H., Villegas, M. E., and Maggard, J. B. 1998. Production Analysis of Linear Flow into Fractured Tight Gas Wells. Presented at the SPE Rocky Mountain Regional/Low-Permeability Reservoirs Symposium, Denver, USA, 5–8 April. SPE-39931-MS. http://dx.doi.org/10.2118/39931-MS.

Paper II

Robust Production Optimization with Capacitance-Resistance Model as Precursor

Hong, A.J., Bratvold, R.B., and Nævdal, G.

Comput Geosci (2017) 21:1423–1442
DOI 10.1007/s10596-017-9666-8

ORIGINAL PAPER

Robust production optimization with capacitance-resistance model as precursor

A. J. Hong[1,2] · **R. B. Bratvold**[1,2] · **G. Nævdal**[1,3]

Received: 13 October 2016 / Accepted: 24 May 2017 / Published online: 24 June 2017

Abstract Many model-based techniques for optimizing hydrocarbon production, especially robust optimization (RO), carry prohibitive computational cost. Ensemble-based optimization (EnOpt) is a promising RO method but is computationally intensive when based on rich grid-based reservoir models with hundreds of realizations. We present a proxy-model workflow where a grid-based model is supplemented by a useful yet tractable proxy model. A capacitance-resistance model (CRM) can be a proxy model for waterflooding systems. We illustrate the use of CRM-based models and investigate their pros and cons using synthetic 2D and 3D models. A selected proxy model is embedded into the proxy-model workflow. The results obtained from the proxy-model and traditional workflows are compared. The impact of any differences is assessed by considering a relevant decision-making context. The main contributions are (1) a general RO workflow that embeds proxy models, (2) a discussion of the desiderata of proxy models, (3) illustration and discussion of the use of CRM-based models in the proxy-model workflow, and (4) a discussion of the impact of using a proxy model for production optimization in a decision-making context. Based on our study, we conclude that CRM-based models have high potential to serve as a cogent proxy model for waterflooding related decision-making context and that the proxy-model workflow, leveraging a faster, but relevant, production model, significantly speeds up the optimization yet gives robust results that leads to a near-optimal solution.

Keywords Model-based hydrocarbon production optimization · Geological uncertainty · Robust production optimization · Ensemble-based optimization · Computational cost · Grid-based reservoir model · Capacitance-resistance model · Proxy model · Water injection · Decision-making · Value of verisimilitude

1 Introduction

The past few decades have seen rapid development in numerical techniques for model-based optimization of subsurface hydrocarbon production (to optimize the reserves, the production over some time frame, the value of the reserves or production, etc.). However, these techniques typically require computation for a numerous reservoir simulations, especially for robust optimization (RO).

Within RO, the geological uncertainties are represented by a set of realizations (i.e., an ensemble). The objective of RO is to find a control vector (e.g., a water injection scheme) that optimizes the expected value (EV) of the objective function, such as net present value (NPV), given the geological uncertainties represented by the realizations. Hence, the optimal solution is robust to geological uncertainties. van Essen et al. [33] used RO to optimize hydrocarbon production under geological uncertainty, where an adjoint-based

✉ A. J. Hong
aojie.hong@uis.no

R. B. Bratvold
reidar.bratvold@uis.no

G. Nævdal
geir.naevdal@iris.no

[1] The National IOR Centre of Norway, Stavanger, Norway

[2] Department of Petroleum Engineering, University of Stavanger, Stavanger, Norway

[3] International Research Institute of Stavanger, Stavanger, Norway

method is used for obtaining the gradient information. The adjoint-based method requires access to the source code of the reservoir simulator, which is seldom available for commercial simulators, and it is computationally intensive.

Chen et al. [13] introduced the ensemble-based optimization method (EnOpt), in which the gradient is approximated by the covariance between the objective function values and the control variables. Regardless of the type of reservoir simulators, EnOpt is easier to implement than other RO methods and is less computationally intensive [17]. Do and Reynolds [15] analyzed the theoretical connections between EnOpt and other approximate gradient-based optimization methods. Having realized that it is unnecessary to approximate the ensemble mean to the sample mean as was done by Chen et al. [13], Do and Reynolds [15] used the ensemble mean in their EnOpt formulation.

Fonseca et al. [17] further modified Do and Reynolds's [15] EnOpt formulation and found that the modified EnOpt formulation[1] both gives a better objective function value and converges more quickly than the original EnOpt and other variants of ensemble-based optimization method. Although it has been demonstrated that EnOpt (both the original formulation and its variants) is a vast improvement over earlier optimization methods, it is nonetheless computationally intensive, typically involving thousands of reservoir simulations [29]. The differences between the original and modified EnOpt formulations will be stated in detail in Section 3.

Yang et al. [35] used a second order polynomial proxy model to reduce the number of required simulations in RO. Denney [14] discussed the pros and cons of applying a proxy model as a substitute for full reservoir simulations in assisted history matching, production optimization and prediction, and found that proxy models do not give an optimal solution in many cases but are less computationally demanding in finding improved solutions. Our paper presents an alternative, and general, workflow where a grid-based reservoir model is supplemented by a material-balance-based proxy model for RO. This proxy model is intended to capture the range of possible production profiles, yet be less computationally intensive than the reservoir simulators currently used in EnOpt. Thus, the choice and usefulness of a proxy model for this purpose will be a function of the reservoir characteristics and drainage scheme. Our main focus is waterflooding systems for which capacitance-resistance model (CRM) is a potential candidate.

Bruce [9] applied the analogy of a capacitance-resistance electric network to analyzing reservoir and well behavior. Albertoni and Lake [1] introduced a model that combined a multivariate linear regression analysis with diffusivity filters and used only production and injection rate data to investigate the connectivity and response time between a producer and an injector in a waterflooded reservoir. Gentil [18] showed that the connectivity is a function of transmissibilities. Yousef et al. [36] introduced CRM by modifying Albertoni and Lake's work [1] to use the weight to quantify the connectivity and to employ the time constant to quantify the fluid storage within each injector-producer pair. Sayarpour [27] used superposition in time to analytically solve the numerical integration in CRM.

To overcome the original CRM's limitation that the total fluid flow is treated as a single phase, a CRM was combined by Sayarpour et al. [28] with a Buckley-Leverett-based fractional flow model, by Gentil [18] with an empirical fractional flow equation, and by Cao et al. [12] with the Koval [22] model. Cao [10] demonstrated that the Koval model might not yield a good match for mature waterfloods because it might approach an abrupt end of 100% water cut. She also noted that the Gentil model might not work well for immature waterfloods because the relationship between the natural logs of water-oil ratio and cumulative water injection might be non-linear at that stage. These concerns prompted the development of a fully coupled two-phase-flow-based CRM, which is applicable in all stages of maturity [10, 11].

To allow for the inherent uncertainty in history matching [31], Sayarpour et al. [28] started with different sets of initial guesses of unknown CRM parameters and used production data to generate numerous history-matched solutions of CRM. Jafroodi and Zhang [21] applied a CRM in a closed loop consisting of ensemble Kalman filter (EnKF) and EnOpt to capture the geological uncertainties and the time-variance of the CRM parameters. CRM has been shown to be quite accurate for waterflooding systems and used as a tool for waterflood production optimization [24, 30].

The above studies on CRM employed it in the production phase with real production data (in the synthetic studies, the "real" production data is mimicked by the synthetic model simulated production data). In the proxy-model workflow, we do not need any real production data for using CRM because the CRM is history-matched to grid-based model simulated production data (i.e., pseudo production data). Thus, the proxy-model workflow can be used when real production data is unavailable (e.g., in the design phase).

In Section 2, we review two CRM-based models, a CRM combined with the Koval model (the CRMP-Koval Model) and the fully coupled two-phase flow based CRM (the Coupled CRMP). Section 3 gives a short introduction on RO and EnOpt. Section 4 illustrates the proxy-model workflow. In Sections 5 and 6, this workflow is tested in two examples; one with a 2D reservoir simulation model and the other with a 3D model. The matching qualities of the two CRM-based

[1] We refer the EnOpt formulation modified by Fonseca et al. [17] to as the modified EnOpt formation in the rest of the paper.

models are compared for screening purposes in the 2D model example. The selected proxy model is embedded into the proxy-model workflow. The results of the proxy-model and traditional workflows are contrasted and discussed. Furthermore, the impacts of these results are investigated in a decision-making context. Finally, we present a discussion and conclusions in Sections 7 and 8, respectively.

The main contributions of this work are (1) a general RO workflow embedding proxy models, (2) a discussion of the desiderata of proxy models for production optimization, (3) illustration and discussion of the use of CRM-based models in the proxy-model workflow, and (4) a discussion of the impact of using a proxy model for production optimization in a decision-making context.

2 Capacitance-resistance model

A capacitance-resistance model (CRM) is based on material balance and derived from total fluid continuity equation. It contains considerably fewer parameters and needs significantly less computation time than does a grid-based reservoir model. Required input data for a CRM are production rates, injection rates, and producers' Bottom-Hole Pressures (BHP). The two main parameters of a CRM are connectivity and time constant. For an oil-water system, connectivity is the proportion of injected water in an injector that contributes to the total fluid production in a producer. The time constant indicates how long a pressure wave from an injector takes to reach a producer.

Based on control volume, CRM can be divided into three categories: single-tank CRM (CRMT), producer-based CRM (CRMP), and injector-producer-pair-based CRM (CRMIP) [30]. As illustrated in Fig. 1, the control volume in a CRMT is the entire drainage volume of a reservoir covering a single pseudo-producer (encompassing all physical producers) and a single pseudo-injector (encompassing all physical injectors), enabling the entire reservoir to be treated as a single tank with one inlet and one outlet. A CRMP has producer-based control volumes, each of which covers all the injectors influencing its corresponding producer. A CRMIP has one control volume for each injector-producer pair. Because the original CRM [36] can predict only total fluid production, it must be combined with a fractional flow model in order to separate oil production from the total production. In the following, we review two CRMP-based models for the case of an oil-water system under waterflooding and constant producer BHPs.

2.1 The CRMP-Koval Model

The Koval factor [22] was introduced into the Buckley-Leverett fractional flow equation in order to capture the fingering effect induced by heterogeneity and the unfavorable mobility ratio between the displacing and displaced fluids. Cao et al. [12] illustrated how CRMP combined with the Koval fractional flow equation (the CRMP-Koval Model) can be used to forecast oil and water production under waterflooding. For constant producer BHPs, the total fluid production equation of CRMP is [27]

$$q_{t_j}^k = q_{t_j}^{k-1} e^{-\Delta t_k/\tau_j} + (1 - e^{-\Delta t_k/\tau_j})(\sum_{i=1}^{n_{inj}} \lambda_{ij} I_i^k), \tag{1}$$

where sub/superscripts i, j, and k are the indices of injector, producer, and time step, respectively; $q_{t_j}^k$ is the total fluid production rate of producer j at time k; Δt_k is the time step length between times k and $k-1$; τ_j is the time constant for producer j; n_{inj} is the number of injectors; λ_{ij} is the connectivity between injector i and producer j; and I_i^k is the water injection rate of injector i during the period of Δt_k. Both λ and τ are assumed to be constant with respect to time [12].[2]

Cao et al. [12] used the following form of the Koval fractional flow equation:

$$f_{w_j}^k = \begin{cases} 0, & \text{if } t_{D\,j} < \frac{1}{K_{val\,j}} \\ \frac{K_{val\,j} - \sqrt{K_{val\,j}/t_{D\,j}}}{K_{val\,j}-1}, & \text{if } \frac{1}{K_{val\,j}} \le t_{D\,j} \le K_{val\,j} \\ 1, & \text{if } t_{D\,j} > K_{val\,j} \end{cases}, \tag{2}$$

where

$$t_{D\,j} = \frac{\sum_{k=1}^{n_T} I_{t_j}^k}{V_{p\,j}}, \tag{3}$$

$f_{w_j}^k$ is the water cut in producer j at time k, $K_{val\,j}$ is the Koval factor of producer j, $t_{D\,j}$ is the dimensionless time or the fraction of cumulative water injected into the drainage pore volume of producer j, n_T is the number of time steps, $I_{t_j}^k$ is the total water injection contribution from all injectors to producer j during Δt_k, and $V_{p\,j}$ is the drainage pore volume of producer j. Assuming that the water injection rate of an injector is constant over Δt_k and equal to I_i^k, $I_{t_j}^k$ can be calculated by

$$I_{t_j}^k = (\sum_{i=1}^{n_i} \lambda_{ij} I_i^k)\Delta t_k. \tag{4}$$

Thus, using the CRMP and the Koval fractional flow equation, the water and oil production rates in producer j at time k ($q_{w_j}^k$ and $q_{o_j}^k$) are calculated, respectively, by

$$q_{w_j}^k = q_{t_j}^k f_{w_j}^k, \tag{5}$$

$$q_{o_j}^k = q_{t_j}^k (1 - f_{w_j}^k). \tag{6}$$

[2] This is equivalent to assuming no dramatic changes in reservoir or well conditions.

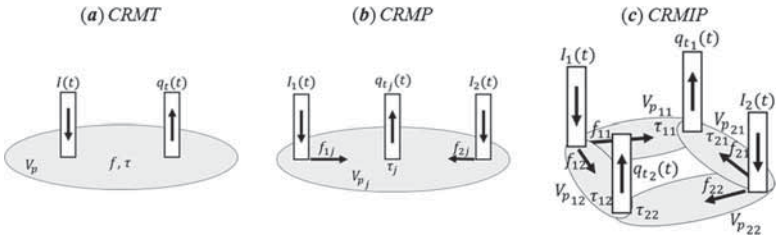


Fig. 1 Illustration of **a** CRMT, **b** CRMP, and **c** CRMIP

In the CRMP-Koval Model, the model parameters are the connectivity λ of each injector-producer pair; and the time constant τ, the Koval factor K_{val}, and the drainage pore volume V_p of each producer. These parameters are determined by solving the least squares problem

$$\min_{\tau_j,\lambda_{ij},K_{val\,j},V_{p_j}} z = \sum_{k=1}^{n_T}\sum_{j=1}^{n_p}[(q_{t_j}^{\,k}-q_{t_j}^{\,k^{obs}})^2+(q_{o_j}^{\,k}-q_{o_j}^{\,k^{obs}})^2] \tag{P1}$$

subject to

$$\tau_j \geq 0, \tag{P1C1}$$

$$\lambda_{ij} \geq 0, \tag{P1C2}$$

$$K_{val\,j} \geq 1, \tag{P1C3}$$

$$\sum_{j=1}^{n_p} V_{p\,j} \leq V_{p\,Field}, \tag{P1C4}$$

$$\sum_{j=1}^{n_p} \lambda_{ij} \begin{cases} < 1, & \text{if injection loss exists} \\ = 1, & \text{if no injection loss} \end{cases} \tag{P1C5}$$

where the superscript *obs* denotes the observed production data, n_p is the number of producers, and $V_{p\,Field}$ is the total pore volume of the field. For an optimization program, a different equation numbering system is used, where "P" denotes a program and "C" denotes a constraint. For example, Eq. P1C5 means constraint number 5 for optimization program 1. For the case with aquifer support, Izgec and Kabir [20] proposed coupling an aquifer model with a CRM. Because no aquifer is considered in the reservoir models used in the following examples, the effect of an aquifer will not be considered in this paper. The total pore volume of the field $V_{p\,Field}$ is defined by the total pore volume of the reservoir simulation model because in the proxy-model workflow, the observed production data are obtained from reservoir simulation using the commercial black oil simulator ECLIPSE^{TM} [16], and the CRM is used to approximate the reservoir simulation model. When $V_{p\,Field}$ is unknown, such as when the CRM is used to match real production data, constraint Eq. P1C4 can be removed.

2.2 The Coupled CRMP

Cao et al. [11] developed the Coupled CRMP: a fully coupled two-phase flow model based on CRMP. In addition to the total fluid balance equation (i.e., the pressure or continuity equation), they considered the oil mass balance equation (i.e., the saturation equation). This model can calculate the average saturation within a producer-based control volume, the outlet saturation of a producer, and the time-variant time constant. The following shows only the numerical solution to the saturation equation. For more details on the derivation, see [10, 11].

Cao et al. [11] derived the semi-analytical explicit solution to the saturation equation as

$$\overline{S}_{o\,j}^{\,k} = \overline{S}_{o\,j}^{\,k-1} - \frac{\Delta t_k}{V_{p\,j}}\left(\frac{\overline{S}_{o\,j}^{\,k-1}(c_f+c_o)}{c_t}\left(\sum_i \lambda_{ij} I_j^k - q_{t\,j}^{\,k}\right) + q_{o\,j}^{\,k}\right), \tag{7}$$

where $\overline{S}_{o\,j}^{\,k}$ is the average oil saturation within the control volume of producer j at time k, and c_f, c_o, and c_t are the pore, oil, and total compressibilities, respectively. Assuming that the major contribution to oil saturation change is from oil production rather than the effect of compressibility [23], Cao [10] eliminated the compressibility term in the saturation equation and derived a simplified solution as

$$\overline{S}_{o\,j}^{\,k} = \overline{S}_{o\,j}^{\,k-1} - \frac{\Delta t_k}{V_{p\,j}} q_{o\,j}^{\,k}. \tag{8}$$

We will be using Eq. 8 rather than Eq. 7. Cao [10] proposed using the Welge equation [10, 23, 34] to calculate the outlet oil saturation of a producer as

$$S_{o2\,j}^{\,k} = \overline{S}_{o\,j}^{\,k} + \frac{\sum_{k=1}^{n_T} I_{t\,j}^{\,k}}{V_{p\,j}}(1 - f_{w\,j}^{\,k}), \tag{9}$$

where $S_{o2\,j}^{\,k}$ is the outlet oil saturation of producer j at time k.

Using the definition of the time constant,

$$\tau_j^k = \frac{V_{pj} c_t}{J_{tj}^k}, \tag{10}$$

where J_{tj}^k is the total productivity index of producer j at time k, Eq. 10 can be rewritten as

$$\tau_j^k = \frac{T_j}{M_j^k}, \tag{11}$$

where

$$T_j = \frac{V_{pj} c_t}{J'_{tj}}, \tag{12}$$

$$M_j^k = \frac{k_{ro}(S_{o2}{}_j^k)}{\mu_o} + \frac{k_{rw}(S_{o2}{}_j^k)}{\mu_w}. \tag{13}$$

Because J'_{tj} is assumed to be constant, T_j is time-invariant [11]. The total relative mobility M_j^k is time-variant because it depends on the oil and water relative permeabilities, k_{ro} and k_{rw}, which are functions of the outlet saturation. The oil and water viscosities, μ_o and μ_w, are assumed to be constant. It should be noted that the time constant is assumed to be time-invariant in the CRMP-Koval Model, whereas it is time-variant in the Coupled CRMP. Thus, the time-invariant time constant τ_j in the solution of the total fluid continuity equation (1) is replaced by the time-variant time constant τ_j^k, giving

$$q_{tj}^k = q_{tj}^{k-1} e^{-\Delta t_k/\tau_j^k} + (1 - e^{-\Delta t_k/\tau_j^k}) \left(\sum_{i=1}^{n_{inj}} \lambda_{ij} I_i^k \right). \tag{14}$$

The model parameters can be determined by solving the least squares problem

$$\min_{T_j, \lambda_{ij}, \overline{S}_{oj}^0, V_{pj}} z = \sum_{k=1}^{n_T} \sum_{j=1}^{n_p} (q_{tj}^k - q_{tj}^{k\,obs})^2 \tag{P2}$$

subject to

$$T_j \geq 0, \tag{P2C1}$$

$$\lambda_{ij} \geq 0, \tag{P2C2}$$

$$S_{or} \leq \overline{S}_{oj}^k, \tag{P2C3}$$

$$S_{o2}{}_j^k \leq 1 - S_{wi}, \tag{P2C4}$$

$$\sum_{j=1}^{n_p} V_{pj} \leq V_{pField}, \tag{P2C5}$$

$$\sum_{j=1}^{n_p} \lambda_{ij} \begin{cases} < 1, & \text{if injection loss exists} \\ = 1, & \text{if no injection loss} \end{cases} \tag{P2C6}$$

where $\overline{S}_{oj}^0$ is the initial average oil saturation of producer j, S_{or} is the residual oil saturation, and S_{wi} is the irreducible water saturation. For the Coupled CRMP, the model parameters are determined by matching only the total fluid production data. This is because the water cut f_{wj}^k in Eq. 9 can be directly obtained from the production data. The fractional flow model is constructed using the water cut data plotted against the average oil saturation calculated from the Coupled CRMP. Cao [10] used a polynomial function of degree 3 to fit the fractional flow curve based on the historical data. We construct the fractional flow curve using a data table consisting of existing water cut data against the calculated average oil saturation from the Coupled CRMP at each time for each producer.

Using the Coupled CRMP for production prediction is not as straightforward as using the CRMP-Koval Model because production rates and saturations at any given time are mutually dependent. In order to predict production with minimal errors, production rates and saturations have to be solved implicitly with iterations. For more details on the iterative algorithm, see [10]. For a given predicted average oil saturation, the water cut can be predicted by extrapolation or interpolation of the water cut against average oil saturation table constructed earlier. We use linear extrapolation or interpolation to find the corresponding water cut for a given average oil saturation, based on the data in the table.

3 Robust optimization (RO) of production

RO is performed over an ensemble of realizations representing the geological uncertainties. For a risk-neutral decision-maker, the objective of RO is to optimize the EV over of the ensemble. The purpose of production optimization is to maximize the NPV. Assuming that all revenues are from oil production and that all costs are induced by water injection and water production, the objective function for a single realization can be defined as

$$\begin{aligned} J(\mathbf{u}) &= NPV(\mathbf{u}) \\ &= \textstyle\sum_{k=1}^{n_T} \frac{(q_o^k(\mathbf{u}) P_o - q_w^k(\mathbf{u}) P_{wp} - I^k(\mathbf{u}) P_{wi}) \Delta t_k}{(1+b)^{t_k/D}}, \end{aligned} \tag{15}$$

where $\mathbf{u}$ is the control vector (i.e., a vector of control variables) defined as $\mathbf{u} = [u_1, u_2, ..., u_N]^{\mathrm{T}}$, where N is the number of control variables; q_o^k is the field oil production rate at time k; q_w^k is the field water production rate at time k; I^k is the field water injection rate at time k; P_o, P_{wp}, and P_{wi} are the oil price, water production cost, and water injection cost, respectively; b is the discount factor; t_k is the cumulative time for discounting; and D is the reference time for discounting ($D = 365$ days if b is expressed as a fraction per year and the cash flow is discounted daily). If

a case involves multiple realizations, the objective function becomes

$$\overline{J}(\mathbf{u}) = \frac{\sum_{r=1}^{n_e} J_r(\mathbf{u})}{n_e}, \tag{16}$$

where $\overline{J}$ is the EV of the objective functions over all realizations, J_r is the objective function (15) of a single realization r, and n_e is the number of realizations (i.e., ensemble size).

EnOpt method can significantly reduce the number of simulations required for gradient calculation and can be easily implemented with any type of simulator. In the original EnOpt, the ensemble of values for control vector $\mathbf{u}_1$ $\mathbf{u}_2$ ... $\mathbf{u}_M$, where M is the ensemble size,[3] is generated from a multivariate normal distribution with predefined mean $\widehat{\mathbf{u}}$ and predefined covariance matrix $\mathbf{C_u}$, which is used to specify the temporal correlation of the controls to limit the frequency of changes in the controls. The predefined mean $\widehat{\mathbf{u}}$ can be approximated by its sample mean $\overline{\mathbf{u}}$, i.e.,

$$\widehat{\mathbf{u}} \approx \overline{\mathbf{u}} = \frac{\sum_{r=1}^{M} \mathbf{u}_r}{M}, \tag{17}$$

and the average objective value $\overline{J}(\widehat{\mathbf{u}})$ can be approximated by

$$\overline{J}(\widehat{\mathbf{u}}) = \frac{\sum_{r=1}^{M} J_r(\widehat{\mathbf{u}})}{M} \approx \frac{\sum_{r=1}^{M} J_r(\mathbf{u}_r)}{M}. \tag{18}$$

Chen et al. [13] used these two approximations to calculate the mean-shifted ensemble matrix $\overline{\mathbf{U}}$ and the mean-shifted objective function vector $\mathbf{j}$, respectively. However, Do and Reynolds [15] did not find any advantage, theoretical or practical, in approximating these two terms. Hence, the mean-shifted ensemble matrix should be calculated directly using the predefined mean $\widehat{\mathbf{u}}$ by

$$\overline{\mathbf{U}} = [\mathbf{u}_1 - \widehat{\mathbf{u}}, \mathbf{u}_2 - \widehat{\mathbf{u}}, ..., \mathbf{u}_M - \widehat{\mathbf{u}}]^{\mathrm{T}}. \tag{19}$$

Moreover, Fonseca et al. [17] suggested calculating the mean-shifted objective function vector with respect to the objective value of the predefined mean for each individual realization instead of the average objective value (18), i.e.,

$$\mathbf{j} = [J_1(\mathbf{u}_1) - J_1(\widehat{\mathbf{u}}), J_2(\mathbf{u}_2) - J_2(\widehat{\mathbf{u}}), ..., J_M(\mathbf{u}_M) - J_M(\widehat{\mathbf{u}})]^{\mathrm{T}}. \tag{20}$$

The cross-covariance matrix is then

$$\mathbf{C_{uJ}} = \frac{1}{M}(\overline{\mathbf{U}}^{\mathrm{T}}\mathbf{j}). \tag{21}$$

Chen et al. [13] approximated the gradient by

$$\mathbf{g} \approx \mathbf{C_{uJ}}. \tag{22}$$

For each iteration, the control vector is updated as

$$\widehat{\mathbf{u}}_{l+1} = \widehat{\mathbf{u}}_l + \alpha_l \frac{\mathbf{g}_l}{\|\mathbf{g}_l\|_\infty}, \tag{23}$$

where the subscript l denotes the iteration number and α is the step length for updating. Equations 19–23 constitute the modified EnOpt formulation [17]. The original EnOpt formulation can be obtained by replacing $\widehat{\mathbf{u}}$ with $\overline{\mathbf{u}}$ in Eq. 19 and $J_r(\widehat{\mathbf{u}})$ with the approximated $\overline{J}(\widehat{\mathbf{u}})$ in Eq. 20. Fonseca et al. [17] showed that the modified EnOpt converges to a higher objective value and more quickly than does the original EnOpt. Therefore, we use the modified EnOpt. The step length α significantly affect convergence speed. A naive line search procedure is to simply reduce α by half if $\overline{J}(\widehat{\mathbf{u}}_{l+1}) < \overline{J}(\widehat{\mathbf{u}}_l)$. We use an interpolation-based line search procedure to find a relatively large value of α that satisfies the Armijo condition [26]. The line search procedure halts when $\overline{J}(\widehat{\mathbf{u}}_{l+1}) > \overline{J}(\widehat{\mathbf{u}}_l)$ or after seven iterations, whichever comes first. The modified EnOpt halts after 30 iterations of control vector updating. In our case, an optimal solution is found during the first 5–15 iterations. However, we continue to iteration 30 in order to observe a clear evidence of convergence.

4 The proxy-model workflow

When geological uncertainty is considered and is represented by hundreds of realizations, EnOpt usually requires thousands of production prediction runs. In the traditional workflow, production is predicted by running grid-based reservoir simulation. The computational intensiveness of a grid-based reservoir model can be reduced by having a proxy model serve as a precursor. A proxy model must be able to serve our purpose (i.e., useful), and its computational time must be low (i.e., tractable). See the Discussion section for the descriptions of "useful" and "tractable". Figure 2 depicts the traditional and proxy-model workflows.

A proxy model should be chosen based on the reservoir characteristics and drainage scheme. More than one potential proxy models may be found. For example, we consider two CRM-based models for waterflooding. Thus, we need to screen these potential proxy models by testing their qualities of matching and prediction. 'Pseudo production data" are generated by reservoir simulation that includes a random injection scheme.[4] A history-matching process is used to tune the parameters of a proxy model to where the production predicted by the proxy model fits the

[3] $M = n_e$ if each realization is coupled with one sample of values for the control vector, i.e., a 1:1 ratio is applied, as in our case.

[4] We use the term "pseudo production data" because the data are simulated rather than measured.

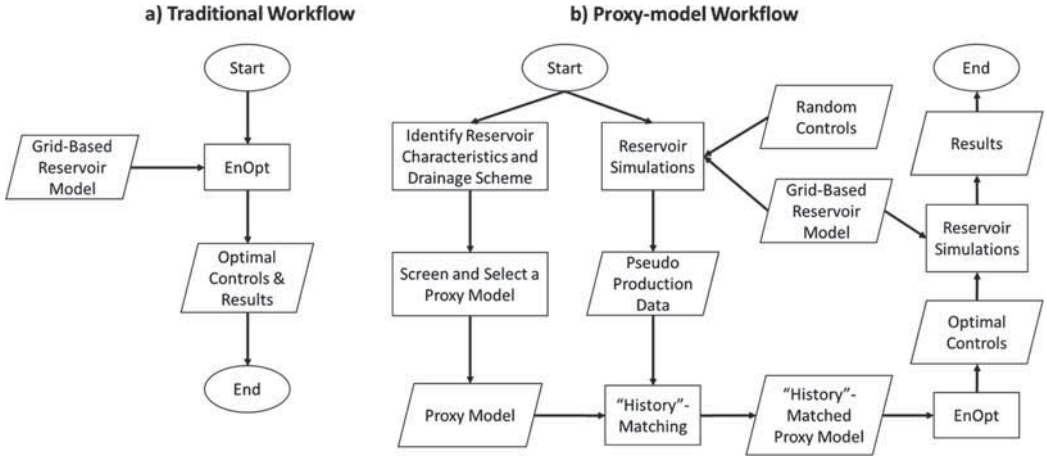


Fig. 2 The traditional and proxy-model workflows

pseudo production data. The production predicted by the history-matched proxy model is then compared to that of the grid-based reservoir model, using another random injection scheme. If the difference is unacceptably large either at the history-matching step or at the validation step, we test another potential proxy model. If the difference is small at both steps, we use the proxy model for production optimization. After using the optimization algorithm with the proxy model, we run simulation again on the grid-based reservoir model with the optimal control vector of the proxy model to get the final results (e.g., NPV). We conduct the last step for evaluating the optimal control vectors of the reservoir simulation and proxy models on the same basis. This will be discussed later.

5 Example with a 2D reservoir model

5.1 Description of the 2D reservoir model

This example uses a synthetic 2D reservoir model ("2D Model") to test the CRM-based models and the proxy-model workflow. The 2D Model is an isotropic heterogeneous model with four injectors and one producer in a five-spot pattern. The heterogeneity pertains to permeability only. Figure 3 shows the permeability field in millidarcy (md) and well locations of the model. The producer BHPs are fixed at 200 bars. The control variables are injection rates for 50 time intervals of 30 days each (i.e., a life cycle of 1,500 days). For this specific example, the injection rates of all injectors for a given day are assumed to be identical. Thus, there are 50 control variables.

5.2 Screening of the CRM-based models

This section examines the suitability of the CRMP-Koval Model and the Coupled CRMP for the proxy-model workflow. However, this screening process is not limited to CRM-based models. For example, we can screen and compare a CRM-based model to other useful and tractable models such as the Interwell-Numerical-Simulation Model (INSIM) [37].

Model parameter determination and matching quality We determine the parameters of a proxy model by solving a least squares problem using the interior-point optimization algorithm in MATLAB^{TM} [25]. Holanda et al. [19] introduced a matrix structure for the CRM representation that does not need the analytical equation (1) for solving the least squares problem using MATLAB. Solving the least squares problem is actually a history-matching process because we minimize the difference between the production profiles predicted by the grid-based reservoir model and the proxy

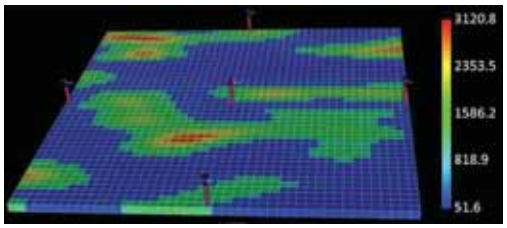


Fig. 3 Illustration of the 2D model and its permeability field (indicated by the colors) for the deterministic case. All values are in md

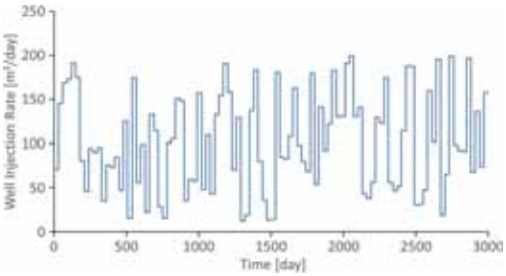


Fig. 4 Random well injection scheme for matching, 2D model example

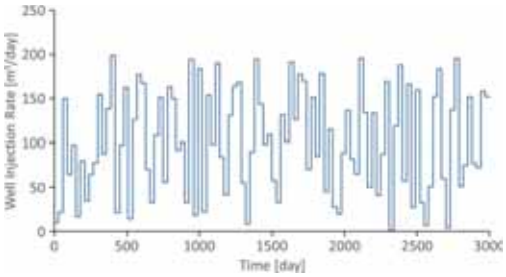


Fig. 5 Random well injection scheme for validation, 2D model example

model given an injection scheme. The well injection rates are randomly generated from 0 to 200 m^3/day (Fig. 4). The closeness of matching is depicted in Figs. 6 and 7 and is quantified by the coefficient of determination R^2, defined as

$$R^2 = 1 - \frac{\sum_k (d_k^{obs} - d_k^{pro})^2}{\sum_k (d_k^{obs} - \overline{d}^{obs})^2}, \tag{24}$$

where d_k^{obs} is the observed production data (the production predicted by the grid-based reservoir model in our case) at time k, d_k^{pro} is the production predicted by the proxy model at time k, and $\overline{d}^{obs}$ is the mean of d_k^{obs} over all k. The closer R^2 to 1, the better the matching. The values of R^2 are listed in Table 1. Both proxy models give almost equally good matches of total fluid production rate. The Coupled CRMP matches water cut better than the CRMP-Koval Model because Koval model gives an abrupt end of 100% water cut for mature waterflood (i.e., when water cut is high), as shown in Fig. 7 and demonstrated in [10, 11]. Therefore, we choose the Coupled CRMP for further application.

Model validation and prediction quality Good matching quality does not guarantee good prediction quality. This section will investigate whether the history-matched coupled CRMP can provide a production prediction close to that of the 2D Model when the injection scheme is changed. We generate a new random injection scheme as shown in Fig. 5. The history-matched Coupled CRMP can provide highly accurate prediction in total production rate, with an R^2 of 0.9998, and satisfactory prediction in water cut, with an R^2 of 0.9984, as illustrated in Figs. 8 and 9, respectively.

Table 1 R^2 of matching, deterministic case, 2D model example

	Total production rate	Water cut
CRMP-Koval	0.9998	0.9847
Coupled CRMP	0.9999	0.9999

Run time The computational time required for production prediction was around 4.1 s for the 2D Model by ECLIPSE but only around 0.17 s for the coupled CRMP by MATLAB. That is, the run time of production prediction is reduced by a factor of 24 by using Coupled CRMP instead of grid-based model. The run time of a grid-based reservoir model is highly dependent on the number of grid blocks, number of wells, and geological complexity, whereas the run time of a CRM-based model depends on only the number of wells and model chosen (CRMT, CRMP or CRMIP). Thus, a finer and/or more geologically complex grid-based reservoir model will have a run time more than 24 times greater than that of its corresponding CRM-based model.

The shorter run time of the coupled CRMP enables faster production prediction than using the 2D Model. Therefore, the run time of production optimization will be significantly reduced if we use the Coupled CRMP instead of the 2D Model for RO.

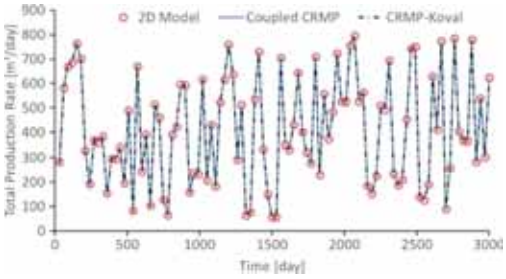


Fig. 6 Matching of total fluid production rate, 2D model example

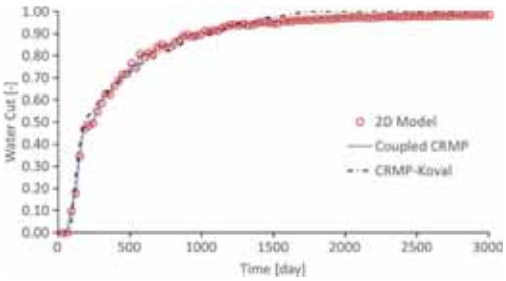


Fig. 7 Matching of water cut, 2D model example

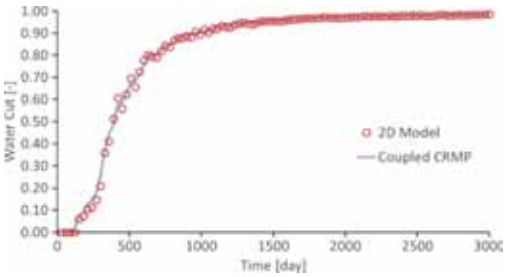


Fig. 9 Prediction of water cut, 2D model example

5.3 Life-cycle production optimization

We embed the Coupled CRMP into the proxy-model workflow to serve as a proxy for production optimization. We consider two cases: a deterministic case where the permeability field of the grid-based reservoir model is certain, and a stochastic case where the permeability is uncertain and represented by an ensemble. For both cases, we optimize NPV for 50 time steps of 30 days each, yielding a period of 1,500 days. In order to get a well-constructed fractional flow curve for the Coupled CRMP, we extend the reservoir simulation to 3,000 days for determining the model parameters so that the full range (0 to near 1) of water cut can be covered. For NPV calculation, the oil price, water production cost, water injection cost, and discount rate are set to 315 \$/m^3, 47.5 \$/m^3, 12.5 \$/m^3, and 8%, respectively.

The deterministic case The grid-based reservoir model used for this case is the single geological realization described earlier (Fig. 3). The model parameters of the Coupled CRMP are determined by matching the production data simulated by the 2D Model. The matching quality is shown in Figs. 6 and 7. The history-matched Coupled CRMP is then validated by contrasting its production prediction to that of the 2D Model given a new injection scheme. The prediction quality is shown in Figs. 8 and 9.

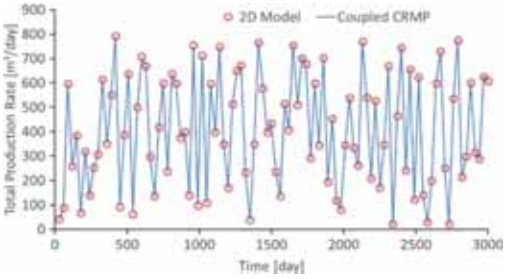


Fig. 8 Prediction of total fluid production rate, 2D model example

The well injection rate is constrained from 0 to 200 m^3/day. As a base case, the well injection rate is set to 100 m^3/day for the whole period of production.

Table 2 lists the base case NPV, the optimal NPV of the Coupled CRMP, and the solutions obtained using the traditional and proxy-model workflows. Figure 10 illustrates the optimal injection schemes of the Coupled CRMP and the 2D Model. For the 2D Model and the Coupled CRMP under the optimal injection scheme of the Coupled CRMP, Fig. 11 shows the predicted total fluid production rate and Fig. 12 shows the water cut. These results indicate that the Coupled CRMP again provides an accurate prediction, with an R^2 of 0.9998 for total production rate and an R^2 of 0.9968 for water cut.

Compared to the 2D Model, the Coupled CRMP underestimates NPV (\$31.960 million vs. \$32.406 million) for the same injection scheme. The results show that the proxy-model workflow (i.e., applying the optimal strategy from the Coupled CRMP to the 2D Model) can provide an optimal NPV (\$32.406 million) very close to that of the traditional workflow (\$32.902 million). Figure 10 shows that the proxy-model and traditional workflows result in different optimal injection schemes. This is due to the stochastic feature of EnOpt and/or the approximation of the proxy model to the grid-based reservoir model.

Table 2 NPVs [\$ million], deterministic case, 2D model example

Model	Base case	Optimal	Improvement
Coupled CRMP	23.767	31.960	34.47%
2D Model	24.476	32.406[a,b]	32.40%
2D Model	24.476	32.902[c]	34.43%

[a]With optimal injection scheme of the Coupled CRMP
[b]The solution of the proxy-model workflow
[c]The solution of the traditional workflow

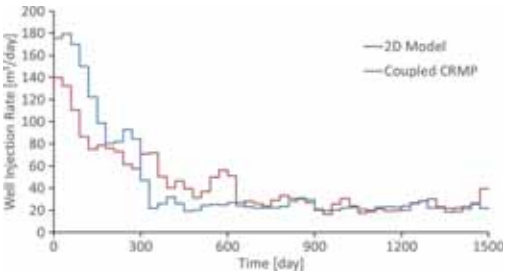


Fig. 10 Optimal injection schemes of the 2D Model and the Coupled CRMP, deterministic case

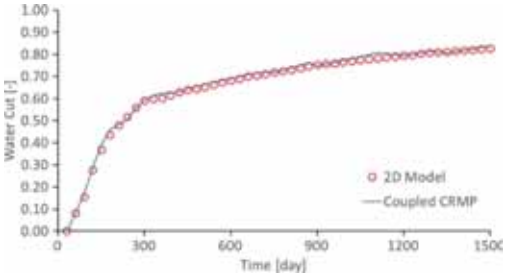


Fig. 12 Water cut under the optimal injection scheme of the Coupled CRMP, deterministic case, 2D Model example

To further compare the Coupled CRMP with the 2D Model, we depict in 3D their objective function spaces using multi-dimensional scaling (MDS) [4, 17]. MDS can project a set of high-dimensional vectors onto a 2D space according to the Euclidean distance between any two of them. In the 3D plot, the control vectors are projected on the x-y plane, and the objective function values (i.e., NPV) are on the z-axis. Because the objective function actually occupies a high-dimensional space, errors in the Euclidean distances and overlaps of vectors can occur in the 3D plot. Nevertheless, the 3D plot provides useful insights into the optimization problem.

