Value of information in closed-loop reservoir management

Gonçalves Dias de Barros, Eduardo

DOI
10.4233/uuid:9667dc41-c736-47e6-b818-78c7c50fb08d

Publication date
2018

Document Version
Final published version

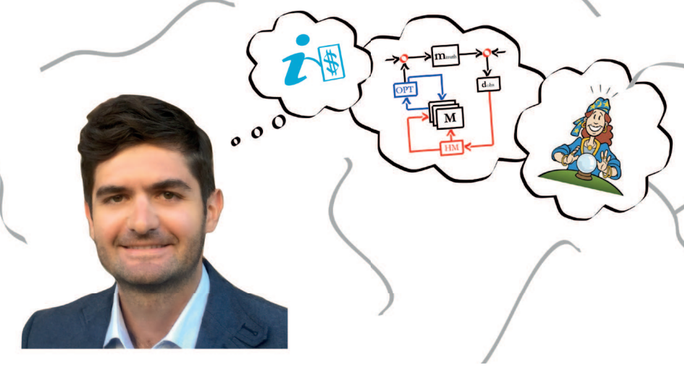
Citation (APA)
Gonçalves Dias de Barros, E. (2018). *Value of information in closed-loop reservoir management*. [Dissertation (TU Delft), Delft University of Technology]. https://doi.org/10.4233/uuid:9667dc41-c736-47e6-b818-78c7c50fb08d

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

Over the past decades, many technological advances have unlocked new opportunities to boost efficiency in the oil and gas industry (e.g., complex well drilling, injection of advanced chemicals, sophisticated instrumentation). The real engineering challenge is to apply these technologies in the best possible way for each particular case. This leads to very difficult decisions to be made, mainly because every oil and gas field is one of its kind and our knowledge of the subsurface is very limited. Many efforts have been made to develop tools to support these decisions by applying a more systematic approach to determine smart exploitation strategies. Yet, very little has been done on the optimization of the reservoir surveillance plans to establish the best observations to monitor the field response to the exploitation strategies, which, in turn, can also contribute to a better exploitation of the reservoir.

In this thesis we propose a methodology to assess the value of future measurements as a first step towards the development of a framework to optimize the design of reservoir surveillance plans. We also investigate alternatives to improve current reservoir management approaches by recommending actions which anticipate the availability of future information and account for the impact of immediate decisions on the decisions to be made in the future.

Throughout the chapters, we discuss how to combine a variety of topics (e.g., model-based optimization, data assimilation, uncertainty quantification) with other unusual ingredients (e.g., plausible truths, clairvoyance, flexible plans) to develop a methodology which can be applied in many problems involving decision making and learning. Despite being motivated by a real application, this research addresses abstract concepts such as value and information, but always from an engineering perspective. This makes us approach the problem in a different way, which, we hope, will inspire innovative solutions in the future.



Value of information in closed-loop reservoir management Eduardo Barros 2018

Value of information
in closed-loop reservoir management

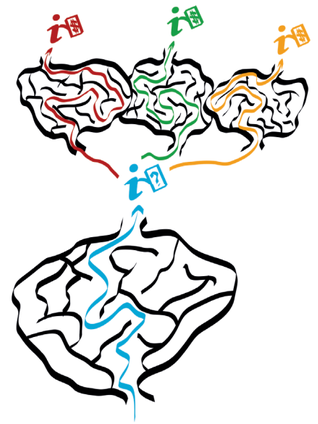
Eduardo Barros

TUDelft

Invitation

You are kindly invited to attend the public defense of my doctoral thesis

Value of information in closed-loop reservoir management



on
Monday, January 22nd 2018 at 10:00 in the Senaatszaal of the Aula of Delft Univerity of Techonology, Mekelweg 5, Delft.

The defense ceremony will be preceded by a short introduction at 9:30.
A reception will be held after the defense.

Eduardo Barros

Propositions

accompanying the dissertation

Value of information in closed-loop reservoir management

by

Eduardo Barros

1. Very little research is done on the optimization of reservoir surveillance strategies. The design of smart surveillance strategies can contribute to the success of reservoir management as much as the optimization of production strategies.
2. A (flexible) strategy only becomes a fixed plan to be put in action. The execution step is the one that allows us to use the closed-loop framework as a decision making environment.
3. Common approaches to assess the usefulness of additional observations are based on the determination of their informativeness defined as either their ability to reduce uncertainty on the model predictions or the extent to which the model predictions are affected by them. However, the most informative observations are not always the most valuable ones.
4. Value of information workflows rely on simulated data to evaluate the contribution of future measurements, i.e. before they are actually gathered. Yet, such a-priori evaluation represents an opportunity to gain insight also into the operations after the new measurements are available.
5. An investment in additional information is one of the alternatives to mitigate the effects of uncertainty. Value of information assessment only has meaning with proper uncertainty quantification.
6. Sometimes the main obstacle to seeing the truth is the choice to ignore the possibility of the wrong. Falsehood as the absence of truth and error as a deviation from the truth are not equal: an error does not necessarily lead to a mistake.
7. A lot of effort is spent on deriving good quality approximations to accelerate computational workflows. However, the usefulness of an approximation depends more on our ability to assess its quality than on its quality itself.
8. Even the most imaginative people are not always able to be creative. Creativity is inversely proportional to the abundance of resources, upper-bounded by hope and propagated through openness.
9. Collaboration goes beyond the execution of tasks in a coordinated effort. It is not possible to have true collaboration without trust or involvement.
10. Not even the most righteous contracts can produce virtue where there is not. The excessive proliferation of professional agreements and codes of conduct in combination with an extreme individualism is a threat to the future of work ethic in our societies.

These propositions are regarded as opposable and defendable, and have been approved by the promotors, Prof. dr. ir. J.D. Jansen and Prof. dr. ir. P.M.J. Van den Hof.

Value of information in closed-loop reservoir management

Value of information in closed-loop reservoir management

Proefschrift

ter verkrijging van de graad van doctor
aan de Technische Universiteit Delft,
op gezag van de Rector Magnificus prof. dr. ir. T.H.J.J. van der Hagen,
voorzitter van het College voor Promoties,
in het openbaar te verdedigen op
maandag 22 januari 2018 om 10:00 uur

door

Eduardo GONÇALVES DIAS DE BARROS

Ingénieur diplômé de l'Ecole Centrale de Lyon, France
geboren te Rio de Janeiro, Brazil.

Dit proefschrift is goedgekeurd door de promotoren:
Prof. dr. ir. J.D. Jansen
Prof. dr. ir. P.M.J. Van den Hof

Samenstelling promotiecommissie:

Rector Magnificus	voorzitter
Prof. dr. ir. J.D. Jansen	Technische Universiteit Delft, promotor
Prof. dr. ir. P.M.J. Van den Hof	Technische Universiteit Eindhoven, promotor

Onafhankelijke leden:

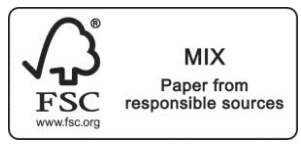
Prof. dr. ir. A.W. Heemink	Technische Universiteit Delft
Prof. dr. W.R. Rossen	Technische Universiteit Delft
Prof. dr. R.B. Bratvold	University of Stavanger, Norway
Dr. A.A. Emerick	Petrobras, Brazil
Dr. ir. O. Leeuwenburgh	TNO
Prof. dr. P.L.J. Zitha	Technische Universiteit Delft, reservelid

This research was carried out within the context of the Integrated Systems Approach to Petroleum Production (ISAPP) knowledge centre. ISAPP is a joint project of TNO, Delft University of Technology (TU Delft), ENI, Statoil and Petrobras.

Cover design by E.G.D. Barros
Printed by Gildeprint

Author's e-mail: egdbarros@hotmail.com

ISBN: 978-94-6366-009-9
An electronic version of this dissertation is available at:
http://repository.tudelft.nl/



"Wherever you go, go with all your heart."
— Confucius

To the loving memory of my father.

"Joy is prayer; joy is strength; joy is love;
joy is a net of love by which you can catch souls."
— Mother Theresa

To the loving memory of my godmother.

Contents

Summary

Efficiency is one of the keys to solve the current and future energy issues in our societies. Improvements in the use of the subsurface will become increasingly necessary to meet the predicted energy demand for the coming decades. According to projections, oil and gas will continue to occupy a large share in our energy mix. In this context, the efficiency with which we exploit our hydrocarbon reserves plays a very important role.

Over the past decades, many technological advances have unlocked new opportunities to boost efficiency in the oil and gas industry (e.g., complex well drilling, injection of advanced chemicals, sophisticated instrumentation). The real engineering challenge is to apply these technologies in the best possible way for each particular case. This leads to very difficult decisions to be made, mainly because every oil and gas field is one of its kind and our knowledge of the subsurface is very limited. Many efforts have been made to develop tools to support these decisions by applying a more systematic approach to determine smart exploitation strategies, like, for example, in reservoir management practices. The focus of these developments has been mostly on production optimization, seeking to determine well settings that result in improved reservoir performance. Yet, very little has been done on the optimization of the reservoir surveillance plans to establish the best observations to monitor the field response to the exploitation strategies, which, in turn, can also contribute to a better exploitation of the reservoir. In this thesis we establish a methodology to assess the value of future measurements as a first step towards the development of a framework to optimize the design of reservoir surveillance plans. We also investigate alternatives to improve current reservoir management approaches by recommending strategies which anticipate the availability of future information and account for the impact of immediate actions on the decisions to be made in the future. This thesis focusses on applications to oil and gas reservoirs, but the topics addressed here are also of relevance to the management of sustainable resources (e.g., geothermal energy) and other uses of the subsurface (e.g., CO_2 and energy storage).

In this thesis, we use state-of-the-art research tools to create an environment for value of information (VOI) assessment for reservoir management applications. The main goal is to

develop a methodology to assess the value of future measurements during the field development planning (FDP) phase of a reservoir, before any actual measurement has been gathered. We propose a workflow to quantify the VOI in closed-loop reservoir management (CLRM), under the assumption that frequent life-cycle optimization will be performed using frequently updated reservoir models (Chapters 2 and 3). The procedure requires extreme amounts of simulations, which makes its application to real-field cases intractable. As a first step to make VOI assessment more practical, we investigate opportunities to apply clustering techniques to select a small subset of representative models and reduce the computational load of the workflow (Chapter 4). The reasoning behind the a-priori VOI analysis unveils an opportunity to improve our approach to reservoir engineering optimization problems by anticipating the fact that additional information will become available in the future. Therefore, we also investigate possible ways to integrate a VOI assessment tool in the optimization framework, with applications to production optimization (Chapter 5) and field development optimization (Chapter 6).

As a result, this thesis covers various aspects to be considered when accounting for the value of future information in reservoir management workflows. Throughout the chapters, we discuss how to combine a variety of topics (e.g., model-based optimization, data assimilation, uncertainty quantification) with more unusual ingredients (e.g., plausible truths, clairvoyance, flexible plans) to develop a methodology which can be applied in many problems involving decision making and learning. Despite being motivated by a real application, this research addresses abstract concepts such as value and information, but always from a practical engineering perspective. This combination contributes to a new way of reasoning that can be useful to support decisions in reservoir management, which, we hope, will inspire innovative solutions in the future.

1

Introduction

Energy plays a fundamental role in our modern societies. Our capacity of changing the world around us to make it a better place depends on its availability. In the pursue of a sustainable development, many efforts have been made to reduce our dependency on traditional energy resources and come up with technologies that allow us to do more with less energy. However, these advances can still not entirely solve all our energy issues: it is estimated that the global energy demand will continue to increase in the near future and the supply of renewable energy will not be able to keep up with it, at least not in the next few decades. In this context, the most sustainable approach is to introduce these emerging technologies while relying on a necessary transition period in which fossil fuels such as hydrocarbons will remain an important part of the energy mix (IEA, 2016).

New technologies can also make the exploitation of hydrocarbon reserves more efficient and increase the recovery of oil and gas. For example, improvements in drilling create opportunities to design more complex wells with better contact to the reservoir formations and the injection of advanced chemicals reduces the residual volumes of hydrocarbons that would otherwise stay trapped in the reservoir. Other efforts are related to the so-called smart fields technologies, with the installation of control devices to regulate the flow of fluids in the wells, and instruments to obtain more information from the reservoir. Next to these and many other technical advances, numerical techniques for reservoir simulation and model-based optimization have developed rapidly over the past decades, providing support to design and operational decisions in order to benefit the most from all the technologies and maximize the recovery of hydrocarbons.

Reservoir management is the set of practices adopted by oil and gas companies to optimize the performance of their reservoir assets. As reservoir engineers and geoscientists, we recognize our inability to fully characterize the reservoir due to our

limited knowledge of the subsurface. Despite the presence of uncertainties, important decisions on how to exploit the reservoir have to be made, such as the elaboration of a field development plan (i.e., where, when, which type and how many wells to drill) and a production strategy (i.e., how to operate the wells). The success of the exploitation of the reservoir is directly related to the quality of such decisions, which involve significant investments. This high-risk decision making process is supported by performance predictions (e.g., simulation forecasts) based on all the knowledge available. Other important decisions concern the design of a reservoir surveillance plan (i.e., where, when, what, with which frequency and precision to measure) to guide the deployment of sensors, which is also costly. As the development of the field starts, the deployed sensors gather measurements that are used to monitor the response of the reservoir and to determine whether the implemented actions (i.e., field development plan and production strategy) have the expected effect in the reservoir performance. When there is a discrepancy between the actual response of the reservoir and the predictions, there is an opportunity for learning and improving the knowledge of the reservoir to make better decisions in the future.

Many efforts have been made to automate the reservoir management process as much as possible so that a more systematic approach can be used to continuously optimize all the decisions throughout the reservoir life-cycle (i.e., closed-loop reservoir management, real-time reservoir management, integrated operations, etc.); see references in section 1.1. However, the focus of these developments has been mostly on production optimization, seeking to determine production strategies that result in improved performance. Very little has been done on the optimization of reservoir surveillance to establish the best observations to monitor the performance of the production strategies, which in turn also contribute to a better reservoir management. Such an optimization requires the ability to quantify the contribution of surveillance plans to the success of the reservoir management before any measurements are gathered. The challenge is on how to estimate the incremental performance that the future measurements (i.e., yet to be gathered) will enable, or, in other words, to assess their value. This thesis addresses this challenge.

1.1. Closed-loop reservoir management

Closed-loop reservoir management (CLRM) is a combination of frequent life-cycle production optimization and data assimilation (also known as computer-assisted history matching). Life-cycle optimization aims at maximizing a financial measure, typically net present value (NPV), over the producing life of the reservoir by optimizing the production strategy. This may involve well location optimization, or, in a more restricted setting, optimization of well rates and pressures for a given configuration of wells, on the basis of one or more numerical reservoir models. Data assimilation involves modifying the parameters of one or more reservoir models, or the underlying geological models, with the aim to improve their predictive capacity, using measured data from a potentially wide variety of sources such as production data or time-lapse seismic. For further information on CLRM see, e.g., Jansen et al. (2005, 2008, 2009), Naevdal et al. (2006), Sarma et al. (2008); Chen et al. (2009); Wang et al. (2009), Foss and Jensen (2011) and Hou et al. (2015).

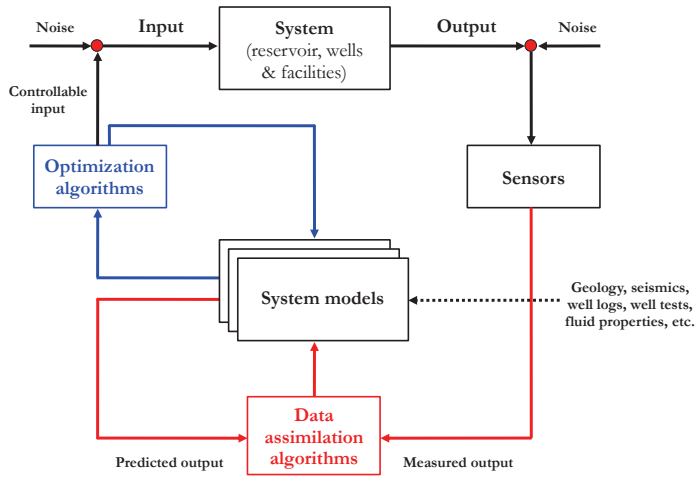


Figure 1.1: Closed-loop reservoir management as a combination of life-cycle optimization and data assimilation.

Figure 1.1 depicts the CLRM framework in a block-diagram representation often used in systems and control theory. We recognize two distinct loops connecting the system predictive models to the real system: the optimization loop shown in blue and the data assimilation loop in red. The idea behind the CLRM framework is to use computer-

assisted workflows to increase the frequency of reservoir management cycles and change it from a batch-type to a near-continuous process. Thus, in other words, the goal of CLRM is to maximize the performance of the "ground truth" reservoir by frequently adjusting the production strategy to be implemented in reality based on predictions from continuoustly updated reservoir models that incorporate all the knowledge and information available. Note that, besides the two loops, we have also the execution step in which the optimized strategy is applied to the real system. This is the step through which the CLRM framework can increase the performance of the real asset and therefore create value. This point will be important for the understanding of the methodology developed in this thesis.

1.2. Optimization

1.2.1. Robust optimization

An efficient model-based optimization algorithm is one of the required elements for CLRM. Because of the inherent uncertainty in the geological characterization of the subsurface and the deterministic character of conventional reservoir simulation, a multi-scenario approach is necessary. Robust life-cycle optimization uses one or more ensembles of geological realizations (reservoir models) to account for uncertainties and to determine the production strategy that maximizes a given objective function over the ensemble; see, e.g., Yeten et al. (2003) or Van Essen et al. (2009). Figure 1.2 schematically represents robust optimization over an ensemble of N realizations $\mathbf{M} = \{\mathbf{m}_1, \mathbf{m}_2, \ldots, \mathbf{m}_N\}$, where $\mathbf{m}$ is a vector of uncertain model parameters (e.g., grid block permeabilities or fault multipliers). Typically, the objective function to be optimized is the average net present value (NPV):

$$J_{NPV} = \mu_{NPV} = \frac{1}{N}\sum_{i=1}^{N} J_i \,, \tag{1.1}$$

where μ_{NPV} is the ensemble mean of the objective function values J_i of the individual realizations. The objective function J_i for a single realization i is defined as

$$J_i = \int_{t=0}^{T} \frac{q_o(t,\mathbf{m}_i)\, r_o - q_{wp}(t,\mathbf{m}_i)\, r_{wp} - q_{wi}(t,\mathbf{m}_i)\, r_{wi}}{(1+b)^{t/\tau}}\, dt \,, \tag{1.2}$$

where t is time, T is the producing life of the reservoir, q_o is the oil production rate, q_{wp} is the water production rate, q_{wi} is the water injection rate, r_o is the price of oil produced, r_{wp} is the cost of water produced, r_{wi} is the cost of water injected, b is the discount factor expressed as a fraction per year, and τ is the reference time for discounting (typically one year). The outcome of the optimization procedure is a vector **u** containing the settings of the control variables over the producing life of the reservoir. Typical elements of **u** are monthly or quarterly settings of well head pressures, water injection rates, valve openings etc. Note that, although the optimization is based on N models, only a single strategy **u** is obtained, under the rationale that only one strategy can be implemented in reality. Note also that, despite being very disseminated among CLRM practitioners, the robust optimization approach presented by Van Essen et al. (2009) is only one way of dealing with uncertainty in production optimization. An alternative approach is to balance risk and return within the optimization by including well-defined risk measures or other utility functions in the objective function; see, e.g., Capolei et al. (2015) and Siraj et al. (2016).

Figure 1.2: Robust optimization: optimizing the objective function of an ensemble of N realizations resulting in a single control vector **u**.

1.2.2. Data assimilation

Efficient data assimilation algorithms are also an essential element of CLRM. Many methods for reservoir-focused data assimilation have been developed over the past years, and we refer to Oliver et al. (2008), Evensen (2009), Aanonsen et al. (2009) and Oliver and Chen (2011) for overviews. An essential component of data assimilation is accounting for uncertainties, and it is generally accepted that this is best done in a Bayesian framework, which updates the probability for a hypothesis on the basis of the availability of additional evidence. This can be formulated in mathematical terms through the Bayes' rule:

$$p(\mathbf{m}\,|\,\mathbf{d}) = \frac{p(\mathbf{d}\,|\,\mathbf{m})\,p(\mathbf{m})}{p(\mathbf{d})}, \tag{1.3}$$

where p indicates the probability density, and $\mathbf{d}$ is a vector of measured data (e.g., oil and water flow rates or saturation estimates from time-lapse seismic). In equation (1.3) the terms $p(\mathbf{m})$ and $p(\mathbf{m}|\mathbf{d})$ represent the prior and posterior probability densities of the model parameters $\mathbf{m}$, which are, in our setting, represented by the prior and posterior ensembles respectively. The underlying assumption in data assimilation is that the assimilation of measured (historical) data leads to an improved (future) predictive capacity of the models, which, in turn, leads to improved decisions. In our CLRM setting, decisions take the form of control vectors $\mathbf{u}$, aimed at maximizing the objective function J.

1.3. Previous work

In order to situate the objectives of this research, we review very briefly in this section some of the previous work related to the topic addressed in this thesis. A more extensive literature review will be presented in Chapter 2.

Previous work on information valuation in reservoir engineering focused on analyzing how additional information impacts the model predictions. Krymskaya et al. (2010) use the concept of observation impact, which provides a measure of the information content in the observations. Le and Reynolds (2014a, 2014b) quantify the usefulness of information as how much the assimilation of an observation contributes to reducing the uncertainty of a variable of interest (e.g., NPV). Both approaches are based on data assimilation, and Figure 1.3 schematically represents how measured data are used to update a prior ensemble of reservoir models, resulting in a posterior ensemble which forms the basis to compute various measures of information valuation. Note that these two studies only measure the effect of additional information on model predictions and do not explicitly take into account how the improved model predictions are used to make better decisions.

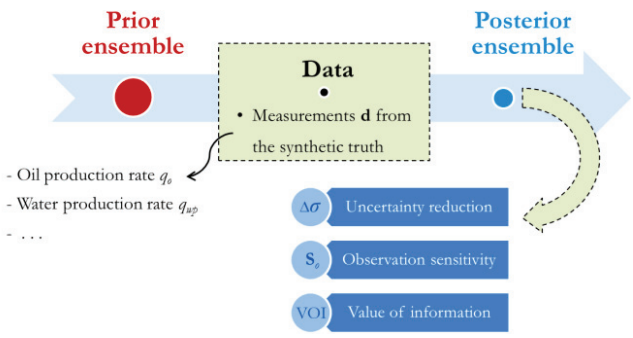


Figure 1.3: Data assimilation and information valuation.

Another way to quantify the value of future information originates from the field of decision theory. Howard (1966) was one of the first to formalize the idea that information could be economically valued within a context of decision making under uncertainties. Bratvold et al. (2009) produce an extensive literature review on VOI analysis in the oil industry, but none of the applications reported by them seems to address reservoir management problems. Their main point is that "one cannot value information outside of a particular decision context". In this context, VOI is defined as the difference between the value achieved by the decision made with the additional information and the value achieved by the decision made without it.

1.4. Research objectives

The main goal of this PhD work is to answer two key questions:

Q1. How to quantify the contribution of future measurements to the success of CLRM?

Finding an answer to this first question has been the object of recent research studies in the reservoir engineering community; see section 1.3. However, these studies were restricted to the use of the data assimilation scheme to quantify the effect of measurements on the model predictions, which was then assumed to be a measure of their usefulness in the view of the inherent geological uncertainties. On the other hand, previous work in decision theory (section 1.3) showed that the value (or usefulness) of additional information depends on the decision context. Inspired by the VOI concepts of decision theory, in this thesis, we investigate opportunities to use the CLRM environment

as a tool to quantify how much future measurements will contribute to better decision making. The improvement in decision making results in incremental performance (or additional value, expressed in terms of the selected performance metric), allowing us to assess the VOI in CLRM. Thus, the VOI should be ultimately understood as a function of the assessed future measurements, the decisions of interest (and the corresponding performance metric) and the (initial) state of uncertainty.

VOI assessment is the first step towards the development of a framework to optimize the design of reservoir surveillance plans. The ability of assessing the value of future measurements is relevant to determine the economic feasibility and support investment decisions on reservoir surveillance. It can be used to establish whether the expected additional value of specific observations is worth the cost to obtain them or to determine how much one should be willing to pay for them. Besides that, improved designs of monitoring strategies represent an opportunity to reduce project expenses on costly observations by allowing us to invest only on the deployment of the measurements that are expected to add the most value. Some of these measurements may be gathered only once or a couple of times (e.g., a repeat seismic survey or a production test) whereas others may be gathered multiple times or even continuously once the sensors are installed (e.g., production data from permanent downhole gauges). In this thesis, we seek a methodology that can be applied in both situations, for single and multiple observation times.

An additional complexity arises when it is attempted to quantify the VOI for CLRM, i.e., under the assumption that frequent life-cycle optimization will be performed using frequently updated reservoir models. Therefore, the objective of the first part of this research is to:

Obj.1: Develop a methodology to assess the VOI in such a CLRM context.

The optimal reservoir surveillance plan is the one that delivers the most valuable measurements throughout the reservoir life-cycle, providing the most useful information for reservoir management purposes. However, it is important to realize that these measurements are observations of the response of the reservoir to the implemented production and development strategies, which therefore have also an impact on the

outcome and value of the measurements. This brings us to the second question that motivates this PhD research:

Q2. How to determine production strategies that, besides optimizing production, can also deliver the most useful information?

The ability of assessing the value of future measurements constitutes an opportunity to reformulate the production optimization problem, allowing us to account for the contribution of future information to optimize production. The current approach for life-cycle optimization under uncertainty (section 1.2.1) assumes that the (geological) uncertainties are static. However, we know that additional measurements will become available throughout the reservoir life-cycle. When we consider only the initial state of uncertainty to optimize the production strategy for the entire reservoir life-cycle, we do not take advantage of the fact that we control when to gather and assimilate the future information. To circumvent this limitation, in the second part of this work we investigate possible approaches to:

Obj.2: Integrate a VOI assessment tool in the CLRM optimization framework.

1.5. Thesis outline

This thesis has two parts corresponding to the two main research objectives described in section 1.4. In the first part (Chapters 2, 3 and 4) we focus on the development of the methodology for VOI assessment in CLRM. In the second part (Chapters 5 and 6), we discuss the use of VOI considerations in the optimization of exploitation strategies.

Chapter 2 presents the base of the methodology by introducing our workflow to determine the VOI given a single observation time in the future. In this chapter, we start by reviewing more extensively the previous work on information valuation with applications in reservoir engineering. We then describe our methodology that uses the entire CLRM framework to include the decision making in the VOI assessment instead of relying only on the data assimilation step to quantify the value of future measurements.

Chapter 3 discusses the extension of the methodology to cases with multiple observation times. In this chapter, we describe how a slight modification can enable the original workflow from Chapter 2 to assess the value of a series of measurements without a

prohibitive increase in computational costs. Next, with the help of a simple illustrative example we show that the results obtained are consistent with our previous findings.

Chapter 4 is dedicated to making our proposed methodology more practical. For that, we use clustering techniques to select subsets of representative models and significantly reduce the computational costs of the original workflow. We repeat the numerical experiments from the previous chapters and we obtain similar results with a reduction of the number of reservoir simulations by approximately two orders of magnitude. After that, we apply the same measures to make the VOI assessment possible in a larger example as a first step towards large-scale applications.

Chapter 5 presents a new approach for production optimization in the context of CLRM by considering the impact of future measurements within the optimization framework. We integrate the reasoning behind the a priori VOI analysis to modify the optimization problem so that it anticipates the fact that additional information (e.g., production measurements) will become available in the future. We illustrate the concept with the simple example from the previous chapters and the results obtained confirm that this new approach can lead to better decisions and increased VOI.

Chapter 6 expands the ideas introduced in Chapter 5 to the field development optimization problem. We combine VOI assessment and well-location optimization in a nested approach which delivers flexible development plans that consider the effect of time-dependent uncertainties. This allows the optimization to benefit of the sequential nature of the drilling activities and to be informed of the impact of current decisions (i.e., the drilling of the first next wells) and future information on subsequent decisions (i.e., the drilling locations of future wells), resulting in better development strategies.

Finally, this thesis is concluded in Chapter 7. This chapter provides an overview of the conclusions of this research, highlighting the main findings of each chapter, followed by a list of recommendations for future research on topics related to VOI assessment in CLRM.

2

Value of information for a single observation time

This chapter [1] *presents our methodology to perform value of information (VOI) analysis within a closed-loop reservoir management (CLRM) framework. The workflow combines tools such as robust optimization and history matching in an environment of uncertainty characterization. The approach is illustrated with two simple examples: an analytical reservoir toy model based on decline curves and a water flooding problem in a two-dimensional five-spot reservoir. The results are compared with previous work on other measures of information valuation, and we show that our method is a more complete, although also more computationally intensive, approach to VOI analysis in a CLRM framework.*

2.1. Introduction

Over the past decades, numerical techniques for reservoir model-based optimization and history matching have developed rapidly, while it also has become possible to obtain increasingly detailed reservoir information by deploying different types of well-based sensors and field-wide sensing methods. Many of these technologies come at significant costs, and an assessment of the associated value of information (VOI) becomes therefore increasingly important (Kikani, 2013). In particular assessing the value of future measurements during the field development planning (FDP) phase of an oil field requires techniques to quantify the VOI under geological uncertainty. An additional complexity arises when it is attempted to quantify the VOI for closed-loop reservoir management (CLRM), i.e., under the assumption that frequent life-cycle optimization will be performed

[1] This chapter is based on 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.

using frequently updated reservoir models. This chapter describes a methodology to assess the VOI in such a CLRM context.

In the Background section (2.2) we review some previous work on information measures. Next, in the Methodology section (2.3), we present the proposed workflow for VOI analysis in CLRM and thereafter, in the Examples section (2.4), we illustrate it with some case studies in which the results of the VOI calculations are analyzed and compared with other information measures. In the Discussion section (2.5), we address some other topics and issues related to the methodology and its use in practice. Finally, in the Conclusions section (2.6), we comment on the computational aspects of applying this workflow to real field cases, and we suggest a direction for further research.

2.2. Background

2.2.1. Information valuation

Previous work on information valuation in reservoir engineering focused on analyzing how additional information impacts the model predictions. One way of valuing information is proposed by Krymskaya et al. (2010). They use the concept of observation impact, which was first introduced in atmospheric modeling. Starting from a EnKF scheme, they derive an observation sensitivity matrix

$$\mathbf{S} = \frac{\partial \hat{\mathbf{y}}}{\partial \mathbf{y}} = \mathbf{HK} = \left(\mathbf{C_d}^{-1} + \left(\mathbf{HC_m H}^T \right)^{-1} \right)^{-1} \mathbf{C_d}^{-1}, \tag{2.1}$$

where $\mathbf{y}$ is the vector of observations, $\hat{\mathbf{y}}$ is the vector of predicted observations of the updated model, $\mathbf{H}$ is the observation operator, $\mathbf{K}$ is the Kalman gain matrix, $\mathbf{C_m}$ is the covariance matrix of the prior model errors and $\mathbf{C_d}$ the covariance matrix of the measurement errors. The observation sensitivity matrix $\mathbf{S}$ is a symmetric matrix that contains self and cross-sensitivities (diagonal and off-diagonal elements of the matrix, respectively). The self-sensitivities, which quantify how much the observation of measured data impacts the prediction of these same data by a history-matched model, provide a measure of the information content in the data. Their joint influence can be expressed with a global average influence index defined as

$$I_{GAI} = \frac{tr(\mathbf{S})}{N_{obs}}, \quad 0 \le I_{GAI} \le 1 \ , \tag{2.2}$$

where N_{obs} is the number of observations (i.e., the number of diagonal elements in **S**).

Another approach is taken by Le and Reynolds (2014a, 2014b) who address the usefulness of information in terms of the reduction in uncertainty of a variable of interest (e.g., NPV). They introduce a method to estimate, in a computationally feasible way, how much the assimilation of an observation contributes to reducing the spread in the predictions of the variable of interest, expressed as the difference between P_{10} and P_{90} percentiles, i.e. between the 10 % and 90 % cumulative probability density levels.

Both approaches are based on data assimilation, and Figure 1.3 schematically represents how measured data are used to update a prior ensemble of reservoir models, resulting in a posterior ensemble which forms the basis to compute various measures of information valuation. In Figure 1.3 the measurements are obtained in the form of synthetic data generated by a synthetic truth. This preempts our proposed method of information valuation in which we will use an ensemble of models in the FDP stage, of which each realization will be selected as a synthetic truth in a consecutive set of twin experiments.

2.2.2. VOI and decision making

The two studies that we referred to above (Krymskaya et al., 2010 and Le and Reynolds, 2014a and 2014b) only measure the effect of additional information on model predictions and do not explicitly take into account how the additional information is used to make better decisions. In these studies it has simply been assumed that history-matched models automatically lead to better decisions. However, there seems to be a need for a more complete framework to assess the VOI, including decision making, in the context of reservoir management. VOI analysis originates from the field of decision theory. It is an abstract concept, which makes it a powerful tool with many potential applications, although implementation can be complicated.

An early reference to VOI originates from Howard (1966) who considered a bidding problem and was one of the first to formalize the idea that information could be economically valued within a context of decision making under uncertainties. Since then, several applications have appeared in many different fields, including the petroleum

industry. Bratvold et al. (2009) produce an extensive literature review on VOI in the oil industry. Their main message is that "one cannot value information outside of a particular decision context". Thus, reducing uncertainty in a model prediction has no value by itself, and VOI is decision-dependent.

Over the last years, the number of publications on VOI in reservoir engineering applications has been growing, along with new approaches which tend to include decision making in the analysis. Bhattacharjya et al. (2010), Trainor-Guitton et al. (2013) and Nakayasu et al. (2016) have proposed methodologies to quantify the value of spatial information to assist in the exploration and development of reservoirs. Sato (2011) has discussed the use of VOI analysis for the design of monitoring strategies in geological CO_2 storage. Bailey et al. (2011) have addressed the problem of valuing future measurements in the context of the optimization of well completions to maximize production. More recently, He et al. (2016) and Chen et al. (2016) have studied the a-priori evaluation of pilot and surveillance programs. For a more complete review of the recent developments on VOI in Earth science related topics, we refer to Eidsvik et al. (2015).

2.3. Methodology

In our setting, the decision is the use of an optimized production strategy as obtained in the CLRM framework (section 1.1). We intend to not only quantify how information changes knowledge (through data assimilation), but also how it influences the results of decision making (through optimization). We express the optimized production strategy in the form of a control vector $\mathbf{u}$ which typically has tens to hundreds of elements (e.g. bottom-hole pressures, injection rates or valve settings at different moments in time) and which needs to be updated when new information becomes available. The proposed workflow is depicted in Figure 2.1. The procedure consists of a sort of twin experiment on a large scale, because the analysis is performed in the design phase – when no real data are yet available. Note that classical CLRM is performed during the operation of the field whereas we are considering here an a-priori evaluation of the value of CLRM (i.e. in the design phase). The workflow starts with an ensemble $\mathbf{M}_{truth}$ of N_{truth} realizations which characterizes the initial uncertainty associated with the model parameters. From this ensemble, one realization is selected to be the synthetic truth $\mathbf{m}_{truth}$. Thereafter, a new

ensemble of N equiprobable members is generated, by sampling from the same distribution as used to create the initial ensemble $\mathbf{M}_{truth}$, to form the prior ensemble $\mathbf{M}_{prior}$ for the robust optimization procedure:

$$\mathbf{u}_{prior}(0:T) = \arg\max_{\mathbf{u}} \mu_{NPV}(\mathbf{u}, \mathbf{M}_{prior}) = \arg\max_{\mathbf{u}} \frac{1}{N}\sum_{i=1}^{N} J^{i}_{NPV}(\mathbf{u}, \mathbf{m}_i). \tag{2.3}$$

Next, synthetic data $\mathbf{d}_{obs}(t)$ are generated by running a reservoir simulation for the synthetic truth $\mathbf{m}_{truth}$ while applying the robust strategy $\mathbf{u}_{prior}(0:T)$. The synthetic data $\mathbf{d}_{obs}(t)$ are perturbed by adding zero-mean Gaussian noise with a predefined standard deviation ε_{obs}. With these, data assimilation is performed, the model realizations of $\mathbf{M}_{prior}$ are updated and a posterior ensemble $\mathbf{M}_{post}$ obtained. As a next step robust optimization is carried out on this posterior to find a new optimal production strategy $\mathbf{u}_{post}(t:T)$ (from the time the data became available to the end of the reservoir life-cycle):

$$\begin{aligned}\mathbf{u}_{post}(t:T) &= \arg\max_{\mathbf{u}} \mu_{NPV}\left(\mathbf{u}_{prior}(0:t), \mathbf{u}(t:T), \mathbf{M}_{post}\right) \\ &= \arg\max_{\mathbf{u}} \frac{1}{N}\sum_{i=1}^{N} J^{i}_{NPV}\left(\mathbf{u}_{prior}(0:t), \mathbf{u}(t:T), \mathbf{m}_i\right).\end{aligned} \tag{2.4}$$

The concept of a twin experiment in data assimilation is in this way extended to include the effects of the model updates on the reservoir management actions.

The strategies obtained for the prior and the posterior ensembles are then tested on the synthetic truth, and their economic outcomes (NPV values $J_{NPV,prior}$ and $J_{NPV,post}$) are evaluated. The difference between these outcomes is a measure of the VOI incorporated through the CLRM procedure for this particular choice of the synthetic truth:

$$J_{NPV,prior} = J_{NPV}(\mathbf{u}_{prior}, \mathbf{m}_{truth}). \tag{2.5}$$

$$J_{NPV,post} = J_{NPV}\left(\mathbf{u}_{prior}(0:t), \mathbf{u}_{post}(t:T), \mathbf{m}_{truth}\right). \tag{2.6}$$

The choice of one of the realizations to be the synthetic truth in the procedure is completely random. In fact, because the analysis is conducted during the FDP phase, any of the N_{truth} models in the initial ensemble $\mathbf{M}_{truth}$ could be selected to be the 'truth'. Note that this also implies that the VOI is a random variable. One of the underlying assumptions of our proposed workflow is that the truth is a realization from the same probability distribution function as used to create the realizations of the ensemble. Hence,

the methodology only allows to quantify the VOI under uncertainty in the form of known unknowns. Obviously, specifying uncertainty in the form of unknown unknowns is impossible, which therefore is a fundamental shortcoming in any VOI analysis. (I.e., we may think that we know the complete reservoir description (as captured in the prior ensemble), but we may have missed "unmodeled" features such as an unexpected aquifer or a sub-seismic fault.)

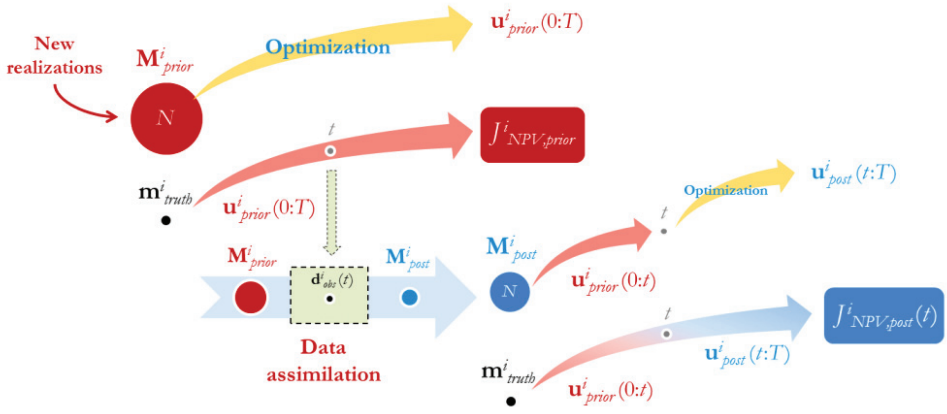


Figure 2.1: Proposed workflow to compute the value of information. (t indicates the observation time and T indicates the end time).

Because any of the N_{truth} models in $\mathbf{M}_{truth}$ could be the truth, the procedure has to be repeated N_{truth} times, consecutively letting each one of those initial models act as the synthetic truth. This allows us to quantify the expected VOI over the entire ensemble:

$$\text{VOI} = \frac{1}{N_{truth}} \sum_{i=1}^{N_{truth}} \left(J^i_{NPV,post} - J^i_{NPV,prior} \right). \tag{2.7}$$

We note that this repetition is similar to the use of multiple plausible truth cases in Le and Reynolds (2014a, 2014b). We also note that in the literature on VOI, most of the times the term VOI is used to refer to the expected VOI. The flowchart in Figure 2.2 shows the complete procedure. Finally we note that, to be absolutely rigorous, we would have to repeat the whole workflow several times with different realizations of the noise in the observation vectors. However, we argue that by far the largest contribution to uncertainty originates from the geology, as captured in the various ensembles of geological realizations. In comparison, the effect of measurement noise is small and sufficiently captured by using a new noise realization for each synthetic measurement.

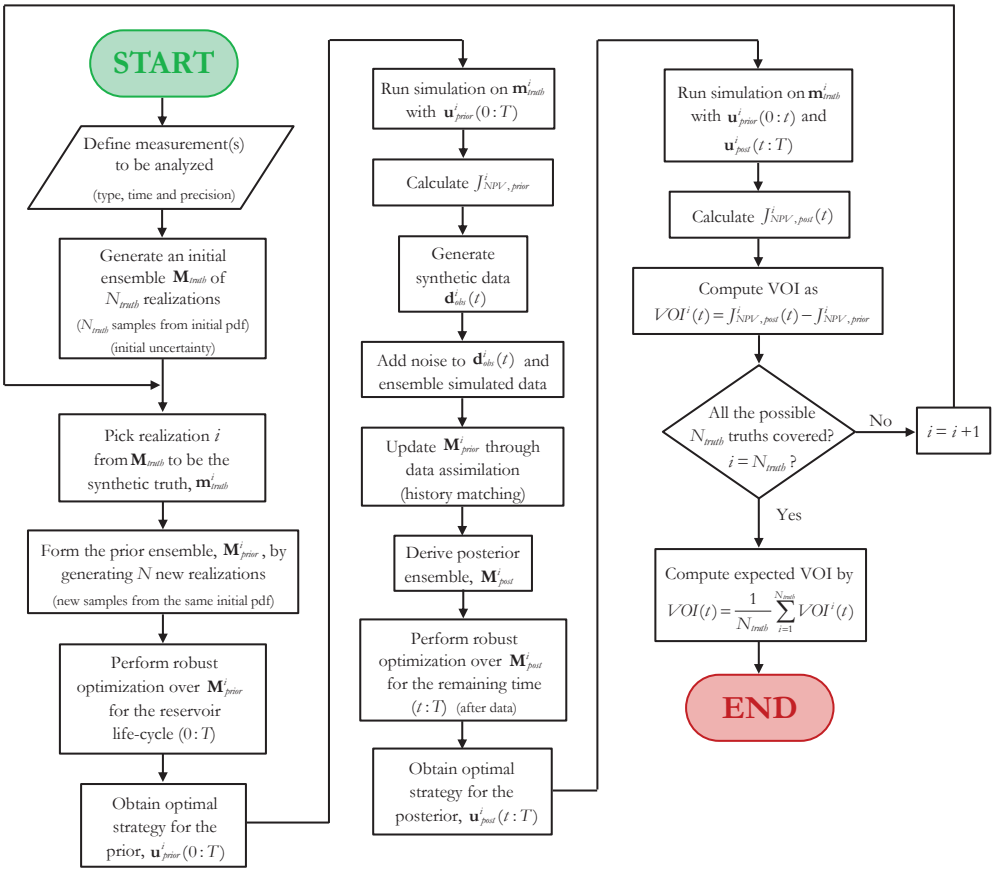


Figure 2.2: Complete workflow to compute the expected VOI.

The workflow can be adapted to compute the expected value of clairvoyance (VOC), which simply means that at some time in the reservoir life we suddenly know the truth so we can perform life-cycle production optimization on the true reservoir model instead of considering the robust optimization over the posterior ensemble (as in equation (2.4)):

$$\mathbf{u}_{post}(t:T) = \arg\max_{\mathbf{u}} J_{NPV}\left(\mathbf{u}_{prior}(0:t), \mathbf{u}(t:T), \mathbf{m}_{truth}\right) \tag{2.8}$$

The estimated expectation of VOC is then computed from equation (2.7) where each posterior NPV is obtained while applying the optimal controls determined for the associated synthetic true model. Such a clairvoyance implies the availability of completely informative data without observation errors, and the expected VOC therefore forms a theoretical upper bound (i.e., a "technical limit") to the expected VOI. Moreover, because

this modified workflow does not require data assimilation, and, after the truth has been revealed, only requires optimization of a single (true) model, it is computationally significantly less demanding.

2.4. Examples

2.4.1. Toy model

As a first step to test the proposed methodology we used a very simple model with only a few parameters, based on reservoir decline curves. It describes oil and water flow rates q_o and q_w as a function of time t and a scalar control variable u according to the following expressions:

$$q_o(u,t) = \left(q_{o,ini} + c_1 u\right) \exp\left(-\frac{t}{a + \frac{1}{c_2}u}\right), \tag{2.9}$$

$$q_w(u,t) = H\left[t_{bt}\left(1 - \frac{1}{c_3}u\right)\right]\left(q_{w,\infty} + u\right)\left[1 - \exp\left(-\frac{t - t_{bt}\left(1 - \frac{1}{c_3}u\right)}{c_4 a - \frac{1}{c_5}u}\right)\right], \tag{2.10}$$

where $q_{o,ini}$ is the initial production rate, t_{bt} is the water breakthrough time, and $q_{w,\infty}$ is the asymptotic water production rate, all for a situation without control, i.e., for $u = 0$. The oil production follows an exponential decline and the water production builds up exponentially from a breakthrough time modelled by a Heaviside step function H. The variables have dimensions as listed in Table 2.1, where L, M and t indicate length, monetary value and time respectively. Some of the parameters are constants, while four uncertain parameters are normally distributed with values indicated in Table 2.1. The scalar control variable u somehow mimics a water injection rate to the reservoir; higher values of u slow down the decline of oil production but accelerate water breakthrough and increase water production, as shown in Figure 2.3. Given the prices and costs associated with oil and water production, there is room for optimization to determine the value of u that maximizes the economics of the reservoir over a fixed producing life-time. To allow for regular updating of the control strategy over the producing life of the reservoir, the

scalar u can be replaced by a vector $\mathbf{u} = [u_1 \quad u_2 \quad \cdots \quad u_M]^T$, where M is the number of control intervals.

Table 2.1: Parameter values for the toy model.

Variables		Constant parameters		Uncertain parameters	
q_o	[L^3 t^{-1}]	$c_1 = 0.1$	[–]	$q_{o,ini} \sim N(100, 8)$	[L^3 t^{-1}]
q_w	[L^3 t^{-1}]	$c_2 = 4$	[L^3 t^{-2}]	$a \sim N(30.5, 3.67)$	[t]
$t \in [0, 80]$	[t]	$c_3 = 150$	[L^3 t^{-1}]	$q_{w,\infty} \sim N(132, 6)$	[L^3 t^{-1}]
$u \in [0, 50]$	[L^3 t^{-1}]	$c_4 = 2$	[–]	$t_{bt} \sim N(32, 6)$	[t]
		$c_5 = 1.33$	[L^3 t^{-2}]		
		$r_o = 70$	M L^{-3}		
		$r_w = 10$	M L^{-3}		
		$b = 0.10$	[–]		

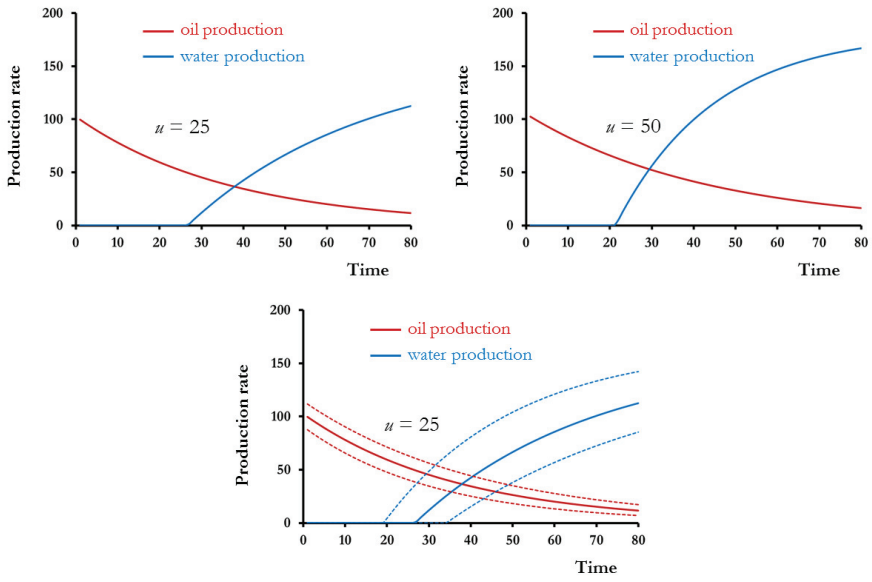


Figure 2.3: Toy model behavior: oil and water production for two fixed values of the control variable u (top); representation of uncertainty in the form of P10 and P90 percentiles (bottom).

