

Computationally efficient techniques for well control and placement optimization under geological uncertainty

Yazan Arouri, B. Eng. (Hons)

A thesis submitted for the degree of Doctor of Philosophy (Ph.D.)

Australian School of Petroleum and Energy Resources

Faculty of Engineering, Computer & Mathematical Sciences

The University of Adelaide



April 2022

In the Name of The Most High and The Most Beneficent

Table of Contents

Abstract.....	i
Declaration.....	iv
Acknowledgement	v
Thesis by publication	vi
1. Introduction.....	1
1.1 Problem Statement	1
1.2 Thesis Structure	4
1.3 Contributions of each publication to this thesis	4
2 Literature Review.....	8
2.1 Well Control Optimization.....	8
2.1.1 Gradient-based techniques	9
2.1.2 Derivative-free techniques	11
2.1.3 Use of surrogates	12
2.2 Well Placement and Trajectory Optimization.....	13
2.2.1 Derivative-free techniques	14
2.2.2 Gradient-based techniques	15
2.2.3 Use of surrogates	16
2.3 Joint Well Placement and Well Control Optimization.....	16
2.3.1 Gradient-based techniques	18
2.3.2 Derivative-free techniques	18
2.3.3 Hybrid techniques	19
2.3.4 Use of proxies	19
2.4 Optimization under Geological Uncertainty	20

3	An accelerated gradient algorithm for well control optimization	24
4	Efficient techniques for well placement optimization	44
4.1	An adaptive moment estimation framework for well placement optimization	44
4.2	A study of simulation-based surrogates in well placement optimization for hydrocarbon production	63
5	Bilevel optimization of well placement and well control settings assisted by capacitance-resistance models	79
6	Adaptive rank-based selection of geological realizations for optimum field development planning	108
7	Summary, Conclusions and Future Work	109
	Bibliography	115

Abstract

Hydrocarbon field development plans outline the specific exploitation strategy with the aim of maximizing economic returns over the lifetime of the resource. These plans encompass a wide array of decision variables, including well location and trajectory, drilling schedule, well type, platform location, well control settings (injection/production rates and/or bottom-hole pressure), amongst other considerations. Given the highly nonlinear relationship between these development variables and the production volumes (and hence the economic returns), optimization techniques are applied to find the most optimal solution (i.e., field development plan).

However, it is essential that any proposed technique is pragmatic and consider the computational cost of the optimization algorithms. Even with improvements in computing hardware, realistic reservoir models can still be computationally demanding. Hence, the implementation of optimization algorithms into reservoir engineering workflows may be hindered if optimization process is computationally intense. To this end, this thesis investigates the development and implementation of state-of-the-art techniques for the efficient optimization of well control settings and well location under geological uncertainty.

To begin, the thesis develops and implements a novel gradient-based optimization algorithm, Adam-SPSA, for high-dimensional well control problems. The proposed algorithm combines the adaptive moment estimation framework with simultaneous perturbation stochastic approximation. The adaptive moment framework utilizes first-order gradient information to generate dimension-wise step-sizes. This allows for faster convergence to an optimum. The proposed algorithm is applied to two well control problems. The first being a two-dimensional heterogeneous reservoir model produced under water-flooding with four producers and four injectors. The second problem is a three-dimensional model produced through 20 production wells and 10 injection wells. The developed algorithm, Adam-SPSA, achieved improvements of up to 91% in convergence speeds and up to 5% improvements in objective function value when compared to the popular steepest descent framework.

The promising results of Adam-SPSA when applied to well control problems encouraged the investigation of applying the algorithm to the well placement and trajectory optimization problem. This investigation was done under the premise of a stringent computational budget and the availability of a heuristic-based initial guess. This takes advantage of the algorithm's ability to efficiently converge to local optimum from a suitable starting point. The performance of Adam-SPSA was demonstrated through the application to two experimental problems. The first problem studied the placement of four vertical wells, resulting in a low-dimensional problem of only eight decision variables. The second problem was the placement of 20 nonconventional (i.e., deviated, horizontal and/or slanted) wells, resulting in a 120-dimensional optimization problem. The proposed algorithm, Adam-SPSA, consistently outperformed the steepest descent framework and a local derivative-free algorithm (i.e., generalized pattern search). The work was then expanded further with an in-depth discussion surrounding the effect of parameterization on optimization performance as well as constraint handling techniques within gradient approximations.

An alternate approach was undertaken for the efficient optimization of well placement under strict computational budgets. The study investigated the use of a surrogate-based optimization approach, utilizing manifold mapping as a two-stage treatment, for well placement optimization. Manifold mapping was coupled with multiple surrogates including analytical – kriging and quadratic approximation – and a physics-based (reduced-order model) surrogate – local grid coarsening. The methodology was applied to two experimental problems, including the placement of four production wells in the presence of two pre-existing production wells and the placement of five production wells in a more complex three-dimensional model undergoing water-flooding. The proposed approach showed a reduction of computational costs of up to 80% compared to a local derivative-free algorithm. The results also gave insights into the most applicable scenario for the use of analytical surrogates and physics-based surrogates. It was shown that analytical surrogates are sufficient for simple reservoir models undergoing primary depletion. However, as the complexity of the reservoir models increases, such as secondary recovery through water-flooding, higher fidelity physics-based surrogates are more applicable as their accuracy is sufficient enough to model the trends of the full-physics simulations.

A natural progression was the joint optimization of well control settings and well placement concurrently. However, this type of problem is very computationally intensive as it requires a large number of full-physics reservoir simulations for convergence. As such, a novel technique was proposed which included the use of capacitance resistance models (CRM) to improve the efficiency of the optimization. This was implemented in a bilevel approach for the simultaneous optimization of well controls and well locations. The outer loop was the well placement optimization solved by particle swarm optimization. The inner loop was the well control optimization solved by Adam-SPSA assisted by CRMs. The proposed approach was tested against the full-physics approach on two reservoir models of varying levels of complexity. The proposed bilevel approach found solutions that were up to 22% higher in objective function value than the conventional full-physics approach and accompanied by a decrease of up to 99% in the number of required reservoir simulations.

The last research gap investigated in this thesis relates to the efficient incorporation of geological uncertainty in field development optimization problems. Typically, a full set of geological realizations (in the order of 100s to 1000s) are available to do this, however; including the full ensemble into an optimization problem makes it intractable. It was argued that previous work focussed on an intermediate goal of ranking base-case scenarios, static properties or a combination of these to select a subset of realizations. A reformulation of the subset selection problem to one which aims at ensuring consistent ranking of development strategies between the full set and the selected subset of realizations was introduced. The developed technique is a more relaxed problem and is not restricted in application. An application to well placement under uncertainty resulted in a reduction of computational costs, on average, by a factor close to 9, compared to the full set optimization. This result did not compromise the quality of the solution either.

Declaration

I certify that this work contains no material which has been accepted for the award of any other degree or diploma in my name, in any university or other tertiary institution and, to the best of my knowledge and belief, contains no material previously published or written by another person, except where due reference has been made in the text. In addition, I certify that no part of this work will, in the future, be used in a submission in my name, for any other degree or diploma in any university or other tertiary institution without the prior approval of the University of Adelaide and where applicable, any partner institution responsible for the joint award of this degree.

The author acknowledges that copyright of published works contained within the thesis resides with the copyright holder(s) of those works.

I give permission for the digital version of my thesis to be made available on the web, via the University's digital research repository, the Library Search and also through web search engines, unless permission has been granted by the University to restrict access for a period of time.

I acknowledge the support I have received for my research through the provision of an Australian Government Research Training Program Scholarship.

Yazan Arouri

17/02/2022

Acknowledgement

In the name of Allah, the Most Gracious and the Most Merciful.

First and foremost, I thank God for giving me the strength and ability to finish this thesis. I thank Him for all the blessings bestowed upon me over the course of my life.

Next, I thank the most important people in my life. I am forever indebted to my family for their constant love and support throughout my PhD. To my Baba and Mama, I thank you both for the never-ending love and support I have felt not only during my PhD but throughout my life. Your belief in me has kept me going in the most difficult of times. To my siblings, Laith, Yasmin and Roshdi, I thank you all for giving me the laughs and support I needed when even I didn't know I needed it. I hope I have made you all proud.

I also extend my sincere gratitude to my supervisors for their consistent support throughout my research. Mohammad has been a reassuring supervisor and has always helped me get out of any research slump I found myself in. He has allowed me to turn my passions into high-quality research that I can be proud of. Steve has been selfless with his time in providing motivating discussions and a unique perspective throughout my research.

A special thank you to those who I've met along this long and tortuous, yet fulfilling journey. All my colleagues at the Australian School of Petroleum: Abdullah, Abofazl, Cuong, Gabriel, Grace, Marat, Mike, Monica, Nassim, Reza, Roozbeh, Shahdad, Shuyan, Thomas and (last but not least) Yutong. We created a real work family. I really enjoyed and looked forward to our coffee time and lunch breaks. We celebrated many occasions together.

Thesis by publication

This is a thesis by publication and is composed of multiple pieces of work. This includes 5 published peer-reviewed journals papers. These are detailed below.

Published Peer-Reviewed Journal Papers

Arouri, Y. and M. Sayyafzadeh, An accelerated gradient algorithm for well control optimization. Journal of Petroleum Science and Engineering, 2020. 190: p. 106872.

Arouri, Y. & Sayyafzadeh, M. An adaptive moment estimation framework for well placement optimization. Computational Geosciences, 2022.

Arouri, Y.; Sayyafzadeh, M. & Begg, S. Adaptive rank-based selection of geological realizations for optimum field development planning. SPE Journal, 2022.

Arouri, Y.; Echeverría Ciaurri, D. & Sayyafzadeh, M. A study of simulation-based surrogates in well-placement optimization for hydrocarbon production. Journal of Petroleum Science and Engineering, 2022.

Arouri, Y.; Lake, L. W. & Sayyafzadeh, M. Bilevel optimization of well placement and control settings assisted by capacitance resistance models. SPE Journal, 2022.

1. Introduction

1.1 Problem Statement

Once a resource has been evaluated to contain economic volumes of hydrocarbons, the next task involves designing the most cost-effective, sustainable, and profitable extraction plan. This culminates in a field development plan (FDP) which includes details regarding decisions on well locations, well trajectories, well control settings, production platform system and surface facility design, amongst other important considerations. All these various decisions have a substantial influence on the recovery of hydrocarbon fluid volumes from the reservoir.

Reservoir simulation is an important predictive tool that is used to assist with the evaluation of proposed FDPs and is used to perform what-if analyses. The governing equations of reservoir simulation encompass the highly nonlinear relationship between the input parameters, such as the development parameters (e.g., well control settings and well locations) and reservoir fluid and rock properties, and the output parameters, such as produced fluid volumes. These simulations are computationally demanding, especially for more complex geological settings, with potential run-times in the order of hours to days for realistic reservoir models.

The interaction and interdependency of the large number of decision variables makes designing and selecting an optimal FDP a non-trivial task. The problem is further complicated with the incorporation of geological uncertainty, which is typically done with the consideration of a large number (in the order of 100s to 1000s) of geological realizations. In essence, each proposed development strategy (e.g., well location and/or well control settings) is evaluated and averaged against all the geological realizations included in the set. For each development strategy, a full-physics reservoir simulation is needed to predict the associated fluid volumes. This increases the complexity and becomes for a large number of realistic reservoir models.

Consequently, mathematical optimization techniques have been proposed to assist in proposing and evaluating potential FDPs that are robust and optimal in a systematic manner. This is accomplished

by transforming the field development planning task into a general optimization problem in which an objective function (e.g., oil recovery or net present value) is, without loss of generality, minimized for a given set of defined decision variables (e.g., well control settings and/or well locations/trajectories) and constraint functions. Depending on the defined field development optimization problem, the number of decision variables will vary. Typically, well location problems involve 10s to 100s of decision variables to define well placements in two-dimensional or three-dimensional. On the other hand, well control optimization problems are composed of 100s to 1000s of decision variables as they are usually defined at specified time intervals across the field's lifetime. When these problems are considered in a joint manner, the decision variables may be different in nature (i.e., continuous and discrete) and be in the order of 1000s in number.

The underlying premise for solving this problem is based on using optimization algorithms which seek iterative improvements in the objective function value as the solution space is traversed. The FDP optimization problem is simulation-based and as such each objective function evaluation entails running one or more computationally expensive reservoir simulation/s. This brings into question the feasibility of seamlessly implementing optimization techniques to the workflows of field development planning. Vitrally, the computational efficiency of any proposed optimization technique for field development planning must be of focus. In other words, any proposed optimization technique should be able to propose optimal solutions for the FDP problem without becoming too computationally demanding. However, this must be done without compromising the quality of the proposed solutions.

In this work, the aim is to develop and implement optimization techniques for various field development planning tasks with a specific focus on computational efficiency. Firstly, a gradient-based optimization algorithm, Adam-SPSA, is developed to accelerate the optimization of well control settings in high-dimensional solution spaces. This is done through the use of an adaptive moment estimation framework which utilises first-order gradient information to calculate dimension-wise search directions. The proposed algorithm was compared to contemporary methods, including a gradient-based algorithm and a local derivative-free algorithm.

Next, two approaches were investigated for the efficient optimization of well placement and trajectories. Both these approaches were developed for situations where a stringent computational budget exists (e.g., limited number of software licenses, short time-frame, etc.). The first approach extended the use of Adam-SPSA to the well placement and trajectory problem to take advantage of its accelerated convergence to suitable solutions in a limited number of reservoir simulations. To do this efficiently, the application of Adam-SPSA is done under the premise of a heuristic (reservoir engineering) based initial guess.

The second approach tackled the well placement problem by simplifying the use of reservoir simulations with fast-to-evaluate surrogates. In this way, a suitable solution is found using a lower number of reservoir simulation calls. Specifically, the use of a two-level surrogate-based technique, known as manifold mapping, was investigated. This technique is a surrogate correction technique that iteratively improves the quality of the surrogate before optimizing it. The use of various types of surrogates is investigated including analytical surrogates and physics-based surrogates.

An intuitive progression is the consideration of both well control settings and well placement in a joint manner. Previous studies have shown an improved solution compared to the consideration of each optimization task separately. However, this comes at the expense of a significant increase in reservoir simulation calls. Consequently, the use of bilevel optimization approach with the outer loop consisting of the well placement problem and the inner loop consisting of the well control problem assisted by capacitance-resistance models (CRMs) was proposed.

Lastly, an adaptive rank-based subset selection technique for the efficient incorporation of geological uncertainty in field development optimization problems was developed. In this approach, the subset selection problem was reformulated to one which aims at ensuring consistent ranking of the development strategies between the full set and the selected subset of realizations. The developed technique is a more relaxed problem and is not restricted in application.

1.2 Thesis Structure

This is a thesis by publication. Chapter 1 begins with a problem statement to provide context and describe the aim of this thesis. It also presents the contribution of each publication to the thesis. Chapter 2 presented an in-depth literature review which encompasses well control optimization, well placement optimization, joint optimization of well control and well placement, and optimization under geological uncertainty. The subsequent chapters, including Chapters 3, 4, 5 and 6 are composed of published journal papers. The thesis is completed with Chapter 7 giving a summary and conclusions. Table 1 gives an overview of the Chapter numbers, their titles, the journal papers contributing to them and their publication status.

Table 1: Thesis structure

Chapter Number	Chapter Title	Paper number	Status
3	An accelerated gradient algorithm for well control optimization	1	Published
4	Efficient techniques for well placement optimization	2	Published
		3	Published
5	Bilevel optimization of well placement and well control settings assisted by capacitance-resistance models	4	Published
6	Adaptive rank-based selection of geological realizations for optimum field development planning	5	Published

1.3 Contributions of each publication to this thesis

The overarching theme of this thesis is the investigation, development and implementation of optimization algorithms to large-scale field development problems with a key focus on computational efficiency. Various algorithms have been proposed in the literature to solve the well control optimization problem, however; there has been a greater focus on finding the best solution and less focus on the computational efficiency of the optimization frameworks. This thesis is comprised of five journal publications each with a specific contribution to the primary goal.

In the paper titled “*An accelerated gradient algorithm for well control optimization*”, a gradient-based algorithm, Adam-SPSA, is developed and implemented to large-scale well control optimization under uncertain conditions. This is the first application of Adam to reservoir engineering problems. The

optimization framework is based on estimates of the first (mean) and second (variance) moments of the gradient approximations. In addition, the variable-specific (dimension-wise) nature of the search direction ensures that the search progression is tailored for each variable. When applied to well control optimization problems, the novel algorithm (Adam-SPSA) was more computationally efficient than the popular steepest descent (SD) SPSA - without compromising the optimal solution found. Furthermore, an in-depth investigation into the effect of different constraint handling techniques on the optimization was undertaken. These investigations and findings constitute Chapter 3 of this thesis.

Following from the success of the developed gradient-based algorithm in Chapter 3, its application was extended to well placement optimization in the paper titled “*An adaptive moment estimation for well placement optimization*”. The objective of this paper was to take advantage of the known characteristics of Adam-SPSA, including the fast convergence to a local optimum and the dependence on the initial starting point, for certain practical applications in well placement optimization. In practical scenarios, the availability of resources, whether that be human resources or hardware/software resources, is limited. This paper makes the underlying assumptions that there is a stringent computational budget (i.e., limited number of reservoir simulations) and there exists a reservoir engineering judgement that is a suitable starting point for the optimization procedure.

In this paper, it is argued that for such scenarios it becomes vital to take advantage of the local efficiency of gradient-based algorithms to produce solutions which maximise the improvement on the initial guess within the allotted computational budget. An investigation is also presented detailing the reasons behind the improvements presented in the search directions of Adam compared to the steepest descent framework, visually showing the ability of Adam-SPSA to evade local minimums in search for a better optimum. To the author’s knowledge, this is the first time this has been done for gradient-based algorithms in well placement optimizations. In the paper the proposed algorithm (Adam-SPSA) is compared with a gradient-based algorithm (steepest descent SPSA) and a local derivative-free algorithm (generalized pattern search) for the placement of vertical and nonconventional wells to show the computational efficiency of the proposed methodology. This paper also included additional discussions

regarding nonconventional parameterization and the effect of constraint handling techniques on simultaneous perturbation gradient approximations.