Each 3D plot of the objective function space consists of 100 samples of normally distributed control vectors around a central point (either the optimal injection scheme of the Coupled CRMP or of the 2D Model), with a standard deviation of 1/10 of the difference between the upper bound (200 m^3/day) and the lower bound (0 m^3/day). In Figs. 13, 14, 15, and 16, the red dots represent the optimal control vector either of the Coupled CRM or of the 2D Model, the rings represent the random control vectors, and the big blue dots represent the control vector corresponding to the highest NPV among all the samples. Figures 13 and 14 have the same random control vectors and thus the same projection on the x-y plane. The same applies to Figs. 15 and 16. However, the control vectors and projection of Figs. 13 and 14 are different from that of Figs. 15 and 16.

Comparing Figs. 13 to 14 and Figs. 15 to 16 confirms that the objective function spaces of the Coupled CRMP and the 2D Model differ because of approximating the 2D Model with the Coupled CRMP. However, they share some common features. For example, the locations of some of their peaks and valleys are the same. Figure 13 shows that, on the Coupled CRMP, EnOpt leads to a higher NPV than all the sample vectors around it, indicating that it might be a global optimum (at least in the control vector space around the optimum found). However, for the 2D Model, Fig. 14 shows that some control vectors can have a higher NPV than the optimal control vector of the Coupled CRMP. Given that the differences between the highest and lowest values are only 1.2% and 2.3% in Figs. 13 and 14, respectively, the spaces can be considered plateaus with small bumps. Performing

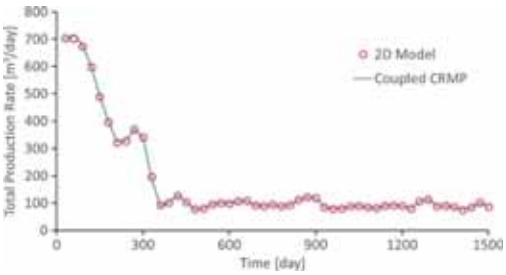


Fig. 11 Total fluid production rate under the optimal injection scheme of the Coupled CRMP, deterministic case, 2D model example

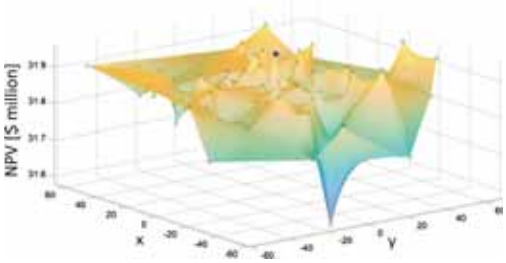


Fig. 13 Objective function space of the Coupled CRMP around the optimal control vector of the Coupled CRMP, deterministic case, 2D model example

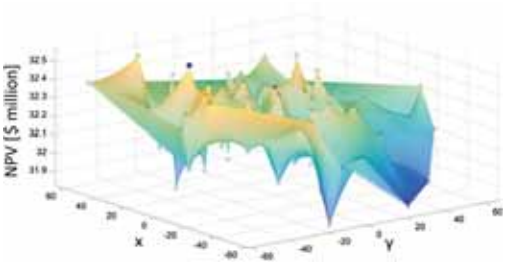


Fig. 14 Objective function space of the 2D Model around the optimal control vector of the Coupled CRMP, deterministic case, 2D model example

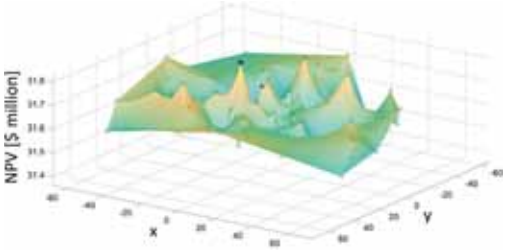


Fig. 16 Objective function space of the Coupled CRMP around the optimal control vector of the 2D Model, deterministic case, 2D model example

EnOpt on the Coupled CRMP brings us to the plateau of the 2D Model's objective function space, but to a bump that is somewhat lower than the highest. In Fig. 15, the optimal control vector of the 2D Model has the highest NPV among all the sample vectors, indicating that it might be a global optimum. The optimal control vector of the 2D Model also does not lead to an optimal NPV on the Coupled CRMP, as shown in Fig. 16.

Figure 17 illustrates the objective function values of the 2D Model vs. that of the Coupled CRMP, where the blue dots correspond to the control vector samples in Figs. 13 and 14, and the red dots do likewise for Figs. 15 and 16. The Coupled CRMP is shown to underestimate NPV with a largest difference of 3.8% for all the samples. This is why, in the proxy-model workflow, we estimate NPV on the 2D Model again after we have obtained the optimal control of the Coupled CRMP.

The stochastic case In order to consider the uncertainty in permeability, we generate 100 realizations of the permeability field from a multivariate normal distribution. Three of these realizations are shown in Fig. 18. The matching procedure Eq. P2 is repeated for all the realizations, yielding 100 sets of the parameters of the Coupled CRMP (i.e., an ensemble of the Coupled CRMP). The injection schemes for matching and validation in this case are the same as those used for the deterministic case (Figs. 4 and 5). The qualities of matching and prediction are shown in Figs. 19 and 20, respectively. The matching quality is good: the minimal R^2 (the worst matching) for total production rate and water cut are 0.9950 and 0.9929, respectively. The prediction quality is not as good as the matching quality, but still satisfactory with a minimal R^2 of 0.9946 for total production rate and 0.9546 for water cut.

For robust production optimization, we use EnOpt to maximize the expected NPV over all the realizations under a single injection scheme. The starting point uses an injection rate of 100 m^3/day for all the producers and time steps. The optimal expected NPV of the 2D Model is \$32.031 million and of the Coupled CRMP is \$31.124 million. The

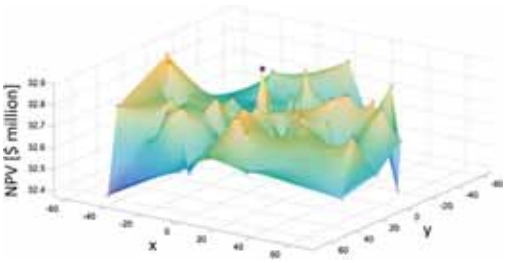


Fig. 15 Objective function space of the 2D Model around the optimal control vector of the 2D Model, deterministic case, 2D model example

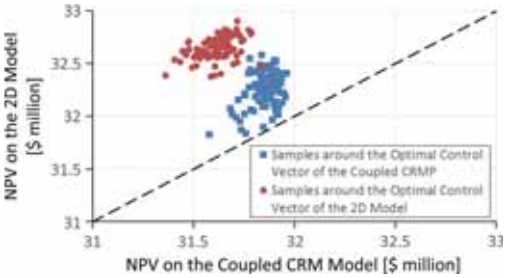


Fig. 17 Objective function values of the 2D Model vs. the Coupled CRMP, deterministic case

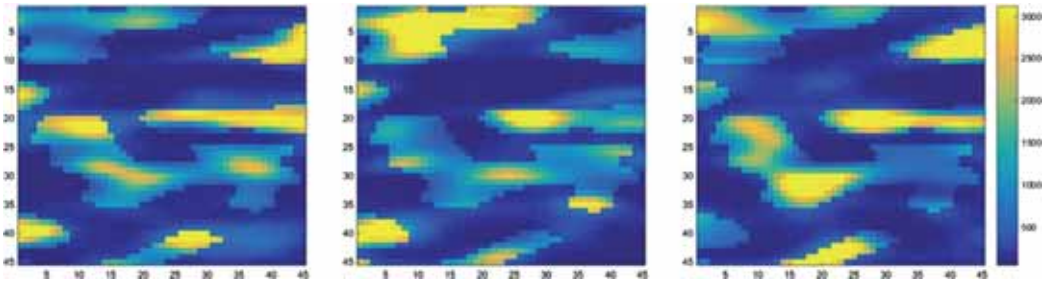


Fig. 18 Three realizations of the permeability field, stochastic case, 2D model example

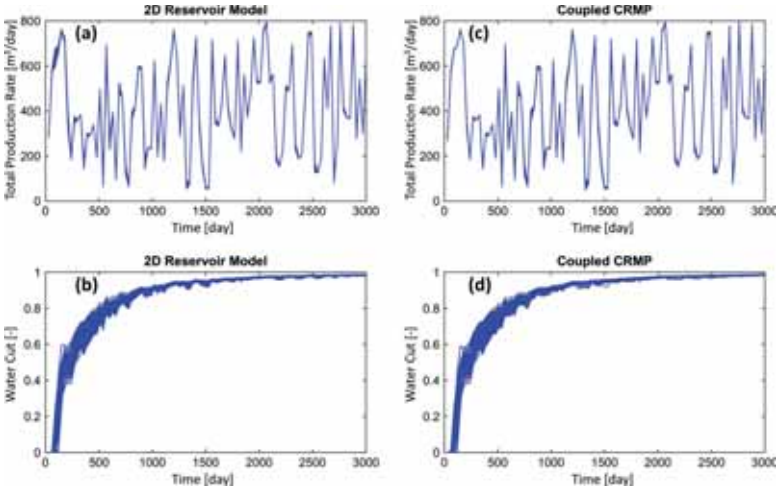


Fig. 19 Matching of total fluid production rate and water cut, stochastic case, 2D model example. **a** total fluid production rate from the 2D Model, **b** water cut from the 2D Model, **c** total fluid production rate from the Coupled CRMP, and **d** water cut from the Coupled CRMP

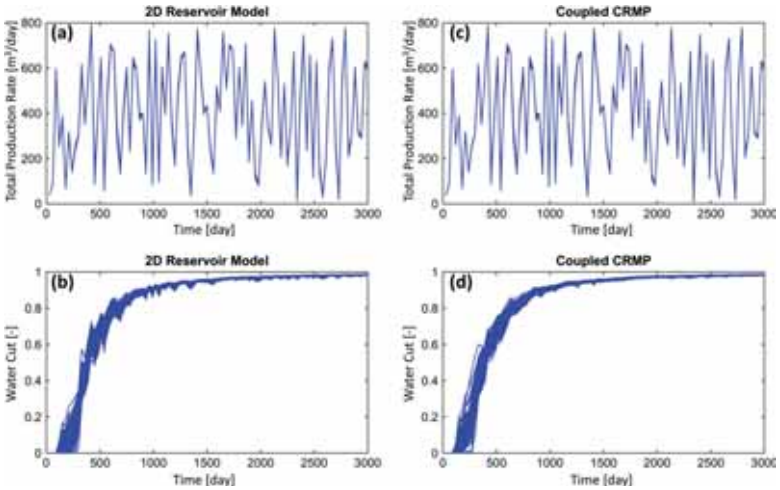


Fig. 20 Validation of total fluid production rate and water cut, stochastic case, 2D model example. **a** total fluid production rate from the 2D Model, **b** water cut from the 2D Model, **c** total fluid production rate from the Coupled CRMP, and **d** water cut from the Coupled CRMP

Table 3 Expected NPVs [$ million], stochastic case, 2D model example

Model	Base case	Optimal	Improvement
Coupled CRMP	22.961	31.124	35.55%
2D Model	23.624	31.837[a,b]	34.76%
2D Model	23.624	32.031[c]	35.59%

[a]With optimal injection scheme of the Coupled CRMP

[b]The solution of the proxy-model workflow

[c]The solution of the traditional workflow

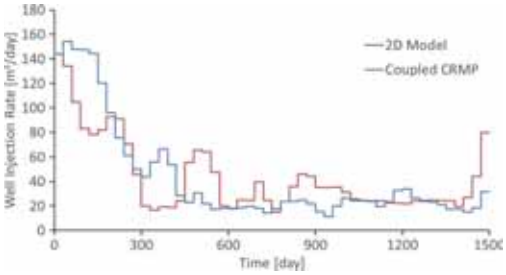


Fig. 22 Optimal injection schemes, stochastic case, 2D model example

expected NPV of the proxy-model workflow is $31.837 million, which is close to that of the traditional workflow. Table 3 lists the expected NPVs. Figure 21 illustrates the cumulative distribution functions (CDF) of NPV corresponding to the optimal solutions of the 2D Model and of the Coupled CRMP, and to the 2D Model under the optimal injection scheme of the Coupled CRMP. The optimal injection schemes are shown in Fig. 22.

As was found in the deterministic case, the proxy-model and traditional workflows lead to similar optimal expected NPVs, but to different optimal injection schemes and hence different NPV distributions (the solid red curve vs. the solid blue curve in Fig. 21). Again, the Coupled CRMP underestimates NPV (the dashed red curve vs. the solid red curve in Fig. 21). The shapes of the cumulative distribution functions (CDF) of NPV are very similar.

The total number of production prediction runs during EnOpt was 21,800 for the Coupled CRMP and 15,600 for the 2D Model. The run times of the Coupled CRMP and the 2D Model for a prediction of 3,000 days were 0.17 and 4.1 s, respectively, and the computational time of one history matching was 1.1 s. The 1,500-day run time of a life-cycle prediction is one half of the run time of a 3,000-day prediction. The total computational time of the traditional workflow is

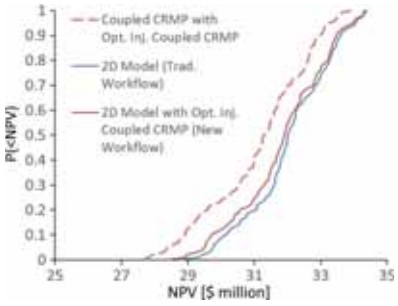


Fig. 21 CDF of NPV, stochastic case, 2D model example

$$\text{Time for optimization} = 15,600 \times \frac{4.1}{2} = 31,980 \text{ [s]},$$

and of the proxy-model workflow is

$$\begin{aligned}&(\text{Time for matching}) + (\text{time for optimization}) \\ &\qquad + (\text{time for expected NPV calculation}) \\ &= 100 \times (4.1 + 1.1) + 21{,}800 \times \frac{0.17}{2} \\ &\qquad + 100 \times \frac{4.1}{2} \\ &= 2{,}578 \text{ [s]}.\end{aligned}$$

The computation time required by the proxy-model workflow is therefore less than 1/10 that of the traditional workflow. The distinction will become more pronounced if the grid-based reservoir model has a finer grid block size or higher geological complexity.

5.4 Impact on decision-making

This section explores the impact of any difference in optimal solutions obtained by the proxy-model and traditional workflows. We consider a simple decision-making context that involves three decisions in series. The first decision has two options: to continue developing the field at a cost of $10 million or to walk away. The second decision is whether to install valves on the injectors at a cost of $2 million, given that water would be injected at a constant rate of 100 m^3/day over the entire production life-cycle without valves or at any rate from 0 to 200 m^3/day with valves. The third decision is to choose a water injection scheme if valves are installed. All the decisions will be based on our current knowledge about the field. We assume a risk-neutral case where the objective is to optimize the EV.

Figures 23 and 24 illustrate the decision trees for the solutions of the proxy-model and traditional workflows, respectively, for the deterministic case. In decision trees, a square, a circle, and a triangle represent a decision node, an uncertainty node, and a payoff node, respectively. (See [6] for a detailed description of using decision trees to solve hydrocarbon production related decision-making problems.) The

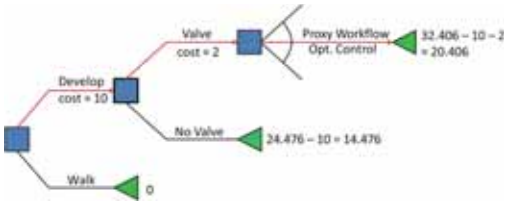


Fig. 23 Decision tree with the optimal solution of the proxy-model workflow, deterministic case, 2D model example. All values are in $ million

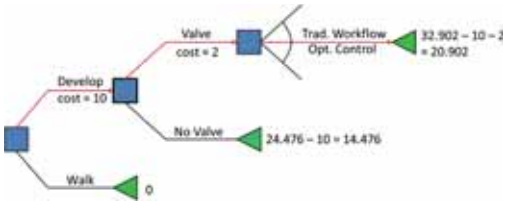


Fig. 24 Decision tree with the optimal solution of the traditional workflow, deterministic case, 2D model example. All values are in $ million

decision trees in Figs. 23 and 24 select the same decisions for the first and second decision nodes. However, the two workflows make the third decision diverge, resulting in a revenue of $20.406 million for the proxy-model workflow (applying the blue line in Fig. 10) and a revenue of $20.902 million for the traditional workflow (applying the red line in Fig. 10).

One reason of this difference is the stochastic feature of EnOpt and the other is the verisimilitude or the richness[5] of the models. If we adjust for the stochastic feature of EnOpt, the only difference between the proxy-model and traditional workflows is in the verisimilitude of the predictive model used for production optimization. The proxy-model workflow uses a model that is less rich but faster than that used in the traditional workflow. To quantify the benefit that the verisimilitude might offer, we introduce the value of verisimilitude (VOV), which is defined as

$$VOV = v(\mathbf{u}_{rich}) - v(\mathbf{u}_{proxy}), \tag{25}$$

where $v(\mathbf{u})$ is the true value of the decision given a control $\mathbf{u}$, $\mathbf{u}_{rich}$ is the control vector found by using the traditional workflow, and $\mathbf{u}_{proxy}$ is that found by the proxy-model workflow. However, being that a given field can support only one production strategy and can not yield the true value until the end of production, the rich model is assumed to represent the truth. The VOV of the deterministic case shown in Figs. 23 and 24 is $20.902 million – $20.406 million = $0.496 million. Similar to the statement for value of information (VOI) that "one cannot value information outside of a particular decision context" [8], using a more verisimilar production prediction model has no value in itself, unless it changes our decision. Thus, the 2D Model (a richer model) adds no value to the first and second decisions in this case. It adds value only to the third decision. The VOV is positive because the 2D Model can capture more minor flow behaviors than can the Coupled CRMP, giving a more accurate production prediction under our assumption that the 2D Model represents the truth. However, the VOV ($0.496 million), compared to the NPV improvement ($20.406 million − $14.476 million = $5.930 million) by installing the valves and conducting the optimal injection scheme found by performing EnOpt on the Coupled CRMP, is relatively small. This is because the Coupled CRMP is already a good approximation to the 2D Model.

For the stochastic case, the decision trees are illustrated in Figs. 25 and 26. The P10, P50, and P90 values are shown in the decision trees. The distributions of the NPV are expected to differ because the proxy-model and traditional workflows lead to different optimal injection schemes. Given our objective of maximizing the expected NPV of a particular decision-making problem, the solution of the proxy-model workflow gives an expected NPV very close to that of the traditional workflow. This observation is consistent with that in the deterministic case. To calculate the VOV, we assume that each realization of the 2D Model is equiprobable to be the truth. Hence, the VOV for one particular realization is calculated as

$$VOV_r = v_r(\mathbf{u}_{rich}) - v_r(\mathbf{u}_{proxy}), \tag{26}$$

where $v_r(\mathbf{u})$ is the value of the decision given a control vector $\mathbf{u}$ if realization r of the rich model is the truth. The expected VOV (EVOV) is the EV of VOV_r over all the geological realizations. The EVOV is $20.031 million − $19.837 million = $0.194 million for the case shown in Figs. 25 and 26. The corresponding CDF of the VOV is illustrated in Fig. 27 and shows that there is around 5.4% chance of a negative VOV. The minimal VOV is −$0.096 million, and the maximal is $0.783 million.

We investigate the sensitivity of EVOV to the valve cost by varying the latter. The results are plotted in Fig. 28 which can be divided into three regions: valve cost lower than $8.21 million (Region 1), between $8.21 million and $8.41 million (Region 2) and greater than $8.41 million (Region 3). In Region 1, the EVOV is a constant $0.194 million because only the third decision is affected by whether the proxy-model or traditional workflow is chose. In Region 2, installing valves is called for by the traditional workflow

[5]This paper uses "richness" interchangeably with "complexity" and "verisimilitude" to indicate a high level of detail built into a model.

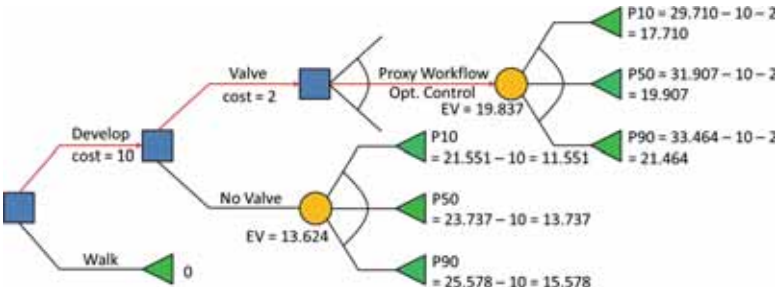


Fig. 25 Decision tree with the optimal solution of the proxy-model workflow, stochastic case, 2D model example

but not by the proxy-model workflow. In Region 3, EVOV drops to 0 because neither the proxy-model nor the traditional workflow calls for installing valves. This illustrates that the value of using a richer production prediction model is decision-dependent.

6 Example with a 3D reservoir model

In this section, we apply the proxy-model workflow to a synthetic 3D reservoir model ("3D Model"). The model was introduced in [33] and was used in [17]. A channelized depositional system is modeled. The possible patterns of highly permeable channels are described by 100 equi-probable geological realizations, three of which are illustrated in Fig. 29. The model has eight water injectors and four producers (Fig. 30). All the wells are vertical and completed in all seven layers. Capillary pressure is ignored. The reservoir rock is assumed to be incompressible.

The life cycle of the reservoir is 3,600 days. It is divided into 40 time intervals of 90 days for water injection rate optimization. Thus, there are $8 \times 40 = 320$ control variables. The water injection rate of an injector can be adjusted from 0 to 60 m^3/day. The base case is production under the maximal allowable injection rate of 60 m^3/day in all the injectors for the entire life cycle. The injectors are operated with no pressure constraint, and the producers are under a minimal BHP of 395 bars without rate constraint. The oil price, water production cost, and water injection cost are set to 126 \$/m^3, 19 \$/m^3, and 5 \$/m^3, respectively. The discount rate is set to 0.

The Coupled CRMP is used to approximate the 3D Model. The resulting NPVs are listed in Table 4. The optimal NPV obtained using the proxy-model workflow is \$43.359 million which is only 2.18% lower than that of the traditional workflow (\$44.326 million). The NPV improvement attributable to the use of the 3D Model (\$0.967 million) is much smaller than that attributable to production optimization (\$14.823 million). The computational time of the 3D Model is 67 seconds, whereas that of the Coupled CRMP is only 0.46 s. The total computational time for production optimization is reduced by a factor of ten: from 408,700 s using the traditional workflow, to 40,582 s using the proxy-model workflow.

For the decision problem illustrated in Figs. 25 and 26 (where the values for the 2D example should be replaced by the values for the 3D example) again as an example, using the 3D Model adds no value to the decision of whether to develop the field, because the optimization results of the proxy-model and traditional workflows both call for development. For a total valve cost of \$3.2 million, using the 3D Model also adds no value to the decision

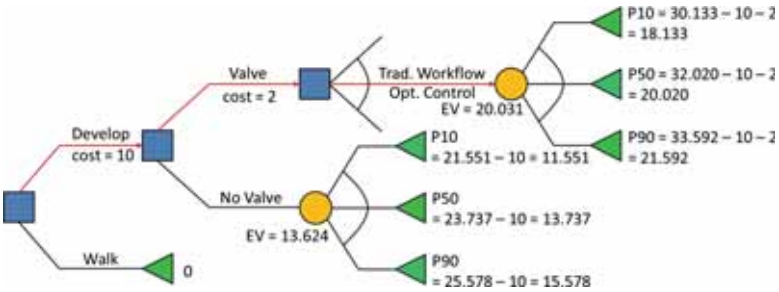


Fig. 26 Decision tree with the optimal solution of the traditional workflow, stochastic case, 2D model example

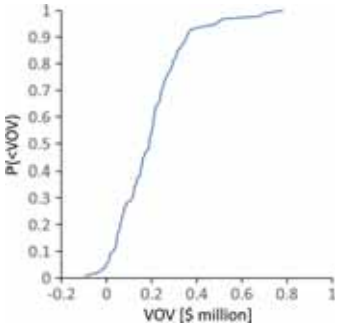


Fig. 27 CDF of the VOV, stochastic case, 2D model example

on valve installation. The 3D Model adds value only to the decision on injection strategy, with an EVOV of $0.967 million.

7 Discussion

7.1 Proxy models

Whether using rich models or proxy models, one should keep in mind the counsel of George E.P. Box [5]: *All models are wrong but some are useful*.

Engineering uses models to support decision-making. Good decision models are both useful and tractable. By useful we mean that the model must be relevant and generate insight to resolve the decision at hand. Usefulness also requires the model to be credible and transparent—will the decision-makers believe the result of the analysis and can the approach be clearly explained and understood? By tractable we mean that the required analysis can be done within the time and resources available.

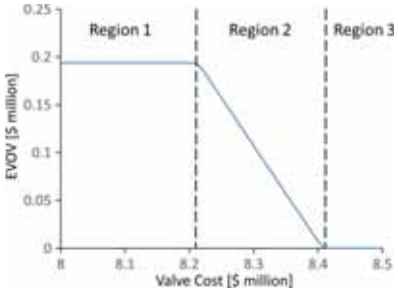


Fig. 28 Sensitivity of EVOV to valve cost, stochastic case, 2D model example

We might argue that both grid-based models and proxy models (e.g., the CRM-based models) satisfy the usefulness requirement although a reservoir simulator such as ECLIPSE, with its voluminous code, is so feature rich and detail oriented that it is not transparent to most users. Furthermore, as we have argued earlier, a rich grid-based model is often not tractable—particularly when the underlying problem is uncertain.

In decision-making contexts, we need cogent (compelling) models. Companies tend to build too much detail into their decision-making models from the start and focus too much energy on specific cases or inputs that do not influence the decision. This level of modeling detail is really a shirking of responsibility on the part of the decision analyst who will and can build a model that includes only the most salient factors. Building in detail may be easy, but building in incisiveness is hard work.

The Coupled CRMP is cogent for the optimization and decision problems discussed in this work. However, other models such as the INSIM [37], streamline models [2], reduced-order models [3], or upscaled grid-based reservoir models [32] can be cogent in this or similar contexts and can serve as proxy models. The purpose of introducing the proxy-model workflow is to make the optimization process tractable by reducing the computational requirements while ensuring optimal or near-optimal decisions. Thus, the desiderata of proxy models is that they are cogent–a proxy model must be useful, tractable, and lead to incisiveness. This, in turn, requires the models to capture the most relevant physics and mechanisms affecting production prediction in reality and be very computationally attractive. As shown in the plots of the objective function spaces (Figs. 13, 14, 15, and 16), even a very good approximation of a proxy model to a rich model can lead to a different objective function space resulting in a different optimal result. The choice and usefulness of a proxy model for this purpose is a function of the reservoir characteristics and drainage scheme. Before screening proxy models, we must first identify main affecting factors. For example, in this paper, we consider a heterogeneous reservoir with water flooding, so a useful proxy model must have the function to mimic the heterogeneity and predict production with various water injection schemes.

A disadvantage of using proxy models is that many simple and fast models are case-specific and hence less flexible than a grid-based reservoir model. For example, to check the impact of adding a new well would entail re-determining all the parameters of a CRM. Another problem is that the possible values of the parameters of a proxy model are usually too abstract to be readily determined. For example, the connectivities and time constants of a CRM depend on the permeability field, porosity field, well pattern, etc. If production data are available, these parameters can be

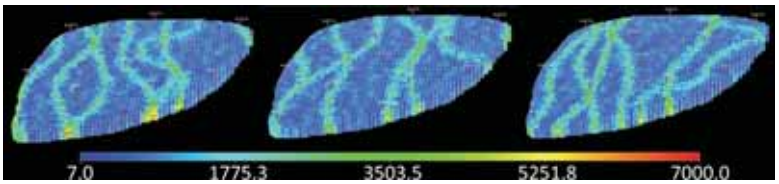


Fig. 29 Three realizations of the permeability field of the 3D model. All values are in md

determined by history matching; if not, there is seldom a reference for them. However, fundamental physical parameters such as permeability in each grid block are much more intuitive to engineers than are abstract ones like connectivities. That is why, in the proxy-model workflow, we propose to first generate pseudo production history by running grid-based reservoir simulations and then determine the parameters of the proxy model by history matching.

7.2 CRM-based models

We have screened two CRM-based models. The Coupled CRMP couples the saturation equation so that it can capture the time-variant behavior of time constants and has a better ability in water cut prediction than does the CRM-Koval Model [10]. Both models are producer based, as illustrated in Fig. 1b. An inherent flaw of a CRMP is that the region in the control volume is treated as homogeneous, which can lead to incorrect water cut prediction when well injection rates change because there is no difference between water injected into a more permeable region and into a less permeable region in the control volume of CRMP. The CRMIP can better capture heterogeneity.

Although the Coupled CRMP includes the time-variant behavior of time constants, it does not consider the time-variant behavior of connectivities. Gentil [18] formulated connectivity as a function of transmissibilities, but we found no literature that introduced an equation to calculate the time-variant connectivities.

Despite the limitations of the Coupled CRMP, it has shown to provide a near-optimal solution. Thus, further improvements in the Coupled CRMP might not create much value for cases similar to the examples shown.

7.3 Performance of the Coupled CRMP in the proxy-model workflow

The performance of the Coupled CRMP in the proxy-model workflow depends on two factors: computation time and optimization result.

The computation for the proxy-model workflow takes about a factor of ten less time than that for the traditional workflow. In a more realistic case involving a much richer grid-based reservoir model, the proxy-model workflow will save even more computation time but could weaken the ability of the Coupled CRMP to approximate the grid-based reservoir model. This tradeoff can be evaluated by analyzing the VOV. A significant computation time reduction with small VOV would suggest using the proxy model.

Because of the difference between the objective function spaces of the 2D Model and the Coupled CRMP (Figs. 13, 14, 15, and 16), the optimal injection schemes of the proxy-model and traditional workflows are different. Nevertheless, the proxy-model workflow provides an optimal expected NPV very close to that of the traditional workflow, with differences of only 0.61% for the 2D Model and 2.18% for the 3D Model. This might be because the locations of the major plateaus and basins of the objective function spaces of the grid-based model and the Coupled CRMP are almost the same, with only the locations of small bumps being different.

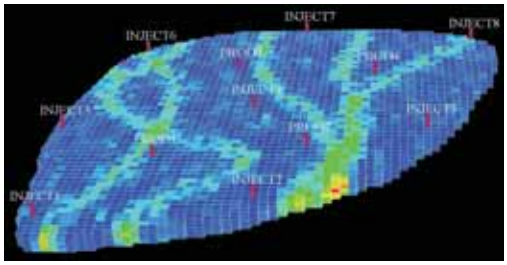


Fig. 30 Well locations of the 3D Model

Table 4 Expected NPVs [$ million], 3D model example

Model	Base case	Optimal	Improvement
Coupled CRMP	29.658	43.255	45.85%
3D model	28.536	43.359[a,b]	51.94%
3D model	28.536	44.326[c]	55.33%

[a]With optimal injection scheme of the Coupled CRMP
[b]The solution of the proxy-model workflow
[c]The solution of the traditional workflow

7.4 The VOV/EVOV

The VOV/EVOV is a measure of the value of using a richer production prediction model for production optimization in a decision-making context. In our examples, because the Coupled CRMP is a good proxy to the grid-based model and because their optimal objective function values (NPV) are very close, the VOV/EVOV is low or zero, indicating that using a grid-based model instead of the Coupled CRMP might not create any value. Therefore, the value of using a richer production prediction model is decision-dependent.

Although EVOV is positive in our examples, VOV for a particular realization can be negative. This indicates that using a richer production prediction model for production optimization cannot guarantee a better control than using a proxy model.

The VOV/EVOV can be compared with the computational cost (converted to monetary terms) saved by using the proxy-model workflow instead of the traditional one: if the computational cost saved exceeds the VOV/EVOV, then the proxy model is preferable.

Bratvold and Begg [7] noted that "companies tend to build too much detail into their decision-making models from the start." This pitfall might be avoided by being aware of and calculating VOV/EVOV. Unfortunately, VOV, as we have defined it, can be estimated only by running both the rich model and the proxy model. Ideally, VOV should be obtained without running the rich model. One possible solution is to build a database of VOV assessments, which can eventually be used to estimate the value of using a richer production prediction model for similar cases.

8 Conclusions

We have presented a proxy-model workflow for reducing the computation time of robust production optimization. The proxy-model workflow uses a proxy model to approximate a rich model. We screened two CRM-based models and investigated the goodness of using the Coupled CRMP to serve as a proxy model, using synthetic 2D and 3D reservoir models as examples. The results showed that the proxy-model workflow can provide an objective function value that is very close to the optimal value found by using the traditional workflow, but a different optimal control vector. The computation time decreased by a factor of more than ten.

We used the MDS approach to visualize the objective function spaces of the 2D Model and the Coupled CRMP. This gives insight on how a proxy model affects optimal results. It was shown that even a very good proxy model can lead to obvious deviations from the objective function space of a rich production prediction model.

VOV/EVOV was used to quantify the value of using a richer model for production prediction, which was shown to be decision-dependent. Using a richer production prediction model does not always provide greater value than does using a proxy model.

We conclude that CRM-based models have high potential to serve as a cogent proxy model for waterflooding related decision-making context and that the proxy-model workflow, leveraging a faster, but relevant, production model, significantly speeds up the optimization yet gives robust results that leads to a near-optimal solution.

Nomenclature

Greek letters

α	Step length for control vector updating
Δt_k	Time step length between times k and $k-1$
λ_{ij}	Connectivity between injector i and producer j
μ_o	Oil viscosity
μ_w	Water viscosity
τ_j	Time-invariant time constant for producer j
τ_j^k	Time-variant time constant for producer j at time k

Roman letters

b	Discount factor
c_f	Pore compressibility
c_o	Oil compressibility
c_t	Total compressibility
$\mathbf{C_u}$	Predefined covariance matrix of $\mathbf{u}$
$\mathbf{C_{uJ}}$	Cross-covariance matrix between $\mathbf{u}$ and J
d_k	Production data at time k
$\overline{d}$	Mean of d_k over all k
D	Reference time for discounting
$f_{w}{}_j^k$	Water cut in producer j at time k
$\mathbf{g}$	Gradient for control vector updating
I_i^k	Water injection rate in injector i during Δt_k
$I_t{}_j^k$	Total water injection contribution from all injectors to producer j over Δt_k
$\mathbf{j}$	Mean shifted objective function vector
J	Objective function of production optimization
J_r	J for realization r
$\overline{J}$	EV of J
$J_t{}_j^k$	Total productivity index of producer j at time k
$J'_t{}_j$	Constant in Equation 12 for producer j
k_{ro}	Oil relative permeability
k_{rw}	Water relative permeability
$K_{val\,j}$	Koval factor of producer j
M	Number of perturbed control vectors (i.e., ensemble size of control vector)
M_j^k	Total relative mobility for producer j at time k
n_e	Number of geological realizations (i.e., ensemble size of geological realization)

n_{inj}	Number of injectors
n_p	Number of producers
n_T	Number of time steps
N	Number of control variables
P_o	Oil price
P_{wi}	Water injection cost
P_{wp}	Water production cost
$q_{o_j}^k$	Oil production rate in producer j at time k
$q_{t_j}^k$	Total fluid production rate in producer j at time k
$q_{w_j}^k$	Water production rate in producer j at time k
R^2	Coefficient of determination
S_{or}	Residual oil saturation
${S_o}_{2_j}^k$	Outlet oil saturation of producer j at time k
S_{wi}	Irreducible water saturation
$\overline{S}_{o_j}^k$	Average oil saturation within the control volume of producer j at time k
t_{Dj}	Dimensionless time of producer j
t_k	Cumulative time for discounting at time k
T_j	Constant in Equation 11 for producer j
u	Control variable
$\mathbf{u}$	Control vector or vector of control variables
$\mathbf{u}_r$	$\mathbf{u}$ for realization r
$\widehat{\mathbf{u}}$	Predefined mean of $\mathbf{u}$
$\overline{\mathbf{u}}$	Sample mean of $\mathbf{u}$
$\overline{\mathbf{U}}$	Mean shifted ensemble matri
v	Value function
V_{p_j}	Drainage pore volume of producer j
$V_{p_{Field}}$	Total pore volume of the field

Superscripts and subscripts

i	Index of injector
j	Index of producer
k	Index of time
l	Index of optimization iteration
o	Oil
r	Index of realization or ensemble member
w	Water

Acknowledgements The authors acknowledge the Research Council of Norway and the industry partners; ConocoPhillips Skandinavia AS, Aker BP ASA, Eni Norge AS, Maersk Oil Norway AS, DONG Energy A/S, Denmark, Statoil Petroleum AS, ENGIE E&P NORGE AS, Lundin Norway AS, Halliburton AS, Schlumberger Norge AS, Wintershall Norge AS of The National IOR Centre of Norway for support.