The question to be answered here was: given an initial ensemble of models describing the geological uncertainties and an initial optimized control vector **u**, what is the value of a production test in the form of a measurements $\mathbf{d} = [q_o(t_{data}) \quad q_w(t_{data})]^T$ of oil and water

production rates at a given time t_{data}, for different measurement errors and observation times? The VOI assessment procedure described in the previous section was applied, and repeated for different observation times, $t_{data} = \{1, 2, \ldots, 80\}$. We used a random measurement error with a standard deviation $\varepsilon_{data} = 5\,\%$ of the measured value, $N_{truth} = 100$ plausible truths, prior ensembles of $N = 99$ model realizations and $M = 8$ control time-steps. Ensemble optimization (EnOpt) and ensemble Kalman filtering (EnKF) were used to perform the robust optimization and the data assimilation respectively. (We used the robust StoSAG implementation of Fonseca et al. (2016) which is a modified form of the original EnOpt formulation proposed by Chen et al. (2009). For general information on EnKF, see, e.g., Evensen (2009) or Aanonsen et al. (2009); we used a straightforward implementation without localization or inflation.) The VOI, the VOC, the observation impact I_{GAI} and the uncertainty reduction $\Delta\sigma_{NPV} = \sigma_{NPV,prior} - \sigma_{NPV,post}$ (i.e., the difference of standard deviations of the prior and posterior NPV distributions) were computed for each of the 80 observation times. The average NPV for the initial ensemble is \$ 108,900 when using base line control (i.e. the average of the upper (50) and lower bounds (10), $\mathbf{u}_{ini} = \{30, 30, \ldots, 30\}$) and \$ 114,300 when using robust optimization over the prior (i.e. without additional information). The initial uncertainty is $\sigma_{NPV,ini}$ = \$ 11,960, computed as the average of the standard deviations in the NPV of the different prior ensembles. We repeated the optimization by starting from a more aggressive initial strategy where the values of $\mathbf{u}_{ini}$ were at their bounds, which gave near-identical results.

The expected VOC as a function of the time of clairvoyance is depicted in Figure 2.4 (top left), where we expressed the monetary value, arbitrarily, in \$. The dashed line represents the expected VOC, i.e. the mean of the ensemble of $N_{truth} = 100$ plausible truths. The dark solid line and the two lighter solid lines represent the P_{50} and P_{10}/P_{90} percentiles respectively. Here, P_x is defined as the probability that x % of the outcomes exceeds this value. The expected VOC is the value one could obtain if the truth could be revealed and all the uncertainty could be eliminated at no costs at time t_{data}. Of course, these results depend on the operation schedule (i.e., the number of control time-steps) and on the initial ensemble of realizations that characterize the uncertainty. As can be seen, the VOC exhibits a stepwise decrease over time, with the steps coinciding with the eight control

time steps. This stepwise behavior occurs because knowing the truth only affects the way one operates the reservoir from the moment of clairvoyance and because the production strategy can only be updated at the defined control time steps. The sooner clairvoyance is available, the more control time steps can be tuned to re-optimize the production strategy based on the truth, and, therefore, the more value is obtained. Thus, this plot demonstrates the importance of timing when collecting additional information to make decisions. Even clairvoyance can be completely useless (VOC = 0) when it is obtained too late (in this case after t_{data} = 40).

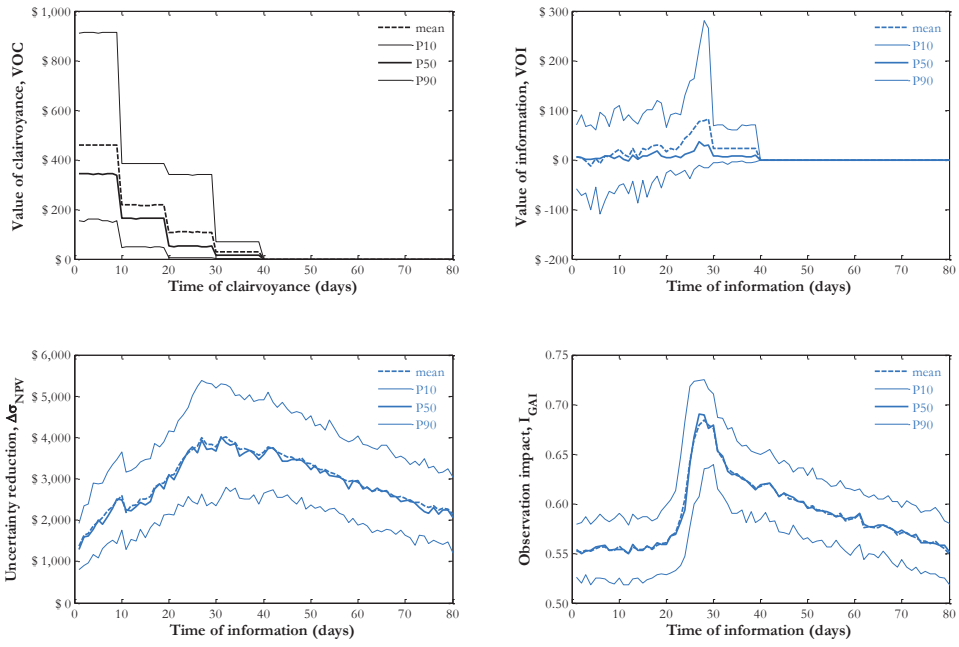


Figure 2.4: Results for the VOI analysis in the toy model: VOC (top left); VOI (top right); uncertainty reduction (bottom left); observation impact (bottom right).

The percentiles of the VOC distribution in Figure 2.4 (top left) illustrate that the VOC is itself a random variable, because, despite knowing that the truth has been revealed, it is not possible to know which of the model realizations is this truth; all members of the initial ensemble are potentially true in the design phase. Hence, the VOC for a particular case may be higher or lower than the expected VOC.

In a similar fashion, Figure 2.4 (top right), Figure 2.4 (bottom left) and Figure 2.4 (bottom right) display the VOI, the uncertainty reduction in NPV and the observation impact as a function of observation time t_{data}. In Figure 2.4 (bottom right), the peak in the observation impact indicates that production data is most informative around t_{data} = 30; in Figure 2.4 (bottom left), the uncertainty reduction follows the same trend; and, in Figure 2.4 (top right), the VOI also increases at the same time. This suggests that, in this example, measurements with a higher observation impact also result in a larger uncertainty reduction in NPV and a higher VOI. However, whereas the observation impact and the uncertainty reduction both peak around t_{data} = 30 and gently decrease afterwards, the VOI exhibits a more abrupt decrease, similar to what is observed for the VOC. This indicates that the VOI depends not only on the information content of the observations but also on their timing, just as was discussed for the VOC. Moreover, these results illustrate that the proposed workflow allows to take both information content and timing into account and, therefore, results in a VOI assessment more complete than the analyses proposed in previous works related to reservoir management applications.

Figure 2.5 (left) shows the same results, but focusing on the expected (or mean) values of VOC (black) and VOI (blue). This plot clearly illustrates that the expected VOC is always an upper bound to the expected VOI, which is an anticipated result provided that the same set of plausible truths is considered in both VOI and VOC analyses. Indeed, production data, no matter how accurate, can never reveal all uncertainties. After water breakthrough, production data is more informative and it is more likely that the uncertainties influencing the optimization of the production strategy be revealed; thus, information more closely approaches clairvoyance. Figure 2.5 (right) illustrates this in a different way by displaying the chance of knowing (COK), defined as the ratio VOI/VOC (Bhattacharjya et al., 2010).

The different information measures suggest in this case that the most valuable measurements are the ones around t_{data} = 30. We conclude that a decision maker analyzing when to obtain a production test to optimally operate this reservoir should take a measurement around this time and should be willing to pay at most approximately $ 80 – and not $ 4,000 as the uncertainty reduction analysis would suggest (Figure 2.4 (bottom left)). Note that the model we used in this example is very simple. The optimal strategies

for the different realizations are quite similar, which means that the robust strategy (the one that maximizes the mean NPV of the ensemble) is already quite good for all the realizations. For that reason, in this case, the additional information does not lead to a significant improvement in the production strategy.

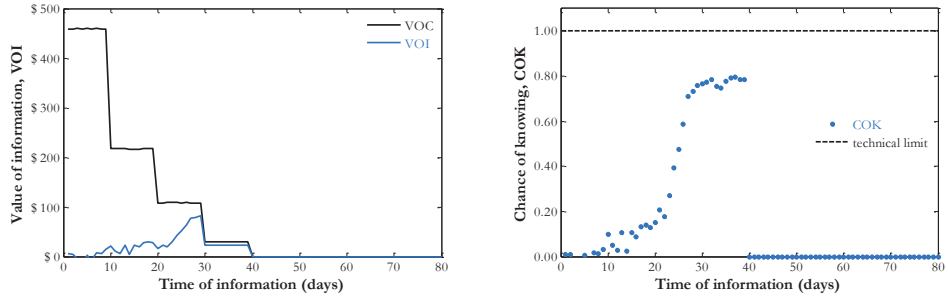


Figure 2.5: Results for the toy model: the expected VOI is upper-bounded by expected VOC (left); the ratio of VOI and VOC results in the COK (right).

2.4.2. 2D five-spot model

As a next step, we applied the proposed VOI workflow to a simple reservoir simulation model representing a two-dimensional (2D) inverted five-spot water flooding configuration; see Figure 2.6. In a 21×21 grid (700×700 m), with heterogeneous permeability and porosity fields, the model simulates the displacement of oil to the producers in the corners by the water injected in the center. Table 2.2 lists the values of the physical parameters of the reservoir model. We used $N_{truth} = 50$ plausible truths and $N_{truth} = 50$ ensembles of $N = 49$ realizations of the porosity and permeability fields, conditioned to hard data in the wells, to model the geological uncertainties. The simulations were used to determine the set of well controls (bottom-hole pressures) that maximizes the NPV. The economic parameters considered in this example are also indicated in Table 2.2. The optimization was run for a 1,500-day time horizon with well controls updated every 150 days, i.e. $M = 10$, and, with five wells, $\mathbf{u}$ has 50 elements. We applied bound constraints to the optimization variables (200 bar $\leq p_{prod} \leq$ 300 bar and 300 bar $\leq p_{inj} \leq$ 500 bar). The initial control values were chosen as the average of the upper and lower bounds. The whole exercise was performed in the open-source reservoir simulator MRST (Lie et al., 2012), by modifying the adjoint-based optimization module to allow for robust optimization and combining it with the EnKF module to create a CLRM

environment for VOI analysis. The average NPV for the initial ensemble is $ 53.5 million when using base line control (i.e. the average of the upper and lower bounds on the bottom hole pressures: 400 bar in the injector and 250 bar in the producers) and $ 55.7 million when using robust optimization over the prior (i.e., without additional information). Just like for the toy model example, the workflow was repeated for different observation times, t_{data} = {150, 300, … , 1,350} days. For this 2D model we assessed the VOI of the production data (total flow rates and water-cuts) with absolute measurement errors (ε_{flux} = 5 m³/day and ε_{wct} = 0.1). The VOI, the VOC, the observation impact I_{GAI}, and the uncertainty reduction $\Delta\sigma_{NPV}$ were computed for each of the nine observation times.

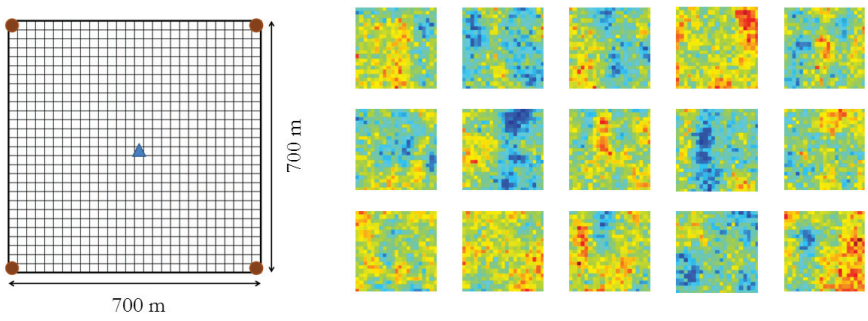


Figure 2.6: 2D five-spot model (left); 15 randomly chosen realizations of the uncertain permeability field (right).

Figure 2.7 depicts the results of the analysis for production data. Again, dashed lines correspond to expected values and solid lines to percentiles quantifying the uncertainty of the information measures. The markers correspond to the observation times at which the analysis was carried out. In Figure 2.7 (top left) we note that, like for the toy model example, clairvoyance loses value with observation time, following the previously described stepwise behavior. In addition, by observing the percentiles, we realize that, in this case, the VOC has a non-symmetric probability distribution. The high values of P_{10} indicate that, for some realizations of the truth, knowing the truth can be considerably more valuable than indicated by the expected VOC; however, the P_{50} values, which are always below those of the expected VOC, indicate what is more likely to occur. The same holds for the VOI, as can be observed in Figure 2.7 (top right). The observation that

provides the best VOI is the one at t_{data} = 150 days. Note that in our example the earliest observation seems to be the most valuable one, but that this may be case-specific.

Table 2.2: Parameter values for the 2D five-spot model.

Rock-fluid parameters		Initial conditions	
ϱ_o = 800	kg/m^3	p_0 = 300	bar
ϱ_w = 1,000	kg/m^3	S_{oi} = 0.8	[–]
μ_o = 0.5	cP	S_{wi} = 0.2	[–]
μ_w = 1	cP		
n_o = 2	[–]	Economic parameters	
S_{or} = 0.2	[–]	r_o = 80	\$/bbl
$k_{ro,or}$ = 0.9	[–]	r_{wp} = 5	\$/bbl
n_w = 2	[–]	r_{wi} = 5	\$/bbl
S_{wc} = 0.2	[–]	b = 0.15	[–]
$k_{rw,wc}$ = 0.6	[–]		

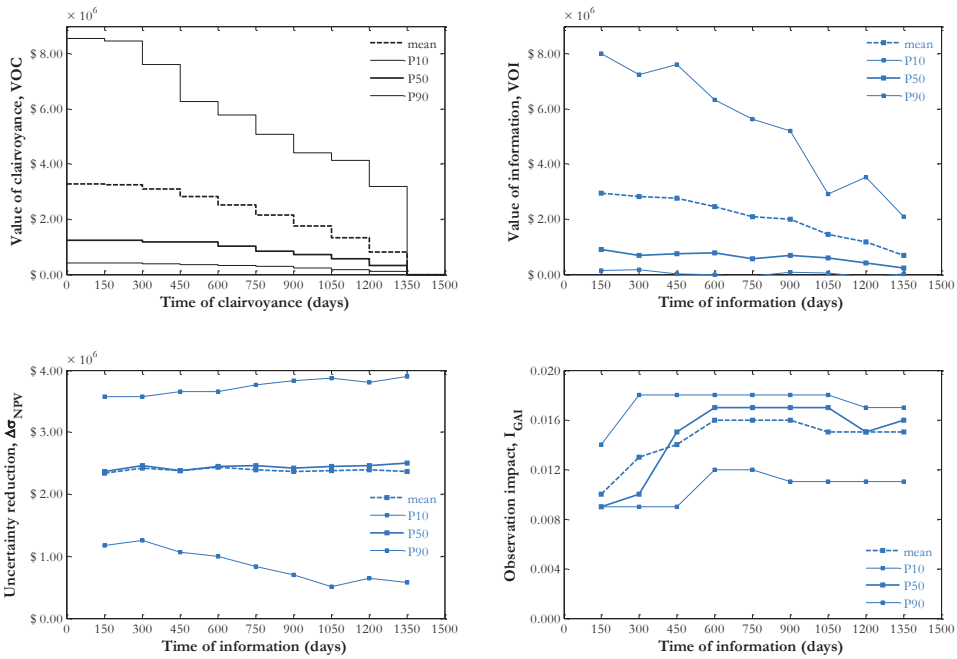


Figure 2.7: Results for the VOI analysis of production data in the 2D five-spot model: VOC (top left); VOI (top right); uncertainty reduction (bottom left); observation impact (bottom right).

Figure 2.7 (bottom right) shows that the information content of the production data increases when water breaks through in the producers but gently decreases thereafter. The observation impact achieves its maximum at $t_{data} = 600$ days; this is the time when, on average, most of the realizations have already experienced first water breakthrough. Figure 2.7 (bottom left) displays the uncertainty reduction in NPV where the initial uncertainty is $\sigma_{NPV,ini} = \$\ 4.1$ million. We observe that, on average, the uncertainty reduction is approximately the same for all the observation times considered in the analysis, which suggest that the measurements gathered at different times are equally useful to reduce the uncertainty in NPV.

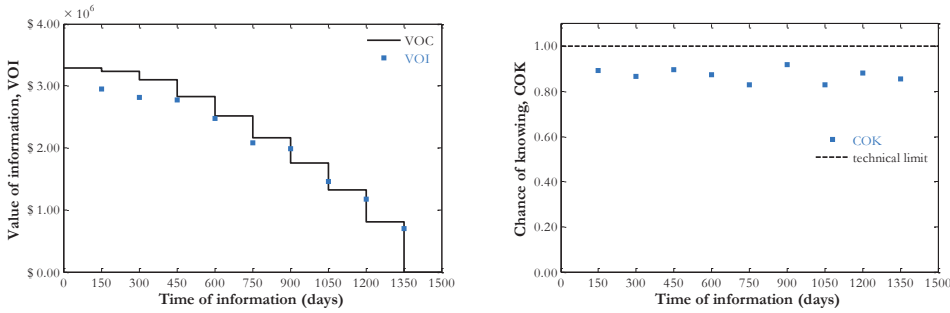


Figure 2.8: Results for the 2D five-spot model: the expected VOI is upper-bounded by expected VOC (left); the COK (right) is less informative than for the toy model (c.f. Figure 2.5).

Figure 2.8 (left) depicts the expected values of VOI (blue dots) and VOC (black line). The plot confirms that clairvoyance can be considered the technical limit for any information gathering strategy and that the expected VOC forms an upper bound to the expected VOI. We also note that the expected VOI comes closer to the expected VOC with time. Indeed, as water breakthrough is observed in more producers, the production data of this five-spot pattern become more effective in revealing the main features of the true permeability and porosity fields. Figure 2.8 (right) displays the COK with time. Although the VOI clearly approaches the VOC, their ratio does not change substantially with time, unlike what was observed for the toy model example.

In contrast to the toy model case, for this example the different information measures indicate different times for the most useful measurements. This shows that taking into account the update of the optimal production strategy can influence the conclusions drawn by this kind of analysis. Using the proposed workflow as the reference for VOI

assessment, for this case, we recommend the production data to be collected at t_{data} = 150 days and we estimate this additional information to be worth $ 2.9 million.

2.5. Discussion

2.5.1. Accelerated procedure

We observe an opportunity to reduce the number of simulations required in the proposed workflow by using the complete initial ensemble to perform a single prior robust optimization (rather than performing the robust optimization N_{truth} times for N_{truth} priors). For instance, in our 2D example, we could reduce the number of prior robust optimizations from 50 to 1, which represents a significant reduction of the computational costs of the VOI assessment procedure: approximately 420,000 simulations for the original formulation and 215,000 for the modified formulation to compute the VOI for one observation time; and approximately 1.7 million simulations for the original formulation and 1.5 million simulations for the modified formulation to compute all the VOI values depicted in Figure 2.7 and Figure 2.8 (9 observation times).

Figure 2.9 displays the results for the VOI analysis of production data in the 2D model using the accelerated procedure. They are nearly identical to those of Figure 2.7, which were obtained using the rigorous procedure, although the uncertainty in the various measures is somewhat under-estimated as can be seen from the difference between P_{10} and P_{90} values, especially at later times. This reduction in uncertainty results from the use of a single control strategy (based on a single prior robust optimization) instead of a set of different control strategies as obtained in the rigorous procedure; see Figure 2.10.

The flowchart in Figure 2.11 describes the accelerated procedure. Note that the use of a single ensemble in the accelerated procedure only concerns the computation of the prior control strategy. The remainder of the procedure (columns 2 and 3 in Figure 2.11) is still performed using N_{truth} different ensembles.

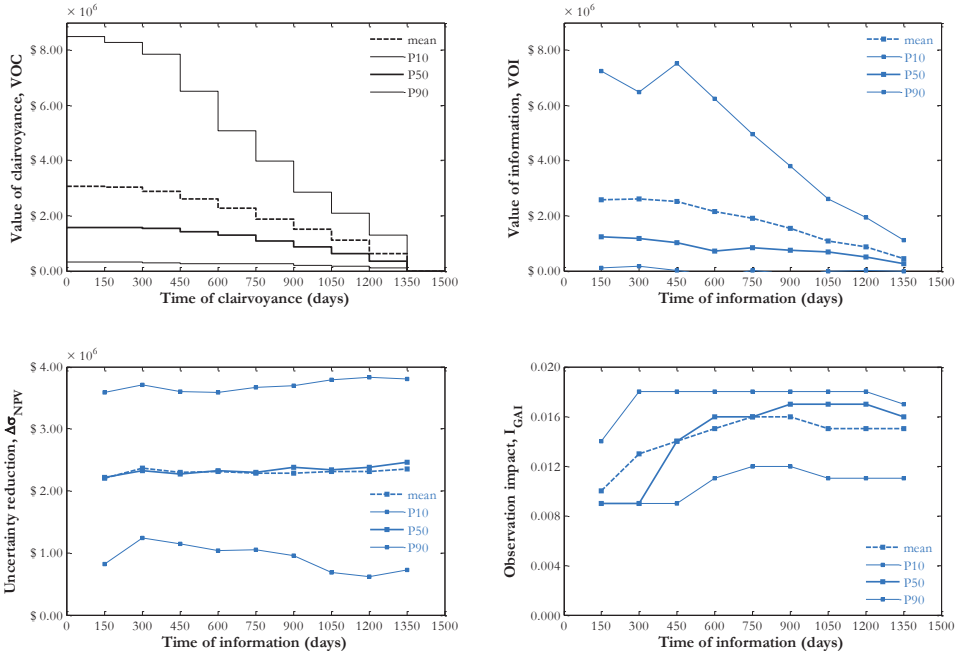


Figure 2.9: Results for the VOI analysis of production data in the 2D five-spot model using an accelerated procedure; VOC (top left); VOI (top right); observation impact (bottom left); uncertainty reduction (bottom right). The results are nearly identical to those of Figure 2.7 although the uncertainty in the various measures is somewhat under-estimated.

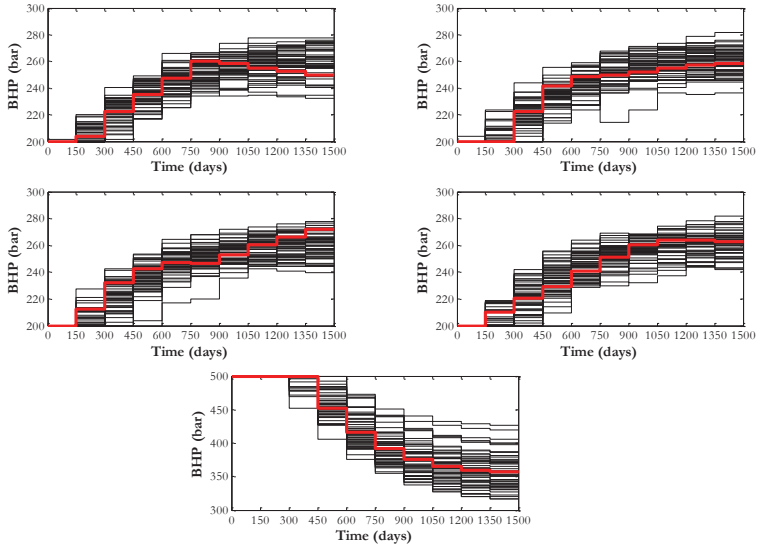


Figure 2.10: Optimal well controls (BHP) for the 50 different prior ensembles in the 2D model example: the black lines indicate the controls for each one of the ensembles (rigorous procedure) and the red line corresponds to the single set of well controls considered in the accelerated procedure.

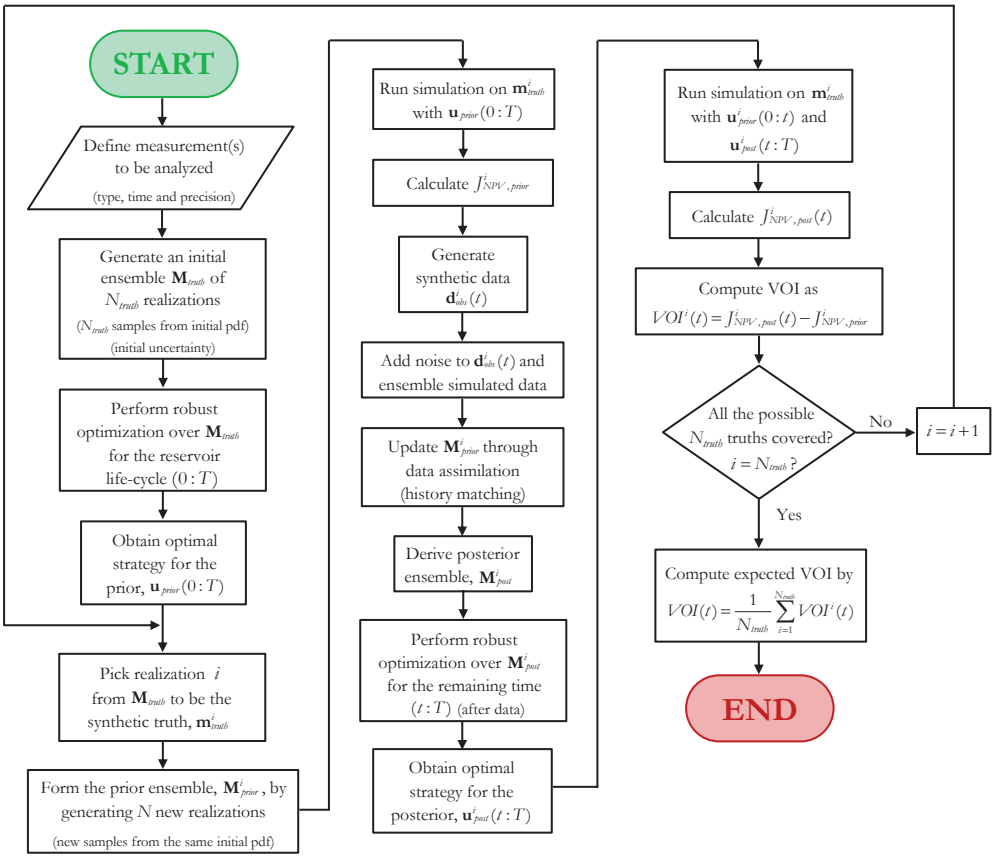


Figure 2.11: Accelerated procedure to compute the expected VOI.

2.5.2. Use of VOI assessment

Our proposed methodology is a first step towards the development of a framework to support the design of reservoir surveillance plans. In the previous sections, we showed how the VOI assessment can be used to determine the best time to collect measurements if we have the opportunity to make a single observation in the future. For that, we have repeated the VOI analysis for different observation times. The design of reservoir surveillance plans also involves other decisions concerning the specification of sensors to be deployed, such as the measurement precision, location, etc. Our methodology presented in section 2.3 can also be used to determine the best option for these choices.

When choosing the precision and accuracy of the measurements, it seems obvious that the most precise observations tend to be the most valuable ones. On the other hand, more

precise measurements will likely be more costly, because they may require the use of more expensive technologies to deliver better precision. Thus, there may be cases in which it is more advantageous to compromise the precision for the sake of costs. In this context, a cost-benefit analysis for measurements with different precision can be useful, and VOI assessment can be used of quantify the benefit. To illustrate such an application, we repeated the exercise from section 2.4.2 for production measurements (i.e., total rates and water-cuts) with three different measurement errors: $\varepsilon_{flux} = \{5,\ 10,\ 15\}$ m³/day and $\varepsilon_{wct} = \{0.1,\ 0.15,\ 0.2\}$. The results, in terms of expected VOI, are displayed in Figure 2.12. We note that, for observations until $t_{data} = 600$ days, the measurements with different measurement errors have approximately the same VOI. The value of these earlier observations is related to the ability of detecting water breakthrough in the producers, and this seems to be equally possible for all the three measurement errors that we considered. This suggests that there is no advantage in collecting highly precise measurements at this early stage. After that (i.e., $t_{data} \geq 750$ days), since water breakthrough has already occurred in most of the wells, the added value of measurements is more related to the ability of determining how much water the wells are producing, and poorer measurement precision results in lower VOI. This means that, depending on the difference in costs, it may be more interesting to invest in higher precision for late measurements.

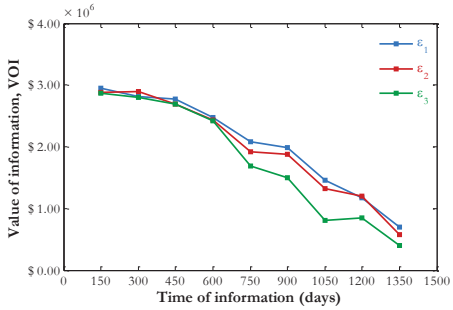


Figure 2.12: Expected VOI of production observations with different measurement precisions (2D five-spot model).

We can also use our methodology to determine where to collect the most valuable information. For that, we studied the VOI for production measurements at different locations. Once again, we used the same 2D five-spot example, but now considering the

case where a single observation for only a single well is available in the future, and then repeating the analysis for each one of the five wells of the model.

Figure 2.13 shows the results in terms of expected VOI. First, we observe that the expected VOI of measurements at the injector are always negative (Figure 2.13 (right)), which suggests that the information obtained from these observations is misleading and the additional knowledge results in worse decisions. In practice the added value of these measurements would be zero, because a decision maker facing the results of this VOI assessment would choose to implement the production strategies determined under prior uncertainty instead of investing in additional information that is expected to result in worse updated strategies.

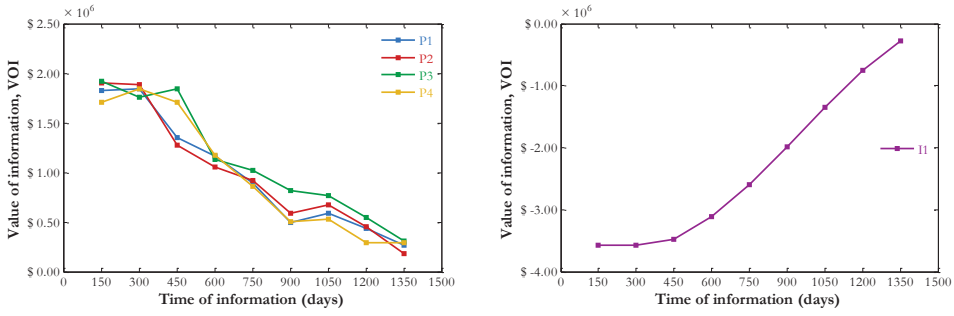


Figure 2.13: VOI of production measurements at different locations (2D five-spot model): total production rates and water-cuts measured at the producers (left) and water injection rates measured at the injector (right).

Figure 2.13 (left) shows that the expected VOI of production measurements at the different producers are about the same. The most significant difference occurs at t_{data} = 450 days, which can be explained by the different water breakthrough times at each one of the producers. Besides that, we note that the VOI of the measurements at individual wells is lower than the VOI of the measurements at all wells together (Figure 2.8). Finally, the results in Figure 2.14 confirm that VOI is not additive: the sum of the VOI of the measurements at different locations is not equal to the VOI of the measurements at all the different locations combined.

Another choice when designing reservoir surveillance plans concerns what to measure. In our original 2D five-spot example, we assumed that both total rates and water-cuts were measured. However, there may be cases where other observations are available. In

principle, we can assess, with our methodology, the VOI of any measurement that can be modeled with the reservoir simulator and incorporated in a history matching exercise. By repeating the VOI assessment for different measurement types, we can determine which sensing technology is likely to be more useful to improve the reservoir management. Here, we considered a new case where only oil production rate measurements are available at the producers. As before, we used the same 2D five-spot model as an example. Figure 2.15 depicts the results obtained in comparison with the results of the original example. The VOI of oil rate measurements can be considerably lower than the VOI of the production measurements considered previously. The low values, especially for the measurements at t_{data} = 300 days, are related to the inability of the oil rate measurements alone to accurately inform about the water breakthrough time at the producers, whose prediction is paramount to achieve an optimal reservoir management in this example.

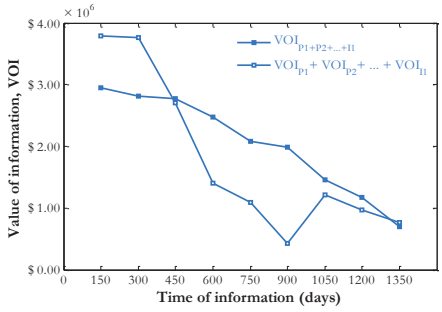


Figure 2.14: VOI is not additive: the sum of the VOI of production measurements at different locations is not equal to the VOI of the measurements at all the different locations combined (2D five-spot model).

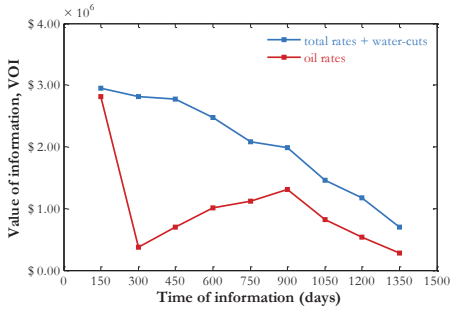


Figure 2.15: VOI of different types of production measurements (2D five-spot model): total production rates and water-cuts (blue) or only oil production rates (red).

2.5.3. Other practical issues

The successful application of our methodology requires attention to some practical issues. One point concerns the use of reactive control to shut-in wells that reach uneconomical production (e.g., high water-cuts causing production costs to exceed the production revenues). In our workflow, the production strategies are determined by optimization over an ensemble of model realizations for a fixed life-cycle T and then applied to the respective plausible truth. These strategies are likely not the optimal ones for the plausible truths and their direct application without any additional consideration (i.e., stopping production before T if necessary) may result in uneconomical production.

The second issue related to the uneconomical production refers to the plausibility of future measurements. In practice, if a given plausible truth reaches uneconomical production before the time t_{data} considered in the VOI assessment, we should consider that the future measurement would never be collected and that the VOI (and VOC) associated with that plausible truth is equal to zero.

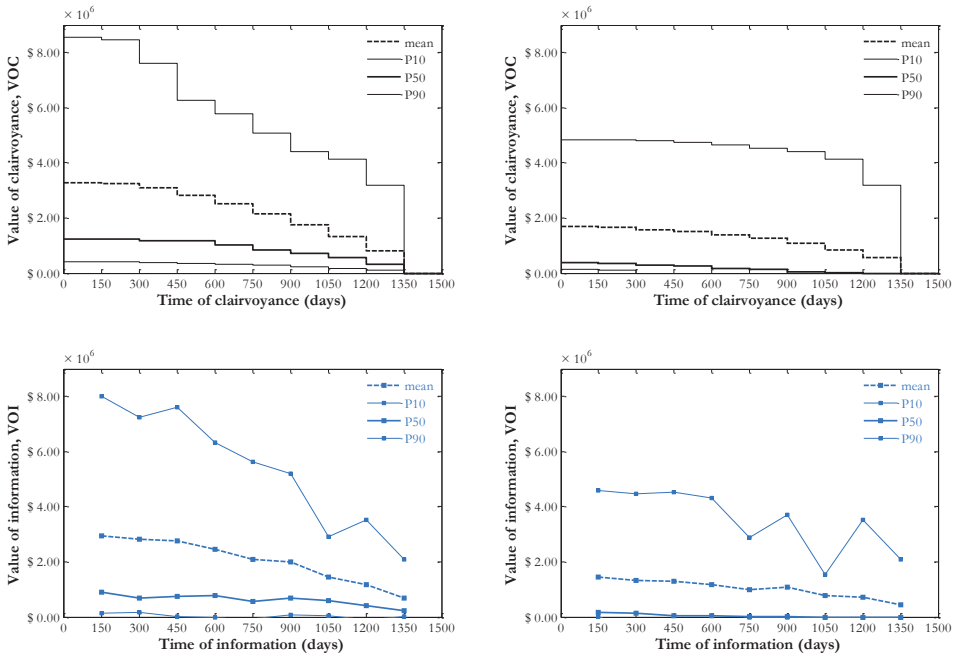


Figure 2.16: Effect of disregarding uneconomic production when assessing the VOI of future measurements (2D five-spot model): including uneconomical production (left) and disregarding it (right).

In Chapter 3 we will address the two points discussed above in more details. For now, we apply these considerations to the example of section 2.4.2 to show their impact on the VOI assessment. Figure 2.16 shows the previous results for VOC and VOI which include the uneconomical production (left) and the new ones which disregard it (right). We note that the new values are lower. This happens because the uneconomical production explained above occurs more often when applying the optimal strategies based on prior knowledge, which correspond to the baseline for the VOI assessment. From the moment we disregard the uneconomical production, the baseline values increase and the VOI decreases. In addition to that, because of the second issue described above, the VOI associated with some plausible truths are set to zero, which also contributes to reduce the VOI and VOC.

2.6. Conclusions

We proposed a new workflow for VOI assessment in CLRM. The method uses elements available in the CLRM framework, such as history matching and robust optimization. First, we identified the opportunity to combine these elements with concepts of information value theory to create a VOI analysis instrument. We then designed a generic procedure that can, in theory, be simply implemented in a variety of applications, including our optimal reservoir management problem. Next, the workflow was illustrated with two examples and the results were compared with previous measures for information valuation. Because we take into account that the production strategy is updated after new information has been assimilated in the models, our proposed method is more complete than previous work to estimate the VOI in a reservoir engineering context. In the end, we also discussed how this approach could be used to support the design of reservoir surveillance plans and we commented on some aspects to be taken into account when applying this methodology in practice.

The main drawback of our proposed VOI workflow is its computational costs; it involves the repeated application of robust optimization and data assimilation, which requires a very large number of reservoir simulations. Depending on the types of optimization and assimilation methods used (e.g., adjoint-based, ensemble-based, or gradient-free) there may be large differences in the computational requirements, but even in case of using the

most efficient (i.e., adjoint-based) algorithms, the computational load of the workflow will be huge. For instance, for the toy model example with ensembles of 100 realizations, we ran more than 50 million forward simulations (8,100 robust optimizations with EnOpt) in order to obtain the 80 values of VOI displayed in Figure 2.4 (top right). Hence, if the method is to be applied to real-field cases, some serious improvements regarding the number of simulations required are necessary. In this chapter, we showed a first step in this direction by suggesting a way to reduce the number of robust optimizations necessary. However, more has to be done. One potential method could be to use clustering techniques to select a few representative realizations rather than a full ensemble, as it will be discussed in Chapter 4. Furthermore, reduced-order modelling techniques to generate surrogate models could facilitate the application of our workflow to larger reservoir models by reducing the number of full reservoir simulations. Despite its computational cost, we conclude that our approach constitutes a rigorous VOI assessment for CLRM. For this reason, we recommend that it be used as the reference for the development of more practical and less computationally demanding tools to be applied in real-field cases.

3

Value of information for multiple observation times

This chapter [1] *extends our methodology for VOI assessment in CLRM to estimate the added value of performing multiple measurements along the producing life of the reservoir. The new procedure is based on the workflow presented in Chapter 2 which allows to quantify the VOI of a single observation under geological uncertainty. Here we show that, by modifying that workflow slightly, it is possible to assess the value of a series of measurements without a prohibitive increase in computational costs. The approach is illustrated with two cases based on a simple water flooding problem in a two-dimensional five-spot reservoir: the first one, in which we assess the value of a series of production measurements, and the second one, in which we estimate the additional value of water front positions tracked by an interpreted time-lapse seismic survey.*

3.1. Introduction

Through the last few decades, new sensing technologies have been boosting the flow of detailed reservoir information from oil and gas fields. Different types of well-based sensors and field-wide data have become available, which, in combination with the development of numerical techniques for reservoir model-based optimization and history matching, increase the possibilities to monitor and improve reservoir management operations. Because many of these sensing technologies come at significant costs, the design of reservoir surveillance strategies has been receiving more attention. The solution of this design problem requires the ability of assessing the value of future measurements

[1] This chapter is based on Barros, E.G.D., Leeuwenburgh, O., Van den Hof, P.M.J. and Jansen, J.D. (2015). Value of multiple production measurements and water front tracking in closed-loop reservoir management. Paper SPE-175608-MS presented at SPE Reservoir Characterization and Simulation Conference and Exhibition, Abu Dhabi, UAE, 14-16 September.

during the field development planning (FDP) phase of an oil field. In this context, techniques to quantify the value of information (VOI) under geological uncertainty become increasingly important.

An additional complexity arises when it is attempted to quantify the VOI associated with the deployment of a sensor that is expected to provide multiple measurements in its lifespan. In Chapter 2 we introduced a methodology to make use of the closed-loop reservoir management (CLRM) framework (i.e., under the assumption that frequent life-cycle optimization will be performed using frequently updated reservoir models) to assess the VOI of a single measurement. The first contribution presented in this present chapter is to show how that methodology can be extended to also assess the value of a series of measurements at the same location.

Another challenge consists of quantifying the value of field-wide sensing methods, such as time-lapse seismic surveys. In Chapter 2 we illustrated the VOI assessment methodology with an example considering only production data from the wells. Here we propose to repeat the workflow in combination with the work of Leeuwenburgh and Arts (2014) so that we are also able to analyze the value of water front tracking measurements, which can be an alternative to assess the VOI of a time-lapse seismic survey.

For a recap of what has been done in terms of VOI assessment of a single measurement, we refer to Chapter 2. In the Background section (3.2), we summarize previous work on the assimilation of water front tracking "measurements" that will be addressed later in this chapter. Next, in the Methodology section (3.3), we explain how to extend our workflow for a series of measurements and thereafter, in the Examples section (3.4), we illustrate it with some case studies and we analyze the results. Finally, in the Discussion (3.5) and Conclusions (3.6) sections, we comment on the main practical issues regarding the application of our methodology, and we suggest directions for future work.

3.2. Background

3.2.1. Water front tracking measurements

Water front tracking "measurements" represent a good alternative to incorporate information from interpreted 4D (or time-lapse) seismics. Time-lapse seismic surveys fall

under the category of field-wide sensing methods and have been showing great potential for reservoir surveillance. With a much higher volume of data, they provide a different type of information than the well-based measurements. A typical 4D seismic dataset gives good insight into the evolution of pressures and saturations in the whole extension of the reservoir. It may be a challenge to assimilate all these data to the reservoir models using classical history matching techniques such as the ensemble Kalman filter (EnKF). Recognizing that one of the main advantages of having 4D seismic data available lies on the possibility of interpreting saturation maps to track the water front in water flooding problems, Leeuwenburgh and Arts (2014) have proposed a distance parametrization to facilitate the history matching of seismic data (i.e., interpreted saturation fields). They use the fact that the water front can be detected in the seismic data, as a jump in the saturation values. Then they apply "a fast marching method (…) to calculate distances between observed and simulated fronts, which are used as innovations in the EnKF". Figure 3.1 depicts the general idea behind their work. This new parametrization allows Leeuwenburgh and Arts (2014) to simplify the assimilation of the time-lapse seismic dataset, without having to use any localization or inflation schemes to ensure the good behavior of the history matching ensemble techniques. For a detailed description of the fast marching method used to track the evolution of interfaces (water fronts in our setting), see Sethian (1996).

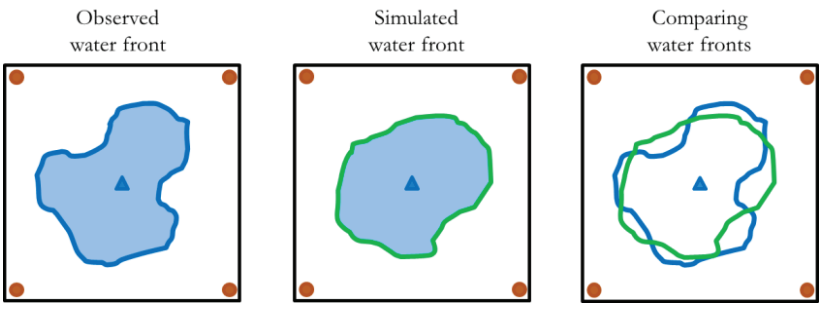


Figure 3.1: General idea behind the work of Leeuwenburgh and Arts (2014): detecting water fronts and comparing them in terms of a certain distance/metric to perform history matching.

3.3. Methodology

The new procedure is based on the workflow from Chapter 2 which allows to quantify the VOI of a single observation under geological uncertainty. Throughout the description of

the method and the illustrative examples presented there, we proposed to "close the loop" once, considering new information to become available at one observation time. Then, in section 2.4, we repeated the procedure for several observation times to show how the VOI changes in time and to determine which moment in time would be best to collect data (if it were to be collected only once). We also showed there the importance of repeating the workflow for different plausible realizations of the truth, which made our procedure very expensive in terms of computational costs.

In order to assess the VOI of a series of future measurements, we propose here to "close the loop" multiple times, repeating the original procedure from Chapter 2 while gradually progressing over the producing life of the reservoir in time steps equal to the specified control time intervals. Every time new data become available, the models are history matched and the production strategy for the remaining control intervals are updated. Figure 3.2 depicts the new workflow for VOI assessment of a series of measurements, showing a simple case with 3 control time intervals only. Note that the repetition of the procedure for an ensemble of plausible truths also holds here.

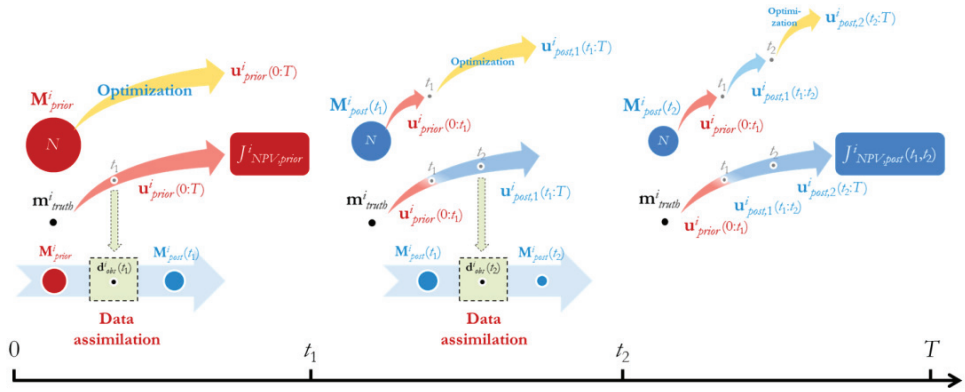


Figure 3.2: VOI assessment workflow extended for multiple observation times. (t_i indicates the observation times and T indicates the end time)

To make it more clear, we describe the steps of the extended workflow using mathematical notation. We consider the same case as described above with 3 control time intervals and two observation times at $\{t_1, t_2\}$ (Figure 3.2). Just like in Chapter 2, the first step is to determine the optimal production strategy $\mathbf{u}_{prior}(0{:}T)$ for the entire reservoir life-cycle given the prior uncertainty (equation (2.3)). Next, we generate the synthetic

measurements at t_1 by applying $\mathbf{u}_{prior}(0{:}t_1)$ to the plausible truth $\mathbf{m}_{truth}$. This allows us to perform history matching of the prior ensemble $\mathbf{M}_{prior}$ and obtain the posterior ensemble $\mathbf{M}_{post,1}$. We can then re-optimize the production strategy over $\mathbf{M}_{post,1}$:

$$\begin{aligned}\mathbf{u}_{post,1}(t_1:T) &= \arg\max_{\mathbf{u}} \mu_{NPV}\left(\mathbf{u}_{prior}(0:t_1), \mathbf{u}(t_1:T), \mathbf{M}_{post,1}\right) \\ &= \arg\max_{\mathbf{u}} \frac{1}{N}\sum_{i=1}^{N} J^i_{NPV}\left(\mathbf{u}_{prior}(0:t_1), \mathbf{u}(t_1:T), \mathbf{m}_i\right).\end{aligned} \tag{3.1}$$

In this way the "loop is closed" once, and we can restart the cycle for the second observation time t_2. Again, we generate the synthetic measurements at t_2 by applying $\left[\mathbf{u}_{prior}(0:t_1)^T\ \mathbf{u}_{post,1}(t_1:t_2)^T\right]^T$ to the same plausible truth $\mathbf{m}_{truth}$. These new measurements are then assimilated and the ensemble $\mathbf{M}_{post,1}$ is updated to derive the new posterior ensemble $\mathbf{M}_{post,2}$. And we re-optimize the production strategy over $\mathbf{M}_{post,2}$ for the remaining of control intervals:

$$\begin{aligned}\mathbf{u}_{post,2}(t_2:T) &= \arg\max_{\mathbf{u}} \mu_{NPV}\left(\mathbf{u}_{prior}(0:t_1), \mathbf{u}_{post,1}(t_1:t_2), \mathbf{u}(t_2:T), \mathbf{M}_{post,2}\right) \\ &= \arg\max_{\mathbf{u}} \frac{1}{N}\sum_{i=1}^{N} J^i_{NPV}\left(\mathbf{u}_{prior}(0:t_1), \mathbf{u}_{post,1}(t_1:t_2), \mathbf{u}(t_2:T), \mathbf{m}_i\right).\end{aligned} \tag{3.2}$$

The strategies obtained for the prior and the posterior ensembles are then tested on the synthetic truth, and their economic outcomes (NPV values $J_{NPV,prior}$ and $J_{NPV,post}$) are evaluated:

$$J_{NPV,prior} = J_{NPV}(\mathbf{u}_{prior}, \mathbf{m}_{truth}), \tag{3.3}$$

$$J_{NPV,post}(t_1) = J_{NPV}\left(\mathbf{u}_{prior}(0:t_1), \mathbf{u}_{post,1}(t_1:T), \mathbf{m}_{truth}\right), \tag{3.4}$$

$$J_{NPV,post}(t_1,t_2) = J_{NPV}\left(\mathbf{u}_{prior}(0:t_1), \mathbf{u}_{post,1}(t_1:t_2), \mathbf{u}_{post,2}(t_2:T), \mathbf{m}_{truth}\right). \tag{3.5}$$

By repeating the procedure for all the plausible truths considered the difference between $J_{NPV,post}$ and $J_{NPV,prior}$ are used, in the end, to determine the cumulative VOI through the multiple observation times:

$$\mathrm{VOI}(t_1) = \frac{1}{N_{truth}}\sum_{i=1}^{N_{truth}}\left(J^i_{NPV,post}(t_1) - J^i_{NPV,prior}\right). \tag{3.6}$$

$$\mathrm{VOI}(t_1,t_2) = \frac{1}{N_{truth}}\sum_{i=1}^{N_{truth}}\left(J^i_{NPV,post}(t_1,t_2) - J^i_{NPV,prior}\right). \tag{3.7}$$

Initially, one might think that this new workflow would be much more computationally demanding than the previous one (for a single observation), but, in practice, it is not.