Another approach to improve the computational efficiency of well placement optimization is through the use of surrogates, as was investigated in the paper titled “*A study of simulation-based surrogates in well placement optimization for hydrocarbon production*”. Here, a surrogate-treatment method, manifold mapping, was used to iteratively correct a surrogate (or proxy) during the optimization procedure. The underlying aim is to optimize the corrected and fast-to-evaluate surrogate rather than using full-physics reservoir simulations, which are significantly more computationally demanding. As the optimization progresses, the surrogate is able to converge to a local optimum using a lower number of reservoir simulation calls. In this study, both analytical and physics-based surrogates were investigated to see the advantages and disadvantages of each type. It was found that in cases of simple recovery mechanisms (i.e., natural depletion), analytical surrogates (e.g., kriging and quadratic approximation) are sufficiently accurate to converge to a suitable solution more quickly than a local derivative-free algorithm. However, as the complexity of the recovery mechanism and the reservoir model increase, for example in water-flooding scenarios of three-dimensional models, a physics-based surrogate (e.g., local grid coarsening) provides the necessary accuracy to find an improved solution with a reduction in computational cost. These two papers make up Chapter 4 of this thesis.

A natural progression for the research was to consider both well placement and well control optimization in a joint fashion. The main impediment of current approaches proposed in the literature for solving joint optimization problems is that they are very computationally demanding. That is, they require a prohibitive number of reservoir simulation calls, which can become intractable for large, complex reservoir models.

In the paper titled “*Bilevel optimization of well placement and control settings assisted by capacitance resistance models*”, a nested approach using capacitance resistance models is proposed to solve the joint optimization problem. The proposed bilevel approach is composed of an outer loop for well placement optimization and an inner loop for well control optimization. In the proposed framework, capacitance resistance models are used as a surrogate for reservoir simulations in the inner

loop (i.e., well control optimization) to significantly reduce the number of required reservoir simulations. This paper is the first known use of capacitance resistance models as surrogates in the joint optimization problem. The proposed approach showed significant reduction in the number of required reservoir simulations to converge to the same solution as the contemporary approach. This paper comprises Chapter 5 of this thesis.

An important aspect of field development planning is the incorporation of geological uncertainty in the optimization problem. However, the incorporation of uncertainty through the use of an ensemble of geological realizations, in the order of hundreds to thousands, becomes computationally prohibitive. To this end, in the paper titled “*Adaptive rank-based selection of geological realizations for optimum field development planning*”, a new subset selection technique for field development optimization under uncertainty is developed and implemented. It is argued that the ultimate goal of any subset selection technique is to ensure the proposed field development plans are consistently ranked compared to the ranking based on the full ensemble of realizations. This reformulation of the subset selection problem is expressed mathematically. A technique is proposed which selects a subset of realizations that minimizes the difference between the rankings of development strategies obtained by the subset and the full set. The underlying idea in the proposed technique is that a subset of realizations selected based on a small batch of development strategies (i.e., solutions) is accurate when ranking other development strategies. Comparison with other subset selection techniques and implementation for well placement optimization demonstrated the computational savings of the proposed approach.

The five papers presented represent an effort to develop and implement computationally efficient techniques for large-scale field development problems including well control and well placement/trajectory optimization under uncertainty. The papers collectively provide insight into state-of-the-art algorithms and approaches that focus on reducing the computational costs of field development optimization to allow practical implementation in everyday workflows.

2 Literature Review

2.1 Well Control Optimization

The well control optimization problem, also known in the literature as production optimization, is the assignment of injection/productions rates and/or bottom-hole pressures (BHPs) over pre-defined or adaptive time-steps (or control intervals) to maximize/minimise an objective function. The decision variables, injection/productions rates and/or BHPs, are continuous, piecewise and temporal. Typically, well control optimization problems are high-dimensional and composed of 100s and 1,000s of decision variables. The well locations are pre-defined and do not change throughout the optimization process. The objective function landscapes of well control optimization problems are characterized by long flat valleys and hence are relatively less rugged and smoother (Do and Reynolds, 2013, Fonseca et al., 2014b, Zhao et al., 2013). The well control optimization problem can be formulated mathematically as the following:

$$\begin{aligned} & \min_{\mathbf{u} \in \mathbb{R}^{n_u}} f(\mathbf{u}) \\ & \text{subject to:} \\ & c_i(\mathbf{u}) < 0, \mathbf{i} \in K \\ & c_i(\mathbf{u}) = 0, \mathbf{i} \in I, \end{aligned} \tag{1}$$

where, $f(\mathbf{u})$ is the objective function to be maximized/minimized, \mathbf{u} is the vector of decision variables (i.e., well control settings), n_u is the number of decision variables, I and K are the sets of indices for equality and inequality constraint function, $c(\mathbf{u})$, respectively. These constraint functions can include simple bound constraints as well as, linear and nonlinear constraints. Examples of bound constraints would be upper and lower limits on injection/production rates or BHPs based on field operational capacities. Examples of equality and inequality constraints include field rate constraints, total liquid production constraints, and total injection rate constraints. To solve the well control optimization problem both gradient-based algorithms and derivative-free algorithms – some assisted with surrogates – have been applied.

2.1.1 Gradient-based techniques

The architecture of the objective function landscape for well control optimization problems lends themselves to be better suited for gradient-based algorithms, due to their relatively smoother landscapes (Zhao et al., 2013, Fonseca et al., 2014b, Do and Reynolds, 2013). A number of gradient-based algorithms have been applied to solve the well control optimization problem. They can be categorized by the type of gradients used and the optimization framework employed.

The two major categories of gradients that have been utilized in well control optimization are adjoint gradients and gradient approximation methods. Adjoint formulations have roots in optimal control theory and are an efficient way to calculate gradients. The efficiency of adjoint-based methods comes from the fact that the Jacobian matrices from the forward simulation are used to find the sensitivities of the objective function to the control (decision) vector. This also means that the gradient calculation of the objective function is independent of the number of control variables (Jansen, 2011). Adjoint-based methods have been applied to tertiary recovery using surfactant flooding (Ramirez et al., 1984, Fathi and Ramirez, 1984). Other early applications include carbon dioxide flooding (Mehos and Ramirez, 1989), steam flooding (Liu et al., 1993), and most commonly water flooding (Asheim, 1988).

A resurgence of interest in adjoint-based well control optimization occurred after the invention and application of smart wells, which allowed the control of downhole well settings (Carvajal et al., 2018). This includes the works of Brouwer and Jansen (2004), Sudaryanto and Yortsos (2001), Brouwer et al. (2004), Sarma et al. (2005), Suwartadi et al. (2012) and (Forouzanfar et al., 2013). Although they are efficient in calculating gradients, adjoint methods are considered intrusive as they require access to reservoir simulation source code, which may be unavailable for commercial simulators. In addition, adjoint gradients are not trivial to implement and maintain.

As an alternative to the intrusive adjoint-based methods, other gradient approximations methods have been proposed. The simplest method used for well control optimization is the finite difference method (Yeten et al., 2004, Asadollahi et al., 2014, Isebor, 2009, Wang et al., 2009). However, such a technique is not computationally efficient for high-dimensional well control problems

as it requires $2n$ function evaluations (reservoir simulations), where n is the number of decision variables. Another gradient approximation which has found success in its application to well control optimization problems is the simultaneous perturbation stochastic approximation (SPSA). These works include Wang et al. (2009), Do and Reynolds (2013) and Foroud et al. (2018). Other techniques such as ensemble-based optimization (EnOpt) (Chen et al., 2009, Chen and Reynolds, 2017, Leeuwenburgh et al., 2010, Dehdari and Oliver, 2012, Chaudhri et al., 2009, Dehdari et al., 2012) and stochastic simplex approximate gradient (StoSAG) (Fonseca et al., 2017, Lu et al., 2017b, Liu and Reynolds, 2020, Fonseca et al., 2014b, Silva et al., 2020) incorporate geological uncertainty in the gradient approximation.

Applications of gradient-based algorithms can also be categorized into the optimization framework used, which includes steepest descent (ascent), conjugate method and sequential quadratic programming. Due to its simplicity in implementation, the most common framework is the steepest descent (ascent) method (Chen et al., 2009, Leeuwenburgh et al., 2010, Brouwer and Jansen, 2004, Wang et al., 2009, Brouwer et al., 2004). This framework utilizes the negative of the gradient as the search direction to progress the search. In addition, backtracking line search is also coupled with steepest descent to find an appropriate step size (Nocedal and Wright, 2006). However, this introduced additional parameters and requires additional objective function evaluations which may become costly. Various gradient calculations/approximations have been combined with the steepest descent (ascent) framework, including adjoint formulations (Brouwer and Jansen, 2004, Brouwer et al., 2004), SPSA (Wang et al., 2009, Foroud et al., 2018, Do and Reynolds, 2013), finite difference (Wang et al., 2009) and EnOpt (Chen et al., 2009, Leeuwenburgh et al., 2010, Forouzanfar et al., 2013).

Another optimization framework that has been implemented in well control optimization is the conjugate gradient framework (Yeten et al., 2004, Asadollahi and Naevdal, 2009, Chaudhri et al., 2009), however; improvements in computational efficiency over steepest descent have been modest. In this technique, a search direction is a weighted combination of the current gradient and the gradient in the previous iteration. Adjoint formulations (Asadollahi and Naevdal, 2009), EnOpt (Chaudhri et al., 2009)

and finite difference (Yeten et al., 2004) have been used as to approximate or calculate the gradient in the conjugate gradient framework.

The sequential quadratic programming (SQP) framework has also been applied to well control optimization problems as a means to improve the handling of nonlinear constraints (Liu and Reynolds, 2020, Dehdari and Oliver, 2012, Asadollahi et al., 2014, Lorentzen et al., 2009). In this method, at each iteration a quadratic sub-problem that incorporates a Lagrangian formulation is solved to obtain a search direction. Although this method incorporates nonlinear constraints through the use of the Lagrangian, it requires a large number of function evaluations, making it less computationally efficient than other methods. A number of different gradient approximations have been used within the SQP framework, including StoSAG (Liu and Reynolds, 2020), EnOpt (Dehdari and Oliver, 2012) and finite difference (Lorentzen et al., 2009, Asadollahi et al., 2014).

2.1.2 Derivative-free techniques

Derivative-free optimization (DFO) algorithms have also been applied – to a lesser degree – to well control optimization problems. These types of algorithms rely only on objective function values to search the solution space. Although they search more globally compared to gradient-based algorithms, it comes with an increased number of function evaluations to converge to a solution. This issue intensifies in high-dimensional settings as is typically the case in well control optimization. Derivative-free optimization techniques can be classified into two main categories: deterministic methods and stochastic methods. Deterministic methods include stencil-based algorithms, such as Hooke-Jeeves direct search (HJDS) (Hooke and Jeeves, 1961), Nelder-Mead (or polytope) (Nelder and Mead, 1965) and generalized pattern search (GPS) (Torczon, 1997). Stochastic methods include particle swarm optimization (Kennedy and Eberhart, 1995), genetic algorithms (GA) (Srinivas and Patnaik, 1994) and covariance matrix adaptation – evolutionary strategy (Hansen, 2006).

Various papers have compared different derivative-free algorithms for well control optimization. Ciaurri et al. (2010) compared HJDS, GPS, GA and SQP for the optimization of a 20-dimensional well control problem with general constraints handled using either the penalty method or

filter method. The gradient-based approach (SQP using finite difference) performed efficiently, however; the authors claim the selection of a suitable perturbation size can be problematic. Hence, the authors recommended HJDS when distributed computing resources are not available. Asadollahi et al. (2014) compared HJDS, GPS, Nelder-Mead, SQP and HJDS with a gradient-based line-search hybrid for the well control optimization of the Brugge benchmark. Wang et al. (2019) compared GA, PSO and CMA-ES for the optimization of 128 well control variables using the PUNQ-S3 benchmark.

2.1.3 Use of surrogates

Surrogates, also known as proxy models or meta-models, have also been applied to well control problems to improve the computational efficiency of the optimization procedure. This is done by replacing expensive high-fidelity reservoir simulations with lower fidelity, but fast-to-evaluate, alternatives. There have been applications of both response surface modelling and reduced order modelling to well control optimization problems.

Response surface surrogates are trained and calibrated based on a number of previously evaluated well control scenarios using reservoir simulations. Zhao et al. (2013) used approximated gradients from ensemble-based techniques and SPSA to optimize a quadratic interpolation model for well controls. Golzari et al. (2015) implemented an artificial neural network (ANN) in combination with GA to optimize the well controls of four production wells in the PUNQ-S3 benchmark. Echeverria Ciaurri and Wagenaar (2016) applied a corrective surrogate treatment, manifold mapping, in conjunction with surrogates based on generalized barycentric coordinates and quadratic approximation for well controls.

Reduced order modelling is typically based on reducing the underlying physics or the numerical calculations. Proper orthogonal decomposition and its variants, such as missing point estimation (Cardoso et al., 2009), trajectory piecewise linearization (Cardoso et al., 2009, Jansen and Durlofsky, 2017) and the discrete empirical interpolation method (Alghareeb, 2015, Suwartadi et al., 2015) have been applied to well control optimization. Other reduced order surrogates include streamline simulations and capacitance resistance models (CRMs). Streamline simulations have also been used to optimize well controls of mature fields in Wen et al. (2014). Capacitance resistance models have been

used to in conjunction with gradient-based ensemble techniques (Hong et al., 2017, Jafroodi and Zhang, 2011). Jafroodi and Zhang (2011) applied CRMs to optimize the well controls of five injectors and four producers for a simple synthetic reservoir, whilst Hong et al. (2017) optimized the control settings of eight injectors and four producers.

2.2 Well Placement and Trajectory Optimization

The well placement and trajectory optimization, also known as well location optimization, is the determination of the optimal well locations/trajectories subject to physical field constraints that maximize an objective function. In these problems, the well control settings (i.e., injection/production rates and/or producer BHPs) are kept constant throughout the optimization process. Due to the highly nonlinear relationship between the well locations and the produced fluid volumes, the objective function landscapes are characterized by nonconvex and highly multimodal architectures. Typically, well placement optimization problems are low-dimensional relative to well control optimization, however; due to the increase nonlinearities they are expected to have more rugged objective function landscapes with more minima. The well placement and trajectory optimization problem can be formulated as follows:

$$\begin{aligned} & \min_{\mathbf{x} \in \mathbb{R}^{n_x}} f(\mathbf{x}) \\ & \text{subject to:} \\ & c_j(\mathbf{x}) < 0, \mathbf{j} \in H \\ & c_j(\mathbf{x}) = 0, \mathbf{j} \in V, \end{aligned} \tag{2}$$

where, $f(\mathbf{x})$ is the objective function to be maximized/minimized, \mathbf{x} is the vector of decision variables (i.e., well locations/trajectories), n_x is the number of decision variables, H and V are the sets of indices for equality and inequality constraint function, $c(\mathbf{x})$, respectively. The constraint functions for well placement optimization are fundamentally different than those for well control optimization.

Bound constraints for well placement problems could include reservoir boundaries, however; more sophisticated reservoir boundaries can be included through piecewise polynomial functions to constrain the placement of well in regions within irregular-shaped boundaries. Other well placement constraints include minimum inter-well distances and maximum well lengths.

Another important consideration for well placement and trajectory optimization problems is the parameterization of the wells. When optimizing the placement of two dimensional wells (i.e., vertical wells) they can be parameterized using continuous x and y coordinates or discrete i and j indices, which largely depends on the reservoir simulators ability to handle one over the other (Güyagüler and Horne, 2004, Bangerth et al., 2006, Badru and Kabir, 2003, Onwunalu and Durlofsky, 2010). When considering wells in three-dimensions, the parameterization can be done using spherical coordinates (i.e., x , y , and z for the heel and L , ϕ , and θ for the length of the well, inclination angle and azimuth angle, respectively) or Cartesian coordinates (i.e., x , y and, z for each of the heel and toe)(Yeten et al., 2003, Ding, 2008, Alrashdi and Sayyafzadeh, 2019, Bouzarkouna et al., 2012, Forouzanfar and Reynolds, 2013).

Global derivative-free optimization algorithms have lent themselves to be better suited for well placement optimization problems given their ability to traverse the landscapes more effectively. Gradient-based algorithms and surrogate-based approaches have also shown promise for well placement optimization in computationally constrained scenarios.

2.2.1 Derivative-free techniques

Derivative-free optimization algorithms use objective function values to direct the search by updating the algorithmic parameters to propose new solutions at each iteration. A number of various derivative-free algorithms have been applied to well placement optimization problems. Genetic algorithms (GA) have been popular for solving the well placement optimization problem for vertical wells (Bangerth et al., 2006, Güyagüler and Horne, 2004, Ozdogan and Horne, 2006, Sayyafzadeh, 2017, Montes et al., 2001, Tupac et al., 2007) and nonconventional wells (Artus et al., 2006, Yeten et al., 2003, Emerick et al., 2009).

Another popular derivative-free optimization algorithm for well placement problems is particle swarm optimization (PSO), first applied by (Onwunalu and Durlofsky, 2010). Their work implemented PSO to a number of example problems including the optimization of 20 vertical wells, four deviated wells and two nonconventional wells with an unknown number of laterals. Their work showed

promising results when compared to GA for the example problems. Ding et al. (2014) applied a modified PSO in conjunction with a quality map to improve the solutions for well placement optimization. Jesmani et al. (2016a) applied PSO and focussed on the incorporation of physical field constraints. Wang et al. (2012) combined PSO with a simplex method for the optimization of well placement using a retrospective framework that incorporated geological uncertainty. Nwankwor et al. (2013) implemented a hybrid procedure combining PSO and differential evolution (DE) for the optimization of vertical wells in two and three-dimensional models.

Covariance matrix adaption – evolutionary strategy (CMA-ES) has been applied to well placement optimization in Ding (2008) and Bouzarkouna et al. (2012). Other algorithms applied to well placement optimization include simulated annealing (Bittencourt and Horne, 1997, Norrena and Deutsch, 2002), evolutionary strategy (Sayyafzadeh and Alrashdi, 2019, Alrashdi and Sayyafzadeh, 2019), imperialist competitive algorithm (Al Dossary and Nasrabadi, 2016), Bound Optimization BY Quadratic Approximation (Forouzanfar and Reynolds, 2013), modified cuckoo algorithm (Alghareeb et al., 2014), a hybrid of cat swarm optimization and mesh-adaptive direct search (Chen et al., 2018) and metaheuristic bat algorithm (Naderi and Khamehchi, 2017). These derivative-free optimization algorithms are characterised as being computationally intensive which typically require 1,000s of objective function evaluations (reservoir simulations) to converge.