References

1. Albertoni, A., Lake, L.W.: Inferring interwell connectivity only from well-rate fluctuations in waterfloods. SPE Reserv. Eval. Eng. **6**(1), 6–16 (2003)
2. Al-Najem, A.A., Siddiqui, S., Soliman, M., Yuen, B.: Streamline Simulation Technology: Evolution and Recent Trends. Paper SPE-160894-MS Presented at the SPE Saudi Arabia Section Technical Symposium and Exhibition, Al-Khobar, Saudi Arabia (2012)
3. Alghareeb, Z.M., Williams, J.: Optimum Decision-Making in Reservoir Management Using Reduced-Order Models. Paper SPE-166305-MS Presented at the SPE Annual Technical Conference and Exhibition, New Orleans, Louisiana, USA (2013)
4. Borg, I., Groenen, P.J.: Modern Multidimensional Scaling: Theory and Applications. Springer Science & Business Media, New York, USA (2005)
5. Box, G.E.P.: Robustness in the strategy of scientific model building. Robustness in Statistics, 201–236 (1979)
6. Bratvold, R.B., Begg, S.: Making Good Decisions. Society of Petroleum Engineers, Texas, USA (2010)
7. Bratvold, R.B., Begg, S.: Would you know a good decision if you saw one? The Way Ahead **5**(2), 21–23 (2009)
8. Bratvold, R.B., Bickel, J.E., Lohne, H.P.: Value of information in the oil and gas industry: past, present, and future. SPE Reserv. Eval. Eng. **12**(4), 630–638 (2009)
9. Bruce, W.A.: An electrical device for analyzing oil-reservoir behavior. Pet. Technol. **151**(1), 112–124 (1943)
10. Cao, F.: Development of a Two-phase Flow Coupled Capacitance Resistance Model. PhD Dissertation, The University of Texas at Austin, USA (2014)
11. Cao, F., Luo, H., Lake, L.W.: Development of a Fully Coupled Two-Phase Flow Based Capacitance Resistance Model (CRM). Paper SPE-169485-MS Presented at the SPE Improved Oil Recovery Symposium, Tulsa, Oklahoma, USA (2014)
12. Cao, F., Luo, H., Lake, L.W.: Oil-rate forecast by inferring fractional-flow models from field data with Koval method combined with the capacitance/resistance model. SPE Reserv. Eval. Eng. **18**(4), 534–553 (2015)
13. Chen, Y., Oliver, D.S., Zhang, D.: Efficient ensemble-based closed-loop production optimization. SPE J. **14**(4), 634–645 (2009)
14. Denney, D.: Pros and cons of applying a proxy model as a substitute for full reservoir simulations. J. Petrol. Tech. **62**(7), 634–645 (2010)
15. Do, S.T., Reynolds, A.C.: Theoretical connections between optimization algorithms based on an approximate gradient. Comput. Geosci. **1717**(6), 959–973 (2013)
16. ECLIPSE: http://www.software.slb.com/products/eclipse. Accessed 12 October 2016 (2014)
17. Fonseca, R., Stordal, A., Leeuwenburgh, O., Van den Hof, P.M.J., Jansen, J.D.: Robust Ensemble-Based Multi-Objective Optimization. ECMOR XIV-14Th European Conference on the Mathematics of Oil Recovery, Catania, Sicily, Italy (2014)
18. Gentil, P.H.: The Use of Multilinear Regression Models in Patterned Waterfloods: Physical Meaning of the Regression Coefficients. Master's thesis, The University of Texas at Austin, USA (2005)
19. Holanda, R.W., Gildin, E., Jensen, J.L.: Improved Waterflood Analysis Using the Capacitance-Resistance Model within a Control Systems Framework. Paper SPE-177106-MS Presented at SPE Latin American and Caribbean Petroleum Engineering Conference, Quito, Ecuador (2015)
20. Izgec, O., Kabir, C.S.: Quantifying nonuniform aquifer strength at individual wells. SPE Reserv. Eval. Eng. **13**(2), 296–305 (2010)
21. Jafroodi, N., Zhang, D.: New method for reservoir characterization and optimization using CRM-EnOpt approach. J. Pet. Sci. Eng. **77**(2), 155–171 (2011)
22. Koval, E.: A method for predicting the performance of unstable miscible displacement in heterogeneous media. Soc. Pet. Eng. J. **3**(2), 145–154 (1963)
23. Lake, L.W.: Enhanced Oil Recovery. Prentice Hall, Englewood Cliffs, New Jersey, USA (1989)

24. Liang, X., Weber, D., Edgar, T.F., Lake, L.W., Sayarpour, M., Yousef, A.A.: Optimization of Oil Production based on a Capacitance Model of Production and Injection Rates. Paper SPE-107713-MS presented at the Hydrocarbon Economics and Evaluation Symposium, Dallas, Texas, USA (2007)
25. MATLAB: http://se.mathworks.com/products/matlab. Accessed 12 October 2016 (2014)
26. Nocedal, J., Wright, S.: Numerical Optimization, pp. 56–57. Springer Science & Business Media, New York, USA (2006)
27. Sayarpour, M.: Development and Application of Capacitance-Resistive Models to Water/CO2 Floods. PhD Dissertation, The University of Texas at Austin, USA (2008)
28. Sayarpour, M., Kabir, C., Sepehrnoori, K., Lake, L.W.: Probabilistic history matching with the capacitance-resistance model in waterfloods: a precursor to numerical modeling. J. Pet. Sci. Eng. **78**(1), 96–108 (2011)
29. Stordal, A.S., Szklarz, S., Leeuwenburgh, O.: A theoretical look at ensemble-based optimization in reservoir management. Math. Geosci. **48**(4), 399–417 (2015)
30. Sayarpour, M., Zuluaga, E., Kabir, C.S., Lake, L.W.: The use of capacitance-resistance models for rapid estimation of waterflood performance and optimization. J. Pet. Sci. Eng. **69**(3–4), 227–238 (2009)
31. Tavassoli, Z., Carter, J.N., King, P.R.: Errors in history matching. SPE J. **9**(3), 352–361 (2004)
32. Vakili-Ghahani, S.A., Jansen, J.D.: A system-theoretical approach to selective grid coarsening of reservoir models. Comput. Geosci. **16**(1), 159–176 (2012)
33. van Essen, G.M., Zandvliet, M.J., Van den Hof, P.M.J., Bosgra, O.H., Jansen, J.D.: Robust waterflooding optimization of multiple geological scenarios. SPE J. **14**(1), 202–210 (2009)
34. Welge, H.J.: A simplified method for computing oil recovery by gas or water drive. J. Petrol. Tech. **4**(4), 91–98 (1952)
35. Yang, C., Card, C., Nghiem, L.X., Fedutenko, E.: Robust Optimization of SAGD Operations under Geological Uncertainties. Paper SPE-141676-MS Presented at the SPE Reservoir Simulation Symposium, The Woodlands, Texas, USA (2011)
36. Yousef, A.A., Gentil, P.H., Jensen, J.L., Lake, L.W.: A capacitance model to infer interwell connectivity from production and injection rate fluctuations. SPE Reserv. Eval. Eng. **9**(6), 630–646 (2006)
37. Zhao, H., Kang, Z., Zhang, X., Sun, H., Cao, L., Reynolds, A.C.: INSIM: a Data-Driven Model for History Matching and Prediction for Waterflooding Monitoring and Management with a Field Application. Paper SPE-173213-MS Presented at the SPE Reservoir Simulation Symposium, Houston, Texas, USA (2015)

Paper III

Fast Analysis of Optimal IOR Start Time Using a Two-Factor Production Model and Least-Squares Monte Carlo Algorithm

Hong, A.J., Bratvold, R.B., and Lake, L.W.

Manuscript

Fast Analysis of Optimal IOR Start Time Using a Two-Factor Production Model and Least Squares Monte Carlo Algorithm

Aojie Hong and **Reidar B. Bratvold**, University of Stavanger and The National IOR Centre of Norway; and **Larry W. Lake**, University of Texas at Austin

Abstract

Although numerical techniques (e.g., reservoir simulation) have developed rapidly for modeling oil production in the past decades, simple production models with a few parameters are still of significant use in the oil and gas industry because of they are computationally attractive. Especially for decision analysis, which can be computationally demanding, a useful and tractable model is essential.

A particular decision for development design is: when is optimal to start an improved-oil-recovery (IOR) process such as waterflooding or gasflooding? This decision relates to the plan of manufacturing, transporting and installing the equipment, allocating both financial and human resources and licensing for production. Solving this problem using a state-of-the-art reservoir management approach—closed loop reservoir management (CLRM)—may result in a sub-optimal solution because the CLRM approach considers only the uncertainties and decisions associated with currently available data but not the uncertainties and decisions associated with future data.

This paper illustrates a method for performing a fast analysis of the optimal IOR start time using a two-factor production model and least-squares Monte Carlo (LSM) algorithm. The two-factor production model is similar to an exponential decline and contains only two parameters for each recovery phases. Thus, it is very computational attractive. The LSM algorithm is an approximate dynamic programming approach, which considers both the impact of the information obtained before a decision is made and the impact of the information that will be obtained on the decisions that will be made in the future. It therefore provides a near-optimal solution for the IOR start time problem.

The main contributions of this work are (1) a simple example illustrating the full structure of the IOR start time problem visualized by a decision tree, (2) demonstration and discussion of the sub-optimality of the CLRM solution, (3) illustration of the detailed steps of applying the LSM algorithm, and (4) an example of the implementation of the two-factor model combined with the LSM algorithm for analyzing the optimal IOR start time.

The LSM algorithm can significantly improve the decisions and lead to a significant increase in a field's economic performance. Using the two-factor model combined with the LSM algorithm can provide useful insight in the problem of deciding the optimal IOR start time with limited computational resource.

Introduction

Although numerical techniques (e.g., reservoir simulation) have developed rapidly for modeling oil production in the past decades, simple production models with a few parameters are still of significant use in the oil and gas industry for the following possible reasons. In the development phase and early stage of production, there is little available information that can provide enough details for building a complex reservoir simulation model with thousands of parameters. Besides, when uncertainties are considered, Monte Carlo simulation is a common practice where a large number of production forecast runs are required. Thus, computationally intensive reservoir simulation models can easily make the use of Monte Carlo simulation computationally prohibitive.

Decision analysis in a reservoir management context may suffer from large computational costs because of many uncertain outcomes and sequential decisions points. Thus, a useful and tractable production model is essential for decision analysis. By "useful", we mean that the model must be relevant and generate insight to resolve the decisions at hand. By "tractable", we mean that the required analysis can be done within the time and resources available. We argue that companies tend to build too many details into their decision-making models from the start and focus too much effort on specific cases or inputs that may not influence the decisions. This means that decision analysts can only include the most salient factors in their models (Bickel and Bratvold2008; Hong et al. 2017). As

Howard has counseled, "the real problem in decision analysis is not making analyses complicated enough to be comprehensive, but rather keeping them simple enough to be affordable and useful."

A particular decision for development design is: when is the best (the optimal) time to start the improved-oil-recovery (IOR) process such as waterflooding or gasflooding? The IOR start time is related to the plan of manufacturing, transporting and installing the equipment, allocating both financial and human resources and licensing for production.

A useful and tractable production model for this problem setting is the two-factor production model proposed by Parra-Sanchez (2010). This model is based on the exponential decline model. For each individual recovery phase, it uses only two parameters to describe the reservoir properties and the effect of the recovery mechanism. A more detailed description of this model follows later. Parra-Sanchaz validated it using field production data and further used this model for life cycle optimization where an integrated depletion strategy of the life times of each recovery phase is optimized under the consideration of multiple recovery phases as a whole. She concluded that, compared to the traditional optimization approach where the life time of each recovery phase is optimized only under the consideration of this individual recovery phase, the life cycle optimization approach significantly improves the field's economic performance in terms of net present value (NPV) and severely reduces the total life time of production.

A big concern that has not been included in Parra-Sanchaz's work is the impact of future information on sequential decision making (i.e., the impact of learning over time). The reservoir properties and the effect of recovery mechanism are uncertain because of our lack of knowledge and their uncertainties are represented by uncertain model parameters. After the production has started, we will obtain additional production data and the data will be used to update knowledge and support our decisions. For example, at the primary recovery phase, the production data informs us about the reservoir properties and the effect of the primary recovery mechanism, and we can then adjust our decision on the life time of primary recovery based on the information obtained.

The closed loop reservoir management (CLRM) approach (Brouwer et al. 2004; Nævdal et al. 2006; Chen et al. 2009; Wang et al. 2009; Jansen et al. 2009) is a state-of-the-art approach of including the impact of information on sequential decision making in reservoir management. In the CLRM approach, the loop of model updating through history matching and production optimization is performed continuously when additional data becomes available. However, this approach is myopic[1] and may lead to a sub-optimal decision policy because it considers only the impact of the information obtained before a decision is made but not the impact of the information that will be obtained on the decisions that will be made in the future.

Decision tree is a useful decision analysis tool for visualizing and communicating the structure of such sequential decisions. In the following, we use "fully structured reservoir management (FSRM) approach" to refer to the reservoir management approach that considers explicitly the full structure of a sequential decision-making problem. A FSRM problem can be solved by rolling back the corresponding decision tree. For a more detailed description of the use of decision tree see Howard and Abbas (2016), and Bratvold and Begg (2010).

The decision tree approach is essentially dynamic programming. For complex problems with many uncertain outcomes and decision points, the decision tree approach will suffer from the "curse of dimensionality". In addition, it can be quite involved to solve these complex problems (semi-) analytically using dynamic programming. Fortunately, efficient approximate dynamic programming methods are available. One of them is the least-squares Monte Carlo (LSM) algorithm developed by Longstaff and Schwartz (2001) originally for solving a real option problem. It won't suffer from the curse of the number of uncertain quantities. However, the computational effort it requires increases exponentially with the number of alternatives. Thus, it suits problems with a few alternatives. Recently, it has gained more uses in the oil and gas industry (Hem et al. 2011; Willigers et al. 2011; Alkhatib and King 2011; Jafarizadeh and Bratvold 2012, 2013; Thomas and Bratvold 2015).

Following this introduction, we first briefly review the two-factor production model and economic model proposed by Parra-Sanchez (2010). Next, we compare and discuss the CLRM and FSRM approaches for integrating the impact of information in production optimization. Thereafter, we present the central steps of applying the LSM algorithm. We then use a simple example to illustrate the difference between the CLRM and FSRM solutions and to show the detailed steps of applying the LSM algorithm. We conclude with a more realistic example where we perform a fast analysis of the optimal IOR start time using the two-factor model combined with the LSM algorithm. Finally, we present a discussion and conclusions.

[1] By "myopic or naïve approach", we mean that a decision model considers only one uncertainty node representing the uncertainties we have learnt so far and does not consider the uncertainty nodes that representing the future uncertainties we will learn in the future. That is, a myopic approach does not allow for learning over time.

Two-Factor Production Model and Economic Model

This section presents the two-factor production model and economic model proposed by Parra-Sanchez (2010). The two-factor production model can integrate multiple recovery phases (primary, secondary, tertiary recoveries and so on) for achieving a life cycle optimization analysis.

Two-factor Model. The two-factor model is based on the exponential decline proposed by Parra-Sanchez (2010). She formulates the exponential decline for the recovery efficiency to be

$$E_R(t) = E_R^0 + (E_R^\infty - E_R^0)(1 - e^{-\frac{t}{\tau}}) \tag{1}$$

where $E_R(t)$ is the recovery efficiency up to time t, E_R^0 is the recovery efficiency at time 0, E_R^∞ is the theoretical ultimate recovery efficiency and τ is a time constant. The recovery efficiency E_R is defined as the fraction of original oil in place (OOIP) that has been produced. The two factors, E_R^∞ and τ, depend on the reservoir properties (e.g., permeability) and production mechanisms, and represent the essence of reservoir engineering. E_R^∞ describes how much recovery efficiency that a recovery mechanism can theoretically and ultimately achieve. τ describes how fast the recovery efficiency increases for a recovery mechanism. They are assumed to be constant for a certain recovery phase, although they can be time-variant.

Thus, for primary recovery, we have

$$E_{R1}(t) = E_{R1}^0 + (E_{R1}^\infty - E_{R1}^0)(1 - e^{-\frac{t}{\tau_1}}) \tag{2}$$

where the subscript 1 denotes the first recovery stage (primary recovery) and the primary recovery efficiency at time 0, E_{R1}^0, is 0. Given that primary recovery will be shifted to secondary at time t_{R1} (i.e. the life time of primary recovery is t_{R1}), the recovery efficiency for secondary recovery is calculated using

$$E_{R2}(t) = E_{R1}(t_{R1}) + \Delta E_{R2}^\infty (1 - e^{-\frac{t - t_{R1}}{\tau_2}}) \tag{3}$$

where $E_{R1}(t_{R1})$ is the recovery efficiency at the end of primary recovery and ΔE_{R2}^∞ is the theoretical ultimate recovery efficiency increment caused by converting from primary to secondary recovery. Similarly, the recovery efficiency equation for n'th recovery phase is

$$E_{R(n)}(t) = E_{R(n-1)}\left(\sum_{i=1}^{n-1} t_{Ri}\right) + \Delta E_{R(n)}^\infty (1 - e^{-\frac{t - \sum_{i=1}^{n-1} t_{Ri}}{\tau_n}}) \tag{4}$$

where t_{Ri} is the life time of the i'th recovery phase, $E_{R(n-1)}(\sum_{i=1}^{n-1} t_{Ri})$ is the recovery efficiency at the end of (n-1)'th recovery phase and $\Delta E_{R(n)}^\infty$ is the theoretical ultimate recovery efficiency increment caused by shifting from (n-1)'th recovery to n'th recovery.

Economic Model. The net present value (NPV) is a commonly used performance indicator for reservoir management. NPV is defined as

$$NPV = \sum_{t=0}^{T} \frac{CF_t}{(1+r)^t} \tag{5}$$

where T is the time length of a production life cycle, CF_t is the cash flow from times t-1 to t, and r is the discount rate. We calculate CF_t as

$$CF_t = P_o N[E_R(t) - E_R(t-1)] - (C_t - O_t) \tag{6}$$

where P_o is the oil price, N is the estimated OOIP, C_t is the capital costs from times t-1 to t, and O_t is the operating costs from times t-1 to t.

When multiple recovery phases are involved, the NPV is a function of the life times of each phase. Our objective for production optimization is to maximize the NPV by adjusting the life times of each phase, i.e. to identify the optimal time of shifting the current recovery phase to the next recovery phase.

This work focuses on only geological uncertainties. Thus, the economic model parameters P_o, C_t, O_t and r are assumed to be certain values. Refer to Thomas and Bratvold (2015) for the implementation of the LSM algorithm with regards to oil and gas price uncertainties for a gas cap blowdown problem. The uncertainty in the reservoir properties is represented by assigning a probability distribution to each of the model parameters E_{R1}^{∞}, $\Delta E_{R(n)}^{\infty}$, τ_n and N. For a risk-neutral decision maker[2], the objective is to optimize the expected NPV (ENPV) over the uncertain model parameters.

Integrating the Impact of Information in Production Optimization

In decision making, the optimization problem on deciding the life times of each recovery phase can be interpreted as: for a series of decision points in time; we must decide whether we should continue with the current recovery phase or shift to the next recovery phase at each decision point. This is a sequential decision-making problem. As production continues, we are continuously getting more information from, for example, production data. These data are used to inform our uncertainty in the reservoir properties through history matching, and support our decision making. Thus, in production optimization, we should include the impact of information. A state-of-the-art approach of integrating information in production optimization is the closed loop reservoir management (CLRM) approach. A more advanced approach is the fully structured reservoir management (FSRM) approach. The following briefly presents these two approaches and illustrates the difference between them.

Closed Loop Reservoir Management. In CLRM (also known as "real time reservoir management", "smart reservoir management" or "closed loop optimization"), the loop of history matching and production optimization is closed by continuously update a production model with several realizations and performing life-cycle optimization whenever new data becomes available (Brouwer et al. 2004; Nævdal et al. 2006; Chen et al. 2009; et al. 2009; Jansen et al. 2009).

Several realizations of an initial production model are built based on the prior knowledge of the reservoir, and an initial production strategy for the whole life cycle is determined by optimizing the expected value (EV) over the initial realizations of the production model. The initial production strategy is applied to the real field until new data becomes available. The new data is used to update the production model, and a new production strategy for the period from the time after the new data has become available to the end time of the life cycle is determined by optimizing the EV over the updated realizations of the production model. The new production strategy is applied to the real field until new data becomes available. Repeating the process keeps the reservoir management up to date.

In the viewpoint of sequential decision making, the CLRM is a myopic decision policy where the uncertainties associated with current available data are considered but the uncertainties associated with future data are not (Kullawan 2016; Thomas 2016). A decision tree[3] representation for this approach is illustrated in **Fig. 1** where D_k denotes a decision made at time t_k and U_k denotes the uncertainties associated with current available data until t_k. U_k is commonly represented by a production model with several realizations. The production strategy D_k, D_{k+1}, …, D_{end} is determined with the consideration of only the immediate relevant uncertainty U_k.

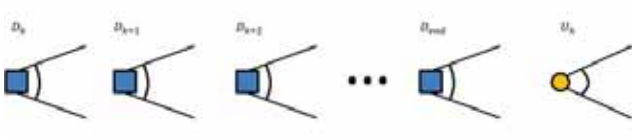


Fig. 1—Illustration of the decision tree for the CLRM.

The advantage of CLRM is that it greatly simplifies the structure of a reservoir management decision problem and consequently, requires less computational cost for solving the problem. Its significant drawback is that it does not reflect the full structure of a reservoir management decision problem and thus, may lead to a sub-optimal production strategy.

Fully Structured Reservoir Management. The full structure of a sequential decision-making problem is illustrated by the decision tree in **Fig. 2**. The FSRM approach aims to solve for the optimal decision policy based on a fully structure decision tree.

[2] Risk-neutrality is a risk attitude of being neither risk-averse nor risk-seeking. A risk-neutral decision maker uses only expected value as his decision criterion, i.e., he is indifferent between choices with equal expected values no matter what the risk of a choice is.

[3] In a decision tree, a square represents a decision node and a circle represents an uncertainty node.

The FSRM explicitly considers both the uncertainties associated with currently available data (U_k) and that with future data (U_{k+1}, …, U_{end}). Therefore, the current decision (D_k) does not depend on only the uncertainties that a decision maker have learnt so far but also the uncertainties that the decision maker will learn in the future. The FSRM includes the decision maker's reactions (D_{k+1}, …, D_{end}) to the uncertainties associated with future data (U_{k+1}, …, U_{end}).

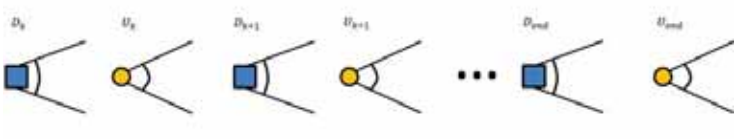


Fig. 2—Illustration of the decision tree for FSRM.

Solving the decision tree in Fig. 2 gives the optimal production strategy. However, the FSRM can be computationally intensive or even prohibitive compared to the CLRM.

Least-Squares Monte Carlo Algorithm for Fully Structured Reservoir Management

The Least-Squares Monte Carlo (LSM) is a promising algorithm for solving a FSRM problem. It was proposed by Longstaff and Schwartz (2001) for an American option problem which involves a yes-no decision: at anytime, an option holder can decide whether to immediately exercise the option at current stock price or to continue the option for exercising it at future stock price. The stock price is an uncertain quantity in an American option problem, whose uncertainty can be modeled as a Markov process, and thus, the future value of an alternative conditions only to the current stock price.

However, geological uncertainty does not only depend on the currently measured data but also the previously measured data and consequently, the future value of an alternative conditions to currently and previously measured data. To deal with our problem at hand, we slightly modify the LSM algorithm to include the dependency of the future value on the currently and previously measured data. The two central steps of our modified LSM algorithm are

1) Monte Carlo Simulation (MCS) Step:
   N independent samples of model parameters representing geological uncertainties are generated using MCS. For one sample of model parameters, forward modeling is performed to provide modeled production data from t_0 (time 0) to t_{end} (end time); and then, random noises generated based on the statistics of the measurement errors[4] are added to the modeled production data to provide a sample of measured data. Because this sample of measured data consists of a series of data points in time, it is also called a path of measured data. Repeating this procedure for each of the N sampled sets of model parameters, we obtain N paths of measured data, i.e. $\boldsymbol{Y} = [\boldsymbol{y}_1, \boldsymbol{y}_2, \dots, \boldsymbol{y}_N]^{\mathrm{T}}$.
2) Least Squares Step:
   For the i'th sample of model parameters, the NPV of alternative a is calculated, giving $NPV_i(a)$. This is repeated for the N sampled sets of model parameters, resulting in $[NPV_1(a), NPV_2(a), \dots, NPV_N(a)]^{\mathrm{T}}$. To estimate the ENPV with alternative a conditional on the measured data, $ENPV(a)|\boldsymbol{Y}$, we regress $[NPV_1(a), NPV_2(a), \dots, NPV_N(a)]^{\mathrm{T}}$ on $[\boldsymbol{y}_1, \boldsymbol{y}_2, \dots, \boldsymbol{y}_N]^{\mathrm{T}}$. This procedure is repeated for each of the alternatives.

More detailed steps of applying the LSM algorithm will be illustrated in the following example.

Illustrative Example

This section uses a simple example to illustrate the fully structured decision tree for the problem of determine the optimal IOR start time, the difference between CLRM and FSRM solutions and how to use the LSM algorithm to solve the problem.

Problem Setting. A field has a life cycle of 15 years. We consider a yes-no decision at Years 0, 5 and 10: whether primary recovery should be shifted to secondary recovery. The shift can only happen once. The production is modeled by a function $f(\cdot)$ with parameters $\boldsymbol{x}$. The geological uncertainty is represented by 3 realizations (R1, R2

[4] In the context of history matching, a common practice is to treat the data fluctuations caused by changes in operating conditions as caused by measurement errors (Jochen and Spivey 1996).

and R3) of the production model parameters ($\boldsymbol{x}_1$, $\boldsymbol{x}_2$ and $\boldsymbol{x}_3$). These realizations are a priori equi-probable. **Fig. 3** illustrates the primary recovery efficiency as a function of time for the realizations.

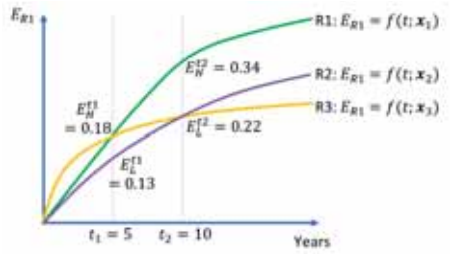


Fig. 3—Primary recovery as a function of time for 3 geological realizations.

The NPVs corresponding to each alternative and each realization are in **Table 1**. All the values are in million USD.

Realization	Shift at Time			No Shift
	t_0	t_1	t_2	
R1	1700	1880	2170	3500
R2	1800	2240	2430	2130
R3	2430	3540	3110	1810

Table 1—NPVs corresponds to each alternative and each realization.

The measured primary recovery efficiency is used to inform the decisions. The likelihood functions, listed in **Table 2**, are used to describe the measurement errors. For example, given that the truth is R1, the probability of the measurement at t_1 saying "high recovery" (i.e. a measured data of 0.18) is 0.75, and the probability of the measurement at t_2 saying "high recovery" (i.e. a measured data of 0.34) is 0.80.

		Given the truth is		
		R1	**R2**	**R3**
Measurement at t_1 says	**"High Recovery" ("E_H^{t1}")**	0.75	0.25	0.75
	"Low Recovery" ("E_L^{t1}")	0.25	0.75	0.25
Measurement at t_2 says	**"High Recovery" ("E_H^{t2}")**	0.80	0.20	0.20
	"Low Recovery" ("E_L^{t2}")	0.20	0.80	0.80

Table 2—Likelihood functions for the measured primary recovery efficiency.

Decision and Expected Value without Information. If the impact of future information on the series of decisions is not considered, the ENPV is calculated based on the a priori probability of each realization (i.e. 1/3). For this case, the optimal decision is referred to as the decision without information (DWOI), and the corresponding optimal EV is referred to as the EV without information (EVWOI). Thus, the DWOI is "shift at time t_2" giving the EVWOI of $2570 million.

Decision and Expected Value with Perfect Information. Perfect information reveals the truth. For example, if the perfect information indicates R1, then R1 is the truth and we can identify the optimal decision by maximizing the NPV of R1. The optimal decision and EV for this case are referred to as the decision and EV with perfect information (DWPI and EVWPI), respectively. The DWPI is "no shift" for R1, "shift at time t_2" for R2 and "shift at time t_1" for R3. The probability of the perfect information indicating each realization is 1/3. This results in the

EVWPI of $3156.7 million. The value-of-perfect-information (VOPI)[5] is thus EVWPI – EVWOI = $3156.7 million – $2570.0 million = $586.7 million.

FSRM Solution: Decision and Expected Value with Imperfect Information. The likelihood functions in Table 2 indicate the reliability of the data and show that the information is imperfect. The optimal decision and EV for such case are referred to as the decision and EV with imperfect information (DWII and EVWII), respectively.

We construct the full decision three for this problem setting, as illustrated in **Fig. 4**, to help in identifying the DWII and EVWII. The conditional probabilities in the tree are calculated using Bayes' theorem. From the tree, we can see that we should continue with primary recovery at t_1 no matter what the measurement at t_1 says, and continue with primary recovery at t_2 if the measurement at t_2 says "high recovery" while shift to secondary recovery at t_2 if the measurement at t_2 says "low recovery". The corresponding EVWII is $2818 million, resulting in a value-of-information (VOI) of EVWII – EVWOI = $2818 million – $2570 million = $248 million.

CLRM Solution. We solve the same problem using the CLRM approach. Unlike that the FSRM approach which solves the decision tree backwards (i.e. rollback from the right nodes to the left-most node in the tree), the CLRM approach solves the tree forwards. For example, given the measurement at t_1 saying "high recovery", the conditional probabilities of R1, R2 and R3 are 0.43, 0.14 and 0.43, respectively; based on these probabilities, the EV is $2642.9 million for "shifting at t_1", $2610 million for "shifting at t_2" and $2580 million for "no shift". Because the alternative of "shifting at t_1" gives the highest EV, the immediate decision at t_1 is to shift to secondary recovery. The decision policy of the CLRM approach is listed in **Table 3**. Applying this decision policy to the decision tree in Fig. 2 results in an EV of $2641.2 million corresponding to a special VOI[6] of $71.2 million.

Compared to the FSRM approach, the CLRM approach gives a difference decision policy and a lower VOI. The CLRM approach realizes only 71.2/248 = 28.7% of the value that the information can create. This is because when the impact of future information is not considered in the CLRM approach, the CLRM approach makes a decision based on the immediate gain rather than evaluating whether it is worth continuing for gathering more information as the FSRM approach does.

Decision at Time		
t_0	t_1	t_2
Continue with primary recovery.	Shift to secondary recovery if measurement at t_1 says "high recovery."	No action because it has been shifted to secondary recovery at t_1.
	Continue with primary recovery if the measurement at t_1 says "low recovery."	Continue with primary recovery if the measurement at t_1 says "low recovery" and at t_2 says "high recovery."
		Shift to secondary recovery if the measurement at t_1 says "low recovery" and at t_2 says "low recovery."

Table 3—Decision policy of the CLRM approach.

[5] The concept of value-of-information (VOI) originates from the decision analysis community. VOI is an indicator of the maximal paying price or cost of an information gathering activity. VOPI is the upper limit of VOI. For a more comprehensive description of the concept of VOI, consult Raiffa and Schlaifer (1961), Howard (1966), and Bratvold et al. (2009). For the illustration and discussion of adopting the concept of VOI in history matching contexts, refer to Hong and Bratvold (2017).

[6] Per the strict definition of VOI, it is referred exclusively to that associated with the optimal decision policy (i.e. the FSRM approach in our case). We relax the requirement here and use "special VOI" to refer to the VOI for a specific sub-optimal decision policy (i.e. the CLRM approach in our case).

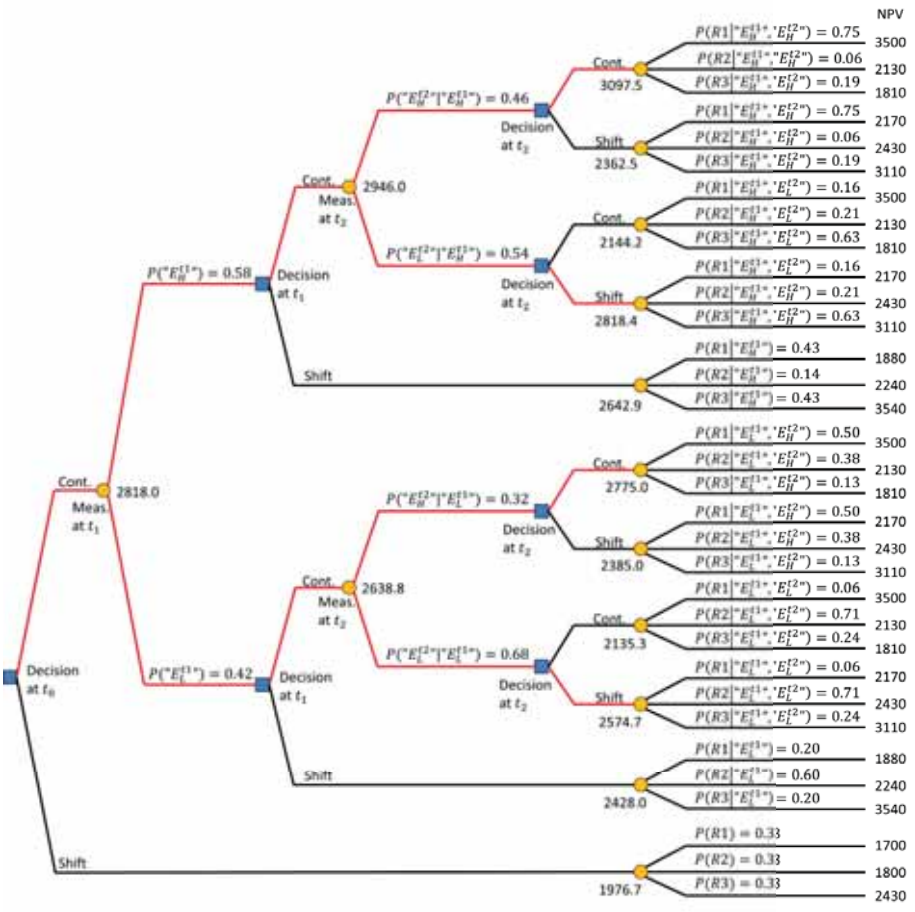


Fig. 4—Fully structured decision tree for the illustrative example.

Solving the FSRM Problem Using the LSM Algorithm. We use the modified LSM algorithm to solve the FSRM problem as shown in Fig. 4. For illustration, we use only 6 Monte Carlo samples. The number of samples will be increased later to verify the modified LSM algorithm. We first sample 6 paths of the measured data as listed in **Table 4**. Then, we start the LSM from the last decision point as we solve the decision tree backwards. **Table 5** lists the NPVs corresponding to alternatives of "continue (with primary recovery)" and "shift (to secondary recovery)" at t_2 for each path.

Path	Geological Realization	Data at	
		t_1 ("E_{R1}^{t1}")	t_2 ("E_{R1}^{t2}")
1	R2	0.18	0.22
2	R1	0.18	0.34
3	R3	0.13	0.34
4	R1	0.13	0.34
5	R3	0.18	0.22
6	R2	0.13	0.22

Table 4—Path of measured data for LSM.

Path	Geological Realization	NPV [million USD]	
		Continue	Shift
1	R2	2130.0	2430.0
2	R1	3500.0	2170.0
3	R3	1810.0	3110.0
4	R1	3500.0	2170.0
5	R3	1810.0	3110.0
6	R2	2130.0	2430.0

Table 5—NPVs for alternatives at time t_2 for LSM.

When we are making the decision at t_2, we have obtained the data at both t_1 and t_2. Therefore, the ENPV is conditional on the data at both t_1 and t_2. We estimate the ENPV by regressing NPV on data. Using $ENPV|data = a_0 + a_1"E_{R1}^{t1}" + a_2"E_{R1}^{t2}"$, the resulting conditional expectation function for "continue" is $ENPV("continue")|data = -979.3 + 6850.0 \cdot "E_{R1}^{t1}" + 8562.5 \cdot "E_{R1}^{t2}"$. The same is done for "shift". **Table 6** lists the resulting ENPV corresponding to alternatives of "continue" and "shift" at t_2 for each path. Now, we can make the optimal decision by taking the alternative that gives the highest ENPV for each path. The optimal decision is colored in red in Table 6.

Path	ENPV [million USD]	
	Continue	Shift
1	2137.5	2635.0
2	3165.0	2440.0
3	2822.5	2505.0
4	2822.5	2505.0
5	2137.5	2635.0
6	1795.0	2700.0

Table 6—ENPVs for alternatives at time t_2 for LSM.

We move one time step backwards to t_1. **Table 7** lists the NPVs corresponding to alternatives of "continue" and "shift" at t_1 for each path. The NPVs for "continue" in Table 7 come from the NPVs in Table 5 for the optimal decisions shown in Table 6.

Path	Geological Realization	NPV [million USD]	
		Continue	Shift
1	R2	2430.0	2240.0
2	R1	3500.0	1880.0
3	R3	1810.0	3540.0
4	R1	3500.0	1880.0
5	R3	3110.0	3540.0
6	R2	2430.0	2240.0

Table 7—NPVs for alternatives at time t_1 for LSM.

When we are making the decision at t_1, the data at only t_1 is available but not t_2. Therefore, the ENPV is conditional on the data at only t_1. Using $ENPV|data = b_0 + b_1"E_{R1}^{t1}"$, the resulting conditional expectation function for "continue" is $ENPV("continue")|data = 2703.7 + 600.0 \cdot "E_{R1}^{t1}"$. The same is done for "shift". **Table 8** lists the resulting ENPV corresponding to alternatives of "continue" and "shift" at t_1 for each path, and the optimal decisions are in red.

Path	ENPV [million USD]	
	Continue	Shift
1	3013.3	2553.3
2	3013.3	2553.3
3	2580.0	2553.3
4	2580.0	2553.3
5	3013.3	2553.3
6	2580.0	2553.3

Table 8—ENPVs for alternatives at time t_1 for LSM.

We move to t_0. **Table 9** lists the NPVs corresponding to alternatives of "continue" and "shift" at t_0 for each path. The NPVs for "continue" in Table 9 come from the NPVs in Table 7 for the optimal decisions shown in Table 8. Because we have no data available at t_0, we calculate the ENPVs for "continue" and "shift" by calculating the mean of the values in columns 4 and 5, respectively, in Table 9. The resulting ENPV for "continue" is \$2796.7 million and that for "shift" is \$1976.7 million. Thus, the optimal decision at t_0 is "continue", and the optimal ENPV is \$2796.7 million.

Path	Geological Realization	NPV [million USD]	
		Continue	Shift
1	R2	2430.0	1800.0
2	R1	3500.0	1700.0
3	R3	1810.0	2430.0
4	R1	3500.0	1700.0
5	R3	3110.0	2430.0
6	R2	2430.0	1800.0

Table 9—NPVs for alternatives at time t_0 for LSM.

The optimal decision policy can be determined by looking back at Table 8 and Table 6, and it is represented by **Table 10**.

Path	Data at		Optimal Decision at		
	t_1 ("E_{R1}^{t1}")	t_2 ("E_{R1}^{t2}")	t_0	t_1	t_2
1	0.18	0.22	Continue	Continue	Shift
2	0.18	0.34	Continue	Continue	Continue
3	0.13	0.34	Continue	Continue	Continue
4	0.13	0.34	Continue	Continue	Continue
5	0.18	0.22	Continue	Continue	Shift
6	0.13	0.22	Continue	Continue	Shift

Table 10—Table representation of optimal decision policy solved using LSM.

The optimal ENPV and optimal decision policy solved using the modified LSM are different from the analytical solution from the decision tree (Fig. 4) because we used only 6 paths. The accuracy can be improved by increasing the number of paths. We increase to 500000 paths and repeat it for 1000 times. This gives a mean of estimated optimal ENPV of \$2818.03 million, which is almost the same as the analytical solution \$2818.00 million, and a standard deviation of \$0.80 million (which is relatively small). This verifies the accuracy of using the modified LSM algorithm to solve the FSRM problem.

Application of the Two-Factor Production Model and the Least Squares Monte Carlo Algorithm

This section uses a more realistic example to illustrate the application of the two-factor production model and the LSM algorithm for analyzing the optimal IOR start time. This example is more complicate than the illustrative

example. Thus, a fully structured decision tree cannot be visualized and the problem cannot be solved by simply rollback a decision tree. The modified LSM algorithm is therefore of significant use for this example.

Problem Setting. In the development phase of an oil field, we consider two recovery phases—primary and secondary recovery. We want to analyze the optimal time of shifting from the primary recovery to the secondary recovery and the optimal time of terminating the production (i.e. to identify the optimal life times of primary and secondary recoveries). This analysis provides useful insights in how the field should be developed and how the leaning over time will impact the decisions. We consider a maximal life cycle of 50 years. The shift can happen at the beginning of any year but only once, and the same for the termination. In other words, we need to decide whether to continue with the primary recovery, to shift to the secondary recovery, or to terminate the production at the beginning of each year from Year 1 to Year 50.

The oil production is modeled using the two-factor production model. We assign probability distributions to the model parameters based on a priori knowledge to represent the uncertainties in reservoir properties and the effects of the primary and secondary recovery mechanism. All the parameters are assumed to be normally distributed and their means and standard deviations are in **Table 11**. The correlation coefficient between the parameters are in **Table 12**.

Parameter	E_{R1}^{∞} [fraction]	τ_1 [years]	ΔE_{R2}^{∞} [fraction]	τ_2 [years]
Mean	0.20	16	0.15	7
Standard Deviation	0.05	2	0.05	1.5

Table 11—Means and standard deviations for the normal distributions of the production model parameters.

	$\boldsymbol{E_{R1}^{\infty}}$	$\boldsymbol{\tau_1}$	$\boldsymbol{\Delta E_{R2}^{\infty}}$	$\boldsymbol{\tau_2}$
$\boldsymbol{E_{R1}^{\infty}}$	1	0.2	-0.7	0.1
$\boldsymbol{\tau_1}$	0.2	1	-0.3	-0.2
$\boldsymbol{\Delta E_{R2}^{\infty}}$	-0.7	-0.3	1	-0.3
$\boldsymbol{\tau_2}$	0.1	-0.2	-0.3	1

Table 12—Correlation coefficients between the production model parameters.

Because we focus on the geological uncertainty, we assume all the parameters of the economic model are known and their values are in **Table 13**. The capital cost is paid at the beginning of each recovery phase. The cashflow is discounted yearly.

Parameter	Value	Unit
OOIP	240	million bbl
Oil Price	55	USD/bbl
Capital Cost (Primary)	50	million USD
Capital Cost (Secondary)	50	million USD
Operating Cost (Primary)	9.1	million USD/year
Operating Cost (Secondary)	21.1	million USD/year
Discount Rate	12%	yearly

Table 13—Values of economic parameters.

The recovery efficiency is measured at the end of each year and is used to support the decisions. The measurement error is assumed to be normally distributed with zero mean and a standard deviation of 0.001.

Results. The DWOI is 6 year's primary recovery followed by 18 year's secondary recovery, leading to a life cycle of 24 years. The corresponding EVWOI is $918.9 million. If we can obtain perfect information on the reservoir properties and the effects of primary and secondary recovery mechanisms before the production starts, the resulting EVWPI is $1057.6 million. Thus, the VOPI is $138.7 million. This value indicates that the total costs of any additional data gathering activity (e.g., conducting a seismic survey), no matter how sophisticated it is, for better understanding the reservoir properties or the effects of primary and secondary recovery mechanisms should not exceed $138.7 million. The FSRM approach leads to an EVWII of $991.8 million for the measure of recovery efficiency. The corresponding VOI is $72.9 million. Therefore, the company should not pay more than $72.9 million to buy and install devices for measuring the recovery efficiency.

Fig. 14 compares the cumulative density functions (CDFs) of NPVs corresponding to the DWOI, DWII and DWPI. We see that the FSRM approach moves the CDF of NPVs corresponding to the DWOI to the right (i.e. the NPV increases) because it allows for learning over time. We notice that some realizations (around 5%) end up with a smaller NPV with DWII than the NPV with DWOI. The possible reasons are: (1) the downside (e.g., the theoretical ultimate recovery efficiency increment turns out to be very small) may show up even though the decision of "shift" is optimal, at the time it is made, based on the EV over possible future events, and/or (2) a sub-optimal decision may be made because the LSM algorithm is an approximation method. The DWPI further moves the CDF curve to the right, leading to a higher ENPV than that of the DWOI and DWII.