Indeed, the assessment of the value of a series of future measurements implies repeating all the steps of the methodology through more observation times. However, we consider only one series of measurements. Thus, the repetition of history matching and optimization procedures while progressing over the producing life of the reservoir requires just as many reservoir simulations as repeating the workflow for a single observation for the different observation times, like we did in Chapter 2. Regarding the repetition of the procedure for different realizations of the truth, the computational cost does not increase either: once a realization $\mathbf{m}_{truth}^{j}$ is selected to be the synthetic truth, it will play the role of truth throughout the whole producing life of the reservoir in our setting. There is no need to consider new plausible truths as we progress over the control time intervals; therefore, the complexity of the procedure does not grow. Note that here we considered, for simplicity, the case where the control, model update and measurement intervals coincide. In practice, the measurements can be more frequent than the control updates; in this case all the measurements gathered over a control period can be assimilated at once or sequentially as they become available, but only re-optimizing the production strategy at the end of the period.

Referring again to Chapter 2, we also showed there that we can adapt the VOI workflow to compute the value of clairvoyance (VOC), which means that at some time of the reservoir life-cycle the true reservoir is suddenly revealed so we can perform optimization with perfect knowledge of the truth. Because clairvoyance implies perfect revelation of the truth, its value represents a "technical limit" to the VOI. The same holds when dealing with a series of measurements, with a small difference: in order to make a fair comparison with the cumulative VOI, the VOC workflow has to assume that we obtain imperfect information while clairvoyance is not yet available. In practice, this means that we must adopt, for the control periods before the time of clairvoyance, the same production strategies as the ones derived within the VOI assessment of imperfect measurement. By considering this, the VOC workflow for a series of measurements requires data assimilation and robust optimization, which makes it somewhat more complex than the original VOC workflow from Chapter 2.

Using again the simple case with 3 control intervals to explain the procedure, in this situation we consider 2 possible moments to obtain clairvoyance $\{t_1, t_2\}$. Note that, unlike

what happens for imperfect information, clairvoyance cannot be cumulated because, once the perfect revelation of the truth takes place, there is nothing else to be revealed. The procedure to compute the VOC at t_1 is exactly the same as the one described in Chapter 2. Because the truth is revealed, we can re-optimize the production strategy using the true reservoir model (equation (2.8)) and no longer a posterior ensemble:

$$\mathbf{u}_{post,1}(t_1:T) = \arg\max_{\mathbf{u}} J_{NPV}\left(\mathbf{u}_{prior}(0:t_1), \mathbf{u}(t_1:T), \mathbf{m}_{truth}\right). \tag{3.8}$$

However, if clairvoyance only becomes available at t_2, we must assume that we have been cumulating imperfect information until that moment. In this case, we consider the assimilation of imperfect synthetic measurements available at t_1, and we derive $\mathbf{u}_{post,1}(t_1{:}T)$ by re-optimizing the production strategy over the updated ensemble $\mathbf{M}_{post,1}$ (equation (3.1)). Next, the truth is perfectly revealed at t_2, and we can re-optimize the remaining of the production strategy using the true reservoir model:

$$\mathbf{u}_{post,2}(t_2:T) = \arg\max_{\mathbf{u}} J_{NPV}\left(\mathbf{u}_{prior}(0:t_1), \mathbf{u}_{post,1}(t_1:t_2), \mathbf{u}(t_2:T), \mathbf{m}_{truth}\right). \tag{3.9}$$

The other steps of the procedure (i.e., the application of the strategies to the truth and the repetition for all the plausible truths) remains the same (equations (3.3)-(3.5)) and the VOC for multiple observation times can be computed using equations (3.6)-(3.7).

3.4. Examples

3.4.1. Multiple production data

To illustrate the new procedure, we used the same 2D five-spot example described in section 2.4.2. The workflow was applied for a series of observation times, t_{data} = {150, 300, … , 1,350} days. As before, for this 2D model we assessed the VOI of the production data (total flow rates and water-cuts) with absolute measurement errors (ε_{flux} = 5 m³/day and ε_{wct} = 0.1). The cumulative VOI, the VOC, and the spread of the ensembles (standard deviation of NPV predictions, σ_{NPV}) were computed for each of the nine observation times.

Figure 3.3 depicts the results of the analysis for production data. The dashed lines represent the expected values, i.e. the mean for the N_{truth} = 50 plausible truths. The dark

solid lines and the lighter solid lines correspond to the P_{50} and P_{10}/P_{90} percentiles respectively. Here, P_x is defined as the probability that x % of the outcomes exceeds this value. The markers correspond to the observation times at which the analysis was carried out. In Figure 3.3 (top left) we can observe that clairvoyance loses value with time, following the stepwise behavior that we also found in Chapter 2. In addition, by observing the percentiles, we realize that, in this example, the VOC has a non-symmetric probability distribution. The high values of P_{10} indicate that, for some realizations of the truth, knowing the truth can be considerably more valuable than indicated by the expected VOC; however, the P_{50} values, which are always below those of the expected VOC, indicate what is more likely to occur. The same holds for the cumulative VOI, as can be observed in Figure 3.3 (top right). The first observation (at t_{data} = 150 days) shows already a significant VOI, but, as expected, this value builds up in time as more observations are incorporated in the models. Note that in our example the earliest observation seems to be the most valuable one and that the incremental values for the following observations are relatively small, but that this may be case-specific.

Figure 3.3 (bottom) shows how the spread of the ensembles of realizations changes as we progress "closing the loop" over time. We express this spread as the standard deviation of the predicted NPV of the ensemble. Note that, because of the repetition of the procedure for different plausible truths, this spread is also a random variable. The initial uncertainty is $\sigma_{NPV,ini}$ = $ 4.1 million, computed as the average of the standard deviations in the NPV of the different prior ensembles. We observe that this spread reduces significantly as the first observations are assimilated, but that later on it reaches a point from which it does not change any more.

Figure 3.4 (left) depicts the expected values of cumulative VOI (blue dots) and VOC (black line), and Figure 3.4 (right) shows the same results using a different scale of the vertical axis. The plots confirm that clairvoyance can be considered the technical limit for any information gathering strategy and that the expected VOC forms an upper bound to the expected cumulative VOI. Here, we can see more clearly that the VOC indeed decreases and that the cumulative VOI increases in time. However, the decrease of VOC over time is less significant than what we observed in the previous chapter, because here we history match the models with production data while clairvoyance is not yet available.

We also note that the expected VOC and expected cumulative VOI converge to the same value at the last observation time (t_{data} = 1,350 days), because after this point there are no controls left to be re-optimized, which means it is too late to benefit from clairvoyance or any additional information. Figure 3.4 also illustrates that collecting production data over the producing life of this reservoir is worth (on average) approximately $ 3.16 million, which represents a gain of 5.7 % compared to relying on prior knowledge to operate the field.

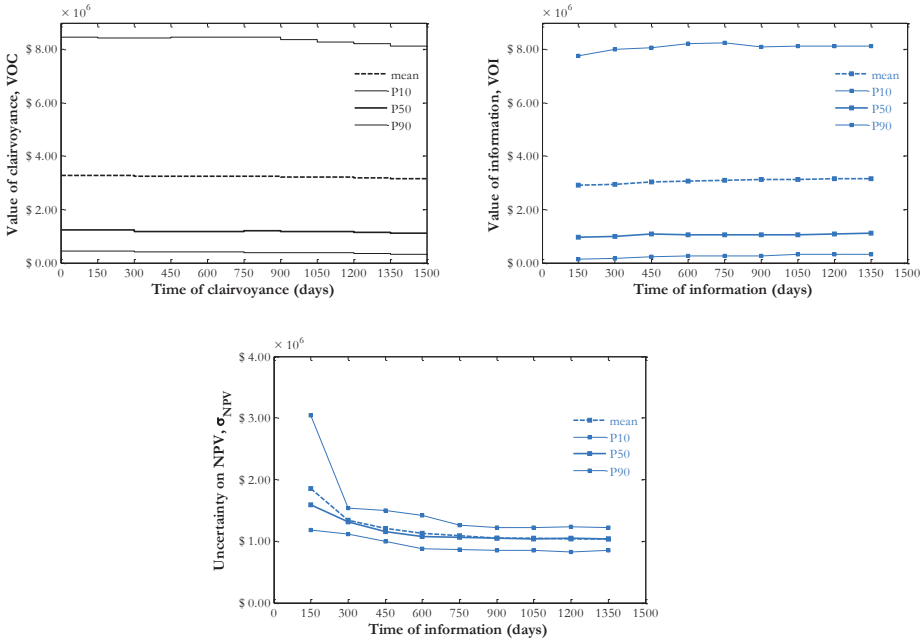


Figure 3.3: Results for the VOI analysis of production data (2D five-spot model): VOC (top left); cumulative VOI (top right); ensemble spread in terms of NPV (bottom).

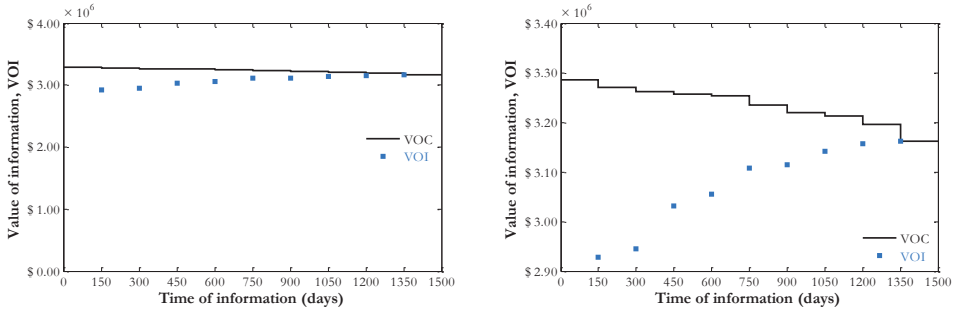


Figure 3.4: Results for the 2D five-spot model: the expected cumulative VOI is upper-bounded by expected VOC (left); same results plotted using a different scale of the vertical axis (right).

3.4.2. Multiple oil rate measurements

Next, we repeat the VOI analysis for the same case, but with more limited production data by considering oil rate measurements only (for the same 2D five-spot model and with an absolute measurement error $\varepsilon_{oil} = 5$ m³/day). By comparing the results with the previous example (total flow rate + water-cut measurements), we can estimate the additional value of also collecting accurate water production data.

Figure 3.5 (left) depicts the expected values of the cumulative VOI for the measurements from the previous example (blue dots) and for oil rate measurements only (red dots), and Figure 3.5 (right) shows the same results using a different scale of the vertical axis. We observe that measuring oil rates only is less valuable than collecting information of total rates and water-cuts, which is an expected result. We also note a difference compared to the previous example: when assimilating multiple oil rate measurements at different times, the value does not increase monotonically; see Figure 3.5 (right). The decrease around $t_{data} = 300$ days can be attributed to the fact that, at low water-cuts (immediately after water breakthrough time in the producers), oil rate measurements are not capable of detecting the presence of water and, therefore, they are not as effective as the water-cut observations in revealing the uncertainties considered here. As a matter of fact, around $t_{data} = 300$ days, most of the realizations have just observed first water breakthrough. Finally, by comparing the values of the last points ($t_{data} = 1{,}350$ days), we estimate that the additional value of collecting reasonably accurate water-cut measurements ($\varepsilon_{wct} = 0.1$) rather than just oil rates is (on average), for this case, approximately $ 440,000, which represents an increase of 1.6 % in terms of NPV.

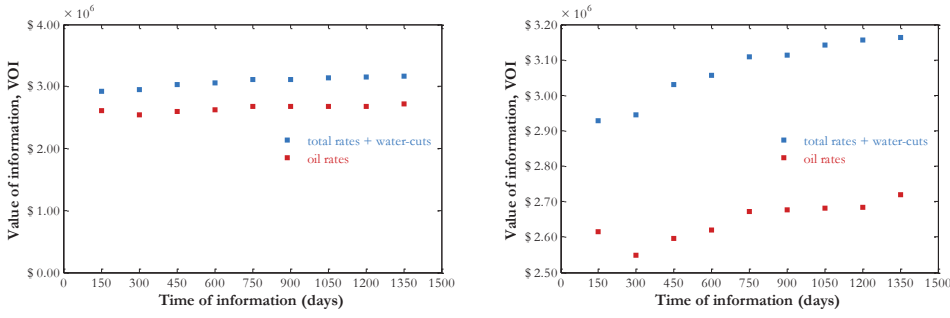


Figure 3.5: Results for the 2D five-spot model: the expected cumulative VOI of total flow rate and water-cut measurements (blue) and oil rate measurements only (red) (left); same results plotted using a different scale of the vertical axis (right).

3.4.3. Value of water front tracking "measurements"

We have taken two approaches to assess the value of an interpreted time-lapse seismic survey. In the first one, we simply repeated the procedure from Chapter 2 but using the water front tracking "measurements" from Leeuwenburgh and Arts (2014) described in section 3.2.1. The methodology for the distance reparametrization of detected water fronts is available in a modified version of the EnKF module for MRST. Once again, we used the same 2D reservoir model from the previous examples, and we adopted an absolute measurement error $\varepsilon_{dst} = 1$ [–], which refers to an error of 1 gridblock when detecting the position of the water fronts. Following the workflow from Chapter 2, we were able to estimate the VOI of these new "measurements" at different observation times (if they were to be collected only once). Figure 3.6 (top left) and Figure 3.6 (bottom left) depict the results for this analysis, showing that these "measurements" are valuable if available at early times, but not so much at later times. This is an expected result. However, by doing the analysis this way, we are actually assessing the VOI of the time-lapse seismic survey by comparing the value of acquiring a new seismic survey with the value of relying on prior knowledge only (not collecting any additional data throughout the producing life of the reservoir), which does not seem to be a realistic practice. Given the economic costs, an oil company will only consider paying to shoot a new seismic survey to monitor a reservoir if it has already decided to invest in sensor deployment to collect production data, which is a much cheaper option.

In order to make a more realistic analysis, in our second approach to assess the VOI of such "measurements", we propose to assume that the time-lapse seismic survey is acquired in addition to the series of production data measurements considered in our first example. Thus, we applied the proposed workflow to analyze the VOI of assimilating production data at multiple observation times $t_{prod} = \{150, 300, \ldots, 1{,}350\}$ days together with a single water front tracking 'measurement' at time $t_{seismic}$. Then, we considered the cumulative VOI obtained after all the observations have been assimilated (at $t_{data} = 1{,}350$ days) and compared it to the same value (at $t_{data} = 1{,}350$ days) obtained in our first example (in which the models were history matched with the production data only). Finally, we repeated the analysis for different moments in time to shoot the seismic: $t_{seismic} = 150$ days, $t_{seismic} = 300$ days, … , $t_{seismic} = 1{,}350$ days. The results obtained with this

second approach are depicted in Figure 3.6 (top right) and Figure 3.6 (bottom right). For this example acquiring a single seismic survey, in addition to performing multiple production measurements, results in an incremental VOI of, at maximum, $ 40,000 (for a survey shot at $t_{seismic}$ = 450 days). (Note that, as in all examples, this is the VOI without accounting for the costs of acquiring the data.) This value is much lower than the VOI observed in the absence of production data, which suggests that the additional seismic survey is not leading to significantly better decisions. Apparently the production data already provide sufficient information in this case.

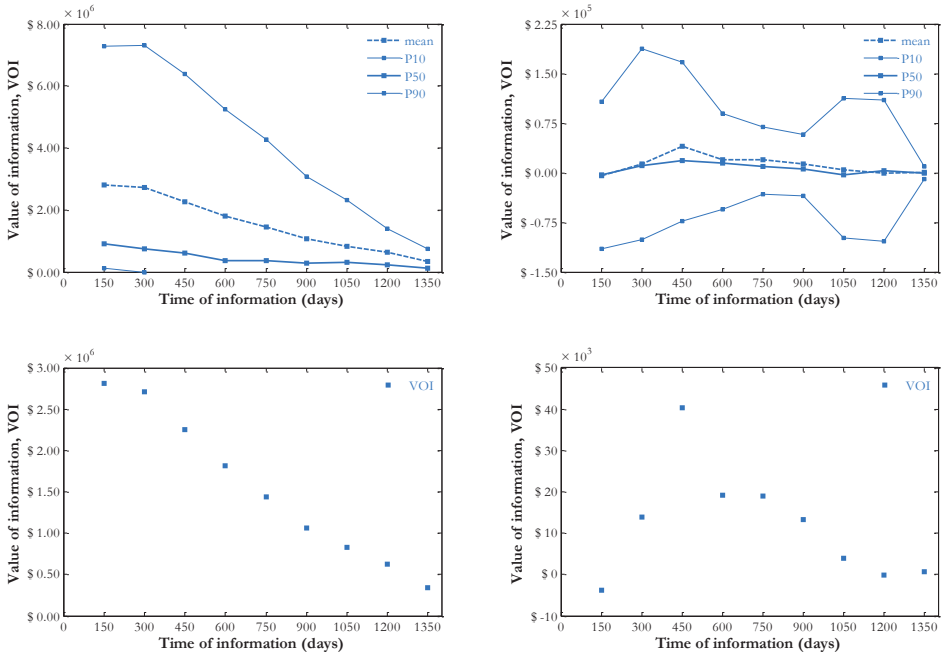


Figure 3.6: Results for the 2D five-spot model: VOI of interpreted 4D seismic data (at a single moment in time) without performing production measurements (top left: mean and percentiles; bottom left: mean only). Incremental VOI of interpreted 4D seismic data (at a single moment in time) in combination with performing production measurements (at multiple observation times) (top right: mean and percentiles; bottom right: mean only).

3.5. Discussion

3.5.1. Additional insights from VOI assessment

Besides assessing the VOI of future measurements, our approach of "closing the loop" in the design phase produces a great amount of data which could be used to gain additional

insight into the (closed-loop) reservoir management problem. Because we consider an ensemble of plausible truths, multiple CLRM sequences are obtained, including optimal production strategies (Figure 3.7), history matched ensembles and their predictions (Figure 3.8 and Figure 3.9). An in-depth analysis of these data may lead to a better understanding on how to manage the reservoir.

Figure 3.7 displays the optimal schedules of well controls (i.e., BHP over the control time intervals) for the five wells of the 2D model example. Figure 3.7 (top left) shows the strategies optimized under prior uncertainty. There are multiple strategies because we generate a different prior ensemble for each one of the plausible truths considered. Figure 3.7 (bottom) shows the production strategies optimized specifically for each one of the plausible truths, under the assumption that clairvoyance is available at $t = 0$. And Figure 3.7 (top right) exhibits the strategies obtained for each plausible truth through CLRM with additional production measurements. We observe that the strategies which benefit from additional knowledge are much more spread over the allowed range of controls compared to the ones derived with prior knowledge only. We also highlight that the strategies obtained with prior knowledge and with multiple production measurements are the same for the first control time interval because additional information only becomes available at the end of this interval.

The production strategies from Figure 3.7 can be applied to the plausible truths to simulate their production response and analyze their economic performance. For the VOI assessment, we were mainly interested in the final NPV values, but, in practice, an analysis of the production response in a broader sense can reveal patterns and mechanisms to explain the improved performance of strategies obtained with additional knowledge. Figure 3.8 displays the entire curves of cumulative NPV obtained for the different production strategies from Figure 3.7. It is clear that those strategies that are more spread over the range of controls produce a different NPV profile which results in higher final NPV values, indicating a more efficient use of the operational flexibility of the reservoir management problem. Figure 3.9 has already been discussed before: it shows the same results but disregarding any uneconomical production.

The workflow for VOI assessment generates even more information than what we show here. We believe that these data can be used in a smart way to support operational

decisions, serving as a sort of pre-computed operation manual. A second purpose for our proposed methodology could help us to promote its practice and justify its high computational costs.

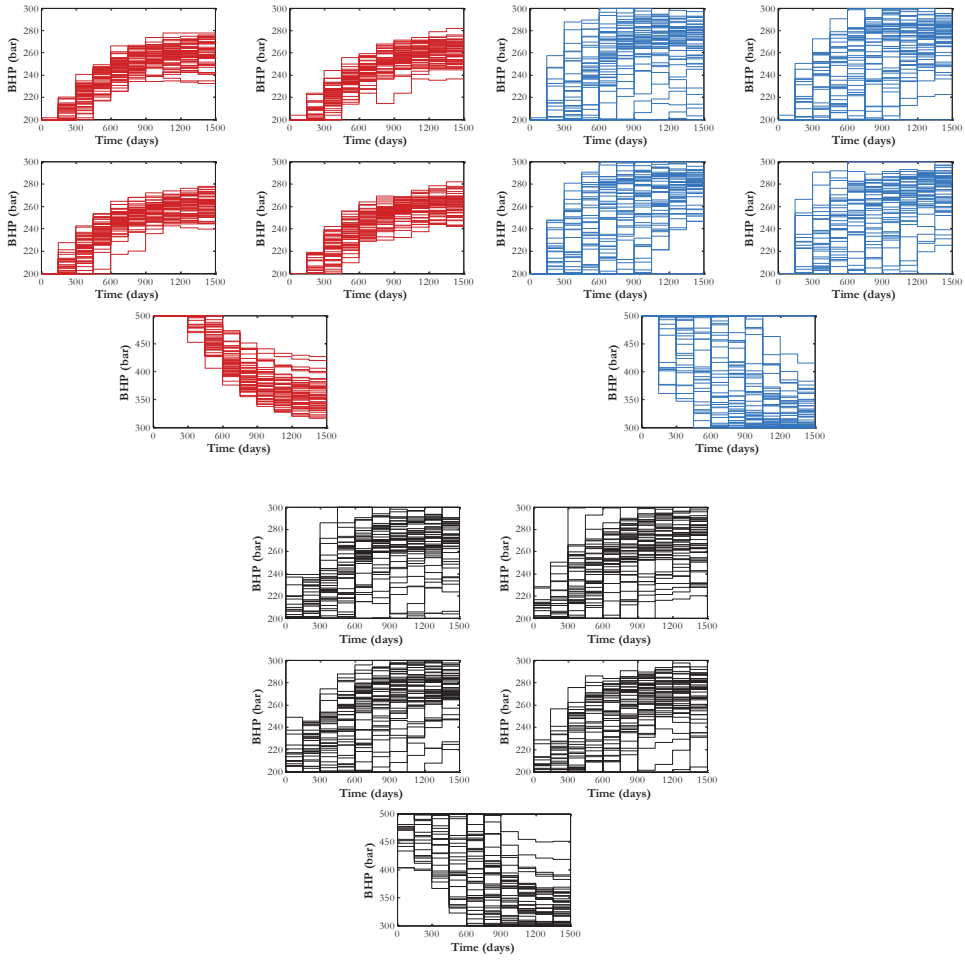


Figure 3.7: Optimal production strategies for the plausible truths considered 2D five-spot model: optimized under prior uncertainty (red – top left); obtained through CLRM with additional production measurements (blue – top right); optimized under the assumption of clairvoyance available at $t = 0$ (black – bottom). Five plots for each case, corresponding to the BHP schedules for each one of the 5 wells of the five-spot pattern.

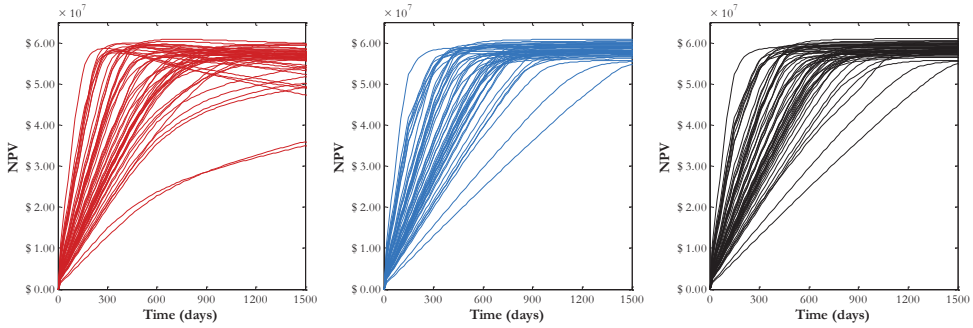


Figure 3.8: Cumulative NPV curves for the 50 plausible truths considered including uneconomical production (2D five-spot model) for the production strategies from Figure 3.7.

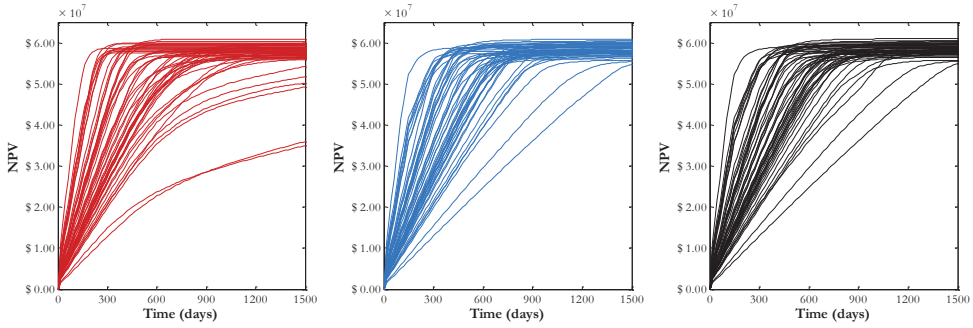


Figure 3.9: Cumulative NPV curves for the 50 plausible truths considered disregarding uneconomical production (2D five-spot model) for the production strategies from Figure 3.7.

3.5.2. Disregarding uneconomical production

In section 2.5.3, we discussed the impact of the uneconomical production of the plausible truths on the VOI assessment. Here, we make again the same considerations, but this time for the VOI assessment with multiple observations. The results are shown in Figure 3.10 and Figure 3.11 for comparison with those in Figure 3.3. Although the values are different, the same conclusions still hold: VOI and VOC are random variables and VOC constitutes an upper bound for VOI. Like we saw in section 2.5.3, we note that the values obtained with these considerations are lower, indicating that the production strategies determined with prior knowledge lead to uneconomical production. This can be confirmed in Figure 3.8 (left): some of the curves of cumulative NPV reach a maximum and start to go down. Figure 3.9 shows the cumulative NPV curves when we disregard the uneconomical production of the plausible truths. We observe that the main changes compared to Figure 3.8 are in the curves obtained with the production strategies

optimized under prior uncertainty. Those curves that before would go down now reach their maximum and stay flat because we force the production to be interrupted at that time. This results in an increase on the final NPV values that form the baseline in the VOI assessment and explains the decrease in VOI observed in Figure 3.10 and Figure 3.11. The cumulative NPV curves with additional measurements or clairvoyance remain the same, suggesting that we are able to determine better production strategies when applied to the plausible truths (i.e., which prevent uneconomical production to occur) once more knowledge becomes available.

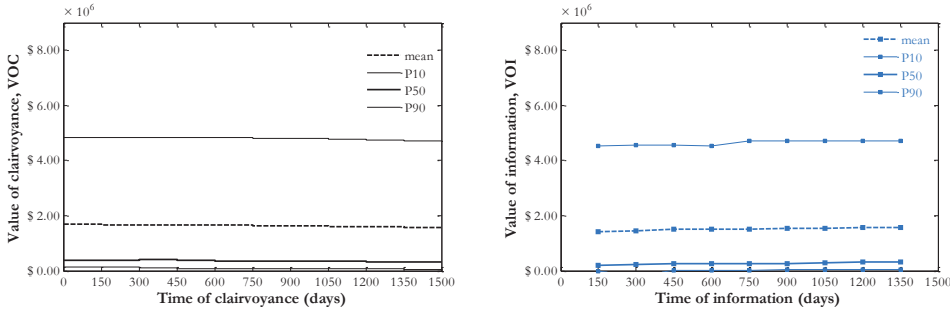


Figure 3.10: Results for the 2D five-spot model disregarding uneconomical production of the plausible truths (for comparison with Figure 3.3).

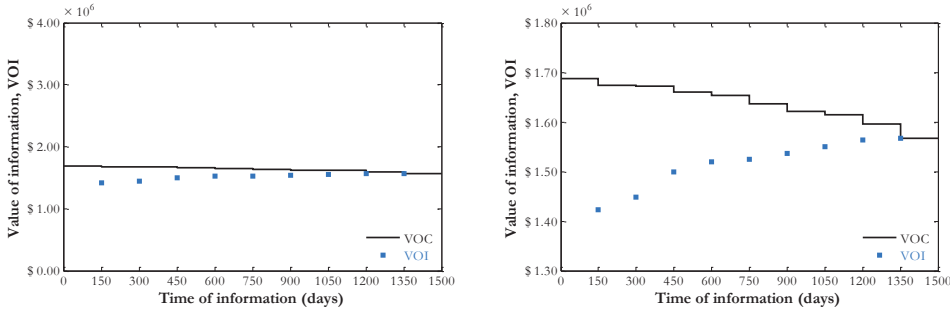


Figure 3.11: Expected VOC and cumulative VOI with multiple production measurements disregarding uneconomical production (2D five-spot model) (for comparison with Figure 3.4). VOI still increases as more observations are gathered and VOC is still an upper bound for VOI.

In Chapter 2, we also discussed a second point related to the interruption of the production when it becomes uneconomical, which could mean that for some plausible truths the future measurements would not be possible. Since, in the current example with multiple observations, the first additional measurements are incorporated at $t_{data} = 150$

days, this point is not an issue because none of the model realizations have reached uneconomical production so early (Figure 3.8 (left)).

3.6. Conclusions

We extended the work from Chapter 2 to create a new workflow that allows the VOI assessment of a series of future measurements. The method uses elements available in the CLRM framework, such as history matching and robust optimization. First, we identified the opportunity to combine these elements with concepts of information value theory to create a VOI analysis instrument. We then designed a generic procedure that can, in theory, be simply implemented in a variety of applications, including our optimal reservoir management problem. Next, the workflow was illustrated with three examples and the results were analyzed. In the third example, we assessed the additional value of a time-lapse seismic survey, which, however, in our example was limited because apparently the production data were already substantially informative. We believe that our proposed workflow is a complete methodology to estimate the VOI in a CLRM context because we take into account that the production strategy is updated periodically after new information has been assimilated in the models. However, the computational complexity of the method is, at present, prohibitively large. Future work is required to reduce the computational load, e.g. through the use of representative model selection (addressed in Chapter 4), proxy models or reduced-order modeling, to allow for the application to real-field cases.

4

Clustering techniques for value of information assessment

The application of CLRM to real-field cases can be computationally demanding. An even higher computational load results from procedures to assess the VOI in CLRM. Such procedures, which are performed prior to field operation, i.e. during the FDP phase, require extreme amounts of simulations. Therefore, we look for alternatives to reduce this computational burden. In particular we study various clustering techniques to select a limited number of representative members from an ensemble of reservoir models [1]. *Using k-means clustering, multidimensional scaling and tensor decomposition techniques, we test the effectiveness of different dissimilarity measures such as distance in parameter space, distance in terms of flow patterns and distance in optimal sets of controls. We apply several of these measures to a VOI-CLRM exercise using a simple 2D reservoir model which results in a reduction of the necessary number of forward reservoir simulations from millions to thousands. Finally, as a first step towards large-scale application, we assess the VOI in a larger benchmark case study.*

4.1. Introduction

Modern reservoir management workflows include uncertainty quantification (UQ) based on reservoir simulation models. An increasingly popular UQ practice in the reservoir engineering community uses ensembles of reservoir model realizations to account for the geological uncertainties, which, however, contributes to increasing the computational costs of these workflows. Closed-loop reservoir management (CLRM) is a combination of life-cycle optimization and computer-assisted history matching, both accounting for uncertainties and demanding a significant amount of simulations. For this reason, the

[1] This chapter is based on Barros, E.G.D., Van den Hof, P.M.J. and Jansen, J.D. (2018). Clustering techniques for value-of-information assessment in closed-loop reservoir management. Submitted for publication.

application of the CLRM framework in combination with UQ can be extremely computationally expensive. Workflows to assess the value of information (VOI) in CLRM during the field development planning (FDP) phase require even more simulations, which, at the current level of hardware development, makes real-field applications unfeasible (Chapter 2). Therefore, we look for alternatives to reduce this computational cost.

The development of more practical ways of a-priori assessing the value of future measurements has been a topic of several studies recently. Some of these have focused on the use of proxy models to reduce the number of high-fidelity reservoir simulations required for the VOI analysis (He et al., 2016, and Chen et al., 2016). Cardoso and Durlofsky (2010), He et al. (2013), Hewson (2015), and Jansen and Durlofsky (2016) investigated the use of reduced-order modeling to speed-up production optimization and history matching procedures. Others have proposed a more approximate definition of VOI which simplifies their procedure (Le and Reynolds, 2014a and 2014b). Eidsvik et al. (2015) have envisaged more sophisticated designs of experiments to be a promising alternative to alleviate the computational costs of VOI assessment workflows. Recently, Shirangi and Durlofsky (2016) presented a general framework that uses clustering techniques to determine representative models to accelerate computations for optimization under uncertainty. Insuasty et al. (2017) also showed how clustering methods based on flow-relevant dissimilarity measures can be used to form reduced ensembles. This paper explores the use of clustering techniques to select subsets of representative model realizations to speed-up production optimization and other computational procedures present in the workflow for VOI assessment introduced in Chapter 2.

In the Background section (4.2) we briefly recap our previously proposed methodology for VOI assessment in CLRM and review some previous work on cluster analysis. Next, in the Methodology section (4.3), we identify opportunities to apply clustering within the original procedure and we describe our approach to reduce the computational costs in different steps of the VOI assessment to come up with a more practical workflow. Thereafter, in the Examples section (4.4), we illustrate the application of the proposed measures to accelerate VOI calculations and we compare the results with those obtained with the original procedure.

4.2. Background

4.2.1. VOI assessment in CLRM

In Chapter 2, we presented our methodology to assess the VOI of future measurements within the CLRM framework. Our approach consists of "closing the loop" in the design phase to simulate how future information, to be obtained during the producing life-time of the reservoir, comes into play in the context of optimal reservoir management. By considering both data assimilation and optimization in the procedure, we are able "to not only quantify how information changes knowledge, but also how it influences the results of decision making" (Chapter 2). This is possible because a new production strategy is obtained every time the models are updated with new information, and the strategies with and without additional information can be compared in terms of the value of the optimization objective function (typically NPV) obtained when applying these strategies to a virtual asset (a synthetic truth).

One of the key aspects of this methodology is the idea of using an ensemble of N_{truth} "plausible truths" to account for the fact that in reality we do not know the true reservoir nor the outcome of the future measurements whose value we would like to assess. This requires extensive use of robust optimization and history matching procedures: for N_{truth} plausible truths we have N_{truth} robust optimizations under prior uncertainty, N_{truth} history matches to assimilate the future measurements and N_{truth} robust optimization given the posterior uncertainty. Note that in section 2.5.1 we proposed also an accelerated procedure where the number of prior robust optimizations is reduced from N_{truth} to 1. We consider this accelerated form of the workflow in the next sections of this chapter.

4.2.2. Model selection

We use multiple ensembles of realizations to account for geological uncertainties. Typical ensembles are formed by tens or hundreds of realizations, making the procedures involved computationally intensive. The cost, in terms of the amount of simulations required, of robust optimization and history matching algorithms tends to scale linearly with the size of the ensemble (i.e., $O(N)$), while the VOI assessment workflow described above scales with the square of the ensemble size (i.e., $O(N \cdot N_{truth}) \approx O(N^2)$). Thus, a decrease in the number of realizations considered in the analysis may lead to significant reduction in the

computational cost and make the VOI assessment problem more tractable. The challenge is how to cleverly select a subset of realizations which can represent the full ensemble to quantify the uncertainty. Others have worked on this problem; e.g., Armstrong et al. (2013) use stochastic programming with recourse to reduce the number of scenarios to be considered and Sarma et al. (2013) recommend the use of a minimax approach to efficiently select representative models from a large ensemble by matching target percentiles. This work focusses on the use of clustering techniques to automate the selection of representative model realizations, along the lines of the work of Shirangi and Durlofsky (2016).

4.2.2.1. Clustering

Cluster analysis aims to group a set of N objects

$$\boldsymbol{\Theta} = \begin{bmatrix} \boldsymbol{\theta}_1 & \boldsymbol{\theta}_2 & \cdots & \boldsymbol{\theta}_N \end{bmatrix} = \begin{bmatrix} \theta_{11} & \theta_{12} & \cdots & \theta_{1N} \\ \theta_{21} & \theta_{22} & \cdots & \theta_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ \theta_{M1} & \theta_{M2} & \cdots & \theta_{MN} \end{bmatrix}, \tag{4.1}$$

into N_{repr} clusters according to the similarity between the objects; see, e.g., Baker (2015). Note that here the objects have been chosen as vectors $\boldsymbol{\theta}_i$, $i = 1, 2, \ldots N$, in an M-dimensional space (e.g., N realizations of M grid block permeability values) but they could also be scalars, matrices or higher-order objects (tensors). Clustering has been widely used in pattern recognition, machine learning and statistics (Arabie and Hubert, 1996) and is broadly classified into partitional and hierarchical categories. As the name suggests, partitional clustering separates the objects into exclusive clusters such that the objects within a cluster are more similar to each other than to the objects in another cluster. On the other hand, hierarchical clustering, also known as connectivity-based clustering, connects objects to form clusters based on their distance. The connected objects in clusters can then be represented using a dendrogram (i.e., a diagram with a tree structure).

K-means clustering is one of the most used partitional clustering methods due to its simplicity (Caers, 2011). The user predefines a number N_{repr} of sets C_j, $j = 1, 2, \ldots N_{repr}$, that each contain a total of N_j indices corresponding to the objects belonging to each cluster, where the clusters are not necessarily of equal size. The algorithm then attempts to

iteratively improve the partitioning to achieve the lowest intra-cluster distance. This minimization problem can be formulated as follows:

$$C_{opt} = \arg\min_{C} \sum_{j=1}^{N_{repr}} \sum_{k \in C_j} d_{jk} \left(\boldsymbol{\theta}_k, \overline{\boldsymbol{\theta}}_j \right)^2 , \tag{4.2}$$

where $C = \{C_1, C_2, \ldots, C_{N_{repr}}\}$ is the set of N_{repr} clusters, i.e. a set of sets of indices, and $d_{jk}(\boldsymbol{\theta}_k, \overline{\boldsymbol{\theta}}_j)$ is the distance between one of the N_j data points within each cluster and the cluster centroid $\overline{\boldsymbol{\theta}}_j$ computed as

$$\overline{\boldsymbol{\theta}}_j = \frac{1}{N_j} \sum_{k \in C_j} \boldsymbol{\theta}_k . \tag{4.3}$$

The first step to use cluster analysis consists of choosing a feature operator F to compare the model realizations $\mathbf{m}_i$. The feature operator could just select a number of parameters (e.g., grid block permeability values) of the vectors of model parameters $\mathbf{m}$, or it could represent a more complex operation like a full simulation to compute the NPV or a sequence of saturation snapshots. Using this operator, the set of features $\Theta = [\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \ldots, \boldsymbol{\theta}_N]$ is formed, where $\boldsymbol{\theta}_i = F(\mathbf{m}_i)$. The clustering algorithm can then generate the distances required to determine C_{opt}. It has been shown that the choice of the appropriate feature operator is extremely case-dependent; thus, there is no one-method-fits-all solution.

Note that the sets $\boldsymbol{\theta}_i$ are elements of an M-dimensional space, where M can be very large. Unfortunately, most clustering algorithms do not work efficiently in higher dimensional spaces because of the inherent sparsity of the data, and as M grows, distance measures become increasingly meaningless (Keim et al., 1997; Parsons et al., 2004). A solution to this problem is to eliminate some of the dimensions of the feature space. However, if done wrongly, this may cause information loss and introduce wrong correlations between the model realizations. Aggarwal et al. (1999) have shown that the projection of high-dimensional data spaces on reduced-order subspaces can lead to improved clustering results.

4.2.2.2. Projection methods

There is more than one method to project datasets onto a reduced-order space. (Note that this is sometimes referred to as reducing the dimensionality of the "feature space".) One of them involves the use of tensor decomposition techniques. Tensor decomposition is strongly related to principal component analysis (PCA) or singular value decomposition (SVD). It enables the transformation of data into a compact representation while honoring their structure (e.g., spatio-temporal correlations) which is usually degraded with the vectorization step in SVD approaches. These techniques can be used to compress large datasets stored as tensors by constructing low-rank approximations with minimal approximation or reconstruction error. For instance we may form a dataset $\boldsymbol{\Theta}$ in the form of a tensor representation of the data, $\boldsymbol{\Theta} = F(\mathbf{m}_1, \mathbf{m}_2, \ldots, \mathbf{m}_N)$, which better preserves their structure than using a vector representation. E.g., we could construct a 3D tensor $\boldsymbol{\Theta}$ by stacking up the two-dimensional permeability fields (matrices) of an ensemble of 2D model realizations. We then perform the following minimization (Insuasty et al., 2017):

$$\min_{\varphi_{1:I},\psi_{1:J},\chi_{1:K}} \left\| \boldsymbol{\Theta} - \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K} \sigma_{ijk} \left(\varphi_i \otimes \psi_j \otimes \chi_k \right) \right\|_{\mathcal{F}} \quad \text{subject to } \varphi_{i'}^T \varphi_{i''} = \delta_{i'i''}, \psi_{j'}^T \psi_{j''} = \delta_{j'j''}, \chi_{k'}^T \chi_{k''} = \delta_{k'k''}, \tag{4.4}$$

where σ_{ijk} is the core tensor (i.e., an all-orthogonal and ordered tensor which is analogous to the coefficients matrix in classical SVD), φ_i, ψ_j, χ_k are orthonormal basis functions, $\|.\|_{\mathcal{F}}$ represents the Frobenius norm, and the symbol $\otimes$ denotes the tensor (outer) product over a vector space. This can be extended to tensors with more dimensions. For more information on tensor-based model-order reduction, we refer to Insuasty et al. (2017) and Afra and Gildin (2016), who also show that the solution to the minimization problem in (4.4) can be approximated by performing a higher-order SVD (HOSVD). In this case, the tensor is flattened (or unfolded) in a planar matrix structure where we can operate similarly to classical SVD. This allows us to determine the basis functions and the coefficients associated with them. Like in classical SVD, a truncation can then be applied to retain only those basis functions that explain the most dominant patterns in the data, thus resulting in the lower-dimensional representation we were aiming for. One of the modes of $\boldsymbol{\Theta}$ in our applications (here we assume mode k) typically refers to the model uncertainty, characterized by the N model realizations. We can apply a truncated SVD to

the covariance of the unfolded form of $\boldsymbol{\Theta}$ in this mode ($\boldsymbol{\Theta}_{(k)}\boldsymbol{\Theta}_{(k)}^T = \mathbf{U}_k\boldsymbol{\Sigma}_k\mathbf{V}_k^T$) and use the obtained coefficients ($\mathbf{U}_k$) to derive the dissimilarity measure to cluster the realizations. Insuasty et al. (2017) showed that this approach allows us to compare model realizations based on very rich datasets, such as the temporal evolution of the spatial distribution of pressures and saturations inside the reservoir. They were able to select a subset of realizations representative in terms of dynamic flow patterns and form reduced ensembles to perform robust production optimization more efficiently (Insuasty et al., 2015).

Another tool to represent model realizations $\mathbf{m}_i$ in a lower dimension is multidimensional scaling (MDS). It refers to techniques that use distance measures to produce a $\tilde{\boldsymbol{\theta}}_i$ representation of data points $\boldsymbol{\theta}_i$ in a reduced M_{MDS}-dimensional space with $M_{MDS} \ll M$. MDS was first introduced for the analysis of proximities in Shepard (1962). In recent years, the machine learning community has applied MDS for nonlinear dimension reduction. Kruskal (1976) argues that MDS can be complementary to clustering techniques. Scheidt and Caers (2009) introduced MDS in the reservoir simulation community, and since then many reservoir applications have been documented; see, e.g., Caers (2011).

When applying MDS, a measure of fit, referred to as "stress" in the MDS literature, can be calculated to quantify the conformance of the representation $\tilde{\boldsymbol{\Theta}}$ to the original data $\boldsymbol{\Theta}$, which can be a criterion to define the number M_{MDS} of dimensions of the reduced space; see Kruskal (1964). Low values of stress (i.e., below 5 %) indicate an excellent fit between dissimilarities and distances, and thus a good representation of the original samples in the reduced-dimensional space. We can then use this stress to determine the appropriate number of dimensions M_{MDS} to proceed with: we start with a small number of dimensions and increase this number until the stress value drops below an acceptable level. For further information on dimensionality reduction, including MDS and PCA approaches, we refer to Cunningham and Ghahramani (2015). Once we obtain the representation $\tilde{\boldsymbol{\Theta}}$, we perform clustering (e.g., also with the k-means method) based on the coordinates of the data points in the reduced-dimensional space.

4.3. Methodology

As discussed in the previous section, the VOI assessment workflow described in Chapter 2 and depicted in Figure 2.11 requires an excessive amount of simulations to be applied in practice. This is mainly due to the extensive use of robust optimization and history matching and to the fact that multiple plausible truths are considered. The most demanding steps constitute opportunities for considerable acceleration of the workflow. The focus of this work is on the use of model selection to achieve this goal. Thus, it is about looking for approximated results by compromising the rigor in UQ for the sake of computational speed-up.

In this section, we describe first how to select representative models to speed-up robust optimization and history matching, and then how to assess the quality of the results with the accelerated procedures. Afterwards, we explain how we can accelerate the VOI analysis by picking representative plausible truths. We also discuss the choice of the most appropriate feature to distinguish model realizations in the different parts of the workflow. Finally, we combine all these measures to come up with a new and faster VOI assessment workflow.

4.3.1. Speeding-up robust optimization

The whole idea behind accelerating robust optimization is to reduce the number of reservoir simulations required. This is done by reducing the number of model realizations in the ensemble used in the optimization.

We start with a full ensemble $\mathbf{M}_{full}$ of N model realizations. The first step is deciding on the number N_{repr} of representative realizations to form the reduced ensemble $\mathbf{M}_{repr}$. This number should reflect the speed-up factor we would like to achieve or the maximum ensemble size we can afford to use with the available computational resources.

The second step is choosing a feature F relevant to the problem to be optimized and building our dataset to apply the clustering algorithms. In our case, we are using reservoir simulation to evaluate the objective function in the water flooding optimization process. Therefore, it seems to be important to distinguish the model realizations on the basis of their simulated dynamical behavior. One option is to rely on the fact that model

parameters (e.g., permeabilities and porosities) tend to correlate with the reservoir flow characteristics and use them as the feature to distinguish the realizations $\mathbf{m}_i$. An alternative is to perform reservoir simulations and work with features associated with the dynamics of the system, in which case we may consider relying on model states or flow patterns (e.g., pressure/saturation snapshots and streamlines), or the model outputs (e.g., well production data and NPV evolution over time), as it has been studied by Shirangi and Durlofsky (2016) and Insuasty et al. (2017).

The next step is preparing the feature data set for clustering, by applying projection methods to reduce the dimensionality of the feature space if necessary. Thereafter, k-means clustering is performed. Once the N_{repr} clusters are formed, one realization is selected as the representative of each cluster, forming the reduced ensemble $\mathbf{M}_{repr}$ of N_{repr} representative realizations. A common choice is to pick the closest realization of the cluster centroid as the representative of that cluster. Note that the derived clusters may have different sizes, as the clustering algorithms are not constrained to form groups containing the same number of realizations. Based on this observation, different weights proportional to the cluster sizes can be assigned to the representative realizations to reflect the number of realizations in their respective clusters. In this case, averages and other statistics of the reduced ensemble $\mathbf{M}_{repr}$ are computed with weights, unlike what is done for the full ensemble $\mathbf{M}_{full}$ where all the N realizations are usually assumed to be equiprobable and therefore equal weights. Finally, robust optimization is performed over $\mathbf{M}_{repr}$ resulting in an optimized strategy $\mathbf{u}_{repr}^{opt}$.