2.2.2 Gradient-based techniques

Although gradient-based algorithms are considered local search methods, various approaches have been applied to well placement optimization. This includes the adjoint-based method, which utilizes an indirect method to approximate the sensitivity of various vertical well locations on the objective function (Zandvliet et al., 2008, Sarma and Chen, 2008, Zhang et al., 2010). Vlemmix et al. (2009) extended this approach to well trajectories in three-dimensional spaces by utilizing pseudo-sidetracks. Volkov and Bellout (2018) used the adjoint formulation to find key partial derivative terms and then used the finite difference method to approximate these terms. Bangerth et al. (2006) applied an integer variant of SPSA for the optimization of vertical wells, whilst Jesmani et al. (2016b) applied a continuous variant of SPSA for the optimization of a single nonconventional well. Wang et al. (2007) compared

SPSA to finite difference methods for the optimization of a water injection well in a two-dimensional case. Leeuwenburgh et al. (2010) implemented EnOpt for the optimization of nine vertical wells in a two-dimensional model.

2.2.3 Use of surrogates

The use of surrogates, also known as proxy models or meta-models, can improve the computational efficiency of well placement optimization by reducing the number of required high fidelity reservoir simulations. Güyagüler et al. (2002) combined GA with artificial neural networks (ANNs) and kriging, whilst Yeten et al. (2003) combined GA with ANN for the optimization of nonconventional wells. Kwon et al. (2021) used convolutional neural networks (CNNs), an advanced type of ANNs, to predict the cumulative oil production at each possible well location. Artus et al. (2006) implemented a statistical proxy that was built using k-means clustering. Bouzarkouna et al. (2012) combined CMA-ES with a local quadratic surrogate built on the k -nearest points to optimize two unilateral wells. In Zubarev (2009), multiple analytical surrogates, including kriging, thin-plate splines and ANNs, were investigated for the optimization of two vertical wells, however; only very simple problems were considered (the areal location of two vertical wells).

2.3 Joint Well Placement and Well Control Optimization

The architectures of the objective function landscapes for well placement and well control optimization problems are intrinsically different. This has to do with the decision variables and their nonlinear relationship with the output from reservoir simulations (i.e., produced fluid volumes). Consequently, as presented thus far, these two problems have been considered in silo of one another. However, more recently, a number of works have presented arguments that such an approach may lead to sub-optimal results (Bellout et al., 2012, Sayyafzadeh and Alrashdi, 2019). It is argued that there is an inherent interdependency between the well locations and their associated well control settings and as such their joint consideration would lead to more optimal solutions. However, this comes at the increased cost of overhead from additional required computationally expensive reservoir simulations.

The joint optimization of well placement and well control can be formulated as follows:

$$\begin{aligned}
& \min_{\mathbf{x} \in \mathbb{R}^{n_x}, \mathbf{u} \in \mathbb{R}^{n_u}} f(\mathbf{x}, \mathbf{u}) \\
& \text{subject to:} \\
& c_r(\mathbf{x}, \mathbf{u}) < 0, \mathbf{r} \in W \\
& c_r(\mathbf{x}, \mathbf{u}) = 0, \mathbf{r} \in T,
\end{aligned} \tag{3}$$

where, $f(\mathbf{x}, \mathbf{u})$ is the objective function to be maximized/minimized, \mathbf{x} is the vector of well placement decision variables (i.e., well locations/trajectories), \mathbf{u} is the vector of well control decision variables, n_x is the number of well placement decision variables, n_u is the number of well control decision variables, W and T are the sets of indices for equality and inequality constraint function, $c(\mathbf{x})$, respectively.

In the literature, there are two main categories of implementation for the joint optimization of well placement and well control. The first category is the sequential optimization of well placement and well controls. In this approach, either the well placement or well control problem is solved followed by the other using the optimal solution found for the first optimization. For example, one may solve the well placement problem to obtain optimal well locations and/or trajectories. Then, using these optimal well locations and/or trajectories, the well controls settings are optimized. This can be done once or in an iterative manner until no improvement in objective function value occurs.

The second approach is the simultaneous optimization of well placement and well controls. As indicated by the name, both well locations and/or trajectories and the well control settings comprise the decision variable vector. By virtue of the curse of dimensionality, this increase in the number of decision variables will increase the complexity of the optimization problem (Bellman, 1966).

Debate still exists in the literature regarding which implementation (sequential or simultaneous) provides more optimal solutions. There are several pieces of literature which argue that a sequential (usually iterative) approach is the preferred method over a simultaneous one (Humphries et al., 2014, Humphries and Haynes, 2015, Wang et al., 2016, Lu and Reynolds, 2020, Lu et al., 2017a). On the contrary, several pieces of work have argued that a simultaneous approach results in more optimal solutions as it incorporates all the decision variables concurrently (Isebor et al., 2014, Bellout et al., 2012, Forouzanfar et al., 2016, Sayyafzadeh and Alrashdi, 2019). Various types of techniques have

been developed and used for joint optimization problem including gradient-based algorithms, derivative-free algorithms or a hybrid of these.

2.3.1 Gradient-based techniques

There has been a number of works which use gradient-based techniques to solve the joint optimization both sequentially and simultaneously. Li and Jafarpour (2012) implemented the gradient-based SPSA algorithm in an iterative sequential manner to optimize the well locations and well control settings for the top layer of the SPE10 model and the PUNQ-S3 model. Salehian et al. (2021) also used the SPSA algorithm for the iterative sequential optimization of well locations and well controls in a multi-solution framework for the Brugge benchmark model. Li et al. (2013) used the SPSA algorithm in a simultaneous fashion and incorporated geological uncertainty in their example problems. Lu et al. (2017a) implemented the stochastic simplex approximate gradient (StoSAG) algorithm in an iterative sequential manner for the optimization of three-dimensional well segments and their associated well controls.

2.3.2 Derivative-free techniques

Derivative-free optimization algorithms have also found success when implemented for joint optimization. Forouzanfar et al. (2016) tested CMA-ES and a hybrid of CMA-ES with an ensemble-based technique for both sequential and simultaneous approaches. Their work found that CMA-ES using a simultaneous approach performed the best on both problem types (i.e., simultaneous or sequential). Wang et al. (2016) found that the use of a multilevel coordinate search algorithm was competitive against other derivative-free algorithms for the sequential and simultaneous problems. Sayyafzadeh and Alrashdi (2019) and Alrashdi and Sayyafzadeh (2019) used an evolutionary strategy (ES) for the simultaneous optimization of well placement in three-dimensions and well controls on various problems, including the three-dimensional Olympus benchmark. Awotunde (2014) implemented a dimensionality reduction technique to the well control problem for the simultaneous optimization using differential evolution.

2.3.3 Hybrid techniques

Hybrid optimization algorithms have found success in applications to joint optimization of well placement and well control. They have the ability to take advantage of both local and global search capabilities to traverse the solution space more efficiently than either gradient-based or derivative-free techniques separately.

Bellout et al. (2012) combined pattern search methods (HJDS and GPS) with sequential quadratic programming (SQP) with using adjoint gradients. They implemented joint optimization in a nested (i.e., bilevel) fashion, where each proposed well location underwent an inner loop of well control optimization. The outer well placement optimization was solved using the pattern search method and the inner well control optimization was solved using SQP. Another popular hybrid optimization algorithm is the combination of PSO with either mesh-adaptive direct search (MADS) or GPS, which has been implemented for both simultaneous and sequential joint optimization (Isebor et al., 2014, Humphries and Haynes, 2015, Humphries et al., 2014). Isebor et al. (2014) and Humphries et al. (2014) considered only vertical wells whilst Humphries and Haynes (2015) studied the placement of three-dimensional wells. Lu and Reynolds (2020) optimized well drilling path, drilling sequences and well type using a GA with mixed encodings combined with StoSAG for well control optimization.

2.3.4 Use of proxies

There is only a small number of literature which investigate the use of surrogates for joint optimization. Møyner et al. (2014) introduced the use of time-of-flight and tracer partition equations for the joint optimization on multiple synthetic reservoir models. The authors state that the aim of flow diagnostics isn't to achieve a fully automated framework that can find the true optimum, rather it is to provide a set of fast tools that can be used in existing workflows to improve on an initial configuration. Aliyev and Durlofsky (2017) introduced a multilevel framework for the joint optimization of well placement and controls. In this work, particle swarm optimization is used to optimize a sequence of models with increasing grid refinement at each iteration of the optimization. The upscaled models are obtained through a global single-phase method that generates upscaled well indices and transmissibilities. More recently, de Brito and Durlofsky (2021) implemented a nested approach, similar to the one presented in

Bellout et al. (2012), for the joint optimization using a set of surrogate treatments. The inner loop of well control optimization is divided further into two sub-problems each with a different objective function based on a different surrogate.

2.4 Optimization under Geological Uncertainty

The incorporation of geological uncertainty is an important consideration for reservoir optimization problems to ensure robust solutions are found and proposed. This is typically done through the use of an ensemble of geological realizations - in the order of hundreds or thousands. The evaluation of each proposed solution (i.e., development strategy) against the full ensemble of realization results in a cumulative distribution function (CDF) of the fitness metric. The fitness metric will depend on the decision-maker, but common ones are Net-Present-Value (NPV) or cumulative oil production, among others. The resulting CDF, and its associated probability distribution function (PDF), serves as a representation of the uncertainty needed for a robust optimization process. This CDF is then represented by a single decision metric value to allow the ranking of the proposed solutions and subsequent selection of the optimal solution.

There has been a number of works which utilize the full set of geological realizations and optimize to maximize/minimize a single decision metric (e.g., expected value of NPV). Chen et al. (2009), Fonseca et al. (2017), Fonseca et al. (2015), Chen et al. (2012), Fonseca et al. (2014a), Dehdari and Oliver (2012), and van Essen et al. (2009) used the full set of realizations for well control optimization with an objective function represented by the expected value of NPV (i.e., the decision metric). Güyagüler and Horne (2004) implemented a decision analysis framework for well placement optimization over the full set of realizations using the expected utility as the objective function to account for the decision-maker's risk attitude.

However, using the full ensemble of realizations to evaluate each proposed solution (i.e., field development plan) makes the optimization problem intractable. Consequently, there has been literature that investigated techniques to reduce the computational cost of optimization under geological uncertainty. These techniques attempt to select the most representative subset of realizations from the

full set to propagate the geological uncertainty in a computationally efficient manner. Previous literature on subset selection techniques can be classified into four main categories: (1) random selection, (2) static property methods, (3) flow-response vector-based methods and (4) a combination of static property and flow-response vector-based methods.

The random subset selection technique has been widely implemented in the literature, mainly due to its simplicity. As the name suggests, in this technique a pre-defined number of realizations are randomly selected from the full set to form the subset. Lorentzen et al. (2009) utilized the random selection technique within a closed loop reservoir management framework for well control optimization of the Brugge benchmark model. Li et al. (2013) applied the random selection subset selection technique at each iteration of the joint optimization of well control and well location. Jesmani et al. (2020) and Jesmani et al. (2016b) implemented the random selection technique in a similar manner for well placement optimization. The authors argued that this would help ensure that all the realizations in the subset are utilized in the optimization and improve the results.

Several works investigated the use of static property based methods for subset realization selection. These techniques rely on the selection of a static geological property, such as average permeability, average porosity, original oil in place (OOIP), or even a combination of these. Once the property is selected, a single value is calculated to represent each realization. This allows the realizations to be ranked accordingly, usually in an ascending order.

Deutsch and Begg (2001) proposed that once the realizations are ranked, the aim should be to select realizations which are equally spaced. The authors argued that this would preserve the quantiles of the full set. McLennan and Deutsch (2005) proposed the combination of a number of static properties including OOIP, averages of water saturation, permeability and porosity, as well as local and global connectivity measures. Li et al. (2012) followed this work up by using a connected hydrocarbon volume measure based on well locations and types. Rahim and Li (2015) used an optimized-based method to find a subset of realizations that minimized a probability distribution distance metric compared to the full set. The technique was then applied to well placement optimization and compared to the ranking

method and the random selection technique. The authors argued that this would ensure the subset would have a similar statistical distribution characteristics to the full set.

Another category of realization selection techniques aims at approximating the CDF of the output (flow-response) for a base case field development plan. In these flow-response vector methods, an inherent assumption is made: the approximated CDF of the base case is representative of the CDFs from all other possible development strategies. In Ballin et al. (1992), fast-to-evaluate alternatives were used to run the full ensemble for a base case development strategy. The realizations were then ranked and selected based on the output from the base case to form a subset of realizations. Mishra et al. (2002) and Odai and Ogbe (2011) used a similar technique, but instead used streamline simulations to evaluate a base case development strategy for the full ensemble of realizations. In Chen et al. (2012), a base case of well control settings was used to create a CDF of NPVs. Then, 11 reservoir models that corresponded to P1, P10, P20, P30... P90 and P99 were selected as the subset. The selected subset was then used for the full well control optimization procedure.

Another implementation of flow-response vector-based subset selection techniques are those that use distance-based clustering. To find the distance between the realizations in the full set a distance (or dissimilarity) matrix is calculated. Each element in the matrix represents the pairwise distance between two realizations based on a defined characteristic. The characteristic could be the recovery factor, cumulative oil production and water breakthrough time. The distance matrix is then converted into the Euclidean space using multidimensional spacing (MDS). A following step is then undertaken using kernel methods to map the data points to a feature space. This allows the application of principal component analysis (PCA) and k-means clustering to select a subset of realizations. This technique was implemented in Scheidt and Caers (2009), Scheidt and Caers (2008) and Alpak et al. (2010) using streamline simulations to obtain the flow-response vectors used to calculate the difference matrix.

The last category of subset selection techniques is one which combines both the static property methods and the flow-response vector methods. In Wang et al. (2012), static measures and flow-response vectors are combined and projected into a two-dimensional space after which a clustering technique is used to select the subset of realizations. The flow-response vectors were based on the

cumulative oil production for an initial set of well locations. The subset was then used within a retrospective optimization framework for well placement optimization. Haghghat Sefat et al. (2016) then used a similar approach to Wang et al. (2012) for well control optimization with the dissimilarity measure being based on the well water cut curves. In Shirangi and Durlofsky (2016), flow-response vectors (from selected well configurations) and static measures were combined in a weighted fashion resulting in a feature vector that can be used for clustering. The technique was then applied for well control optimization for a simple channelized two-dimensional reservoir model. Salehian et al. (2021) also used the technique introduced in Wang et al. (2012) for joint optimization of well control and well placement. The static measure used was permeability and the flow-response vector was the cumulative oil production for the initial well configuration.

3 An accelerated gradient algorithm for well control optimization

Aroui, Y. and M. Sayyafzadeh,

Journal of Petroleum Science and Engineering, 2020. 190: p. 106872.

Statement of Authorship

Title of Paper	An accelerated gradient algorithm for well control optimization
Publication Status	<input checked="" type="checkbox"/> Published <input type="checkbox"/> Accepted for Publication <input type="checkbox"/> Submitted for Publication <input type="checkbox"/> Unpublished and Unsubmitted work written in manuscript style
Publication Details	Arouri, Y. and M. Sayyafzadeh, An accelerated gradient algorithm for well control optimization. Journal of Petroleum Science and Engineering, 2020. 190: p. 106872.

Principal Author

Name of Principal Author (Candidate)	Yazan Arouri		
Contribution to the Paper	conceptualization, development of algorithm, experimentation, analysis, coding, original manuscript, corresponding author duties		
Overall percentage (%)	80%		
Certification:	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.		
Signature		Date	10/04/2022

Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

- the candidate's stated contribution to the publication is accurate (as detailed above);
- permission is granted for the candidate to include the publication in the thesis; and
- the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author	Mohammad Sayyafzadeh		
Contribution to the Paper	Analysis of results, discussions, revisions, supervision		
Signature		Date	06/06/2022

Name of Co-Author			
Contribution to the Paper			
Signature		Date	

Please cut and paste additional co-author panels here as required.



An accelerated gradient algorithm for well control optimization

Yazan Arouri^{*}, Mohammad Sayyafzadeh

Australian School of Petroleum, The University of Adelaide, 5005, SA, Australia

ARTICLE INFO

Keywords:

Well control optimization
Optimization under geological uncertainty
Brugge model
Simultaneous perturbation stochastic approximation (SPSA)
Accelerated gradient
Adaptive moment estimation

ABSTRACT

The simulation-assisted optimization of injection and production control settings can be considered as a way for improved management of hydrocarbon reservoirs. Although well control optimization problems are expected to have a relatively smooth landscape, they suffer from having a high number of dimensions. As a result, gradient-based algorithms are typically the preferred choice over derivative-free algorithms. Since the calculation of the exact gradient is not always feasible, recent studies into gradient-based algorithms have focused on ameliorating the accuracy of gradient approximation methods. Less attention has been paid to the framework in which these gradient approximations are utilized. The most common framework is the steepest descent with a backtracking line-search. This study aims to provide an improved framework using an adaptive moment estimation technique. The proposed method is compared to the steepest descent framework in two case studies of increasing complexity. In both frameworks, the gradient is approximated by the simultaneous perturbation stochastic approximation. The first case study is a two-dimensional heterogeneous reservoir model produced under water-flooding. The second case study investigates the injection and production settings of the three-dimensional benchmark case, the Brugge model, whilst incorporating geological uncertainty. The proposed framework showed improvements of up to 91% in convergence speeds and up to 5% in optimal value for the first case study. In addition, the proposed framework showed an improvement of up to 81% in convergence speeds and up to 2% in optimal value for the second case study over the steepest descent framework. For both case studies, the effect of the gradient approximation accuracy (perturbation number) was also investigated. The results indicated that the proposed algorithm is less sensitive to the gradient approximation accuracy than the steepest descent framework. In addition, this study investigated the effect of two popular bound constraint handling techniques in both frameworks. The results indicated that the projection method outperforms the logarithmic transformation method when applied to production optimization problem.

1. Introduction

The cyclical nature of the industry has put greater emphasis on improving hydrocarbon reservoir management processes. Optimization in reservoir management has been a topic of interest to the industry over the past few decades due to its ability to potentially increase returns. However, this topic has garnered major interest more recently due to the advances in common computational power (Hou et al., 2015), allowing practitioners to employ high-order mathematical models for forecasting system's response.