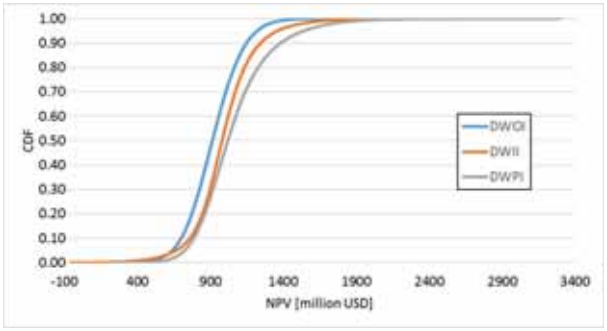


Fig. 14—CDFs of NPVs corresponding to DWOI, DWII and DWPI.

The probability density functions (PDFs) and CDFs of the primary recovery, secondary recovery and total life times corresponding to the DWII are in **Figs. 15**–**17** where the DWOI is marked in red.

Fig. 15 shows that the P10–P90 interval of the primary recovery life time ranges from 1 to 20 years (i.e. there is 80% chance that the recovery mechanism should be shifted from primary to secondary anytime from Year 1 to Year 20). This is a wide range. The specific timing of shifting depends on the measure of recovery efficiency. There is only 4% chance that it is optimal to shift after 6 year's production as the DWOI indicates while the most probable time of shifting is after 1 year's production with a probability of more than 20%. This informs that the facilities for secondary recovery should be ready on site after 1 year's production rather than postponed to 6 years. Otherwise, the company will lose the flexibility to shift before 6 year's production when the measured data indicating "shift", which has a chance of about 35% to happen.

Fig. 16 shows that the P10–P90 interval of the secondary recovery life time ranges from 12 to 22 years. The 18 year's secondary recovery life time of the DWOI has about 11.5% chance to happen while the most probable (13%) secondary recovery life time is 19 years. Besides, there is 40% chance that the secondary recovery life time is shorter than 18 years and 48.5% chance longer.

Fig. 17 shows that the P10–P90 interval of the total life time ranges from 19 to 36 years. This informs that it would better to has a license for production up to 36 years and a flexibility to terminate the production or sell their contract any time from Year 19 to Year 36. If the license for production lasts only 24 years according to the DWOI, the company would have a chance of more than 60% to lose the opportunity of producing longer.

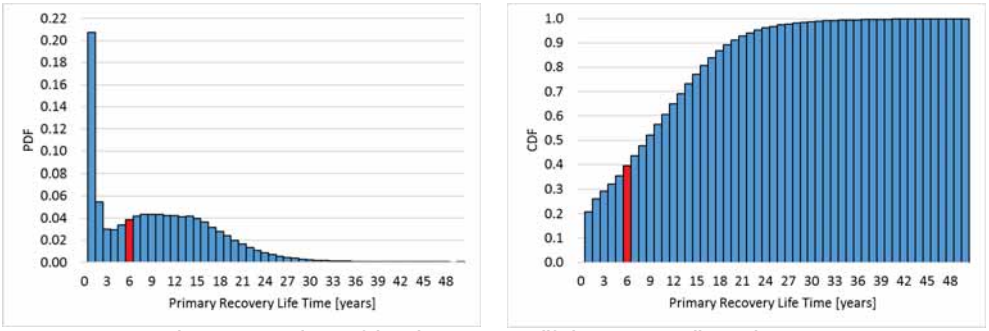


Fig. 15—PDF and CDF of the primary recovery lifetime corresponding to the DWII.

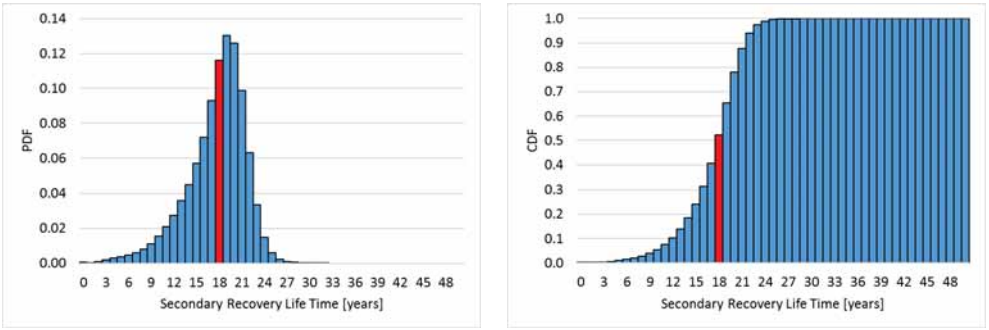


Fig. 16—PDF and CDF of the secondary recovery lifetime corresponding to the DWII.

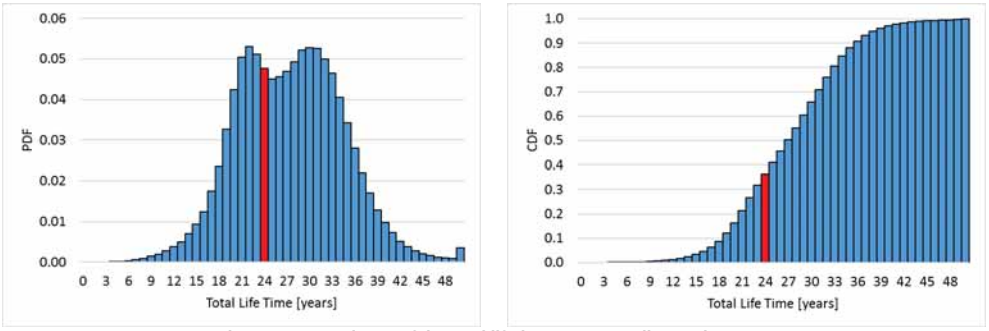


Fig. 17—PDF and CDF of the total lifetime corresponding to the DWII.

Fig. 18 illustrates the probabilities for different combinations of primary and secondary recovery life times corresponding to the DWII. The more probable combinations are primary recovery lifetime for 1 year combined with secondary recovery life time for 17 to 23 years. These combinations have a probability more than twice than the other combinations. This may because we assigned 16 years to the mean value of τ_1 and 7 years to the mean value of τ_2, which means that the secondary recovery mechanism will recovery oil faster than the primary recovery mechanism. Per the NPV calculation, the higher the cash flow is at the early times, the higher the NPV is. Shifting primary recovery to secondary recovery at an early time will results in a higher NPV than doing that at a later time.

The highly probable (>0.5%) combinations can be separated into two groups: (1) small span of primary recovery life time (1–2 years) with large span of secondary recovery life time (13–24 years), and (2) small span of secondary recovery life time (17–21 years) with large span of primary recover life time (5–17 years). The correlation

coefficient between the primary and secondary recovery life time is -0.44. This negative linear correlation is because there is a strong negative linear correlation between the theoretical ultimate primary recovery efficiency and the theoretical ultimate recovery efficiency increment as assigned in Table 12.[7] That is, if the primary recovery can already lead to a relatively high ultimate recovery efficiency, then the recovery efficiency increment of the secondary recovery will be relatively low (i.e. more effective the primary recovery is, less effective the secondary recovery is). Consequently, we can expect a relatively short secondary recovery life time if the data inform us to continue with primary recovery for a relatively long time.

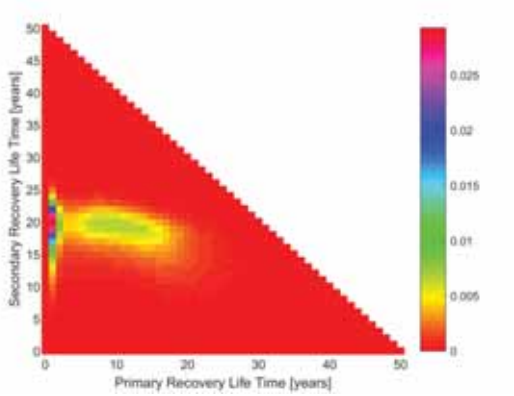


Fig. 18—Probabilities for different combinations of primary and secondary recovery life times.

Discussion

The two-factor production model is useful and tractable. When it is combined with the LSM algorithm, a fast analysis of optimal IOR start time and life time can be realized with the consideration of the impact of future information. Although the model is very simple, the analysis provides many useful insights in the decision problem. The FSRM solution contains a decision policy, which reflects a decision maker's optimal reaction to information, rather than a single value of the decision variable. The solutions of this fast analysis can then be interrogated with sophisticated simulations if adding details in the production model is deemed to be relevant and material for the decisions at hand.

The LSM algorithm is an approximate dynamic programming method. Its accuracy depends on two factors: the number of Monte Carlo samples and how well the regression function approximates the real EV. Because we used very many Monte Carlo samples (500000), we believe that the sampling error is negligible. We used a first order linear regression function, the simplest form of a regression function. The illustrative example tested this regression function and showed that it provided high accuracy. It seems from the results that this regression function performed well for the more complex application of the LSM algorithm. Although a more complicated regression function (for example, with higher order) can improve the accuracy, we think the simple fist order linear regression function is sufficient to extract high quality insight for our problem setting.

This work focused only on the geological uncertainty represented by the two-factor model parameters. However, the uncertainty in economic model parameters can be included by adding these variables into the regression function. It is easier to handle the economic variables than the production data. The economic variables are usually assumed to follow a Markov chain. Thus, the EV is conditional on only their most recent value unlike that the EV, which is conditional on the full path of the production data.

We considered only two recovery phases—the primary and secondary recovery phases. More recovery phases should be included for a more comprehensive analysis of the production life cycle. The inclusion of more recovery phases (i.e. introducing more alternatives at each decision point) will require more computational time for the LSM algorithm. Generally, the computational time increases exponentially with the number of alternatives. Thus,

[7] We assign a negative linear correlation between the theoretical ultimate primary recovery efficiency and the theoretical ultimate recovery efficiency increment for illustrative purpose only. The correlation can be any value from -1 to 1, depending on reservoir properties, well locations, recovery mechanisms, production strategies, etc.

including too many recovery phases may become computationally prohibitive even when a computationally attractive production model is used.

Conclusions

This paper reviewed a useful and tractable model for modeling various recovery phases involved in a life cycle of oil production and illustrated the detailed steps of implementing the LSM algorithm to solve a sequential decision-making problem.

We demonstrated the value of considering the impact of future information in the analysis of the optimal IOR start time. We used an illustrative example to discuss the full structure of the relevant decision problem. Although the state-of-the-art reservoir management approach—CLRM—can lead to larger EV than the case without considering the impact of information, its solution may be sub-optimal. It is because it considers only the information obtained before a decision is made but not the information obtained after that decision. The optimal solution can be guaranteed using the FSRM approach.

A more realistic example demonstrated the application of the two-factor model and the LSM algorithm. The results showed that the FSRM approach can significantly improve the decisions and leads to a significant increase in ENPV. The resulting decision policy represents the optimal action a decision maker should take according to his/her observation.

We conclude that using the two-factor model and the LSM algorithm can, with limited computation resources, provide useful insights in the problem of deciding the optimal IOR start time.

Acknowledgements

The authors acknowledge the support of the Research Council of Norway and industry partners: ConocoPhillips, Skandinavia AS, BP Norge AS, Det Norske Oljeselskap AS, Eni Norge AS, Maersk Oil Norway AS, DONG Energy A/S, Denmark, Statoil Petroleum AS, ENGIE E&P NORGE AS, Lundin Norway AS, Halliburton AS, Schlumberger Norge AS, and Wintershall Norge AS of The National IOR Centre of Norway.

Aojie Hong would like to thank the staffs and colleagues at the University of Texas at Austin for their support during his visiting. Larry W. Lake holds the Shahid and Sharon Ullah Chair at the University of Texas.

References

Alkhatib, A. M., and King, P. R. 2011. Applying Real Options Theory in Determining Optimal Policies for a Surfactant Flood. Presented at the SPE Enhanced Oil Recovery Conference, Kuala Lumpur, Malaysia, 19–21 July. SPE-144869-MS. http://dx.doi.org/10.2118/144869-MS.

Bickel, J. E., and Bratvold, R. B. 2008. From Uncertainty Quantification to Decision Making in the Oil and Gas Industry. *Energy Exploration & Exploitation* **26** (5): 311–325. https://doi.org/10.1260/014459808787945344.

Bratvold, R. B., and Begg, S. 2010. *Making Good Decisions*, first edition. Texas, USA: Society of Petroleum Engineers.

Bratvold, R. B., Bickel, J. E., and Lohne, H. P. 2009. Value of Information in the Oil and Gas Industry: Past, Present, and Future. *SPE Reservoir Evaluation & Engineering* **12** (4): 630–638. SPE-110378-PA. http://dx.doi.org/10.2118/110378-PA.

Brouwer, D. R., Nævdal, G., Jansen, J. D. et al. 2004. Improved Reservoir Management Through Optimal Control and Continuous Model Updating. Presented at the SPE Annual Technical Conference and Exhibition, Houston, USA, 26–29 September. SPE-90149-MS. https://doi.org/10.2118/90149-MS.

Chen, Y., Oliver, D. S., Zhang, D. 2009. Efficient Ensemble-Based Closed-Loop Production Optimization. *SPE Journal* **14** (4): 634–645. SPE-112873-PA. https://doi.org/10.2118/112873-PA.

Howard, R. A. 1966. Information Value Theory. *IEEE Transactions on Systems Science and Cybernetics* **2** (1): 22–26. http://dx.doi.org/10.1109/TSSC.1966.300074.

Howard, R. A., and Abbas, A. E. 2016. *Foundations of Decision Analysis*, global edition. Harlow, England: Pearson Education Limited.

Hong, A. J., and Bratvold, R. B. 2017. Value of Information from History Matching—How Much Information is Enough? Presented at the IOR NORWAY 2017—19th European Symposium on Improved Oil Recovery, Stavanger, Norway, 24–27 April. https://doi.org/10.3997/2214-4609.201700327.

Hong, A. J., Bratvold, R. B., and Nævdal, G. 2017. Robust Production Optimization with Capacitance-Resistance Model as Precursor. *Computational Geosciences*. https://doi.org/10.1007/s10596-017-9666-8.

Hem, Ø. D., Svendsen, A., Fleten, S. E. et al. 2011. The Option to Switch from Oil to Natural Gas in Active Offshore Petroleum Fields. Presented at the Annual International Real Options Conference, Turku, Finland,

15–18 June.

Jafarizadeh, B., and Bratvold, R. B. 2012. Two-Factor Oil-Price Model and Real Option Valuation: an Example of Oilfield Abandonment. SPE Economics & Management **4** (3): 158–170. SPE-162862-PA. http://dx.doi.org/10.2118/162862-PA.

Jafarizadeh, B., and Bratvold, R. B. 2013. Sell Spot or Sell Forward? Analysis of Oil-Trading Decisions with the Two-Factor Price Model and Simulation. SPE Economics & Management **5** (3): 80–88. SPE-165581-PA. http://dx.doi.org/10.2118/165581-PA.

Jansen, J. D., Brouwer, R., Douma, S. G. 2009. Closed Loop Reservoir Management. Presented at the SPE Reservoir Simulation Symposium, The Woodlands, USA, 2–4 February. SPE-119098-MS. https://doi.org/10.2118/119098-MS.

Jochen, V. A., and Spivey, J. P. 1996. Probabilistic Reserves Estimation Using Decline Curve Analysis with the Bootstrap Method. Presented at the SPE Annual Technical Conference and Exhibition, Denver, USA, 6–9 October. SPE-36633-MS. http://dx.doi.org/10.2118/36633-MS.

Kullawan, K. 2016. *A Bayesian Framework for Real-Time Optimization of Well Placement.* PhD Thesis. The University of Stavanger, Norway.

Longstaff, F., and Schwartz, E. 2001. Valuing American Options by Simulation: a Simple Least-Squares Approach. *Review of Financial Studies* **14** (1): 113–147. http://dx.doi.org/10.1093/rfs/14.1.113.

Nævdal, G., Brouwer, D. R., Jansen, J. D. 2006. Waterflooding Using Closed-Loop Control. *Computational Geosciences* **10** (1): 37–60. https://doi.org/10.1007/s10596-005-9010-6.

Raiffa, H., and Schlaifer, R. 1961. *Applied Statistical Decision Theory*. Boston: Division of Research, Graduate School of Business Administration, Harvard University (Reprint).

Parra-Sanchez, C. 2010. *A Life Cycle Optimization Approach to Hydrocarbon Recovery*. Master Thesis. The University of Texas at Austin, USA.

Thomas, P. 2016. *Managing Uncertainty with Decision Analysis in The Oil and Gas Industry*. PhD Thesis. The University of Stavanger, Norway.

Thomas, P, and Bratvold, R. B. 2015. A Real Options Approach to the Gas Cap Blowdown Decision. Presented at the SPE Annual Technical Conference and Exhibition, Houston, USA, 28–30 September. SPE-174868-MS. https://doi.org/10.2118/174868-MS.

Willigers, B. J. A., Begg, S. H., and Bratvold, R. B. 2011. Valuation of Swing Contracts by Least-Squares Monte Carlo Simulation. SPE Economics & Management **3** (4): 215–225. SPE-133044-PA. http://dx.doi.org/10.2118/133044-PA.

Wang, C., Li, G., Reynolds, A. C. 2009. Production Optimization in Closed-Loop Reservoir Management. *SPE journal* **14** (3): 506–523. SPE-109805-PA. https://doi.org/10.2118/109805-PA.

Paper IV

Robust Discretization of Continuous Probability Distributions for Value-of-Information Analysis

Bratvold, R.B., Thomas, P., Begg, S.H., and Hong, A.J.

Manuscript

Robust Discretization of Continuous Probability Distributions for Value-of-Information Analysis

R.B. Bratvold, P. Thomas, SPE, University of Stavanger, S. Begg, SPE, University of Adelaide, and A.J. Hong, SPE, Univeristy of Stavanger

Abstract

One of the most useful features of decision analysis is its ability to distinguish between constructive and wasteful information gathering. Value-of-Information (VoI) analysis evaluates the economic benefits of collecting additional information before making decisions.

VoI models describe the relationship between the uncertain quantities of interest, the reliability of the information and the decision criteria. Many uncertainty quantities are continuous in nature and the probability distributions that describe their uncertainty are discretized, and presented in decision trees, to simplify analysis. A three-point discretization of a continuous distribution is standard, and often preserves the main characteristics (central tendency, spread) of distributions that are close to symmetric. However, VoI studies rarely, if ever, include an analysis on the sensitivity of the VoI to the quality of the discretization.

An alternative approach to the three-point discretization, is to use Monte Carlo (MC) based methods for VoI analysis. The growth of computing power has contributed to the proliferation of MC based VoI methods as observed in several recent papers. Nevertheless, MC based VoI methods always raise questions concerning its robustness and its sensitivity to the number of samples.

In this work we investigate a variety of discretization techniques, for a range of typical information gathering situations. We also investigate the accuracy of MC based methods in calculating the VoI. The investigation utilizes a robust model that accurately calculates the VoI for any combination of continuous distributions. The key criterion for assessing a discretization techniques is whether or not it has a significant impact on the decision to collect the information. The goal of the work is to provide practical guidance on the level and technique of discretization required.

Introduction

Most of what petroleum engineers or geoscientists do involves "acquiring" information, with the aim of improving decision-making. "Information" is used here in a broad sense to cover such activities as acquisition of data, performing technical studies, hiring consultants, or performing diagnostic tests. In fact, other than to meet applicable regulatory requirements, the main reason for collecting any information, or doing any technical analysis, should be to make better decisions. The fundamental question for any information-gathering process is then whether the likely improvement in decision-making is worth the cost of obtaining the information. This is the question that the Value-of-Information (VoI) technique is designed to answer.

Schlaifer (1959) was the first to define VoI in the context of business decisions. Since then, the VoI approach has been refined and several important publications has appeared both outside and inside the petroleum industry. VoI analysis is also introduced and discussed in a number of textbooks (Schlaifer, 1961; Raiffa and Schlaifer, 1961; Raiffa, 1968; McNamee and Celona, 1987; Clemen and Reilly, 2014, Winkler, 2010; Newendorp and Schuyler, 2000; Bratvold and Begg, 2010; Howard and Abbas, 2015).

In any information-gathering activity, we are concerned with two fundamental uncertainties: (1) the uncertainty we hope to learn about, but cannot directly observe, which we call the distinction (or event) of interest, and (2) the test result, which we refer to as the observable distinction. The distinction of interest

might be the amount of oil in a particular reservoir and the observed distinction might be the interpretation of a seismic survey.

The VoI is defined as the most the decision maker should pay for additional information on the distinction of interest. If the decision maker is risk neutral, then

$$VoI = \begin{bmatrix} Expected\ value\ with \\ additional\ information \end{bmatrix} - \begin{bmatrix} Expected\ value\ without \\ additional\ information \end{bmatrix}$$

If the additional information is perfect (that is; the information provides perfect knowledge of the state of the world), then we refer to the VoI as the value of perfect information, which places an upper bound on any information-gathering activity. No test, no matter how sophisticated, is worth more than perfect information.

Although the notion of perfect information is hypothetical, calculating its value is a common and useful initial assessment of information value. Calculating the value of perfect information is relatively simple compared with calculating the value of imperfect information which requires the application of Bayes' theorem. Furthermore, if the value of perfect information is equal to or less than the cost of getting the information, there is no need to assess the value of imperfect information.

Any imperfect information analysis requires a Bayesian calculation. The technique is used in almost every VoI publication is to construct a discrete probability tree, assess the probabilities, and "flip the tree" to perform the Bayesian calculation. However, as will be demonstrated in this paper, this technique does not work well if one wants to model learning about continuous reservoir properties, such as porosity or in-place-volumes. For an introduction to Bayesian calculations, see Clemen and Reilly (2014) or Bratvold and Begg (2010).

In the context of decision making in the oil & gas industry, many uncertainties; e.g., in-place volumes, are derived and represented as continuos distributions. However, as decision trees, which typically are constructed with 2 or 3 possible outcomes (degrees) of relevant and material uncertainties, are commonly used to model decision contexts, there is a need to discretize continuous distributions, with an inifinite number of possible outcomes, in order to use them with decision trees. Discretization methods have been suggested to develop representations of contiuous uncertainties (probability density functions - PDFs) with a small number of discrete outcomes (probability mass functions – PMFs.) The focus of these methods is to approximate the continuous distribution with a discrete distribution as shown in Figure 1.

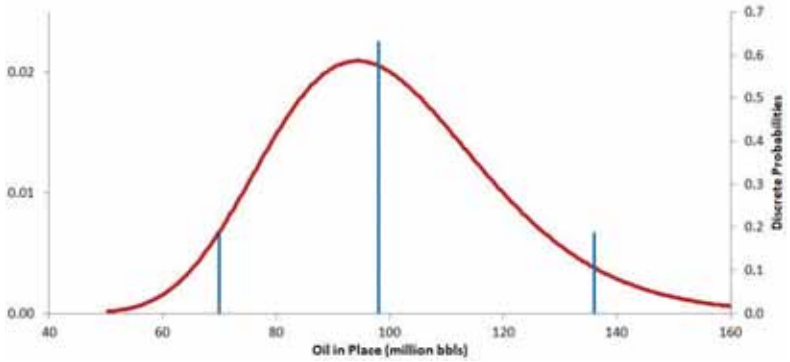


Figure 1 - Discrete PMF approximation (blue) to the continuous (red) PDF.

Miller and Rice (1983) suggests that the cireterin for the accuracy of the discrete approximation is that it preserves the relevant moments of the original distribution. In decision making, the mean is a common decision metric but, depending on the problem, we may also want the approximation to preserve the

standard deviation, the skewness, and kurtosis (and, possibly, more) moments. Furthermore, as the moments of a PDF do not uniquely determine the underlying PDF, a moment matching approximation may still fail to match particular percentiles or the extreme value of the PDF (Bickel, 2011). As we will see later in this paper, this lack of uniqueness may lead to very different VoI assessments when using the underlaying PDF versus a PMF that matches many of the moments of the PDF. Commonly used discretization methods include the Extended Swanson-Megill (ESM), Extended Pearson-Tukey (EPT), or the McNamee-Celona (MCS) approaches. In this paper, we refer to the discretization of a PDF into a small number (2 – 5) of possible outcomes as the Low Resolution Probability Tree (LRPT) approach. Bickel et al (2011) and Hammond and Bickel (2013) provide a rich disussion of discretization methods.

There are other approaches for calculating the VoI for continuous distributions that do not use a LRPT: the analytical methods (Bickel, 2012), and MC based methods (Chavez & Henrion, 2004; Arild, et al., 2008; Bickel, 2012; Barros, et al. 2015a, 2015b, 2016). However, the analytical approach is limited to a small subset of TALL-N (two-action linear loss with normal prior) problems whilst the papers applying MC approaches often requires a very large number of iteration to achieve a satisfactory accuracy and have very different representations of the reliability or quality of the information gathered. Despite the fact the LRPT approach can have significant errors and lead to different decisions, VoI calculations are rarely conducted using the full continuous representation of the uncertain event of interest. We suspect that the reason lies in the issue discussed here.

This paper introduces a general framework for calculating the VoI based on continuous distributions. VoIs calculated using this framework can be made arbitrarily close to the VoIs calculated with analytical solutions for TALL-N problems and are general in the sense that, for all practical purposes, they work with any continuous prior and likelihood as well as with any value function. We call this the High Resolution Probability Tree (HRPT) approach (we use “Probability Tree” rather than “Decision Tree” because the increase in resolution is confined to the uncertainty nodes in the tree). This paper also investigates the approximation error induced by the LRPT approach, and the MC based methods. Furthermore, the HRPT approach is more accurate (requires significantly fewer samples for the same accuracy level), and more straight-forward to implement than the MC based methods.

The are multiple contributions of this paper. We illustrate how to extend the common LRPT approach to an arbitrarily rich HRPT model for the purpose of accurately assessing the VoI based on continuous distributions for both the prior and the test result. We also investigate and discuss the accuracy of common discretization methods in the context of VoI calculations. This goes beyond Bickel (2012) in that we investigate the accuracy for more realistic problems, where the underlying uncertainty and the test results can be represented by any continuous probability density function and the value function is not constrained to be linear. Furthermore, we introduce and illustrate several practical approaches for quantifying continuous likelihood functions. The HRPT method is also compared with the Ensemble Kalman Filter (EnKF) and the Bayes Monte Carlo (BMC) approaches. This comparison gives insights regarding the robustness and accuracy of the MC based methods. We believe the results of our investigation of various VoI approaches and discretizations provide a basis for choosing an appropriate VoI approach for any decision context. Finally, by illustrating how continuous-variables techniques can be used in VoI analysis, our paper has some tutorial value that may help inspire and increase the applications of VoI analysis where the uncertainties are represented by continuous distribution.

Value-of-Information

A number of papers (e.g. Bratvold et al, 2009) and books (e.g. Clemen, 2013; Howard and Abbas, 2015; Bratvold and Begg, 2010) describe the common VoI decision tree structure and solution method. Figure 2 shows the decision tree structure used to value information using three discrete outcomes.

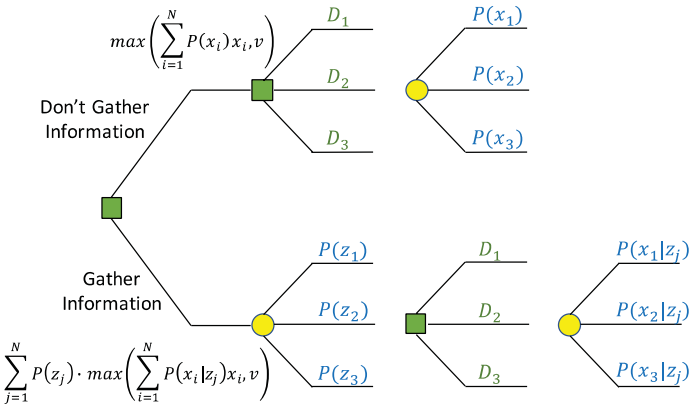


Figure 2 - Decision tree for valuing information for the case with three alternatives and three discrete outcomes.

The first step is to calculate $E[X|Z = z_j]$ for each possible outcome of the information source. With the $E[X|Z = z_j]$ for each z, we exercise our decision if $\mathrm{E}[X|Z = z_j] > v$. We then multiply these values by the corresponding preposterior probabilities to obtain the value with information. As the decision maker in this case is risk neutral, the VoI is given by

$$\boldsymbol{VoI} = \sum_{\boldsymbol{j=1}}^{\boldsymbol{N}} \boldsymbol{P(z_j)} \cdot \boldsymbol{max}\left(\sum_{\boldsymbol{i=1}}^{\boldsymbol{N}} \boldsymbol{P(x_i|z_j)x_i, v}\right) - \boldsymbol{max}\left(\sum_{\boldsymbol{i=1}}^{\boldsymbol{N}} \boldsymbol{P(x_i)\, x_i, v}\right) \tag{1.}$$

Analytical Method of Calulating VoI for Continuous Uncertain Quantities

Before introducing our HRPT method for analyzing VoI, we will discuss an analytical solution (Bickel, 2008) for a simple example to use as a truth case with which to test the accuracy of the HRPT approach. The example is kept simple to facilitate intuitive understanding of both VoI methods, but is still relatively realistic.

We will use the following notation:

- X - Random variable for the possible values of the underlying uncertainty of interest (e.g., OIP).
- x - A particular realization of X.
- μ - Mean of X.
- Z - A random variable that represents the possible outcomes of the observed signal provided by our test (e.g., the OIP indication resulting from drilling an appraisal well or running a pressure transient test).[1]
- z - A particular realization of Z.
- σ - Standard deviation of X.
- ρ - Correlation between X and Z.

It is convenient for comparison purposes to use the example presented by Bickel (2012). Suppose a risk-neutral oil company is considering drilling a well in an undeveloped area where the net present value (NPV) is normally distributed with a mean of \$10 million and standard deviation of \$20 million. If the company does not drill, it will earn a sure $v = \$0$. In an effort to improve its decision, the company is considering the acquisition of a seismic survey. The geophysicist in the company believe the seismic

[1] In common O&G context, this variable is based on expert interpretation of the measured data.

results are correlated with the true value of the well with correlation coefficient, $\rho = 0.6$. This information signal is represented by a standard normal distribution.

The problem described above is known as the two-action-linear-loss (TALL) problem in the decision analysis literature (Schlaifer, 1959). In the TALL problem, a risk-neutral decision maker chooses between two alternatives with uncertain outcomes, which are linear functions of some underlying uncertainty. If we also impose that the prior distribution must be normally distributed, a variant of the TALL problem with normal priors (TALL-N) is obtained. The example described above fits with the TALL-N definition and can thus be used to compare the various VoI approaches that involve continuous distributions.

Analytical Solution for the VoI of a TALL-N problem with ρ-Information System (ρ-IS)

Schlaifer (1959) derived a closed form solution for the expected value of imperfect information (EVII) of the TALL-N problem. Bickel (2008) reviewed and expanded this work to account for risk aversion and discussed (Bickel, 2012) the solution in an O&G context. As argued by Bickel (2008), the advantage of this relatively simple formulation is that the VoI can be calculated exactly. This allows us to compare any numerical approximations with the true value.

The closed form VoI solution for the TALL-N problem for a ρ Information System (ρ-IS), that is, one where the information content in the signal is expressed by the correlation coefficient between the prior and the signal, is:

$$\boldsymbol{EVII} = \begin{cases} \boldsymbol{\rho\sigma[\phi(\rho^{-1})c - \rho^{-1}c\Phi(-\rho^{-1}c)], \mu \geq v} \\ \boldsymbol{\rho\sigma[\phi(\rho^{-1})c + \rho^{-1}c\Phi(\rho^{-1}c)], \mu < v} \end{cases} \quad (2.)$$

where: μ is the mean of the prior; σ is the standard deviation of the prior; v is the value of alternative decision (the best decision without gathering more information); and $c = (\mu - v)/\sigma$ which is also known as "coefficient of divergence"; ϕ is the standard normal probability density function (PDF); Φ is the standard normal cumulative distribution function (CDF); and ρ denotes the positive correlation between the prior and the observed signal obtained from the information system.

We denote the VoI given by Eq. 1 as the Expected Value of Imperfect Informaton (EVII) as it is the expected value of the distribution that describes the uncertainty in the VoI.

In Figure 3, we show, using Eq. 1, the EVII as a function of the correlation coefficient, ρ using the example introduces earlier. As expected, the EVII increases with increasing correlation coefficient. It can be seen that a correlation coefficient of 1 yields an Expected Value of Perfect Information, EVPI, of \$3.95 million. If the correlation between prior and signal is 0.6, the EVII is \$1.36 million. If the correlation coefficient is 0, meaning the signal is independent of the prior, then, as should be expected, EVII = 0.

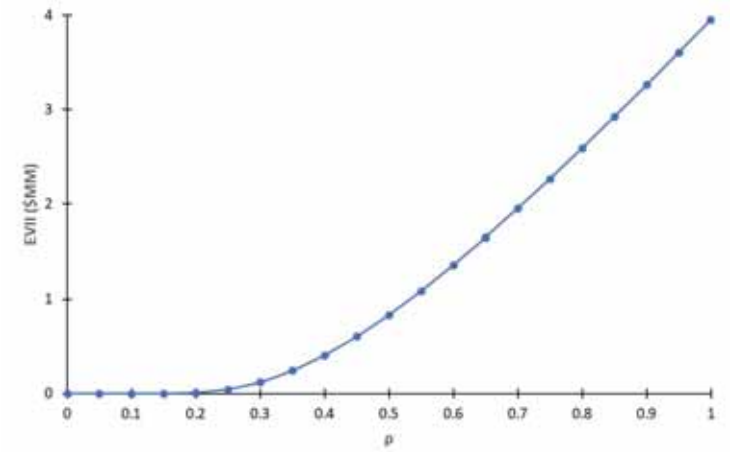


Figure 3 - EVII as a function of correlation between the prior and the observed signal.

Discretization shortcuts

Bickel (2012) also derived a closed form formula to calculate the EVII using 3-point discretizations

$$EVII = \begin{cases} p_1\sigma\left(-\rho\Phi^{-1}(\theta_1) - c\right), \frac{c}{\rho} < -\Phi^{-1}(\theta_1) \\ 0, \frac{c}{\rho} \geq -\Phi^{-1}(\theta_1) \end{cases} \quad (3.)$$

where θ_1 is the smallest fractile and p_1 the weight placed on that fractal. The above equation onlydepends upon $\Phi^{-1}(\theta_1)$ because $\Phi^{-1}(\theta_2) = 0$ and $\theta_1 = \theta_3$.

The discretization shortcuts do not require any calculation of percentiles or probabilities. These shortcuts were originally developed to approximate normal distributions. Bickel investigated the use of three common discretization shortcuts: the Extended Pearson-Tukey (EPT), McNamee-Celona (MCS), and Extended Swanson-Megill (ESM). Table 1 shows the percentiles and probabilities for these discretization shortcuts.

Table 1 - Discretization shortcuts

Discretization Shortcut	Probability	Percentile
ESM	0.3	P10
	0.4	P50
	0.3	P90
MCS	0.25	P10
	0.5	P50
	0.25	P90
EPT	0.185	P5
	0.63	P50
	0.185	P95

Using the exact and approximate VoI solutions in Eqs. 1 and 2, respectively, Bickel compared the VoI calculated using the most common discretization methods with the exact solution for the example introduced earlier. The results are shown in Figure 4.

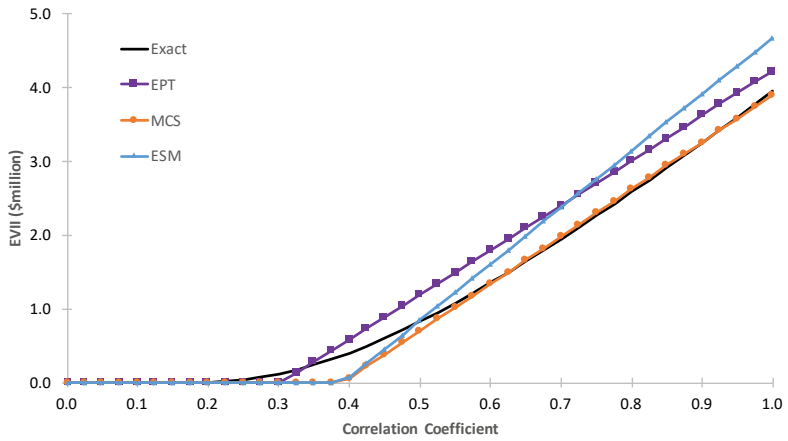


Figure 4 - Comparison of approximate EVII to exact EVII for commonly used discretization methods.

The approximations can give very different VoI estimates. These differences may or may not lead to different decisions regarding information gathering depending on the cost of the information. As Bickel (2012) explains, for example with a cost of $1.5 million, both EPT and ESM would recommend gathering the information, while MCS would not (consistently with the exact EVII).

Discussion

A key assumption built into the analytical approach is that the observed signal is jointly normally distributed with the prior with a positive linear correlation coefficient ρ. This differs from the more common decision tree approach in that the likelihood is not explicitly specified. Nor is there a need to calculate the preposterior or posterior. The correlation function embodies the amount of information contained in the information system, that is, the reliability of the information. If the reliability of the information-gathering experiment is small, the correlation coefficient will be close to zero, whereas the correlation coefficient will be close to 1 if the information content is large.

Although not required for calculating the VoI for a ρ-IS, Appendix A shows the analytical expression for the joint distribution $P(Z, X)$ (Bertsekas, 2002) – Eq. A1. We can use Eq. A1 or the linear relationships between the mean and variance of the likelihood and the prior shown in Eqs. A2 and A3 to calculate the likelihood and posterior. The key elements of the method are summarized below:

- Assumptions
  - The observed information signal is jointly normally distributed with the prior with the positive linear correlation coefficient ρ which can be quantified. Typically, the assessment of ρ will come from prior observations and knowledge, the same way the likelihood function is estimated.
  - Linear value function
- Pros
  - Analytical
  - Provides an intuitive means of specifying the relationship between the gathered information and the underlying uncertainty
- Cons
  - Limited to jointly normal distributions and linear value functions

- Limited to a single underlying uncertainty

High-Resolution ProbabilityTree Method (HRPT)

The analytical method has limitations in terms of the choice of prior and likelihood distributions and in the specification of the value function. In this section, we present a general method that can be used to calculate the VoI for any set of prior and likelihood pairs and any value function as well as any number of underlying uncertainties and decision alternatives. We use a simple[2] example to introduce and illustrate the method and compare the resulting VoI with the analytical solution.

As indicated by our choice of terminology, the HRPT method is simply a higher-level resolution (bigger N) version of the LRPT method, in which continuous distributions are approximated by a specified number of discrete outcomes. As well as being able to use any probability distribution for the prior or the likelihoods, we don't need to do any analytical integration to get the posterior or the VoI itself. The preposterior is a sum over many discrete outcomes instead of an integral.

Deriving the Signal or Likelihood Distributions

In using the previously discussed $\rho - IS$ approach, we worked with the prior and signal distributions and there was no need to specify the likelihood function as we could calculate the posterior distribution using the joint distribution (Eq. A1) and the signal (preposterior). As we will discuss below, when working with a more general approach to calculate the VoI based on continuous input distributions, we can choose whether to specify either the signal distribution or the likelihood function.

To refresh, a likelihood is a conditional probability that represents our degree of belief that the information (e.g. seismic survey or appraisal well) will indicate a specific signal, z_j, for a given prior outcome x_i, (e.g. HIP). That is, $P(Z = z_j|X = x_i)$. The collection of $P(Z = z_j|X = x_i)$ over all z_j is the likelihood function given x_i. Prior to gathering the information, a likelihood function is needed for each x_i. The main challenge in the HRPT approach is to specify an appropriate likelihood function (this is no less a challenge in the LRPT approach). However, once specified, if it is continuous, it is relatively straightforward to discretize it using one of the methods suggested below. The collection of all likelihood functions can be stored as an $N\ by\ N$[3] of conditional probabilities that make calculation of the posteriors, and thus the EVII, efficient. The likelihoods and priors are used, via Baye's theorem, to infer the conditional probabilities required in a VoI calculation, that is, probabilities of the form, $P(X = x_i|Z = z_j)$:

$$\boldsymbol{P(x_i|z_j) = \frac{P(z_j|x_i)P(x_i)}{\sum_i^N P(z_j|x_i)P(x_i)}} \tag{4.}$$

Specifying and discretizing the likelihood function becomes increasingly demanding the larger N is, particularly if we are working with 100 or 1,000 prior intervals and 100 or 1,000 "signal" intervals for each of the prior interval. Thus, we need an efficient mechanism to specify the likelihood whilst ensuring that it still is representative of the technology, test, or model being used to gather the data and interpret the information.

Discretization

Different discretization methods were briefly discussed in the introduction of this paper. In using the HRPT method, all distributions must be discretized and, in principle, any discretization method could be used. However, as the resolution increases and the number of degree gets large, it should not matter which

---

2 As several of the VoI methods discussed have severe constraints on both the probability distributions and value functions, any example on which we would want to apply the methods must fit with these constraints.