In order to assess the ability of reduced ensembles to reproduce the results obtained by performing the optimization over the full ensemble $\mathbf{M}_{full}$, we compare the performance of the derived production strategies over $\mathbf{M}_{full}$ (Figure 4.1). Note that this is done only for validation purposes; once we start using the approach to accelerate the optimization, we proceed only with the steps in Figure 4.1 (b) and rely on optimization over the reduced ensemble $\mathbf{M}_{repr}$.

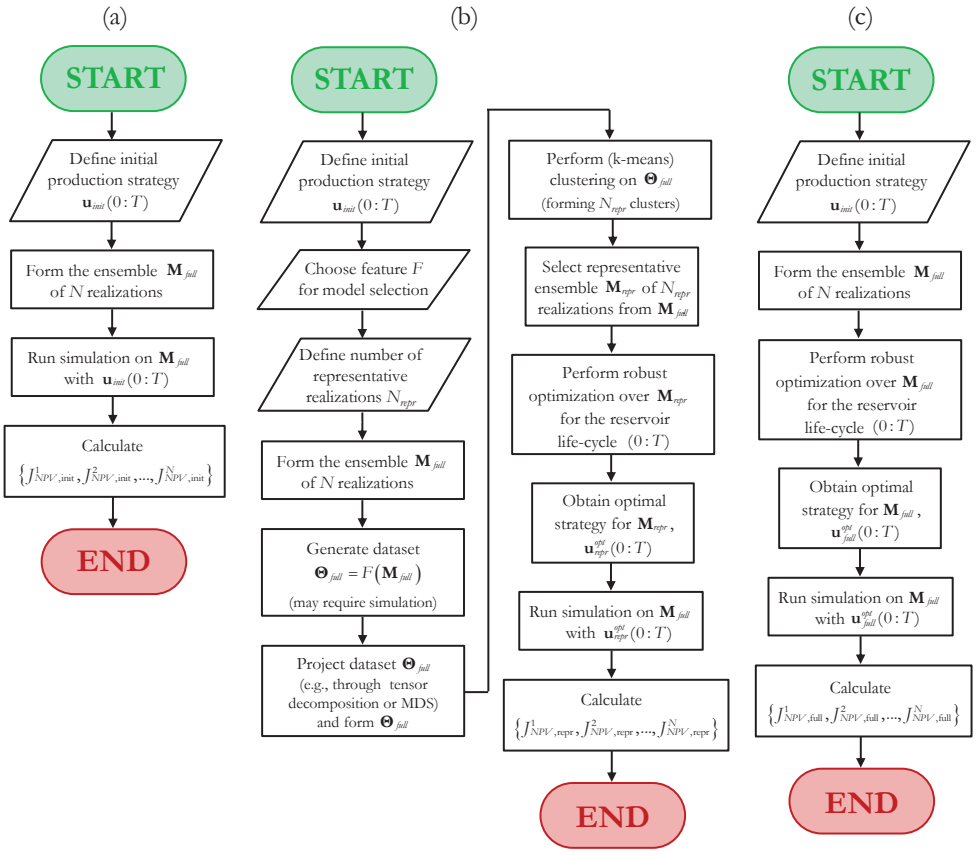


Figure 4.1: Workflow to evaluate the use of representative realizations for efficient robust optimization. (a) Computation of objective function values for unoptimized production, (b) robust optimization using representative realizations and (c) robust optimization using the full ensemble (reference).

4.3.2. Speeding-up history matching

History matching procedures (section 1.2.2) can also be accelerated by considering a reduced ensemble of realizations. Figure 4.2 illustrates how to compare the performance of a representative ensemble to the full ensemble based on the history-matched models. The first step is to create the representative ensemble using the techniques described in section 4.2.2. This step is the same as what is done to form the reduced ensemble for robust optimization. Next, history matching is performed on both full and representative ensembles, resulting in two different posterior ensembles. We then perform simulations on the realizations of the posterior ensembles to generate data for comparison and validation of the representative ensemble. Note that, when applying representative model

selection to make the history matching procedure more efficient, only the workflow in Figure 4.2 (right) is carried out.

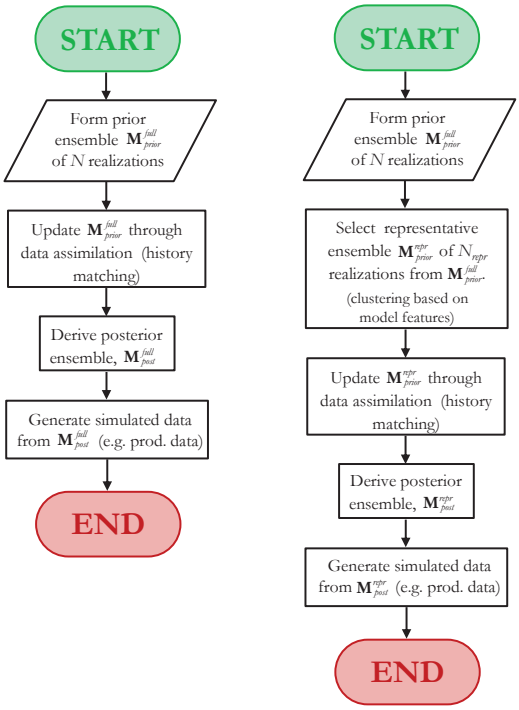


Figure 4.2: Workflow to evaluate the use of representative realizations for efficient history matching: procedure over the full ensemble (reference for validation) (left) and over the reduced ensemble (right).

Note also that, in this section, we discuss the selection of representative model realizations only as a principle to make history matching computationally more efficient. This is a mechanism that will later fit in the VOI workflow (section 4.3.4), where important choices not specified in Figure 4.2 will be determined according to the application considered (e.g., whether to assimilate data measured at a single time or during a time interval, which production strategy **u** is used to generate the data).

4.3.3. Representative plausible truths

The selection of fewer plausible truths for the VOI analysis can help reducing the computational cost of the workflow at a different level than the acceleration of robust optimization and history matching. The plausible truths are model realizations which we

pick to play the role of truth in the CLRM framework. Thus, the goal remains the same: to select representative model realizations.

The challenge is to find relevant features to distinguish these realizations considering their role in the workflow. Although we are still interested in the reservoir management problem (i.e., the water flooding process in our case), the plausible truths are not directly involved in the optimization procedure; we perform the optimizations on the realizations of the prior and posterior ensembles. Due to this difference in roles, the features which are relevant to select representative realizations for the robust optimization and history matching may be not the most appropriate to distinguish plausible truths. As we mentioned before, literature suggests there is no one-method-fits-all solution for choosing the selection features and, therefore, we look for fit-for-purpose solutions.

The methodology for VOI assessment presented in the previous chapters accounts for the decision making process, which in CLRM takes the form of optimized production strategies. When we perform the VOI assessment following the workflow presented in Chapter 2 (considering N_{truth} plausible truths), we obtain the solution schematically depicted in Figure 4.3. There is a different production strategy $\mathbf{u}^i_{post}$ corresponding to each plausible truth $\mathbf{m}^i_{truth}$. Also, each plausible truth has its own pair $\{J^i_{NPV,prior}, J^i_{NPV,post}\}$ that is directly related to the VOI calculation (equation (2.7)). These (possibly unique) associations unveil a mechanism to distinguish plausible truths according to the decisions made (or their consequences) in each scenario. This suggests that we should look for features that carry a similar structure, attributing a different production strategy (or model input) to each plausible truth. We refer to these features as decision-based features.

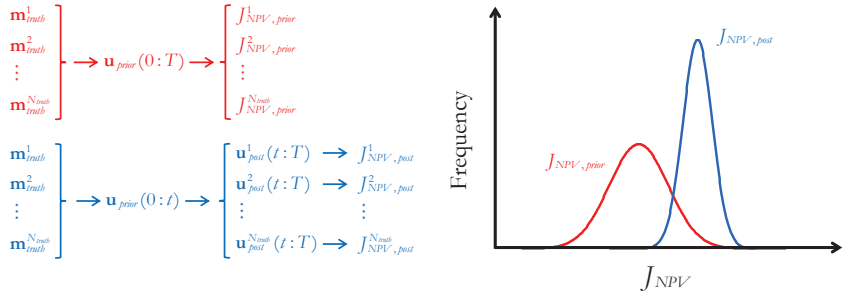


Figure 4.3: Typical solution obtained with the VOI workflow presented in Chapter 2: a different decision for each plausible truth (left) resulting in improved performance. The NPV plots show prior (red) and posterior (blue) distributions sampled by the plausible truths (right).

In sections 4.3.1 and 4.3.2, we discussed possible features to distinguish model realizations for robust optimization and history matching purposes. It is important to highlight that in both applications all the model realizations are submitted to the same production strategy. In this context, the features used to select representative realizations rely on the fact that we can compare them through their inherent characteristics (i.e., model parameters) or response to a given strategy (i.e., model states and model outputs). From here on, we refer to these features as model-based features. Figure 4.4 summarizes the main characteristics of both model-based and decision-based features. Model-based features do not account for the fact that the different scenarios imply different decisions, while the decision-based ones do, connecting them to the VOI assessment setting introduced in Chapter 2.

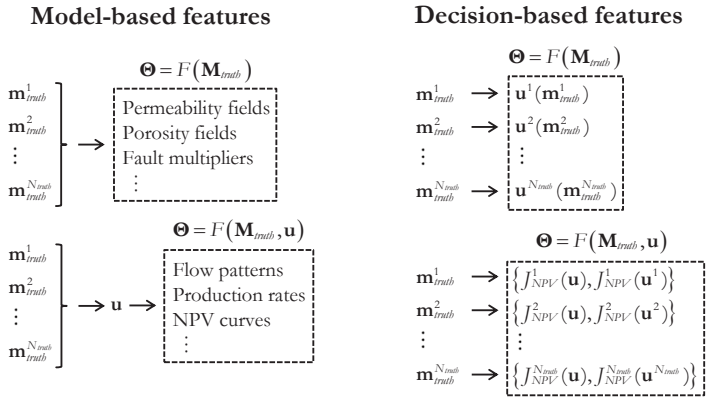


Figure 4.4: Schematic comparison between model-based and decision-based features. Model-based features rely exclusively on the characteristics of the models and their response to a fixed input, while decision-based features rely on the distinction of models through the decisions associated with each scenario.

The problem becomes then how to determine a set of decisions (or production strategies) $\mathbf{u}^i$, $i = 1,\ldots,N_{truth}$ that can uniquely identify each plausible truth $\mathbf{m}^i_{truth}$. Very importantly, in addition to that, we seek ways of obtaining these decisions $\mathbf{u}^i(\mathbf{m}^i_{truth})$ in a computationally attractive way so that we can accelerate the VOI workflow when selecting representative plausible truths based on them.

As mentioned before, in the CLRM context the decisions take the form of optimized production strategies. Thus, a good way of identifying characteristics of the optimal configurations related to each of the plausible truths is to perform separate optimizations on them. If we consider an initial ensemble of N_{truth} plausible truths, this means we need

to perform N_{truth} optimizations. (Note that the computational cost associated with these N_{truth} optimizations is significantly lower than the cost of the full VOI workflow, which makes this approach suitable for accelerating the VOI assessment.) As a result, we obtain a set of N_{truth} optimal production strategies $\mathbf{u}_{opt}^{i}$ and N_{truth} optimal objective function values $J_{NPV}^{i}(\mathbf{u}_{opt}^{i})$.

Typically, the optimal production strategies $\mathbf{u}_{opt}^{i}$ tend to be very different from each other because the plausible truths are different model realizations. Following this reasoning, the optimal production strategies seem to be an appropriate feature to support the clustering of plausible truths. The main potential problem of this approach is the risk of non-uniqueness of optimal solutions for the production optimization problem, due to the possible presence of redundant degrees of freedom in the high-dimensional space of control variables (Van Essen et al., 2009). This may result in multiple production strategies that are equally optimal for a given plausible truth, which could put the validity of this approach at stake. A possible way to avoid the redundancy is to perform an a-priori tensor- or SVD-based decomposition of a large set of possible controls and perform the optimization in a reduced control space. Alternatively, one could impose temporal and/or spatial correlations on the controls which also reduces the degrees of freedom in the control space. A downside of such a-priori measures is that they may lead to lower NPV values. As an alternative, we therefore apply an a-posteriori tensor decomposition of the set of optimal production strategies and retain a fraction of the basis functions by truncation based on their energy. By doing so, we intend to capture only the main trends of the data and reduce the effects of the non-uniqueness of the optimal production strategies, although we note that this a-posteriori decomposition of the controls does not guarantee an improvement of the situation. As a final remark on this issue, we note that for these N_{truth} nominal optimizations would in principle be performed with the same optimization methods that will be used for the VOI and CLRM exercises and that the risk of non-uniqueness of optimal production strategies will also be present there. Ultimately, the selection of representative plausible truths based on these strategies would still be consistent with the VOI workflow, even in its limitations.

On the other hand, the optimal objective function values $J_{NPV}^{i}(\mathbf{u}_{opt}^{i})$ typically tend to be close to each other because the different optimal production strategies compensate for the

differences in the model realizations (i.e., given the same well locations, the optimal sweep of the reservoir tends to be similar for all the realizations); see, e.g., the final NPV values in Figure 3.8 or more schematically the distributions in Figure 4.3. Thus, the dissimilarities between the realizations in terms of NPV are less pronounced for their optimal configurations, and, because of that, they are less suitable to help in the selection of representative plausible truths. However, in combination with the objective function values $J^i_{NPV}(\mathbf{u}_{prior})$ obtained with a robust strategy $\mathbf{u}_{prior}$ (i.e., optimized to maximize the mean objective function given the initial uncertainty), these data reveal how much we may benefit if we learn or observe the truth for each one of the plausible truths. We can then distinguish plausible truths according to the gains associated with their optimal configurations, which are directly related to the VOI, and this can be useful for our purposes. One of the advantages of using these $\left\{J^i_{NPV}(\mathbf{u}_{prior}), J^i_{NPV}(\mathbf{u}^i_{opt})\right\}$ data features is avoiding the problem of non-unique optimal solutions discussed in the previous paragraph.

We introduce the selection of representative plausible truths to the original VOI assessment workflow (Chapter 2) and we obtain the procedure depicted in Figure 4.5 for cases with a single observation time. The main difference compared to the original workflow is that, before entering the loop where each one of the N_{truth} realizations of the initial ensemble $\mathbf{M}_{truth}$ is picked to be the truth $\mathbf{m}^i_{truth}$, we have a few more pre-processing steps. First, a step where we optimize each one of the realizations and then a step where we perform the clustering to select N^{repr}_{truth} representative plausible truths based on the decision-based features as explained above. Another minor change in the workflow refers to the computation of the statistics of VOI: before, the plausible truths were considered (for simplicity) to be equiprobable, but, now, the selected plausible truths in $\mathbf{M}^{repr}_{truth}$ may have different weights w_i assigned by our selection procedure (i.e., weights proportional to the number of realizations in each cluster). Note that there is a computational cost associated with the additional N_{truth} nominal optimizations required, but that this extra cost is minor when compared to the cost of the full workflow. Another point to realize is that these N_{truth} nominal optimizations would be performed anyway if we carry out a value of clairvoyance analysis (VOC; see Chapter 2) to determine the upper bound for the VOI assessment.

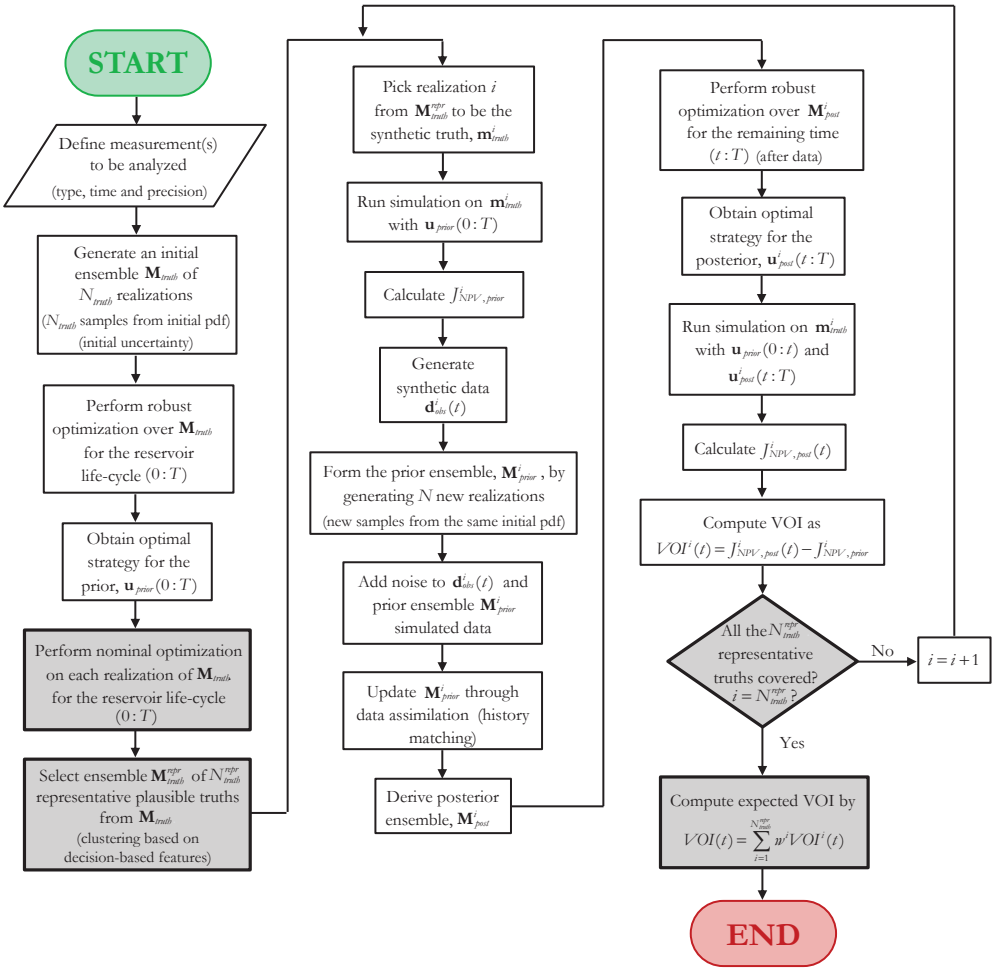


Figure 4.5: Workflow to compute the expected VOI for a single observation time using representative plausible truths and full ensembles for robust optimization and history matching. The main modifications in the workflow with respect to the original procedure (Chapter 2) are highlighted in grey bold boxes.

4.3.4. Accelerated VOI assessment

We combine the ideas of the three previous sections to the original VOI workflow of Figure 2.11. The result is a new workflow for accelerated VOI assessment, shown in Figure 4.6.

We note that, like in Chapter 2, in this flowchart we consider the sizes of the ensemble of plausible truths and the prior ensembles to be N_{truth} and N. Note that these ensembles may be chosen to have the same size, but that this is not a necessity. The same holds for the

number of representative realizations for robust optimization (and history matching) and the number of representative plausible truths, which here are chosen as N_{repr} and N_{truth}^{repr} respectively.

Given these considerations, we can expect a speed-up factor of the order of $O((N_{truth} / N_{truth}^{repr}) \cdot (N / N_{repr}))$ by using this accelerated procedure. This means that, if we are able to select reduced ensembles with 10 times fewer realizations, we can reduce the number of required reservoir simulations by a factor of 100.

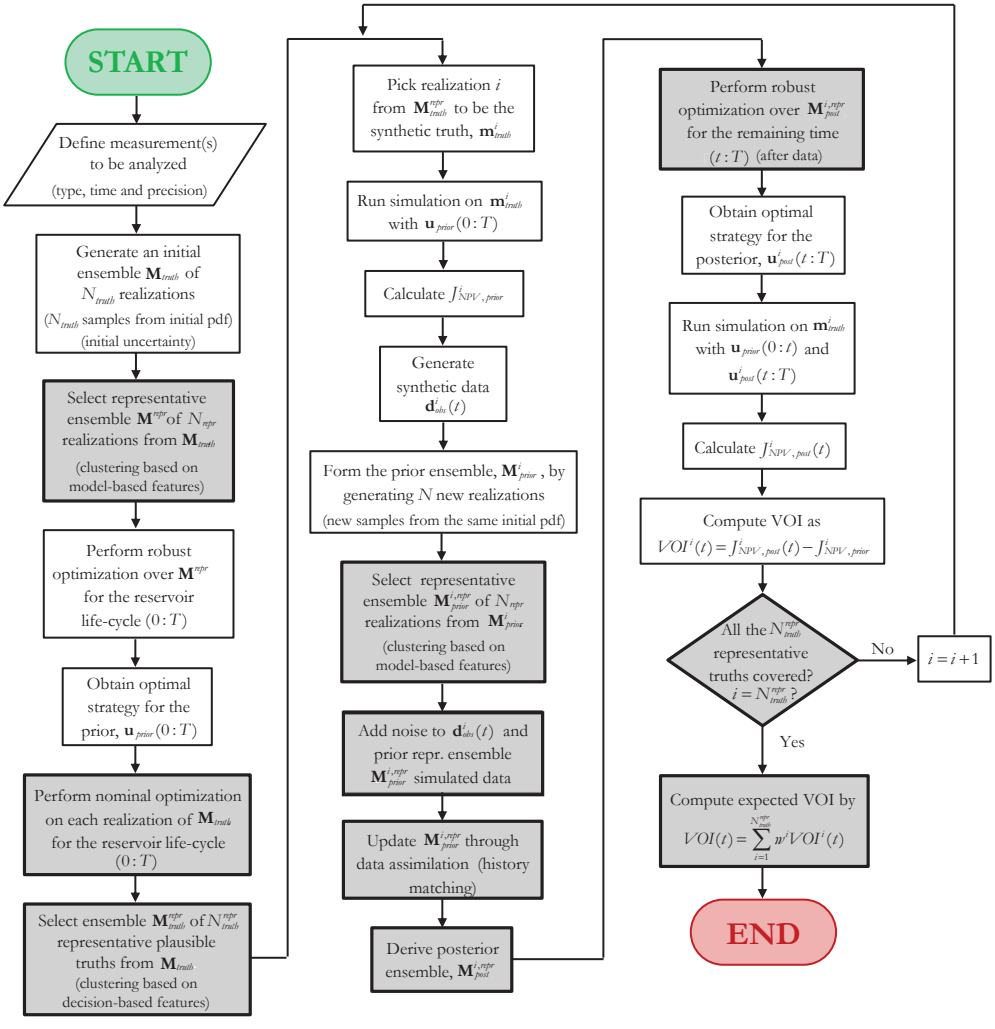


Figure 4.6: Workflow to compute the expected VOI for a single observation time using representative plausible truths and reduced ensembles of representative model realizations for robust optimization and

history matching. The main modifications in the workflow with respect to the original procedure (Chapter 2) are highlighted in grey bold boxes.

4.4. Examples

4.4.1. Robust optimization with reduced ensembles

4.4.1.1. 2D five-spot model

As a first step to illustrate our approach, we used the same small 2D five-spot example (441 grid blocks) from the previous chapters; see section 2.4.2 for the complete description of the example. Originally, we had ensembles of $N = 50$ model realizations to characterize the geological uncertainties. To accelerate the robust optimizations, we considered a reduced number of representative realizations, $N_{repr} = \{3, 5, 10\}$, representing approximately 5 %, 10 % and 20 % of the number of realizations in the full ensembles. We evaluated the performance of different features for clustering: permeability field, oil saturation snapshots at every control time interval (every 150 days) and NPV time-series. We also studied the effect of different projection methods by using both MDS and tensor decomposition to perform the projection of the feature space before clustering. For MDS we used the standard implementation available in Matlab with the previously mentioned stress criterion (here 5 %) to determine the dimension of the reduced space. For the tensor decomposition we used the implementation of the HOSVD described by Insuasty et al. (2017) with a cut-off of 95 % in terms of the energy of the eigenvalues to determine the number of basis functions to be retained for the uncertainty dimension. The production strategy used to generate the features for clustering was fixed as the initial production strategy chosen as the starting point of the optimization (here mid in-between bounds: $p_{prod} = 250$ bar for the producers and $p_{inj} = 400$ bar for the injector).

By applying the steps as in Figure 4.1 to compare the performance of the robust optimization with reduced and full ensembles, we obtained results in terms of NPV. Figure 4.7 depicts the results for this 2D five-spot example expressed as probability distribution function (pdf) plots. The NPV values for the unoptimized production (Figure 4.1 (a)) are shown in grey and the reference results (Figure 4.1 (c)) in black. The results obtained using the reduced ensembles (Figure 4.1 (b)) are represented by the colored lines

according to the number of representative realizations. We repeated the procedure with several ensembles $\mathbf{M}_{full}$ and we obtained results similar to the ones shown here.

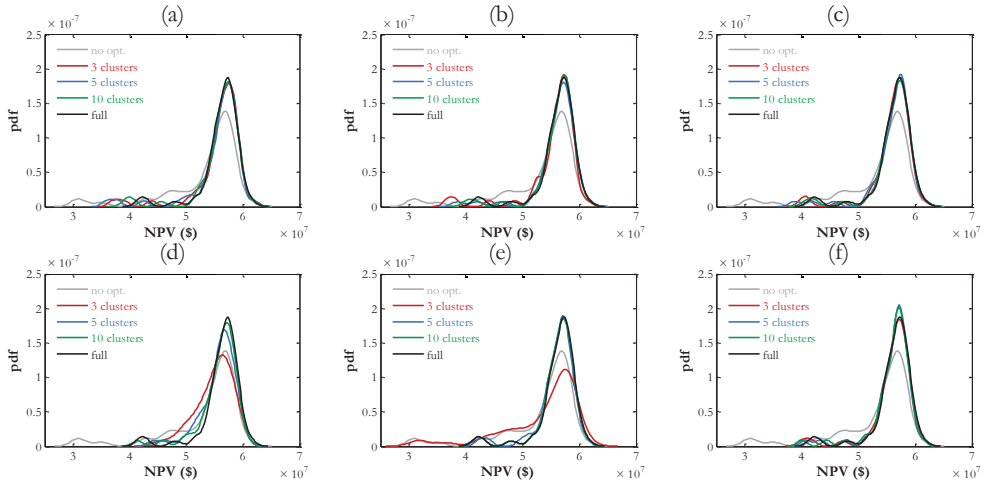


Figure 4.7: Results of robust optimization over an ensemble of the 2D five-spot model expressed in terms of NPV pdf plots. (a) Permeability as feature and MDS as projection method, (b) oil saturation snapshots as feature and MDS as projection method, (c) NPV time-series as feature and MDS as projection method, (d) permeability as feature and tensor decomposition as projection method, (e) oil saturation snapshots as feature and tensor decomposition as projection method, and (f) NPV time-series as feature and tensor decomposition as projection method.

Generally, all the production strategies $\mathbf{u}_{repr}^{opt}$ optimized with the representative ensemble performed very well when compared to the strategy $\mathbf{u}_{full}^{opt}$ for the full ensemble. We notice that the selection with three representative realizations (5 %) from the ensemble performs poorer in most cases, and fails badly in the case depicted in Figure 4.7 (e). This shows that, for this example, taking an amount of representative realizations corresponding to only 5 % of the original ensemble size is insufficient whereas taking 10 % and 20 % give good results in all cases. Also, we notice that, for this example, the MDS transformation helped the selection of representative realizations better than the tensor decomposition, especially when using the permeability field as feature (Figure 4.7 (a) and (d)) but also when using the oil saturation snapshots (Figure 4.7 (b) and (e)). We believe the better performance of MDS here to be related to the small size of the model (441 grid blocks) and the (relatively) smooth character of the geological realizations (Figure 2.6). For larger models and cases with more complex geological features, in which there is more spatial-temporal structure to be preserved, it is expected that tensor decomposition should perform the best

(Insuasty et al., 2017). Overall, based on the results from the 2D five-spot model, oil saturation snapshots and NPV time-series seem to be the most suitable selection features, which is in alignment with the conclusions drawn by Shirangi and Durlofsky (2016) after comparing several model-based features within their general framework.

4.4.1.2. Egg model

The Egg model is a synthetic reservoir model created to serve as a benchmark for water flooding optimization, computer-assisted history matching and CLRM applications. The model consists of 100 realizations of a channelized reservoir with 60 × 60 × 7 grid cells. Its 18,553 active cells give it the shape of an egg (Figure 4.8). The field is produced through water flooding, with four producers and eight injectors in defined locations, and has been used for several studies; see, e.g., Van Essen et al. (2009). For further information on a standardized version of this model, we refer to Jansen et al. (2014).

For this case study, we considered the optimization of the water injection rates for a production period of 3,600 days. The rates can be adjusted every 360 days (i.e., $M = 10$ control time intervals) in the range of $0 \leq q_{inj} \leq 79.5$ m³/day, and the maximum injection pressure allowed is $p_{inj,max} = 420$ bar. The bottom-hole pressure of the producers is kept constant at $p_{prod} = 395$ bar. The robust optimization experiments were carried out with the help of the AD-GPRS simulator to obtain the required predictions and gradients (Zhou, 2012, and Bukshtynov et al., 2015) for each model realization which were used in a simple implementation of the steepest ascent algorithm.

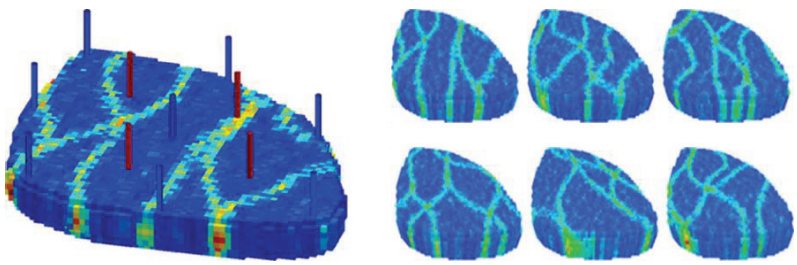


Figure 4.8: The Egg model: a channelized reservoir with 8 injectors (blue) and 4 producers (red) (left). Six randomly chosen realizations of the permeability distribution (right). (after Van Essen et al., 2009, and Jansen et al., 2014)

In this example, to illustrate the use of representative ensembles, we considered a reduced number of representative realizations, $N_{repr} = \{5, 10, 20\}$, representing again 5 %, 10 %

and 20 % of the number of realizations in the full ensemble. We applied clustering based on permeability distributions and oil saturation snapshots, with MDS and tensor decomposition as projection techniques using the same settings as described for the 2D five-spot model (section 4.4.1.1) to determine the dimension of the reduced space and the number of basis functions to be retained.

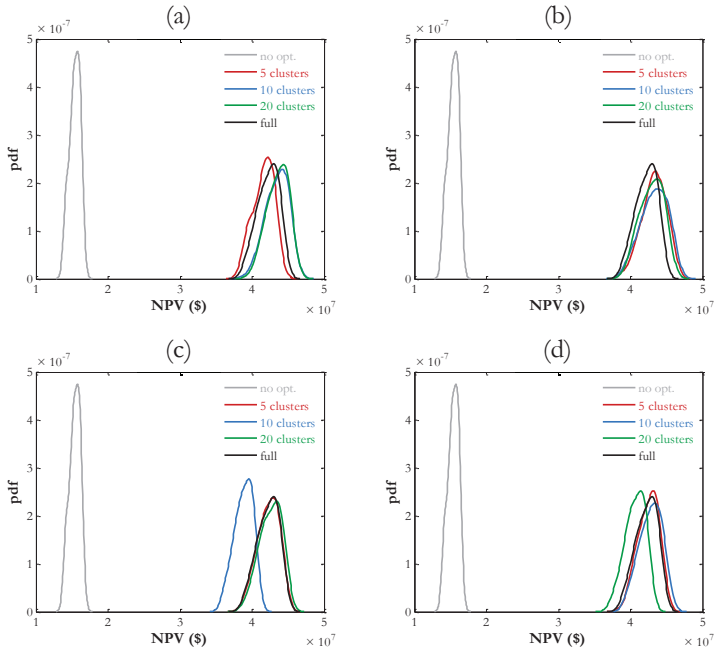


Figure 4.9: NPV pdf plot of an ensemble from the Egg model. (a) Permeability as feature and MDS as projection method, (b) permeability as feature and tensor decomposition as projection method, (c) oil saturation snapshots as feature and MDS as projection method, and (d) oil saturation snapshots as feature and tensor decomposition as projection method.

Figure 4.9 depicts the results obtained. The color scheme is the same as in Figure 4.7. First, we observe that, in this case, robust optimization results in significantly higher NPV values compared to those obtained with the unoptimized strategy. We can also see that the optimizations over the reduced ensembles were able, in most of the cases, to achieve a similar performance compared to the optimization performed over the full ensemble. In some cases, the optimization with the representative realizations managed to outperform the reference results, which reminds us that the optimization techniques used here are not perfect and cannot guarantee globally optimal solutions for the production optimization

problem. Moreover, we observe some inconsistent results: in Figure 4.9 (c), the robust optimization based on $N_{repr} = 10$ selected models performed significantly worse than the one using $N_{repr} = \{5, 20\}$ realizations, and in Figure 4.9 (d) the optimization based on $N_{repr} = 20$ realizations was the one that performed the worst. We expected to see an increase in the performance of the robust optimization as N_{repr} gets closer to N. Although these unexpected results could be related to limitations of our implementation (e.g., imperfections within the optimization algorithms and simulations), we have not yet been able to find a conclusive explanation for it.

4.4.2. History matching with reduced ensembles

4.4.2.1. 2D five-spot model

To illustrate the workflow from Figure 4.2, we first tested it in a history matching twin experiment on the 2D five-spot example. For this exercise, we used AD-GPRS to obtain the required gradients to update the uncertain parameters. As the version of AD-GPRS made available to us only provided the gradients with respect to permeability multipliers, this example was slightly modified with respect to the original model used in section 4.4.1.1 and previous chapters: here we assumed the porosity field to be homogeneous and known ($\phi = 0.2$), and thus not to be updated throughout the history matching. We considered the availability of measurements of water and oil field production rates at a single observation time, and measurement errors of $\varepsilon_{prod} = 5$ m³/day. We repeated the history matching for different observation times. The production strategy was fixed at $p_{prod} = 250$ bar for the producers and $p_{inj} = 400$ bar for the injector.

First, we performed the history matching over the full ensemble of $N = 49$ realizations, with one additional realization used to generate the synthetic measurements. Next, we considered only the representative realizations. Like in section 4.4.1.1, we repeated the procedure for several choices, by varying the number of representative models, the selection feature and the projection method. Figure 4.10 displays the results obtained with $N_{repr} = 10$ representative models selected based on oil saturation snapshots projected by tensor decomposition. The results for other clustering settings, which are reported in Yap (2016), were comparable to the ones showed here.

Figure 4.10 (left) shows the simulated forecasts of field production rates for the prior ensemble. The thin dashed lines represent each one of the $N = 49$ realizations of the full prior ensemble while the thicker dashed lines correspond to the $N_{repr} = 10$ representative ones. The thick solid lines show the forecast generated with the synthetic truth, and the yellow circles indicate the measurements available. We observe that, although the synthetic truth seems to be captured by the ensemble, there is a large spread in the predictions. Figure 4.10 (right) shows the predictions a posteriori, after the history matching was performed over the full and reduced prior ensembles. We note that the assimilation of the available measurements contributed to a significant reduction in the spread of the curves, but, more importantly, we observe a reasonably good agreement between the uncertainty characterized by the full and reduced posterior ensembles. We observe a slight reduction in the spread of the curves, which was expected because smaller ensembles tend to underestimate uncertainty in the forecasts. There are measures such as ensemble inflation that could possibly minimize this undesirable effect, but they were not considered in this work. We also emphasize that different weights have been assigned to the representative realizations forming the reduced ensemble while the realizations of the full ensemble are equiprobable, and that, for this reason, the visual comparison may not be the most appropriate way of assessing the quality of the approximation here. Despite these remarks, the results suggest that representative realizations can be used to make history matching procedures more efficient without compromising the uncertainty quantification to an unacceptable level.

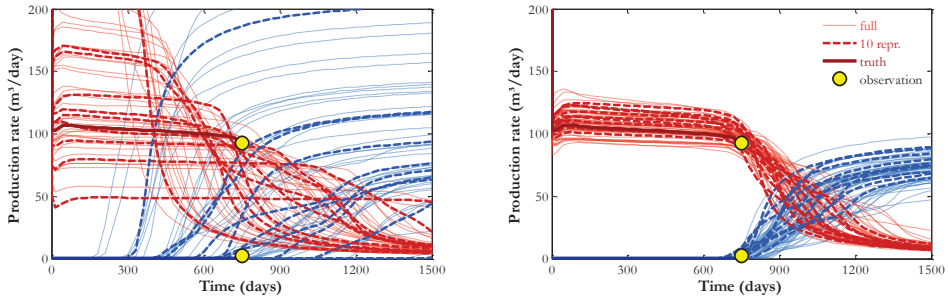


Figure 4.10: History matching results for the 2D five-spot model. Solid lines represent the prediction from the synthetic truth, the dots correspond to the synthetic data to be matched and the dashed lines represent the predictions of the realizations to be updated. Red lines correspond to oil production and blue lines to water production.

4.4.2.2. Egg model

Based on the learning from the 2D five-spot example, we repeated the same procedure for the Egg model to confirm that it is possible to make the history matching more efficient also in larger case studies. Once again, we used AD-GPRS to obtain the required gradients to update the uncertain permeability field. The settings are the same as the ones described in section 4.4.1.2, but here the injection rates were fixed to $q_{inj} = 79.5$ m³/day. The observations considered were field production rates and bottom-hole pressures measured in the injectors, available at $t_{data} = 1{,}800$ days, with measurement errors of $\varepsilon_{prod} = 5$ m³/day and $\varepsilon_{BHP} = 10$ bar. Here we used the full ensemble with $N = 99$ realizations, plus one synthetic truth.

Figure 4.11 displays the results obtained with $N_{repr} = 10$ representative models selected based on oil saturation snapshots projected by tensor decomposition. Figure 4.11 (left) shows the simulated forecasts of field production rates for the prior ensembles, and Figure 4.11 (right) for the posterior ensembles. The color scheme is the same as in Figure 4.10. The differences between the prior and posterior predictions of production rates are minor, indicating that the main contribution to the mismatches is probably related to the bottom-hole pressure measurements in the injectors, displayed in Figure 4.12. The black dashed lines represent the realizations of the full prior ensemble while the blue lines correspond to the representative ones. The red lines show the forecast generated with the synthetic truth, and the yellow circles indicate the measurements available. Like in the 2D five-spot example, we observe that the reduced ensemble constitutes a good approximation of the full ensemble. As before, we repeated the exercise considering other choices for the selection of representative realizations, but we displayed only one set of the results here. The results for other clustering settings can be found in Yap (2016). Those results showed that $N_{repr} = 10$ seems to be close to the limit as the minimum number of representative realizations to approximate the uncertainty characterized by the full ensemble. We also noticed that MDS and tensor decomposition lead to similar performance, but that the latter technique resulted in reduced ensembles which provided slightly larger spreads in the production forecasts.

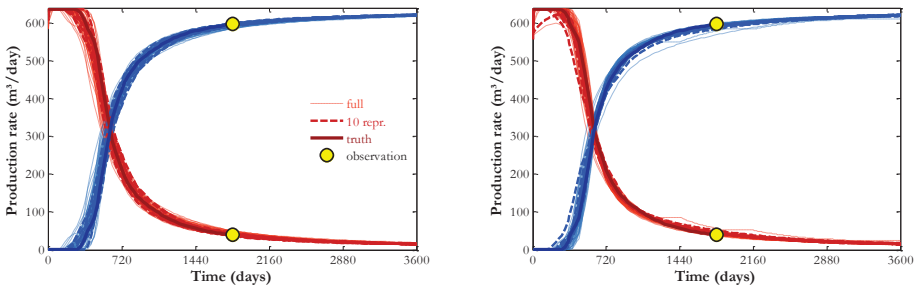


Figure 4.11: History matching results for the Egg model in terms of the predictions of field production rates: prior (left) and posterior (right). Red lines correspond to oil production and blue lines to water production. Solid lines represent the prediction from the synthetic truth, the yellow dots correspond to the synthetic data to be matched and the dashed lines represent the predictions of the realizations to be updated (full and reduced ensembles; see legend).

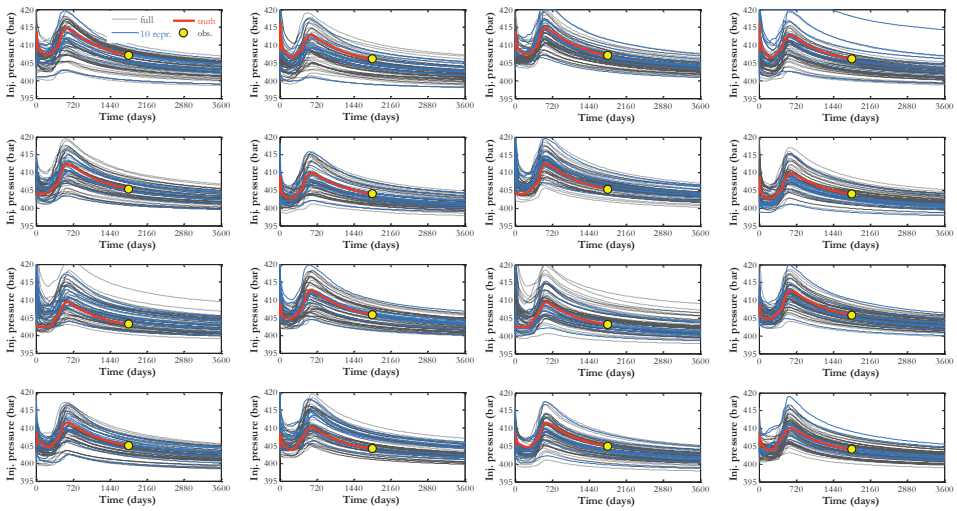


Figure 4.12: History matching results for the Egg model in terms of the predictions of BHP at the 8 injectors: prior (left) and posterior (right). Solid red lines represent the prediction from the synthetic truth, the yellow dots correspond to the synthetic data to be matched and the dashed lines represent the predictions of the realizations to be updated (full ensemble in black and reduced ensemble in blue; see legend).

4.4.3. Accelerating VOI assessment

4.4.3.1. 2D five-spot model

To illustrate the selection of representative plausible truths, we applied to the 2D five-spot model the workflow to assess the VOI for a single observation time with a reduced number of plausible truths (Figure 4.5). The workflow was repeated for different

observation times, t_{data} = {150, 300, … , 1,350} days. The history matching step was performed with the EnKF module of MRST (Lie et al., 2012). We assessed the VOI of the production data (total flow rates and water-cuts) with absolute measurement errors (ε_{flux} = 5 m³/day and ε_{wct} = 0.1). The VOC and the VOI were computed for each of the nine observation times, and we compared the results against those obtained using the original workflow (Figure 2.11) with the full ensemble $\mathbf{M}_{truth}$ of plausible truths (N_{truth} = 50), which serves as a reference.

First, we checked our hypothesis that the model-based features are not suitable for selecting plausible truths. For that, we selected N_{truth}^{repr} = 5 plausible truths through clustering based on: permeability fields, NPV evolution profiles and flow patterns (i.e., pressure and saturation snapshots). No projection methods (e.g., MDS) were used. We also tested a random selection. The results are shown in Figure 4.13, including the reference results. The different lines represent percentiles and mean of the VOC and VOI distributions as a function of the time when the additional information (or clairvoyance) becomes available. We do not interpret or explain the VOI and VOC results here; we refer to Chapter 2 for this purpose. Here we assess the ability to obtain similar results with fewer plausible truths.

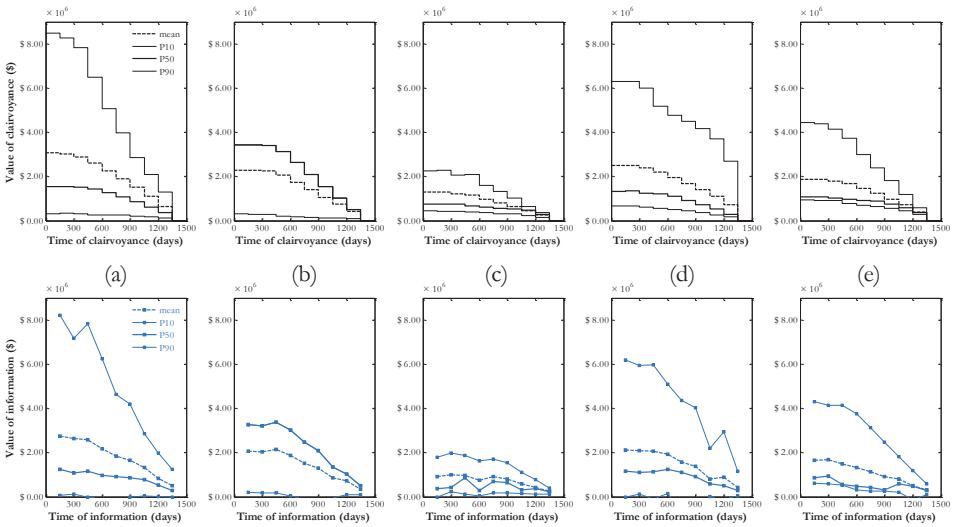


Figure 4.13: Results of the VOI (and VOC) assessment for the 2D five-spot model using a selection of 5 plausible truths based on model-based features. (a) Reference obtained using the original workflow, (b) random selection, (c) selection based on permeability field, (d) selection based on NPV time series and (e) selection based on flow patterns.

Overall, none of the selections in Figure 4.13 is able to satisfactorily reproduce the reference results, shown in Figure 4.13 (a). Although a selection based on some feature (Figure 4.13 (c), (d) and (e)) is clearly better than a random selection (Figure 4.13 (b)), these results seem to support our hypothesis that the model-based features are not the most appropriate means to select representative plausible truths.

Next, we repeated the same procedure with our proposed decision-based features. This required a nominal optimization on each of the $N_{truth} = 50$ plausible truths initially considered. Figure 4.14 shows the data we obtain from these optimizations, which can be used for clustering and model selection. It becomes clear that these features create a space in which we can distinguish the samples and select those instances of $\mathbf{M}_{truth}^{repr}$ that can better represent the entire population $\mathbf{M}_{truth}$. Figure 4.14 (left) displays each of the plausible truths plotted in a two-dimensional space: the first dimension corresponds to the objective function values (i.e., NPV) before the nominal optimizations and the second to the objective function values after the optimizations. The clusters are shown in different colors and the selected plausible truths are marked as squares. Figure 4.14 (right) exhibits the optimal production strategy for each plausible truth: the plots show the BHP controls at each of the five wells of the 2D five-spot model.

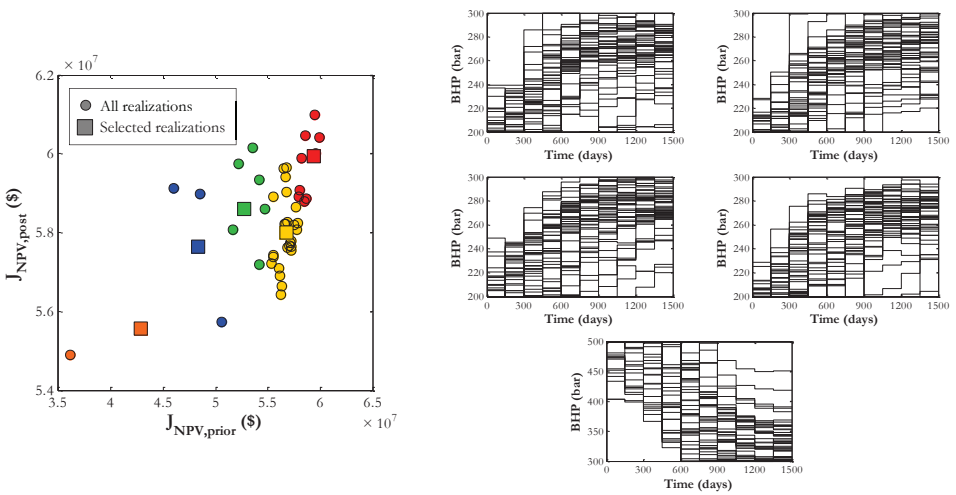


Figure 4.14: Proposed decision-based features for the selection of representative plausible truths. Objective function values, before and after nominal optimizations (left). Optimal production strategies (bottom-hole pressures at all the wells for every control interval) (right).

Figure 4.15 depicts the results obtained by picking $N_{truth}^{repr} = 5$ representative plausible truths according to these new features. This time we observe better selections, which are able to repeat the reference results (Figure 4.15 (a)) almost perfectly. We note that the selection based on the objective function values (Figure 4.15 (b)) succeeds in reproducing the reference results for VOC, but less for VOI. In contrast, the selection based on optimal production strategies (Figure 4.15 (c)) performs fairly well for both VOC and VOI. Therefore, the optimal production strategy is the feature we chose for selecting representative plausible truths.

After that, we investigated the impact of the non-uniqueness of optimal production strategies in the model selection. For that, we carried out nominal optimizations on each of the plausible truths starting from three different initial solutions. Figure 4.16 shows the results: Figure 4.16 (b), (c) and (d) correspond to the selection obtained with the three different starting points and Figure 4.16 (e) to the selection based on the data of the three optimizations altogether. We observe that the results are not identical, which confirms that the optimal production strategies may be nonunique and, thus, not suitable for model selection purposes.