Production optimization, also known as well control optimization, focuses on assigning values to the control settings of wells (bottom-hole pressure (BHP) or flow rate) over a pre-defined or adaptive control intervals. The aim is to optimize an objective function over the reservoir lifespan. These settings determine the internal boundary conditions of

the system (reservoir), and adjusting them can allow the optimal control of the system. Production optimization problems are highly nonlinear with no closed-form solutions. This is a result of the complexity of the governing equations describing multi-phase fluid flow in porous media (which are typically solved using numerical methods). To evaluate the goodness of any proposed setting (solution-candidate), one numerical reservoir simulation should be performed, which is computationally intensive. In addition, the number of control (decision) variables can be in the range of 100's to 1000's, depending on the number of wells, the reservoir lifespan and the number of control intervals. Further complexities arise with the incorporation of linear and nonlinear constraints into the problem, such as minimum oil production, maximum water production, and water cuts in producer wells. Moreover, it is important to incorporate geological uncertainty to ensure a solution that is robust against the subsurface possibilities. This leads to a very computationally

^{*} Corresponding author.

E-mail addresses: yazan.arouri@adelaide.edu.au (Y. Arouri), mohammad.sayyafzadeh@adelaide.edu.au (M. Sayyafzadeh).

demanding iterative optimization problem which can be time-consuming for real life problems. Consequently, the development of an algorithm that is suitable for such problems has been a matter of research.

There are two broad categories of optimization algorithms: gradient-based and derivative-free optimization (DFO) algorithms. These algorithms attempt to improve the fitness value by iteratively exploring and/or exploiting the solution space. Gradient-based algorithms rely on first and/or second-order information about the objective function with respect to the variables to direct the search in each iteration. Conversely, derivative-free algorithms solely rely on the objective function values as a means of guiding the search progression and are typically aided by a form of stochasticity. This makes them more suitable for low-dimensional problems with multimodal landscapes (Nocedal and Wright, 2006). However, as the number of control variables increases, DFO algorithms become computationally inefficient as they require a large number of function evaluations to converge (Hou et al., 2015). Hence, gradient-based algorithms have been popular for use in production optimization due to their ability to traverse the search space in a computationally efficient manner (Do and Reynolds, 2013). In addition, it is thought that the objective function search space is relatively smooth, characterized by long flat valleys (Do and Reynolds, 2013; Fonseca et al., 2014; Zhao et al., 2013).

Many gradient-based algorithms have been applied to production optimization problems which typically differ in the method of gradient calculation/approximation. One such method is the adjoint method, which efficiently computes the gradient of the objective function with change in the input regardless of the number of control variables (Jansen, 2011). It was first implemented to optimize the tertiary recovery process of surfactant flooding (Fathi and Ramirez, 1984; Ramirez et al., 1984). Subsequent studies investigated its application in carbon dioxide flooding (Mehos and Ramirez, 1989), steam flooding (Liu et al., 1993), and water flooding (Asheim, 1988; Brouwer and Jansen, 2004; Brouwer et al., 2004; Sarma et al., 2005; Sudaryanto; Yortsos, 2001; Virnovsky, 1991). Adjoint-based methods are considered intrusive, as they require access to the source code of the reservoir simulator, which is unavailable for the abundantly used commercial simulators. Furthermore, the effort required for the development and maintenance of adjoint-based optimizers is not trivial (Jansen, 2011). A detailed review of adjoint-based optimization is given by Jansen (2011) for the interested reader.

Accordingly, other methods have been proposed in the literature which do not require access to the simulator source code. The simplest method is the finite-difference gradient approximation. This method requires $2n$ function evaluations (reservoir simulations), where n is the number of variables, to approximate the gradient vector for one iteration. As such, for typical problems which involve many control variables, this technique becomes inefficient. Other gradient approximations which have been implemented in production optimization and are computationally less demanding include the simultaneous perturbation stochastic approximation (SPSA), introduced by Spall (1992) and applied to production optimization by Wang et al. (2009) and Foroud et al. (2018), ensemble-based optimization (EnOpt) (Chen et al., 2009) and the stochastic simplex approximate gradient (StoSAG) (Fonseca et al., 2017). These techniques approximate the gradient by a simultaneous perturbation of all the variables, whereas the finite-difference only perturbs one variable at a time. The EnOpt and StoSAG techniques are gradient approximations which aim to incorporate geological uncertainty (noise in the objective function) more efficiently. Do and Reynolds (2013) investigated the theoretical similarities between G(Gaussian)-SPSA, a SPSA-type gradient approximation, simplex gradient, and EnOpt within a steepest descent (SD) framework. They concluded that there were minimal differences between gradient

approximations, including SPSA.

A common feature of many gradient-based methods employed in production optimization is that these gradient approximation techniques have been utilized within the SD framework. The SD framework utilizes first-order information about the objective function to direct the search in the direction of descent (Nocedal and Wright, 2006). In addition, a line search is typically implemented within the SD framework to find a suitable step size which controls the magnitude of movement in the search direction (Nocedal and Wright, 2006). Majority of research in production optimization has focused around improving the gradient approximation techniques with little emphasis on the framework which utilizes the approximated gradients. Asadollahi and Naevdal (2009) compared the SD and conjugate gradient frameworks for water-flooding optimization of the Brugge model using the adjoint method for gradient calculation. Chaudhri et al. (2009) also compared the SD and conjugate frameworks using EnOpt for gradient approximation. Both these studies reported slight improvements in optimal values over SD. Other studies have aimed at improving the handling of nonlinear constraints through the use of the Sequential Quadratic Programming (SQP) framework. This method aims at using the Lagrangian formulation to iteratively solve quadratic sub-problems. Asadollahi et al. (2014) implemented SQP for the production optimization of the Brugge model and compared it to Hooke-Jeeves direct search (HJDS), the Nelder-Mead (NM) method, and the generalized pattern search (GPS) method. Dehdari and Oliver (2012) also implemented SQP for the production optimization of the Brugge model with EnOpt as the gradient approximation. Although the SQP framework can handle nonlinear constraints, the additional computational requirement is considerable.

This study presents the development of a first-order optimization algorithm with application to production optimization. The developed algorithm is a combination of SPSA as a gradient approximation within an adaptive moment estimation framework. The adaptive moment estimation framework has found significant success in optimization problems in machine learning applications. This is due to the incorporation of additional information from previous gradients to calculate variable-specific search directions. The search direction in the proposed framework is a vector composed of an average of previous gradients that is normalized element-wise by an estimated variance. As a result, this assists the search progression to be adjusted further for each control variable and allow a convergence speed-up. The developed algorithm is compared to the original SD-SPSA on two production optimization problems of increasing complexity.

The outline of the paper is as follows. The problem statement, including the formulation of the optimization problem, as well as the objective function are presented in Section 2. Section 3 begins by introducing the steepest descent (ascent) framework and the adaptive moment estimation framework. Then, the development of a new algorithm, Adam-SPSA, is presented. The selection of algorithm-specific parameters is also discussed. Section 4 begins by studying the effect of bound constraint handling techniques on both algorithms. Following this, numerical results for production optimization on two different case studies of differing complexity are presented. The significance and implication of these results are discussed in Section 5. A summary and final conclusions are made in Section 6.

2. Problem statement

2.1. Optimization problem formulation

The optimization problem involves the maximization of a defined objective function where the well controls are the variables of interest. Typical variables considered in the literature are the liquid rates and/or

the bottom-hole pressure (BHP) of wells over a predefined or an adaptive set of control intervals. Other approaches have been studied, including explicit switching times (Sudaryanto and Yortsos, 2001), the halving-interval method (Hasan and Foss, 2013), and switching time intervals (Fonseca et al., 2015). The more common method is to use liquid rates and/or BHP values as the control variables. These well controls are piecewise constant between two control intervals which are typically pre-determined.

The production optimization problem can be formulated as a general optimization problem as follows:

$$\begin{aligned} & \min_{x \in \mathbb{R}^n} f(x) \\ & \text{subject to:} \\ & c_i(x) \leq 0, \quad i \in K \\ & c_i(x) = 0, \quad i \in I, \end{aligned} \quad (1)$$

where, $f(x)$ represents the objective function (defined later) to be optimized, x is the vector of control (decision) variables, n is the number of control variables, I and K are sets of indices for equality and inequality constraint functions, $c(x)$, respectively. These constraint functions may include linear and nonlinear constraints, as well as bound constraints. The control variables considered are continuous bottom-hole pressures (BHP) with upper and lower bounds for a pre-determined number of control intervals.

2.2. Objective function

In this paper, we consider the objective function to be the negative of the Net-Present-Value (NPV) for a given lifespan. The NPV is defined as:

$$f(x) = -NPV(x, m) = -\sum_{t=1}^{N_t} \frac{r_o Q_{o_t}(x, m) + c_{wp} Q_{wp_t}(x, m) + c_{wi} Q_{wi_t}(x, m)}{(1+b)^t}, \quad (2)$$

where N_t is the reservoir lifespan in years, r_o is the price of oil (per unit volume), Q_{o_t} , Q_{wi_t} and Q_{wp_t} are cumulative oil production, cumulative water injection and cumulative water production, over time t and $t + 1$, respectively. c_{wp} and c_{wi} are the cost of handling produced water and the cost of water injection (per unit volume), respectively, b is the discount rate (%), and t is the number of years passed since start of production. To obtain an objective function value for a specific set of control variables x , a reservoir simulation must be run for a specified reservoir model, m . When no geological uncertainty is incorporated into the optimization problem, the deterministic objective function is defined as above for a specific reservoir model, m . On the other hand, when geological uncertainty is incorporated the expected value of an ensemble of reservoir models is used as the objective function. The robust objective function for N_e equally probable geological realizations can be formally stated as:

$$E[NPV] = \frac{\sum_{i=1}^{N_e} NPV(x, m_i)}{N_e}, \quad (3)$$

where, m_i represents a reservoir model. Previous literature investigated incorporating risk attitude and other statistics when determining the objective function value (Chen et al., 2017; Yeten et al., 2004), however; these are not the focus of this paper. Furthermore, to ensure a computationally efficient implementation a subset of realizations is used to calculate the robust objective function value, which is discussed in more detail later.

3. Methodology

We first present the two optimization frameworks that are the focus of this paper and then outline the gradient approximation utilized. The

combination of each framework with the gradient approximation leads to the two optimization algorithms that are investigated in this study. The formulations and modifications made to each algorithm are also presented and discussed.

3.1. The steepest descent framework

The steepest descent (ascent) framework is the simplest and most widely implemented framework. This framework utilizes the gradient as the search direction to progress the optimization (Nocedal and Wright, 2006). The movement of the search depends on the value of the gradient in the direction of the local downhill (uphill) gradient. The search direction can also be the approximated gradient normalized by its infinity norm. This enables a more appropriate guess for the initial step size (Do and Reynolds, 2013). The framework is presented in Framework 1.

Framework 1

0. Initialize iteration counter, $k = 0$
1. Select initial guess $x_{k=0}$
2. Approximate the gradient as the search direction, g_k
3. Calculate proposed iterate: $x_{k+1} = x_k - a \times g_k$, where a is a pre-defined step size.
4. Return to step 1 and repeat until the stopping criterion is met.

This framework has been abundantly used within production optimization while utilizing different methods for computing or approximating the gradient in step 2. Brouwer and Jansen (2004) implemented the adjoint method within the steepest descent framework to optimize the water-flooding in three two-dimensional reservoir models. Chen et al. (2009) utilized an ensemble-based gradient approximation to incorporate geological uncertainty within a steepest descent (ascent) framework for the production optimization of a two-dimensional model. Other gradient approximations implemented within a steepest descent framework include stochastic simplex approximate gradient (Fonseca et al., 2017) and finite difference (Wang et al., 2009).

3.2. The adaptive moment estimation framework

The adaptive moment estimation (Adam) framework was first introduced by Kingma and Ba (2014) as a stochastic gradient-based algorithm which utilizes first-order information. Although the framework was built for general stochastic optimization in science and engineering, its main application has been in computer science, including Google's translation system (Wu et al., 2016) and image processing (Gregor et al., 2015; Ledig et al., 2017). To the best of the authors' knowledge, application of Adam to engineering applications, in particular production optimization, has yet to be investigated.

The efficiency of Adam comes from its requirement for only first-order gradient information and its ability to compute unique search directions for different control variables (dimension) (Kingma and Ba, 2014). Adam adaptively calculates the search direction at each iteration depending on the approximation of first-order information, which allows the search progression to be adjusted with regards to each control variable. The proposed search direction can be thought of as the signal-to-noise ratio (SNR), where a small SNR value represents a greater uncertainty about whether the approximated gradient corresponds to the direction of the true gradient (Kingma and Ba, 2014). As the search progresses and an optimum is approached, the SNR will typically approach a value closer to zero. The Adam framework is based upon two algorithms previously used in machine learning: AdaGrad (Duchi et al., 2011) and RMSProp (Hinton et al., 2012). AdaGrad introduced the idea of utilizing per-parameter search directions, whilst RMSProp bases these search directions on the average of the magnitudes of recent gradients (i.

e. how quickly these gradients are changing). Adam not only combines these ideas, but the adaptive per-parameter search direction are also calculated based on an estimate of the second moment, too.

The adaptive nature of Adam is dependent on the calculation of two main values: estimation of the first and second moment. By definition, the mean is the first moment of a random variable X about the origin defined (Lefebvre, 2006). The random variable is the gradient approximation. Hence, the estimation of the first moment is an estimation of the expected value of the gradient. In addition, the second central moment of X is the variance about the mean (Lefebvre, 2006) and is defined by:

$$Var(X) = Diag(E[(X - E[X])^2]) \quad (4)$$

with simple algebraic manipulation, the variance formula above can be also defined element-wise, as below, in which it is assumed variables of X are independent:

$$Var(X) = Diag(E[X^2] - (E[X])^2) \quad (5)$$

In the long-run during optimization, the mean of the gradients tends to approach zero as a local optimum is found. As a result, the second term in Equation (5) provides no additional information leading to the uncentred variance:

$$unVar(X) = Diag(E[X^2]) \quad (6)$$

Hence, since the random variable is the gradient approximation, the uncentred variance is the expected value of the element-wise gradient squared. Adam uses an exponential moving average to estimate the first moment and the second raw moment of the gradient, defined in Framework 2 (Kingma and Ba, 2014). By using an exponential moving average greater emphasis is put on the most recent approximated gradient to guide the search, yet utilizing previous gradients as additional information. The convergence theory for Adam is presented in Kingma and Ba (2014), whilst Chen et al. (2018) go into further detail of the convergence of adaptive gradient algorithms in general.

Framework 2

0. Initialize iteration counter, select initial guess x_0 , assign nonnegative constant parameters: step size a and exponential decay rate for moment estimates β_1, β_2 . Typical values for β_1, β_2 are 0.9 and 0.999, respectively (Kingma and Ba, 2014). A small positive value $\epsilon = 10^{-8}$, is used to avoid the division by zero (Kingma and Ba, 2014).
1. Initialize first and second moment vectors, m_0 and v_0 , which are $n \times 1$ column vectors
2. Approximate the gradient
3. Update the first moment vector using update rule:
 - a. $m_k = \beta_1 \times m_{k-1} + (1 - \beta_1) \times \hat{g}_k$, where m_k, m_{k-1} , and \hat{g}_k are $n \times 1$ column vectors
 - b. $\hat{m}_k = \frac{m_k}{1 - \beta_1^k}$, where \hat{m}_k is a $n \times 1$ column vector
4. Update the second raw moment vector:
 - a. $v_k = \beta_2 \times v_{k-1} + (1 - \beta_2) \times (\hat{g}_k \odot \hat{g}_k)$, where \odot is the element-wise multiplication and v_k and v_{k-1} are $n \times 1$ column vectors
 - b. $\hat{v}_k = \frac{v_k}{1 - \beta_2^k}$, where \hat{v}_k is a $n \times 1$ column vector
5. Update the iterate: $x_{k+1} = x_k - a \times \frac{\hat{m}_k}{\sqrt{\hat{v}_k + \epsilon}}$

It should be noted that all operations are done in an element-wise manner. The \odot represents the element wise multiplication of two vectors. The approximated average gradient was used when calculating v_k for a perturbation number greater than 1. Also, the operation on the hyper-parameters (β_1^k, β_2^k) are denoted as β_1 and β_2 raised to the power k (iteration number). The gradient approximation method was not

explicitly stated in Kingma and Ba (2014). Also, to the best of the authors' knowledge, there is no investigation on the effect of bound constraint handling techniques on the performance of Adam.

3.3. Gradient approximation by SPSA

The SPSA gradient approximation was presented as an efficient alternative for problems whose analytical derivative are unavailable and measurements are considered to have noise (Spall, 1992). The efficiency of SPSA lies in its requirement of only two function evaluations (for the central-difference method) to calculate a gradient approximation. This property allows the gradient approximation to be very efficient for high dimensional problems, as is typically the case in production optimization. SPSA has also been compared to other leading optimization algorithms, including simulated annealing, evolution strategies, and genetic algorithms (Spall et al., 2006). The stochastic gradient approximation is based on a central finite-difference written as:

$$\hat{g}_k(x_k) = \begin{bmatrix} \frac{f(x_k + c_k \Delta_k) - f(x_k - c_k \Delta_k)}{2c_k \Delta_{k1}} \\ \vdots \\ \frac{f(x_k + c_k \Delta_k) - f(x_k - c_k \Delta_k)}{2c_k \Delta_{kp}} \end{bmatrix} \quad (7)$$

$$= \frac{f(x_k + c_k \Delta_k) - f(x_k - c_k \Delta_k)}{2c_k} \times [\Delta_{k1}^{-1}, \Delta_{k2}^{-1}, \dots, \Delta_{kp}^{-1}]^T,$$

where the mean-zero p -dimensional random perturbation vector $\Delta_k = [\Delta_{k1}^{-1}, \Delta_{k2}^{-1}, \dots, \Delta_{kp}^{-1}]^T$, has a user-specified distribution and c_k is a positive scalar. Other finite difference methods, such as forward and background difference, can also be used (Do, 2012; Spall, 1992). The convergence theory of the SPSA algorithm can be found in Spall (1992). A major argument is the conditions of the user-specified distribution used to select the perturbation vector. The convergence condition requires the randomly sampled perturbation vectors to be independent for all k , identically distributed at each k , symmetrically distributed about zero and uniformly bounded in magnitude for all k (Spall, 1992). As a result, Normal (Gaussian) and Uniform distributions are not valid for use in SPSA, according to Spall (1992). Do and Reynolds (2013) utilized the Gaussian distribution in their SPSA implementation and reported suitable results. However, a distribution which does satisfy these conditions, and the one used in Spall (1998), is the Bernoulli distribution. The implementation of SPSA as a gradient approximation is presented in Gradient Approximation 1.