3 Or an $N_P x N_S$ table if we use N_P intervals for the prior and N_S intervals for the signal.

method is being used. In this work, we are employing a straightforward bracket-median method where the value axes are divided into N discrete intervals, or brackets, and each interval is being represented by a single value-probability pair $\{x_i, p_i\}$ where x_i is the median of the interval and p_i is the assigned from the underlying continuous distribution. We use the same number of intervals for the prior and the signal (or likelihood function). This is convenient but not required.

General Methods for Discretizing the Likelihood

In principle, the HRPT works with any likelihood function. However, in light of the discussion above, we have found it useful to work with three different methods for discretizing the likelihood function.

Likelihood derived through the correlation coefficient

As per the TALL-N problem, Bickel (2008) introduces the correlation coefficient as a means of specifying the relationship between the prior and the preposterior (observed signal). He was applying an information system whose data is jointly normally distributed with the underlying uncertainty with a positive correlation coefficient ρ. The specific value of ρ that is used could come from theoretical considerations, observed correlations or it may be intuitively assessable by a subject matter expert. Given the prior and the preposterior, we can calculate their joint probabilities and thence the posterior. This differs from the LRPT approach in that there is no need to specify the likelihood function explicitly. However, to enable comparison with the HRPT method it is useful to specify the implied likelihood matrix directly.

Bickel's (2008) approach is limited to jointly normally distributed functions. It assumes there is a linear relationship between the mean of the probability distribution of the observed signal, $E[Z]$, and the mean of the prior, $E[X]$. The linear relationship allows us to derive the analytical form of the mean and variance of the information, $E[Z]\ and\ Var[Z]$, as well as the joint distribution between the information and the prior, $P(Z, X)$. This, in turn, simplifies the calculation of the posterior.

Drawing on the approach presented by Bickel (2008), we can specify the information system for any prior distribution by simply specifying the relationship between the expected value of the likelihood function of the observed signal and expected values of the signal and the prior.

$$\boldsymbol{E(Z|X) = f\{E(Z), E(X)\}} \tag{5.}$$

For example, if we assume that the prior and the observed signal are given by the same distribution, we can approximate the linear relationship between the expected values of the prior and observed signal for any two distributions by using a positive correlation coefficient ρ. Drawing from the joint normal approach, we let the mean and variance of the likelihood be

$$\boldsymbol{E(Z|X = x_i) = E(Z) + \rho(x_i - E[X])} \tag{6.}$$

and

$$\boldsymbol{VAR[Z|X] = (1 - \rho^2)\sigma_X^2} \tag{7.}$$

respectively. Assuming a functional form, with the above mean and variance, we can now discretize it using any of the methods described previously.

In Figure 5 we present plots of the likelihood, joint and posterior probabilities between a normal prior and normal information system (each with mean=10, standard deviation=20), using the relationships presented in Eqs. 6 and 7. It can be seen that for $\rho = 0$ (no relationship between the prior and the observed signal) the signal carries no information that would make us change our beliefs, i.e. posterior = prior for

any value of Z, as it should. For $\rho \sim 1$ (observed signal represents perfect information), our beliefs should be updated to exactly what the signal tells us, i.e. posterior = likelihood, as it should. For $0 < \rho < 1$ the observed signal represents imperfect information and our beliefs should be updated accordingly, such as shown in the middle panels which are for $\rho = 0.6$.

General likelihood for any functional form

As mentioned above, the HRPT method is entirely general. That is, the subject matter expert (SME) or analyst can specify any $P(Z|X = x_i)$ at each x_i. Pragmatically, this can be achieved by specifying a functional form for the PDF (e.g. Normal, Lognormal, Triangular (Arild et al.,2008), Pert, etc.) where the parameters that define the PDF are constant, or could be a function of X if there is evidential or theoretical support for that. That is, the method permits the use of any function that "makes sense" to the SME in terms of defining the reliability of the data for a given x_i - like any other subjective probability, it describes the SME's degree of belief. Care must be taken in defining the function so that is does not imply probabilities for non-permissible values of the uncertain quantity, e.g. negative volumes.

We suggest that in many cases it makes sense to think of the likelihood at a given x_i. in terms of the reliability of the signal, with the parameters of the conditional PDFs independent of X. For perfect information, the diagonal must have values of 1, and all off-diagonal terms be 0. As the signal becomes less reliable the off-diagonal terms become non-zero and the diagonal terms become less than 1, with the constraint that the sum of the probabilities equals 1 at each x_i. Thinking in terms of reliability, it seems reasonable that (again, for a given x_i) the highest probability would be on the diagonal and so should be set to the mode of the, discretized, chosen distribution (expected value if symmetric).

If the parameters of the distribution are thought to be a function of X, then a specified functional relationship can be imposed on these parameters that define the conditional PDF. If the SME has evidence or argument that the relationship between the conditional PDF parameters and X does not follow a specific functional form, then the bounding envelop method (Vose, 2008) can be employed.

Validation of the HRPT approach

Given that the analytical solutions presented in the VoI literature, and summarized here, are constrained to a few special cases, we cannot provide a general VoI validation. Thus, we will validate the HRPT method by (1) comparing with the TALL-N analytical solution and (2) validating the Bayesian updating step; i.e., the calculation of the posterior for a lognormal - lognormal prior - likelihood pair where the result of the Bayesian updating can be compared with existing analytical solutions.

Eq. 1 gives $EVPI_{exact} = \$3.956$ million and $EVII_{exact} = \$1.3596$ million. Meanwhile using $N = 100$ in the HRPT approach, results in $EVPI_{100} = \$3.955$ million and $EVII_{100} = \$1.358$ million, so correct to 3 significant digits. Figure 6 compares the VoI results using the HRPT approach with the analytical solution in Eq. 1 for different values of ρ and N.

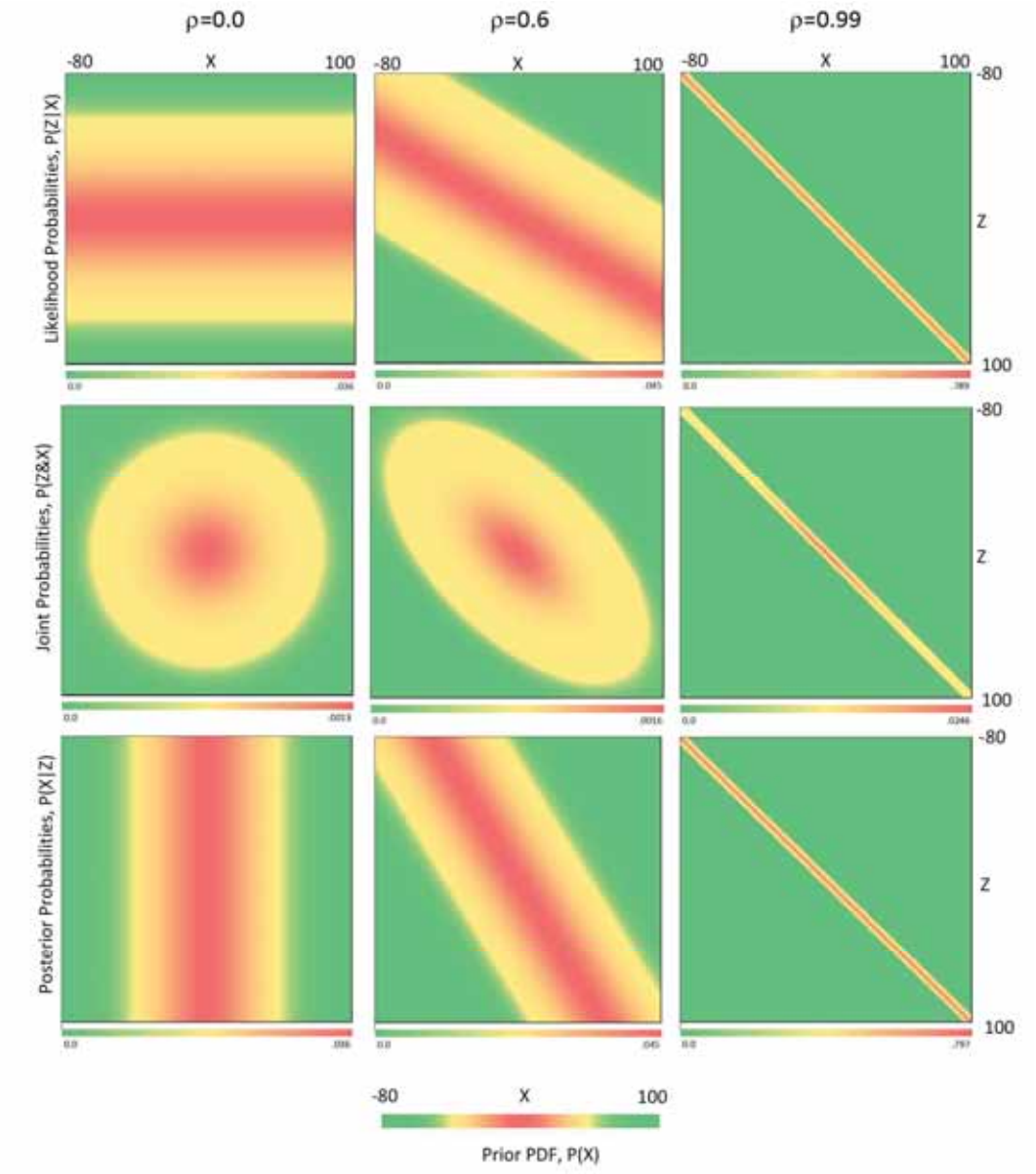


Figure 5 - Likelihood, Joint and Posterior probability density functions for different correlation coefficients between the uncertain quantity of interest, X, and the signal, Z.

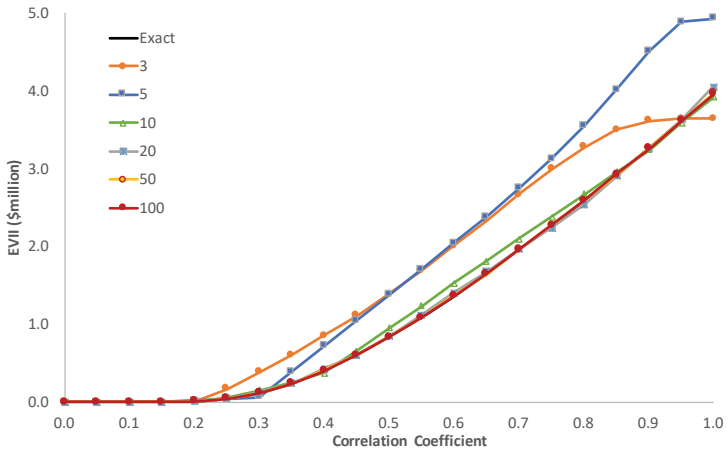


Figure 6 - Comparison of approximate EVII to exact EVII for different values of ρ and N.

From the graph we see that for $N < 20$ there can be significant errors in the EVII. Table 2 provides the detailed results.

Table 2 – Percent error when using the HRPT approach in a TALL-N, $\rho - IS$ problem for different ρ and N.

		N					
		3	5	10	20	50	100
ρ	0.3	220.79	-47.75	26.58	6.08	2.30	0.36
	0.6	47.22	49.79	12.08	2.94	1.25	0.38
	0.9	10.79	38.03	-0.62	0.08	0.71	0.07

For $N \leq 20$, there are significant errors in EVII. For $N = 50$ and $N = 100$, the percent errors are less than 2.5% and 0.5%, respectively. For all practical purposes, using $N \geq 50$ will ensure a good EVII estimate resulting in the same information gathering decision as the closed form solution (Eq. 1).

Recall that what has been calculated so far and shown in the above graphs is the value of free information. As what is important in an approximate VoI assessment is not the VoI itself but whether it will lead to the same information gathering decision as the exact solution. From Figure 6 we see that regardless of N, the value of free information becomes positive at $\rho \sim 15$.

In order to investigate the value of information that comes at a cost, we need to introduce a cost function. Let the cost of the information be a linear function of its quality so that $c = 0.2 + 3\rho$ in $million. We can then add the cost to the EVII for different N's as shown in Figure 7.

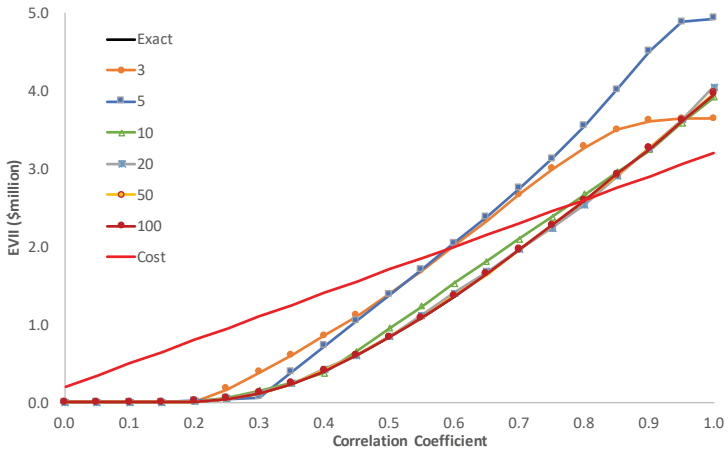


Figure 7 - Comparison of approximate EVII to exact EVII for different value of N when the information gathering cost is included.

From the graph, we see that irrespective of N, the information will not add value if ρ is less than about 0.6 and will add value if ρ is greater than about 0.8. If ρ is between 0.6 and 0.8, the commonly used 3 and 5 point discretization methods will result in different information gathering decisions than the HRPT assessment.

Non-Normal Prior – Likelihood Combinations

In this section we investigate the accuracy of the LRPT approach with different discretization methods for more general and commonly used continuous distributions and value functions. We draw on the results from the validation section and treat the VoI calculated using the HRPT approach with $N = 100$, $HRPT_{100}$, as the correct one.

Comparison Cases

We investigate the VoI accuracy using the PERT – PERT prior – likelihood pair and the lognormal – lognormal prior – likelihood pair.

PERT – PERT prior – likelihood pair

Suppose an oil company is facing a decision whether to invest or not in a project where the uncertainty in the prior NPV has been assessed as a PERT(-20,-10,50) distribution. This NPV distribution is representative of typical exploration projects where the upside is large but with a low probability and the downside is relatively small but with a high probability. The E[NPV] of the prior is -$1.67 million. If the company chooses not to invest, they will get $0 for sure. As a part of the analysis the company is considering gathering information to reduce their uncertainty. The company experts believe the reliability (likelihood function) of the information is normally distributed where the parameters are described using the conditional expectation method, as in Eq. 6. The cost of the information is a linear function of its quality and is $c = 0.5 + 3\rho$ in $million. Figure 8 shows the cost and the $HRPT_{100}$ as functions of ρ.

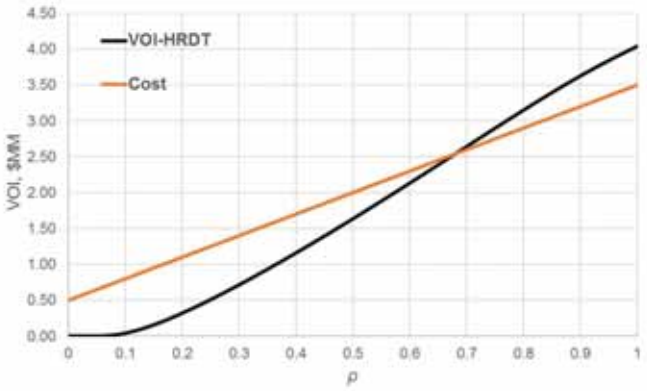


Figure 8 - VoI from HRPT vs cost of information gathering – PERT – PERT prior – likelihood pair.

Lognormal – lognormal prior – likelihood pair

In this case the distinction of interest is reserves. Let the prior assessment of reserves be lognormally distributed with a mean of 130 million and a standard deviation of 30 million; i.e., LN(130,30). The NPV is a linear function of reserves and is given by $NPV = -150 + X$, where X is the reserves and the NPV is in \$million. As before, the cost of information is a function of its reliability and is given by $c = 0.5 + 3\rho$ in \$million. The results are shown in Figure 9.

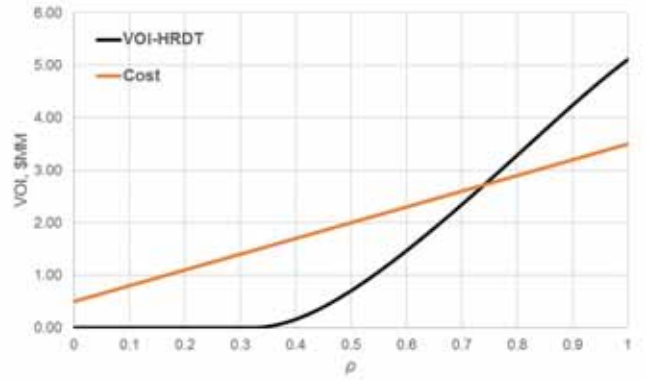


Figure 9 - VoI from HRPT vs cost of information gathering – lognormal – lognormal prior – likelihood pair.

General Discretization Methods

We now provide a brief overview of some more general different discretization methods. The interested reader should consult Bickel et al. (2011) for more details.

Bickel et al. (2011) present and describe three general discretization methods: Moment Matching, Bracket Mean and Bracket Median.

Moment Matching

In the moment matching (MM) method the probabilities and values for each of the degrees (discretization points) are calculated so that the moments[4] of the distributions are preserved. These probabilities and values can be calculated using Gaussian Quadrature (GQ) as well as by using an optimization formulation.

[4] The moments are the mean, variance, skewness, kurtosis, etc.

Using the MM method an n-degrees discretization match 2n-1 moments. We use the MM method with 3 and 5 degrees in our investigation.

Bracket Mean and Bracket median

The intuition behind the Bracket Mean (BMn) and Bracket Median (BMd) methods is quite straightforward. The CDF of the probability distributions is divided into N intervals. For each N intervals, we have the probability and value that representing those intervals. The probability at interval i is given by $CDF_i - CDF_{i-1}$, whilst the value is described with the mean (BMn) or median (BMd) of x_i to x_{i-1}. Table 3 shows the percentile and probability of the BMn and BMd for both PERT prior and lognormal prior case.

Table 3- PERT prior and lognormal prior discretization methods

Percentile and Probability used for PERT Prior Case						Percentile and Probability used for Lognormal Prior Case					
	Percentile	Probability		Percentile	Probability		Percentile	Probability		Percentile	Probability
BMn 3	0.1601	0.3330	BMn 5	0.0948	0.2000	BMn 3	0.1444	0.3330	BMn 5	0.0844	0.2000
	0.5044	0.3340		0.3001	0.2000		0.5028	0.3340		0.2985	0.2000
	0.8614	0.3330		0.5015	0.2000		0.8697	0.3330		0.5010	0.2000
				0.7035	0.2000					0.7037	0.2000
				0.9173	0.2000					0.9230	0.2000
BMd 3	0.1665	0.3330	BMd 5	0.1000	0.2000	BMd 3	0.1665	0.3330	BMd 5	0.1000	0.2000
	0.5000	0.3340		0.3000	0.2000		0.5000	0.3340		0.3000	0.2000
	0.8335	0.3330		0.5000	0.2000		0.8335	0.3330		0.5000	0.2000
				0.7000	0.2000					0.7000	0.2000
				0.9000	0.2000					0.9000	0.2000
MM 3	0.0500	0.1830	MM 5	0.0500	0.1415	MM 3	0.1172	0.3740	MM 5	0.0287	0.1121
	0.5000	0.6337		0.2500	0.2313		0.7155	0.5755		0.3582	0.5336
	0.9500	0.1834		0.5000	0.2547		0.9900	0.0505		0.8473	0.3230
				0.7500	0.2322					0.9921	0.0310
				0.9500	0.1403					1.0000	0.0003

Results

We use two measures to investigate the VoI accuracy achieved with LRPT discretization methods. The first measure, introduced by Bickel (2012), is the ratio of the approximate to the exact VoI: VOI_{LRDT}/VOI_{Exact}. As an example, a $VOI_{LRDT}/VOI_{Exact} = 1.2$ tells us that the VoI resulting from LRPT is overestimated by 20%.

The second measure we use is to capture the VoI decision impact; i.e., are there cases where the exact solution would have resulted in yes or no information gathering decision whilst the LRPT concludes that more information has no value or is valuable? For example, if the exact solution indicates that information is valuable for all $\rho \geq 0.75$ whilst the LRPT result in a positive information gathering decision for $\rho \geq 0.5$, we say that the LRPT solution is more "aggressive" in terms of information gathering activities compared with the exact solution. We investigate this measure by looking at the ratio ρ_{LRDT}/ρ_{HRDT}. If the ratio is greater (less) than one, the LRPT method is more (less) aggressive than the exact solution; i.e., using the LRPT with the given discretization method will result in a positive (negative) information gathering decision for lower (higher) values of ρ.

Figure 10 and Figure 11 show the VoI ratio for the five points general discretization methods using the PERT and lognormal prior, respectively. Figure 12 and Figure 13 show the VoI ratio for the three points general discretization methods using the PERT and lognormal prior, respectively. The black line represents the exact solution. Clearly, for this case none of the discretization methods investigated are very accurate across the entire range of ρ although the MM5 method is closer to the exact VoI for ρ > 0.4.

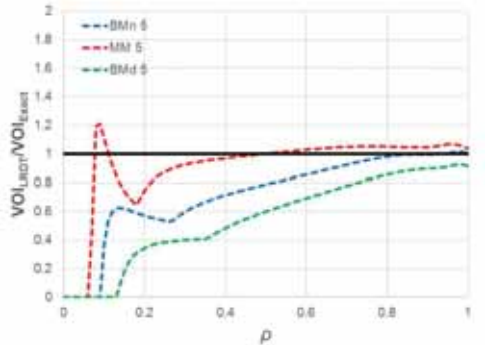


Figure 10 - VoI error for the general discretization methods with PERT Prior, 5 points discretization

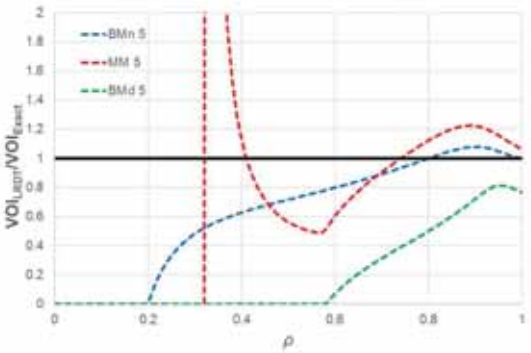


Figure 11 - VoI error for the general discretization methods with lognormal Prior, 5 points discretization

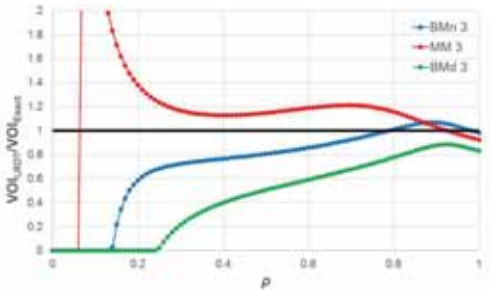


Figure 12 - VoI error for the general discretization methods with PERT Prior, 3 points discretization

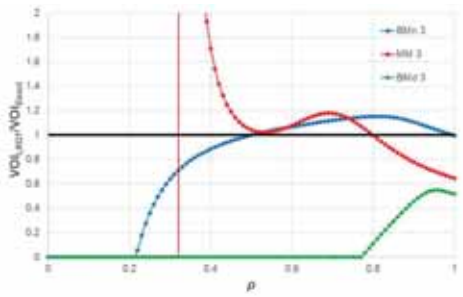


Figure 13 - VoI error for the general discretization methods with lognormal prior, 3 points discretization

Figure 14 and Figure 15 provide the same comparison for the three short cut discretization methods. The errors resulting from the use of short cuts compared with the exact solution with a PERT prior are relatively constant for $\rho > 0.3$. For both priors the EPT short-cut adds value at a lower ρ than the other two short-cuts.

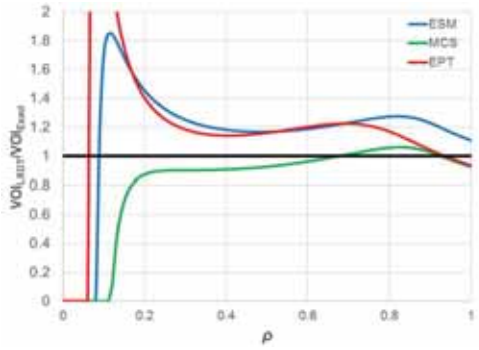


Figure 14 - VoI error for the discretization shortcuts with PERT prior

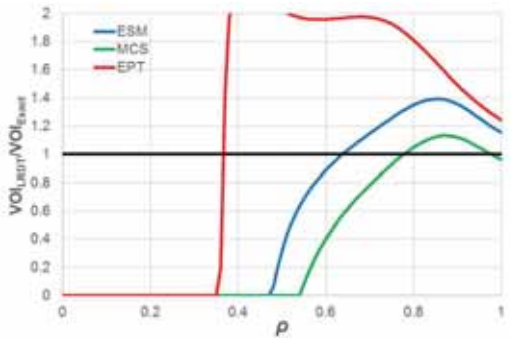


Figure 15 - VoI error for the discretization shortcuts with lognormal prior

Most of the three to five discretization methods above have VoI when $\rho > 0.6$, and thus we decided to see the minimum and maximum of the relative errors for this range[5]. It would be unreasonable to compare the range of error for each discretization methods at any value of ρ. Figure 16 and Figure 17 show the range of relative errors for $\rho \geq 0.6$. The bars show the range from the maximum relative underestimate to the maximum relative overestimate for each of the discretization methods investigated.[6] Again, we cannot draw any general conclusions except that none of the discretization methods are doing a particularly good job in estimating the correct VoI.

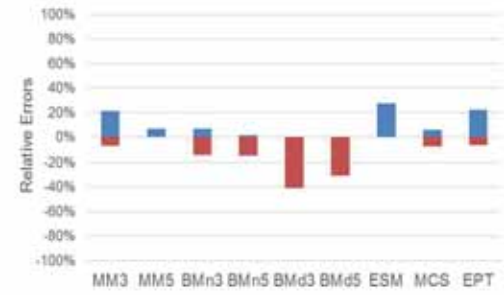


Figure 16 - Range of relative errors using the PERT prior

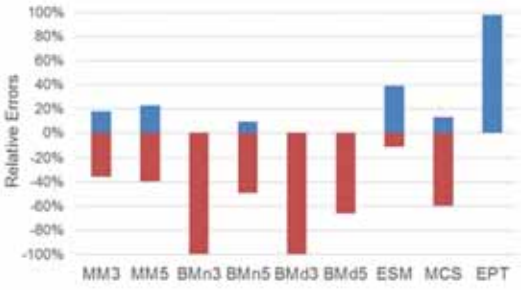


Figure 17 - Range of relative errors using the lognormal prior

In general, increasing the number of discretization does help to reduce the error. However, it is interesting to note that in MM-lognormal case, an increase in the number of degrees does not, lead to an increase in VoI accuracy. Indeed, for this case, matching 9 (2*5 – 1) moments of the prior uncertainty leads to a poorer estimate than matching only 5 (2*3 – 1) moments. We address this question in the discussion section below.

Nonetheless, the VoI value error is not important in and of itself. What we should be concerned about is its potential impact on the information gathering decision. Even though the LRPT VoI in many cases has a large error range as witnessed by Figure 16 and Figure 17, the impact of this error on the decision to gather information may be less pronounced.

Next, we look at the decision impact of using LRPT with the discretization methods. For a given cost

[5] Relative errors investigation when $\rho \geq 0.6$ is also more realistic, as there should be no reason for us to consider the information source if they have low accuracy (low ρ) settings.

[6] Relative errors is defined as $(VOI_{LRDT}/VOI_{exact} - 1) \times 100\%$

C ($million) of information, we can identify the minimum ρ such that VoI > C; i.e., the point where the decision changes from "do not gather information" to "gather information." In Figure 8 and Figure 9, we show the HRPT-VoI and cost, C, of the information in the same graph. Figure 18 shows the minimum $\rho-$ ratio (ρ_{LRDT}/ρ_{HRDT}) to get VoI > C for each of the discretization methods for the PERT prior.

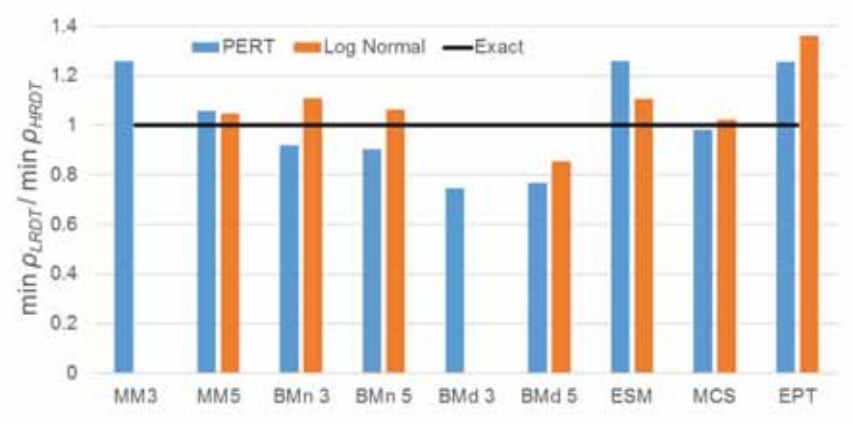


Figure 18 - Effect of VoI error on the information gathering decision

The black line corresponding to a ratio of 1 which corresponds to the exact solution. Whenever the $\rho-$ ratio is greater (less) than 1, the LRPT method will recommend a "gather information" decision at a lower (higher) ρ than the exact solution. If a method has no bar for a given prior, such as the MM 3 and BMd 3 for the lognormal distribution, not even perfect information, $\rho = 1$, is worth enough to invest in information gathering.

As before, the main observation is that, in general, the approximate methods used in LRPT will result in making the "gather information" decision with either less or more reliable information (lower or higher ρ) than what we get from the exact VoI computation. Furthermore, there is no fixed pattern in these errors. A given discretization method may not be consistently "early" or "late" in triggering the "gather information" decision for different priors.

Accuracy of Monte Carlo Methods

So far, we have been focusing on VoI analysis where decision tree is used as a main engine of the VoI analysis. Alternatively, few have introduced MC methods for VoI analysis. Arild et al. (2008) introduce VoI analysis where they using MC sampling for computing the Bayes' Rule. We called the method by Arild et al. (2008) as Bayes Monte Carlo (BMC). Barros et al. (2015a, 2015b, 2016) introduced MC based VoI analysis using EnKF algorithm. Later Hong et al. (2016) improve the formulation of Barros et al. (2015a) to make it more robust and consistent with VoI analysis in decision analysis.

These MC based methods do have its benefit. In oil & gas contexts, VoI analysis involving a large number of uncertainties is often required. Since the computanional costs of the HRPT approach are growing exponentially,[7] it is not suitable for solving problems with a large number of uncertainties. MC methods are less sensitive to the number or uncertainties involved. Nevertheless, there are always choice with MC methods that is not easy to answer. How many MC samples is required to get a good VoI approximation? Or in the case of EnKF, How many member of ensembles are required. This is the question we would like to address in this section.

This section investigate the accuracy of the BMC approach by Arild, et al. (2008) and EnKF approach,

[7] HRPT is growing exponentially at the rate of $O(n^k)$ where n is the size of the discretization points, and k represents the number of uncertainty sources.

using Formulation 2I of Hong, et al. (2016), which is an improvement of Barros et al. (2015a, 2015b, 2016) methods. Specifically, we will compare analytical methods solution to the results from HRPT, BMC, and EnKF, on the TALL-N problems. The results is then used to compare the effectiveness of MC based approach to the HRPT solution.

Bayes Monte Carlo (BMC). Arild et al. (2008) introduced a MC simulation based method for VoI calculations with continuous priors. They introduce an error function (EF) that represents the imperfect information embedded in the observed signal. The EF is then combined with the prior to generate the likelihood function which in turn is used to calculate the preposterior and posterior distributions. Introducing an error function is simply another means of specifying the functional relationship between the prior and the likelihood so, in that respect, this BMC approach is similar to the HRPT method. However, the BMC approach is using Monte Carlo sampling to "discretize" the distributions. Figure 19 illustrates the discretization of a probability distribution $P(X)$ where $N = 9$ as well as a set of 9 Monte Carlo samples. The continuous probability distribution is represented by the blue curve, the HRPT discretized outcomes by the orange dots and the Monte Carlo samples by the black rings. The shape of the continuous probability distribution is better preserved by the HRPT discretized outcomes than the Monte Carlo samples because the HRPT discretized outcomes distribute evenly throughout the range of X whilst most of the Monte Carlo samples cluster around the mode. If the number of degree for HRPT and the number of BMC samples are the same, the computational time for VoI estimation using HRPT and BMC method is the same. Thus, compared to the BMC method, HRPT can estimate VoI more accurately within a given computational time.

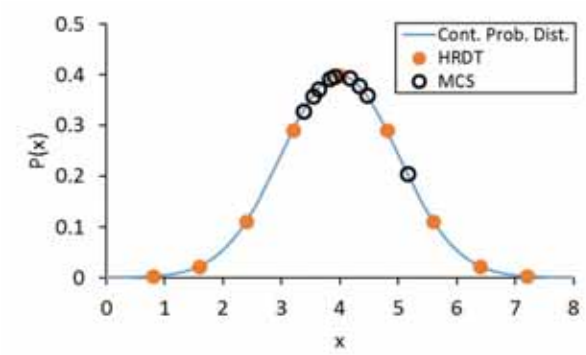


Figure 19 - Illustration of HRPT discretized outcomes and Monte Carlo samples for a continuous probability distribution.

The details of the algorithm to solve TALL-N problem using BMC is described in Appendix C.

Ensemble Kalman Filter (EnKF). The EnKF approach is based on a MC implementation of the Kalman filter (KF), which was pioneered by Evensen (1994) and discussed in details by Oliver et al. (2008). The key feature of KF is that it provides a closed-form solution for updating the probability, through a parameter called Kalman gain. Kalman gain is derived from the Bayes' rule (see Barker et al., 1995). Due to the Kalman gain, the probability updating in EnKF is very efficient and suitable to be used for large problems.

The EnKF is commonly implemented in a dynamical system where the parameters are time dependent. As in the original Kalman filter approach, the EnKF is implemented through the use of two equations: the prediction equation and the observation equation. The observation equation needs measurement error described as a normal distribution. This means, the measurement error in EnKF is the same with EF in the BMC, but is restricted to normal distribution. Since we are dealing with TALL-N problem where

everything is described to be normally distributed, then EnKF can be used to calculate the VoI. The algorithm that provides the detail on calculating the TALL-N with EnKF is provided in Appendix D.

Comparison to HRPT and Analytical VoI. The VoI calculated using EnKF with an ensemble size of 10,000 for imperfect information, say $\sigma_{z|x} = 5$, is 2.246 which is very close to the exact solution (2.267) with an error of only 0.52%. Meanwhile, for the same case, BMC with 10,000 samples resulting in 2.22, an error of 2.1% error.

In applying BMC and EnKF to VoI calculations for the normal distribution we can compare the number of ensemble members (MC samples) with the number of degrees in the HRPT approach and ask: How many ensemble members or number of samples are required to match the analytical VoI and compare this with the number of degrees required by the HRPT approach for the same purpose. The same thing also applies to BMC approach, i.e. How many MC samples are required to match the analytical VoI. We show this comparison in Figure 20 below, where the relative error from the exact VoI are calculated for each of the VoI methods.

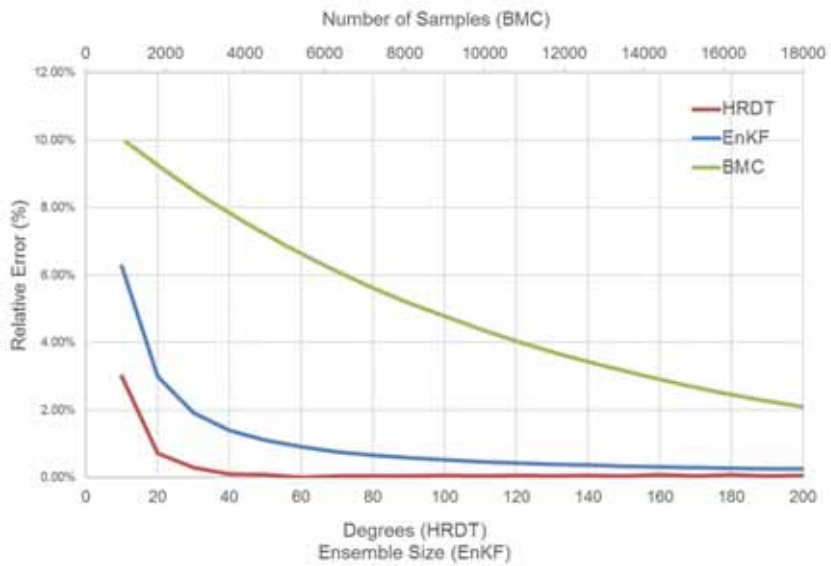


Figure 20 - VoI Mean Squared Error of VoI for HRPT and EnKF

As we can see from Figure 20, the HRPT is far more efficient than EnKF as it requires only higher than 20 degrees to get the same answer as the closed form solution within 1% range of error. However, using the EnKF requires higher than 70 ensemble members to get the same precision. And as for BMC approach, the MC samples required are extensive. Even with more than 10,000 samples, the relative error are more than 4% and still worse than the HRPT with 10 degrees, and worse than EnKF with ensemble size of 20. This explained that for VoI analysis with small number of uncertainties, the HRPT approach is preferred. However, as the number of uncertainties increases, so does the computational costs which exhibits linear growth and the EnKF will be more efficient.

Discussion

The focus of this paper has been on supporting information gathering decisions by using the VoI concept in decision contexts where the underlying uncertainties are represented by continuous distributions.

All of the VoI calculations discussed here, have inaccuracies, limitations or constraints. However, the

HRPT method provides complete flexibility in the choice of prior and likelihood function as well as in the choice of value function. It also quickly converges to the exact analytical VoI for the TALL-N problem. Thus, the HRPT approach provides an accurate, easy to implements, and very flexible approach to calculating VoI values based on continuouos distributions for priors and likelihood functions.

If the problem being evaluated fits with the constraints of the $\rho - IS$, the closed form solution is an obvious choice due to its accuracy and simplicity. Similarly, if the VoI calculation is concerned with a large number of uncertainties and fits within the constraints of the EnKF approach, this is a natural and robust approach for calculating the VoI. However, if the priors and likelihoods are not normal, and the decision involves a large number of uncertainties, the only choice is the BMC approach despite the fact that it has high error rates.

Neither the general nor the shortcut based discretization methods will result in VoI values that match the decisions or values of the exact VoI calculations based on continuouos distributions. The fact that the moment matching discretization methods are superior in matching the statistics of the prior does not necessarily mean that they will be more precise is valuing information or in triggering information gathering decisions. Indeed, in several cases the discretization shortcuts result in smaller VoI error than the moment matching methods.

The shape and properties of the distribution of the prior, particularly in $x < v$ region, has a large impact on the VoI. Figure 21 provides an example of a lognormal value distribution. The Value of Perfect Information is defined as $VOPI = \int_{-\infty}^{v} xp(x)dx$ which is the integration across $x < v$ range (red area in Figure 21). What would be required for a highly accurate VoI calculation using discretization methods is to preserve the shape and properties of the $x < v$ range. None of the discretization methods in common use have been developed with the intent of preserving the shape and properties of this area and so it should be no surprise that using these discretization does not lead to a good estimate of VoI value or choice.

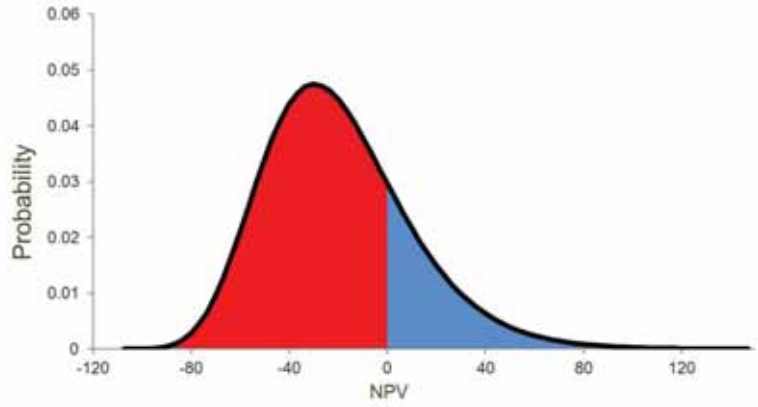


Figure 21 - Linear loss integral of NPV distribution for lognormal case.

Conclusions

The contribution of this paper are been five-fold:

- We have introduced a practical, flexible, efficient, and very general approach for analyzing the VoI when the underlying uncertainties are represented by continuous variables.
- We have investigated and discussed the accuracy of common and general discretization methods in the context of VoI calculations. From the discretization methods that we investigated, we found no common or general discretization methods that is accurate in calculating the VoI. More importantly, the discretization methods will often lead to different information gathering decisions

compared with those relusting from a more accurate VoI assessment. We cannot conclude that any of the discretization methods are better or worse than any other.

- We have also investigated the accuracy of MC based VoI methods. This investigation gives insight on the accuracy level for each MC based VoI methods.
- Given the availability of the $\rho - IS$, EnKF, BMC, and HRPT approaches, thorough investigation from this paper provide basis of recommendations for choosing the appropriate VoI approach.
- By illustrating how continuous-variables techniques can be used in VoI analysis, our paper also has some tutorial value that may help inspire and increase the use of these methods.