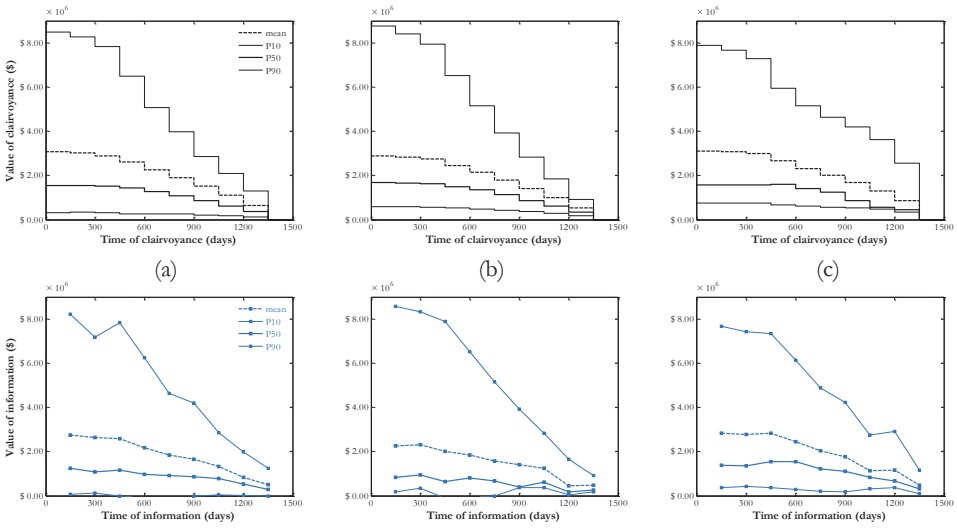


Figure 4.15: Results of the VOI (and VOC) assessment for the 2D five-spot model using selection of 5 plausible truths based on decision-based features. (a) Reference obtained using the original workflow, (b) selection based on objective function values and (c) selection based on optimal production strategies.

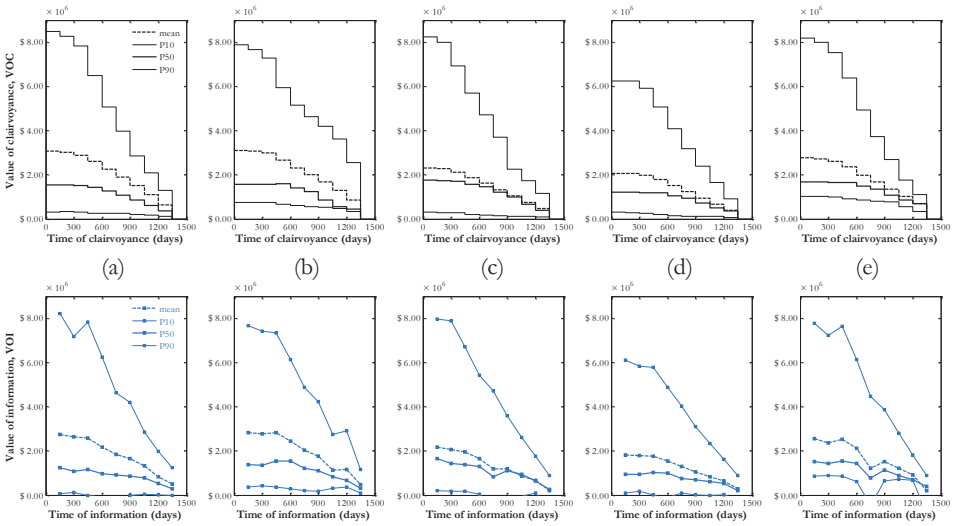


Figure 4.16: Results of the VOI (and VOC) assessment for the 2D five-spot model selecting representative plausible truths based on optimal production strategies obtained with different starting solutions. (a) Reference results, (b) starting from robust optimal solution, (c) starting from greedy controls (maximum injection and maximum drawdown in the producers), (d) starting from mid in-between bounds and (e) selection based on data from all the three optimizations.

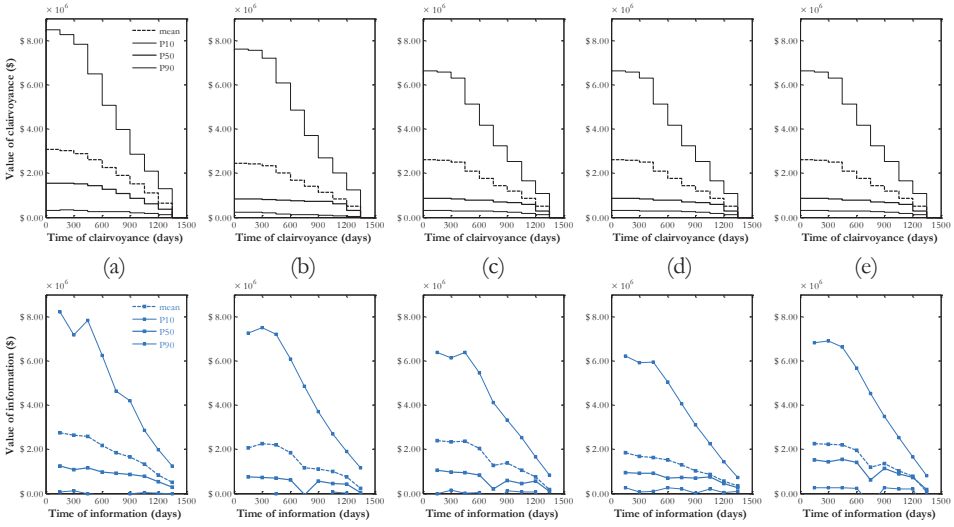


Figure 4.17: Results of the VOI (and VOC) assessment for the 2D five-spot model selecting representative plausible truths based on optimal production strategies projected by tensor decomposition. (a) Reference results, (b) starting from robust optimal solution, (c) starting from greedy controls (maximum injection and maximum drawdown in the producers), (d) starting from mid in-between bounds and (e) selection based on data from all the three optimizations.

As a measure to remediate this problem, we applied a projection based on a truncated tensor decomposition to the optimal production strategies dataset. By retaining the fraction of the basis functions that preserves 90 % of the energy of the singular values, we hope to capture only the main trends and reduce the effect of non-uniqueness of the optimal solutions. Figure 4.17 shows the results obtained with such a projection for the same optimal production strategies. The difference between the results for the three different optimization starting points is smaller than in Figure 4.17.

Finally, we applied all the measures we discussed so far (i.e., selection of representative ensembles for robust optimization, history matching and plausible truths) following the accelerated procedure for VOI assessment depicted in Figure 4.6. Note that here the history matching was performed on the full ensembles because our implementation with MRST used the EnKF method, which is not reliable for very small ensembles. After the history matching steps, we selected representative realizations to accelerate the robust optimization over the posterior ensembles. Figure 4.18 presents the results. Again, Figure 4.18 (a) exhibits the reference results with $N_{truth} = 50$ plausible truths and robust optimization over ensembles of $N = 49$ realizations. Figure 4.18 (b) corresponds to the results obtained by using $N_{truth}^{repr} = 5$ representative plausible truths and full ensembles with $N = 49$ for the robust optimizations. Figure 4.18 (c) shows the results when considering all the $N_{truth} = 50$ plausible truths and reduced ensembles for robust optimizations. And Figure 4.18 (d) displays the results obtained with $N_{truth}^{repr} = 5$ plausible truths and ensembles of $N_{repr} = 5$ realizations for the optimizations.

We see that the acceleration measures allow us to obtain similar results by considering only 10 % of the original number of realizations. We also observe that even the combination of the two acceleration measures described is still able to correctly approximate the main trend of the reference results. Note that the lines plotted in Figure 4.18 (b) and (d) represent percentiles and mean values of VOI based on $N_{truth}^{repr} = 5$ samples, while in Figure 4.18 (a) and (c) these values are computed with $N_{truth} = 50$ samples, and this should be taken into account when interpreting the quality of the results.

In terms of computational cost, the results in Figure 4.18 (a) require approximately 1.5 million simulations. This number includes forward and backward simulations. The results shown in Figure 4.18 (b) and (c) require 150,000 and 170,000 simulations, respectively.

And the results in Figure 4.18 (d) need 17,000 simulations to be computed. Thus, by applying all measures to reduce the number of model realizations considered in the assessment, we were able to alleviate the computational cost of the workflow by a factor of 88, which is quite significant.

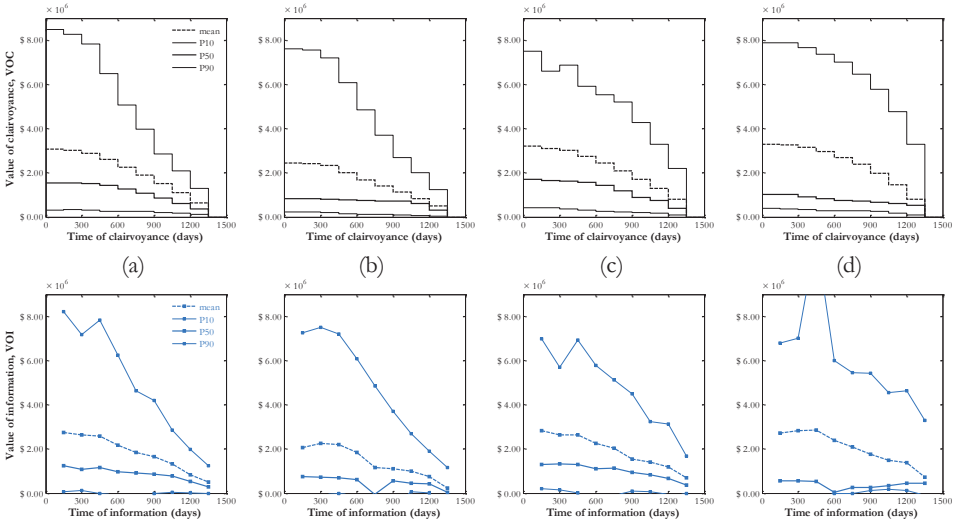


Figure 4.18: Results of the new accelerated VOI (and VOC) assessment for the 2D five-spot model. (a) Reference results, (b) only selecting representative plausible truths, (c) only reducing the ensembles for robust optimization, and (d) combining both measures.

4.4.3.2. Egg model

As a final test, we applied all the model selection measures discussed in this chapter (Figure 4.6) to make the VOI assessment possible for the Egg model. We considered $N_{truth}^{repr} = 10$ representative plausible truths and ensembles of $N_{repr} = 10$ representative realizations for the history matching and robust optimization. The workflow to assess the VOI of a single observation time was repeated for different observation times, t_{data} = {360, 720, … , 3,240} days. The entire exercise was performed with AD-GPRS to evaluate the objective functions and obtain the required gradients for the production optimization and history matching steps. We assessed the VOI of the production data (field production rates and pressures in the injectors) with absolute measurement errors ($\varepsilon_{prod} = 5$ m³/day and $\varepsilon_{BHP} = 10$ bar). The VOC and the VOI were computed for each of the nine observation times.

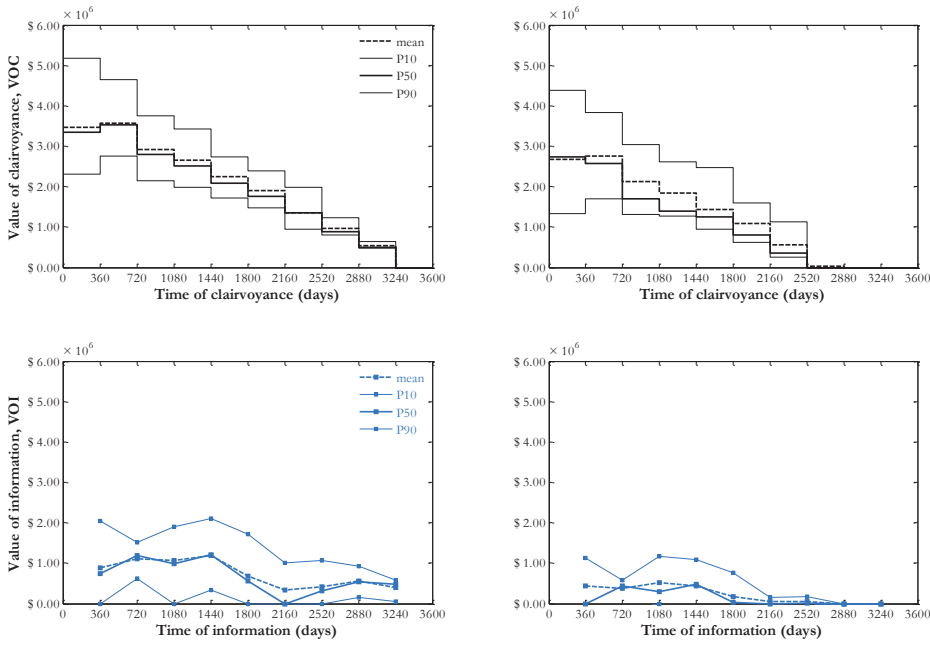


Figure 4.19: Results for the Egg model with a single observation time: VOC and VOI including (left) and disregarding (right) the uneconomical production of the plausible truths.

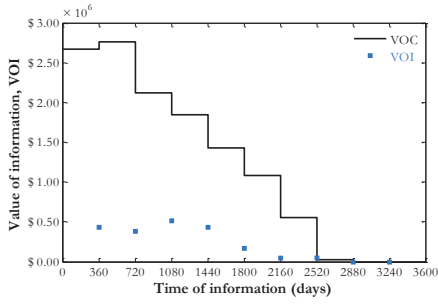


Figure 4.20: Results for the Egg model with a single observation time disregarding the uneconomical production of plausible truths: expected VOC and VOI.

Figure 4.19 and Figure 4.20 display the results obtained. We note that the type and trend of these results share similarities with those observed for the 2D five-spot model. As before, the VOC decreases in a stepwise fashion and serves as an upper bound for the VOI. On the other hand, this time the VOI values are significant lower than the VOC values, most likely because the field production rate measurements are not as informative as well-per-well measurements and also because of the larger number of uncertain

parameters. Figure 4.19 (right) shows the results obtained with the considerations discussed in sections 2.5.3 and 3.5.2 (i.e., disregarding uneconomical production from the plausible truths).

We also repeated the VOI assessment for the case with multiple observation times (Chapter 3). Figure 4.21 and Figure 4.22 display the results obtained, which, once again, confirm the conclusions drawn in the previous chapters. However, unlike the results for the 2D five-spot model in section 3.4.1, here the increase in VOI after the first observation time is more important. Figure 4.21 (right) shows the results obtained when disregarding uneconomical production from the plausible truths, and we note that this causes the VOI and VOC to decrease. The explanation for that is the same as what was discussed in section 3.5.2 and can be understood by comparing Figure 4.24 and Figure 4.25: when disregarding the uneconomical production, the baseline values for the VOI assessment (i.e., the final NPV achieved with the strategies optimized under prior uncertainty) increase, while this consideration does not have a major impact on the values obtained with additional knowledge.

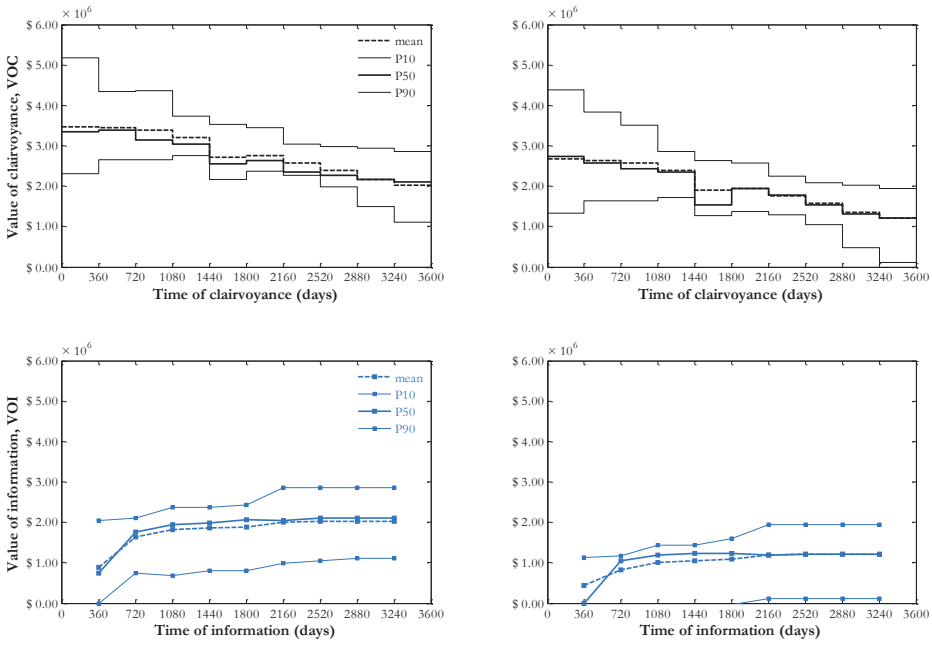


Figure 4.21: Results for the Egg model with multiple observation times: VOC and VOI including (left) and disregarding (right) the uneconomical production of the plausible truths.

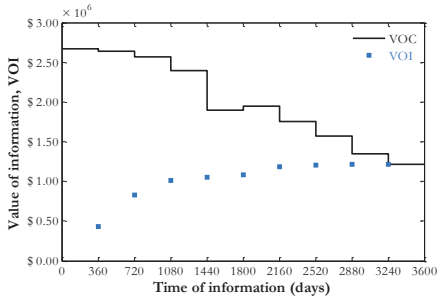


Figure 4.22: Results for the Egg model with multiple observation times disregarding the uneconomical production of plausible truths: expected VOC and VOI.

Finally, we end up with a great amount of simulated data whose analysis may reveal opportunities to improve field operations. Although we consider ensembles with a reduced number of model realizations and plausible truths, we still obtain multiple CLRM sequences, including optimal production strategies (Figure 4.23) and simulation forecasts (Figure 4.24 and Figure 4.25). An in-depth analysis of these data may lead to a better understanding on how to manage the reservoir, serving as a sort of pre-computed operation manual.

In Figure 4.23, we note that the production strategies optimized with perfect geological knowledge already from $t = 0$ make a much better use of the operational flexibility available, allowing a better control over the reservoir system to improve its performance. The strategies derived through CLRM manage to use some of this flexibility, but not as efficiently as in the case with clairvoyance because the learning obtained from production measurements is more limited. The robust strategy determined under prior uncertainty is not able to capture the possibility of reacting to future measurements.

Figure 4.24 and Figure 4.25 depict the performance of the production strategies discussed above in terms of the NPV predictions for the $N_{truth}^{repr} = 10$ representative plausible truths considered. First, we note that the considerations on the uneconomical production concerns mainly the production strategies derived under prior uncertainty. Besides that, we observe that the CLRM strategies result in higher NPV values, but also in a larger spread. Finally, we notice a distinct behavior of the NPV curves derived with CLRM and under clairvoyance: the curves of the different plausible truths cross each other, while the curves derived under prior uncertainty never do that. The reason for this behavior is that

we have a single production strategy optimized with the prior knowledge and one optimal strategy for each plausible truth in the other cases. These multiple optimal strategies have been tailored to their respective plausible truths, which allows them to achieve higher NPV values. This also makes them achieve their maximum at different times, causing the curves to cross.

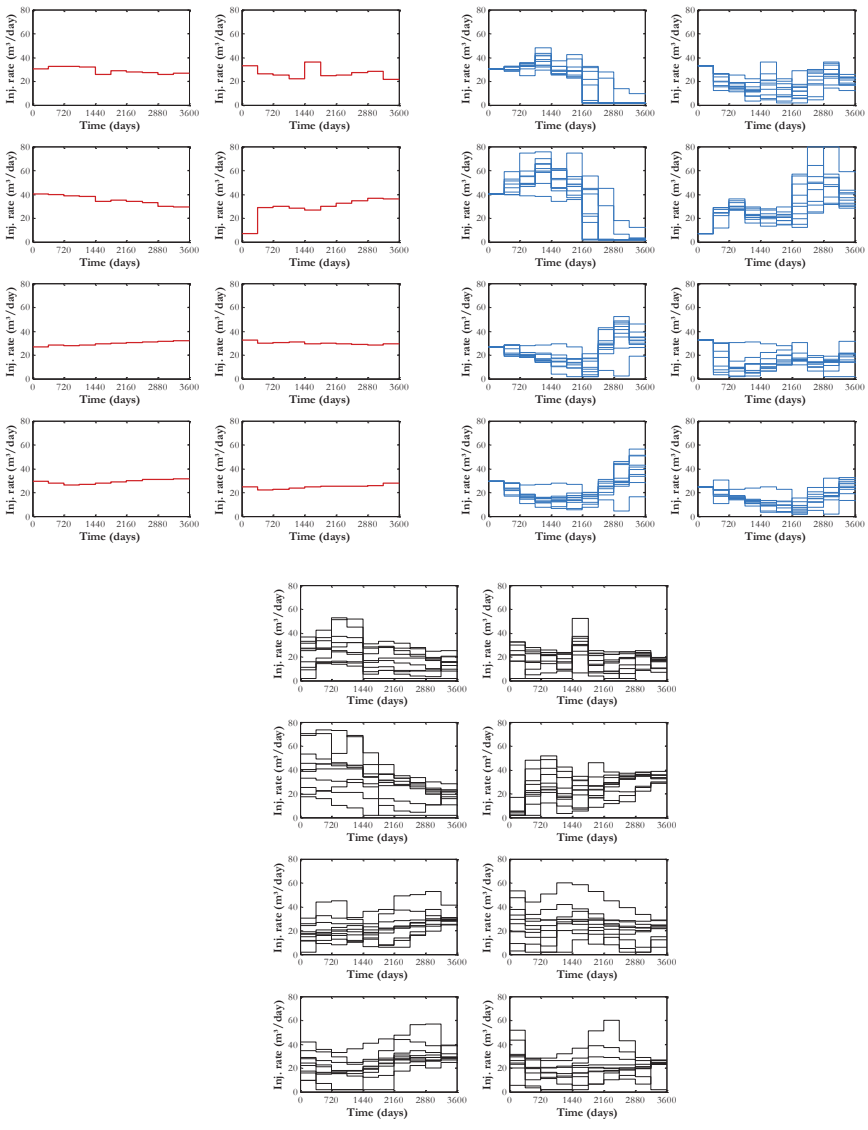


Figure 4.23: Optimal production strategies for the plausible truths considered Egg model: optimized under prior uncertainty (top left); obtained through CLRM with additional production measurements (top right); optimized under the assumption of clairvoyance available at $t = 0$.

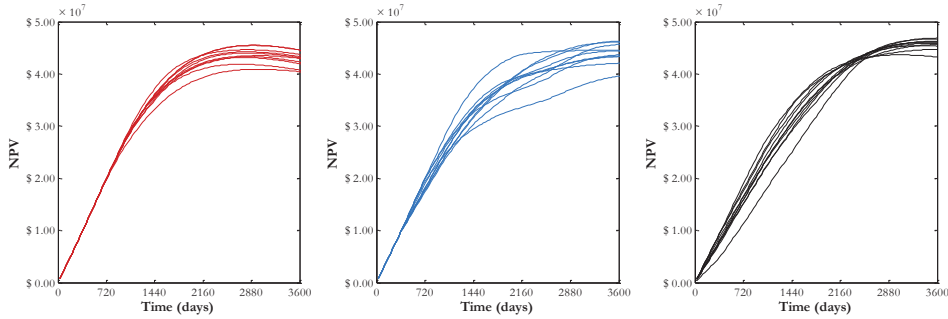


Figure 4.24: Cumulative NPV curves for the 10 representative plausible truths including uneconomical production (Egg model) for the production strategies from Figure 4.23.

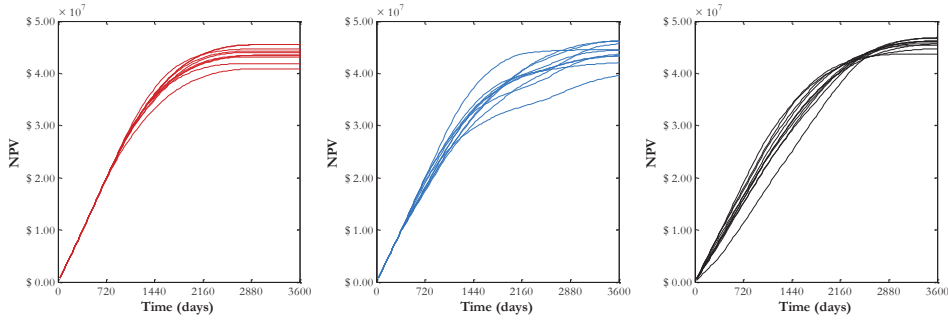


Figure 4.25: Cumulative NPV curves for the 10 representative plausible truths disregarding uneconomical production (Egg model) for the production strategies from Figure 4.23.

4.5. Discussion and conclusions

We applied two different measures to decrease the amount of simulations required in VOI assessment workflows. First, we showed how to make the robust optimization and history matching procedures more efficient by reducing the size of the ensembles considered. We used several model-based features to select representative realizations through clustering and compared them against reference results. For the small 2D five-spot example (441 grid blocks), we concluded that oil saturation snapshots data transformed by tensor decomposition or MDS provide a good basis for the selection of reduced ensembles to speed-up robust optimization and history matching. Second, we introduced two new decision-based features to support the selection of representative plausible truths as another alternative to speed-up the VOI assessment procedure. The results obtained for the 2D example suggest that optimal production strategies are most suitable for this

purpose, confirming that the choice of the selection feature is case-dependent even within the same workflow. A disadvantage of this approach is the challenge in deriving meaningful statistics of the VOI given the reduced number of plausible truths. Finally, we combined both aforementioned acceleration measures to design a new procedure for faster VOI assessment. For the 2D example, we were able to reduce the amount of required reservoir simulations from millions to tens of thousands. This significant reduction in computational costs represents an important step for the use of the VOI assessment in larger examples.

Based on the learning from the experiments with the small 2D model, we tested the acceleration measures on the medium-sized Egg model (18,553 grid blocks). The results confirmed the main conclusions obtained for the small model, reassuring us that clustering based on the appropriate features can support the selection of representative models to approximate the uncertainty characterized by the full ensembles of realizations and make our workflows much more efficient also in larger examples. In the end, by combining all the acceleration measures, we were able to apply our methodology for VOI assessment to the Egg model, which would otherwise have been computationally intractable. The results obtained supported our main conclusions from the previous chapters, which reassures the consistency of our approach.

However, there is still scope for future research to further accelerate these workflows. In particular, the combination of the ideas presented in this paper with the use of surrogate models (e.g., proxies), reduced-physics models or multiscale methods may be necessary to develop practical tools for VOI assessment to be applied in realistic, large-scale applications.

5

Informed production optimization

This chapter [1] introduces a new approach for production optimization in the context of the closed-loop reservoir management (CLRM) by considering the impact of future measurements within the optimization framework. CLRM enables instrumented oil fields to be operated more efficiently through the systematic use of life-cycle production optimization and computer-assisted history matching. In Chapter 2, we introduced a methodology to assess the value of information (VOI) of measurements in such a CLRM approach a-priori, i.e. during the field development planning phase, to improve the planned history matching component of CLRM. The reasoning behind the a-priori VOI analysis unveils an opportunity to also improve our approach to the production optimization problem (i.e., optimization of well rates and pressures) by anticipating the fact that additional information (e.g., production measurements) will become available in the future. Here, we show how the more conventional optimization approach can be combined with VOI considerations to come up with a novel workflow, which we refer to as informed production optimization (IPO). We illustrate the concept with a simple water flooding problem in a two-dimensional five-spot reservoir and the results obtained confirm that this new approach can lead to significantly better decisions in some cases.

5.1. Introduction

Recent tools developed to support decision making in reservoir operations rely on model-based optimization and uncertainty quantification to find the best production strategy. Closed-loop reservoir management (CLRM) goes a step further and takes advantage of the frequent measurements collected throughout the reservoir producing life-time to determine the optimal set of controls. This is achieved through the systematic use of life-cycle optimization in combination with computer-assisted history matching. This

[1] This chapter is based on Barros, E.G.D., Van den Hof, P.M.J. and Jansen, J.D. (2018). Informed production optimization. Paper in preparation.

combination provides the ability to react to the measurements from the true reservoir (i.e., through the designed surveillance plan), offering the opportunity to benefit from the remaining flexibility of the production strategy and compensate for possibly wrong previous decisions, which are deemed to be suboptimal due to the presence of uncertainty.

A major challenge in determining the optimal production strategy arises from the nature of the main uncertainties inherent to the reservoir management problem. The geological uncertainties are endogenous (Jonsbråten, 1998): they only get revealed at the cost of decisions. In other words, the production measurements that can help reducing the uncertainty can only be gathered when the field is operated, which requires a decision to be made (i.e., an operational strategy to be defined). In practice this means that the future state of uncertainty depends on current and previous decisions. Therefore, a truly optimal production strategy can only be determined by considering its own impact in the future measurement outcomes and their consequences to the future decisions to be made under the future state of uncertainty.

Recently, we have proposed a methodology in which the CLRM framework is used to assess the value of future measurements (Chapter 2). Such a methodology allows one to quantify the expected value achievable with the improved decision making enabled by the selected surveillance plan. Thus, it can be used as a tool to assist in the design of the optimal surveillance plan (i.e., measurement type, location, frequency, precision, …).

In order to estimate the VOI for a given surveillance plan, we calculate the additional value of the future measurements in terms of the value enabled by the production strategies re-optimized in a closed-loop fashion with the new information. Because it considers the availability of future information, the VOI assessment framework can potentially also help eliminating the shortcoming of the traditional optimization approach related to the endogenous nature of the uncertainties. In this chapter, we therefore propose to integrate VOI considerations into the production optimization framework to come up with a novel workflow, which we refer to as informed production optimization (IPO).

In the Background section (5.2) we recap the key concepts for the development of our new method and relate to previous work addressing problems of similar nature. Next, in the Two-stage IPO section (5.3), we introduce the proposed workflow for the simplest

application of IPO, showing how it builds upon the VOI assessment framework presented in Chapter 2. Thereafter, we illustrate it with a simple case study, describing details of our implementation and analyzing the improved performance obtained with IPO. After that, we discuss in the Multistage IPO section (5.4) how the methodology can be extended to a more general case with more decision stages. Finally, in the Discussion and conclusion section (5.5), we reflect on the advantages of this new concept to optimize production strategies, and we suggest directions for future work.

5.2. Background

5.2.1. VOI assessment in CLRM

Recently, we have proposed a new methodology to assess the VOI of future measurements by making use of the CLRM framework (Chapter 2). Our approach presented there consists of "closing the loop" in the design phase to simulate how information obtained during the producing life-time of the reservoir comes into play in the context of optimal reservoir management. By considering both data assimilation and optimization in the procedure, we are able "to not only quantify how information changes knowledge, but also how it influences the results of decision making" (Chapter 2). This is possible because a new production strategy is obtained every time the models are updated with new information, and the strategies with and without additional information can be compared in terms of the value of the optimization objective function (typically NPV) obtained when applying these strategies to a virtual asset (a synthetic truth); see the flowchart in Figure 2.11.

We define the VOI of future measurements as the additional value realized by the asset when the future information is utilized for the optimization of the subsequent well controls. Since we do not have perfect knowledge of the true asset, we use an ensemble of plausible truths. For a problem with M control time intervals and considering the i^{th} plausible truth $\mathbf{m}_{truth}^{i}$, the baseline for VOI calculation is the value J_{prior}^{i} obtained without the future information. This involves optimization of all the well controls $\mathbf{u}_{prior} = \left[\mathbf{u}_{prior,1}^{T}\ \mathbf{u}_{prior,2}^{T}\ \ldots\ \mathbf{u}_{prior,M}^{T}\right]^{T}$ under the initial state of uncertainty (i.e., over the prior ensemble of model realizations). Note that here we consider the approximation

presented in Chapter 2 to accelerate the procedure: although multiple prior ensembles are used to perform the history matching for the different plausible truths, a single ensemble is used to determine the strategy $\mathbf{u}_{prior}$ optimal under initial uncertainty. Starting from this baseline, the additional value VOI_j^i of the $\mathbf{d}_j^i$ measurements gathered during the j^{th} control interval is the difference between the value $J_{post,j}^i$ obtained through the re-optimization of the well controls in the subsequent control time intervals $\mathbf{u}_{post}^i = \left[\mathbf{u}_{post,j+1}^{i\,T}\ \mathbf{u}_{post,j+2}^{i\,T}\ \ldots\ \mathbf{u}_{post,M}^{i\,T}\right]^T$ compared to the baseline value J_{prior}^i. Note that the values J_{prior}^i and $J_{post,j}^i$ are not statistics of the ensembles used in the analysis, but the value produced by implementing the strategies $\mathbf{u}_{prior}$ and $\mathbf{u}_{post}^i$ to the plausible truth $\mathbf{m}_{truth}^i$. Note also that the well controls prior to the $\mathbf{d}_j^i$ measurements, i.e. $\mathbf{u}_{prior}(0:t_j) = \left[\mathbf{u}_{prior,1}^{T}\ \mathbf{u}_{prior,2}^{T}\ \ldots\ \mathbf{u}_{prior,j}^{T}\right]^T$, are the ones determined by optimization without future information (i.e., under initial uncertainty).

5.2.2. Future information and optimization

The concept of accounting for the availability of future information within the optimization is not new and has been investigated for several applications in different scientific communities. In operations research, such an optimization problem is addressed from a more mathematical perspective, being referred to as stochastic programming. Two-stage or multistage models are used to account for the sequential nature of the decisions, which allows the optimization to be expressed in a nested formulation. For more information, see Birge and Louveaux (1997), Ruszczyński and Shapiro (2003), and Georghiou et al. (2011). In the formalism from decision and game theory, non-myopic decision rules are the ones where the decision makers look ahead and consider future information, opposed to the myopic approach in which the influence of current decisions on the future state of uncertainty (i.e., conditional posterior distributions) is ignored (Mirrokni et al., 2012). The artificial intelligence community refers to this class of problems as partially observable Markov decision processes (POMDPs), related to applications where direct observations of the state of the uncertain processes are not available (Smallwood and Sondik, 1973; Hauskrecht, 2000). In these cases, the decisions are optimized not only based on their direct contribution to produce value but also to maximize the expected pay-off (or reward) of subsequent decisions (Krause and Guestrin, 2007 and 2009). In systems and control theory, the dual control introduced by Fel'dbaum

(1960 and 1961) seeks to determine the optimal trade-off between excitation and control to promote a more active learning from the measurements while directing the system to its optimal state. This is only possible through the definition of control policies that anticipate the availability of future measurements and their learning effect. In this respect, Van Hessem (2004) discussed the steps to be taken to turn the traditional open-loop MPC (i.e., model predictive control) methods into feedback mechanisms that know how to respond to future measurements. More recently, Hanssen (2017) proposed an implicit dual MPC controller that explicitly includes the feedback mechanism in the optimization problem. In system identification, Forgione et al. (2015) investigated the use of different model update strategies to enable batch-to-batch improvement in the control of industrial processes.

The main applications of these ideas are in: logistics and supply chain problems; network problems such as traffic control and power grids; medical decision making; planning and scheduling. In the oil and gas upstream sector, Jonsbråten (1998) applied stochastic programming to drilling sequence optimization in a simplified setting. Goel and Grossmann (2004) used stochastic models in the planning of offshore gas fields with uncertainties related to the reserves, but without considering reservoir simulation models. Foss and Jensen (2011) claimed that a conscious exploitation of the dual effect of controls can be significant in a reservoir system given the typical large uncertainties, but highlighted the fact that the dual-control problem is unsolvable in practice. Zenith et al. (2015) investigated the use of sinusoidal oscillations as a means of exciting the reservoir to obtain more informative data during well testing, but without analyzing the influence on the subsequent decisions. More recently, Abreu et al. (2015) have proposed a methodology to optimize under uncertainty the valve settings of smart wells considering the possibility of acquiring additional information in the future. Their approach uses the ideas of approximate dynamic programming to make the problem tractable.

5.3. Two-stage IPO

In the context of production optimization, the decision in question is the use of an optimized production strategy that maximizes the value of the asset. Despite the limited potential for learning due to the poor information content of typical production

measurements, the knowledge about a reservoir changes along its producing life-time. Thus determining a single production strategy by performing robust optimization based on the initial state of uncertainty (Van Essen et al., 2009) is equivalent to ignoring the endogenous nature of the reservoir management problem. It is a myopic approach which will likely lead to suboptimal solutions. The stochastic programming ideas discussed in the previous section can be a solution to overcome this limitation.

In this chapter we present a workflow for non-myopic production optimization within the CLRM framework. The proposed procedure is a combination of classical stochastic programming and our previously introduced workflow for VOI assessment in CLRM (Chapter 2). The idea is to model the decision process related to the reservoir management problem with the help of elements of the CLRM framework (i.e., ensemble-based uncertainty quantification, model-based optimization and computer-assisted history matching), which allows us to represent the sequential character of the decisions (corresponding to a Markov decision process) while accounting for future measurement data with very limited information content. Because this approach considers future information when performing production optimization, we name it informed production optimization (IPO).

The ultimate goal of reservoir management is to maximize the value delivered by the asset with the implementation of production and surveillance strategies. Our methodology for VOI assessment (Chapter 2) provides a framework to quantify the value $J^i_{post,j}$ to be produced by the asset (here, the i^{th} plausible truth) with the incorporation of future information gathered through the designed surveillance strategy (here, measurements collected during the j^{th} control time interval). For simplicity, we initially address the case with only a single observation time. Later, we discuss how these ideas can be extended to cases with multiple observation times.

In the conventional approach for production optimization, we seek to maximize the predicted objective function for the ensemble of models anticipating that in a closed-loop setting, once more information becomes available, we will have the opportunity to improve our predictive models and adjust our strategies to achieve the best possible $J^i_{post,j}$. By following this reasoning, the optimization is making use of the flexibility of the remaining degrees of freedom $\mathbf{u}^i_{post}(t_{j+1}:T) = \left[\mathbf{u}^i_{post,j+1}{}^T \ \mathbf{u}^i_{post,j+2}{}^T \ \ldots \ \mathbf{u}^i_{post,M}{}^T\right]^T$, while the

part of the strategy prior to the future information $\mathbf{u}_{prior}(0:t_j)=\left[\mathbf{u}_{prior,1}{}^T\ \mathbf{u}_{prior,2}{}^T\ \ldots\ \mathbf{u}_{prior,j}{}^T\right]^T$ is deemed to be suboptimal. We propose here to use $J^i_{post,j}$ as the cost function for our production optimization problem, as a way of reflecting the true goal of reservoir management when optimizing $\mathbf{u}_{prior}(0:t_j)$. In fact, since we consider an ensemble of N_{truth} equiprobable plausible truths, we still perform robust optimization and our new objective function is defined as

$$\mu_{post,j}\left(\mathbf{u}(0:t_j),t_j\right)=\frac{1}{N_{truth}}\sum_{i=1}^{N_{truth}}J^i_{post,j}\left(\mathbf{u}(0:t_j),\mathbf{u}^i_{post}(t_{j+1}:T),t_j\right), \tag{5.1}$$

where $\mu_{post,j}$ is the ensemble mean of the objective function values $J^i_{post,j}$ for the individual plausible truths. The objective function $J^i_{post,j}$ for a single realization i of the plausible truth is calculated according to the VOI assessment workflow for a single observation time (Figure 2.11 in Chapter 2) and involves the solution of history matching with the future information gathered during control interval j and re-optimization of the production strategy for the subsequent control intervals $\mathbf{u}^i_{post}(t_{j+1}:T)$. We can also see the optimization of this new cost function as a two-stage stochastic model or a nested optimization problem where the outer optimization concerns the production strategy up to control interval j and the inner optimization determines the remaining part of the strategy, which will be different for each one of the N_{truth} plausible truths. Thus, the IPO problem can be formulated as

$$\mathbf{u}_{IPO}(0:t_j)=\arg\max_{\mathbf{u}}\frac{1}{N_{truth}}\sum_{i=1}^{N_{truth}}J^i_{post,j}\left(\mathbf{u}(0:t_j),\mathbf{u}^i_{post}(t_{j+1}:T),t_j\right), \tag{5.2}$$

where the outcome of the optimization is a single optimal production strategy $\mathbf{u}_{IPO}(0:t_j)$ until the observation time t_j, and N_{truth} optimal strategies $\mathbf{u}^i_{post}(t_{j+1}:T)$, $i=1,\ldots,N_{truth}$, for the remaining producing time. The proposed workflow to solve this optimization problem iteratively is displayed in Figure 5.1. The procedure resembles the workflow of Figure 2.11, but contains an outer iterative loop (shown in blue) to keep updating the part of the production strategy prior to the observation time. The inner part of the workflow (indicated in yellow) remains the same as in the original VOI assessment workflow presented in Chapter 2 (Figure 2.11); the only difference being that, for the optimization problem, we are only interested in evaluating the new cost function $J^i_{post,j}$ for each plausible truth and not in computing the associated VOI. Note that, for the case with a

single observation time, the optimizations required in this part of the workflow reduce to the conventional robust optimization problem, because there are no more future measurements to be considered after t_j. Therefore, the optimal strategies $\mathbf{u}^i_{post}(t_{j+1}:T)$ can be determined by following the conventional robust optimization approach proposed by Van Essen et al. (2009). Note also that, just like the VOI workflow can be simplified to quantify the value of clairvoyance (VOC) (Chapter 2), the IPO procedure can be modified to perform optimization under the assumption that perfect revelation of the truth will be possible in the future.

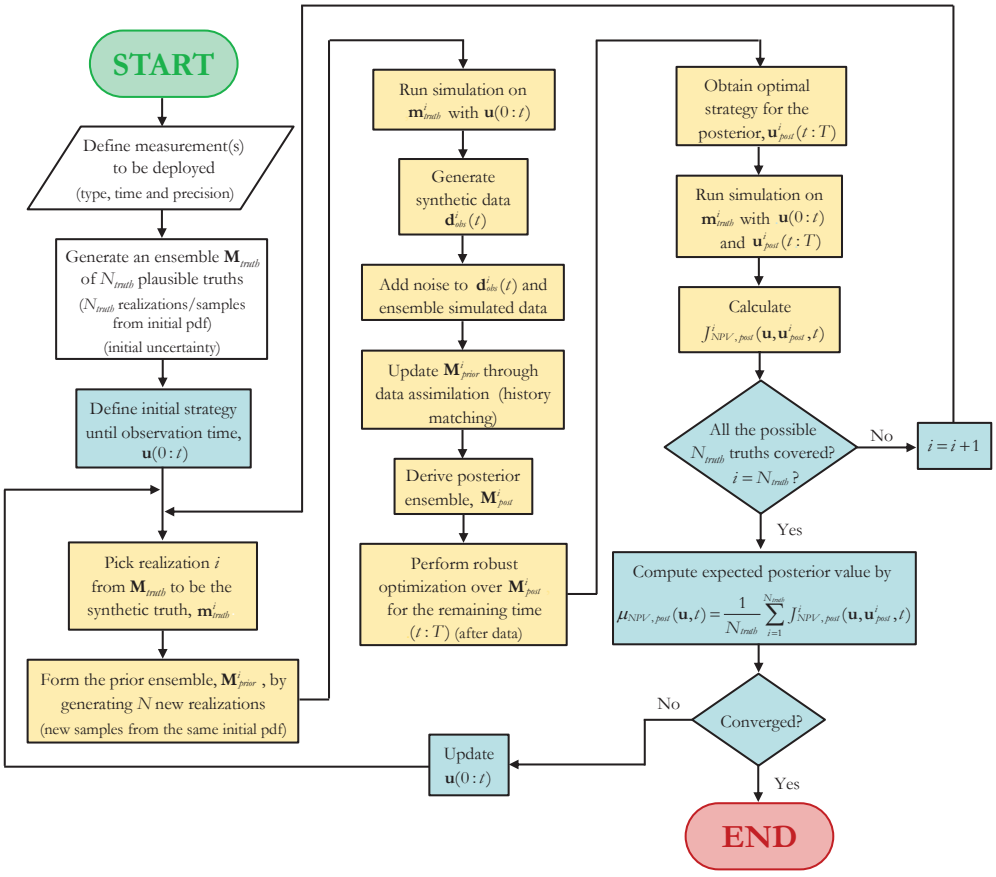


Figure 5.1: Workflow for IPO with a single observation time.

By recommending a single $\mathbf{u}_{IPO}(0:t_j)$ and multiple $\mathbf{u}^i_{post}(t_{j+1}:T)$ we design a flexible production strategy that accounts for the fact that we have the opportunity to update well controls after collecting the measurements at t_j. Because we consider an ensemble of

plausible truths, we recognize that the outcome of the future measurements (or future clairvoyance) is unknown, and that the ensemble of $\mathbf{u}^{i}_{post}(t_{j+1}:T)$ controls should reflect this uncertainty in the production strategy to be implemented after t_j. Only when we proceed with the reservoir operations (i.e., by truly implementing $\mathbf{u}_{IPO}(0:t_j)$) and the actual measurements take place at t_j, the uncertainty model can be updated and the optimal strategy to be implemented for the remaining time can be determined.

From the algorithmic point of view, this nested optimization can be solved just as any other optimization problem. The same methods apply but the cost function to be evaluated is more complex. Here, we use gradient-based methods for both the inner and the outer optimizations. In this setting, the main challenges concern the outer optimization, in particular for the definition of the starting point and the computation of the gradient for such a complex cost function. We explain the details of our implementation when we describe our case study in the section 5.3.1.

5.3.1. Example

To test our approach, we used the same 2D five-spot example from the previous chapters. See Chapter 2 for a more detailed description. We used $N_{truth} = 50$ plausible truths and $N_{truth} = 50$ ensembles of $N = 49$ realizations of the porosity and permeability fields, conditioned to hard data in the wells, to model the geological uncertainties. The simulations were used to determine the set of well controls (bottom-hole pressures) that maximizes the NPV. The optimizations were run for a $T = 1{,}500$-day time horizon with well controls updated every 150 days, i.e. $M = 10$, and, with five wells, $\mathbf{u}$ had 50 elements. We applied bound constraints to the optimization variables (200 bar $\leq p_{prod} \leq$ 300 bar and 300 bar $\leq p_{inj} \leq$ 500 bar). The whole exercise was performed in the open-source reservoir simulator MRST (Lie et al., 2012). For the inner part of the IPO workflow (shown in yellow in Figure 5.1), we used the same setup as the one used in our previous work: a CLRM environment created by combining the adjoint-based optimization and EnKF modules available with MRST (see Chapter 2).

To perform the outer optimization (shown in blue in Figure 5.1), we used an implementation of the StoSAG method (Fonseca et al., 2016). The ensemble-based gradient computation was very convenient for our IPO problem, allowing us to use a

black-box approach to obtain the gradients for the complex and expensive cost function while benefiting from the computational advantages for cases with the presence of uncertainties. The main difference with the use of StoSAG in a conventional production optimization problem is that here the uncertainties are characterized by the ensemble of plausible truths. Thus we pair each of the $N_{pert} = 50$ perturbations of the well controls to one of the $N_{truth} = 50$ plausible truths and we estimate the search direction by an approximate linear regression of their $J^i_{post,j}$ values. We used a standard deviation of $\sigma_{pert} = 0.01$ to generate the perturbations of the well controls. The starting point for the outer optimization is the solution obtained with the conventional robust optimization procedure.

For this case study, we considered the availability of oil flow rate measurements in the producers with absolute measurement errors ($\varepsilon_{oil} = 5$ m³/day). Just like in Chapter 2, the proposed workflow considers a single observation time but was repeated for different observation times, $t_{data} = \{150, 300, \ldots, 1{,}350\}$ days. After applying the IPO procedure, we computed the VOI for each of the nine observation times to compare with the VOI obtained when using conventional robust optimization. In addition to that, we repeated the experiments by considering future availability of clairvoyance instead of imperfect measurements.

Figure 5.2 illustrates the motivation for us to select oil rate measurements for this example. It compares the VOC with the value of observing total rates and water-cuts and the value of measuring only oil rates, all obtained by following the conventional optimization approach and displayed in terms of their mean values. We observe that the VOI of total rate and water-cut measurements is very close to the VOC, where the latter represents an upper bound. On the other hand, the VOI of oil rate measurements is considerably lower than the VOC, which suggests that there is room for improvement. The low values, especially for the measurements at $t_{data} = 300$ days, are related to the inability of the oil rate measurements alone to accurately provide information about the water breakthrough time in the producers, whose prediction is paramount to achieve optimal reservoir management in this example. Since our goal here was to verify the potential of our proposed IPO approach to enhance the value produced by the asset, we considered the case with oil rate measurements to be a more suitable case study.

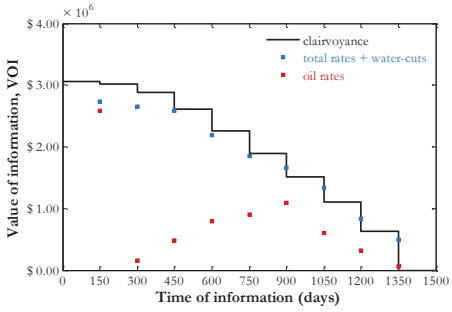


Figure 5.2: Results for the VOI analysis of production data (2D five-spot model): expected VOC, VOI of total rate and water-cut measurements, and VOI of oil rate measurements.