Gradient Approximation 1

0. Initialize non-negative coefficients (α and c) based on guidelines provided in Spall (1998)
1. Generate the p -dimensional random perturbation vector Δ_k based on the Bernoulli distribution.
2. Use the perturbation vector to calculate the forward and backward steps at $x_k^+ = x_k + c_k \Delta_k$ and $x_k^- = x_k - c_k \Delta_k$.
3. Calculate the corresponding objective function values: $f(x_k^+)$ and $f(x_k^-)$
4. Compute the stochastic gradient approximation: $\hat{g}_k(x_k) = \frac{f(x_k^+) - f(x_k^-)}{2c_k} [\Delta_{k1}^{-1}, \Delta_{k2}^{-1}, \dots, \Delta_{kp}^{-1}]^T$

Furthermore, the accuracy of gradient approximation can be improved with by averaging a number of individual stochastic gradients (Spall, 1992), as shown in Equation (8). This is referred to as the perturbation number and is investigated in the remainder of this paper.

An increase in the number of averaged gradients could lead to a more accurate gradient approximation that requires a lower number of iterations, hence leading to a faster convergence to the optimum. However, an increase in perturbation number may not necessarily lead to an increase in computational efficiency. This occurs if the associated increase in function evaluations outweighs the increase in gradient accuracy. As such, if there little or no increase in accuracy, the search is not necessarily going to progress faster.

Wang et al. (2009) compared one gradient and an average of 10 gradients in SD-SPSA on a simple synthetic model made up of an 11 by 11 grid with 32 control variables. The results showed that an average of 10 gradients converged to a higher optimum value. They also tested an average of 10 and 20 gradients on a 25 by 25 synthetic reservoir model with a total of 48 control variables. In this case, the results showed that using an average of 20 gradients outperformed using an average of 10 gradients. Do and Reynolds (2013) compared an average of 5 and 10 individual gradients for their SPSA-type gradient approximation on a 25 by 25 reservoir model with 130 control variables. The results showed that an average of 5 gradients outperforms 10 for their simple two-dimensional model. These results reflect the problem dependency with regards to the optimal number of perturbations to be used. As such, this is investigated for each of the case studies of this paper.

$$\bar{g}_k(x_k) = \frac{1}{N_p} \sum_{j=1}^{N_p} \hat{g}^{(j)}(x_k) \quad (8)$$

3.4. The steepest descent SPSA algorithm

The algorithm presented in Spall (1992) utilizes the SPSA gradient approximation within a steepest descent framework. The combination of these two components results in the steepest descent SPSA (SD-SPSA) algorithm, Algorithm 1. This algorithm uses a gain sequence to determine the step size within each iteration using Equation (9). In addition, typical implementation of gradient-based algorithms are accompanied by a backtracking line-search to help with the determination of the step size (Do and Reynolds, 2013; Fonseca et al., 2017). This aims to ensure an increase in the objective function value at each iteration (Nocedal and Wright, 2006). A backtracking line-search involves reducing the step size by a constant factor, typically by half, if a decrease in objective function value is not obtained. A backtracking line-search with a maximum of 5 cuts was employed in this paper. If the line-search is unsuccessful, the gradient is recalculated and the backtracking is repeated. The backtracking line-search is used in addition to the gain sequence mentioned above, leading to the following update formula:

$$x_{k+1} = x_k - \rho a_k \times \hat{g}_k(x_k), \quad (9)$$

where, ρ represents the step size reduction factor. For each iteration, the gain sequence (Spall, 1992) is used to calculate the step size a_k , before the backtracking line-search is undertaken.

Algorithm 1

0. Initialize iteration counter, select initial guess x_0 and nonnegative coefficients (a and c) based on guidelines provided in Spall (1998)
1. Calculate gradient approximation using Gradient Approximation 1
2. Calculate proposed iterate: $x_{k+1} = x_k - a_k \times \hat{g}_k(x_k)$
3. Update a_k and c_k according to Spall (1998)
4. Return to step 1 and repeat until stopping criterion is reached.

The SD-SPSA algorithm has been utilized for production optimization problems of varying levels of complexity. Wang et al. (2009) implemented SPSA within a closed-loop reservoir management workflow for a low order 2-dimensional model. More recently, Foroud et al.

(2018) compared the performance of gradient-based algorithms to DFO algorithms for production optimization of the Brugge model. The results showed that gradient-based algorithms outperformed the DFO algorithms (GPS, particle swarm optimization, covariance matrix adaptation evolutionary strategy, differential evolution and self-adaptive differential evolution).

The application of SD-SPSA to constrained optimization problems has also been investigated. SD-SPSA has been paired with the projection method as a means to deal with constraints (Fu and Hill, 1997; Sadegh, 1997). This technique maps proposed solutions back into the feasible domain if they are infeasible. This technique is typically implemented when the constraints are bound constraints, with lower and upper bounds. Another approach presented by Wang and Spall (2003) is the penalty function. Briefly, the objective function is modified with the addition of a penalty term that accounts for any constraint violations by penalizing the objective function value. This technique is typically used to handle inequality constraints, which are not considered in this paper.

3.5. The Adam-SPSA algorithm

In this paper, the gradient approximation of SPSA is combined with the framework of Adam to produce an algorithm that aims to be more efficient than the steepest descent (ascent) framework. A modification was made to the original Adam (Kingma and Ba, 2014) to make it more efficient when combined with SPSA. Our experiments, not shown, found that as the number of perturbations used to calculate the average SPSA gradient increased, there was no effect on the initial iteration of Adam while there was a considerable effect on steepest descent (ascent). This was because Adam does not use the approximated gradient to direct the search, rather it used the search direction vector made up of ± 1 . To overcome this deficiency, the first iteration was altered to an iteration of steepest descent (ascent). This ensures the first proposed solution is directed by the search direction of SPSA and the subsequent iterations are based on the Adam update rule. As opposed to SD-SPSA, this algorithm does not require a backtracking line-search, as the results (not presented) showed no improvements. The proposed algorithm, Adam-SPSA, is presented in Algorithm 2.

Algorithm 2

0. Initialize iteration counter. Select initial guess x_0 . Assign nonnegative constant parameters: initial step size α and exponential decay rate for moment estimates β_1, β_2 , and c_k . Typical values for β_1, β_2 are 0.9 and 0.999, respectively (Kingma and Ba, 2014). A small positive value $\epsilon = 10^{-8}$, is used avoid the division by zero (Kingma and Ba, 2014).
1. Initialize first and second moment vectors, m_0 and v_0
2. Compute the gradient approximation (\hat{g}_k) using Gradient Approximation 1
3. If $k = 1$, proceed with steepest descent search direction
 - a Calculate proposed iterate:

$$x_{k+1} = x_k - a_k \hat{g}_k(x_k)$$
- Elseif $k > 1$
 - b Update the first moment vector using update rule:
 - (i) $m_k = \beta_1 \times m_{k-1} + (1 - \beta_1) \times \hat{g}_k$
 - (ii) $\hat{m}_k = \frac{m_k}{1 - \beta_1^k}$
4. Update the second raw moment vector:
 - a $v_k = \beta_2 \times v_{k-1} + (1 - \beta_2) \times (\hat{g}_k \odot \hat{g}_k)$
 - b $\hat{v}_k = \frac{v_k}{1 - \beta_2^k}$
5. Update the iterate: $x_{k+1} = x_k - \alpha \times \frac{\hat{m}_k}{\sqrt{\hat{v}_k + \epsilon}}$
6. If termination condition not met, return to Step 2 and continue

We efficiently parallelize the gradient calculation using the on-

premises cluster by enabling the reservoir simulation to run on as many cores as there was perturbations. Since Adam-SPSA does not have a backtracking line-search, parallelization is more effective than SD-SPSA. Furthermore, each experiment was run 5 times to account for the stochastic nature of the algorithms.

3.6. Selection of parameters

There are multiple parameters which need to be set. The main parameter which relates to the gradient calculation is the parameter c_k , which is the perturbation size when taking the forward and backward difference. This parameter has a decreasing sequence given by the following equation (Spall, 1998):

$$c_k = \frac{c}{(\text{iteration number} + 1)^\gamma}, \quad (10)$$

where c is the initial value that is chosen by the user, and γ is typically equal to 0.101 (Spall, 1998). A sensitivity analysis was done to select a suitable value for c (in the normalized domain). It was found that an initial value of $c = 0.1$ resulted in a perturbation size in the range of ~ 1 psi. Arguably, the effect of a perturbation size that is infinitesimally small may not be captured properly by the numerical simulation.

Another parameter that is utilized by SPSA is the step size gain sequence, a_k (Spall, 1998). The formula is given by:

$$a_k = \frac{a}{(\text{iteration number} + 1 + A)^\alpha}, \quad (11)$$

where $A = 0.1 * \text{maximum number of iterations}$, $\alpha = 0.602$, and a is defined by the user. These values are suggested by Spall (1998), however choosing the maximum number of iterations and a are problem dependent. For the purpose of our research, we elected to keep the maximum number of iterations constant regardless of the perturbation number. This allowed us to perform a sensitivity analysis on the value of a . This value can have major impact on the search progression, as it controls the step size used at each iteration. As such, an ideal step size is one which ensures that the search does not swing from one bound to another. The average of the absolute difference between the first two iterations was taken as an indication of this movement towards the bounds. A suitable value of $a = 10$ for SD-SPSA was found, whilst the step size of $a = 0.3$ gave a similar effect in Adam, in terms of update step, which allows delivering a fair comparison between the two. This is an initial step size that is constant throughout the optimization run. The

decay factors used for β_1 and β_2 are taken from Kingma and Ba (2014) as 0.9 and 0.999, respectively.

4. Numerical experiments

This section presents two numerical experiments used to evaluate and compare the performance of the SD-SPSA algorithm and the proposed Adam-SPSA algorithm for production optimization. The first case study under investigation is a simple two-dimensional (2D) heterogeneous model (Isebor, 2009), with a high water handling cost to motivate efficient water-flooding. In addition, we investigate the effect of perturbation number and constraint handling technique. The second case study is a more complex three-dimensional (3D) reservoir model, known as the Brugge model, that was utilized to test the robustness of the proposed algorithm by incorporating geological uncertainty (Peters et al., 2010). The second case study has different economics parameters to test the response of the proposed algorithm to different situations. Both these models are readily used for benchmarking production optimization algorithms.

4.1. Case study 1 – two-dimensional heterogeneous reservoir

4.1.1. Problem description

The first case study involves a 2D channelized reservoir model discretized into $40 \times 40 \times 1$ grid with four injectors and four producers (Isebor et al., 2014). Fig. 1 shows the line drive pattern and the x -direction permeability field. This reservoir involves two-phase flow of oil and water. The objective is to maximize NPV through finding the optimal BHP settings of the injectors and producers. The project lifespan is assumed to be 3000 days and is divided into 100 control steps, with the well's BHPs being updated every 30 days. As a result, the dimension of this problem is 800, composed from 100 controls steps for 8 wells. The reservoir and economic parameters are given in Table 1. The optimization variables are normalized and bounded between 0 and 1. To handle this bound constraint, we investigate both the projection method and the logarithmic transformation method. In addition, a full sensitivity analysis on the effect of the number of perturbations is also presented.

4.1.2. Algorithm parameter settings and constraint handling

As previously presented, parameters related to the SD-SPSA algorithm need to be selected. The results of this preliminary study are presented in Table 2.

The optimal value of a for this case study seems to be between 10 and 100. From the results of the last row in Table 2, it is clear that a value of 100 leads to a large movement towards to boundary indicated by Δx . If this occurs, the gradient-based techniques may have difficulty moving away from the boundaries and are more inclined to get stuck without thoroughly searching the landscape. On the other hand, using a value of 1 leads to small movement, which slows down the search considerably. As such, all optimization runs were conducted using a value of 10 for a . As discussed in Section 3.6, the step size used in the Adam-SPSA is 0.3.

The first investigation undertaken using this case study examined the effect of the bound constraint handling technique on production optimization. There are two common methods that are utilized to deal with bound constraints: projection method (Asadollahi et al., 2014; Dehdari

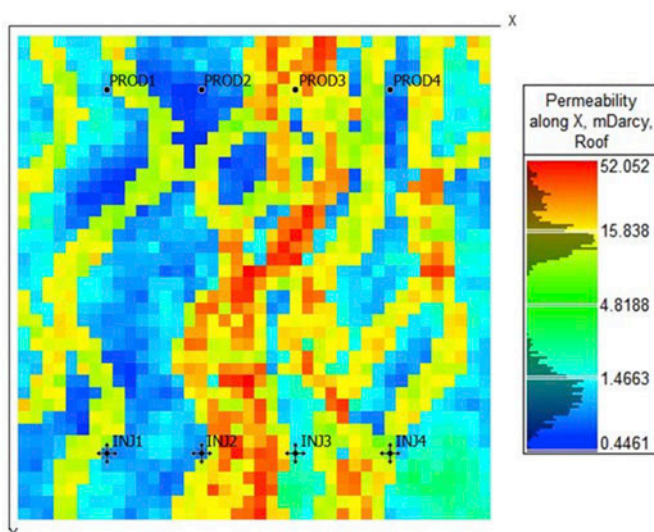


Fig. 1. Two-dimensional channelized heterogeneous model.

Table 1

Economic and reservoir parameters for case study 1.

Porosity	0.3
Oil Price (USD/STB)	80
Water injection costs (USD/STB)	36
Water handling costs(USD/STB)	18
Injector BHP range (psi)	6000–9000
Producer BHP range (psi)	2500–4500

Table 2
Sensitivity analysis on SPSA step size.

a	0.1	1	10	100
a_0	0.00621	0.0621	0.621	6.21
Δx	0.0020	0.0217	0.2034	0.4601

et al., 2012; Fonseca et al., 2017) and the logarithmic transformation method (Do and Reynolds, 2013; Lu et al., 2017; Wang et al., 2009). The projection method involves projecting control variables which are outside the bounds to their closest respective bound. Furthermore, the problem can be made simpler by normalizing the bounds to a hypercube to have bounds of $[0, 1]^n$, where n is the number of control variables. On the other hand, the logarithmic transformation method aims at converting a bounded constrained optimization problem to an unconstrained optimization problem. This is achieved through a transformation of variables using the following equation:

$$s_i = \ln\left(\frac{x_i - x_{lower}}{x_{upper} - x_i}\right), \quad (12)$$

where s_i is the i th transformed control variable, x_i is the i th control variables, x_{lower} is the lower bound, and x_{upper} is the upper bound. As a result, the bounds are extended to $[-\infty, \infty]$, essentially making the problem unconstrained. As such, as the control variables in the untransformed form x_i , approaches the lower bound, x_{lower} , the transformed variable, s_i , approaches $-\infty$. On the other hand, as the control variable x_i , approaches the upper bound x_{upper} , the transformed variable, s_i , approaches ∞ . When implemented, all operations are done in the transformed domain, and when required, the original control

variables are obtained through the inverse function:

$$x_i = \frac{\exp(s_i) \times x_{upper} + x_{lower}}{1 + \exp(s_i)} \quad (13)$$

To investigate the effect of the two bound constraint handling techniques, three different values for perturbation number were used: 1, 5 and 10. All experimental runs were obtained using the mid-point of the hypercube, i.e. $[0.5]^n$, as the initial guess. The results are presented in the appendix. The results reflect a detrimental impact on the convergence speed and the final optimal value of both algorithms, steepest descent SPSA (SD-SPSA) and Adam-SPSA, when using the logarithmic transformation method compared to the projection method. This result is common irrespective of the perturbation number used. There were differences of up to 4.8%, 4.9%, and 5.2% between the optimal values when using projection method and the logarithmic method for perturbation values of 1, 5, and 10, respectively.

Although these results are not extensive with regards to the problems tested, they give an indication that there can be a significant effect on the performance of the algorithm (SD-SPSA and Adam-SPSA) when using different constraint handling techniques. Based on the results presented, the logarithmic transformation method may induce gradient-based algorithms into getting entombed on the bounds. Consequently, the projection method is employed throughout the remainder of the paper.

4.1.3. Results of case study 1

An extensive set of experiments were conducted to assess the performance of Adam-SPSA against SD-SPSA. Firstly, three different initial guesses were used to investigate the robustness of the algorithms with regards to the starting point. The three initial guesses used were $[0.25]^n$, $[0.5]^n$ and $[0.75]^n$, where in this case n is equal to 800 control variables.

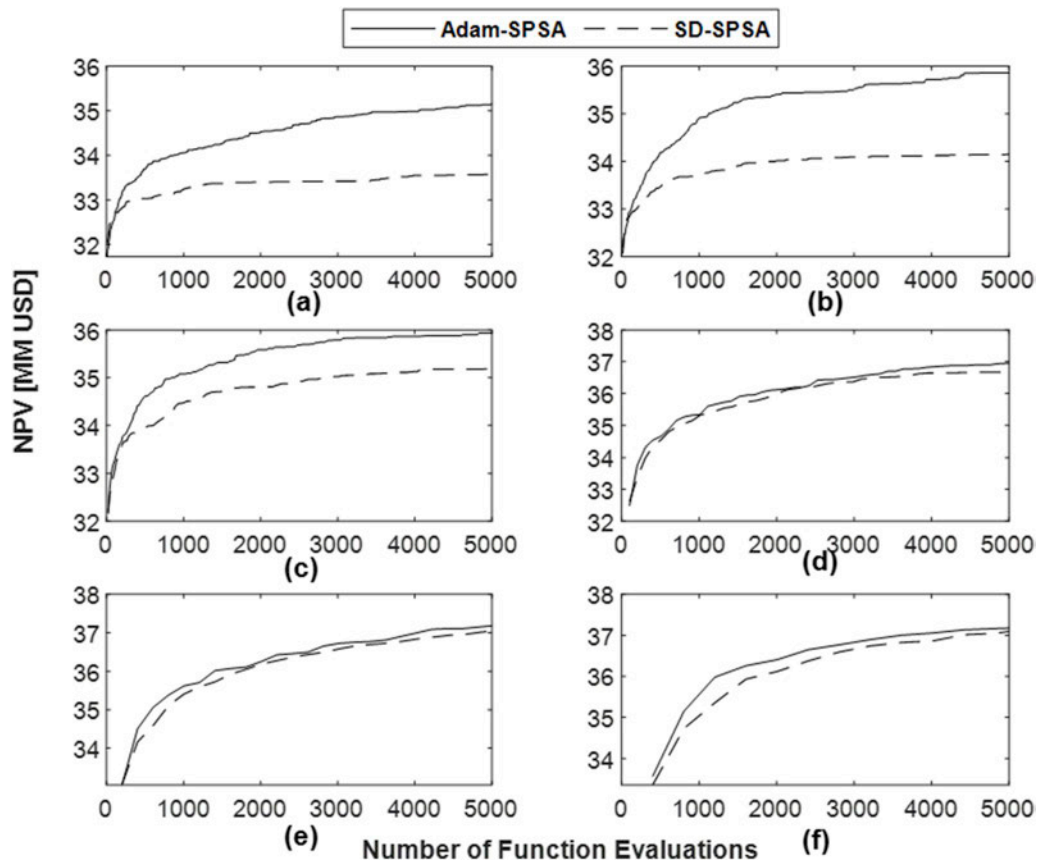


Fig. 2. Results of production optimization for an initial guess of 0.5 using (a) 1, (b) 5, (c) 10, (d) 50, (e) 100, and (f) 200 perturbations.