Acknoledgement
Aojie Hong would like to thank the National IOR Centre of Norway for supporting his research.

References

Arild, O.,Lohne, H. P. and Bratvold, R. B. 2008. A Monte Carlo Approach to Value of Information Evaluation. Presented at International Petroleum Technology Conference, Kuala Lumpur, Malaysia, 3-5 December. http://dx.doi.org/10.2523/IPTC-11969-MS.

Barker, A. L.,Brown, D. E. and Martin, W. N. 1995. Bayesian estimation and the Kalman filter. Computers & Mathematics with Applications. **30** (10): 55-77. http://www.sciencedirect.com/science/article/pii/089812219500156S.

Barros, E. G. D.,Jansen, J. D. and Van den Hof, P. M. J. 2015. Value of Information in Parameter Identification and Optimization of Hydrocarbon Reservoirs. IFAC-PapersOnLine. **48** (6): 229-235.

Barros, E. G. D., Leeuwenburgh, O., Van den Hof, P. M. J. et al. Value of Multiple Production Measurements and Water Front Tracking in Closed-Loop Reservoir Management. Society of Petroleum Engineers.

Barros, E. G. D.,Van den Hof, P. M. J. and Jansen, J. D. 2016. Value of Information in Closed-Loop Reservoir Management. Computational Geosciences. **20** (3): 737–749.

Begg, S.,Bratvold, R. and Campbell, J. 2002. The Value of Flexibility in Managing Uncertainty in Oil and Gas Investments. Presented at SPE Annual Technical Conference and Exhibition, San Antonio, Texas, 29 September-2 October. http://dx.doi.org/10.2118/77586-MS.

Bertsekas, D. P. and Tsitsiklis, J. N. 2002. Introduction to probability. Belmont, Mass., Athena Scientific.

Bickel, J. E. 2008. The Relationship Between Perfect and Imperfect Information in a Two-Action Risk-Sensitive Problem. Decision Analysis. **5** (3): 116-128. http://pubsonline.informs.org/doi/abs/10.1287/deca.1080.0118.

Bickel, J. E. 2012. Discretization, Simulation, and the Value of Information. SPE Economics & Management. **4** (04): 198 - 203.

Bickel, J. E. 2014. Decision Analytics, Value of Information, and Piloting in Unconventional Reservoirs. Presented at SPE Hydrocarbon Economics and Evaluation Symposium, Houston, Texas, 19-20 May. http://dx.doi.org/10.2118/169841-MS.

Bickel, J. E., Lake, L. W. and Lehman, J. 2011. Discretization, Simulation, and Swanson's (Inaccurate) Mean. SPE Economics & Management. **3** (03).

Bratvold, R. B. and Begg, S. 2010. Making good decisions. Richardson, TX, Society of Petroleum Engineers.

Bratvold, R. B.,Bickel, J. E. and Lohne, H. P. 2009. Value of Information in the Oil and Gas Industry: Past, Present, and Future. SPE Reservoir Evaluation & Engineering. **12** (04): 630-638.

Burgers, G.,Jan van Leeuwen, P. and Evensen, G. 1998. Analysis Scheme in the Ensemble Kalman Filter. Monthly weather review. **126** (6): 1719-1724.

Chavez, T. and Henrion, M. 1994. Efficient estimation of the value of information in Monte Carlo models. Presented at Proceedings of the Tenth international conference on Uncertainty in artificial intelligence, Seattle, WA.

Clemen, R. T. and Reilly, T. 2014. Making hard decisions with decision tools (3rd Edition). Mason, Oklahoma, Cengage Learning.

Evensen, G. 1994. Sequential Data Assimilation with a Nonlinear Quasi-Geostrophic Model Using Monte Carlo Methods to Forecast Error Statistics. Journal of Geophysical Research: Oceans. **99** (C5): 10143-10162.

Fink, D. 1997. A compendium of conjugate priors. 46.

Gelb, A. and The Analytic Sciences Corporation, T. S. 1974. Applied optimal estimation. Cambridge, MA, The M.I.T. Press.

Robert K. Hammond, J. Eric Bickel, (2013) Reexamining Discrete Approximations to Continuous Distributions. Decision Analysis 10(1):6-25. https://doi.org/10.1287/deca.1120.0260

Hong, A., Bratvold, R., Thomas, P. et al. 2016. Value-of-Information for Model Parameter Updating through History Matching. Working Paper-University of Stavanger.

McNamee, P. and Celona, J. 1987. Decision analysis for the professional—with supertree. Redwood City, California, Scientific Press.

Newendorp, P. D. and Schuyler, J. R. 2000. Decision analysis for petroleum exploration. Aurora, Colorado, Planning Press.

Pratt, J. W.,Raiffa, H. and Schlaifer, R. 1995. Introduction to statistical decision theory. http://search.ebscohost.com/login.aspx?direct=true&scope=site&db=nlebk&db=nlabk&AN=11378.

Raiffa, H. and Schlaifer, R. 1961. Applied statistical decision theory. Boston, Division of Research, Graduate School of Business Adminitration, Harvard University.

Schlaifer, R. 1959. Probability and statistics for business decisions; an introduction to managerial economics under uncertainty. New York, McGraw-Hill.

Stengel, R. F. 1994. Optimal control and estimation. New York, Dover Publications.
Vose, D. 2008. Risk analysis : a quantitative guide. Chichester [u.a., Wiley.
Winkler, R. L. 2010. An introduction to Bayesian inference and decision. Sugar Land, TX, Probabilistic Publishing.

Appendix A—Equations for Joint Normal Distributions

$$f(x,z)=\frac{\exp\left\{-\frac{1}{2(1-\rho^2)}\left[\left(\frac{x-\mu_x}{\sigma_x}\right)^2-2\rho\left(\frac{x-\mu_x}{\sigma_x}\right)\left(\frac{z-\mu_z}{\sigma_z}\right)+\left(\frac{z-\mu_z}{\sigma_z}\right)^2\right]\right\}}{2\pi\sigma_x\sigma_z\sqrt{1-\rho^2}} \quad \text{(A1)}$$

$$E[Z \mid X=x_i]=E[Z]+\rho\frac{\sigma_Z}{\sigma_X}(x_i-E[X]) \quad \text{(A2)}$$

$$VAR[Z \mid X]=(1-\rho^2)\sigma_X^2 \quad \text{(A3)}$$

Appendix B—Conjugate Prior for Lognormal Prior and Lognormal Likelihood

If X is lognormally distributed prior with $LN(m,v)$, then $ln(x)$ is normally distributed with:

$$N\left(\mu=\frac{m^2}{\sqrt{v+m^2}},\sigma=\sqrt{\ln\left(\frac{v}{m^2+1}\right)}\right) \quad \text{(B1)}$$

If we do one time measurement, and observe $ln(z)$ where the measurement variance is known and equal to $\sigma^2_{\ln(z)}$, then by the conjugate prior relationship, the posterior is determined as the following.

$$N\left(\frac{\frac{\mu}{\sigma^2}+\frac{\ln(z)}{\sigma^2_{\ln(z)}}}{\frac{1}{\sigma^2}+\frac{1}{\sigma^2_{\ln(z)}}},\frac{1}{\sigma^2}+\frac{1}{\sigma^2_{\ln(z)}}\right) \quad \text{(B2)}$$

Appendix C— Bayes Monte Carlo Approach for TALL-N VoI

As the name implies, BMC use a MC implementation of the Bayes' Rule. Intuitively, BMC is a Monte Carlo implementation of HRPT. HRPT sampling the probability distribution into fine grids of outcome, and in the case of BMC, the sampling is done via MC methods.:

The steps of using BMC calculate the VoI is as follows:

(1) Draw N samples from the prior distribution, such that we get a set of prior samples $[x^1 \quad x^2 \quad \dots \quad x^N]$ where x^i is a sample drawn from the prior $p(x)$.
(2) For each x^i, we draw one sample from the error function distribution and add them to x^i. This results in N samples of the measurement z, $[z^1=x^1+\varepsilon^1 \quad z^2=x^2+\varepsilon^2 \quad \dots \quad z^N=x^N+\varepsilon^N]$. These samples represent the preposterior distribution.
(3) We then calculate the likelihood, $z^i|x$ using preposterior distribution that we obtain from step 2. For example, the likelihood $z^1|x = [x^1-z^1, x^2-z^1, \dots x^1-z^1]$.
(4) Convert everything into probability density, to be used into the Bayes' Rule.

(5) Then we calculate the posterior density using Bayes' Rule, for each z^i. $f(x|z^i) \propto f(x)f(z^i|x)$
(6) For each z^i, we then calculate the value of the optimal alternative with the following arguments: $v_{opt}^i = max(E[x|z^i], v)$. In our case, $v = 0$
(7) VoI is then equal to: $VOI = v_{opt}^i/N - E[x]$

We are using the latin-hypercube sampling as our sampling algorithm to create the comparison to analytical solution and HRPT. Interested readers who wish to see general implementation of BMC approach should consult Arild, et al. (2008).

Appendix D— Ensemble Kalman Filter Approach for TALL-N VoI

Ensemble Kalman Filter (EnKF). The purpose of discussing the EnKF in this paper is to demonstrate the connection with VoI analysis and to discuss the accuracy we can expect from the EnKF model. Readers who are interested in application of EnKF for multivariate VoI analysis should consult Hong et al. (2016).

The EnKF approach is based on a recursive Monte Carlo implementation of the Kalman filter (KF) which is a solution to estimate the behavior of a process given observations. In O&G, the EnKF was pioneered by Evensen (1994) and discussed in details by Oliver et al. (2008). The EnKF is commonly implemented in a dynamical system. As in the original Kalman filter approach, the EnKF is implemented using two equations: the prediction equation (Eq. D1), and the observation equation (Eq. D2).

$$\mathbf{x}_k = \mathbf{f}(\mathbf{x}_{k-1}, \mathbf{m}) + \mathbf{w}_{k-1} \tag{D1}$$

$$\mathbf{z}_k = \mathbf{g}(\mathbf{x}_k, \mathbf{m}) + \mathbf{v}_k \tag{D2}$$

The definition of these equations is as follow: Let $\mathbf{x_k}$ be the value that we want to predict at time k, and $\mathbf{x_k}$ is then predicted through a function f which describes the relationship among $\mathbf{x_k}$, the value from previous predictions $\mathbf{x_{k-1}}$ and model parameters m, while the model error is described by normal distribution $\mathbf{w_{k-1}}$. The observation equation then links the observation $\mathbf{z_k}$ to our prediction $\mathbf{x_k}$ using a function g where the observation noise $\mathbf{v_k}$ is normally distributed. The equations are expressed in the state-space form. Gelb (1974) and Stengel (1994) provide a detail discussion in state-space modelling for engineering model.

The analysis scheme of EnKF describes how the prior is updated given an observation and is formulated as (Burgers et al. 1998)

$$\mathbf{Y}_a = \mathbf{Y}_f + \mathbf{K}(\mathbf{z}_p - \mathbf{H}\mathbf{Y}_f) \tag{D3}$$

where $\mathbf{Y}_f$ is the forecasted **Y**-matrix whose entries are based on the prior, $\mathbf{Y}_a$ is the analyzed or updated **Y**-matrix with updated information given an observation, **K** is the Kalman gain which acts as a weighing factor, $\mathbf{z}_p$ is a matrix containing the perturbed observations[8], and **H** is a matrix with entries of either 0 or 1 such that $\mathbf{HY} = [\mathbf{z}^1 \quad \mathbf{z}^2 \quad \dots \quad \mathbf{z}^N]$. The **Y**-matrix is constructed as

$$\mathbf{Y} = \begin{bmatrix} \mathbf{m}^1 & \mathbf{m}^2 & \cdots & \mathbf{m}^N \\ \mathbf{x}^1 & \mathbf{x}^2 & \cdots & \mathbf{x}^N \\ \mathbf{g}(\mathbf{x}^1, \mathbf{m}^1) & \mathbf{g}(\mathbf{x}^2, \mathbf{m}^2) & \cdots & \mathbf{g}(\mathbf{x}^N, \mathbf{m}^N) \end{bmatrix}$$

[8] For the practical issue of using the EnKF, an observation has to be perturbed with its corresponding statistics in order to avoid too low variance (Burgers et al. 1998).

where the superscript is the sample index and N is the total number of samples. The Kalman gain K is calculated as

$$\mathsf{K} = \mathbf{P}_\mathrm{f}\mathbf{H}^\mathrm{T}\left(\mathbf{H}\mathbf{P}_\mathrm{f}\mathbf{H}^\mathrm{T} + \mathbf{C}_\mathrm{y}\right)^{-1} \tag{D4}$$

where $\mathbf{P}_\mathrm{f}$ is the covariance matrix of $\mathbf{Y}_\mathrm{f}$, and $\mathbf{C}_\mathrm{y}$ is the covariance matrix of the observations.

EnKF formulation to Analytical VoI for TALL-N

The following formulation described the workflow to obtain the EnKF solution for the TALL-N problem presented and discussed by Bickel (2008). We drop the vector form because we are only concerned with a single source of uncertainty. For TALL-N problem, there is no temporal update of x, and therefore we don't need to predict the next estimate using Eq. D1. Our observation z for the TALL-N problem is equal to the value of interest (underlying uncertainty) plus a normally distributed error. Therefore, we can write the observation equation for the TALL-N problem as follows

$$\mathsf{z} = \mathsf{x} + \varepsilon \tag{D5}$$

Let the observation noise ε be normally distributed with a zero mean and a standard deviation σ_ε, i.e. $p(\varepsilon) \sim N(0, \sigma_\varepsilon)$. Note that this is the likelihood function expressed as an error function; i.e., the same as the one used in the MCVoI approach. That is, if we assign a distribution over the measurement noise, $p(\varepsilon) \sim N(0, \sigma_{z|x})$, we can describe the likelihood of an observation given the real values as $p(z|x) \sim N(x, \sigma_{z|x})$.

The step of using EnKF to calculate the VoI is as follows:

(8) N samples are drawn from the prior distribution, so we get a prior ensemble $[x^1 \quad x^2 \quad \dots \quad x^N]$ where x^i is a sample drawn from the prior $p(x)$.

(9) For each x^i, N samples are drawn from the error function and add them to x^i. This results in N samples of z, $[z^{i1} = x^i + \varepsilon^1 \quad z^{i2} = x^i + \varepsilon^2 \quad \dots \quad z^{iN} = x^N + \varepsilon^N]$. In total, we have N^2 samples of z. These samples represent the preposterior distribution.

(10) Given each z^{ij}, we calculate the posterior ensemble, $[\hat{x}^1|z^{ij} \quad \hat{x}^2|z^{ij} \quad \dots \quad \hat{x}^N|z^{ij}]$, using EnKF.

(11) The VoI is calculated as

$$\mathsf{VOI} = \frac{1}{\mathsf{N}^2}\sum_{\mathsf{i}=1}^{\mathsf{N}}\sum_{\mathsf{j}=1}^{\mathsf{N}} \max\left[\mathsf{v}, \frac{1}{\mathsf{N}}\sum_{\mathsf{k}=1}^{\mathsf{N}}\left(\hat{\mathsf{x}}^{\mathsf{k}} \mid \mathsf{z}^{\mathsf{ij}}\right)\right] - \max\left[\mathsf{v}, \frac{1}{\mathsf{N}}\sum_{\mathsf{k}=1}^{\mathsf{N}}\left(\mathsf{x}^{\mathsf{i}}\right)\right] \tag{D6}$$

where v is a certain threshold, the same as the one used throughout the examples, $v = 0$.

For our specific example, we construct $\mathbf{Y}_\mathrm{f}$ as $\mathbf{Y}_\mathrm{f} = \mathbf{X}_\mathrm{f} = [x^1 \quad x^2 \quad \dots \quad x^N]$. Then, the Kalman gain is

$$\mathsf{K} = \frac{\sigma^2_{\mathrm{x,N}}}{\sigma^2_{\mathrm{x,N}} + \sigma^2_{\mathrm{z_p,N}}} \tag{D7}$$

where $\sigma^2_{x,N}$ is the variance of $\mathbf{X}_\mathrm{f}$ and $\sigma^2_{\mathbf{z}_\mathrm{p},N}$ is the variance of $\mathbf{z}_\mathrm{p}$. Given an observation z, the $\mathbf{z}_\mathrm{p}$ is constructed as $[z + \varepsilon^1 \quad z + \varepsilon^2 \quad \dots \quad z + \varepsilon^N]$ where ε^i is a sample drawn from the error function $p(\varepsilon)$. The updated $\mathbf{X}$, i.e. $\mathbf{Y}_\mathrm{a} = \mathbf{X}_\mathrm{a} = [\hat{x}^1|z \quad \hat{x}^2|z \quad \dots \quad \hat{x}^N|z]$, is calculation using the analytical scheme of EnKF (Eq. D3).

Now, let our value of interest $p(x)\sim N(\bar{x},\sigma_x)$ and the error function or measurement noise $p(\varepsilon)\sim N(0,\sigma_{z|x})$. When the number of samples approaches infinity, the sample means $E[\mathbf{X}_\mathrm{f}]$ and $E[\mathbf{z}_\mathrm{p}]$ approach $\bar{x}$ and z, respectively; and the sample variances $\sigma^2_{x,N}$ and $\sigma^2_{\mathbf{z}_\mathrm{p},N}$ approach σ^2_x and $\sigma^2_{z|x}$, respectively. Therefore, we have $\mathbf{X}_\mathrm{a}$ being normally distributed with

$$E[\mathbf{X}_\mathrm{a}] = \mathrm{E}[\mathrm{x}\,|\,\mathrm{z}] = \bar{\mathrm{x}} + \left(\mathrm{z} - \bar{\mathrm{x}}\right)\frac{\sigma_\mathrm{x}^2}{\sigma_\mathrm{x}^2 + \sigma_{(\mathrm{z}|\mathrm{x})}^2}$$

$$Var[\mathbf{X}_\mathrm{a}] = \mathrm{Var}[\mathrm{x}\,|\,\mathrm{z}] = \sigma_\mathrm{x}^2\left(1 - \frac{\sigma_\mathrm{x}^2}{\sigma_\mathrm{x}^2 + \sigma_{(\mathrm{z}|\mathrm{x})}^2}\right) = 1\Big/\frac{1}{\sigma_\mathrm{x}^2} + \frac{1}{\sigma_{(\mathrm{z}|\mathrm{x})}^2}$$

as the number of samples approaches infinity. We see that the equations above is identical to the result obtained from the conjugate Bayesian analysis for normal distributions. The Kalman gain is essentially a correction factor between prior and observation and, thus, we can think of the Kalman gain as the Bayesian factor in the conjugate prior formulation.

The EnKF is a means to conduct the Bayesian inference using Monte Carlo methods. In EnKF, the prior is represented by the realizations stored in $\mathbf{Y}_\mathrm{f}$ and the posterior is represented by the updated realizations (i.e. realizations stored in $\mathbf{Y}_\mathrm{a}$). Although the EnKF (or KF in general) was initially designed for dynamic systems, it is also applicable for static systems. The advantage of using EnKF is that a large number of uncertain quantities can be updated efficiently. What makes the EnKF approach different from the BMC is that for a given observation, EnKF updates the values instead of updating the probabilities.

Paper V

Value-of-Information for Model Parameter Updating through History Matching

Hong, A.J., Bratvold, R.B., Thomas, P., and Hanea, R.G.

Manuscript

Value-of-Information for Model Parameter Updating through History Matching

A.J. Hong and **R.B. Bratvold**, University of Stavanger and The National IOR Centre of Norway; **P. Thomas**, University of Stavanger; and **R.G. Hanea**, University of Stavanger and Statoil ASA

Abstract

The oil and gas (O&G) industry spends billions of dollars on data (e.g., production data, seismic data and tracer data) gathering and analysis for the purpose of reducing uncertainty and improving their understanding of the most salient features of the subsurface. Yet, the O&G industry spends minimal effort and investment to assess whether the benefit of this data gathering and analysis exceeds the cost. One form of data gathering and analysis is history matching (HM), which has been an essential reservoir management tool for decades. This paper addresses the value of HM by applying the value-of-information (VOI) framework originally developed in the decision sciences.

There are several challenges involved in assessing the VOI in the HM context. Although reservoir management (HM and production optimization) and decision analysis (DA) use many of the same methods, the two domains involve different terminology used in fluid flow modeling and application of state-of-the-art HM and optimization methods. Furthermore, most applications of VOI analysis have focused on static, as opposed to time-dependent, analysis. Finally, some recent publications in the O&G industry that have illustrated and discussed VOI from HM have not been consistent with the original definition of VOI.

In this paper, we illustrate and discuss the use of a consistent, DA-based, VOI analysis framework to assess the VOI in HM contexts. In order to make the VOI framework understandable and accessible to both the reservoir management and DA communities, we provide a "bridge" between the nomenclature and terminology used in VOI calculations and that used in state-of-the-art HM and optimization methods.

The paper includes four VOI analysis examples. The first illustrates the implementation of the general VOI framework for a simple HM problem. The second illustrates and discusses the difference between the calculations presented by other authors and the standard VOI definition used in the DA community. The third illustrates the implementation of VOI calculations in more realistic settings, including a sensitivity analysis of measurement noise. The fourth illustrates the application of VOI assessment in a case where a reservoir simulation model is involved.

Introduction

Petroleum engineers and geoscientists involved in reservoir management continually "acquire" information, with the aim of improving decision making. "Information acquisition" is broadly defined here, to cover such activities as acquiring data, performing technical studies, hiring consultants, and performing diagnostic tests. In fact, other than to meet applicable regulatory requirements, the main reason for collecting any information or doing any technical analysis should be to make better decisions. The fundamental question for any information-gathering process is then whether the likely improvement in decision making is worth the cost of obtaining the information. This is the question that value-of-information (VOI) technique is designed to answer.

VOI analysis is an a priori[1] analysis that evaluates the benefits of collecting additional information before one actually gathers the data and makes a decision. Such information gathering might be worthwhile if it could change the decision that would have been made without further information. Although many engineers and geoscientists tend to believe that more information or data is always better, VOI assigns no value to "uncertainty reduction" or "increased confidence" per se. Rather, value is added by enabling the decision maker (DM) to "tune" his/her choice to the underlying uncertainty. Thus, information value is forever an entanglement of uncertainty and decision making; one cannot value information outside of a particular decision context (Bratvold et al. 2009).

[1] "A priori" means "before the data are gathered and interpreted".

The concept of VOI[2] originates from the decision analysis (DA) community. Schlaifer (1959) was the first to define VOI in the context of business decisions. Other early references to VOI analysis can be found in Raiffa and Schlaifer (1961) and Howard (1966). Its use has been documented in wide areas of real-world application, from nuclear waste storage assessment (Eppel and von Winterfeldt 2008) to biosurveillance (Willis and Moore 2013). Grayson (1960) introduced this concept in the oil and gas (O&G) industry. Bratvold et al. (2009) presented the definition of VOI as well as an overview of its use in the O&G industry. Eidsvik et al. (2015) provided a more recent exposition of VOI in the earth sciences.

Although it is more than half a century since the VOI concept was first introduced in the petroleum literature, it has seen limited use in the O&G industry. Most of the published papers that discuss and apply VOI analysis focus on common exploration and development decisions, and most of these applications focus on static, as opposed to time-dependent, analysis. Only recently has the concept been discussed in the context of history matching (HM) which involves data gathered at different times (Barros et al. 2015a, 2015b, 2016a).

HM has been an essential reservoir management tool for decades. Originally, HM referred to the adjustment of the production model parameters to reflect the historical production data (rates and pressures) as closely as possible. Today, the term HM is often used in a broader context and includes model calibration[3] using any relevant data and information (seismic data, log data, tracer behavior, etc.). Once the measured data have been "matched", the model is used to generate possible future production profiles and their probabilities; i.e., the future production uncertainty, and thus support reservoir management decisions.

With the increase in available computing power over the past several decades, numerical techniques for model-based optimization of subsurface hydrocarbon production have evolved rapidly. The traditional manual approach to HM has in many companies been replaced by automatic or semi-automatic approaches based on robust and efficient numerical algorithms that allow for the inclusion of geological and petrophysical uncertainties. A particularly promising method for HM is the ensemble Kalman filter (EnKF) approach introduced by Evensen (1994). EnKF is based on a Bayesian approach where data is used to quantify geological and petrophysical uncertainties, ensuring consistency in the updated probabilities. The resulting production forecast captures these uncertainties and provides possible production scenarios with associated probabilities; i.e., the uncertain future production.

Although operators in the O&G industry commit significant effort and cost to calibrating their production models, few publications have discussed the a priori value of gathering the data and conducting the calibration. The usefulness of information from HM has been quantified in terms of uncertainty reduction (Le and Reynolds 2014) or by the fraction of information extracted from the measured data (Krymskaya et al. 2010). However, neither concept constitutes an information valuation approach as defined by the DA community, nor is either of them an a priori assessment of the value of the information.

Barros et al. (2015a, 2015b, 2016a) proposed a workflow to assess the VOI in the HM context using ensemble-based[4] HM and optimization methods, and later on, Barros et al. (2016b) applied clustering techniques to speed up the previous workflow. Although they have considered the impact of additional information on decision making, the VOI assessed through their workflow is conceptually inconsistent with the definition of VOI as it will be discussed later. Furthermore, they employed terminology and notation that differ significantly from those commonly used in VOI analysis. Indeed, the work by Barros and co-authors is one of the few to claim the use of decision analytic methods to assess VOI in the context of HM. Their workflow mimics in the best possible way the decision process that will happen during reservoir management, but their work is lack of the DA touch and the clarity about why the value they are calculating is VOI. It is important not only to show how the VOI should be assessed but also to illustrate and discuss whether or not their approach is providing the VOI from HM. In this paper, we refer to their workflow (particularly, the workflow introduced in Barros et al. (2015a)) as Barros's workflow. Since transparency is a key feature of DA community, we try to put HM and production optimization in the "world" of DA by setting the "right" scene, and to provide a transparent connection between Barros's workflow and VOI's definition by first deriving a workflow (referred to as our proposed workflow) based on VOI's definition and formulation, and then comparing our proposed workflow with Barros's workflow.

The main contributions of this paper are three-fold. First, we illustrate how VOI analysis can be applied to assess the information value of gathering production data and performing HM. This requires reconciling the terms used in VOI analysis versus those in state-of-the-art HM methods, especially EnKF for HM.[5] Second, we discuss the

[2] This paper uses "value-of-information" as the term is understood in the DA community.
[3] In other modeling contexts, the "matching" of models to measured data is usually referred to as model calibration. We will use "history matching" and "model calibration" interchangeably.
[4] An ensemble means a set of realizations.
[5] Some of the terminology used in EnKF also appears in other approaches to HM in the O&G industry.

differences and similarities between the analysis suggested by Barros et al. (2015a) and VOI analysis. Finally, we present four VOI analysis examples: the first is illustrating the calculation steps of VOI analysis on a simple HM case with the use of decision trees; the second illustrates the use of EnKF with a static model whose corresponding VOI can be calculated analytically; the third uses a decline curve based dynamic model to illustrate a practical case where EnKF combined with robust optimization (RO) is used to calculate the VOI from HM, followed by a sensitivity analysis of the measurement noise—the sensitivity analysis provides a means for illustrating the two VOI extremes: value of perfect information and the case where the information provides no value; and the fourth uses our proposed workflow in a more realistic case where a reservoir simulation model is involved.

Following this introduction, we first clarify the distinction between VOI analysis and terminal analysis and then specify the definition of VOI and present the general steps in VOI assessment. Next, we briefly review EnKF and RO. Thereafter, we propose a procedure for assessing the VOI in HM contexts using ensemble-based methods, which is compared and contrasted with Barros's workflow. The implementation of VOI analysis is then illustrated in a decision-tree example and three Monte Carlo examples. Finally, we present a discussion and conclusions.

Value-of-Information Analysis versus Terminal Analysis

Bratvold et al. (2009) identified several papers in the O&G literature that present cases where the information value is calculated after the information has been gathered. This might take the form of historical lookbacks to document the impact of the information (Aylor 1999; Waggoner 2000). Raiffa and Schlaifer (1961) called this "terminal analysis." Terminal analysis involves the evaluation of and choice among alternatives after a test (actual or hypothetical) has been conducted and the data gathered, whereas VOI analysis (which Raiffa and Schlaifer (1961) called "preposterior analysis") considers the decision problem as it appears before a test has been conducted. [6]

Fig. 1 illustrates the stages of VOI and terminal analysis. **Fig. 2** depicts the decision-tree elements of VOI analysis versus terminal analysis, where the circles represent uncertainty nodes and the squares represent decision nodes. The data of concern in VOI analysis (Fig. 2a) are future data, therefore unknown and treated as uncertain. In contrast, the data of concern in terminal analysis (Fig. 2b) are historical data, therefore already known and treated as certain.

Fig. 1—Stages of (a) VOI analysis, and (b) terminal analysis.

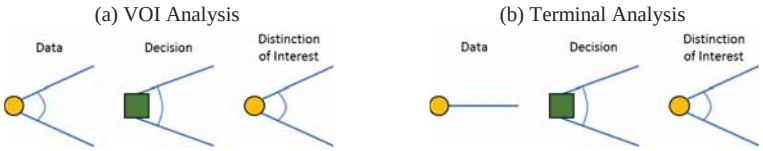


Fig. 2—Decision-tree elements of VOI analysis versus terminal analysis.

Although it might be deemed valuable to conduct terminal analysis, it is not a replacement for VOI analysis. Furthermore, it introduces a bias for two reasons. First, from a communication and publishing perspective, there is a strong incentive not to publish or communicate unsuccessful (in the sense of not being able to demonstrate any value creation) information gathering activities. Second, it ignores cases in which information was not gathered, but should have been.

See Thomas et al. (2016) for a detailed discussion of how terminal analysis can be applied to historical production data.

[6] Another case, which might be called, "post-hoc analysis," is where information is valued after the data have been gathered and the decision has been made (e.g., Coopersmith et al. 2006).

Value-of-Information (VOI) Analysis

HM is the process of acquiring information through updating uncertain model parameters on static and dynamic data. The aim of acquiring information is to improve hydrocarbon production-related decision making. This raises the question of whether the likely improvement in decision making is worth the cost of obtaining the information. Because VOI analysis is designed to answer this, it is useful in the design phase in deciding whether a certain type of data should be gathered for HM. VOI analysis considers a series of actions: data gathering, updating our beliefs based on the gathered data (in our specific case, this is done by model parameter updating through HM), and using the updated beliefs for decision making. If any one of these actions is missing, there will be no point in doing VOI analysis.

VOI analysis is concerned with two fundamental uncertainties: (1) the uncertainties we hope to learn about but cannot directly observe, which we call the distinctions (or events) of interest; and (2) the test results, referred to as the observable distinctions (Bratvold et al. 2009). In reservoir management, the distinction of interest is future production after time t and the observable distinction is the data gathered until time t. Future production is predicted using a dynamic model (e.g., a decline curve model or a reservoir simulation model) with uncertain model parameters and assuming that once we have established values for these parameters, the dynamic model itself will correctly "predict" future production. Therefore, the observed data until time t is the observable distinction which is used to establish values for the uncertain model parameters and the distinctions of interest we actually work on are the uncertain model parameters. For a given set of uncertain model parameters, the future production is a deterministic calculation.

The remainder of this paper uses VOI terminology. **Table 1** lists terms common in VOI analysis, along with the corresponding terms used in HM and optimization.

Notation	VOI Analysis	History Matching / Optimization	Ensemble-based Methods (e.g., EnKF and RO)
$\boldsymbol{x}$	Distinctions of Interest	Model Parameters	Model Parameters
$\tilde{\boldsymbol{y}}$	Observable Distinctions (without measurement noise)	Predicted Observations	Predicted Observations
$\boldsymbol{y}$	Observable Distinctions (with measurement noise)	Observations (with measurement noise) / Measurement Data	Observations (with measurement noise)
$p(\boldsymbol{x})$	Prior	Probability Distribution of Model Parameters	Prior Ensemble—a set of realizations from the prior
$p(\tilde{\boldsymbol{y}}\|\boldsymbol{x})$	Reliability of Model*	Probability Distribution of Predicted Observations Given Model Parameters	Model Noise (higher noise level implies lower reliability)
$p(\boldsymbol{y}\|\tilde{\boldsymbol{y}})$	Reliability of Measurements*	Probability Distribution of Observations Given Predicted Observations	Measurement Noise (higher noise level implies lower reliability)
$p(\boldsymbol{y}\|\boldsymbol{x})$	Likelihood	Probability Distribution of Observations Given Model Parameters	Assessed implicitly by assessing model and measurement noise
$p(\boldsymbol{y})$	Preposterior	Probability Distribution of Observations	Realizations generated by forward modelings plus realizations of model and measurement noise
$p(\boldsymbol{x}\|\boldsymbol{y})$	Posterior	Probability Distribution of Model Parameters Given Observations	Updated/Posterior Ensemble—a set of realizations from the posterior
	Bayesian Inference	Bayesian Calibration	Using e.g., EnKF
$\boldsymbol{a}$	Alternative	Control Variables	Control Variables
	Decision Making	Optimization	Using e.g., RO

*The term is not used in common VOI analysis. It is introduced here for the HM context.

Table 1—Correspondence between terms used in VOI analysis and those used in model calibration and optimization.

The VOI is defined as the most that the DM should pay for additional information on the distinctions of interest. If the DM is risk neutral, then[7]

$$VOI = \left[\begin{matrix}\text{Expected value } with \\ \text{additional information}\end{matrix}\right] - \left[\begin{matrix}\text{Expected value } without \\ \text{additional information}\end{matrix}\right].$$

[7] This is not the general definition of VOI. This equation is true only if the decision maker is risk neutral or is risk-averse with an exponential utility function (Bratvold et al. 2009).

If the additional information is perfect (i.e., the information reveals the truth), we refer to the VOI as the value-of-perfect-information (VOPI) or the value-of-clairvoyance (VOC). No test, no matter how sophisticated, can be worth more than the VOPI (Bratvold et al. 2009). In the context of HM, obtaining perfect information is usually precluded by model noise[8], measurement noise, or non-uniqueness of inverse modelling.

In mathematical form,

$$VOI = \max\{0, \Delta\} \tag{1}$$
$$\Delta = EVWI - EVWOI \tag{2}$$

where $EVWOI$ is the expected value (EV) without additional information and $EVWI$ is the EV with additional information. The lower bound of VOI is always 0 because if Δ is negative when $EVWOI > EVWI$, one can always choose to not gather the information. VOI is an indicator of the maximal buying price or cost of an information gathering activity. If the VOI is greater than the cost, the DM should gather the information; otherwise, he/she should not do so.

In a decision making context, the decision without information (DWOI) is the alternative that optimizes EV over the prior, and the EVWOI is the optimal EV over the prior, i.e.,

$$\boldsymbol{a}_{pri}^{opt} = \underset{\boldsymbol{a}\in\boldsymbol{A}}{\operatorname{argmax}}\{EV[v(\boldsymbol{x},\boldsymbol{a})]\} = \underset{\boldsymbol{a}\in\boldsymbol{A}}{\operatorname{argmax}}\left\{\int_{\boldsymbol{x}} v(\boldsymbol{x},\boldsymbol{a})p(\boldsymbol{x})d\boldsymbol{x}\right\} \tag{3}$$
$$EVWOI = EV[v(\boldsymbol{x},\boldsymbol{a}_{pri}^{opt})] = \int_{\boldsymbol{x}} v(\boldsymbol{x},\boldsymbol{a}_{pri}^{opt})p(\boldsymbol{x})d\boldsymbol{x} \tag{4}$$

where $\boldsymbol{a}$ is an alternative from the decision space $\boldsymbol{A}$, $\boldsymbol{a}_{pri}^{opt}$ is the DWOI, $\boldsymbol{x}$ is the distinctions of interest, $v(\boldsymbol{x},\boldsymbol{a})$ is the value function that assigns a value to each alternative-outcome pair for a given $\boldsymbol{x}$, and $p(\boldsymbol{x})$ is the prior probability distribution of $\boldsymbol{x}$. Similarly, for given observations, the decision with information (DWI) is the alternative that optimizes EV over the posterior, i.e.,

$$\boldsymbol{a}_{pos}^{opt}(\boldsymbol{y}) = \underset{\boldsymbol{a}\in\boldsymbol{A}}{\operatorname{argmax}}\{EV[v(\boldsymbol{x},\boldsymbol{a})|\boldsymbol{y}]\} = \underset{\boldsymbol{a}\in\boldsymbol{A}}{\operatorname{argmax}}\left\{\int_{\boldsymbol{x}} v(\boldsymbol{x},\boldsymbol{a})p(\boldsymbol{x}|\boldsymbol{y})d\boldsymbol{x}\right\} \tag{5}$$
$$EV_{pos}^{opt}(\boldsymbol{y}) = EV\left[v(\boldsymbol{x},\boldsymbol{a}_{pos}^{opt}(\boldsymbol{y}))|\boldsymbol{y}\right] = \int_{\boldsymbol{x}} v(\boldsymbol{x},\boldsymbol{a}_{pos}^{opt}(\boldsymbol{y}))p(\boldsymbol{x}|\boldsymbol{y})d\boldsymbol{x} \tag{6}$$

where $\boldsymbol{a}_{pos}^{opt}(\boldsymbol{y})$ is the DWI given observations $\boldsymbol{y}$, $p(\boldsymbol{x}|\boldsymbol{y})$ is the posterior probability distribution (i.e., the probability distribution of the model parameters given the observations), and EV_{pos}^{opt} is the optimal EV over the posterior. The posterior probability distribution is assessed by using Bayes' rule

$$p(\boldsymbol{x}|\boldsymbol{y}) = \frac{p(\boldsymbol{x},\boldsymbol{y})}{p(\boldsymbol{y})} = \frac{p(\boldsymbol{y}|\boldsymbol{x})p(\boldsymbol{x})}{p(\boldsymbol{y})} \tag{7}$$

where $p(\boldsymbol{y}|\boldsymbol{x})$ is the likelihood function that encodes the reliability of model calibration,[9] $p(\boldsymbol{x},\boldsymbol{y})$ is the joint probability distribution, and $p(\boldsymbol{y})$ is the preposterior probability distribution calculated as

$$p(\boldsymbol{y}) = \int_{\boldsymbol{x}} p(\boldsymbol{x},\boldsymbol{y})d\boldsymbol{x} = \int_{\boldsymbol{x}} p(\boldsymbol{y}|\boldsymbol{x})p(\boldsymbol{x})d\boldsymbol{x}. \tag{8}$$

Then, the EVWI is

[8] By "model noise," we mean any disturbance that affects the dynamics. We don't know the exact details of the disturbance, so we have to treat it as a random process. The lower the model noise, the better that the model represents the reality, and the more reliable the model is. In reservoir management, the model noise is usually ignored.

[9] The reliability of model calibration includes the reliability of the model, the reliability of the measurements, and the non-uniqueness of inverse modelling.

$$EVWI = \int_{\boldsymbol{y}} EV_{pos}^{opt}(\boldsymbol{y})p(\boldsymbol{y})d\boldsymbol{y} = \int_{\boldsymbol{y}}\int_{\boldsymbol{x}} v\big(\boldsymbol{x}, \boldsymbol{a}_{pos}^{opt}(\boldsymbol{y})\big)p(\boldsymbol{x}|\boldsymbol{y})d\boldsymbol{x}\, p(\boldsymbol{y})d\boldsymbol{y}. \tag{9}$$

Eq. 9 can be rewritten as

$$\begin{aligned} EVWI &= \int_{\boldsymbol{y}}\int_{\boldsymbol{x}} v\big(\boldsymbol{x}, \boldsymbol{a}_{pos}^{opt}(\boldsymbol{y})\big)p(\boldsymbol{x}|\boldsymbol{y})d\boldsymbol{x}\, p(\boldsymbol{y})d\boldsymbol{y} = \int_{\boldsymbol{y}}\int_{\boldsymbol{x}} v\big(\boldsymbol{x}, \boldsymbol{a}_{pos}^{opt}(\boldsymbol{y})\big)p(\boldsymbol{x}|\boldsymbol{y})p(\boldsymbol{y})d\boldsymbol{x}\, d\boldsymbol{y} \\ &= \int_{\boldsymbol{y}}\int_{\boldsymbol{x}} v\big(\boldsymbol{x}, \boldsymbol{a}_{pos}^{opt}(\boldsymbol{y})\big)p(\boldsymbol{x}, \boldsymbol{y})d\boldsymbol{x}\, d\boldsymbol{y} = \int_{(\boldsymbol{x},\boldsymbol{y})} v\big(\boldsymbol{x}, \boldsymbol{a}_{pos}^{opt}(\boldsymbol{y})\big)p(\boldsymbol{x}, \boldsymbol{y})d(\boldsymbol{x}, \boldsymbol{y}) \end{aligned} \tag{10}$$

where $(\boldsymbol{x}, \boldsymbol{y})$ denotes a joint space of $\boldsymbol{x}$ and $\boldsymbol{y}$, and $\int_{(\boldsymbol{x},\boldsymbol{y})} \cdot\, d(\boldsymbol{x}, \boldsymbol{y})$ stands for the integration over all combinations of $\boldsymbol{x}$ and $\boldsymbol{y}$. The expressions in Eqs. 1–10 do not show explicitly the time-dependent components. Each component of the decision vector $\boldsymbol{a}_{pos}^{opt}(\boldsymbol{y})$ is a function of all the observations available at the time of evaluating the decision. Refer to Bellman equation for a complete expression for a time-dependent system (Eidsvik et al. 2015).