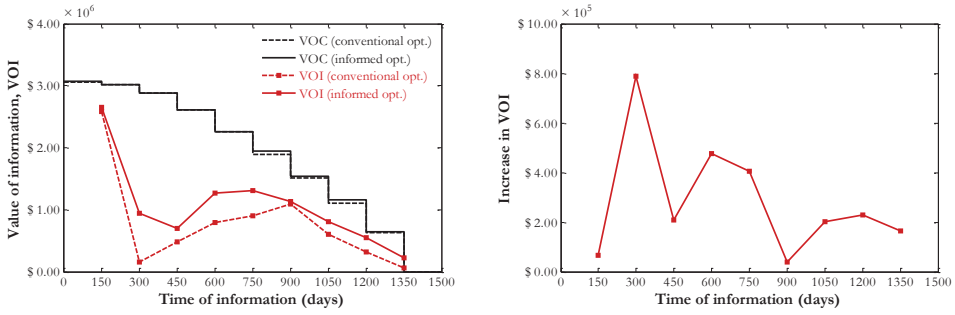


Figure 5.3: Results for the VOI analysis (2D five-spot model) with the conventional robust optimization and the IPO approaches (expected VOI).

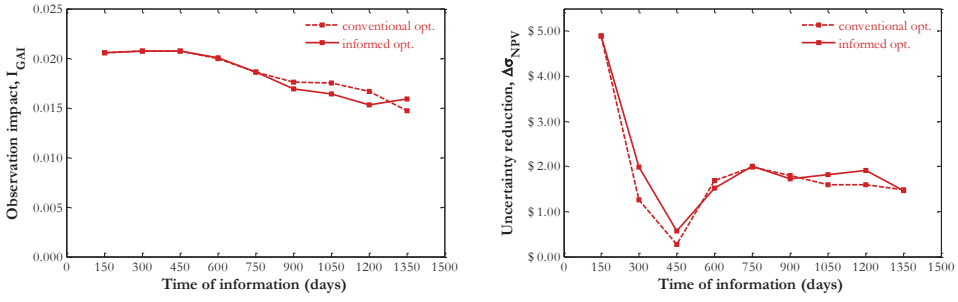


Figure 5.4: Observation impact (left) and uncertainty reduction (right) obtained for the 2D five-spot model with the conventional robust optimization and the IPO approaches (mean values).

Figure 5.3 (left) shows the results obtained with the IPO approach in comparison to those obtained when using the conventional robust optimization approach. First, we notice that the IPO does not improve significantly the VOC. This happens because, in the unrealistic case of clairvoyance, the revelation of the truth is, by definition, perfect, irrespective of the production strategy $\mathbf{u}_{prior}(0:t_j)$ that is implemented prior to it. Thus, the gain of using

the IPO approach would be more important in cases where the sub-optimality of $\mathbf{u}_{prior}(0:t_j)$ cannot be compensated by the re-optimization of the remaining degrees of freedom $\mathbf{u}^i_{post}(t_{j+1}:T)$ with perfect revelation of the truth, which does not seem to be the case for this example.

On the other hand, for the case with imperfect oil rate measurements, we observe a higher VOI for the IPO approach (Figure 5.3). Note that, for the VOI calculation, the values without information J^i_{prior} are identical for both approaches. This means that the IPO was able to determine production strategies that create more value when applied to the plausible truths. Besides that, this shows that there is indeed room to achieve improved reservoir management by accounting for the availability of future information when searching for the optimal well controls. We highlight the significant increase in VOI obtained for the measurements at $t_{data} = 300$ days, resulting in an improvement of approximately \$ 0.8 million.

In order to explain why the strategies determined by the IPO approach are better, we looked also at their effect in terms of changing the observation impact I_{GAI} and the uncertainty reduction $\Delta\sigma_{NPV}$ associated with the measurements (see Chapter 2). Figure 5.4 displays these results for comparison with the VOI in Figure 5.3. The observation impact plot indicates that the production strategies obtained with both optimization approaches produce oil rate measurements with similar information content, in terms of predicting its own outcome (i.e. observation self-sensitivity). The results in Figure 5.4 (right) also suggest that both optimization approaches perform similarly in terms of producing measurements to reduce the uncertainty in NPV forecasts. We note, however, an increase in the uncertainty reduction obtained with the IPO approach for the measurements at $t_{data} = 300$ and $t_{data} = 450$ days, which is consistent with the improvement observed in terms of VOI for those observation times (Figure 5.3).

Figure 5.5 shows the production strategies obtained with both optimization approaches for the case considering measurements at $t_{data} = 300$ days, which correspond to the largest improvement in terms of VOI. For both approaches we observe a single production strategy $\mathbf{u}_{prior}(0:300)$ defined until $t_{data} = 300$ days and multiple strategies $\mathbf{u}^i_{post}(300:1500)$ after that. We notice small differences between the $\mathbf{u}_{prior}(0:300)$ strategies for the two optimization approaches but a reasonable difference in the range of

the $\mathbf{u}^{i}_{post}(300:1500)$ strategies. Thus, for this example, small changes in $\mathbf{u}_{prior}$ seem to have a significant impact on the optimization of the remaining controls $\mathbf{u}^{i}_{post}$.

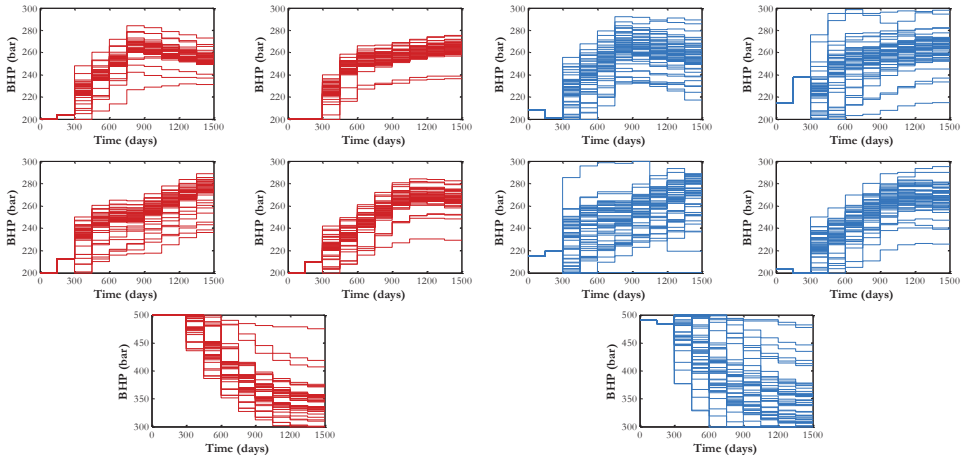


Figure 5.5: Production strategies obtained for oil rate measurements at t_{data} = 300 days (2D five-spot model): conventional optimization (left) and IPO (right).

We conclude that, for this example, the increase in VOI due to the IPO approach can be partially attributed to the ability of the improved production strategies to yield measurements that result in additional uncertainty reduction in terms of the quantity of interest (NPV). Nevertheless, there seem to be other effects that enable the IPO strategies to perform better for the plausible truths. This could be related to an enhanced controllability of the reservoir states once the measurements reveal some of the initial uncertainty. Overall, these results suggest that, when relying on measurements with limited information content, wrong reservoir management decisions cannot be entirely compensated by re-optimizing the subsequent decisions. In such cases, an approach which is merely reactive to measurements is more likely to result in poorer decisions than an approach that anticipates the availability of future information.

5.4. Multistage IPO

The results of the example above show that the IPO approach can be valuable for cases with imperfect measurements. However, the procedure depicted in Figure 5.1 only addresses the case considering a single observation time, which is not realistic for well

production measurements. Therefore we discuss here what can be done to apply the same reasoning in problems with multiple observation times.

In Chapter 3, we showed how our original methodology for VOI assessment can be modified to account for multiple observation times. We demonstrated that, contrary to what one would expect, the computational costs of assessing the value of multiple measurements do not increase beyond reasonable levels. The reason is that, although one might think that we need to consider new additional plausible truths for every observation time, this is not necessary. Once a realization of the initial ensemble is picked to be the plausible truth, the same realization plays the role of truth and the "loop is closed" as many times as the number of observation times, because that is the only way of ensuring consistency of the synthetic future measurements that we need to generate in order to assess their VOI. Thus, the computational costs of assessing the VOI for multiple observation times is the same as those of the VOI assessment for a single observation time repeated for different observation times.

Unfortunately, the extension of IPO to cases with multiple observation times is not that simple. In VOI assessment only, the plausible truths are not directly involved in the optimization, but they do play a role in the IPO. Thus, sticking to one of the realizations as the plausible truth throughout all the observation times would imply that we can identify the truth with the measurements obtained at the first observation time, which corresponds to the availability of clairvoyance. Since we are dealing with imperfect measurements, in order to be realistic, we need to update the uncertainty model given by our ensemble of plausible truths. This means that, in principle, the size of the problem grows exponentially with the number of observation times considered, if it is to be solved rigorously. In other words, the number of branches on the scenario tree associated with the problem would be multiplied for every new observation. Figure 5.6 displays a decision tree for a problem with four observation times where the uncertainty is represented by three new plausible truths each time. We notice the rapid increase in the number of scenarios, which would be even more dramatic if we consider tens of plausible truths and more observation times.

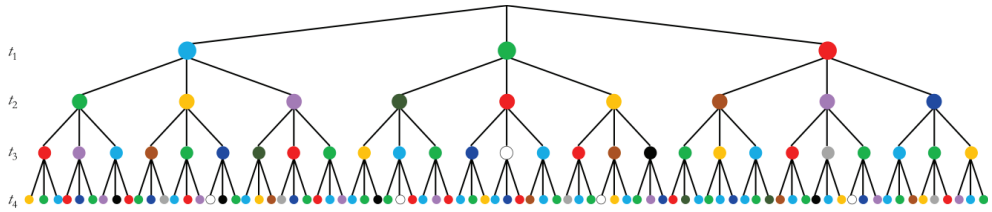


Figure 5.6: Full plausible truth scenario tree for rigorous solution of IPO with four observation times and ensembles of three realizations.

Given the already prohibitive computational costs of VOI assessment, it is safe to say that the rigorous solution of this scenario tree is not feasible for problems relying on reservoir simulation models. Gupta and Grossmann (2011) and Tarhan et al. (2013) have proposed strategies to solve multistage stochastic programming models more efficiently through approximate solutions involving decomposition and constraint relaxation methods.

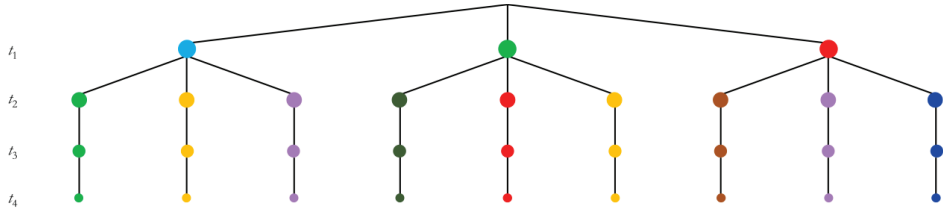


Figure 5.7: Pruned plausible truth scenario tree for approximate solution of IPO with four observation times and ensembles of three realizations.

Here, we propose to use a more intuitive approach to simplify this problem. The first branching before t_1 is the most important one, as it introduces the different plausible truths that will be used to generate the synthetic measurements and to evaluate the performance of the production strategies derived. The second branching before t_2 is also important because it establishes that the respective plausible truth is not perfectly revealed by the imperfect observations made at t_1. However, the branching events associated with the remaining observation times are of smaller importance, because the one at t_2 already guarantees that there are always multiple branches below each plausible truth, reflecting the fact that at any time the uncertainty around the plausible truths is never revealed through unrealistic clairvoyance. This allows us to prune the scenario tree and reduce the size of the problem to be solved for IPO with multiple observation times. Figure 5.7 shows the pruned scenario tree. Note that after t_2 the number of branches does not

increase, which means that, by applying this pruning strategy, we can obtain approximate solutions for three-stage and *M*-stage stochastic problem at similar computational costs.

To put this idea into practice and extend the two-stage IPO (Figure 5.1) to *M*-stage IPO (with $M - 1$ observation times), we need to make a few changes in the workflow. Figure 5.8 depicts the modified procedure. First, we introduce an additional ensemble of N_{truth} truths $\mathbf{M}^i_{truth}$ for each one of the original plausible truths $\mathbf{m}^i_{truth}$ of our initial ensemble $\mathbf{M}_{truth}$. This corresponds to the second branching of the scenario tree before t_2 (Figure 5.7). From there on, the tree does not develop new branches but, for every observation time, $\mathbf{M}^i_{truth}$ is updated through history matching of the synthetic measurements generated with $\mathbf{m}^i_{truth}$. After the last observation, we derive the posterior ensemble of truths $\mathbf{M}^i_{truth,M-1} = \{\mathbf{m}^{i,1}_{truth}, \mathbf{m}^{i,2}_{truth}, \ldots, \mathbf{m}^{i,N-1}_{truth}\}$ and we reach the last decision stage. Like for the two-stage IPO (Figure 5.1), the decisions in the last stage are optimized following the conventional robust optimization approach. We use the posterior truths $\mathbf{m}^{i,j}_{truth}$ to generate synthetic measurements and derive the posterior ensembles $\mathbf{M}^{i,j}_{post,M-1}$ through history matching. And finally we perform robust optimization over $\mathbf{M}^{i,j}_{post,M-1}$. Because we do not introduce additional truth scenarios after the second stage, the nested optimization can group the intermediate stages into a single one. Thus, the *M*-stage IPO reduces to a nested optimization at three levels: the outer optimization (shown in blue in Figure 5.8), the intermediate optimization (in yellow) and the inner optimization (in purple). In the end, the production strategies obtained ($\mathbf{u}(0:t_1)$, $\mathbf{u}^i_{post}(t_1:t_{M-1})$ and $\mathbf{u}^{i,j}_{post}(t_{M-1}:T)$) are applied to the respective original plausible truth $\mathbf{m}^i_{truth}$. We calculate their performances $J^{i,j}_{post}$ in terms of our objective function and, finally, we compute the cost function as the mean over all the realizations. Thus, the *M*-stage IPO problem can be formulated as

$$\mathbf{u}_{IPO}(0:t_1) = \\ = \arg\max_{\mathbf{u}} \frac{1}{(N_{truth})^2} \sum_{i=1}^{N_{truth}} \sum_{j=1}^{N_{truth}} J^{i,j}_{post,j}\left(\mathbf{u}(0:t_j), \mathbf{u}^i_{post}(t_1:t_{M-1}), \mathbf{u}^{i,j}_{post}(t_{M-1}:T), t_1\right), \tag{5.3}$$

where the outcome of the optimization is a single optimal production strategy $\mathbf{u}_{IPO}(0:t_1)$ until the first observation time t_1, N_{truth} optimal strategies $\mathbf{u}^i_{post}(t_1:t_{M-1})$, $i = 1, \ldots, N_{truth}$, for the period between the first and the last observation, and $(N_{truth})^2$ optimal strategies $\mathbf{u}^{i,j}_{post}(t_{M-1}:T)$, $i = 1, \ldots, N_{truth}, j = 1, \ldots, N_{truth}$, for the remaining producing time.

The *M*-stage IPO workflow can be realized in a similar implementation as the one described for our case study (section 5.3.1). The inner optimization can be performed with the help of the adjoint-based gradients, while the intermediate and outer optimizations can make use of the StoSAG method to estimate the required gradients.

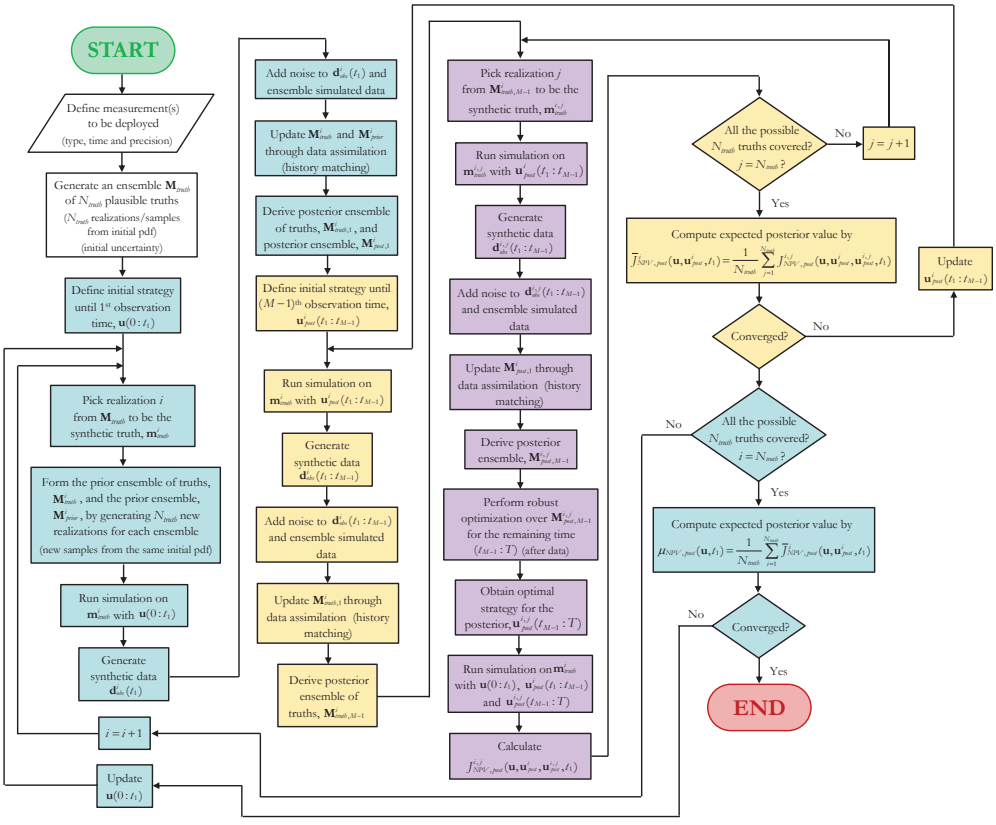


Figure 5.8: Workflow for IPO with multiple observation times.

5.4.1. Example

To test the IPO approach with multiple observation times, we used again the 2D five-spot example, but with a small modification: instead of $M = 10$, we considered $M = 3$ control time intervals. We assumed oil production measurements to be available at $t_{data} = \{500, 1{,}000\}$ days, with the same measurement error as before. In order to make the problem more tractable, we used the acceleration measures presented in Chapter 4, and we reduced the number of plausible truths to $N^{repr}_{truth} = 10$ representative ones and the size of the

ensembles for robust optimization to $N_{repr} = 5$. The history matching step with the EnKF was performed over the full ensembles.

Figure 5.9 displays the VOI of the multiple oil rate measurements obtained through the procedure from Chapter 3 with the conventional optimization approach. Figure 5.9 (bottom) shows that the VOI for the last observation time (t_{data} = 1,000 days) is significant lower than the VOC, which suggests that there might be scope to improve them with the IPO approach. Thus, the goal of the IPO here can be understood as an optimization of the VOI at t_{data} = 1,000 days, by maximizing the $J_{post}^{i,j}$ achievable through the CLRM framework with the measurements at t_{data} = {500, 1,000} days.

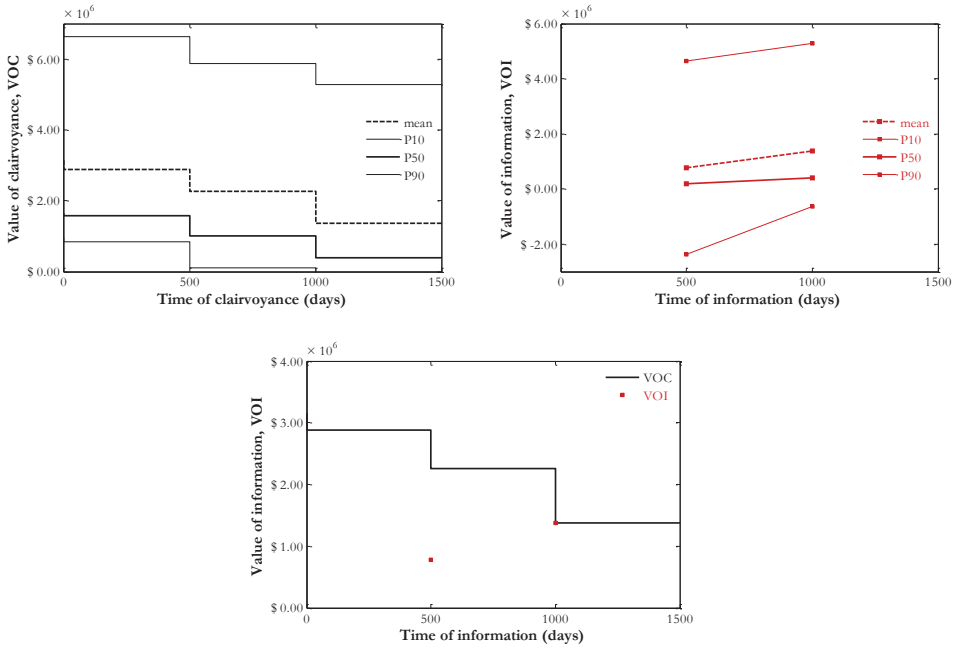


Figure 5.9: Results for the VOI analysis with the conventional robust optimization (2D five-spot model with M = 3 control time intervals).

Figure 5.10 displays the results obtained in terms of the NPV achieved by the plausible truths. Figure 5.10 (top left) compares the mean NPV for different cases: in red, we have the value attained by the production strategies optimized under prior uncertainty; in blue and in black, the NPV values corresponding to the VOI and VOC at t_{data} = 1,000 days (Figure 5.9); and, in green, the value for the IPO approach. We observe an improvement of approximately $ 55,000 for the mean NPV achieved with the IPO (green) in

comparison with the mean NPV for the CLRM using the conventional optimization approach, which represents an increase of 0.1 %, or of 4 % if expressed in terms of VOI. We note that the incremental gain is small but should be understood as an attempt to improve a solution already close to optimal for this example. This suggests that, in this example with more observations, the CLRM framework with the conventional optimization approach was able to sufficiently compensate for previous suboptimal decisions, resulting in production strategies almost as good as the ones determined through IPO.

Figure 5.10 (top right) depicts the same results but showing the empirical pdf curves derived with the NPV of the $N_{truth}^{repr} = 10$ plausible truths considered. The differences between the curves for the CLRM and IPO are small. However, the distribution for the IPO approach seems to achieve higher values at its peak, which suggests a slight reduction in the spread of the NPV values. Note that the pdf curves displayed here are the result of curve fitting with a small number of samples ($N_{truth}^{repr} = 10$), but that the empirical histograms exhibit similar trends (Figure 5.10 (bottom)).

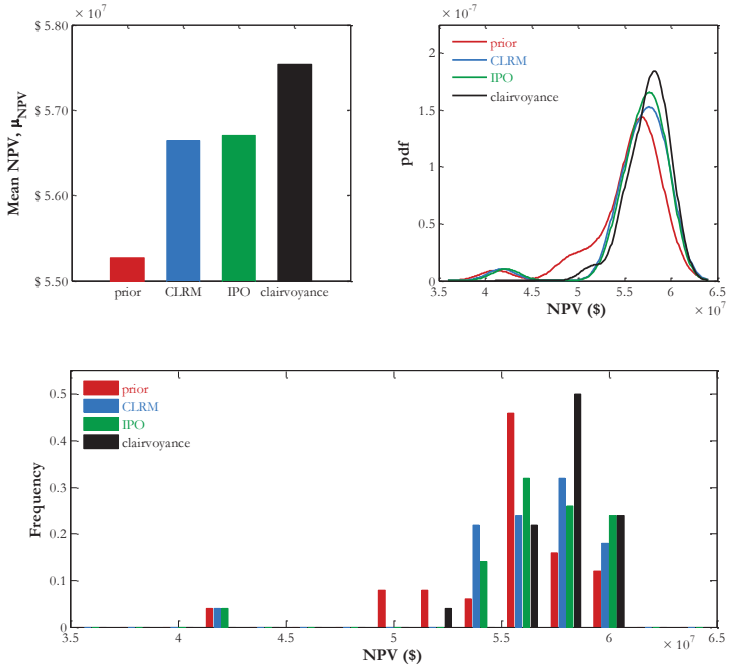


Figure 5.10: Results for the production optimization with the different approaches (2D five-spot model with $M = 3$ control time intervals): mean NPV values for the plausible truths (top left) and their respective pdf-fitted plots (top right) and empirical histograms (bottom).

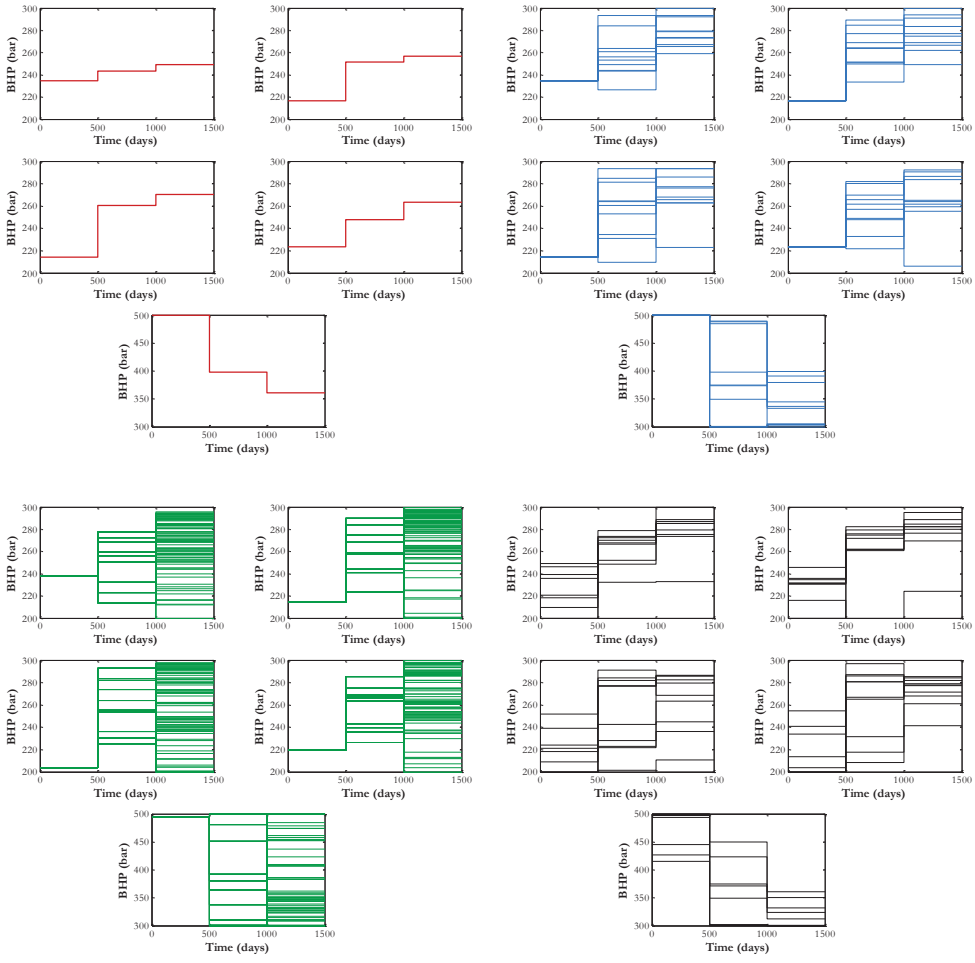


Figure 5.11: Optimal production strategies for the plausible truths considered (2D five-spot model with $M = 3$ control time intervals): optimized under prior uncertainty (top left); obtained through CLRM with additional production measurements (top right); determined by the IPO approach with future production measurements (bottom left); optimized under the assumption of clairvoyance available at $t = 0$ (bottom right).

Figure 5.11 shows the production strategies determined by the different approaches compared in Figure 5.10. A single production strategy $\mathbf{u}_{prior}(0:500)$ is derived for the conventional robust optimization approach. On the other extreme, a different optimal strategy is determined for each of the $N_{truth}^{repr} = 10$ plausible truths considered, under the assumption of clairvoyance at $t = 0$. In between, for the CLRM, the solutions obtained consist of the same single strategy as $\mathbf{u}_{prior}(0:500)$ for the first control time interval, and $N_{truth}^{repr} = 10$ different strategies $\mathbf{u}_{post}^{i}(500:1500)$ for the remaining of the time. Finally, we

see the outcome of the IPO for this three-stage problem: a single optimal production strategy $\mathbf{u}_{IPO}(0:500)$ until the first observation time $t_1 = 500$ days, $N_{truth}^{repr} = 10$ optimal strategies $\mathbf{u}_{post}^{i}(500:1000)$, $i = 1, \ldots, N_{truth}^{repr}$, for the period between $t_1 = 500$ and $t_2 = 1{,}000$ days, and $(N_{truth}^{repr})^2 = 100$ optimal strategies $\mathbf{u}_{post}^{i,j}(1000:1500)$, $i = 1, \ldots, N_{truth}^{repr}$, $j = 1, \ldots, N_{truth}^{repr}$, from $t_2 = 1{,}000$ until $T = 1{,}500$ days. The larger number of possibilities associated with future control time intervals allows the IPO strategies to account for the flexibility of the reservoir management problem. We also note that $\mathbf{u}_{IPO}(0:500)$ is not the same as $\mathbf{u}_{prior}(0:500)$ and that the difference between them enables the IPO approach to result in improved NPV values, as seen in Figure 5.10.

5.5. Discussion and conclusions

We have proposed a new approach for production optimization under geological uncertainty. The informed production optimization (IPO) approach considers the endogeneous nature of these uncertainties and includes the availability of future information to circumvent the limitations of the conventional robust optimization. To achieve this, the approach accounts for the fact that there is uncertainty about the optimal well settings to be implemented after additional measurements are processed in the future, and that this uncertainty depends on the outcome of these measurements. As a consequence, we perform robust IPO over an ensemble of plausible truths resulting in a single optimal production strategy until the moment that the future measurements take place and a range of optimal production strategies for the period after them. We applied IPO to a simple case study and observed a significant increase in the VOI of imperfect measurements, reflecting the improvement in the production optimization. After demonstrating the potential value of the IPO approach for a case with a single observation time, we discussed how it can be extended to situations with multiple observation times and we applied it to the same simple example in a three-stage decision problem. The improvements achieved were more modest, which suggests that, with more measurements, the CLRM framework with the conventional optimization approach was able to make up for previous suboptimal decisions, at least for this example. Based on the results of these examples, we conclude that IPO can improve the way we approach the

reservoir management problem, especially in situations where wrong decisions cannot be entirely compensated by the remaining degrees of freedom in the control problem.

An important point that has not been addressed in this work concerns the generation of future synthetic measurements and, in particular, the contribution of the measurement noise. Here we assumed white noise with a pre-defined standard deviation to be added to the simulated measurements, but, in practice, other choices may be more appropriate. Although, in most of the cases, small when compared to the geological uncertainties, this noise contribution may have an impact on the results obtained with the proposed IPO approach. Further research is required for a better understanding of its impact.

Another point for future work concerns the computational costs of such an approach, which are comparable or higher than those of VOI assessment workflows. The IPO ideas only make sense if uncertainty quantification is addressed carefully, and this may require a prohibitive amount of reservoir simulations. In order to be useful, the IPO needs to be tractable. In Chapter 4 we managed to reduce the number of reservoir simulations required in VOI workflows by selecting representative models, a solution that can also be applied here. Besides that, the use of proxy models or multiscale methods may help to further reduce the number of high-fidelity reservoir simulations needed and make the IPO procedure more practical. Still regarding the tractability of the problem, the procedure we proposed here to extend the IPO approach to cases with multiple observation times represents an approximation of the full multistage stochastic model. There may be other solutions to prune the decision tree and effectively reduce the problem to fewer relevant scenarios, as, e.g., suggested by Gupta and Grossmann (2011) and Tarhan et al. (2013). Therefore, also in this direction there is scope for future research.

6

Informed field development optimization

In this chapter [1,2] we present a methodology for field development optimization where we include in the well-location optimization procedure the added value of future information.

The optimization considered in current field development workflows does not explicitly account for the availability of future information resulting from drilling new wells. Here we combine (closed-loop) field development and value of information to create a novel approach, which we refer to as informed field development optimization (IFDO). This method makes use of a nested well-location optimization to quantify not only how much the decision of interest contributes to the selected cost function but also how this decision affects uncertainty in subsequent optimization runs (i.e., to determine the drilling locations of future wells).

We illustrate IFDO by means of two relatively simple examples where we assume that perfect information is obtained after the drilling of the first wells and we observe increases of 1 to 6 % in the cost function values obtained in comparison with the conventional optimization approach. We believe our proposed methodology represents an improvement in the state-of-the-art approach to field development. Due to its computational cost, we recommend that it be used to inspire future research on this topic and the development of more practical tools to be applied in real-field cases.

[1] This chapter is based on work done during an internship at IBM T.J. Watson Research Center.

[2] The work presented in this chapter is covered under the following Spanish patent:

Embid Droz, S.M., Rodríguez Torrado, R., Echeverría Ciaurri, D. and Barros, E.G.D. (2017). Method for determining a drilling plan for a plurality of new wells in a reservoir. Spanish Patent ES P201700314, filed March 29, 2017.

6.1. Introduction

The development of an oil field involves a complex process of decision making under uncertainty, with plenty of scope for optimization. In the age of the smart fields, more and more measurements become available throughout the development of this type of fields. Nevertheless, traditional field development optimization workflows do not consider the availability of future information. By performing optimization based on the current state of knowledge only, we are actually not solving the right problem, which means we are possibly making suboptimal decisions. In order to make better decisions, we need ways of accounting for the added value of future information within our optimization workflows.

In Chapter 2, we presented a workflow to assess the value of information (VOI) in a context where decisions regarding the well controls are supported by reservoir simulation models updated by measurements such as production or time-lapse seismic data (e.g., closed-loop reservoir management). Shirangi and Durlofsky (2015) have suggested to optimize both well placement and controls and update the reservoir simulation models with data collected along the development of the field (e.g., well drilling and production data). However, the optimization considered in these works does not consider the availability of future information.

In this chapter, we propose to combine (closed-loop) field development and VOI assessment to create a novel workflow, which we refer to as informed field development optimization (IFDO). This workflow uses a nested optimization approach to ensure that the decision of interest is optimal in terms of both its direct and indirect contributions to create value (e.g., actual oil production and additional knowledge to facilitate future development decisions, respectively). We illustrate this new approach with two case studies. The first one is based on a two-dimensional reservoir model with uncertain permeability and porosity fields, where we optimize the location of two wells considering the availability of future perfect information. The second example is built upon a larger two-dimensional reservoir model comprising flow barriers with uncertain transmissibility factors, for which we optimize the location of six wells considering the availability of perfect information. In order to quantify the added value of IFDO, we compare the results to those obtained using existing methodologies.

This chapter is structured as follows. In the Previous work section (6.2) we introduce the main concepts that inspire our proposed methodology and revisit previous work on decision making under uncertainty. Thereafter, in the Methodology section (6.3), we present the IFDO workflow, showing how it builds upon the VOI assessment framework introduced in Chapter 2. In the Examples section (6.4), we illustrate it with two simple case studies, explaining in detail how the IFDO methods can be put into practice and comparing the results to those obtained with current state-of-the-art in field development optimization. Finally, in the Discussion and conclusion section (6.5), we analyze the improvements trying to understand where IFDO works the best, and we suggest directions for future work.

6.2. Previous work

6.2.1. Field development optimization under uncertainty

One of the most important decisions during the field development planning (FDP) phase of an asset concerns the placement of the wells to be drilled. Because this usually involves a considerable number of wells and complex reservoir response, there is significant room for optimization. In a conventional setting, the control variables (e.g., well locations) are optimized simultaneously in an open-loop fashion, based on the initial state of uncertainty (e.g., a set of realizations for some geological parameters). This direct approach for optimization under geological uncertainties is the so called robust optimization (section 1.2.1). Robust life-cycle optimization uses one or more ensembles of geological realizations (reservoir models) to account for uncertainties and to determine the development strategy that maximizes a given objective function over the ensemble. For recent applications of robust optimization to field development, see Wang et al. (2012), Leeuwenburgh et al. (2016), Jesmani et al. (2016) and Hanea et al. (2016).

The idea is the same as the one described in section 1.2.1: although the optimization is based on an ensemble of models, only a single strategy **u** is obtained, under the rationale that only one strategy can be implemented in reality. In the context of field development optimization, typical elements of **u** are well locations, trajectories, types, drilling sequence and times, sometimes also jointly with well operating points (i.e., settings of well head

pressures, water injection rates etc.) (Bellout et al., 2012). However, this approach does not make use of the fact that more knowledge becomes available as the wells are drilled, logged and tested, i.e. it is assumed that the quantification of uncertainty does not change over time.

6.2.2. Closed-loop field development (CLFD)

An alternative to the suboptimal open-loop solution is to make use of the data which become available during the development of the field, by performing a closed-loop field development (CLFD) as, for example, in Shirangi and Durlofsky (2015). This means that, when new data are gathered, the simulation models can be adjusted to honor the acquired data. The subsequent decisions (e.g., locations of wells yet to be drilled) are then re-optimized based on the updated knowledge (see Figure 1.1). The claim is that, by "closing the loop" frequently, improved field development can be achieved, when compared to the traditional open-loop optimization approach. Shirangi and Durlofsky (2015) illustrate their CLFD workflow with a few synthetic examples, in which they generate synthetic data from a model considered to represent the truth. They also suggest a way of evaluating the improvements due to the closed-loop approach: by applying the closed-loop and open-loop optimal solutions to the truth model used to generate the synthetic data, and comparing their performances in terms of the objective function used. The same framework is also used to determine optimal well controls to maximize NPV or recovery over the producing life of an asset for a given configuration of wells (CLRM in section 1.1), although we note that early definitions of CLRM also include the possibility to optimize well locations (Jansen et al., 2005). It is important to note that the optimization underlying CLFD is based on current knowledge only, even though it is known that new data will become available in the future.

6.2.3. Time-dependent uncertainty

As mentioned above, the main limitation of the current approach for optimal field development under uncertainty has to do with the assumption that the uncertainties are static. For cases where we know that additional measurements will be gathered throughout the development of the field, we need to consider how the quantification of uncertainty will change and this will allow us to make better decisions. The concept of accounting for the availability of future information within the optimization is not new and has been

investigated in different scientific communities. In operations research, such an optimization problem is addressed by means of stochastic programming. Two-stage or multistage stochastic models are used to deal with the sequential nature of the decisions, and this yields optimization problems expressed through a nested formulation. For more information, see, for example, Birge and Louveaux (1997) and Ruszczyński and Shapiro (2003). In the systems and control community, the dual control introduced by Fel'dbaum (1960 and 1961) seeks to determine the optimal trade-off between excitation and control to promote a more active learning from the measurements which, in turn, may enable the remaining controls to direct the system to its optimal state. This is only possible with the definition of control policies that are able to anticipate the learning effect of future measurements and use it to improve the performance of a closed-loop system (Foss and Jensen, 2011). More recently, Hanssen (2017) proposed an implicit dual model predictive control (MPC) that explicitly includes the feedback mechanism from the measurements in the optimization problem.

These concepts that leverage availability of forthcoming information have been applied in logistics and supply chain, general planning and scheduling, power grids and medicine. In the oil and gas upstream sector, Jonsbråten (1998) considered stochastic programming for drilling sequence optimization in a simplified setting. Goel and Grossmann (2004) used stochastic models in the planning of offshore gas fields with uncertainty related to the reserves, but without including reservoir simulation.

More related to what we address in this chapter, Cunningham and Begg (2008) proposed to use VOI to define the optimal well drilling order. Güyagüler and Horne (2004) investigated how to propagate the effect of geological uncertainty to future well locations. Tarhan et al. (2013) investigated strategies to efficiently solve problems with decision-dependent uncertainty, also for cases with nonlinearities in which the uncertainties are gradually resolved as decisions are made. In order to account for time-dependent uncertainty, Özdogan and Horne (2006) introduced the concept of pseudohistory to estimate the outcome of future measurements and thus incorporate data assimilation in the optimization to determine the well locations that maximize the expected utility of the field production. Lyons and Nasrabadi (2013) proposed to reduce the computational costs of Özdogan and Horne's procedure by relying on the mean of the history matched

ensemble as a single "best estimate" model to evaluate the objective function instead of simulating the entire ensemble to compute the expected utility.

6.2.4. Value of information

The problem of assessing the value of information (VOI) associated with future observations has been receiving noticiable attention in the Earth Sciences in recent years. Due to our usually limited knowledge of the subsurface, one of the main challenges in this assessment consists in determining which measurements are the most useful for the decisions that have to be made. In a situation where new measurements incur additional costs, it is very important to quantify the value of potential measurements before investing in their deployment. Many studies in different contexts have focused on developing methods to make an a-priori assessment of the VOI, see, e.g., Bratvold et al. (2009), Bhattacharjya et al. (2010), Trainor-Guitton et al. (2013) and Eidsvik et al. (2015).

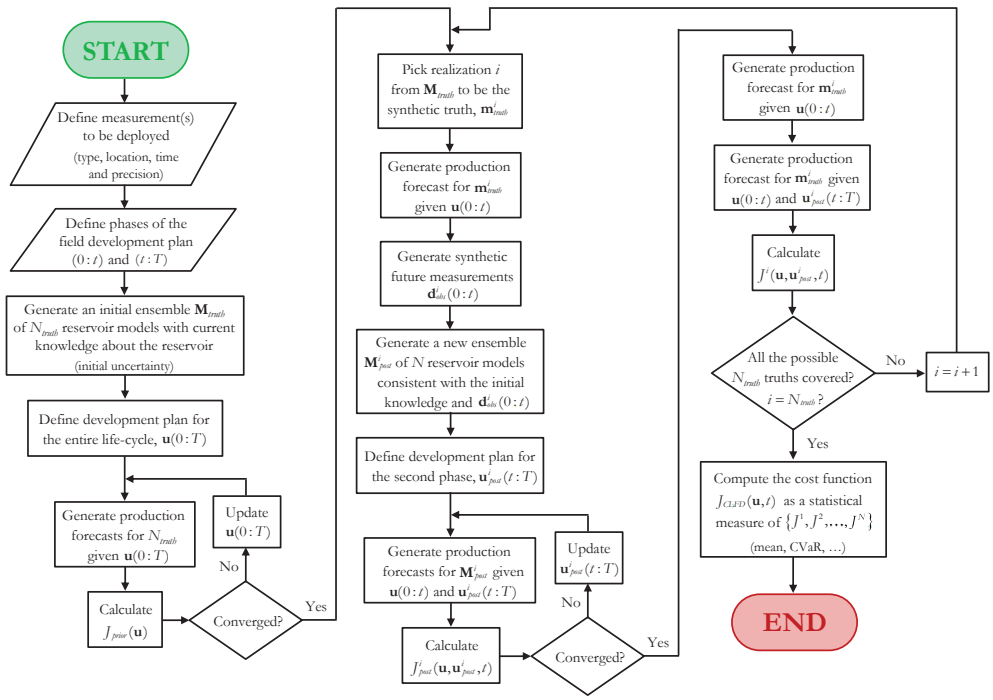


Figure 6.1: Workflow for VOI assessment considering multiple plausible truths in CLFD.

In Chapter 2, we proposed a new methodology to assess the VOI of future measurements by leveraging the CLRM framework. Our approach is based on "closing the loop" in the

design phase to simulate how information obtained during the producing life-time of the reservoir is used in the context of optimal reservoir management. A new production strategy is obtained every time the models are updated with new information, and the strategies with and without additional information can be compared in terms of the value of the optimization objective function (typically NPV) obtained when applying these strategies to a virtual asset (a synthetic truth). The VOI of future measurements is defined as the additional value realized by the asset when the future information is considered for the optimization of the subsequent well controls. This methodology for VOI assessment (Chapter 2) can be adapted to the field development problem (i.e., in the context of CLFD), resulting in the procedure depicted in Figure 6.1 which shows the repetition of the CLFD exercise for an ensemble of plausible truths to quantify the VOI of the analyzed future measurements.

6.3. Methodology

In field development planning, the decisions in question concern the implementation of an optimized development plan which maximizes the value of the asset. Field development plans (FDP) include a variety of design variables, but here we focus on the well locations. Restricting the development to cases with vertical wells only, the development strategy $\mathbf{u} = [\mathbf{u}_1^T \ \mathbf{u}_2^T \ \dots \ \mathbf{u}_M^T]^T$ comprises the list of well locations (e.g., areal coordinates of well head positions) of the M wells to be drilled throughout the entire field life-cycle T (typically tens of years). As new wells are drilled and the field starts to be produced, our knowledge about the reservoir changes. Thus determining a single deterministic development plan by performing robust optimization based on the initial state of uncertainty is equivalent to ignoring the flexibilities that we have to react on the realization of uncertainties, which may lead to suboptimal decisions.

Here we investigate if the stochastic programming ideas discussed in the previous section can be a solution to overcome this limitation within the CLFD framework. The proposed procedure is a combination of classical stochastic programming and our previously introduced workflow for VOI assessment in CLRM (Chapter 2). We model the field development decision making process with the help of elements of the closed-loop framework (i.e., ensemble-based uncertainty quantification, model-based optimization and

computer-assisted history matching), which allows us to represent the sequential character of the decisions while accounting for future information. For that, we refer to this approach as informed field development optimization (IFDO).

The methodology for VOI assessment (Chapter 2) provides a framework to quantify the value J^i to be produced by the asset (here, the i^{th} plausible truth) with the incorporation of future information gathered through the designed surveillance strategy (here, the measurements collected until drilling the j^{th} well) in a CLFD setting (Figure 6.1). For simplicity, we address the case in which the development is carried out in two phases: before and after drilling the j^{th} well.

In the direct approach for robust optimization, we seek to maximize the predicted objective function for the ensemble of models with the hope that, once more information becomes available, we will have the opportunity to improve our predictive models and adjust our strategies to achieve the best possible J^i. In this way the optimization benefits from the flexibility to adjust the remaining control variables $\mathbf{u}^i_{post}(t_{j+1}:T)=\left[\mathbf{u}^i_{post,j+1}{}^T\ \mathbf{u}^i_{post,j+2}{}^T\ \ldots\ \mathbf{u}^i_{post,M}{}^T\right]^T$, while the part of the development plan prior to the future information $\mathbf{u}_{prior}(0:t_j)=\left[\mathbf{u}_{prior,1}{}^T\ \mathbf{u}_{prior,2}{}^T\ \ldots\ \mathbf{u}_{prior,j}{}^T\right]^T$ is fixed and consistent with the initial state of uncertainty. We propose here to use J^i as the cost function for our field development optimization problem, in order to quantify the true performance that we expect to maximize when optimizing $\mathbf{u}_{prior}(0:t_j)$. In fact, since we consider an ensemble of N_{truth} equiprobable plausible truths, we still perform robust optimization and our new objective function is defined as

$$J_{IFDO}\left(\mathbf{u}(0:t_j),t_j\right)=\frac{1}{N_{truth}}\sum_{i=1}^{N_{truth}}J^i\left(\mathbf{u}(0:t_j),\mathbf{u}^i_{post}(t_{j+1}:T),t_j\right), \tag{6.1}$$

where J_{IFDO} is a statistical measure (e.g., the mean in (6.1) and (6.2), or the conditional value at risk (CVaR) introduced by Rockafellar and Uryasev (2000) and also used in Siraj et al. (2016)) of the ensemble of the objective function values J^i for the individual plausible truths. The objective function J^i for the i^{th} plausible truth is calculated according to the workflow in Figure 6.1 and involves the solution of the parameter estimation problem given the future information to be gathered until drilling the j^{th} well and re-optimization of the well locations of the subsequent wells $\mathbf{u}^i_{post}(t_{j+1}:T)$. We can also see the optimization of this new cost function as a two-stage stochastic model or a nested optimization

problem where the outer optimization concerns the well locations of all the wells until the j^{th} one (the decision of interest now) and the inner optimization determines the locations of the future wells, which will be different for each one of the N_{truth} plausible truths (Figure 6.2). Thus, the IFDO problem can be formulated as

$$\mathbf{u}_{IFDO}(0:t_j) = \arg\max_{\mathbf{u}} \frac{1}{N_{truth}} \sum_{i=1}^{N_{truth}} J^i\left(\mathbf{u}(0:t_j), \mathbf{u}^i_{post}(t_{j+1}:T), t_j\right), \tag{6.2}$$

where the outcome of the optimization is a single optimal development plan $\mathbf{u}_{IFDO}(0:t_j)$ until the drilling time of the j^{th} well t_j, and N_{truth} optimal plans $\mathbf{u}^i_{post}(t_{j+1}:T)$, $i = 1, \ldots, N_{truth}$, for the rest of the life-cycle.

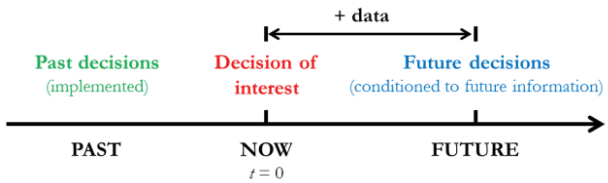


Figure 6.2: Timeline for IFDO: past decisions, decisions of interest and future decisions.