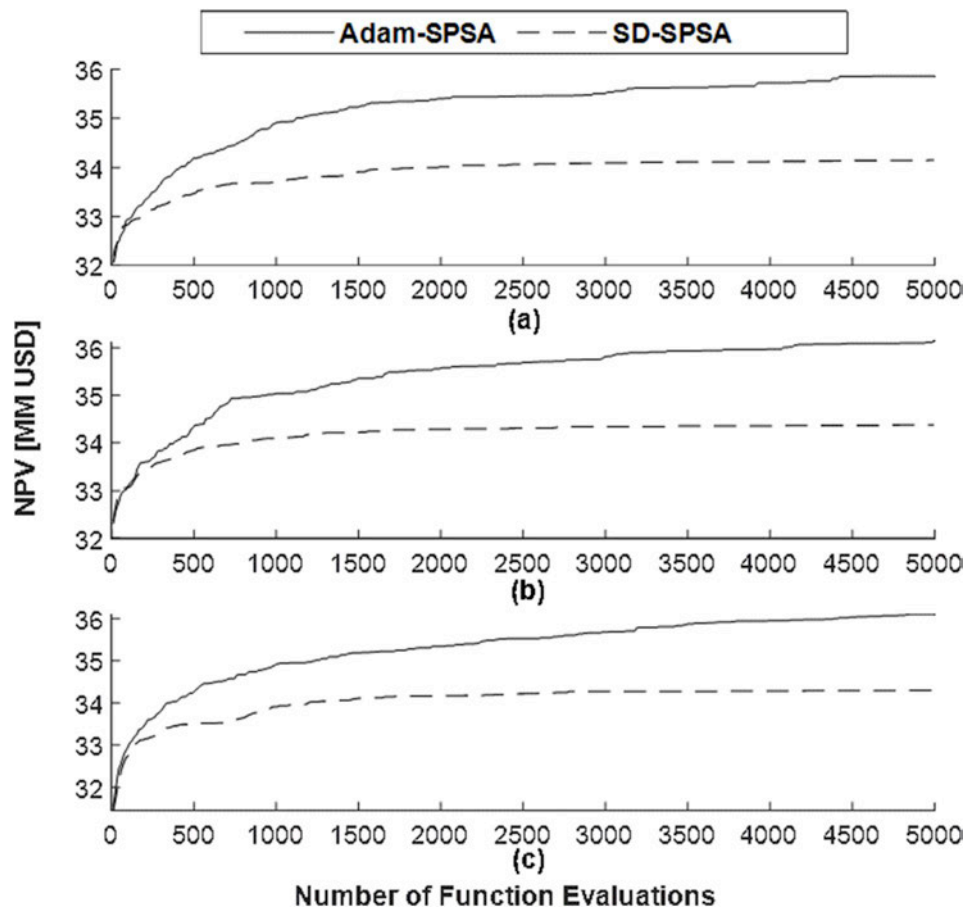


Fig. 3. The effect of initial guess on the performance of Adam-SPSA (solid line) and SD-SPSA (dashed line).

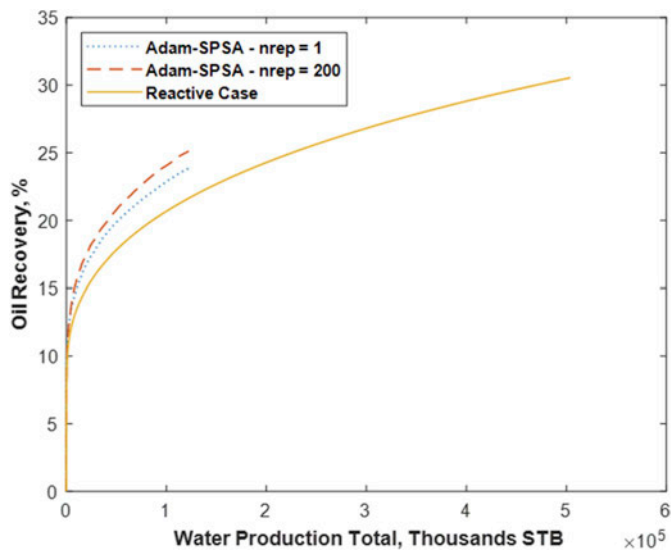


Fig. 4. Oil recovery as a function of total water production for reactive case and Adam-SPSA.

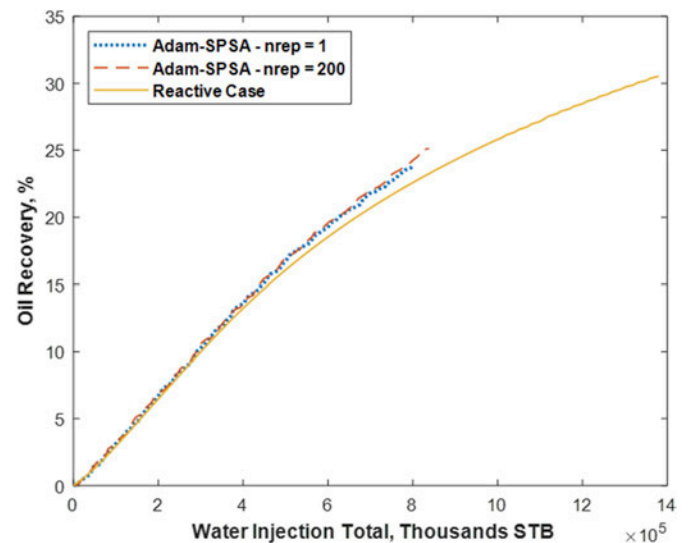


Fig. 5. Oil recovery as a function of total water injection for reactive case and Adam-SPSA.

For each initial guess, the number of perturbations used to calculate the gradient was also altered. Six different values were used: 1, 5, 10, 50, 100, and 200. This range of values was used to test the effect of perturbation number on the improvement in accuracy of the gradient approximation. The results are presented in Figs. 2-5. The results shown are an average of 5 individual optimization runs. The stopping criterion

for all the experiments, unless otherwise stated, is a maximum number of function evaluations of 5000.

In Fig. 2, the Adam-SPSA algorithm outperforms SD-SPSA for all values of perturbation number. The difference between Adam-SPSA and SD-SPSA ranges from, on average, 4.6% for one perturbation to less than 1% for 200 perturbations. However, it is clear that in all six perturbation numbers, Adam-SPSA outperforms SD-SPSA with regards to

convergence speed. That is, for any desired value of NPV, Adam-SPSA requires a lower number of function evaluations. For example, in Fig. 2f, for a desired NPV of 36 MM USD Adam-SPSA requires 1203 function evaluations whilst, SD-SPSA requires 1604. This is a 25% reduction in the required computation to achieve a similar NPV. Similar trends are achieved for 50 and 100 perturbations with reductions of 12% and 43%, respectively. It is also evident that SD-SPSA is more prone to converging to a local minimum as depicted in Fig. 2a and b, when a small number of perturbations is employed. On the other hand, Adam-SPSA does not converge to the same local solution, rather it is able to find a higher NPV at the end of 5000 function evaluations. Also, Adam-SPSA does not seem to fully converge after 5000 function evaluations.

Similar results were obtained when utilizing two other initial values of $[0.25]^n$ and $[0.75]^n$. That is, regardless of the initial guess Adam-SPSA is able to outperform SD-SPSA in regards to convergence speed and optimal value. As an example, Fig. 3 shows the results for all three initial values for a perturbation number of 5. These results reflect the consistent robust performance of Adam-SPSA across all three initial guesses used. These three initial guesses were selected to test the performance of the two algorithms when they are close to the upper and lower bounds, as well as in the center of the feasible domain. In addition, all three initial guesses see a similar trend with regards to the perturbation number. As the perturbation number increases, the difference in optimal value between the two algorithms decreases. This indicates that Adam-SPSA performs well even with a low number of perturbations.

The reactive case NPV for Case Study 1 is 25.66 MM USD. The oil recovery factor (%) is plotted against the total water production (STB) for the reactive case, Adam-SPSA using perturbation of 1, and Adam-

SPSA using a perturbation of 200 in Fig. 4. These results represent the best solution found from the optimization runs. Both solutions from Adam-SPSA are much more efficient than the reactive case, as they obtain higher oil recoveries for the same volume of produced water. For Adam-SPSA using perturbation number of 200, to reach an oil recovery of 25%, 122,385 STB of water is produced. On the other hand, for the same oil recovery the reactive case produces a volume of 229,538 STB of water. This solution results in a reduction of water production by 47%. By the same token, when using a perturbation number of 1, the optimized solution results a reduction of 36% in total volume of produced water when compared to the reactive case. This indicates that the results when using a higher perturbation number are slightly higher. Yet, both optimized solutions are significantly superior to the reactive case.

These observations are supplemented by the results shown in Fig. 5, which depicts oil recovery (%) as a function of water injection total (STB). The solution obtained by Adam-SPSA using 200 perturbations resulted in a 12% reduction in the total water injected (838,394 STB) for a 25% oil recovery. Fig. 5 also shows that for any volume of injected water, Adam-SPSA results in a higher oil recovery regardless of perturbation number. This indicates that Adam-SPSA is able to find a solution that results in a more efficient water-flooding by reducing the volumes of produced and injected water. This is owing to the economics of the case study which has relatively high water injection and handling costs. As such, it is more favorable to reduce these volumes, which the reactive case does not consider. This indicates that the water-flooding is more efficient with the optimized solution as it is delaying water breakthrough and reduces the volume of produced water, hence reducing the costs associated with handling the produced water.

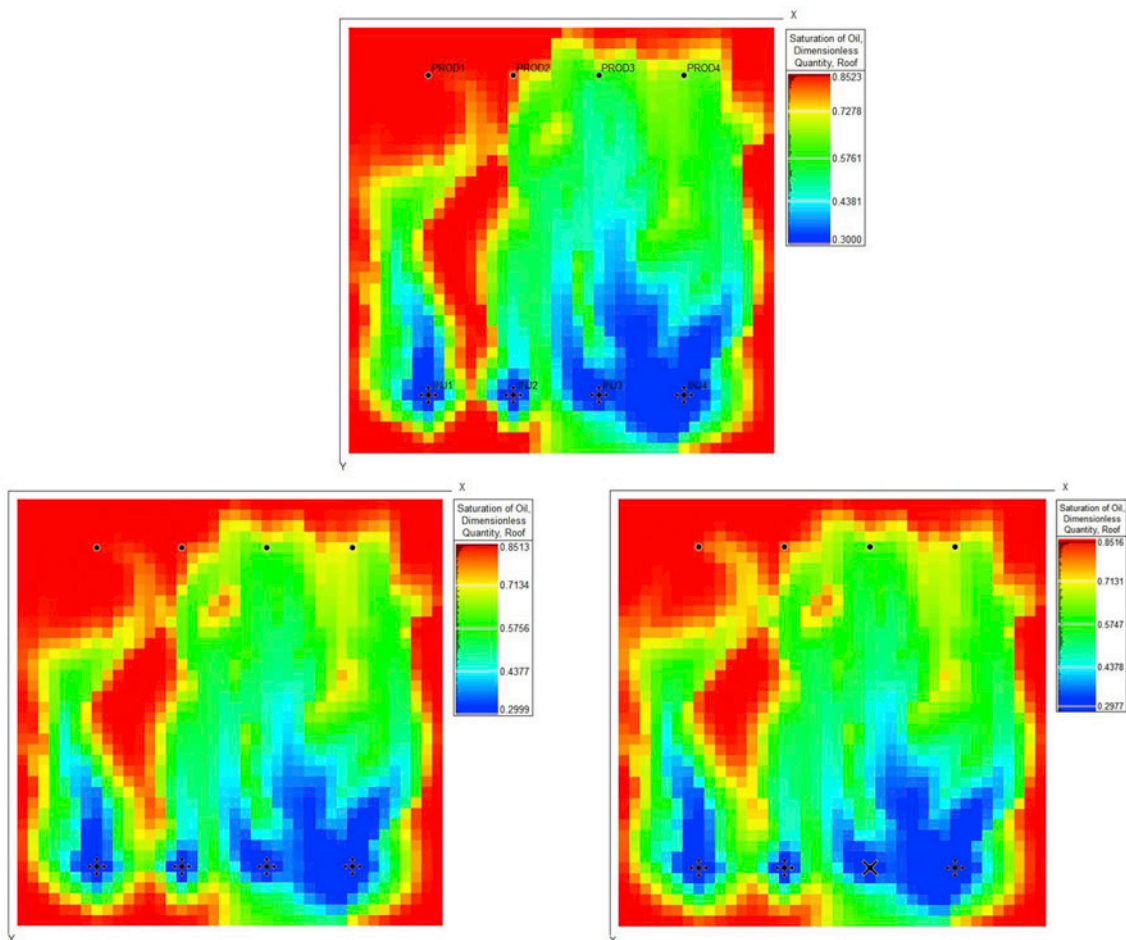


Fig. 6. Oil saturation for case study 1 for the (a) reactive case, (b) Adam-SPSA using a perturbation number of 1, and (c) Adam-SPSA using a perturbation number of 200.

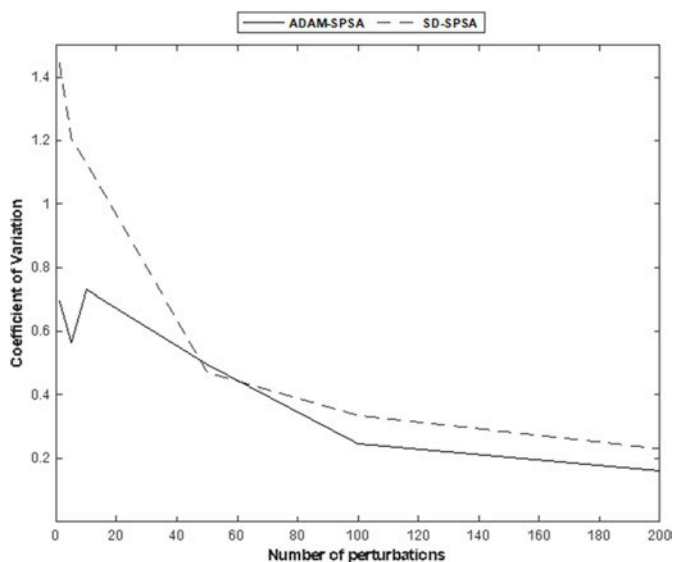


Fig. 7. The coefficient of variation for both algorithms at different values perturbation number.

The improved efficiency of the optimization solutions can be seen on a field scale in Fig. 6. Fig. 6a shows the sweep efficiency of the reactive case once it reaches an oil recovery of 25%. This allows the comparison to Fig. 6b and c which represent the field at the end of the reservoir lifespan when using perturbations of 1 and 200, respectively. The increased water production seen in Fig. 4, can be attributed to the differences in sweep efficiency between Fig. 6a–c. The sweep efficiency between injectors 1 and 2 are at various stages in all three figures. The section of bypassed oil increases in size from Adam-SPSA using 200 perturbations, to 1 perturbation, and then reactive case. This indicates that the optimized solutions ensure a more efficient water-flood is undertaken to allow for a better sweep of the field.

A trend between the number of perturbations and the final optimal value can be extracted from the results. Fig. 7 shows the coefficient of variation for the results presented in Fig. 2 for both algorithms. The coefficient of variation (CV) is defined as the ratio of standard deviation to the mean (average). The points used to calculate the CV are the final optimal solutions obtained by each algorithm for each of the 5 individual runs. The results indicate that for a low number of perturbations,

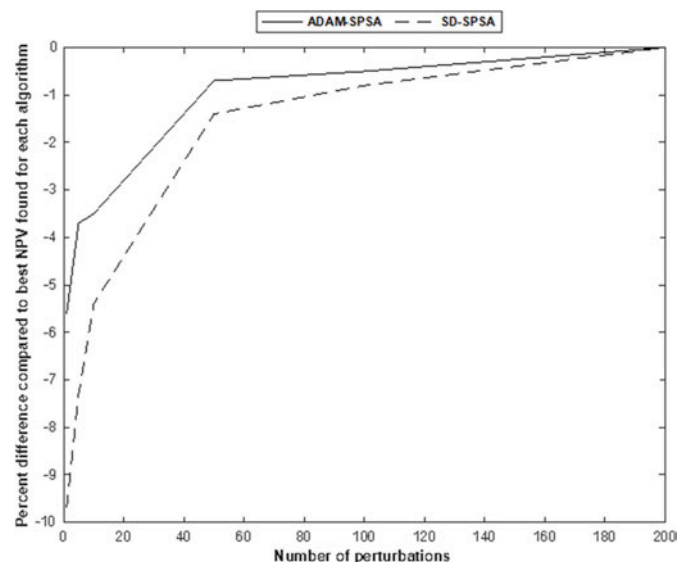


Fig. 8. Percent difference between the best found NPV and best NPV at each perturbation number for each algorithm.

SD-SPSA illustrates the effect of having a stochastic gradient approximation which results in a high coefficient of variation. As the number of perturbations increases, the coefficient of variation decreases. This is expected as the approximation of the gradients tends to be closer to the true gradient for higher perturbation numbers. On the other hand, Adam-SPSA has a relatively low coefficient of variation even for a low number of perturbations. Similarly, it shows a decreasing trend with an increase in the number of perturbations. This reflects that the Adam update rule, which takes into account the variances in gradients for each variable has a positive effect on the search progression. From Fig. 7, it is evident that the update rule utilized by Adam-SPSA, which includes estimating the first and second order moments of the gradient help improve the gradient approximation. As a result, the stochasticity decreases, yet this improves the optimum NPV, as reflected in Fig. 2.