Bayes' rule (Eq. 7) is central to VOI analysis. One must assess the prior and the likelihood function in order to draw the inference, so it is assumed that the prior accurately reflects the DM's uncertainty about the distinction of interest and that the likelihood function correctly indicates the strength of the evidence from the model calibration.

The general workflow of VOI analysis is illustrated in **Fig. 3**. The main steps of VOI analysis are (1) to identify the DWOI (Eq. 3) and calculate the EVWOI (Eq. 4); (2) to conduct Bayesian inference (Eq. 7), identify the DWI (Eq. 5) and calculate the EVWI (Eq. 9 or 10); and (3) to calculate the VOI (Eqs. 1 and 2).

EVWI can be calculated using either Eq. 9 or Eq. 10. In Formulation 1 (F1) (Eq. 9), the optimal EV over the posterior is first calculated by integrating over all possible distinctions of interest given observations, and the EVWI is then calculated by integrating over all possible observations. In Formulation 2 (F2) (Eq. 10), EVWI is calculated by integrating over all possible combinations of the distinctions of interest and the observations. Mathematically, these two formulations give the same value of EVWI. However, when Monte Carlo methods are used, the EVWIs calculated using these two formulations might be very different, leading to different VOI estimates. This will be further discussed later.

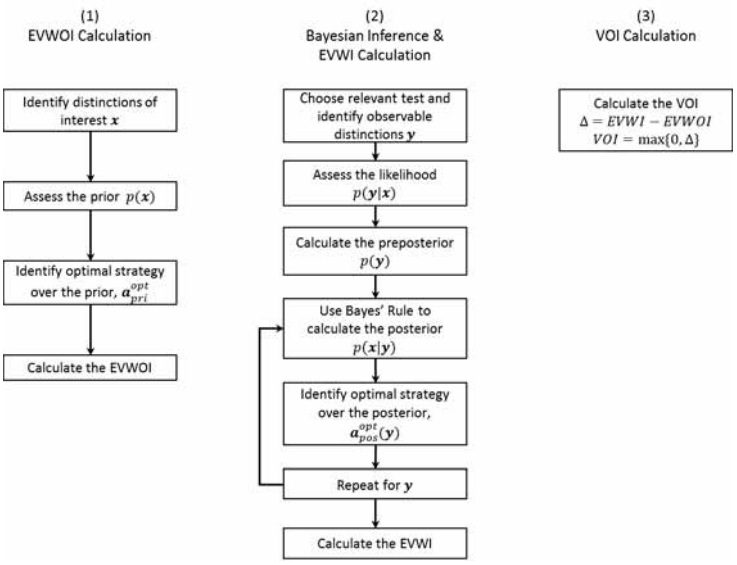


Fig. 3—General workflow of VOI analysis.

There are two essential operations in VOI analysis: Bayesian inference (i.e., calculating the posterior probability distribution) and decision making (i.e., identifying the optimal strategy). Reservoir management decision problems that involve continuous probability distributions and continuous decision spaces require methods other than decision trees for these two steps. EnKF for Bayesian inference and RO for decision making from closed-loop reservoir management are more appropriate methods for building a VOI analysis environment in HM context (Barros et al. 2016a). The following sections will briefly present these two methods.

Bayesian Inference with Ensemble Kalman Filter (EnKF)

EnKF was introduced by Evensen (1994) as a means to conduct Bayesian inference (Eq. 7) using Monte Carlo methods. The prior probability distribution is represented by a set of realizations of model parameters (i.e., the prior or initial ensemble). The posterior probability distribution is represented by the updated realizations of model parameters (i.e., the posterior or updated ensemble), which are obtained by (Burgers et al. 1998)

$$\boldsymbol{Y}_a = \boldsymbol{Y}_f + \boldsymbol{K}(\boldsymbol{y}_p - \boldsymbol{H}\boldsymbol{Y}_f) \tag{11}$$

where matrix $\boldsymbol{Y}_a$ consists of the vectors containing the updated states,[10] updated model parameters, and updated observations corresponding to each realization in the posterior ensemble; matrix $\boldsymbol{Y}_f$ contains the predicted states by forward modelling, model parameters, and predicted observations by forward modelling corresponding to each realization in the prior ensemble; $\boldsymbol{K}$ is the Kalman gain matrix, which weighs the influences of the prior predicted observations and the real-time observations (i.e., measurement data); $\boldsymbol{y}_p$ is a matrix containing the perturbed observations;[11] and $\boldsymbol{H}$ is an operator that links $\boldsymbol{Y}_f$ to the predicted observations. Eq. 11 describes a linear combination of the prior and the observations. The Kalman gain matrix $\boldsymbol{K}$ is calculated as

$$\boldsymbol{K} = \boldsymbol{P}_f\boldsymbol{H}^T(\boldsymbol{H}\boldsymbol{P}_f\boldsymbol{H}^T + \boldsymbol{C}_y)^{-1} \tag{12}$$

where $\boldsymbol{P}_f$ is the covariance matrix of $\boldsymbol{Y}_f$ encoding the covariance matrix of the prior predicted observations, and $\boldsymbol{C}_y$ is the covariance matrix of the observations. As the measurements become noisier (i.e., the variance of an observation increases) and/or the variance of a prior predicted observation decreases, more weight is given to the prior; otherwise, more weight is given to the observations.

The EnKF embodies the prior in $\boldsymbol{Y}_f$, the likelihood in $\boldsymbol{C}_y$ (when the model noise is ignored), and the posterior in $\boldsymbol{Y}_a$; and the preposterior is a normalizing constant of the posterior. Thus, Bayes' rule describing the relationship among the prior, the likelihood, the preposterior, and the posterior is no longer shown explicitly as in Eq. 7, but is implicitly included in Eq. 11 and 12. A detailed description on the relationship between the formulation of Kalman filter and Bayesian formulation can be found in Meinhold and Singpurwalla (1983). In our case, the inputs of EnKF are the initial guess of the model parameters (i.e., the prior ensemble) together with a model that can predict both the production and observations given a production strategy, observations, and their associated statistics; and the output is the EnKF updated model parameters (i.e., the posterior ensemble). Using EnKF, Bayes' inference is conducted implicitly.

For a comprehensive introduction to EnKF, refer to Evensen (2009). Aanonsen et al. (2009) provided an extensive review of the application of EnKF in reservoir engineering.

Decision Making with Robust Optimization (RO)

The purpose of HM, and any other information gathering activity, is to acquire information with the aim of improving decision making. This generally requires explicit modeling of the decisions affected by the information obtained from the calibration. This section considers how to optimize the decisions of what water injection strategy to employ. The variables control water injection rates are commonly referred to as the control variables.

The goal of applying RO approach to water injection is to find an optimal strategy (i.e., a set of decisions) that is robust to geological uncertainties (Van Essen et al. 2009). A commonly used measure of the value of a hydrocarbon production project is net present value (NPV). We consider only the revenues from oil production and the costs of

[10] Given a dynamic system, a state is an unobservable (in most cases) quantity that must be known currently in order to predict the system behavior (i.e., the future state). A state can be a scalar or a vector.

[11] When EnKF is used, an observation has to be perturbed with its corresponding statistics in order to avoid insufficient variance (Burgers et al. 1998).

water injection and production for a waterflooding case. The NPV for a single geological realization can be calculated by

$$v(\boldsymbol{x},\boldsymbol{a}) = NPV(\boldsymbol{x},\boldsymbol{a}) = \sum_{k=1}^{n_T} \frac{(q_o^k(\boldsymbol{x},\boldsymbol{a})P_o - q_{wp}^k(\boldsymbol{x},\boldsymbol{a})P_{wp} - q_{wi}^k(\boldsymbol{x},\boldsymbol{a})P_{wi})\Delta t_k}{(1+b)^{t_k/\tau}} \tag{13}$$

where $v(\cdot)$ is the value function for decision making; $\boldsymbol{x}$ is a vector of uncertain model parameters for a geological realization; $\boldsymbol{a}$ is a vector of control variables (i.e., an alternative); n_T is the number of time steps; k is the index for time step; the subscripts o, wp, and wi denote oil, water production, and water injection, respectively; q^k is the injection or production rate over time step length Δt_k; P_o, P_{wp}, and P_{wi} are the oil price, water production cost, and water injection cost, respectively; b is the discount rate; t_k is the cumulative time for discounting; and τ is the reference time for discounting. For a risk-neutral DM, the objective function for RO is

$$\bar{v}(\boldsymbol{a}) = \frac{1}{n_e}\sum_{i=1}^{n_e} v(\boldsymbol{x}_i,\boldsymbol{a}) \tag{14}$$

where $\bar{v}(\cdot)$ is the EV of the value function over all realizations, i is the index of a realization, and n_e is the number of realizations (i.e., the ensemble size). The objective of RO is thus to find an optimal $\boldsymbol{a}$ that maximizes the expected NPV over an ensemble of geological realizations.

For computationally attractive models, the FMINCON-function of MATLAB® (2014) can be used to solve the RO problem. For computationally intensive models, such as a grid-based finite difference or finite element model, ensemble-based optimization approaches such as the EnOpt (Chen et al. 2009) are commonly used to significantly reduce the number of reservoir simulations required for RO and hence reduce the computational costs.

VOI Analysis Using Ensemble-based Methods

In ensemble-based methods, a probability distribution is represented by an ensemble, which is a set of realizations. Each realization in the ensemble (i.e., an ensemble member) is equi-probable with a probability of $1/n_e$. This is far more tractable than it would be to weight all possible values according to the probability distribution.

The workflow of VOI analysis using ensemble-based methods is illustrated in **Fig. 4**. This workflow is a modification of the general VOI analysis workflow shown in Fig. 3. The difference between Figs. 3 and 4 is that the probability distribution appears in Fig. 3 are represented by ensembles in Fig. 4.

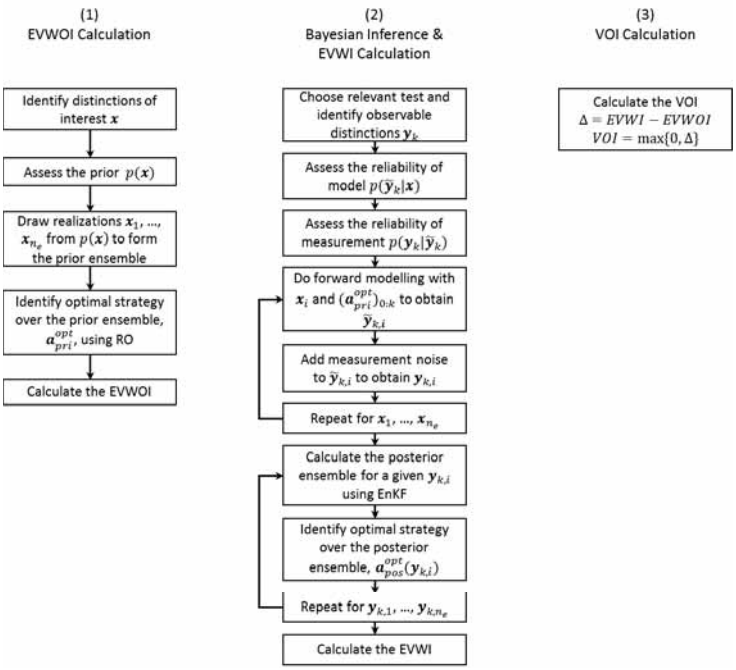


Fig. 4—Workflow of VOI analysis using ensemble-based methods.

The prior probability distribution $p(\boldsymbol{x})$ is represented by the prior ensemble $[\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_{n_e}]$, which consists of the realizations drawn from $p(\boldsymbol{x})$.

The likelihood function $p(\boldsymbol{y}|\boldsymbol{x})$ is assessed through a measurement noise distribution. A dynamic system with model and measurement noise involves the following mathematical relationships:

$$\boldsymbol{\theta}_k = g(\boldsymbol{x}, \boldsymbol{a}_k, \boldsymbol{\theta}_{k-1}) + \boldsymbol{\varepsilon}_{g,k} \quad (15)$$

$$\tilde{\boldsymbol{y}}_k = h(\boldsymbol{\theta}_k) \quad (16)$$

$$\boldsymbol{y}_k = \tilde{\boldsymbol{y}}_k + \boldsymbol{\varepsilon}_{y,k} \quad (17)$$

where the subscript k is the index of time, $\boldsymbol{\theta}$ is a vector of the states, $\boldsymbol{\varepsilon}_{g,k}$ is a vector of the model noise at time k, $g(\cdot)$ is a function that describes the relationship between current states $\boldsymbol{\theta}_{k-1}$ and future states $\boldsymbol{\theta}_k$ given uncertain parameters $\boldsymbol{x}$ and control variables $\boldsymbol{a}_k$ for the period between time $k-1$ and k, $\tilde{\boldsymbol{y}}_k$ is a vector of predicted observations at time k, $h(\cdot)$ is a function that describes the relationship between states and predicted observations, $\boldsymbol{y}_k$ is a vector of observations (with measurement noise) at time k, and $\boldsymbol{\varepsilon}_{y,k}$ is a vector of the measurement noise at time k. Combining Eq. 15 and Eq. 16 yields

$$\tilde{\boldsymbol{y}}_k = f(\boldsymbol{x}, \boldsymbol{a}_{0:k}, \boldsymbol{\theta}_0, \boldsymbol{\varepsilon}_{f,k}) \quad (18)$$

where $f(\cdot)$ denotes a forward model that embodies $g(\cdot)$ and $h(\cdot)$, $\boldsymbol{a}_{0:k}$ is a time-series of control variables from time 0 to time k, $\boldsymbol{\theta}_0$ is the initial state, and $\boldsymbol{\varepsilon}_{f,k}$ encodes the model noise $\boldsymbol{\varepsilon}_g$ from time 0 to time k. If $\boldsymbol{\theta}_0$ is uncertain, it can be included in $\boldsymbol{x}$; if it is certain, it can be included in $f(\cdot)$. Thus, Eq. 18 can be simplified to

$$\tilde{\boldsymbol{y}}_k = f\big(\boldsymbol{x}, \boldsymbol{a}_{0:k}, \boldsymbol{\varepsilon}_{f,k}\big). \quad (19)$$

For example, when using EnKF together with a reservoir simulation model, standard practice is to assume that the model is accurate (i.e. $\boldsymbol{\varepsilon}_f = \mathbf{0}$) and to assign a multivariate Gaussian distribution to $\boldsymbol{\varepsilon}_{y,k}$ with zero mean and a diagonal covariance matrix $\boldsymbol{C}_{y,k}$ (i.e., $p(\boldsymbol{\varepsilon}_{y,k}|\tilde{\boldsymbol{y}}_k) \sim N(\mathbf{0}, \boldsymbol{C}_{y,k})$); this yields $p(\tilde{\boldsymbol{y}}_k|\boldsymbol{x}) \sim N(f(\boldsymbol{x}, \boldsymbol{a}_{0:k}), \mathbf{0})$ and $p(\boldsymbol{y}_k|\tilde{\boldsymbol{y}}_k) \sim N(\tilde{\boldsymbol{y}}_k, \boldsymbol{C}_{y,k})$, where $p(\tilde{\boldsymbol{y}}_k|\boldsymbol{x})$ is referred to as the reliability of the model and $p(\boldsymbol{y}_k|\tilde{\boldsymbol{y}}_k)$ as the reliability of measurements, thus $p(\boldsymbol{y}_k|\boldsymbol{x}) \sim N(f(\boldsymbol{x}, \boldsymbol{a}_{0:k}), \boldsymbol{C}_{y,k})$. Therefore, as long as the model and measurement noise are assessed and a forward model is given, the likelihood function can be calculated.

The realizations representing the preposterior probability distribution (i.e., the realizations of observations) can be generated by Monte Carlo simulation: (1) sample a realization of the uncertain model parameters $\boldsymbol{x}_i$ from $p(\boldsymbol{x})$ (i.e., the same as generating a realization in the prior ensemble), (2) conduct forward modelling (Eq. 19 with $\boldsymbol{\varepsilon}_f = \mathbf{0}$ if the model noise is ignored) with $\boldsymbol{x}_i$ to time k to obtain a realization of the predicted observations $\tilde{\boldsymbol{y}}_{k,i}$, (3) sample a realization of the measurement noise $\boldsymbol{\varepsilon}_{y,k,i}$ from $p(\boldsymbol{\varepsilon}_{y,k}|\tilde{\boldsymbol{y}}_{k,i})$, (4) add $\boldsymbol{\varepsilon}_{y,k,i}$ to $\tilde{\boldsymbol{y}}_{k,i}$ (Eq. 17) to obtain a realization of the observations with noise $\boldsymbol{y}_{k,i}$, and (5) repeat (1)–(3) to obtain a set of $\boldsymbol{y}_k$, $[\boldsymbol{y}_{k,1}, \boldsymbol{y}_{k,2}, \ldots, \boldsymbol{y}_{k,n_e}]$, which represents the preposterior. Moreover, the realizations of $(\boldsymbol{x},\boldsymbol{y}_k)$-pairs $[(\boldsymbol{x}_1, \boldsymbol{y}_{k,1}), (\boldsymbol{x}_2, \boldsymbol{y}_{k,2}), \ldots, (\boldsymbol{x}_{n_e}, \boldsymbol{y}_{k,n_e})]$ represent the joint probability distribution $p(\boldsymbol{x}, \boldsymbol{y}_k)$. In this specific case, we use a 1:1 ratio between $\boldsymbol{x}$ and $\boldsymbol{y}_k$, i.e., for a given $\boldsymbol{x}_i$, only one realization of $\boldsymbol{y}_k$ is drawn from $p(\boldsymbol{y}_k|\boldsymbol{x}_i)$. One can alternatively use a 1:m ratio, i.e., for a given $\boldsymbol{x}_i$, m realizations of $\boldsymbol{y}_k$ are drawn from $p(\boldsymbol{y}_k|\boldsymbol{x}_i)$. This results in a set of $\boldsymbol{y}_k$ with a total of $m \cdot n_e$ realizations $[\boldsymbol{y}_{k,1}, \boldsymbol{y}_{k,2}, \ldots, \boldsymbol{y}_{k,m \cdot n_e}]$ and a set of $(\boldsymbol{x}, \boldsymbol{y}_k)$-pairs with a total of $m \cdot n_e$ realizations $[(\boldsymbol{x}_1, \boldsymbol{y}_{k,1}), \ldots, (\boldsymbol{x}_1, \boldsymbol{y}_{k,m}), (\boldsymbol{x}_2, \boldsymbol{y}_{k,m+1}), \ldots, (\boldsymbol{x}_2, \boldsymbol{y}_{k,2m}), \ldots, (\boldsymbol{x}_{n_e}, \boldsymbol{y}_{k,m(n_e-1)+1}), \ldots, (\boldsymbol{x}_2, \boldsymbol{y}_{k,m \cdot n_e})]$. Using a 1:1 ratio is a simplification that can lead to inaccuracies in the VOI estimate. However, using a 1:m ratio requires m times more Monte Carlo samples, m times increasing the computational cost for the EVWI calculation accordingly. In practice, the measurement noise tends to be quite small (i.e., the standard deviation of $\boldsymbol{y}_k|\boldsymbol{x}_i$ is small compared to that of the prior prediction); thus, a 1:1 ratio suffices.

The posterior ensemble $[\hat{\boldsymbol{x}}_1|\boldsymbol{y}_k, \hat{\boldsymbol{x}}_2|\boldsymbol{y}_k, \ldots, \hat{\boldsymbol{x}}_{n_e}|\boldsymbol{y}_k]$ representing the posterior probability distribution $p(\boldsymbol{x}|\boldsymbol{y}_k)$ is generated by using EnKF to update the prior ensemble with the given observations and assessed measurement (and model) noise.

With all the probability distributions represented by their corresponding realizations, the VOI analysis can be conducted following the steps presented earlier. The following equations omit the time index k for convenience. First, we identify the DWOI and calculate the EVWOI. Converting Eqs. 3 and 4 into their corresponding Monte Carlo forms gives

$$\boldsymbol{a}_{pri}^{opt} = \underset{\boldsymbol{a} \in \boldsymbol{A}}{\operatorname{argmax}} \left\{ \frac{1}{n_e} \sum_{i=1}^{n_e} v(\boldsymbol{x}_i, \boldsymbol{a}) \right\}, \tag{20}$$

$$EVWOI = \frac{1}{n_e} \sum_{i=1}^{n_e} v(\boldsymbol{x}_i, \boldsymbol{a}_{pri}^{opt}). \tag{21}$$

Second, we conduct Bayesian inference, using EnKF to calculate the posterior ensemble members and identify the DWI for the given observations, using the Monte Carlo form of Eq. 5,

$$\boldsymbol{a}_{pos}^{opt}(\boldsymbol{y}) = \underset{\boldsymbol{a} \in \boldsymbol{A}}{\operatorname{argmax}} \left\{ \frac{1}{n_e} \sum_{i=1}^{n_e} v(\hat{\boldsymbol{x}}_i|\boldsymbol{y}, \boldsymbol{a}) \right\}. \tag{22}$$

Again, the EVWI can be calculated using either of two formulations: the Monte Carlo form of F1 (Eq. 9),

$$EVWI = \frac{1}{n_e} \sum_{j=1}^{n_e} \left[\frac{1}{n_e} \sum_{i=1}^{n_e} v\left(\hat{\boldsymbol{x}}_i|\boldsymbol{y}_j, \boldsymbol{a}_{pos}^{opt}(\boldsymbol{y}_j)\right) \right] \tag{23}$$

or the Monte Carlo form of F2 (Eq. 10),

$$EVWI = \frac{1}{n_e}\sum_{i=1}^{n_e} v\left(\boldsymbol{x}_i, \boldsymbol{a}_{pos}^{opt}(\boldsymbol{y}_i)\right). \quad (24)$$

Based on the law of large numbers, Eqs. 23 and 24 give the same limit when the ensemble size n_e approaches infinity. However, the convergence rate can be different, resulting in different estimates when the ensemble size is small. More details on this will be presented with an example later. Finally, we calculate VOI using Eqs. 1 and 2.

Care must be taken when calculating EVWOI and EVWI. If a production optimization problem entails sequential decision making; i.e., the alternative $\boldsymbol{a}$ consists of a series of decisions, only the value corresponding to decisions that can be affected by the observations should be considered. For example, if decisions a_1 and a_2 will be made sequentially at time t and $t + \Delta t$, respectively, then an observation obtained somewhere between time t and $t + \Delta t$ will affect only a_2 and not a_1,. Therefore, when calculating the value of this observation, only the NPV corresponding to the period from time $t + \Delta t$ forward should be considered.

Comparison with Barros's Analysis Workflow

Barros et al. (2015a) proposed a workflow for calculating VOI from HM using ensemble-based methods. The authors have extensive experience in reservoir management. However, their paper did not elaborate on the connection between the workflow they have applied and the definition of VOI. This section seeks to accomplish that.

Barros's workflow is based on a twin experiment where the DWI and DWOI are valued based on a synthetic truth and where the difference between the values corresponding to the DWI and DWOI, respectively, is calculated. A synthetic truth in Barros's workflow is a realization drawn from the prior probability distribution. n_e realizations are first drawn from the prior probability distribution, and one of them is then chosen as a synthetic truth and the rest form the prior ensemble with $n_e - 1$ realizations. The procedure of conducting a twin experiment is repeated for all n_e synthetic truths. For each synthetic truth, there is a corresponding difference between its values corresponding to the DWI and DWOI. Barros et al. then calculated the VOI as the EV over all of these differences associated with synthetic truths. Thus, in Barros's workflow, the concept of synthetic truth is central to the analysis.

However, the concept of "synthetic truth" has no meaning in VOI analysis. In reservoir engineering, a synthetically true model is used to mimic the actual state of a reservoir, by which reservoir engineers can investigate whether the actual outcome falls within the range that was predicted by updated models. This is not the case of VOI analysis where the importance is to quantify uncertainty and how the changes in uncertainty will change decisions. Moreover, VOI as defined by Barros et al. implies a distribution.[12] However, as shown by Bratvold et al. (2009), the VOI is not a distribution but the difference between two expected values. Thus, Barros's workflow is inconsistent with the definition of VOI analysis. However, two small modifications to Barros's workflow can address this.

In VOI analysis, the prior reflects the DM's lack of knowledge, which means that the truth is included in the prior.[13] Therefore, the first modification is to not exclude any synthetic truth from the prior ensemble. Because the VOI is defined as the difference between the EVWI and EVWOI, the second modification is to calculate the VOI by first calculating the EV over the values of the DWIs valued on the synthetic truths and the EV over the values of the DWOI valued on the synthetic truths, and then calculating the difference between these two EVs instead of by first calculating the difference between the values of DWI and DWOI valued on a synthetic truth, and then calculating the EV over the synthetic truths. Although the second modification is only formula rewriting and causes no difference in mathematical result, it makes conceptual difference. The modified Barros's workflow is identical to calculating the VOI using F2. Because F2 is derived without introducing the concept of synthetic truth, there is no need to include the concept of synthetic truth in VOI analysis.

Barros's workflow is from the standpoint of reservoir engineers, whilst F2 is from the original definition of VOI. Despite the inconsistent VOI definition implemented through Barros's workflow, the similarity of it to F2 provides an easier understanding of the VOI from HM for reservoir engineers.

The differences among F1, F2, and Barros's approaches are summarized in **Table 2**. The impact of using the three approaches will be illustrated and discussed later.

[12] Barros et al. (2015a, 2015b, 2016a) illustrated the percentiles of VOI.

[13] As is common in VOI analysis, this work assumes that the probability assessors are unbiased.

Index	Formulation 1	Formulation 2	Barros et al.
1	DWOI is identified and EVWOI is calculated over all n_e prior realizations (Eqs. 20 and 21).	DWOI is identified and EVWOI is calculated over all n_e prior realizations (Eqs. 20 and 21).	DWOI is identified by optimizing the EV over $(n_e - 1)$ prior realizations after a synthetic truth is excluded. The DWOI is valued on the synthetic truth, i.e., value without information for the synthetic truth ($VWOI_i$) is calculated. This procedure is repeated for all synthetic truths.
2	All n_e prior realizations are updated, resulting in n_e posterior realizations.	All n_e prior realizations are updated, resulting in n_e posterior realizations.	$(n_e - 1)$ prior realizations are updated, resulting in $(n_e - 1)$ posterior realizations.
3	DWI is identified over n_e posterior realizations (Eq. 22).	DWI is identified over n_e posterior realizations (Eq. 22).	DWI is identified over $(n_e - 1)$ posterior realizations.
4	EVWI is calculated using Eq. 23.	EVWI is calculated using Eq. 24.	The value with information for each particular synthetic truth (VWI_i) is calculated.
5	VOI is calculated as the difference between EVWI and EVWOI.	VOI is calculated as the difference between EVWI and EVWOI.	The VOI for each synthetic truth (VOI_i) is calculated as $VOI_i = VWI_i - VWOI_i$, and the VOI is calculated as the EV of VOI_i.

Table 2—Distinctions among Formulation 1, Formulation 2, and Barros's approaches.

Decision Tree Example

The following is a simple illustration of how decision trees can be used to assess the VOI in the HM context. For an introduction to decision trees and tree flipping, refer to Clemen (1991), Bratvold and Begg (2010), and Howard and Abbas (2016).

Assume that the prior has been assessed as three equi-probable sets of uncertain model parameters of a given production model: $\boldsymbol{x}_1$, $\boldsymbol{x}_2$, and $\boldsymbol{x}_3$. We want to estimate the value of measuring and calibrating to the oil production rate at time t. The predicted oil production rates $\tilde{q}_o$ by these three realizations are shown in **Fig. 5**, where $\tilde{q}_{o,i}^{0:T}$ is the oil production rate predicted for the period from time 0 to T, which is obtained from the production model $f(\cdot)$ with parameters $\boldsymbol{x}_i$ and production strategy $\boldsymbol{a}$. It is assumed that there is no model noise. For example, if $\boldsymbol{x}_1$ describes the true nature of the reservoir, $\boldsymbol{a}_2$ is the production strategy and there is no measurement noise, the observed oil production rate will definitely be $\tilde{q}_{o,L}^t$ at time t (i.e., any given set of values of model parameters $\boldsymbol{x}$ maps to only one set of predicted observations $\tilde{\boldsymbol{y}}$—a one-to-one relationship). Thus, we have the following model reliabilities: $p(\tilde{q}_{o,L}^t|\boldsymbol{x}_1) = p(\tilde{q}_{o,L}^t|\boldsymbol{x}_2) = p(\tilde{q}_{o,H}^t|\boldsymbol{x}_3) = 1$, else 0. However, the relationship from predicted observations to model parameters might not be one-to-one, because of non-uniqueness of inverse modelling. For example, given $\tilde{q}_{o,L}^t$, the corresponding set of model parameters can be either $\boldsymbol{x}_1$ or $\boldsymbol{x}_2$. The reliabilities of the measurements are listed in **Table 3**. For example, given the oil production rate is low ($\tilde{q}_{o,L}^t$), the probability of the measurement providing an exact result "low" ($q_{o,L}^t$) is 4/5 and that of a deviated result "high" ($q_{o,H}^t$) is 1/5. Given the reliabilities of the model and measurements, we can calculate the likelihoods. For example, $p(q_{o,L}^t|\boldsymbol{x}_1) = p(q_{o,L}^t, \tilde{q}_{o,L}^t|\boldsymbol{x}_1) + p(q_{o,L}^t, \tilde{q}_{o,H}^t|\boldsymbol{x}_1) = p(q_{o,L}^t|\tilde{q}_{o,L}^t)p(\tilde{q}_{o,L}^t|\boldsymbol{x}_1) + p(q_{o,L}^t|\tilde{q}_{o,H}^t)p(\tilde{q}_{o,H}^t|\boldsymbol{x}_1) = \frac{4}{5}\cdot 1 + \frac{1}{5}\cdot 0 = 4/5$. The likelihood function is listed in **Table 4**. The preposterior and posteriors are calculated using Eq. 8 and 7, respectively.

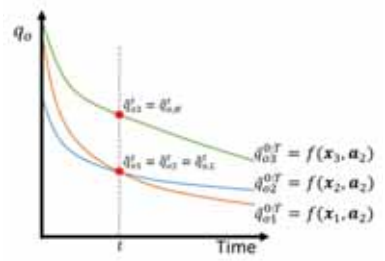


Fig. 5—The oil production rate profiles of three equi-probable realizations for the decision tree example.

Reliabilities of Measurement p(**says**|**is**)		Given q_o at time t **is** … Low ($\tilde{q}_{o,L}^t$)	High ($\tilde{q}_{o,H}^t$)
… the measurement at time t **says**	"Low" ($q_{o,L}^t$)	4/5	1/5
	"High" ($q_{o,H}^t$)	1/5	4/5
	Total	1	1

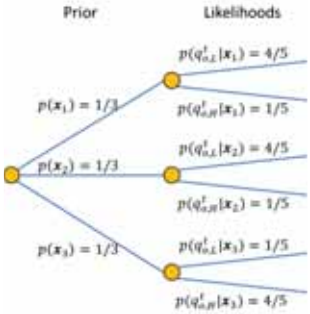
Table 3—Reliabilities of measurement for the decision tree example.

Likelihood Function p(**says**|**is**)		Given the set of model parameters **is** … $\boldsymbol{x}_1$	$\boldsymbol{x}_2$	$\boldsymbol{x}_3$
… the measurement at time t **says**	"Low" ($q_{o,L}^t$)	4/5	4/5	1/5
	"High" ($q_{o,H}^t$)	1/5	1/5	4/5
	Total	1	1	1

Table 4—Likelihood function for the decision tree example.

Fig. 6a shows the probability tree with the prior and likelihoods in the assessed form. Its corresponding inferential form (flipped tree) is shown in Fig. 6b, where the preposterior and posteriors are listed.

(a) Assessed Form: Information we have



(b) Inferential Form: Information we need

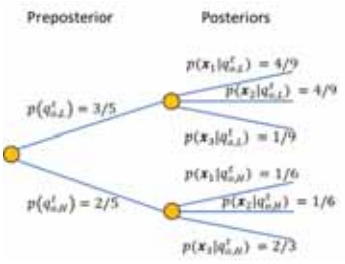


Fig. 6—The uncertainty trees in (a) assessed form, and (b) inferential form for the decision tree example.

Because information cannot add value if it is not material, we must define the decision context before the VOI can be calculated. This example will assume that we can choose one of three production strategies, denoted by $\boldsymbol{a}_1$, $\boldsymbol{a}_2$ and $\boldsymbol{a}_3$, that the strategies are the same for the time interval through t, and that they differ for the remaining period. Thus, we have two options: (1) to choose from $\boldsymbol{a}_1$, $\boldsymbol{a}_2$ and $\boldsymbol{a}_3$ based on the prior (i.e. the case without information), and (2) to measure the oil production rate at time t and then, to choose from $\boldsymbol{a}_1$, $\boldsymbol{a}_2$ and $\boldsymbol{a}_3$ based on the measurement (i.e. the case with information). Let $v(\boldsymbol{x}, \boldsymbol{a})$ be the NPV of the project for the period from time t to the end of the production life-cycle, calculated with realization $\boldsymbol{x}$ and decision $\boldsymbol{a}$. The NPVs for different alternatives and realizations are listed in **Table 5**. The decision trees for the cases with and without information are shown in

Fig. 7, where the triangles represent payoffs or values. These trees show only the branches for the optimal decision, i.e., the alternative that maximizes the expected NPV. Recall that the VOI is assessed before any information is gathered. The decision tree in Fig. 7b includes the observation uncertainty to the left of the decision nodes.

$v_{ij} = NPV(\boldsymbol{x}_i, \boldsymbol{a}_j)$	$\boldsymbol{x}_1$	$\boldsymbol{x}_2$	$\boldsymbol{x}_3$
$\boldsymbol{a}_1$	89	56	49
$\boldsymbol{a}_2$	73	64	65
$\boldsymbol{a}_3$	75	6	88

Table 5—NPVs for different alternatives and realizations [$ millions] for the decision tree example.

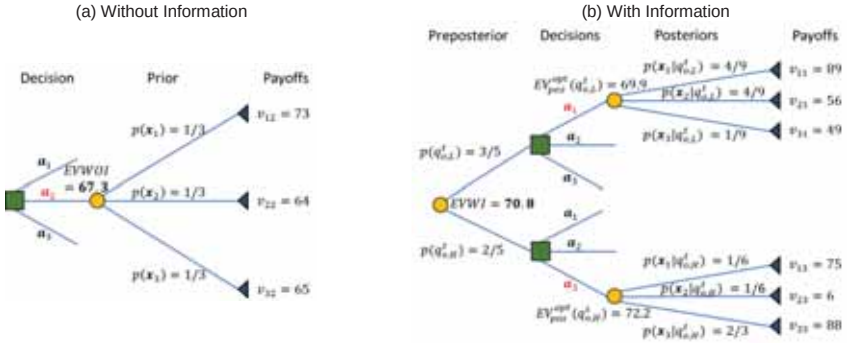


Fig. 7—The decision trees for the cases (a) with information, and (b) without information for the decision tree example.

As shown in Fig. 7a, the DWOI is $\boldsymbol{a}_2$, the EVWOI is \$67.3 million, and the EVWI is \$70.8 million. The EVWI can be calculated using either F1 (Eq. 9):

$$\begin{aligned} EVWI &= p\left(q_{o,L}^t\right) \cdot \left(v(\boldsymbol{x}_1, \boldsymbol{a}_1) p\left(\boldsymbol{x}_1 \middle| q_{o,L}^t\right) + v(\boldsymbol{x}_2, \boldsymbol{a}_1) p\left(\boldsymbol{x}_2 \middle| q_{o,L}^t\right) + v(\boldsymbol{x}_3, \boldsymbol{a}_1) p\left(\boldsymbol{x}_3 \middle| q_{o,L}^t\right)\right) + \\ &\quad p\left(q_{o,H}^t\right) \cdot \left(v(\boldsymbol{x}_1, \boldsymbol{a}_3) p\left(\boldsymbol{x}_1 \middle| q_{o,H}^t\right) + v(\boldsymbol{x}_2, \boldsymbol{a}_3) p\left(\boldsymbol{x}_2 \middle| q_{o,H}^t\right) + v(\boldsymbol{x}_3, \boldsymbol{a}_3) p\left(\boldsymbol{x}_3 \middle| q_{o,H}^t\right)\right) \\ &= \frac{3}{5} \cdot \left(89 \cdot \frac{4}{9} + 56 \cdot \frac{4}{9} + 49 \cdot \frac{1}{9}\right) + \frac{2}{5} \cdot \left(75 \cdot \frac{1}{6} + 6 \cdot \frac{1}{6} + 88 \cdot \frac{2}{3}\right) = 70.8 \end{aligned}$$

or F2 (Eq. 10):

$$\begin{aligned} EVWI &= v(\boldsymbol{x}_1, \boldsymbol{a}_1) p\left(q_{o,L}^t \middle| \boldsymbol{x}_1\right) p(\boldsymbol{x}_1) + v(\boldsymbol{x}_2, \boldsymbol{a}_1) p\left(q_{o,L}^t \middle| \boldsymbol{x}_2\right) p(\boldsymbol{x}_2) + v(\boldsymbol{x}_3, \boldsymbol{a}_1) p\left(q_{o,L}^t \middle| \boldsymbol{x}_3\right) p(\boldsymbol{x}_3) + \\ &\quad v(\boldsymbol{x}_1, \boldsymbol{a}_3) p\left(q_{o,H}^t \middle| \boldsymbol{x}_1\right) p(\boldsymbol{x}_1) + v(\boldsymbol{x}_2, \boldsymbol{a}_3) p\left(q_{o,H}^t \middle| \boldsymbol{x}_2\right) p(\boldsymbol{x}_2) + v(\boldsymbol{x}_3, \boldsymbol{a}_3) p\left(q_{o,H}^t \middle| \boldsymbol{x}_3\right) p(\boldsymbol{x}_3) \\ &= 89 \cdot \frac{4}{5} \cdot \frac{1}{3} + 56 \cdot \frac{4}{5} \cdot \frac{1}{3} + 49 \cdot \frac{1}{5} \cdot \frac{1}{3} + 75 \cdot \frac{1}{5} \cdot \frac{1}{3} + 6 \cdot \frac{1}{5} \cdot \frac{1}{3} + 88 \cdot \frac{4}{5} \cdot \frac{1}{3} = 70.8 \end{aligned}$$

The VOI is \$70.8 – \$67.3 = \$3.5 million. If the cost of measurement is greater than \$3.5 million, the measurement of oil production rate should not be conducted; otherwise, the oil production rate should be measured. The value increment of \$3.5 million comes from the possible change in DWI. The optimal decision is alternative $\boldsymbol{a}_2$ for the case without information, whereas it is $\boldsymbol{a}_1$ if the measurement says "Low" ($q_{o,L}^t$) or $\boldsymbol{a}_3$ if the measurement says "High" ($q_{o,H}^t$). If irrespective of the information, $\boldsymbol{a}_2$ is retained for the decision nodes in Fig. 7b, this will yield $EVWI = EVWOI$ and thus $VOI = 0$, which means that the information has no value if it cannot change our decision.

The VOI of the case without noise in q_o^t measurement (i.e., $p\left(q_{o,L}^t \middle| \tilde{q}_{o,L}^t\right) = p\left(q_{o,H}^t \middle| \tilde{q}_{o,H}^t\right) = 1$, else 0) is \$10.4 million. The VOI of the case with perfect information (VOPI) is \$13.0 million. The VOI increment from \$3.5 million to \$10.4 million is due to the elimination of measurement noise. The difference between the VOI without measurement noise (\$10.4 million) and VOPI (\$13.0 million) is because having q_o^t data does not provide perfect

information about the future production even though the q_o^t data have no noise. This imperfection can be ascribed to the non-uniqueness of inverse modelling. If a given dataset mapped to only one set of model parameters (i.e., $p(\boldsymbol{x}_i|\boldsymbol{y}_i) = 1$ for all i), then the dataset would provide perfect information.

Monte Carlo Example with a Static Model

Assessing the VOI for the example above is straightforward because it is a discretized case having a single uncertainty with only three possible outcomes. In practice, many hydrocarbon production relevant decision making contexts involve continuous probability distributions (e.g., an uncertain permeability field can be a multivariate Gaussian distribution). Monte Carlo method uses a number of equi-probable realizations to represent a continuous probability distribution. A disadvantage of Monte Carlo method is that it can result in sampling errors, which could for example cause the EVWIs calculated using F1 (Eq. 23) and F2 (Eq. 24) to differ. This section presents an example where a simple static model is used so that the VOI can be easily calculated analytically. This VOI will then be compared with those calculated using Barros's workflow, F1, and F2.