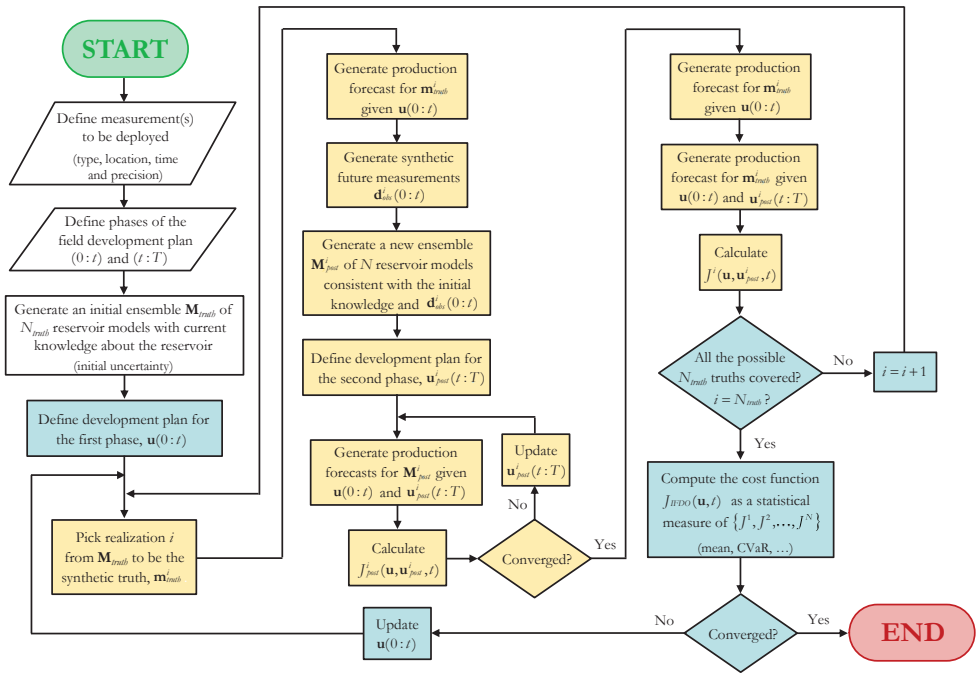


Figure 6.3: Workflow for IFDO for a field development in two phases.

The proposed workflow to solve this optimization problem iteratively is displayed in Figure 6.3. The procedure resembles the workflow of Figure 6.1, but contains an outer iterative loop (shown in blue) to keep updating the first phase of the development plan.

6.4. Examples

6.4.1. Simple 2D model

To test our approach, we used a simple two-dimensional (2D) reservoir simulation model to be developed by water flooding. The reservoir (1,000 × 1,000 × 10 m) is an oil-water system with a represented 2D Cartesian discretization (40 × 40 × 1 grid blocks). The permeability and porosity fields are heterogeneous and present channels whose geometry and position are assumed to be uncertain. The plan is to drill four wells to produce this reservoir over a period of T = 3,000 days, with two producers (denoted as P1 and P2) and two injectors (denoted as I1 and I2). Figure 6.4 shows the reservoir considered in this example and also includes some of the model realizations used together with the training image needed to generate them (through multi-point statistics with SGeMS; see Strebelle, 2002). Table 6.1 lists other parameters used in the example. P1 and I1 are drilled simultaneously in (previously) optimized locations, and the other two wells have to be drilled sequentially after them. The field starts production after P2 is drilled and I2 comes on stream 300 days later. We are interested in optimizing the locations of P2 and I2. The wells are operated by constant bottom-hole pressure (BHP) with reactive control (i.e., the producers are shut in once they reach uneconomical water-cuts). The dashed line delimits the active set for the well placement optimization: we remove the possibility of placing wells close to the edge of the reservoir to make the optimization problem more interesting.

The production forecasts used to compute the NPV were generated by simulation performed with the AD-GPRS reservoir simulator (Zhou, 2012). We used an open-source implementation of the particle swarm optimization (PSO) algorithm available for Matlab. The PSO settings (i.e., swarm size, number of iterations, etc.) were chosen according to the number of optimization variables for each case.

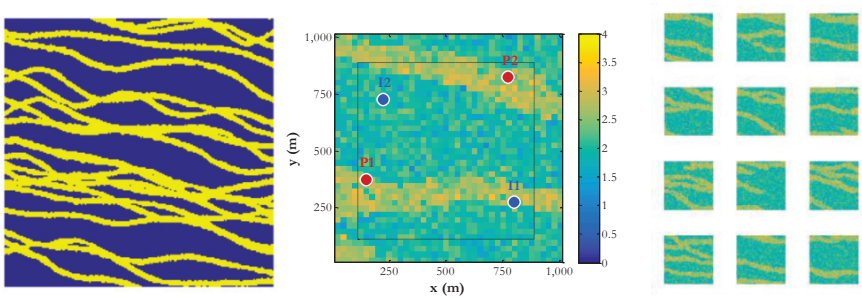


Figure 6.4: Simple 2D example: training image used to generate realizations (left), one realization of the permeability field with illustrative well locations (middle) and 12 realizations of the permeability field (right).

Table 6.1: Parameter values for the 2D model.

Rock-fluid parameters		Initial conditions	
$\varrho_o = 800$	kg/m^3	$p_0 = 300$	bar
$\varrho_w = 1{,}000$	kg/m^3	$S_{oi} = 0.9$	[–]
$\mu_o = 2$	cP	$S_{wi} = 0.1$	[–]
$\mu_w = 1$	cP		
$B_o = 1$	[–]		
$c_w = 10^{-5}$	bar^{-1}		
$c_f = 10^{-5}$	bar^{-1}		
$n_o = 2$	[–]	Economic parameters	
$S_{or} = 0.2$	[–]	$r_o = 50$	\$/bbl
$k_{ro,wc} = 0.8$	[–]	$r_{wp} = 15$	\$/bbl
$n_w = 2$	[–]	$r_{wi} = 15$	\$/bbl
$S_{wc} = 0.2$	[–]	$b = 0$	%/year
$k_{rw,or} = 0.75$	[–]		

We note that, although the location of I2 was also to be optimized, the decision of immediate interest in this example concerned the location of P2. With that in mind, we applied the various approaches presented in the previous section, which are expected, in principle, to yield different solutions. We used $N_{truth} = 30$ model realizations honoring the data obtained after drilling P1 and I1. These $N_{truth} = 30$ realizations together formed the $\mathbf{M}_{truth}$ ensemble of plausible truths, which is the starting point of our approach. The goal was to optimize the location of P2 and I2 given the uncertainty described by $\mathbf{M}_{truth}$ to

maximize the conditional value at 25 % risk ($CVaR_{25}$) of the NPV of the realizations. $CVaR_{25}$ is computed as the average of the NPV of the samples below the $P_{(100-25)} = P_{75}$ percentile of the NPV distribution (here, P_x is defined as the probability that x % of the outcomes exceeds this value). This optimization problem had four variables (i.e., two sets of areal coordinates for P2 and I2).

For this example, we considered the situation where we deal with perfect information ("clairvoyance") after knowing the location of P2. This consideration simplifies the computation procedures (Figure 6.1 and Figure 6.3) because, with perfect revelation of the truth, the solution of the parameter estimation problem is known and deterministic (after defining the location of P2, we assume to automatically know which plausible truth is the actual truth). This also means that the inner-level optimizations (to determine the location of I2) are deterministic, which reduces the computational cost of the whole procedure. Moreover, as we showed in Chapter 2, the assumption of clairvoyance represents the "technical limit" for any case with future imperfect information, serving as an upper bound for the analysis.

We then considered the following cases to illustrate the value of our proposed methodology:

(a) Open-loop optimization. We optimize the location of P2 and I2 under uncertainty, in a robust open-loop fashion. We simply seek the best placement given the initial state of uncertainty (we do not consider the fact that clairvoyance will become available after drilling P2).

(b) CLFD with clairvoyance. We first optimize the location of P2 and I2 under uncertainty, in a robust open-loop fashion. Then, clairvoyance becomes available after drilling P2 and we take advantage of that by re-optimizing the location of I2 knowing in all cases the plausible truth for the spatial distribution of permeability and porosity.

(c) IFDO with clairvoyance. We optimize the location of P2 and I2 under uncertainty, but considering in the optimization that clairvoyance will become available after drilling P2, which then allows us to optimize the location of I2 having access in all cases to the spatial distribution of permeability and porosity of the plausible truth.

(d) Prior clairvoyance. We include the utopic case where there is no geological uncertainty; this allows us to optimize the location of P2 and I2 having access in all cases to the spatial distribution of permeability and porosity of the plausible truth.

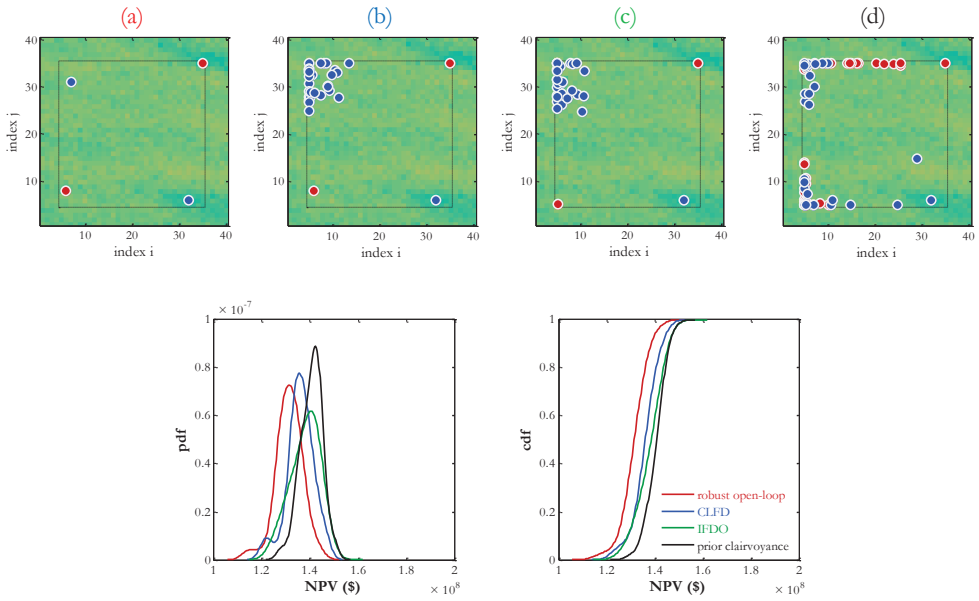


Figure 6.5: Well locations obtained for the 2D model: (a) robust open-loop optimization, (b) CLFD with clairvoyance for multiple plausible truths, (c) IFDO with clairvoyance and (d) optimization under prior clairvoyance. At the bottom, pdf and cdf distribution for the different cases.

Figure 6.5 depicts the results we obtained for this 2D model. The maps on the top show the optimal locations of the wells obtained with the different approaches, displayed on top of the mean permeability field of the plausible truths considered. Figure 6.5 (bottom) also shows the probability density function (pdf) and cumulative distribution function (cdf) corresponding to the solutions found. We notice that the proposed IFDO approach results in higher NPV values than those obtained with the conventional robust open-loop optimization. We also observe that the CLFD is able to improve the development plan by re-optimizing the location of I2 once the uncertainty is revealed. These results confirm that, as expected, the more operational flexibility is included in the optimization, the more value can be obtained with additional knowledge (clairvoyance in this case). Here, the flexibility lies in the fact that the location of I2 does not have to be determined immediately, allowing us to consider a variety of scenarios for this decision.

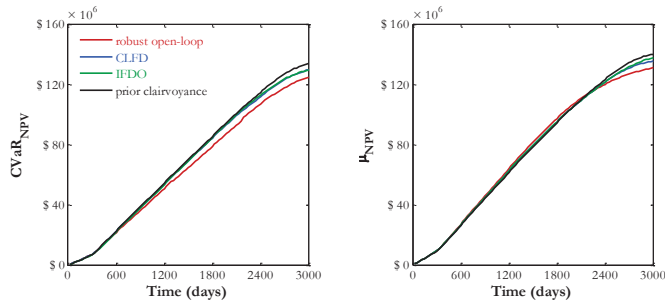


Figure 6.6: Results for the 2D model: $CVaR_{25}$ (left) and mean (right) of NPV for the different cases.

Table 6.2: Summary of results for the 2D model.			
Robust open-loop	CLFD with clairvoyance	IFDO with clairvoyance	Prior clairvoyance
$\mathbf{u}_{P2}$ = (6,8)	$\mathbf{u}_{P2}$ = (6,8)	$\mathbf{u}_{P2}$ = (5,5)	$\mathbf{u}_{P2}$ = multiple
$\mathbf{u}_{I2}$ = (7,31)	$\mathbf{u}_{I2}$ = multiple	$\mathbf{u}_{I2}$ = multiple	$\mathbf{u}_{I2}$ = multiple
$CVaR_{25}$ = 125.2 ×10⁶ $	$CVaR_{25}$ = 129.8 ×10⁶ $	$CVaR_{25}$ = 129.6 ×10⁶ $	$CVaR_{25}$ = 134.3 ×10⁶ $
	$\Delta CVaR_{25}$ = +3.7 %	$\Delta CVaR_{25}$ = +3.5 %	$\Delta CVaR_{25}$ = +7.3 %
μ = 131.3 ×10⁶ $	μ = 135.9 ×10⁶ $	μ = 137.9 ×10⁶ $	μ = 140.4 ×10⁶ $
	$\Delta\mu$ = +3.5 %	$\Delta\mu$ = +5.0 %	$\Delta\mu$ = +6.9 %

Figure 6.6 displays the results in terms of the $CVaR_{25}$ and the mean of cumulative NPV over the field life-cycle for all the approaches, and Table 6.2 shows the same results but for the final cumulative NPV only. We notice that, although the CLFD and the IFDO approaches result in virtually the same $CVaR_{25}$ values, the latter produces higher mean values, which is reflected in the pdf and cdf distributions observed in Figure 6.5. For this example, the gain obtained with the IFDO approach is of approximately $ 4.5 and $ 6.7 million for the $CVaR_{25}$ and mean values respectively, an increase of approximately +3.5 % and +5.0 % in comparison to conventional robust optimization. We also observe that, while conventional robust optimization indicates the optimal location of P2 to be at (6,8) (in grid block coordinates), IFDO finds the optimum to be at (5,5). These two locations are about 80 m apart from each other. Although small, this change seems to have an impact on the value that the asset can produce.

Although the performance of the CLFD approach is slightly inferior, the number of reservoir simulations required is significantly smaller. Besides that, the output of the

CLFD procedure also consists of a single location for P2 and an ensemble of well locations for I2. Thus, the CLFD solution for multiple plausible truths can be seen as an approximation for IFDO and a means of deriving flexible field development plans. This can be useful for cases where we cannot afford the amount of simulations required by IFDO.

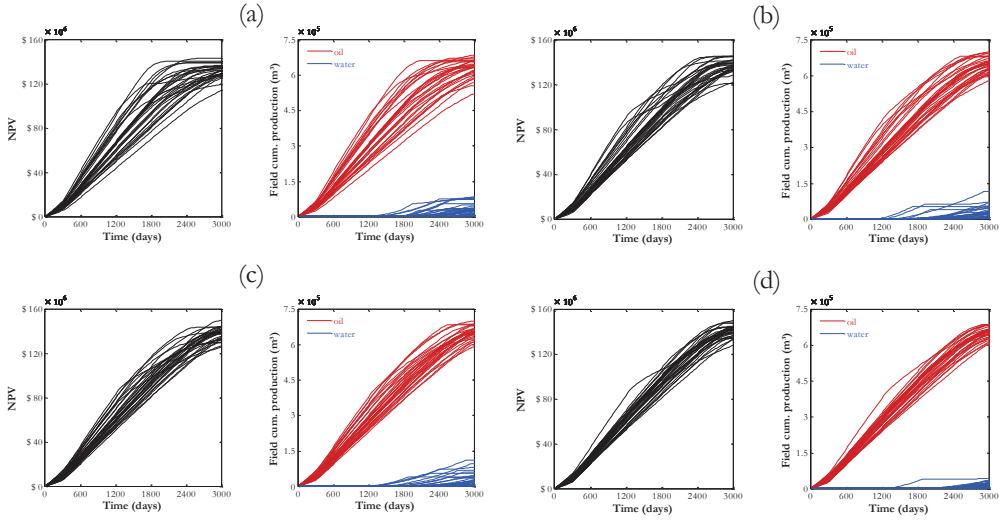


Figure 6.7: NPV and field cumulative production for the 2D model: (a) robust open-loop optimization, (b) CLFD with clairvoyance for multiple plausible truths, (c) IFDO with clairvoyance and (d) optimization under prior clairvoyance.

As an attempt to explain the improvements achieved by the IFDO, we can look at the cash flow and production profiles obtained for the plausible truths by following the different approaches. Figure 6.7 depicts these results. We observe a larger spread in the cash flow and oil production curves for the robust open-loop optimization (Figure 6.7 (a)). We see for the same case that for some of the plausible truths the optimal strategy turns out to be an aggressive one resulting in uneconomical production prematurely, with a maximum NPV before T = 3,000 days. The CLFD approach (Figure 6.7 (b)) seems to be successful in compensating the effect of the wrongly chosen for P2 location, and extending the economic life of the field production by adjusting the location of I2. But still we observe (based on the water production curves) that certain producers are shut-in by reactive control quite early in time. By proposing a slightly different location for P2, IFDO (Figure 6.7 (c)) is able to postpone the shut-in times, making the field operate at

economic levels for a longer period. The optimization under prior clairvoyance (Figure 6.7 (d)) performs even better by postponing the water breakthrough times for most of the plausible truths and reducing even further water production while boosting oil production and reducing the uncertainties (i.e., smaller spreads for the NPV and oil production curves).

6.4.2. Bean model with faults

As a second case study, we tested the different approaches in a synthetic field. This 2D reservoir model is named the "Bean model" due to the characteristic shape formed by its 3,398 active cells. The reservoir (approximately 2,325 × 1,800 × 10 m) is an oil-water horizontal system with heterogeneous permeability and porosity fields based on a two-dimensional Cartesian discretization (93 × 72 × 1 grid blocks) of a channelized formation. The permeability and porosity fields are assumed to be known. The field is split in 4 sectors by 3 faults, whose sealing capacities are unknown: they can be perfectly sealing or perfectly permeable. Given the binary character of the uncertainty on the sealing capacity of the 3 faults, we have $2^3 = 8$ possible realizations, which are assumed to be equiprobable. The sealing faults are modeled as flow barriers by modifying the transmissibility multipliers of the neighboring grid blocks. Figure 6.8 shows the Bean model with faults. The colorful cells (i.e., not the faded ones close to the fault lines or the edge of the reservoir) represent the active set for the well placement optimization: we avoid placing wells close to the flow barriers and to the edge of the reservoir to make the problem more realistic.

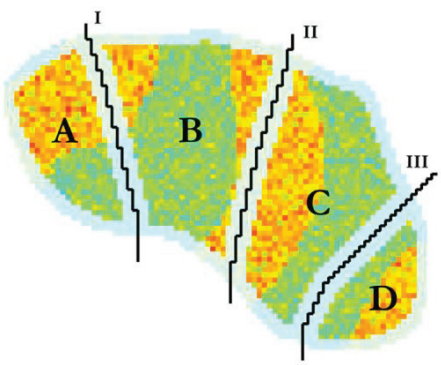


Figure 6.8: Bean model with faults: known permeability and porosity fields and split in 4 sectors by 3 faults with unknown sealing capacity.

Table 6.3: Parameter values for the Bean model.

Rock-fluid parameters		Initial conditions	
$\varrho_o = 800$	kg/m^3	$p_0 = 300$	bar
$\varrho_w = 1{,}000$	kg/m^3	$S_{oi} = 0.9$	[–]
$\mu_o = 5$	cP	$S_{wi} = 0.1$	[–]
$\mu_w = 1$	cP		
$B_o = 1$	[–]		
$c_w = 10^{-5}$	bar^{-1}		
$c_f = 0$	bar^{-1}		
$n_o = 2$	[–]	Economic parameters	
$S_{or} = 0.2$	[–]	$r_o = 300$	\$/m^3
$k_{ro,wc} = 0.8$	[–]	$r_{wp} = 30$	\$/m^3
$n_w = 2$	[–]	$r_{wi} = 15$	\$/m^3
$S_{wc} = 0.2$	[–]	$b = 0$	%/year
$k_{rw,or} = 0.75$	[–]		

This green field is to be developed by water flooding with 3 producers (P1, P2 and P3) to be drilled and to start production at $t = 0$ and 3 injectors (I1, I2 and I3) to be drilled and to start injection after one year ($t = 360$ days). The producers are operated by constant BHP with reactive control and the injectors by prescribed injection rates with a fixed upper limit for the injection pressure. Table 6.3 lists other parameters used in the example. The idea here was to create an example with features that could change dramatically the way well locations are planned, and to engineer a case where an initial wrong decision could hardly be compensated by future decisions. In this example the decision of immediate interest concerns the locations of the producers, and the flexibility available is the opportunity of fixing the locations of the injectors in the future. The $N_{truth} = 8$ model realizations representing the initial uncertainty form the $\mathbf{M}_{truth}$ ensemble of plausible truths, which is the starting point of our approach. The goal was to optimize the location of all the wells given the uncertainty described by $\mathbf{M}_{truth}$ to maximize the $CVaR_{25}$ of the NPV of the realizations. This optimization problem has 12 variables (6 sets of areal coordinates for the 3 producers and the 3 injectors). Like for the 2D model, the production forecasts were generated with the AD-GPRS reservoir simulator. The

optimizer was also the same PSO implementation used in the previous example. In order to show the value of IFDO we considered the cases below:

(a) Open-loop optimization. We optimize the locations of all the wells under uncertainty, in a robust open-loop fashion. We simply seek the best placement given the initial state of uncertainty (we do not consider the fact that clairvoyance will become available after drilling P1, P2 and P3).

(b) CLFD with clairvoyance. We first optimize the locations of all the wells under uncertainty, in a robust open-loop fashion. Then, clairvoyance becomes available after drilling the producers and we take advantage of that by re-optimizing the locations of the injectors knowing in all plausible truth cases whether the faults are sealing or not.

(c) IFDO with clairvoyance. We optimize the locations of all the wells under uncertainty, but considering in the optimization that clairvoyance will become available after drilling the producers, which then allows us to optimize the locations of the injectors knowing in all plausible truth cases whether the faults are sealing or not.

(d) Prior clairvoyance. We include the utopic case where there is no geological uncertainty; this allows us to optimize the locations of all the wells having knowledge of the sealing capacity of the faults.

(e) Engineering solution. For comparison, we also include a likely solution that a reservoir engineer would find to this problem, by placing the wells in 3 producer-injector pairs in high permeable areas on each one of the 3 largest sectors with the producers in the north part of the field and injectors in the south.

Figure 6.9 shows the results obtained for the Bean model with faults. The maps at the top contain the optimal well locations for the different approaches and the bar chart at the bottom shows a comparison of the different solutions for each one of the 8 realizations considered. The first point to note is that the robust open-loop optimization already represents a great improvement (+16.3 % in terms of $CVaR_{25}$ or +17.6 % in terms of mean) to the typical engineering solution and that we test a new approach to improve this already optimized solution. Another point to highlight is that, like for the previous

example, the development plans that embrace flexibility (i.e., with multiple well location scenarios) are expected to deliver more value.

In the search of the 8 CLFD solutions, we identified two different solutions: the $CLFD_1$ solution was optimal for 6 out of the 8 realizations and the $CLFD_2$ solution was the best one for the other 2 realizations. When testing these two CLFD solutions on the entire set of initial realizations, we realized that the $CLFD_1$ solution resulted in a higher $CVaR_{25}$ value than the one achieved with the open-loop approach. This means that the robust open-loop optimization was not able to find the global optimum for this problem, despite the use of PSO which is considered by many to be one of the most efficient global search methods. Therefore, the $CLFD_1$ solution is also included in the analysis of the results (case (f) in Figure 6.9) as the best robust solution found.

Back to Figure 6.9, we observe that the producer to be drilled in sector B has a completely different location in the IFDO (c) solution compared to the location determined with the conventional approach (a) and (b). This change allows the locations of the injectors to be optimized differently resulting in NPV values significantly higher for 4 realizations and slightly lower for 2 realizations. We also notice that the optimizations with initial clairvoyance (d) lead to well locations which produce a lot more value for most of the realizations. Finally, we draw the attention to the fact that realizations 5 and 8 behave as worst case scenarios and that there does not seem to be a lot of room for increasing their NPV compared to the robust open-loop solution.

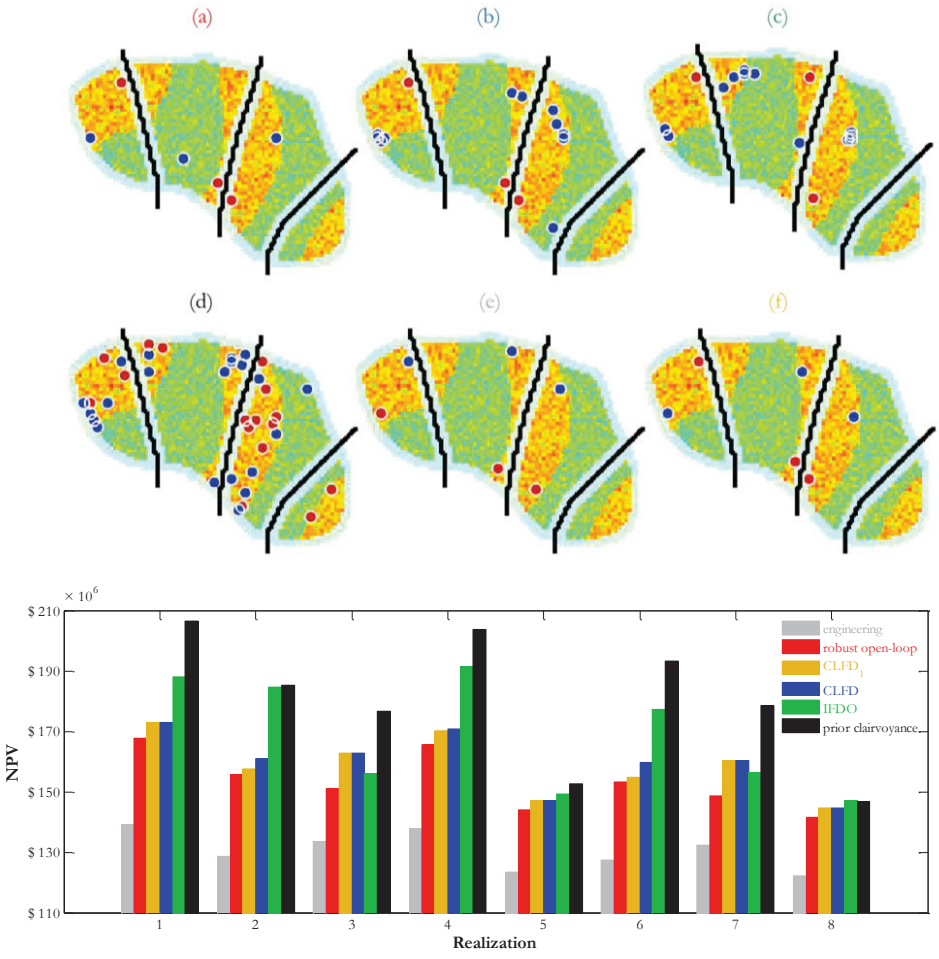


Figure 6.9: Optimized well locations for Bean model and their respective performance in terms of NPV values for the 8 plausible truths considered: (a) robust open-loop optimization, (b) CLFD with clairvoyance for multiple plausible truths, (c) IFDO with clairvoyance and (d) optimization under prior clairvoyance, (e) engineering solution and (f) one of the CLFD solutions which outperformed (a) as robust solution.

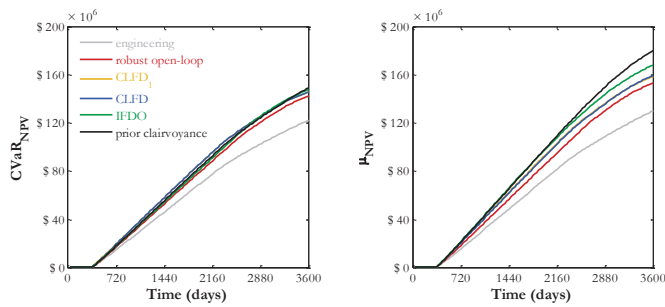


Figure 6.10: Results for the Bean model: $CVaR_{25}$ (left) and mean (right) of NPV for the different cases.

Figure 6.10 displays the same results but in terms of $CVaR_{25}$ and mean values of NPV over time. Except for the engineering solution all the other $CVaR_{25}$ curves are very close to each other, reflecting the presence of the 2 realizations behaving as worst case scenarios. The mean values however reveal more differences between the different solutions because the mean also accounts for the best realizations. This brings us to the discussion on what we should aim for in optimizing cases like this where the uncertainties are of a binary nature. Here we maximize $CVaR_{25}$ of NPV but we could also optimize other statistical metrics of the NPV distribution, such as the mean. In the end we would like to determine development plans that improve entire probability distributions, and we do not seem to be able to achieve that by focusing on a single statistical metric as the cost function.

In order to understand what causes the IFDO strategies to result in more value, we try to diagnose the efficiency of the field water flooding by analyzing the production profiles and the sweep of the reservoir. Figure 6.11 shows the cumulative NPV and field cumulative production curves for all the 8 realizations in the same six cases from Figure 6.9. We note that the cases with higher NPV values correspond to the cases with larger oil production volumes, meaning that the produced water has a smaller effect given the reservoir life-cycle of $T = 3{,}600$ days. We also observe that the solutions for IFDO (c) and prior clairvoyance (d) perform the best and exhibit the largest spreads for the oil production and NPV curves. This is the opposite to what was observed in the 2D example (section 6.4.1). In Figure 6.12 we compare the final water saturation distributions (at $t = 3{,}600$ days) for all the cases. We can see that the approaches considering more operational flexibility with multiple well locations result in larger flooded areas, which means that larger volumes are being produced.

Table 6.4 summarizes all the results obtained for the Bean model example. It shows the well locations optimized for the different cases and it compares the final cumulative NPV obtained for the plausible truths expressed in $CVaR_{25}$ and mean values. If we consider the $CLFD_1$ solution as the baseline for comparison, we observe a modest gain of $\Delta CVaR_{CLFD1} = +1.3\,\%$ with IFDO. The improvement is $\Delta\mu_{CLFD1} = +6.1\,\%$ if we compare the results in terms of mean NPV values. Table 6.4 contains also the percentage gains when the baseline for comparison is the engineering solution or the solution

obtained from the PSO robust open-loop optimization. We also highlight the significant change in location of one of the producers by following the IFDO approach: conventional optimization would place P3 at (50,29) and IFDO at (55,58), more than 700 m far apart. Such a change could have a major impact on decisions regarding the design of surface facilities and the allocation of drilling resources.

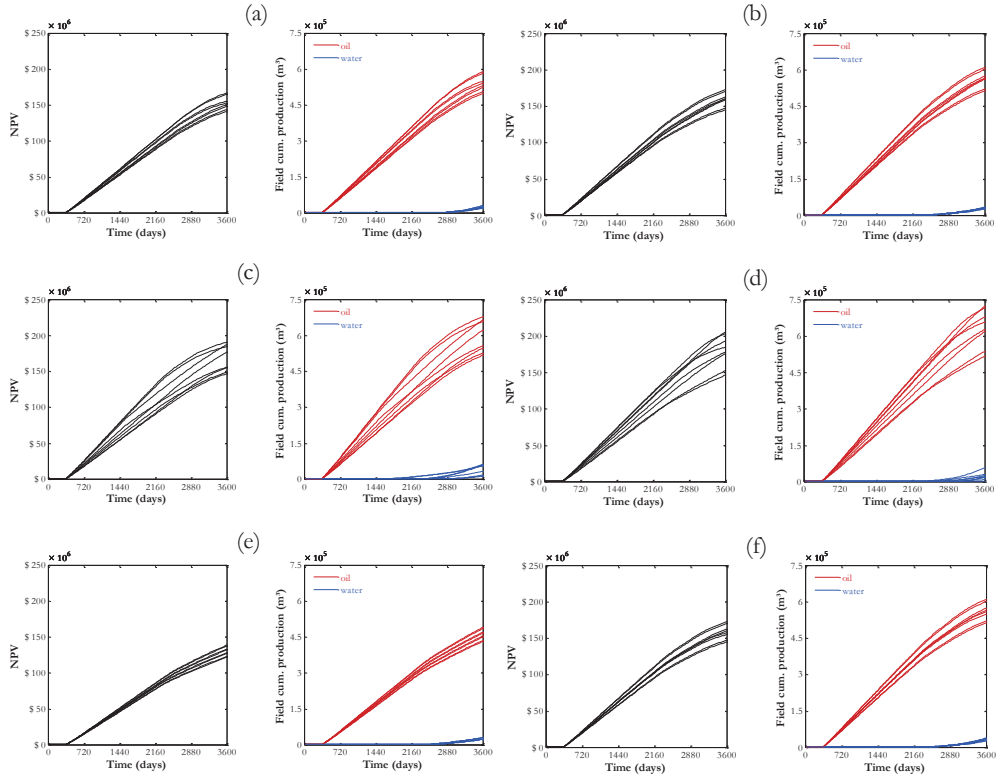


Figure 6.11: NPV and field cumulative production for the Bean model for the same six cases from Figure 6.9.

Based on these results, we started questioning whether the choice of using $CVaR_{25}$ as a cost function was an appropriate choice for this case where clearly the 2 worst realization dominate this statistical metric. We then repeated the same experiments but this time maximizing the mean instead. The new optimizations resulted in very similar solutions for the CLFD and IFDO approaches, meaning that the producers ended up being placed in the same locations for both the conventional robust optimization and our proposed

method. In other words, there was no advantage in pursuing the computationally expensive IFDO when maximizing the mean of NPV for this example.

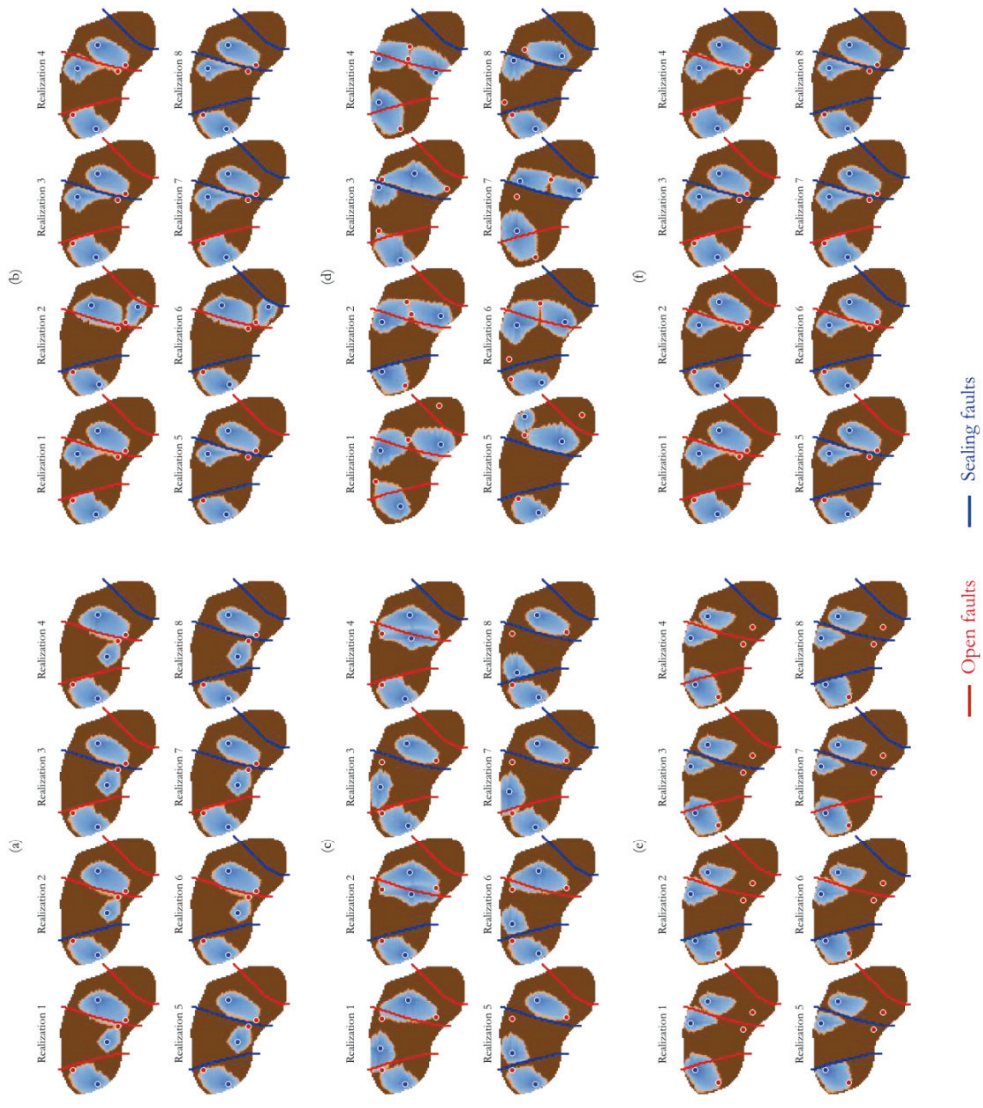


Figure 6.12: Final saturation distributions for the same six cases from Figure 6.9 (Bean model). Permeable faults are indicated in red and sealing faults in blue.

Table 6.4: Summary of results for the Bean model (values in $\times 10^6$ \$).

Engineering	Robust open-loop	$CLFD_1$	CLFD	IFDO	Prior clairvoyance
$\mathbf{u}_{P1}$ = (14,43)	$\mathbf{u}_{P1}$ = (54,24)	$\mathbf{u}_{P1}$ = (54,24)	$\mathbf{u}_{P1}$ = (54,24)	$\mathbf{u}_{P1}$ = (56,23)	$\mathbf{u}_{P1}$ = multiple
$\mathbf{u}_{P2}$ = (48,27)	$\mathbf{u}_{P2}$ = (22,58)	$\mathbf{u}_{P2}$ = (22,58)	$\mathbf{u}_{P2}$ = (22,58)	$\mathbf{u}_{P2}$ = (22,58)	$\mathbf{u}_{P2}$ = multiple
$\mathbf{u}_{P3}$ = (59,21)	$\mathbf{u}_{P3}$ = (50,29)	$\mathbf{u}_{P3}$ = (50,29)	$\mathbf{u}_{P3}$ = (50,29)	$\mathbf{u}_{P3}$ = (55,58)	$\mathbf{u}_{P3}$ = multiple
$\mathbf{u}_{I1}$ = (22,58)	$\mathbf{u}_{I1}$ = (13,42)	$\mathbf{u}_{I1}$ = (52,55)	$\mathbf{u}_{I1}$ = multiple	$\mathbf{u}_{I1}$ = multiple	$\mathbf{u}_{I1}$ = multiple
$\mathbf{u}_{I2}$ = (52,61)	$\mathbf{u}_{I2}$ = (40,36)	$\mathbf{u}_{I2}$ = (13,43)	$\mathbf{u}_{I2}$ = multiple	$\mathbf{u}_{I2}$ = multiple	$\mathbf{u}_{I2}$ = multiple
$\mathbf{u}_{I3}$ = (66,50)	$\mathbf{u}_{I3}$ = (67,42)	$\mathbf{u}_{I3}$ = (67,42)	$\mathbf{u}_{I3}$ = multiple	$\mathbf{u}_{I3}$ = multiple	$\mathbf{u}_{I3}$ = multiple
$CVaR_{25}$ = 122.7	$CVaR_{25}$ = 142.7	$CVaR_{25}$ = 146.2	$CVaR_{25}$ = 146.2	$CVaR_{25}$ = 148.2	$CVaR_{25}$ = 149.7
	$\Delta CVaR_{eng}$ = +16.3 %	$\Delta CVaR_{eng}$ = +19.2 %	$\Delta CVaR_{eng}$ = +19.2 %	$\Delta CVaR_{eng}$ = +20.8 %	$\Delta CVaR_{eng}$ = +22.0 %
		$\Delta CVaR_{rob}$ = +2.5 %	$\Delta CVaR_{rob}$ = +2.5 %	$\Delta CVaR_{rob}$ = +3.9 %	$\Delta CVaR_{rob}$ = +4.9 %
			$\Delta CVaR_{CLFD1}$ = 0.0 %	$\Delta CVaR_{CLFD1}$ = +1.3 %	$\Delta CVaR_{CLFD1}$ = +2.4 %
μ = 130.5	μ = 153.4	μ = 159.1	μ = 159.9	μ = 168.8	μ = 180.4
	$\Delta\mu_{eng}$ = +17.6 %	$\Delta\mu_{eng}$ = +21.9 %	$\Delta\mu_{eng}$ = +22.5 %	$\Delta\mu_{eng}$ = +29.3 %	$\Delta\mu_{eng}$ = +38.3 %
		$\Delta\mu_{rob}$ = +3.7 %	$\Delta\mu_{rob}$ = +4.2 %	$\Delta\mu_{rob}$ = +10.0 %	$\Delta\mu_{rob}$ = +17.6 %
			$\Delta\mu_{CLFD1}$ = +0.5 %	$\Delta\mu_{CLFD1}$ = +6.1 %	$\Delta\mu_{CLFD1}$ = +13.4 %

6.5. Discussion and conclusions

We have identified an opportunity to improve simulation-based field development optimization. We developed a workflow which allows us to design flexible drilling plans that account for the fact that uncertainties are not static throughout the decision making process. We believe this work to be a first attempt in changing the way we approach the field development problem in order to benefit as much as possible from the operational flexibility. Moreover, our proposed procedure provides a framework to generate multiple FDP scenarios which can be helpful to communicate this flexibility. Our methodology makes use of many complex elements (e.g., simulation-based optimization, model inversion, value of information and high-performance computing) that endows the research and the corresponding software implementations with a relatively high level of complexity.

We have tested the IFDO approach in two synthetic case studies with the assumption of availability of future clairvoyance and we obtained improved results compared to conventional optimization, which considers the uncertainties to be static. The additional values were in the range of $ 2 to $ 9 million in terms of $CVaR_{25}$ and mean of NPV, representing gains in the order of 1 % to 6 %. This confirms the potential of the IFDO approach.

In this work, we applied the methodology only to examples considering future clairvoyance. This assumption implies that the revelation of the truth is always perfect and does not depend on the current decisions. We believe IFDO to be even more advantageous for cases with imperfect information where the outcome of future measurements depends on current decisions, leading to a larger impact on the optimization of the subsequent decisions. However, this is yet to be confirmed by applying IFDO to suitable case studies.

One of the biggest challenges of this research is related to the fact that there are no simple benchmark cases truly designed to demonstrate the value of accounting for time-dependent uncertainties. Here we tried to use our intuition and experience to come up with synthetic examples where future information could make a difference while remaining realistic and simple enough to allow a proof of concept within reasonable computation times. The truth is that the decision making process becomes quite elaborate even for cases with just a few well locations to be determined. In practice, we are not certain whether the examples presented here are good candidates to illustrate the advantages of IFDO and an effort is necessary to find an appropriate benchmark case for further research in simulation-based field development optimization under uncertainty.

As recommendations for future work in the same research line, we believe the focus should be in making our methodology more practical by reducing the computational costs. The use of approximations for the reservoir simulation models and at the optimization level will be essential to apply the ideas we propose here to real-field cases. Recently, Shirangi and Durlofsky (2016) presented new ideas on how to select subsets of representative models in a smart way to reduce the computational burden of similar workflows, a topic which we also addressed in Chapter 4. Another possible alternative could be to combine the ideas presented here with the approach of Özdogan and Horne

(2006) which also account for time-dependent uncertainties and look for more tractable solutions which can also deliver flexible well placement strategies. Finally, although conceptually it should not be difficult, there is still work to be done to implement our IFDO approach for other types of decision variables besides well locations (e.g., number of wells, well type, drilling sequence and time).

7

Conclusion

This chapter provides an overview of the conclusions of this research, highlighting the main findings of each chapter, followed by a list of recommendations for future research on topics related to VOI assessment in CLRM.

7.1. Conclusions

Chapter 1

- An a-priori assessment of the value of future measurements is relevant to the design of reservoir surveillance plans. Very few studies have addressed this topic in the reservoir engineering community. Previous works have proposed to quantify the value of observations in terms of their ability to improve our knowledge (i.e., reservoir models and simulation forecasts). In this thesis we have identified the opportunity to take a step further and include the decisions into the assessment, as it is recommended in the field of decision analysis.

Chapter 2

- The CLRM framework can be used for VOI assessment in reservoir management applications. By combining the elements of the CLRM with synthetic truth and data it is possible to "close the loop" in the FDP phase to determine the value of future measurements before deploying them. The data assimilation step serves as a Bayesian environment to quantify the updates in the knowledge (i.e., reservoir characterization, simulation forecasts) based on the future observations while the optimization step constitutes a tool to quantify how the changes in knowledge affect the reservoir management decisions. Finally, the implementation step provides the

opportunity to predict the performance of the decisions with and without the additional information from future measurements when applied to the synthetic truth. Hence, the proposed methodology determines the VOI that accounts for both the information content of the measurements and the information timing with respect to the decisions to be made, which makes it a more complete approach to estimate the VOI in a reservoir engineering context.

- VOI assessment only makes sense in the presence of uncertainties. Because of that, the value of future measurements is uncertain itself: although we control the design of the surveillance plans, we cannot know the outcome of the measurements in advance. The methodology described in this work addresses this point by considering an ensemble of plausible truths, allowing us to obtain samples of the VOI as a random variable.
- The proposed methodology for VOI assessment is generic and flexible:
  - It can be implemented with different algorithms and tools.
  - It can assess the value of measurements of different type (e.g., well data, field-wide sensing data), location, time and frequency, precision and accuracy, etc.
  - It can be useful in a variety of applications of decision making under uncertainty, not only in reservoir management.
- The proposed methodology can also be used to quantify the VOC. Despite being unrealistic, the assumption of clairvoyance provides a way of estimating the technical limit for the achievable added value for any surveillance strategy. Besides that, the computational costs are significantly lower for VOC calculation because there is no need to perform history matching and the optimizations are deterministic once the clairvoyance becomes available. The results obtained confirmed that the VOC constitutes an upper bound for the VOI of imperfect observations. The VOC tends to decrease over time, following the decrease in the number of degrees of freedom or flexibility left to enable value. Moreover, VOC is positive for all the plausible truths, whereas the VOI of imperfect measurements can be negative in some cases.
- As main limitations of our proposed method, we can list:

- It cannot assess the VOI in the presence of "unmodeled" uncertainties (i.e., unknown unknowns). Nevertheless, we believe it to be a limitation of any approach for a priori analysis of future measurements.
- The VOI assessment is susceptible to the limitations of the tools used to implement it. In particular, the VOI obtained can be affected by the inability of the optimizers to guarantee global optimal solutions. This, however, is a drawback of any framework relying on optimization algorithms. In practice, one should use the same tools that would be used to implement the CLRM in an operational setting when assessing the VOI in the FDP phase.
- We have not considered the effect of the decision maker's risk preference in the VOI assessment. Through the entire thesis we have assumed that the decision maker is risk-neutral.
- A rigorous VOI assessment relying on high-fidelity reservoir simulations and a typical number of model realizations can be computationally prohibitive for large-scale applications. An adequate use of approximations is necessary to make the approach tractable in practice.

Chapter 3

- The proposed methodology can be easily extended to assess the VOI for multiple observation times, without a significant increase of computational costs. For that, the "loop is closed" multiple times for each plausible truth; there is no need to consider new plausible truths as the workflow goes through the observation times. In this setting, the cumulative VOI tends to increase over time, as more and more information is gathered.
- The workflow for VOC calculation becomes more complex when considering multiple observation times. Because imperfect information becomes available at every observation time prior to the moment of clairvoyance, the procedure requires the solution of history matching and robust optimization to "close the loop" multiple times. This means that VOC still decreases over time, but at a slower rate, until getting equal to the cumulative VOI of the multiple imperfect measurements at the last observation time. Thus, the trend of VOC over time is no longer the same for

any surveillance plan because it becomes dependent on the imperfect observations considered in the analyzed plan. However, VOC still constitutes an upper bound for VOI at any time.

Chapter 4

- The VOI assessment can be significantly accelerated by reducing the size of the ensembles of model realizations considered in the analysis. Subsets of representative models can be selected with the help of clustering techniques. There are two main opportunities to reduce the computational costs of the VOI workflow, namely by reducing: the size of the ensembles used in the robust optimization and history matching steps, and the number of plausible truths considered. Typically, we are able to reduce the size of the ensembles from 50-100 to 5-10 realizations (i.e., a factor of 10 approximately). Because of the loop structure of the VOI workflow, the total speed-up is of nearly two orders of magnitude to keep a fair approximation of the results obtained with the rigorous VOI procedure. These measures allowed us to apply the VOI assessment methodology to a larger case study, showing that, with further improvements, it may become more tractable even for real-field applications.
- Clustering techniques are powerful tools but there is not a one-fits-all solution. These methods rely on dissimilarity measures to distinguish and cluster the model realizations, and the choice of these measures is extremely case dependent. There are different opportunities to apply clustering within the VOI workflow, and the dissimilarity measures used must be fit for these different purposes:
  - When selecting representative models for history matching and robust production optimization, dissimilarity measures based on model-based features (e.g., reservoir parameters, flow patterns and production predictions) work well.
  - To identify representative plausible truths, it is more appropriate to consider dissimilarity measures determined by decision-based features (e.g., optimal sets of controls, improvements in objective function).
- In order to be effective in practice, the use of clustering techniques needs to be done carefully:

- When dealing with high-dimensional datasets, projection and decomposition techniques are useful to reduce the dimensionality of the data and avoid spurious correlations between the samples which can lead to meaningless clustering results.
- When selecting few representative models (e.g., 5-10), weighting the realizations has a significant impact on the ability to approximate the results of the full ensemble.