In addition, as the number of perturbations increases the final optimum tends to increase up until a ‘turning point’. At this turning point, an increase in number of perturbations does not necessarily result in a higher optimum NPV due to a decrease in efficiency. Fig. 8 summarizes these results by reflecting the percent difference between the best optimum NPV (of the 5 individual runs) for each number of perturbations. For example, the percent difference between the optimum NPV using 1 perturbation and 200 perturbations when using SD-SPSA is almost 10%. Conversely, the percent difference for the same comparison for Adam-SPSA is 5.6%. This indicates that Adam-SPSA is a superior algorithm since for any number of perturbations it is closer to the most optimum NPV. By contrast, SD-SPSA results reflect it is more sensitive to the number of perturbations, which cannot be ascertained beforehand. This is revealed by the lower slope in Fig. 7 for Adam-SPSA compared to the slope for SD-SPSA, especially for lower values of number of perturbations.

4.2. Case study 2 –Brugge model

4.2.1. Problem description

The Brugge model is a synthetic three-dimensional (3D) reservoir model analogous to those found in the North Sea (Peters et al., 2010). The field is characterized by an elongated half-dome stretching East to West bounded by a large fault on the northern edge. An internal fault is also present with a throw angle of 20° to the boundary fault. The up-scaled reservoir model is composed of 60,048 grid blocks with 44, 550 active cells. The field’s dimensions are roughly 10 km × 3 km with an average thickness of 50 m. The reservoir is composed of four stratigraphic units, from top to bottom: Schelde, Mass, Waal, and Schie. The field is populated with 30 wells, 20 producers and 10 injectors. All injectors are perforated and have open completions in all nine layers. On the other hand, producers are perforated in the top eight layers and have completions detailed in Table 3. The reservoir structure and well locations are shown on the oil saturation distribution in Fig. 9. The leasing life of the field is 30 years. A suite of 104 geological realizations were generated as a part of the available data set provided by TNO to allow for the incorporation of geological uncertainty into the optimization problem. Further information regarding the details of the Brugge model can be found in Peters et al. (2010).

The Brugge model was created as part of an exercise focusing on closed-loop reservoir management. As such, the exercise was composed

Table 3
Brugge model well completions.

Well	Stratigraphic Units			
	Schelde, layers 1-2	Mass, layers 3-5	Waal, layers 6-8	Schie, layer 9
Producers 1-4,6-8,11-13,16-20	Open	Open	Open	Closed
Producers 5,10,14,15	Open	Open	Closed	Closed
Producer 9	Open	Closed	Closed	Closed
Injectors	Open	Open	Open	Open

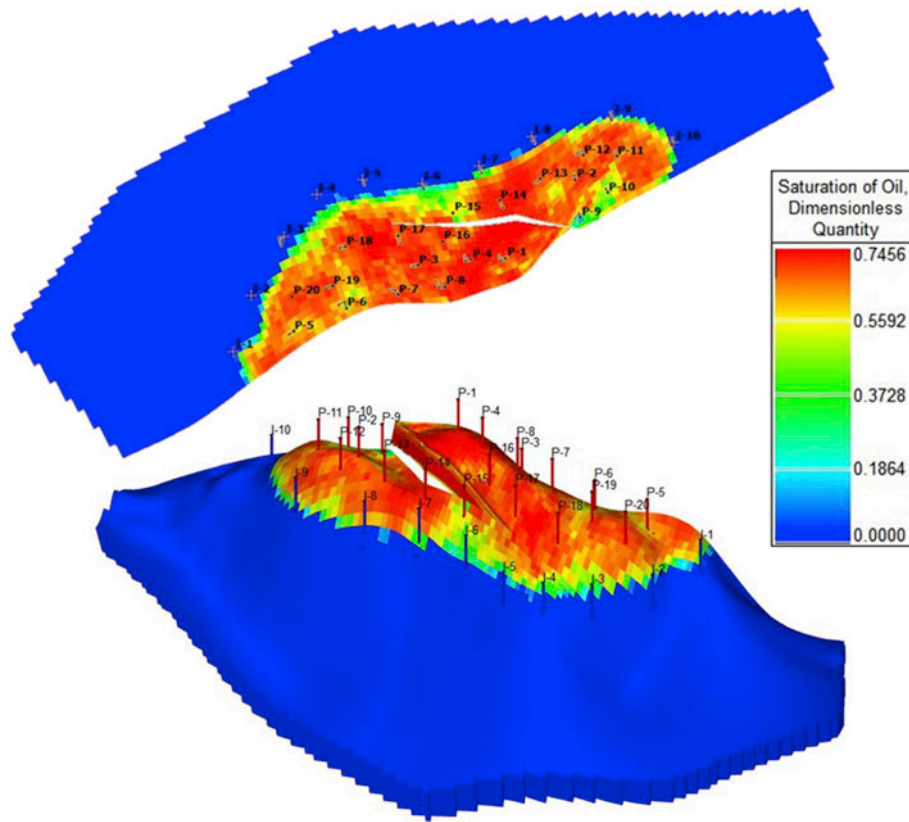


Fig. 9. Brugge model in 2D (top) and 3D (bottom) views showing initial oil saturation distribution with 20 producers and 10 injectors.

Table 4
Economic and reservoir simulation parameters for Brugge model.

Parameter	Value
Initial conditions	170 bars (2611 psi) at 1700 m (5577 ft) depth Free water level: 1678 m (5505 ft)
Pore compressibility	3.5×10^{-6} 1/psi
Fault multiplier	1
Liquid rate constraints	Producer: 477 m ³ /day Injector: 636 m ³ /day
BHP Range	Producers: 50–100 bars Injectors: 100–180 bars
Economic parameters	Oil price: \$503/m ³ (\$80/bbl) Water cost (injection and handling): \$31.45/m ³ (\$5/bbl) Discount rate: 10%

of two individual tasks: a history matching of the model for 10 years of production followed by production optimization for 20 years based on these updated models. However, as this paper's focus is on production optimization, the history matching task was not considered. Consequently, the reservoir lifespan considered in the production optimization problem encompassed a total of 30 years.

The production optimization problem involves the optimization of the BHP's of the 30 wells over a period of 10,800 days (~30 years). The control variables are controlled every 180 days (~6 months) during the simulation lifespan leading to a total of 60 control steps. As a result, there are a total of 1800 (30 wells * 30 years * 2 steps/year) optimization (decision) variables. The optimization variables are bounded by

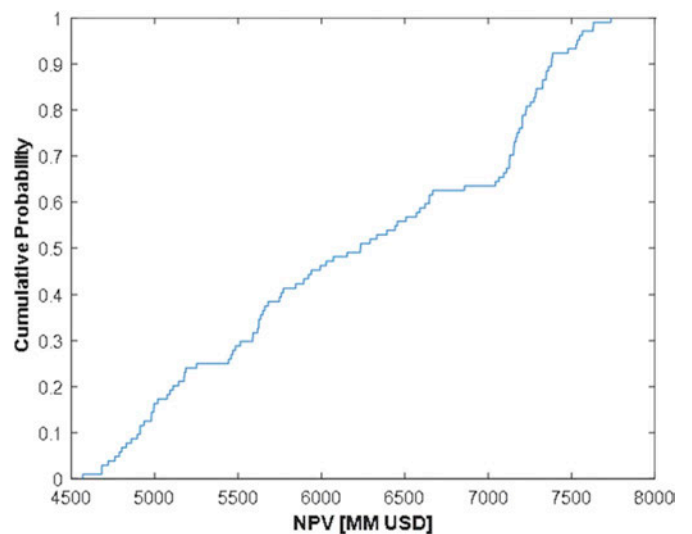


Fig. 10. Cumulative distribution function of the reactive case for all 104 geological realizations.

lower and upper limits. The lower and upper limits for the producers are 50 bars (725 psi) and 100 bars (1450 psi). On the other hand, the lower and upper limits for the injectors are 100 (1450 psi) and 180 bars (2611 psi). These values were chosen based on the initial reservoir pressure of

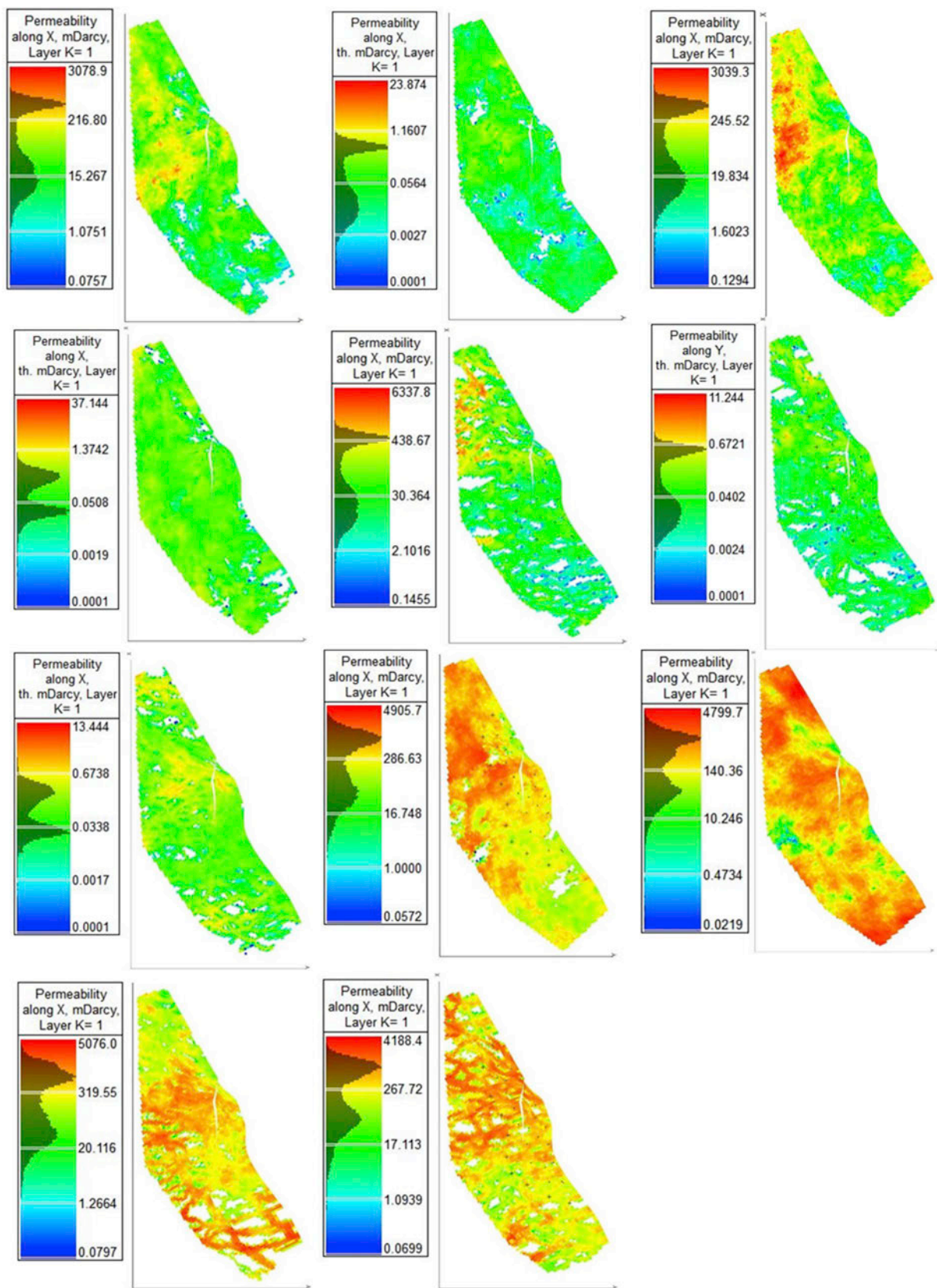


Fig. 11. Permeability fields in x-direction of the top layer for the 11 selected geological realizations.

170 bars (2466 psi). The production wells are constrained by a liquid production rate of $477 \text{ m}^3/\text{day}$ (3000bbl/day), whilst the injectors are constrained by a liquid injection rate of $636 \text{ m}^3/\text{day}$ (4000bbl/day). The reservoir simulation and economic parameters are summarized in Table 4.

The aim of this problem is to investigate the application of Adam-SPSA to a situation where multiple geological realizations are available. To decrease the effect of the stochastic nature of the algorithms, the results are an average of 5 individual optimization runs. To allow this, the computational demand was reduced by selecting a subset of 11 geological realizations from the suite of 104 that are available. Furthermore, the aim was to compare the algorithms behavior across a wide array of geological scenarios in order to challenge the algorithms. As such, the selection process was not random, rather based on the cumulative distribution function (CDF) of the NPV from the reactive case of all the 104 realizations. The reactive case in Peters et al. (2010) used a water-cut of 94%. The liquid rate constraints are similar to those listed in Table 4. Fig. 10 shows the CDF of the reactive case NPVs. The realizations corresponding to $P_1, P_{10}, P_{20}, P_{30}, P_{40}, P_{50}, P_{60}, P_{70}, P_{80}, P_{90}$ and P_{99} were selected. This is similar to other ranking methods in the literature used to reduce the number of geological realizations (Ballin et al., 1992; Deutsch and Begg, 2001; McLennan and Deutsch, 2005).

These 11 percentiles were chosen to reflect the full spectrum of NPVs and geological realizations, including the maximum (P_{99}) and minimum (P_1) percentiles. The 11 selected geological realizations are shown in Fig. 11, where the x-directional permeability field in the top layer are shown. Note the palette to the left of the permeability field is the corresponding color map. As is reflected visually from Fig. 11, there is a wide array of permeability fields of varying levels.

4.2.2. Algorithm parameter settings and constraint handling

A similar sensitivity analysis was undertaken for the second case study to find suitable values for the step sizes in SD-SPSA and Adam-SPSA. Values of $\alpha = 1$ for SD-SPSA and $\alpha = 0.03$ for Adam-SPSA were selected. The value for the perturbation size is similar to the previous case. The projection method was used as the constraint handling technique. The results shown are an average of 5 individual optimization runs. The stopping criterion for all the experiments, unless otherwise stated, is a maximum number of function evaluations of 5000.

4.2.3. Results of case study 2

The numerical experiments that were run for the Brugge model

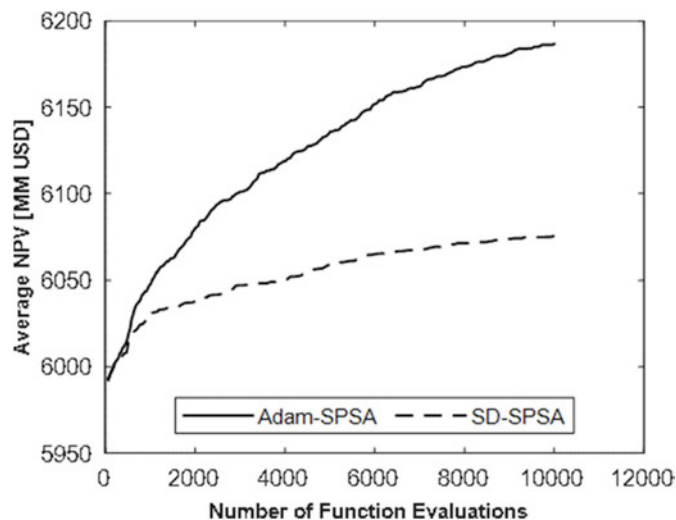


Fig. 12. Production optimization for Brugge Model using a perturbation number of 1.

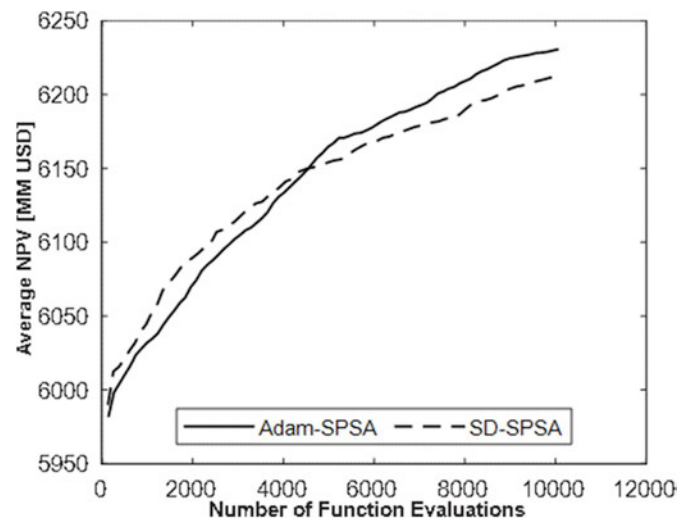


Fig. 13. Production optimization for Brugge Model using a perturbation number of 5.

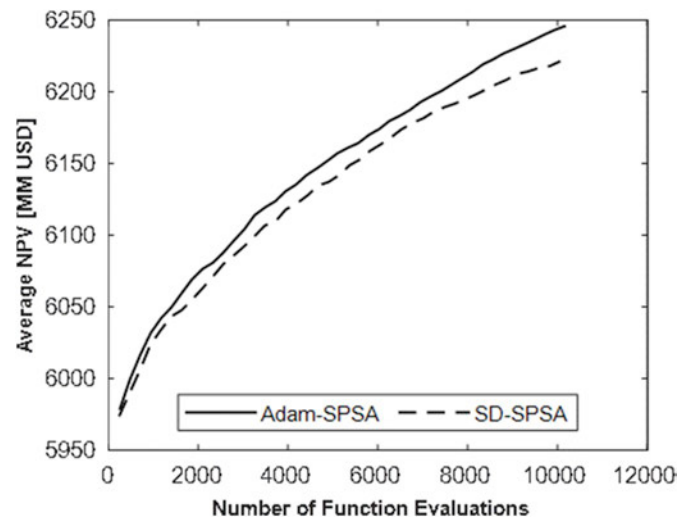


Fig. 14. Production optimization for Brugge model using a perturbation number of 10.

consisted of a sensitivity analysis with regards to the number of perturbations for a given initial guess. As before, five individual optimization runs are averaged to account for the stochastic nature of the optimization algorithms. The reactive case defined in Peters et al. (2010) was used for comparisons in this study. This resulted in a NPV of $\$ 6.14 \times 10^9 \text{ USD}$ for the reactive case for the subset realizations and an NPV of $6.22 \times 10^9 \text{ USD}$ for the full set. All NPVs that are stated in this section are an average NPV for the 11 selected realizations, unless otherwise stated.