Consider that the distinction of interest x is the revenue of a project and that the observable distinction y is the information on the revenue. The relationship between x and y is

$$y = x + \varepsilon \tag{25}$$

where x is normally distributed with mean $\bar{x}$ and standard deviation σ_x, and ε is normally distributed with zero mean and standard deviation σ_ε: the prior $p(x) \sim N(\bar{x}, \sigma_x)$ and the likelihood $p(y|x) \sim N(x, \sigma_\varepsilon)$. The error term in Eq. 25 describes the reliability of the information. Using the relationships for the normal conjugate prior, the preposterior is $p(y) \sim N(\bar{x}, \sqrt{\sigma_x^2 + \sigma_\varepsilon^2})$ and the posterior is $p(x|y) \sim N(\bar{x} + (y - \bar{x})\frac{\sigma_x^2}{\sigma_x^2+\sigma_\varepsilon^2}, \sqrt{1/(\frac{1}{\sigma_x^2} + \frac{1}{\sigma_\varepsilon^2})})$. The decision a to be made is whether the project will be conducted; $a = 1$ and $a = 0$ correspond to "conduct" and "not conduct," respectively. If the expected revenue is positive, the project should be conducted; otherwise, it should not. Thus, the value function is $v(x, a) = ax$. The DM wants to know whether the information should be obtained against a cost of 1 (the monetary unit can be arbitrary here).

The VOI for this case can be calculated analytically:

$$VOI = \int_y \max\{0, \bar{x} + (y - \bar{x})\frac{\sigma_x^2}{\sigma_x^2 + \sigma_\varepsilon^2}\}\frac{1}{\sqrt{2(\sigma_x^2 + \sigma_\varepsilon^2)\pi}} e^{-\frac{(y-\bar{x})^2}{2(\sigma_x^2+\sigma_\varepsilon^2)}} dy - \max\{0, \bar{x}\}. \tag{26}$$

With $\bar{x} = 3$, $\sigma_x = 10$, and $\sigma_\varepsilon = 5$, the VOI is 2.267. Because the cost of information is lower than the VOI, the DM should buy the information.

For comparison, we use Monte Carlo method to calculate the VOI. First, we sample n_e realizations from the prior $p(x)$ to get the prior ensemble $[x_1, x_2, \ldots, x_{n_e}]$, with the probability of each x_i being $1/n_e$. Second, the DWOI is obtained using Eq. 20, and the EVWOI is calculated using Eq. 21. Third, for each x_i, we sample one realization from the likelihood $p(y|x_i)$ to get the preposterior ensemble $[y_1, y_2, \ldots, y_{n_e}]$, with the probability of each y_i being $1/n_e$ as well as an ensemble of (x, y)-pairs, $[(x_1, y_1), (x_2, y_2), \ldots, (x_{n_e}, y_{n_e})]$, with the probability of each (x, y)-pair being $1/n_e$ (the 1:1 ratio mentioned earlier applies here). Fourth, we update the prior ensemble using EnKF for each y_i, yielding the posterior ensemble $[\hat{x}_1|y_i, \hat{x}_2|y_i, \ldots, \hat{x}_{n_e}|y_i]$, and we identify the DWI using Eq. 22; this results in n_e posterior ensembles and n_e DWIs. Fifth, the EVWI can be calculated using either F1 (Eq. 23) or F2 (Eq. 24). Finally, we calculate the VOI using Eqs. 1 and 2.

We first test Barros's workflow, F1 and F2 using a large ensemble size of 10,000. The results are listed in **Table 6**, where Barros's workflow gives a result almost identical to that of F2. This is because the impact of excluding a synthetic truth (i.e., an x_i) from the prior ensemble is very small when the ensemble size is large. The VOIs calculated using F1 and F2 are slightly different, for a reason to be discussed later. All three methods estimate the VOI with an error smaller than 1%.

Method	VOI	Error
Analytical	2.267	-
Barros's Workflow	2.283	0.71%
Formulation 1	2.246	0.93%
Formulation 2	2.283	0.71%

Table 6—VOI estimate with $\bar{x} = 3$, $\sigma_x = 10$,$\sigma_\varepsilon = 5$, and an ensemble size of 10,000 for the Monte Carlo example with a static model.

Then, a small ensemble size of 50 is used. The VOI calculation is repeated 10,000 times for each method. The statistics of the VOI estimates are listed in **Table 7**, and the probability density functions (PDFs) are plotted in **Fig. 8**.[14] Both F1 and F2 are shown to produce a VOI estimate with an error smaller than 0.5% on average. F2 leads to a smaller standard deviation (SD) of the VOI estimate than does F1. This indicates that F2 is more stable than F1 with respect to Monte Carlo sampling for small ensemble size because of the former's sole use of the prior ensemble for calculating EVWI, thereby excluding from that calculation the sampling errors associated with the preposterior and posterior ensembles. In contrast, F1 uses both the preposterior and posterior ensembles for the EVWI calculation, thereby propagating sampling errors from the preposterior and posterior ensembles. For Barros's workflow, the resulting range and SD of VOI estimate are larger than in the other two methods, and the PDF includes a long tail to the right as shown in Fig. 8, resulting in a larger error for the average VOI estimate. This long tail is induced by the exclusion of the synthetic truths from the prior ensemble. This suggests that the impact of excluding the synthetic truths from the prior ensemble can be large when the ensemble size is small. In reservoir management, small ensemble size is usually used for EnKF, typically 100, as the computational costs increase with the ensemble size.

Method	Average	Error	SD	[Min, Max]
Barros's Workflow	2.323	2.47%	0.933	[0, 10.421]
Formulation 1	2.257	0.44%	0.776	[0, 5.497]
Formulation 2	2.259	0.35%	0.695	[0, 5.279]

Table 7—VOI estimate with $\bar{x} = 3$, $\sigma_x = 10$,$\sigma_\varepsilon = 5$, and an ensemble size of 50 (repeated 10,000 times) for the Monte Carlo example with a static model.

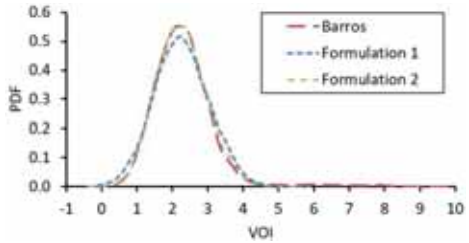


Fig. 8—The PDF of VOI estimate for the Monte Carlo example with a static model.

This example has demonstrated that the three methods give almost equivalent results for large ensembles. However, these methods can lead to quite different results when the ensemble size is small. As there are no significant computational increases resulting from F2 (Eq. 24) for calculating the VOI from HM when Monte Carlo is used, this would be our method of choice.

[14] The uncertainty in the VOI estimate is due to the sampling error associated with Monte Carlo method. As discussed earlier, the VOI itself is not a distribution but the difference between two expected values.

Monte Carlo Example with a Toy Dynamic Model

This section illustrates how the VOI from HM is calculated. We adopt the decline curve model used by Barros et al. (2016a) for modeling oil and water production. The predicted oil and water production rates, $\tilde{q}_o$ and $\tilde{q}_{wp}$, respectively, are functions of time t and a control (decision) variable a,:

$$\tilde{q}_o(a,t) = \left(\tilde{q}_{o,ini} + c_1 a\right) e^{-t/(\beta + a/c_2)} \tag{27}$$

$$\tilde{q}_{wp}(a,t) = H\left[t - t_{bt}\left(1 - \frac{\bar{a}}{c_3}\right)\right]\left(\tilde{q}_{wp,\infty} + a\right)\left[1 - e^{-(t - t_{bt}(1-\bar{a}/c_3))/(c_4\beta - a/c_5)}\right] \tag{28}$$

where $\tilde{q}_{o,ini}$ is the initial oil production rate for $a = 0$; $\tilde{q}_{wp,\infty}$ is the ultimate water production rate for $a = 0$; t_{bt} is the water breakthrough time for $a = 0$; β, c_1, c_2, c_3, c_4 and c_5 are coefficients; $\bar{a}$ is the average of a in all the control intervals before t_{bt}; and $H[\cdot]$ is the Heaviside step function. The water injection rate $\tilde{q}_{wi}$ is the sum of $\tilde{q}_o$ and $\tilde{q}_{wp}$. The constants and uncertain model parameters are listed in **Table 8**. The production life-cycle is 80 months, and $t \in [1, 2, \ldots, 80]$. We include eight control intervals, and the single control variable a is thus replaced by a control vector $\boldsymbol{a} = [a_{1:10}, a_{11:20}, \ldots, a_{71:80}]^T$. That is, when a decision is made at time $t \in [0, 10, \ldots, 70]$, it cannot be changed over the next 10 time intervals. The upper and lower bounds of a are 10 bbl/mo and 50 bbl/mo, respectively. The observations are oil and water production rates at a certain time. In the form of Eqs. 19 and 17, we have the distinctions of interest $\boldsymbol{x} = [\tilde{q}_{o,ini}, \beta, \tilde{q}_{wp,\infty}, t_{bt}]^T$, predicted observations $\tilde{\boldsymbol{y}} = [\tilde{q}_o, \tilde{q}_{wp}]^T$, observations (with noise) $\boldsymbol{y} = [q_o, q_{wp}]^T = \tilde{\boldsymbol{y}} + \boldsymbol{\varepsilon}_y$, and a forward model to time t, $f_t(\boldsymbol{x}, \boldsymbol{a})$ described by Eqs. 27 and 28.

Constant Parameters				Uncertain Parameters		
c_1	0.1	[-]		$\tilde{q}_{o,ini}$	N(100, 8)	[bbl/mo]
c_2	4	[bbl/mo^2]		β	N(30.5, 3.67)	[mo]
c_3	150	[bbl/mo]		$\tilde{q}_{wp,\infty}$	N(132, 6)	[bbl/mo]
c_4	2	[-]		t_{bt}	N(32, 6)	[mo]
c_5	1.33	[bbl/mo^2]				
P_o	70	[$/bbl]				
P_{wp}	10	[$/bbl]				
P_{wi}	10	[$/bbl]				
b	0.1	[-]				
τ	8	[mo]				

Table 8—Constant and uncertain parameters for the decline curve model.

We use a Gaussian distribution for the reliability of measurements, i.e., $p(\boldsymbol{\varepsilon}_y|\tilde{\boldsymbol{y}}) \sim N(\boldsymbol{0}, \mathbf{C}_y)$ where $\mathbf{C}_y = \begin{bmatrix} \sigma_o^2 & 0 \\ 0 & \sigma_{wp}^2 \end{bmatrix}$, with σ_o and σ_{wp} being the SDs of oil and water rate measurements, respectively. For this example, we assume that the SD σ is a percentage Pct_{σ_ε} of the EV of the prior predicted observation (i.e., $\sigma = Pct_{\sigma_\varepsilon} \cdot \frac{1}{n_e}\sum_{i=1}^{n_e} \tilde{q}_i$) for both the oil and water production rates, and set $Pct_{\sigma_\varepsilon} = 5\%$. Model noise is ignored.

An ensemble size of 100 is used. Realizations are sampled from the prior $p(\boldsymbol{x})$ (Gaussian distributions listed in Table 8) to form the prior ensemble $[\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_{100}]$. RO is performed over the prior ensemble to identify the optimal strategy $\boldsymbol{a}_{pri}^{opt}$ (i.e., DWOI) that optimizes the EV over the prior ensemble (Eq. 20).

The VOI calculation will be for information gathered at time t. For a sequential decision making context, the information gathered at time t will affect the decisions made only after the time when we gather the information. Thus, we care about the value only for the period where decisions are changeable, and the value function should be defined as the NPV for that period. For example, decisions that can be affected by information gathered at time $t = 55$, are $\{a_{61:70}, a_{71:80}\}$, so we consider the value function for only the period after time $t = 60$. All monetary values are discounted to time 0.

We first do forward modeling to time t with the prior ensemble members and DWOI to generate the realizations of the predicted observations $[\tilde{\boldsymbol{y}}_1^t = f_t(\boldsymbol{x}_1, \boldsymbol{a}_{pri}^{opt}), \tilde{\boldsymbol{y}}_2^t = f_t(\boldsymbol{x}_2, \boldsymbol{a}_{pri}^{opt}), \ldots, \tilde{\boldsymbol{y}}_{100}^t = f_t(\boldsymbol{x}_{100}, \boldsymbol{a}_{pri}^{opt})]$ and then sample one realization of measurement noise from $p(\boldsymbol{\varepsilon}_y|\tilde{\boldsymbol{y}}_i^t)$ and add it to $\tilde{\boldsymbol{y}}_i^t$ to obtain $[\boldsymbol{y}_1^t, \boldsymbol{y}_2^t, \ldots, \boldsymbol{y}_{100}^t]$ which is the

preposterior ensemble. For each $\boldsymbol{y}_i^t$, we use EnKF to update the prior ensemble to obtain the posterior ensemble $[\hat{\boldsymbol{x}}_1|\boldsymbol{y}_i^t, \hat{\boldsymbol{x}}_2|\boldsymbol{y}_i^t, \ldots, \hat{\boldsymbol{x}}_{100}|\boldsymbol{y}_i^t]$ and perform RO over the posterior ensemble to identify the optimal strategy after time t for the posterior ensemble. Again, using information gathered at time $t = 55$ as an example, the DWI after time $t = 55$ is $\left(\boldsymbol{a}_{pos}^{opt}\right)_{61:80} = \left[\left(a_{pos}^{opt}\right)_{61:70}, \left(a_{pos}^{opt}\right)_{71:80}\right]^T$.

The EVWOI can then be calculated using Eq. 21, and the EVWI using F1 (Eq. 23) or F2 (Eq. 24). The VOI is calculated using Eqs. 1 and 2.

Both F1 and F2 are used to calculate the Δ defined in Eq. 2. From Eq. 1, $VOI = \max\{\Delta, 0\}$. The calculation of Δ at time $t = 10$ using our proposed workflow is repeated 100 times for each formulation. The ensembles used in any two runs are different. The statistics of the Δ estimates are listed in **Table 9**.[15] As was found in the previous example, the average of the estimates calculated using F1 is close to that of F2; however, the SD of the estimates calculated using F1 is much larger than that of F2, due to the small ensemble size. Different VOI estimates can lead to different decisions. Taking one of the 100 runs as an example, we get a VOI estimate of \$621.4 using F1 and \$107.6 using F2, even though the same set of realizations is used; if the cost of gathering data and conducting model calibration is \$150, the decision made based on the VOI calculated using F1 is "to gather data and conduct HM" but the decision made based on that using F2 is "to not gather data". In practice, we won't repeat the calculation of Δ because we want to minimize the computational cost. Thus, the formulation which has the higher chance to produce a good estimate (i.e. smaller SD in estimate) should be used.

Method	Average	SD	[Min, Max]
Formulation 1	64.7	480.3	[-1181.9, 1049.6]
Formulation 2	58.4	31.5	[-64.6, 148.8]

Table 9—Statistics of the Δ estimates (repeated 100 times) for the Monte Carlo example with a dynamic model.

We use F2 to calculate the VOI at different times using the same prior ensemble, to answer the following questions: (1) If we can conduct the measurements of oil and water production rates only once at some time during the production life-cycle (i.e., we can get only a single value of $\boldsymbol{y} = [q_o, q_{wp}]^T$), when should this be done? (2) How much value can be created by the measurements? Repeating the VOI calculation for data gathered at different times yields the graph shown in **Fig. 9**. It indicates that the value of imperfect information with $Pct_{\sigma_\varepsilon} = 5\%$ peaks at time 30 with a value of \$217.5. If the cost of gathering data and conducting model calibration is less than \$217.5, we should gather data and conduct model calibration at time 30; otherwise, we should not gather data at all. The value that can be created by the measurements is the difference between its VOI and cost. For example, if the cost is \$118, then the created value is \$217.5 – \$118 = \$99.5. The VOPI is the upper bound of VOI. Not surprisingly, the graph indicates that if we have access to perfect information, we should gather data as soon as possible (between time 0 and 10) if the information can be obtained at a cost of no more than \$674.3. There are sharp drops in both VOI with measurement noise and VOPI at time 11, 21, 31 and 41. This is because one more decision in the series of decisions of production strategy cannot be affected by the data gathered at these times (Barros et al. 2016a). For example, the injection rate decision for the period from time 20 to 30 is made at time 20, so the data gathered before time 20 can affect this decision whilst the data gathered after time 20 cannot affect this decision but only the decisions made after that.

[15] See footnote 14.

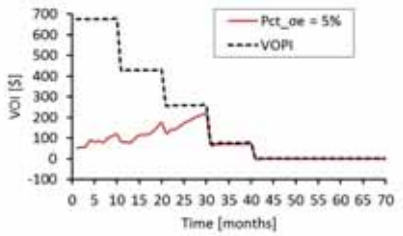


Fig. 9—VOI with $Pct_{\sigma_\varepsilon} = 5\%$ and VOPI at different times for the Monte Carlo example with a dynamic model.

How would different levels of measurement noise impact VOI? To investigate this, we conducted a sensitivity analysis of the measurement noise (i.e., the value of Pct_{σ_ε}) on VOI. VOI at different times is shown for three values of Pct_{σ_ε} (1%, 5%, and 9%) in **Fig. 10**. The smoothness of the curve decreases with increasing measurement noise. We believe this is because of the sampling errors associated with sampling the measurement noise. As the measurement noise increases, larger sampling errors will be introduced when the measurement noise is sampled using the same ensemble size for Monte Carlo method, resulting in larger errors in VOI estimate. Fig. 10 shows that VOI increases as Pct_{σ_ε} decreases, and that the VOI peaks at time 30 no matter how large Pct_{σ_ε} is. This is because water break-through will have occurred over the majority of the realizations at time 30, and the water rate data gathered at this time provides relevant and material information on the future water production. Thus, we focus on VOI at time 30 and calculate it for different values of Pct_{σ_ε}. The results, shown in **Fig. 11**, indicate that as measurement noise increases, VOI approaches 0 (i.e., the information is immaterial). Because the measurement noise tends to be quite small in practice, we consider Pct_{σ_ε} only between 0% and 22% for further analysis. The VOI estimates are indicated by blue dots in **Fig. 12,** and the red dashed line is the second-order polynomial trendline corresponding to the VOI estimates. The cost of reducing measurement noise is represented by the green solid line in Fig. 12, showing that VOI and the cost intersect at $Pct_{\sigma_\varepsilon} = 1.2\%$. For $Pct_{\sigma_\varepsilon} < 1.2\%$, the measurement cost is greater than the VOI, indicating that paying for a more accurate measure does not add value. For $Pct_{\sigma_\varepsilon} > 1.2\%$, the measurement cost is lower than the VOI. The difference between VOI and cost is shown in **Fig. 13**. The information value increases as Pct_{σ_ε} increases until Pct_{σ_ε} it reaches 9.1%, corresponding to an optimal value of $105.8 (the red dot in Fig. 13). The sensitivity analysis indicates that a more accurate measure does not necessarily create more value and that a tradeoff exists between measurement accuracy and costs.

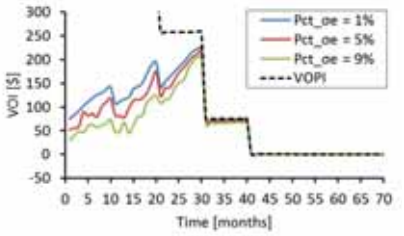


Fig. 10—VOI at different times for different levels of measurement noise for the Monte Carlo example with a dynamic model.

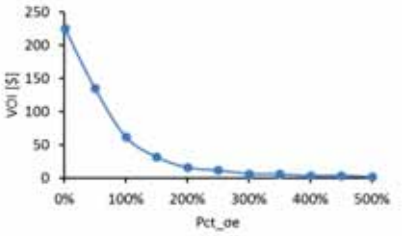


Fig. 11—VOI at time 30 for different levels of measurement noise for the Monte Carlo example with a dynamic model.

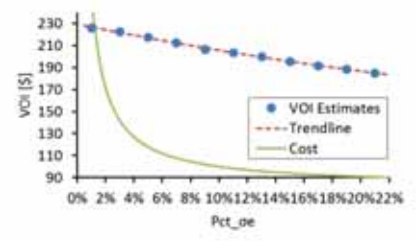


Fig. 12—VOI at time 30 and measurement cost for different levels of measurement noise for the Monte Carlo example with a dynamic model.

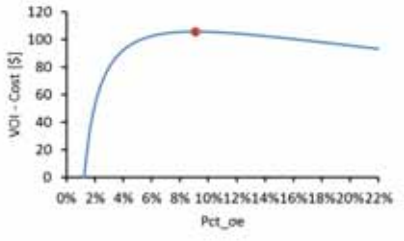


Fig. 13—VOI at time 30 less measurement cost for different levels of measurement noise for the Monte Carlo example with a dynamic model.

Monte Carlo Example with a Reservoir Simulation Model

We use our proposed workflow to estimate VOI from HM for a more realistic case where a reservoir simulation model is involved. The reservoir simulation model (a synthetic 2D model) is shown in **Fig. 14**. Eight injectors are placed in line at one end of the reservoir and eight producers in line at the opposite end. The permeability field is uncertain and its prior uncertainty is described by a multivariate normal distribution. The prior ensemble consists of 50 realizations sampled from the prior distribution, 3 of which are illustrated in **Fig. 15**.

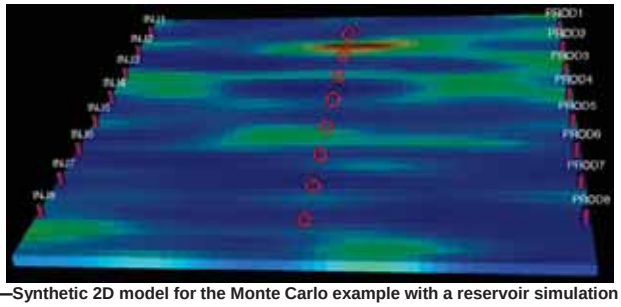


Fig. 14—Synthetic 2D model for the Monte Carlo example with a reservoir simulation model.

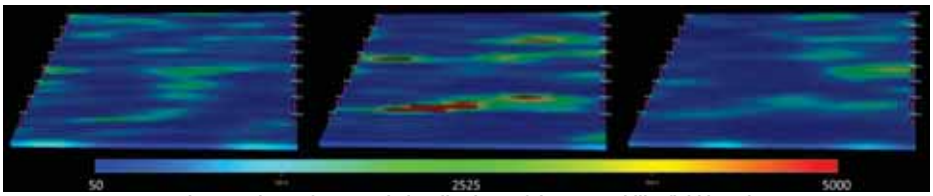


Fig. 15—Three of 50 sampled realizations of the permeability field in md.

We assume that the periods of water pre-flushing, polymer injection and water post-flushing have already been set: 1,200 days' water pre-flushing followed by 1,800 days' polymer injection and then followed by 1,200 days' water post-flushing. The injection rate is fixed at 15 m^3/day for each injector and the producer bottom hole pressure is fixed at 200 bars for each producer. At the end of water pre-flushing, the DM needs to decide what polymer concentration should be used in each injector for the whole period of polymer injection. Thus, there are totally 8 control variables. The polymer concentration is bounded between 0 kg/m^3 and 6.5 kg/m^3.

Oil and water production rates in each producer will be measured every 240 days during the water pre-flushing period. These data can be used for HM and inform the decision on polymer concentration. In addition to oil and water production rate data, the DM is considering to gather another type of data: water saturations in the grid-blocks circled by the red rings in Fig. 14 by placing saturation measuring devices at these locations. If the devices are placed, water saturations in these grid-blocks will be measured every 240 days as well. The questions are: What is the value of additional water saturation data? Should the saturation measuring devices be placed?

The value function is defined as the NPV with revenue from oil production and costs for water production, water injection and polymer injection, i.e.,

$$v(\boldsymbol{x},\boldsymbol{a}) = NPV(\boldsymbol{x},\boldsymbol{a}) = \sum_{k=1}^{n_T} \frac{(q_o^k(\boldsymbol{x},\boldsymbol{a})P_o - q_{wp}^k(\boldsymbol{x},\boldsymbol{a})P_{wp} - q_{wi}^k(\boldsymbol{x},\boldsymbol{a})P_{wi} - q_{ci}^k(\boldsymbol{x},\boldsymbol{a})P_{ci})\Delta t_k}{(1+b)^{t_k/\tau}} \qquad (29)$$

where q_{ci}^k is the total polymer injection rate over time step length Δt_k and P_{ci} is the polymer injection cost. The parameter values used in this specific example are listed in **Table 10**.

P_o	220	\$/m^3
P_{wp}	47.5	\$/m^3
P_{wi}	12.5	\$/m^3
P_{ci}	12	\$/kg
b	8%	–
τ	365	days
Δt	30	days

Table 10— Parameter values for NPV calculation in the Monte Carlo example with a reservoir simulation model.

The measurement noise of oil and water production rate is assumed to be normally distributed with zero mean and a SD of 1 m3/day, and the measurement noise of water saturation to be normally distributed with zeros mean and a SD of 0.01. Model noise is ignored.

VOI is calculated using our proposed workflow combined with EnKF and EnOpt. The resulting VOIs for rate data only, saturation data only, and both rate and saturation data, and VOPI are listed in **Table 11**.

Data Type	**Rate**	**Saturation**	**Rate and Saturation**	**Perfect Information**
VOI [\$·10^5]	3.10	3.36	4.11	6.12

Table 11—VOI estimates for Monte Carlo Example with a Reservoir Simulation Model.

It can be see that the values of rate data and saturation data are very close to each other. Although the saturation data has a value of \$3.36·10^5 when it is considered individually, it adds a value of only \$1.01·10^5 (= \$4.11·10^5 – \$3.10·10^5) in addition to rate data. This indicates that VOI is not additive (Samson et al. 1989). The VOI analysis informs the DM that he/she should not gather saturation data in addition to rate data if the cost of placing saturation measuring devices is greater than \$1.01·10^5.

Discussion and Conclusions

This paper has introduced a workflow for calculating VOI from HM and that is consistent with the original definition of VOI in the decision sciences. The workflow employs ensemble-based methods and Monte Carlo sampling. Two formulations (F1 and F2), both of which are consistent with VOI analysis, can be used to calculate the EVWI. In the limit where the ensemble size approaches infinity, the EVWIs calculated using these two formulations are identical. However, the numerical results may be different, and consequently different VOI estimates will be calculated when the ensemble size is small. The VOI calculated using F2 is less sensitive with regard to ensemble size than is F1. F2 is thus our method of choice for calculating the VOI from HM using Monte Carlo method.

F2 was compared with Barros's workflow. We noted that Barros's workflow, which relies on twin experiments, is inconsistent with VOI analysis because VOI is a value, not a distribution. Nevertheless, two small modifications to Barros's workflow will make it consistent with VOI analysis and consequently, equivalent to F2.

A decision tree was used to illustrate the implementation of the general workflow for VOI analysis in HM contexts. Bayesian inference was explicitly performed by applying Bayes' rule. The use of decision trees provides clarity in communicating the detailed steps for VOI analysis.

We conducted a sensitivity analysis of the impact of different levels of measurement noise. Due to the non-uniqueness inherent in the HM inverse problem, a zero measurement noise will not result in perfect information about future production. As the measurement noise increases, VOI is reduced; and once measurement noise exceeds a threshold, the gathered production data will be immaterial and thus worthless. The sensitivity analysis thus

confirmed that applying the ensemble-based approach for HM in a VOI framework is consistent with classical VOI analysis. The sensitivity analysis also illustrated that the value of a measure (its VOI less its cost) does not necessarily increase as measurement accuracy is improved.

Two Monte Carlo examples demonstrated the application of F2 for VOI assessment. The first example analyzed the optimal timing of conducting data measurement, and the second example investigated the value of obtaining additional type of data.

Although we used EnKF for HM (Bayesian updating) and EnOpt for RO (decision making), the proposed VOI workflow could employ any Bayesian updating method or optimization method. The performance of the calibration method can affect the accuracy of VOI assessment. EnKF is derived from the Kalman filter, which is designed for Gaussian distributions. Thus, any inherent non-Gaussianity (e.g., value truncation) can introduce errors in Bayesian updating and consequent errors in the VOI calculation. When VOI is calculated using Monte Carlo method, the number of realizations used to represent the probability distributions affects the accuracy of the VOI calculation. We used an ensemble size of 100, which has been regarded as a reasonable compromise between fairly representing the probability distribution and limiting computational costs for EnKF. However, given that Bickel (2012) showed that more than 1,000 realizations are required for a good estimate of VOI in a case where only one uncertain quantity is involved, an ensemble size of 100 is unlikely to provide an accurate assessment of VOI from HM. The accuracy of VOI estimate using EnKF and EnOpt should be further investigated in future reseach.

A significant constraint on the routine use of VOI in the HM context is its high computational cost (Barros et al. 2016a). An ensemble size of n_e, would entail n_e ensemble-based HM and optimizations over all n_e realizations. If a production model needs hours to run, as in many real cases, the CPU time for assessing the VOI for observations obtained at a certain time will be years. Thus, using parallel computing for VOI calculation can significantly reduce the waiting time because both the EnKF and RO parts of the workflow can be done in parallel across realizations. If parallel computing is unavailable, one possible approach is to use fast proxy models (e.g., an upscaled model or a decline curve model) for production modeling. If this can reduce model run time from hours to seconds, the CPU time would be reduced from years to hours, making it acceptable for a practical setting. Another possible solution is to reduce the number of realizations using an efficient discretization method (Bratvold and Thomas 2014). This works well when only a few uncertain quantities are involved. However, this approach can also be computationally very expensive when numerous uncertain quantities are involved because the number of discretization points increases exponentially with the number of uncertain quantities. Barros et al. (2016b) reduced the number of realization needed using clustering techniques to select representative models from an ensemble.

The computational cost is even higher for a sequential decision making context. Because of this, we only illustrated VOI assessment for the case with only one decision point. Although Barros et al. (2015b) has extended their workflow to incorporate sequential decisions using the closed loop reservoir management (CLRM) approach, we argue that CLRM is a myopic or naïve decision policy where the uncertainties associated with current available data is considered but the uncertainties associated with future data is not. **Fig. 16** illustrates the decision tree representation for CLRM. The advantage of CLRM is that it greatly simplifies the structure of a sequential decision making problem and consequently, requires less computational cost for solving the problem. Its significant drawback is that it does not reflect the full structure of a sequential decision making problem and thus, may lead to a sub-optimal production strategy. **Fig. 17** illustrates the decision tree representation for the full structure of a sequential decision making problem. The full structure explicitly considers both the uncertainties associated with current available data and that with future data, i.e., it allows for learning over time. Therefore, the current decision does not depend on only the uncertainties that a decision maker have learnt so far but also the uncertainties that the decision maker will learn in the future. Solving the fully structured decision tree gives the optimal production strategy. Per the strict definition of VOI, it is referred exclusively to that associated with the optimal decision policy (i.e. the solution of the fully structured decision tree). The VOI assessed using the CLRM approach is a special case and only for the CLRM production strategy. The VOI associated with the CLRM production strategy may be lower than the generally defined VOI (i.e. the VOI associated with the optimal production strategy). Therefore, using the CLRM approach is a simplification of solving a sequential decision making problem but not a replacement. Solving for the optimal production strategy may create more value than the CLRM production strategy but requires more computational cost.

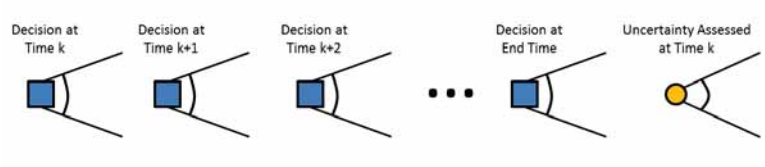


Fig. 16—Decision tree representation for CLRM.

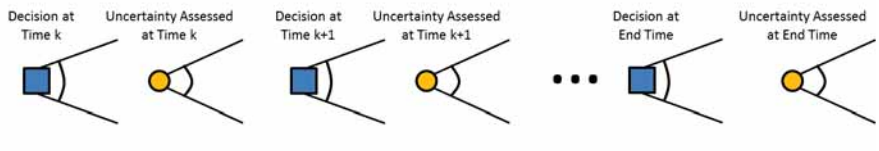


Fig. 17—Decision tree representation for the full structure of a sequential decision making problem.

Another significant constraint is the difficulty in Bayesian inference when a large number of uncertain parameters (e.g., the porosity and permeability in each grid block of a reservoir simulation model) is involved. Although many publications have shown promising results of updating a large number of uncertain parameters using EnKF, it may lead to statistic bias and consequently reduce the accuracy of VOI calculation. Any other Bayesian updating methods, e.g., the ensemble smoother (Skjervheim and Evensen 2011) or Markov-chain Monte Carlo methods (Oliver et al. 1997), can be adopted into our proposed workflow. Thus, a possible direction for further research is to investigate the impact of different updating methods on VOI estimate.

Many producing fields and most new fields gather some types of data (e.g., water, oil, and gas production rate and well bottom-hole pressure) for monitoring and other purposes. VOI analysis plays no role for these types of measurements, because the data already have been or definitely will be gathered. However, VOI analysis is very relevant and useful for other types of costly data, such as 4D seismic surveys.

Although issues remain in the context of developing a broad and deep understanding and routine application of the concept, we remain optimistic about the value of VOI analysis in the O&G industry.

Acknowledgements

The authors acknowledge the support of the Research Council of Norway and industry partners: ConocoPhillips, Skandinavia AS, BP Norge AS, Det Norske Oljeselskap AS, Eni Norge AS, Maersk Oil Norway AS, DONG Energy A/S, Denmark, Statoil Petroleum AS, ENGIE E&P NORGE AS, Lundin Norway AS, Halliburton AS, Schlumberger Norge AS, and Wintershall Norge AS of The National IOR Centre of Norway.

References

Aanonsen, S. I., Nævdal, G., Oliver, D. S. et al. 2009. The Ensemble Kalman Filter in Reservoir Engineering—a Review. *SPE Journal* **14** (3): 393–412. SPE-117274-PA. http://dx.doi.org/10.2118/117274-PA.

Aylor, W. K. Jr. 1999. Measuring the Impact of 3D Seismic on Business Performance. *Journal of Petroleum Technology* **51** (6): 52–56. SPE-56851-JPT. http://dx.doi.org/10.2118/56851-JPT.

Barros, E. G. D., Jansen, J. D., and Van den Hof, P. M. J. 2015a. Value of Information in Parameter Identification and Optimization of Hydrocarbon Reservoirs. *IFAC-PapersOnLine* **48** (6): 229–235. http://dx.doi.org/10.1016/j.ifacol.2015.08.036.

Barros, E. G. D., Leeuwenburgh, O., Van den Hof, P. M. J. et al. 2015b. Value of Multiple Production Measurements and Water Front Tracking in Closed-Loop Reservoir Management. Presented at the SPE Reservoir Characterisation and Simulation Conference and Exhibition, Abu Dhabi, UAE, 14–16 September. SPE-175608-MS. http://dx.doi.org/10.2118/175608-MS.

Barros, E. G. D., Van den Hof, P. M. J., and Jansen, J. D. 2016a. Value of Information in Closed-Loop Reservoir Management. *Computational Geosciences* **20** (3): 737–749. http://dx.doi.org/10.1007/s10596-015-9509-4.

Barros, E. G. D., Yap, F. K., Insuasty, E., Van den Hof, P. M. J., and Jansen, J. D. 2016b. Clustering Techniques for Value-of-information Assessment in Closed-loop Reservoir Management. Presented at ECMOR XV–15th European Conference on the Mathematics of Oil Recovery, Amsterdam, Netherlands, 29 August–1 September. http://dx.doi.org/10.3997/2214-4609.201601858.

Bickel, J. E. 2012. Discretization, Simulation, and the Value of Information. *SPE Economics & Management* **4** (4): 198–203. SPE-145690-PA. http://dx.doi.org/10.2118/145690-PA.

Bratvold, R. B., and Begg, S. 2010. *Making Good Decisions*, first edition. Texas, USA: Society of Petroleum Engineers.

Bratvold, R. B., Bickel, J. E., and Lohne, H. P. 2009. Value of Information in the Oil and Gas Industry: Past, Present, and Future. *SPE Reservoir Evaluation & Engineering* **12** (4): 630–638. SPE-110378-PA. http://dx.doi.org/10.2118/110378-PA.

Bratvold, R. B., and Thomas, P. 2014. Robust Discretization of Continuouos Probability Distributions for Value-of-Information Analysis. Presented at the International Petroleum Technology Conference, Kuala Lumpur, Malaysia, 10–12 December. IPTC-17975-MS. http://dx.doi.org/10.2523/IPTC-17975-MS.

Burgers, G., Jan van Leeuwen, P., and Evensen, G. 1998. Analysis Scheme in the Ensemble Kalman Filter. *Monthly Weather Review* **126** (6): 1719–1724. http://dx.doi.org/10.1175/1520-0493(1998)126<1719:ASITEK>2.0.CO;2.

Chen, Y., Oliver, D. S., and Zhang, D. 2009. Efficient Ensemble-Based Closed-Loop Production Optimization. *SPE Journal* **14** (4): 634–645. SPE-112873-PA. http://dx.doi.org/10.2118/112873-PA.

Clemen, R. T. 1991. *Making Hard Decisions: An Introduction to Decision Analysis*. Boston, USA: PWS-Kent.

Coopersmith, E. M., Burkholder, M. K., and Schulze, J. H. 2006. Value-of-Information Lookbacks—Was the Information You Gathered Really Worth Getting? Presented at the SPE Annual Technical Conference and Exhihition, San Antonio, Texas, USA, 24–27 September. SPE-101540-MS. http://dx.doi.org/10.2118/101540-MS.

Eidsvik, J., Mukerji, T., and Bhattacharjya, D. 2015. *Value of Information in the Earth Sciences: Integrating Spatial Modeling and Decision Analysis*, first edition. Cambridge, United Kingdom: Cambridge University Press.

Eppel, T., and von Winterfeldt, D. 2008. Value-of-Information Analysis for Nuclear Waste Storage Tanks. *Decision Analysis* **5**(3): 157–167. http://dx.doi.org/10.1287/deca.1080.0121.

Evensen, G. 1994. Sequential Data Assimilation with a Nonlinear Quasi-Geostrophic Model Using Monte Carlo Methods to Forecast Error Statistics. *Journal of Geophysical Research: Oceans* **99** (C5): 10143–10162. http://dx.doi.org/10.1029/94JC00572.

Evensen, G. 2009. *Data Assimilation: The Ensemble Kalman Filter*, second edition. Heidelberg, Germany: Springer Science & Business Media.

Grayson, C. J. Jr. 1960. *Decisions under Uncertainty: Drilling Decisions by Oil and Gas Operators*. Boston, Massachusetts: Harvard University Press.

Howard, R. A. 1966. Information Value Theory. *IEEE Transactions on Systems Science and Cybernetics* **2** (1): 22–26. http://dx.doi.org/10.1109/TSSC.1966.300074.

Howard, R. A., and Abbas, A. E. 2016. *Foundations of Decision Analysis*, global edition. Harlow, England: Pearson Education Limited.

Krymskaya, M. V., Hanea, R. G., Jansen, J. D. et al. 2010. Observation Sensitivity in Computer-Assisted History Matching. Presented at the 72nd EAGE Conference and Exhibition incorporating SPE EUROPEC 2010. http://dx.doi.org/10.3997/2214-4609.201400961.

Le, D. H., and Reynolds, A. C. 2014. Optimal Choice of a Surveillance Operation Using Information Theory. *Computational Geosciences* **18**(3–4): 505–518. http://dx.doi.org/10.1007/s10596-014-9401-7.

MATLAB® Software 2014. Version 8.4.0.150421 (R2014b).

Meinhold, R. J., and Singpurwalla, N. D. 1983. Understanding the Kalman Filter. *The American Statistician* **37** (2): 123–127.

Oliver, D. S., Cunha, L. B., and Reynolds, A. C. 1997. Markov Chain Monte Carlo Methods for Conditioning a Permeability Field to Pressure Data. *Mathematical Geology* **29**(1): 61–91. https://doi.org/10.1007/BF02769620.

Raiffa, H., and Schlaifer, R. 1961. *Applied Statistical Decision Theory*. Boston: Division of Research, Graduate School of Business Administration, Harvard University (Reprint).

Schlaifer, R. 1959. *Probability and Statistics for Business Decisions*. New York: McGraw-Hill.

Skjervheim, and Evensen, G. 2011. An Ensemble Smoother for Assisted History Matching. Presented at the SPE Reservoir Simulation Symposium, The Woodlands, Texas, USA, 21–23 Feburary. SPE-141929-MS. https://doi.org/10.2118/141929-MS.

Samson, D., Wirth, A., and Rickard, J. 1989. The Value of Information From Multiple Sources of Uncertainty in Decision Analysis. *European Journal of Operational Research* **39** (3): 254–260. https://doi.org/10.1016/0377-2217(89)90163-x.

Thomas, P., Bratvold, R. B., and Lake, L. W. 2016. Valuing Historical Data in Water Injection Optimization Using

the Capacitance Resistance Model. *Working Paper, University of Stavanger*.

Van Essen, G., Zandvliet, M., Van den Hof, P., Bosgra, O., and Jansen, J. D. 2009. Robust Waterflooding Optimization of Multiple Geological Scenarios. *SPE Journal* **14** (1): 202–210. SPE-102913-PA. https://doi.org/10.2118/102913-PA.

Waggoner, J. R. 2000. Lessons Learned from 4D Projects. *SPE Reservoir Evaluation & Engineering* **3** (4): 310–318. SPE-65369-PA. http://dx.doi.org/10.2118/65369-PA.

Willis, H. H., Moore, M. 2013. Improving the Value of Analysis for Biosurveillance. *Decision Analysis* **11** (1): 63–81. http://dx.doi.org/10.1287/deca.2013.0283.