Chapter 5

- The current research state-of-the-art for optimization under uncertainty in reservoir engineering applications does not account for the availability of future information throughout the producing life of the reservoir, considering that uncertainties are static. This assumption is a limitation of the robust optimization approach and may lead to suboptimal production strategies.
- We have identified the opportunity to integrate our proposed VOI methodology within the optimization framework to determine production strategies that maximize not only the primary objective (e.g., NPV or recovery) but also the value of future measurements. As a matter of fact, these two objectives are not in conflict because, implicitly, the VOI procedure itself also aims to maximize the primary objective. Therefore, we address this optimization problem more naturally in a nested formulation rather than in a multi-objective setting. The new IPO approach also allows us to overcome some of the limitations of the conventional robust optimization, such as its inability to embrace the flexibility of production strategies and to account for time-dependent uncertainty.
- The proposed IPO approach takes advantage of the fact that we have control on which measurements will be available in the future because we design the surveillance plan. A single production strategy has to be determined for the period of time prior to the future observations so that it can be implemented in reality, but a range of strategies can be derived for the remaining of the reservoir life-cycle because these decisions have to wait to be determined after the future information is gathered. This allows us to capture the future flexibility of the production strategies. As a result, the

methodology delivers a flexible strategy which is optimal for the initial uncertainty and the future measurements planned, unlike the myopic strategy obtained by the conventional robust optimization approach which ignores completely the fact that future information will affect the operational decisions in the future.

- Results from our case study confirm that IPO can improve the way we approach the reservoir management problem, especially in situations where wrong decisions cannot be entirely compensated by the remaining degrees of freedom of the control problem. We attribute the improved performance to a combination of multiple effects, which cannot be easily decoupled and require further investigation to be fully understood.
- The extension of the IPO approach to cases with multiple future observations requires simplification of the structure of the problem to be tractable. Although previous work in other fields investigated the use of decomposition and constraint relaxation techniques for this purpose, we still need solutions for frameworks that rely on reservoir simulation models. In this thesis, we have proposed a possible solution that prunes the scenario tree by ignoring the branching which takes place after the two first stages, but the implications of this choice have not been studied in details.

Chapter 6

- The closed-loop and VOI frameworks are also relevant to field development optimization. In fact, they may be even more important than in reservoir management applications because of the larger uncertainties and risks present during the FDP phase. Although FDP activities involve other types of decisions and measurements, the structure of the problem is similar: the decisions (e.g., drilling locations, drilling order) are also made sequentially and under uncertainty, and additional information (e.g., well logs, production data) also becomes available along time.
- The sequential nature of drilling activities combined with the endogenous nature of the geological uncertainties make the current approaches for optimization under uncertainty suboptimal. Because there is opportunity for learning by obtaining data

as new wells are drilled and start production, the uncertainties are time- and decision-dependent. The common practice in FDP is to outline development plans based on the current state of uncertainty only. In this thesis, we have proposed a method to combine optimization with VOI considerations to design development plans which account for time-dependent uncertainty.

- Our IFDO methodology is an attempt to change the way we design and communicate field development plans. By proposing a single location for the wells which are to be drilled immediately and multiple locations for the wells to be drilled in the future, we propagate the geological uncertainties to the development recommendation and deliver plans that reflect the flexibility existent in the FDP phase. This can provide insight to support field development decisions.
- We have applied the IFDO approach to synthetic case studies with the assumption of future clairvoyance available, and the results showed increases of up to 6 % in terms of objective function values. We believe that IFDO can lead to even larger improvements in more realistic cases where only imperfect measurements are available, because, then, the revelation of uncertainties also becomes decision-dependent and conventional myopic approaches are more likely to lead to wrong decisions.
- We need suitable case studies to demonstrate the value of accounting for time-dependent uncertainties. In this thesis we have designed synthetic examples where future information could make a significant difference. However, they had to be simple enough to allow a proof of concept with reasonable computation time. For this reason, we recognize that these examples may not be the best ones to illustrate the full potential of IFDO for real applications and we believe that an effort should be made to find an appropriate benchmark case for research in field development optimization.

7.2. Future perspectives

Further acceleration of VOI assessment

In Chapter 4, we presented some measures to accelerate our proposed workflow for VOI assessment. We explored the use of clustering techniques to smartly select representative model realizations which can approximate the uncertainty characterized by the original (and usually large) ensembles of realizations. There are, however, other alternatives to reduce the computational costs of such a workflow. Techniques such as response surfaces, multiscale methods, reduced-order and reduced-physics modeling can be used to construct fast-evaluation models to generate approximations of the required forecasts. This could facilitate the application of our VOI procedure to larger reservoir models by reducing the number of high-fidelity reservoir simulations needed. We could also consider approximations at the optimization level, such as changes in the stopping criteria and other optimization settings to reduce the number of iterations required, or the use of appropriate heuristics to simplify the problem. Thus, there is scope for research on how to combine approximations at the simulation, optimization and uncertainty quantification levels to come up with more practical procedures for VOI assessment. The main challenge is to find a good balance between the accuracy of all these approximations and the computational costs. For that, it is important to derive measures to quantify the approximation errors so that we can determine the optimal compromise.

VOI assessment for unknown unknowns

In Chapter 2, we recognized that one of the limitations of our method is related the inability to assess the VOI in cases with "unmodeled" uncertainties. More recently, some studies on unregularized history matching have demonstrated that it is possible, in some cases, to identify under modeled features from dynamic data; see Joosten et al. (2011), Kahrobaei et al. (2014) and Kahrobaei et al. (2015). This is achieved by history matching synthetic data obtained with a synthetic truth containing model features that are not present in any of the model realizations of the initial ensemble. This reasoning could be used to extend our VOI procedure to cases with unknown unknowns, by introducing two or more ensembles of plausible truths generated based on different geological

assumptions. Such a modification could also allow us to extend the methodology to assess the value of (prior) geological knowledge.

Different risk preferences

Different individuals may have different attitudes towards risk. This is an example of personal preferences which may cause different decision makers to assess the VOI differently. The methodology presented in this thesis assumes an assessment in the perspective of a risk-neutral decision maker, and this can be seen as a limitation. In other applications, risk preferences are addressed by using utility theory. We believe there is scope for future research on how to integrate utility calculations into our framework for VOI assessment to overcome this limitation, along the lines of what is discussed in Bailey et al. (2011).

Additional insights from closed-loop solutions

In Chapter 3, we saw that the VOI assessment for multiple observation times requires us to "close the loop" multiple times. By doing that, we perform many history matches and optimizations to determine the VOI in the FDP phase. All these calculations produce also a large amount of data: optimal production strategies, production forecasts, synthetic measurements, model updates, simulated data, etc. Although synthetic, this information could be useful to understand the CLRM better, or to make it more efficient when applied in the operational phase, serving as a sort of pre-computed operation manual. As we mentioned before, one of the limitations of our methodology is related to its high computational load. Finding more uses for the information obtained from it can help us to promote its practice.

Formulation for reservoir surveillance optimization

The ability of assessing the VOI is just the first step to determine optimal reservoir surveillance plans. In Chapter 2, we showed how we could use our methodology to determine which measurements result in higher VOI. However, there is more to be taken into account within a reservoir surveillance optimization problem, such as the deployment costs of the measurements or how to model the degrees of freedom of the problem (i.e., how to parametrize a reservoir surveillance program). To the best of our knowledge, there

are currently no studies establishing a proper formulation for the reservoir surveillance optimization problem, for different types of measurements and control variables. We believe there is scope for further work in this direction to formalize an accurate problem description addressing the decisions made in the design of reservoir surveillance plans before we can work on techniques to efficiently solve this optimization problem.

Joint reservoir surveillance and production optimization

In Chapter 5, we showed that we can improve production optimization by accounting for the fact that we design the reservoir surveillance plans, meaning that we know which measurements will become available in the future. However, we did not account for the fact that this design can itself be optimized. For every possible surveillance strategy, there is an optimal production strategy, which can be determined by following the IPO approach described in Chapter 5. A similar structure is present when optimizing well locations and well controls simultaneously: for every possible well location configuration, there is an optimal set of well controls. This type of problem can be solved with a nested optimization approach. Thus, it would also be possible to solve the joint reservoir surveillance and production optimization problem. This would require the solution of a nested optimization (i.e., the IPO) within another nested optimization to optimize reservoir surveillance plans and production strategies simultaneously. The solution of such a problem requires further investigation.

Informed optimization with multiple decision stages

As discussed in Chapter 5, the informed optimization approach cannot be easily extended for cases with multiple stages of decisions without leading to exponential increase of computational costs. To make the problem more tractable, researchers in other areas have studied the structure of multi-stage stochastic models and how to solve them with the help of a variety of computational strategies and approximation measures. Nonetheless, very little has been done in adapting these solutions for applications which rely on reservoir simulations, opening room for future research.

Link with controllability and observability

As discussed throughout the chapters of this thesis, our proposed methodology for VOI assessment in CLRM combines production optimization and history matching, allowing us to quantify how much future measurements impact our knowledge and our decisions in terms of enabled value. From modern systems and control theory, it is possible to determine the controllability and observability within a system model. In the CLRM context, these concepts can be relevant for the optimization of flooding processes and computer-assisted history matching; see Zandvliet et al. (2008) and Van Doren et al. (2013). Most likely there is a way to interpret the results of our VOI assessment in terms of controllability and observability measures. We recommend further research to verify this hypothesis, which can lead to a better understanding of the problem and potentially to alternative ways to assess the value of future measurements.

Cross observation sensitivities

In Chapter 2, we referred to previous work on quantifying the impact of observations in reservoir engineering applications (Krymskaya et al., 2010). They showed that a measure of the information content of the observations can be obtained from the self-sensitivities (i.e., diagonal elements of the observation sensitivity matrix). The observation sensitivity matrix also contains information on the cross-sensitivities, the off-diagonal elements which measure the impact of the observations on the prediction of other observations. A better understanding of these sensitivities can lead to powerful tools. For example, one could use the correlations between a certain observation and future model predictions to establish a smarter framework for VOI assessment.

Value of flexibility

The methodology presented in this thesis is above all a methodology to assess value. Within the CLRM framework, we have the possibility of quantifying the VOI of measurements which are incorporated to the models through the history matching step. However, acquiring additional information is only one of the strategies to cope with uncertainties. Investing in flexibility can also be a valuable strategy in uncertain situations and our methodology can be used to assess its value. The CLRM framework also handles the update of the production strategies through the optimization step and we can use it to

quantify the added value of providing additional flexibility to the optimization. There are many different types of additional flexibility that can be considered and the moment in which it becomes available is a choice of the decision maker. We can model the additional flexibility by removing some of the constraints or by introducing new degrees of freedom to the optimization problem after the new capacity is available.

References

Aanonsen, S.I., Naevdal, G., Oliver, D.S, Reynolds, A.C. and Valles, B. (2009). The ensemble Kalman filter in reservoir engineering – a review. SPE Journal, 14 (3) 393-412.

Abreu, A.C.A; Booth, R.; Bertolini, A.; Prange, M.; Bailey, W.J., Teixeira, G.; Romeu, R.K.; Emerick, A.A. and Pacheco, M.A.C. (2015). Proactive and reactive strategies for optimal operational design: an application in smart wells. Paper OTC-26209-MS presented at the Offshore Technology Conference Brazil, Rio de Janeiro, Brazil, 27-29 October.

Afra, S. and Gildin, E. (2016). Tensor based geology preserving reservoir parameterization with Higher Order Singular Value Decomposition (HOSVD). *Computers & Geosciences* **94**, 110–120.

Aggarwal, C.C., Wolf, J.L., Yu, P.S., Procopiuc, C., Park, J.S. (1999). Fast algorithms for projected clustering. *SIGMOD Record*, **28** (2), 61-72.

Arabie, P. and Hubert, L.J. (1996). An overview of combinatorial data analysis. *In* Arabie, P., Hubert, L.J., and Soete, G.D. (eds.), *Clustering and Classification*, 5-63. World Scientific, River Edge, NJ.

Armstrong, M., Ndiaye, A., Razanatsimba, R. and Galli, A. (2013). Scenario reduction applied to geostatistical simulations. *Mathematical Geosciences*, **45**, 165-182.

Bailey, W.J., Couët, B. and Prange, M. (2011). Forecast optimization and value of information under uncertainty. *In* Uncertainty analysis and reservoir modeling: *AAPG Memoir*, **96**, 217-233.

Baker, M. (2015). Use of cluster analysis to improve representative model selection: a case study. Paper SPE 176408-MS presented at SPE/IATMI Asia Pacific Oil & Gas Conference and Exhibition, Nusa Dua, Bali, Indonesia, 20-22 October.

Barros, E.G.D., Leeuwenburgh, O., Van den Hof, P.M.J. and Jansen, J.D. (2015a). Value of multiple production measurements and water front tracking in closed-loop reservoir management. Paper SPE-175608-MS presented at the SPE Reservoir Characterization and Simulation Conference and Exhibition, Abu Dhabi, UAE, 14-16 September.

Barros, E.G.D., Jansen, J.D. and Van den Hof, P.M.J. (2014). Value of information in closed-loop reservoir management. Paper presented at 14th European Conference on Mathematics in Oil Recovery, Catania, Italy, 8-11 September.

Barros, E.G.D., Jansen, J.D. and Van den Hof, P.M.J. (2015b). Value of information in parameter identification and optimization of hydrocarbon reservoirs. Paper presented at 2nd IFAC Workshop on Automatic Control in Offshore Oil and Gas Production, Florianópolis, Brazil, May, 27-29.

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.

Barros, E.G.D., Yap, F.K., Insuasty Moreno, E., Van den Hof, P.M.J. and Jansen, J.D. (2016b). Clustering techniques for value-of-information assessment in closed-loop reservoir management. Paper presented at the 15th European Conference on Mathematics in Oil Recovery, Amsterdam, The Netherlands, 29 August - 1 September.

Bellout, M.C., Echeverría Ciaurri, D., Durlofsky, L.J., Foss, B., Kleppe, J. (2012). Joint optimization of oil well placement and controls. *Computational Geosciences*, **16** (4), 1061-1079.

Bhattacharjya, D., Eidsvik, J., Mukerji, T. (2010). The value of information in spatial decision making. *Mathematical Geosciences,* **42** (2), 141-163.

Birge, J.R., and Louveaux, F. (1997). *Introduction to stochastic programming*. Springer, New York.

Bratvold, R.B., Bickel, E.J., and Lohne, H.P. (2009). Value of information: the past, present, and future. SPE Reservoir Evaluation and Engineering, 12 (4), 630-638.

Bukshtynov, V., Volkov, O., Durlofsky, L.J. and Aziz, K. (2015). Comprehensive framework for gradient-based optimization in closed-loop reservoir management. *Computational Geosciences*, **19** (4), 877-897.

Caers, J. (2011). Modeling uncertainty in the earth sciences. John Wiley & Sons.

Capolei, A., Suwartadi, E., Foss, B. and Jørgensen, J.B. (2015). A mean-variance objective for robust production optimization in uncertain geological scenarios. *Journal of Petroleum Science and Engineering*, **125**, 23–37.

Cardoso, M.A. and Durlofsky, L.J. (2010). Use of reduced-order modeling procedures for production optimization. *SPE Journal*, **15** (2), 426-435.

Castro, S.A., Caers, J. and Mukerji, T. (2005). The Stanford VI reservoir. 18th Annual Report. Stanford Center for Reservoir Forecasting. Stanford University, May 2005.

Chen, B., He, J., Xie, J., Sarma, P., Wen, X.-H., Chen, W.H. and Reynolds, A.C. (2016). Pilot design analysis using proxies and Markov Chain Monte Carlo method. Paper presented at the 15th European Conference on Mathematics in Oil Recovery, Amsterdam, The Netherlands, 29 August - 1 September.

Chen, Y., Oliver, D.S. and Zhang, D. (2009). Efficient ensemble-based closed-loop production optimization. *SPE Journal*, **14** (4), 634-645.

Cunningham, J.P. and Ghahramani, Z. (2015). Linear dimensionality reduction: survey, insights, and generalizations. *Journal of Machine Learning Research*, **16**, 2859-2900.

Cunningham, P. and Begg, S. (2008). Using the value of information to determine optimal well order in a sequential drilling program. *AAPG Bulletin*, **92** (10), 1393-1402.

Eidsvik, J., Mukerji, T. and Bhattacharjya, D. (2015). Value of information in the earth sciences – Integrating spatial modeling and decision analysis, 1st ed, Cambridge University Press.

Embid Droz, S.M., Rodríguez Torrado, R., Echeverría Ciaurri, D. and Barros, E.G.D. (2017). Method for determining a drilling plan for a plurality of new wells in a reservoir. Spanish Patent ES P201700314, filed March 29, 2017.

Evensen, G. (2009). *Data assimilation – The ensemble Kalman filter*, 2nd ed., Springer, Berlin.

Fel'dbaum, A.A. (1960). Dual control theory I-II. *Automation Remote Control*, **21** (9), 874-880.

Fel'dbaum, A.A. (1961). Dual control theory III-IV. *Automation Remote Control*, **22** (11), 109-121.

Fonseca, R.M., Kahrobaei, S., Van Gastel, L.J.T., Leeuwenburgh, O., and Jansen, J.D. (2015). Quantification of the impact of ensemble size on the quality of an ensemble gradient using principles of hypothesis testing. Paper 173236-MS presented at the 2015 SPE Reservoir Simulation Symposium, 22-25 February, Houston, USA.

Fonseca, R.M., Chen, B., Jansen, J.D., and Reynolds, A.C. (2016). A Stochastic Simplex Approximate Gradient (StoSAG) for optimization under uncertainty. Accepted for publication in *International Journal for Numerical Methods in Engineering.*

Forgione, M., Birpoutsoukis, G., Bombois, X., Mesbah, A., Daudey, P.J. and Van den Hof, P.M.J. (2015). Batch-to-batch model improvement for cooling crystallization. *Control Engineering Practice*, **41**, 72-82.

Foss, B. and Jensen, J.P. (2011). Performance analysis for closed-loop reservoir management. *SPE Journal*, **16**, 183-190.

Georghiou, A., Wiesemann, W. and Kuhn, D. (2011). The decision rule approach to optimisation under uncertainty: methodology and applications in operations management. Available on *Optimization Online*, submitted for publication.

Goel, V., and Grossmann, I.E. (2004). A stochastic programming approach to planning of offshore gas field developments under uncertainty in reserves. *Computers and Chemical Engineering*, **28** (8), 1409-1429.

Gupta, V. and Grossmann, I.E. (2011). Solution strategies for multistage stochastic programming with endogenous uncertainties. *Computers and Chemical Engineering*, **35** (11), 2235-2247.

Güyagüler, B. and Horne, R.N. (2004). Uncertainty assessment of well-placement optimization. *SPE Reservoir Evaluation & Engineering*, **7** (1), 24-32.

Hanea, R.G., Fonseca, R.M., Pettan, C., Iwajomo, M.O., Skjerve, K., Hustoft, L., Chitu, A.G. and Wilschut, F. (2016). Decision maturation using ensemble based robust optimization for field development planning. Paper presented at the 15th European Conference on the Mathematics of Oil Recovery, Amsterdam, The Netherlands, 29 August - 1 September.

Hanssen, K.G. (2017). Optimal control under uncertainty – Applied to upstream petroleum production. PhD thesis, Norwegian University of Science and Technology.

Hauskrecht, M. (2000). Value-function approximations for partially observable Markov decision processes. *Journal of Artificial Intelligence Research*, **13**, 33-94.

He, J., Sarma, P. and Durlofsky, L.J. (2013). Reduced-order flow modeling and geological parameterization for ensemble-based data assimilation. *Computers and Geosciences*, **55**, 54-69.

He, J., Xie, J., Sarma, P., Wen, X.H., Chen, W.H. and Kamath, J. (2016). Proxy-based work flow for a priori evaluation of data-acquisition programs. *SPE Journal*, **21** (4), 1400-1412.

Hewson, C.W. (2015). Reduced-order modelling for production optimisation. MSc thesis, Delft University of Technology.

Hou, J., Zhou, K., Zhang, X.S. Kang, X.D. and Xie, H. (2015). A review of closed-loop reservoir management. *Petroleum Science*, **12**, 114-128.

Howard, R.A. (1966). Information value theory. *IEEE Transactions on Systems, Science and Cybernetics*, **SSC-2** (1), 22-26.

IEA (2016). World Energy Outlook 2016, IEA, Paris.

Insuasty, E., Van den Hof, P.M.J., Weiland, S. and Jansen, J.D. (2015). Spatial-temporal tensor decompositions for characterizing control-relevant flow profiles in reservoir models. Paper SPE-173238-MS presented at the SPE Reservoir Simulation Symposium, Houston, USA, 23-25 February.

Insuasty, E., Van den Hof, P.M.J., Weiland, S. and Jansen, J.D. (2017). Flow-based dissimilarity measures for reservoir models: a spatial-temporal tensor approach. *Computational Geosciences*, **21** (4), 645-663.

Jansen, J.D., Brouwer, D.R., Nævdal, G. and van Kruijsdijk, C.P.J.W. (2005). Closed-loop reservoir management. *First Break*, January, **23**, 43-48.

Jansen, J.D., Bosgra, O.H. and van den Hof, P.M.J. (2008). Model-based control of multiphase flow in subsurface oil reservoirs. *Journal of Process Control* **18**, 846-855.

Jansen, J.D., Douma, S.G., Brouwer, D.R., Van den Hof, P.M.J., Bosgra, O.H. and Heemink, A.W. (2009). Closed-loop reservoir management. Paper SPE 119098 presented at the SPE Reservoir Simulation Symposium, The Woodlands, USA, 2-4 February.

Jansen, J.D. and Durlofsky, L.J. (2016). Use of reduced-order models in well control optimization. *Optimization and Engineering*. Published online.

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

Jesmani, M, Jafarpour, B., Bellout, M.C., Hanea, R.G. and Foss, B. (2016). Application of simultaneous perturbation stochastic approximation to well placement optimization under uncertainty. Paper presented at the 15th European Conference on the Mathematics of Oil Recovery, Amsterdam, The Netherlands, 29 August - 1 September.

Jonsbråten, T.W. (1998). Optimization models for petroleum field exploitation. PhD thesis, Norwegian School of Economics and Business Administration.

Joosten, G.J., Altintas, A. and De Sousa, P. (2011). Practical and operational use of assisted history matching and model-based optimisation in the Salym field. Paper SPE 146697-MS presented at the SPE Annual Technical Conference and Exhibition. Denver, USA, 30 October - 2 November.

Kahrobaei, S., Mansoori, M., Joosten, G.J.P., Van den Hof, P.M.J. and Jansen, J.D. (2014). Hidden information in ill-posed inverse problems. Paper presented at 14th European Conference on Mathematics in Oil Recovery, Catania, Italy, 8-11 September.

Kahrobaei, S., Mansoori, M., Joosten, G.J.P., Van den Hof, P.M.J. and Jansen, J.D. (2015). Identifiability of location and magnitude of flow barriers in slightly compressible flow. *SPE Journal*, **21** (3), 899-908.

Keim, D., Berchtold, S., Böhm, C., Kriegel., H.P. (1997). A cost model for nearest neighbor search in high dimensional data space. *In Proc.* of the 16th Symposium on Principles of Database Systems, 78-86.

Kikani, J. (2013). *Reservoir Surveillance*, SPE, Richardson.

Krause, A. and Guestrin, C. (2007). Nonmyopic active learning of Gaussian processes: an exploration-exploitation approach. *Proc.* 24th Annual International Conference on Machine Learning, Corvallis, USA, 20-24 June.

Krause, A. and Guestrin, C. (2009). Optimal value of information in graphical models. *Journal of Artificial Intelligence Research*, **35**, 557-591.

Kruskal, J.B. (1964). Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. *Psychometrika*, **29** (1), 1-27.

Kruskal, J.B. (1976). The relationship between multidimensional scaling and clustering. *In Proc.* of Advanced Seminar Conducted by the Mathematics Research Center, the University of Wisconsin-Madison, 3-5 May, 17-44.

Krymskaya, M.V., Hanea, R.G., Jansen, J.D. and Heemink, A.W. (2010). Observation sensitivity in computer-assisted history matching. Paper presented at the 72nd EAGE Conference and Exhibition, Barcelona, Spain, 14-17 June.

Le, D.H. and Reynolds, A.C. (2014a). Optimal choice of a surveillance operation using information theory. *Computational Geosciences*, **18**, 505-518.

Le, D.H. and Reynolds, A.C. (2014b). Estimation of mutual information and conditional entropy for surveillance optimization. *SPE Journal*, **19** (4), 648-661.

Leeuwenburgh, O. and Arts, L.J. (2014). Distance parameterization for efficient seismic history matching with the ensemble Kalman filter. *Computational Geosciences*, **18**, 535-548.

Leeuwenburgh, O., Chitu, A.G., Nair, R., Egberts, P.J.P., Ghazaryan, L., Feng, T. and Hustoft, L. (2016). Ensemble-based methods for well drilling sequence and time optimization under uncertainty. Paper presented at the 15th European Conference on the Mathematics of Oil Recovery, Amsterdam, The Netherlands, 29 August - 1 September.

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

Lyons, J. and Nasrabadi, H. (2013). Well placement optimization under time-dependent uncertainty using an ensemble Kalman filter and a genetic algorithm. *Journal of Petroleum Science and Engineering*, **109**, 70-79.

Mirrokni, V., Thain, N. and Vetta, A. (2012). A theoretical examination of practical game playing: lookahead search. In: Serna, M. (ed.) SAGT 2012. LNCS, **7615**, 251-262. Springer, Heidelberg.

Nakayasu M., Goda, T., Tanaka, K. and Sato K. (2016). Evaluating the value of single-point data in heterogeneous reservoirs with the expectation-maximization algorithm. *SPE Economics & Management*, **8** (1), 1-10.

Naevdal, G., Brouwer, D.R. and Jansen, J.D. (2006). Waterflooding using closed-loop control. *Computational Geosciences*, **10** (1), 37-60.

Oliver, D.S., Reynolds, A.C. and Liu, N. (2008). *Inverse theory for petroleum reservoir characterization and history matching*, Cambridge University Press, Cambridge.

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

Özdogan, U. and Horne, R.N. (2006). Optimization of well placement under time-dependent uncertainty. *SPE Reservoir Evaluation & Engineering*, **9** (2), 135-145.

Parsons, L., Haque, E. and Liu, H. (2004). Subspace clustering for high dimensional data: a review. *ACM SIGKDD Explorations Newsletter*, **6** (1) 90-105.

Rockafellar, R.T. and Uryasev, S. (2000). Optimization of conditional valueat-risk. *Journal of Risk*, **2**, 21-42.

Ruszczyński, A. and Shapiro, A. (2003). *Stochastic programming.* Handbooks in Operations Research and Management Science, Elsevier, Amsterdam.

Sarma, P., Durlofsky, L.J. and Aziz, K. (2008). Computational techniques for closed-loop reservoir modeling with application to a realistic reservoir. *Petroleum Science and Technology*, **26** (10 and 11), 1120-1140.

Sarma, P., Chen, W.H. and Xie, J. (2013). Selecting representative models from a large set of models. Paper SPE 163671 presented at the SPE Reservoir Simulation Symposium, The Woodlands, USA, 18-20 February.

Sato, K. (2011). Value of information analysis for adequate monitoring of carbon dioxide storage in geological reservoirs under uncertainty. *International Journal of Greenhouse Gas Control*, **5** (5):1294-1302

Scheidt, C. and Caers, J. (2009). Representing spatial uncertainty using distances and kernels. *Mathematical Geosciences*, **41** (4), 397-419.

Sethian, J.A. (1996). A fast marching level set method for monotonically advancing fronts. *Proc. Nat. Acad. Sci.*, **93** (4), 1591-1595.

Shepard, R.N. (1962). The analysis of proximities: Multidimensional scaling with an unknown distance function. I. *Psychometrika*, **27** (2):125-140.

Shirangi, M.G. and Durlofsky, L.J. (2015). Closed-loop field development under uncertainty by use of optimization with sample validation. *SPE Journal*, **20** (5), 908-922.

Shirangi, M.G. and Durlofsky, L.J. (2016). A general method to select representative models for decision making and optimization under uncertainty. *Computers and Geosciences*, **96**, 109-123.

Siraj, M.M., Van den Hof, P.M.J. and Jansen, J.D. (2016). Robust optimization of water-flooding in oil reservoirs using risk management tools. Paper presented at 11th IFAC Symposium on Dynamics and Control of Process Systems, Trondheim, Norway, 6-8 June.

Smallwood, R.D. and Sondik, E.J. (1973). The optimal control of partially observable Markov processes over a finite horizon. *Operations Research*, **21** (5), 1071-1088.

Strebelle, S. (2002). Conditional simulation of complex geological structures using multiple-point statistics. *Mathematical Geology*, **34** (1), 1-22.

Tarhan, B., Grossmann, I.E. and Goel, V. (2013). Computational strategies for non-convex multistage MINLP models with decision-dependent uncertainty and gradual uncertainty resolution. *Annals of Operations Research*, **203**, 141-166.

Trainor-Guitton, W. J., Mukerji, T., and Knight, R. (2013). A methodology for quantifying the value of spatial information for dynamic earth problems. *Stochastic Environmental Research and Risk Assessment*, **27** (4), 969-983.

Van Doren, J.F.M., Van den Hof, P.M.J., Bosgra, O.H. and Jansen, J.D. (2013). Controllability and observability in two-phase porous media flow. *Computational Geosciences*, **17** (5), 773-788.

Van Essen, G.M., Zandvliet, M.J., Van den Hof, P.M.J., Bosgra, O.H. and Jansen, J.D. (2009). Robust waterflooding optimization of multiple geological scenarios. *SPE Journal*, **14** (1), 202-210.

Van Hessem, D.H. (2004). Stochastic inequality constrained closed-loop model predictive control with application to chemical process operation. PhD thesis, Delft University of Technology.

Wang, C., Li, G. and Reynolds, A.C. (2009). Production optimization in closed-loop reservoir management. *SPE Journal*, **14** (3), 506-523.

Wang, H., Echeverría Ciaurri, D., Durlofsky, L.J. and Cominelli, A. (2012). Optimal well placement under uncertainty using a retrospective optimization framework. *SPE Journal*, **17** (1), 112-121.

Yap, F.K. (2016). Representative models for history matching and robust optimization. MSc thesis, Delft University of Technology.

Yeten, B., Durlofsky, L.J. and Aziz, K. (2003). Optimization of nonconventional well type, location and trajectory. *SPE Journal*, **8** (3), 200-210.

Zandvliet, M.J., Van Doren, J.F.M., Bosgra, O.H., Jansen, J.D. and Van den Hof, P.M.J. (2008). Controllability, observability and identifiability in single-phase porous media flow. *Computational Geosciences*, **12** (4), 605-622.

Zenith, F., Foss, B.A., Tjønnås, J. and Hasan, A.I. (2015). Well testing by sinusoidal stimulation. *SPE Reservoir Evaluation & Engineering*, **18** (3), 441-451.

Zhou, Y. (2012). Parallel General-Purpose Reservoir Simulation with Coupled Reservoir Models and Multi-Segment Wells. PhD thesis, Stanford University.

Nomenclature

AD-GPRS = Automatic Differentiation General Purpose Research Simulator

b = discount factor

B = formation volume factor

bbl = oil barrel, a unit of volume

BHP = bottom-hole pressure

c = compressibility

C = set of indices (cluster) or partition

cdf = cumulative probability distribution

CLFD = closed-loop field development

CLRM = closed-loop reservoir management

COK = chance of knowning

CVaR = conditional value at risk

d = distance

$\mathbf{d}$ = vector of measured data

EnKF = Ensemble Kalman filter

EnOpt = Ensemble Optimization

F = feature operator

FDP = field development planning

H = Heaviside step function

HOSVD = high-order singular value decomposition

i = index or counter

I_{GAI} = global average influence index

IFDO = informed field development optimization

IPO = informed production optimization

j = index or counter

J = objective function value

k = index, counter or relative permeability

$\mathbf{m}$ = vector of model parameters

M = number of control time intervals or data dimensionality

$\mathbf{M}$ = ensemble of model realizations

MDS = multidimensional scaling

MRST = Matlab Reservoir Simulator Toolbox

n = Corey exponent

N = number of model realizations

N_{obs} = number of observations

NPV = net present value

O = workflow complexity

p = pressure

P_x = x % percentile

pdf = probability density function

PCA = principal component analysis

POMDP = Partially observable Markov decision process

PSO = particle swarm optimization

q = volumetric flow rate

r = price per unit volume

S = saturation

$\mathbf{S}$ = observation sensitivity matrix

SGeMS = Stanford Geostatistical Modeling Software

SVD = singular value decomposition

StoSAG = Stochastic Simplex Approximate Gradient

t = time

T = producing lifetime of the reservoir

$\mathbf{u}$ = vector of control variables

UQ = uncertainty quantification

VOI = value of information

VOC = value of clairvoyance

wct = water-cut

ε = measurement error

$\boldsymbol{\theta}$ = data vectors

$\boldsymbol{\Theta}$ = set of objects (dataset)

μ = viscosity or mean value

ϱ = volumetric mass density

σ = standard deviation

τ = reference time for discount factor

$\boldsymbol{\varphi}_i$, $\boldsymbol{\psi}_j$, $\boldsymbol{\chi}_k$ = basis functions

Acknowledgements

"Appreciation is a wonderful thing: it makes what is excellent in others belong to us as well."

— Voltaire

After a very long journey of more than five years, this thesis would not be complete without some acknowledgements to express my sincere gratitude to all the people that contributed to this achievement. Although here the acknowledgements take the form of words, my appreciation goes way beyond them. I believe in showing gratitude throughout the entire process rather than just expressing it in a look-back step. Thus, besides what is written in the next lines, I hope the groups and individuals listed below can relate to the connections and moments we shared, because there lie my kindest feelings for each one of them and the experiences that I value the most.

I would like to thank my promotor Jan Dirk Jansen for his guidance and supervision in both the technical and non-technical aspects of my PhD process. Jan Dirk, it has been a real pleasure to work and interact with you for the past five years. From the moment I had that Skype interview for the PhD position until the final struggles to publish the thesis, I learned a lot from you. I am really grateful for the trust you deposited in me: together with other ingredients such as challenge, criticism and support, all in the right dose, you found a good balance that helped me stay motivated to reach my objectives while developing myself as an independent researcher and keeping things under control (at least for the most part of it). Thank you very much for the countless opportunities you provided me with during the PhD: all the courses, conferences and contacts, so many new experiences from which I tried to get as much as possible. I appreciate your scientific involvement in my research: thank you for all the ideas we developed together and for always being engaged in the (sometimes) too convoluted discussions on VOI workflows. Thank you also for always being frank with me, for encouraging my initiatives, for being a more strict "boss" when necessary and for defending my interests when I could not. You helped me become a better professional in so many levels and I truly believe we formed a great team.

I would like to thank Paul Van den Hof for his role as my copromotor. Paul, your guidance and input were very important in crucial moments of my PhD. I learned a lot

from our technical discussions: the exposure to concepts (and language) from the systems and control community helped me find more clarity to understand and communicate the philosophy of my work. From the discussions, I also learned from you that it is possible to debate and oppose while remaining always kind and positive. I also appreciate your commitment on reviewing my dissertation in detail: your sharp comments contributed to considerably improve many parts of this thesis. In particular, they enlightened me on how to take a step back in order to frame my research questions more explicitly and how to support my reasoning with clarity by strengthening my scientific arguments and critical analysis.

I would like to acknowledge Petrobras who were the main sponsors of my PhD project. In particular, I would like to thank Régis Romeu and Flávia Pacheco for giving me the opportunity of doing an internship in their team at Cenpes. That internship kicked off a series of events that lead me to start (and conclude!) the PhD in Delft. Régis and Flávia, thank you also for believing in my potential and presenting me as a candidate to one of the ISAPP positions sponsored by Petrobras. Régis, thank you for your mentorship during the internship and even after that, and for the always stimulating conversations about the most diverse topics in life. I am extremely grateful to have crossed paths with someone with your profile; everything I learned from you truly inspires me to become a more complete professional and person.

I would like to thank TNO for the coordination of the ISAPP program. In particular, I would like to thank Olwijn Leeuwenburgh for following the progress of my PhD from close and with interest during these past years, always with precise discussions to share his view on my work and understand its implications. I would also like to thank Richard Braal for his role in the ISAPP community and for the pleasant interactions, which have helped me to keep thinking about all the business aspects of developing new technology even when I was immersed in my computational workflows. I would like to thank Dries Hegen for impeccably keeping track of the progress of the ISAPP projects (including mine), for reviewing this thesis while reading to approve it for clearance and for his sincere effort on engaging all the involved parts to do their best. I thank also Remus Hanea who was involved in my selection process for the PhD position and my daily supervisor during my first months in Delft before he left to Statoil in Norway. Remus, even if we did not have

the chance of working a lot together, you were important for me during the process for the strong human component you add to the mix: your relational skills, charisma and attitude to make things happen keep inspiring my personal development. I would also like to thank TNO for the new opportunity that they have offered me: my first real job after the PhD cycle! For that, besides those already mentioned above, I thank Philippe Steeghs, Frank Wilschut and my good friend Rahul Fonseca who opened the doors to help me find a spot in their team. I look forward to the new challenges and to working with my new colleagues.

I would like to thank Prof. Marco Aurélio Pacheco from PUC-Rio, my former university back in Brazil, for the additional financial support provided during my PhD in the form of a supplementary scholarship and coverage of travel expenses to the Netherlands. I thank also Flávia Pacheco, Régis Romeu and Marco Antonio Cardoso from Petrobras for this arrangement with Marco Aurélio's group through the Octopus project; this help was very important for me.

I would like to thank IBM for organizing a summer internship at IBM T.J. Watson Research Center, which gave me the opportunity to experience how research is done in a large corporation, with technological objectives for their products, organizational challenges and a client-oriented strategy, and how it is to be taken on board as an expert with a specific mission and a limited time frame in a different culture. In particular, I would like to thank David Echeverría, my internship mentor, for providing me an experience of true collaboration. David, I truly enjoyed working with you; thank you for the opportunity. I must say that the time we spent together contributed significantly to make me grow as an independent researcher. It was very gratifying to bring in some new knowledge and see how very quickly you understood the ideas and joined me to develop (and defend) them together. Thank you also for sharing your view and advice on everything that goes beyond the content of the work, such as your professional experience as a researcher, life in New York City, and, of course, the footballistic discussions! I would also like to thank my other colleagues at IBM, for their help and involvement during my internship.

I would like to thank Sintef for making available the MRST reservoir simulator, which was fundamental for this research. A special thanks to Olwijn Leeuwenburgh (TNO) for the

development of the EnKF module for MRST which was used in most of the applications presented in this thesis. I would also like to thank Lou Durlofsky and Oleg Volkov from Stanford University for sharing a working version of the AD-GPRS reservoir simulator and its optimization module. I acknowledge Stanford University also for the use of SGeMS.

I would like to thank my doctoral defense committee members, Prof. Arnold Heemink, Prof. Bill Rossen, Prof. Reidar Bratvold, Dr. Alexandre Emerick, Dr. Olwijn Leeuwenburgh and Prof. Pacelli Zitha for their availability to be on my PhD committee and for reviewing this thesis.

I would like to thank TU Delft for the excellent conditions it provided me with during the PhD cycle, which allowed me to focus almost exclusively on my research work and my development as a researcher. I would like to thank the Department of Geoscience and Engineering that hosted me during these years. To the professors, lecturers and research staff, thank you for the amazing amount of knowledge that I could absorb from you. To the MSc students, thank you for the lively atmosphere and for all the opportunities to learn by teaching and interacting with you. A special thanks to the MSc students who I had the pleasure to coach: Chris and Fu Kai, you were very important in this journey, in ways that you are not even aware of. It was a great experience to contribute to your learning curves and follow your progress; it has really taught me a lot about personal development and working relationships. Thank you, obviously, for helping me get a lot of work done, but thank you mainly for your interest and enthousiasm about our research, for your respect and trust on my supervision and, of course, for your friendship. I would also like to thank the SPE student chapter in Delft for all the organized activities, which helped me stay in touch with the oil and gas industry, and for the opportunity to join the board, which taught me a lot about serving and team work in a voluntary context. I would also like to thank the support staff of the department, all the team of secretaries and floor manager for being always friendly, approachable and helpful regarding my requests and for solving so many of my administrative problems. A special thanks to Guus and Joost for their excellent job in setting up and maintaining the remote servers, which have worked flawlessly and helped me so much with all my simulations. I thank also the facilities team and the cleaning staff for keeping a clean, safe and comfortable work environment.

I would also like to thank all the friends that I have met along this journey. I consider myself very lucky for having crossed paths with so many interesting people, making new friends from the first day when I walked in the office without knowing anyone until the last moments of my PhD. Some friends come to hopefully stay in my life for a long time, others were just there for some time, but each one of you were very important for me in your own way and at your own time. I thank you from the bottom of my heart for all the moments we shared together, all the fun moments and all the lessons learned. You have certainly made this journey more pleasant and helped me stay focused on living the present and taking the most out of it. I would like to thank my friends from the Petroleum engineering section for the amazing atmosphere in the office: I enjoyed so much being a part of this group, with a constantly evolving mix of different cultures, personalities and backgrounds. You are the best! The discussions (even the serious ones), the fun moments, the great laughs, the unpredictable conversations during lunch and coffee breaks, the birthday cakes outside the MV office, the group dinners and drinks, the summer barbecues, the inside jokes, stories and nicknames, all these will be always remembered because you have made them special. I would also like to thank the friends of Tuesday's football. That group started small with a few guys from the Petroleum group but grew a lot along the years to embrace colleagues from other groups, departments and faculties in the university, introducing me to so many good friends. Thank you for being there with me every week, almost religiously, even under adverse weather conditions. During the most difficult times of the PhD, knowing that Tuesday's evenings would come with nice football really helped me to keep going. I would also like to thank the friends that I met during my time at IBM. Even though it was for a short period, it was great spending time with you. Thank you for the nice atmosphere in the interns' office, for the conversations in the shuttle car, for the games during the breaks, for the movie nights and for your company in discovering the amazing New York City during that summer. I would also like to thank all the friends that I met in other contexts: from courses and conferences, from SPE and MV, neighbors, friends of friends and also those who I met by accident. Thank you very much for your friendship.

I would like to thank my girlfriend Jessica for all her love and support. Jessica, I do not know how I would have made it without you by my side at all times. If anyone knows

what it took me to get this degree, this person is you. You are truly my "partner in crime" for this one. Thank you for celebrating my achievements with me, for cheering me up when I could not stay positive, for listening to all my PhD related stories, for trying to understand my research crazy stuff despite not having the right background, for your patience when I had to work odd hours, when I only complained about things and when the next steps seemed so uncertain, and, of course, for making my life more complete. I hope that you have enjoyed the journey so far and that we continue to discover together what the future holds for us. I would also like to thank Jessica's family and friends for embracing me as one of yours so naturally. You were all fundamental for my adaptation in the Netherlands and you continue to make me feel like I belong here, which is so important for me as my family is not close by. Thank you also for your patience with my (so far) limited Dutch.

Last but most important, I would like to thank my mother and my brother for the continous love, support and encouragement. You are the ones who took the hardest part of this deal: giving me always the best of you to only see my victories from distance. I missed you so much during this time away. Mom, thank you for the solid education you provided me with and for shaping my character based on the good and love. I had to learn a lot of other things to get to this point, but the values that I learned from you formed the basis that allowed me to find my own way to accomplish everything that I have accomplished so far. Gabriel, thank you for reminding me of who I am in my essence. Simply by being yourself, you are like a mirror that gives me strength and confidence when I forget what I am capable of. Thank you also for sharing with me your wisdom and maturity to make good arguments, even if I am the older brother. I would also like to thank the rest of my family with whom I am really proud to share this achievement. To my grandparents, uncles, aunts and cousins, thank you for your love, for always rooting for and being proud of my success, and for caring about me. Life took me to distant places, but I remain the same Dudu that you know, just in a slightly more experienced version.

Delft, January 2018

Eduardo Barros

Thank you!


Baptiste
Olwijn Leeuwenburgh
Srivats
Paula
Monica
Bas
Koen
Roderik
Edwin
TNO
Rafael
Matteo
Brian Gaucher
Cees
Tom
Irene
SPE
Oleg Volkov
Geraldo
Jinyu
Quint
Petrobras
Cleaning staff
Lou Durlofsky
Manuel Reinout
ENI
Amin
Julien
Negar
Jessica
Lydia
Mark
Jane
Denis Voskov
Ralf
Bill Rossen
Dries Hegen
Reinaldo
IBM
Karel
MV
Marco Cardoso
Dirce
Carlos
Anneke
Jiakun
Anna
Nívia
Pierre-Olivier
Cesar
Mojtaba
Richard Braal
Saskia
Siavash
ICA
Paul
Maryam
Jakolien
Marja
Andreas
Tudor
Marc
Meeneesh
Mohammed
Fu Kai
Dana
TU Delft
Ricardo Ahmed
David Kremer
Diego
Pacelli Zitha
Paul Gelderblom
Laura
Siep Weiland
Marcia
Guus
Axel
Elisa
Hannie
Jan Dirk Jansen
Mehdi
Manu
Barry
Liene
Bets
Nikita
Christian
Michel
Alexis
Roozbeh
PUC-Rio
Steven
Rasoul
Statoil
Chris
Sintef
Gosia Kaleta
Marlijn
Graduate School
David Echeverría
Pedro Cesar
ISAPP
Remus Hanea
Claudia
Kevin
Kees
Marijke
Patrícia
Rémi
Ellis
Bander
Cândida
Dave
Boris
Suzanne
Martijn
Ernst
Daniel
Mohsin Yohei
Vinícius
Joost
Flávia Pacheco
Hadi Hajibeygi
Rahul
Maryia Krymskaya
Marcella
Sian
Tianqi
ECL
Elham
Régis Romeu
Vinh
Repsol
Brandon
Leon
Gabriel
Matei
Frank Wilschut
Femm
Petra
Rodrigo
Stefan
Durgesh
Paul Van den Hof
Nathalino
Jorge
Anne
Gabriela
Pedro
Marco Aurélio Pacheco
Swej
Sören
Femke Vossepoel
Pascalle
Michele
Margot
Colégio de São Bento
Mishal Dholakia

List of publications

Journal publications

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.

Barros, E.G.D., Van den Hof, P.M.J. and Jansen, J.D. (2018). Clustering techniques for value-of-information assessment in closed-loop reservoir management. Submitted for publication.

Barros, E.G.D., Van den Hof, P.M.J. and Jansen, J.D. (2018). Informed production optimization in hydrocarbon reservoirs. Paper in preparation.

Barros, E.G.D., Echeverría Ciaurri, D. and Jansen, J.D. (2018). Informed field development optimization. Paper in preparation.

Patents

Embid Droz, S.M., Rodríguez Torrado, R., Echeverría Ciaurri, D. and **Barros, E.G.D.** (2017). Method for determining a drilling plan for a plurality of new wells in a reservoir. Spanish Patent ES P201700314, filed March 29, 2017.

Conference proceedings

Barros, E.G.D., Jansen, J.D. and Van den Hof, P.M.J. (2014). Value of information in closed-loop reservoir management. Paper presented at the 14th European Conference on Mathematics in Oil Recovery, Catania, Italy, 8-11 September.

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. Paper presented at the 2nd IFAC Workshop on Automatic Control in Offshore Oil and Gas Production, Florianópolis, Brazil, May, 27-29.

Barros, E.G.D., Leeuwenburgh, O., Van den Hof, P.M.J. and Jansen, J.D. (2015). Value of multiple production measurements and water front tracking in closed-loop reservoir management. Paper SPE-175608-MS presented at the SPE Reservoir Characterization and Simulation Conference and Exhibition, Abu Dhabi, UAE, 14-16 September.

Barros, E.G.D., Yap, F.K., Insuasty Moreno, E., Van den Hof, P.M.J. and Jansen, J.D. (2016). Clustering techniques for value-of-information assessment in closed-loop reservoir management. Paper presented at the 15th European Conference on Mathematics in Oil Recovery, Amsterdam, The Netherlands, 29 August - 1 September.

Curriculum vitæ

Eduardo Barros

26-10-1987 Born in Rio de Janeiro, Brazil

Education

1997-2005 Secondary school
Colégio de São Bento do Rio de Janeiro, Brazil

2006-2012 Double-degree program
BSc in Petroleum Engineering
Pontifícia Universidade Católica do Rio de Janeiro, Brazil
MSc in Engineering
Ecole Centrale de Lyon, France

2012-2017 PhD in Petroleum Engineering
Technische Universiteit Delft

Professional experience

2017-present Reservoir Engineering scientist
TNO