It is important to note that all results presented in Peters et al. (2010) that used one inflow control valve (ICV), albeit for a shorter reservoir life span, resulted in NPV values lower than the reactive case. Also, it is evident from the results presented for the production optimization in Peters et al. (2010) that the reactive case is already close to the optimal. As such, the initial guess selected for this study is, on a normalized scale, 0.8 for injectors and 0.2 for producers. It should also be noted that a stopping criterion of 10,000 function evaluations was used for all the following numerical experiments, given this study has more variables.

Subsequently, the results of the production optimization using the Adam-SPSA and SD-SPSA algorithms are presented in Figs. 12-16. Fig. 12 presents the results when using only 1 perturbation for the gradient calculation for both algorithms. It is clear that the proposed algorithm, Adam-SPSA, outperforms the standard SD-SPSA algorithm with regards to both optimal value obtained and convergence speed. The optimal value found for Adam-SPSA is $\$ 6.19 \times 10^9 USD$, whilst for SD-SPSA this value is $\$ 6.08 \times 10^9 USD$. This indicates that Adam-SPSA, on average, outperformed SD-SPSA by almost 2% with regards to the optimal value. In addition, Adam-SPSA required 1903 function evaluations to reach the optimal value found by SD-SPSA, which required 10,000 function evaluations. For this problem, on average, one function evaluation requires 3 min to run. Hence, a reduction of 8097 function evaluations translates to a reduction of 404 CPU-hours.

In Fig. 13, the result for the production optimization using 5 perturbations are presented. The difference in performance of both algorithms has decreased in comparison to the results in Fig. 12, however; both algorithms have improved on their respective optimal values. The optimal value found by Adam-SPSA and SD-SPSA are $\$ 6.23 \times 10^9 USD$ and $\$ 6.21 \times 10^9 USD$, respectively. Although the difference is minimal, Adam-SPSA still outperforms SD-SPSA. With regards to convergence speed, Adam-SPSA reaches the optimal value found by SD-SPSA in 8129 function evaluations. This takes SD-SPSA 10,120 function evaluations to reach this value. This is a speed up of 20% with respect to the required number of function evaluations.

The results for production optimization of the Brugge model when using 10 perturbations are presented in Fig. 14. The optimal value found by Adam-SPSA is $\$ 6.25 \times 10^9 USD$, whilst for SD-SPSA the optimal value is $\$ 6.22 \times 10^9 USD$. This is a larger difference than the results for 5 perturbations, however; the difference is still not as large as when using 1 perturbation. In addition, albeit minimally, there is an increase in the optimal value found for both algorithms compared to using 5 perturbations. With regards to convergence, Adam-SPSA required 8338 function evaluations to achieve the optimal value of SD-SPSA, which took 10,070 function evaluations, resulting in a 17% savings.

It is evident from the results presented in Figs. 13 and 14 that a turning point with regards to the most optimal number of perturbations has been reached. Further experiments were conducted using an increasing number of perturbations, specifically 50, 100 and 200. However, it was evident that these large numbers of perturbation

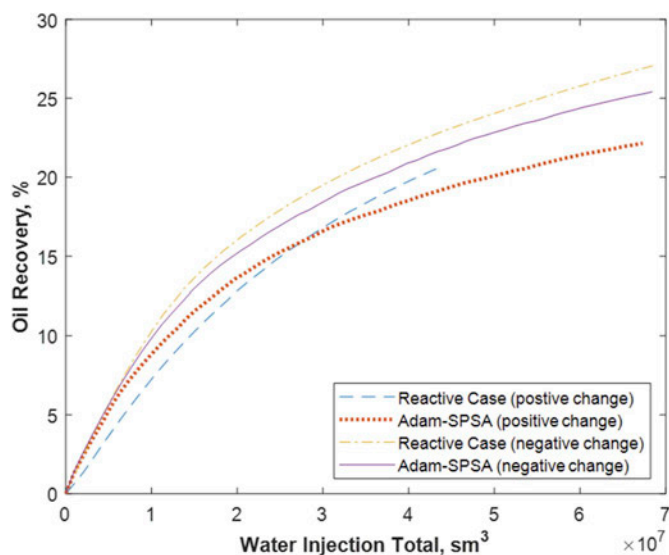


Fig. 16. Oil recovery (%) as a function of total water injected (m^3) for the realizations showing greatest change in NPV for reactive case and Adam-SPSA solution.

rendered both algorithms less efficient for this complex case study. As such, these results were omitted from this paper.

The different production strategies can have a significant effect on the economic performance of the Brugge model. Contrary to case study 1, the economic parameters in case study 2, such as oil price, water handling costs, and discount rate incentivize an increase in oil production, especially at earlier times. The expected NPV of the optimized solution results in a 2% improvement over the reactive case subset NPV. However, it does not necessarily reflect an improvement in the NPV of each realization. The change in NPV across the individual realizations range from an increase of 23% to a decrease of 7%. It is interesting to note that the order of these changes matches the order of the reactive case NPV, i.e. from P_1 to P_{99} . This is an indication to the robustness of the solution across the 11 selected realizations. The changes in individual NPVs can be described with the aid of Fig. 15 and Fig. 16.

Fig. 15 shows the oil production over the production life for the four cases of interest. Firstly, the realization which showed the largest decrease in NPV is represented by the yellow and purple lines for the reactive case and Adam-SPSA solution, respectively. It is clear that the decrease in NPV for this realization is a result of a 6% decrease in total oil production. On the other hand, the realization which showed the largest improvement consistently produces more oil under the optimized solution (orange dotted line) when compared to the reactive case (blue dashed line). There is an 8% difference in the volume of total oil production between the two solutions.

The water-flooding efficiency of the two solutions are compared in Fig. 16. As expected, there is an efficiency decrease between the solutions for the realization that showed a decrease in NPV. On the other hand, there is a turning point when comparing the two solutions for the realization that showed an increase in NPV. As opposed to case study 1, the water injection costs are low in comparison to the oil price. This motivates the early production of oil even if it comes at the expense of sweep efficiency, as shown in Fig. 16. However, the final oil recovery (%) of the Adam-SPSA (orange line) is still greater than the reactive case (blue line).

5. Discussion

The selection of parameters used in optimization algorithms can have a major impact on the search progression and hence the results.

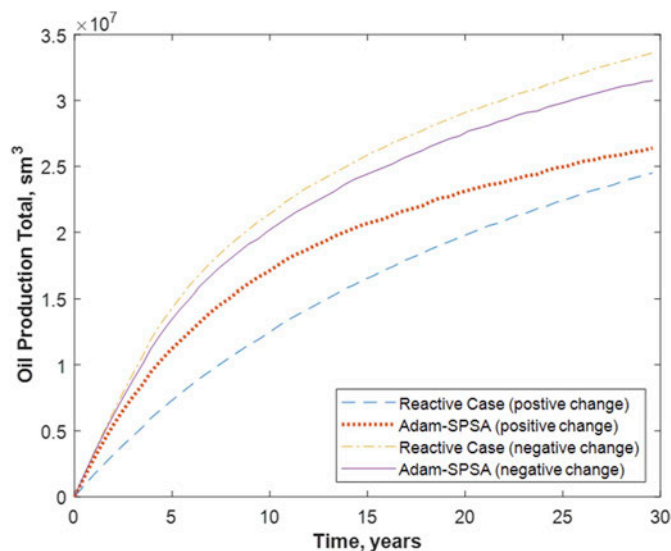


Fig. 15. Oil production for the individual realizations which showed the greatest improvement and decrease for the reactive case and Adam-SPSA solution.

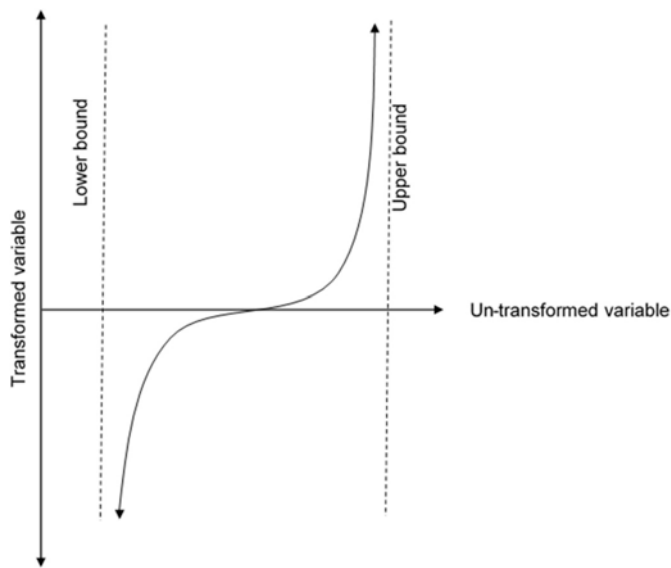


Fig. 17. Typical shape of the logarithmic transformation method.

Our aim when comparing the two algorithms was to ensure a fair comparison, allowing solid conclusions to be made. As such, we wanted to ensure that the driving force of the optimization algorithms, which is the product of the step size and the search direction, are relatively similar. This allows the conclusions that are made to be an effect of the algorithm themselves rather than a difference in step size or perturbation size. The first case study was used to perform preliminary sensitivity analysis on the driving force. To find the effect of the driving force on the proposed solutions, the average absolute difference between the second and third iterations was used. This gives an indication of the movement towards the boundaries at the beginning of the optimization run. If a step size accelerates the movement towards the boundaries, then it is considered to be inefficient since the search space is not explored before reaching the boundaries. The opposite is also true for a small step size which does not progress rapidly enough in the search space. This preliminary study is not computationally expensive, yet can have significant impact on the full optimization runs, especially when comparing the efficiency of the algorithms.

Another significant finding was the difference in effect of the bound constraint handling technique on both algorithms. The logarithmic transformation method relies on transforming the search domain to infinite bounds from a lower and upper bound. Although this changes a constrained optimization problem into an unconstrained one, it comes with its significant challenges. *Sisser (1981)* introduced the Local Mapping Conditions to guarantee that a local minimum of the transformed problem satisfies the first-order necessary conditions for the local minimum of the original problem. One of these conditions is the requirement for the transformed variable to be greater than or equal to zero for the search space domain. However, this condition is not met when utilizing the logarithmic transformation method since its lower bound is negative infinity.

Furthermore, if the Local Mapping Conditions are assumed to be met, there is still the possibility that the search would get stuck at the boundary and converge to a non-optimal stationary point even if indeed the solution does not lie on the boundary (*Sisser, 1981*). Based on our experiments there are two reasons this can occur when using the logarithmic transformation method. Firstly, as the search approaches the boundaries a large perturbation size is required when approximating the gradient that will reflect a reasonable change in objective value. If the perturbation size is too small, this leads to no considerable change in objective value, which results in the gradient prematurely approaches

zero. Secondly, if the search direction is steered towards the boundaries, it will then require a very large step size to return from the boundaries. This is indicated in *Fig. 17* by the sharp increase in the slope of the curve closer to the bounds. Although both algorithms were negatively affected when using the logarithmic transformation method, Adam-SPSA was still able to outperform SD-SPSA for all runs. As a result, the projection method resulted in improvements of up to 5% with Adam-SPSA and up to 4% with SD-SPSA over the logarithmic method. This indicates that the adaptive search direction employed in Adam-SPSA can help limit the negative effect of the logarithmic transformation method, but not completely eliminate it. On the other hand, when using a projection method, both algorithms have the ability to search the boundaries without the same risk of getting trapped. Hence, the optimization runs utilizing the projection method outperformed those of the logarithmic transformation method. Although this may be problem dependent, it indicates the importance of selecting an appropriate bound constraint handling technique.

The advantages of Adam-SPSA over SD-SPSA are reflected in the results presented for both case studies. Adam-SPSA outperforms SD-SPSA in both case studies, especially when the perturbation number is small. As the perturbation number increases, the approximation of the gradient improves leading to a significant improvement in convergence speed. However, this improvement is less significant for Adam-SPSA than it is for SD-SPSA. In other words, the effect of changing perturbation number is less substantial when using Adam-SPSA. This can be accredited to the fact that Adam-SPSA accounts for previous gradients, as well as variances, before computing the next search direction. This in itself can act as an averaging of gradients, albeit not a replacement for it, which seems to produce more representative search directions which improve the results over SD-SPSA. This indicates that regardless of the perturbation number used for Adam-SPSA, it is more likely to find a similar quality solution to the most optimal solution when compared to SD-SPSA. The percent difference between the optimum value using 1 perturbation and 200 perturbations when using SD-SPSA is almost 10%. Conversely, the percent difference for the same comparison for Adam-SPSA is 5.6%. This can be especially useful when resources are limited and preliminary experiments cannot be performed to find the most ideal number of perturbations. In addition, the optimal value for perturbation number can be problem dependent as shown by the differences in results for the case studies presented. These results are in line with literature (*Do and Reynolds, 2013; Wang et al., 2009*). As such, the insensitivity of Adam-SPSA to perturbation number, when compared to SD-SPSA, becomes a pivotal advantage when selecting a preferable optimization algorithm.

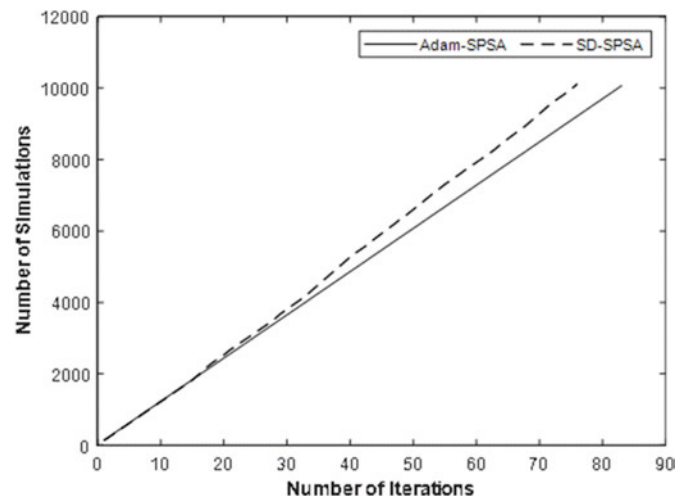


Fig. 18. Computational time reduction per iteration due to parallelization.

In addition, the proposed algorithm outperformed the reactive cases for each of the case studies. The economics of the first case study required a more efficient water-flooding scenario, to increase the oil recovery for a given volume of injected water. On the other hand, the second case study rewarded solutions which prioritized early oil production over sweep efficiency. The superior performance of Adam-SPSA was reflected by up to 5% improvement over SD-SPSA for case study 1 and up to 2% for case study 2 with regards to optimal value. In addition, Adam-SPSA almost always required a lower number of function evaluations to reach a predefined objective value. Adam-SPSA showed an improvement of up to 91% and 81% in convergence speeds for case studies 1 and 2, respectively. Furthermore, Adam-SPSA constantly outperformed the reactive cases by 44% and 2% for case studies 1 and 2, respectively. However, SD-SPSA was not able to outperform the reactive case of case study 2 for low values of perturbation number.

Another advantage of Adam-SPSA is its ability to be easily parallelized. As previously mentioned, all gradient approximations were parallelized. Hence, each iteration of Adam-SPSA was an equivalent of only 2 reservoir simulations, one for the forward and backward perturbation in the gradient approximation, and one for evaluating the new iterate solution. On the other hand, due to the presence of a line-search, each iteration of SD-SPSA had an equivalent of up to 7 reservoir simulations. Fig. 18 shows the number of reservoir simulations for each iteration as the optimization progresses. This figure is an optimization run for case study 2 when using 5 perturbations. At the beginning of the optimization run, SD-SPSA did not require a line-search as indicated by the overlap of both lines. This indicates that the approximated gradients were in an uphill direction, hence not requiring any change to the step size. However, as the search progresses, there is a deviation of the lines around iteration number 30. This divergence is a clear symptom of the SD-SPSA algorithm requiring a backtracking line-search to find a better objective function value. This divergence gap widens as backtracking line-search is consistently undertaken at the later stages of the search. Since the reservoir simulations that are conducted during a backtracking line-search are dependent, this section of the process cannot be parallelized.

Although the Adam framework was applied to the SPSA gradient approximation, there is no reason for it not to show similar results with other aforementioned gradient approximations. Do and Reynolds (2013) showed the theoretical similarities between common gradient approximations, including SPSA, and as such it is not a stretch to assume that the advantages of Adam would be seen with other gradient approximations.

6. Conclusions

This work introduced an accelerated gradient descent framework to solve well control optimization problems for which approximated gradients are required. In this work, the gradients are approximated by SPSA. The proposed framework utilized estimates of the first and second moments of gradient approximations to improve progression during the optimization process. This study compared the proposed algorithm (Adam-SPSA) with the conventional steepest descent SPSA (SD-SPSA) using two case studies. The comparison included studying the effect of different settings on the performance of each algorithm. The following settings were investigated: (i) the number of perturbations, (ii) initial guess, (iii) bound constraint handling technique and (iv) randomness (seed number). The following conclusions can be made from the study.

- Adam-SPSA consistently outperformed SD-SPSA, both in terms of optimal value, up to 5% for case study 1 and up to 2% for case study 2, and computational savings, up to 91% and 81% for case studies 1 and 2, respectively.

- Furthermore, Adam-SPSA outperformed the reactive strategy by 44% and 2% for case studies 1 and 2, respectively.
- Both algorithms improved with an increase in perturbation number up to a turning point. However, SD-SPSA was more sensitive to the increase.
- The results showed that the optimal number of perturbations is problem-dependent, and given the optimal number cannot be known priori, Adam-SPSA is a more superior choice.
- Results also showed that neither algorithm was sensitive to the initial guess. However, Adam-SPSA outperformed SD-SPSA regardless of starting point.
- Both algorithms performed better with projection constraint handling technique compared to the logarithmic transformation technique.
- Adam-SPSA was less sensitive to the randomness of the gradient approximation reflected by values of the coefficient of variation (CV). The CV of Adam-SPSA for low values of perturbation number was half of those for SD-SPSA.

Acknowledgements

The authors would like to thank Steve Begg for his support and fruitful discussion. The authors are also grateful to Rock Flow Dynamics for providing licenses for tNavigator. This work was supported with supercomputing resources provided by the Phoenix HPC service at the University of Adelaide.

Appendix

A.1 Comparison of constraint handling techniques

In Fig. A-1a, the optimal value obtained by Adam-SPSA using projection method was 35.02 MM USD, whilst this value was 33.42 MM USD when using the logarithmic transformation. This is around a 4.8% difference between the optimal values. When considering the results of SD-SPSA, optimal values of 33.33 MM USD and 32.98 MM USD were obtained for projection method and logarithmic transformation method, respectively. This represents, on average, a 1% difference between the two methods when using only 1 perturbation for gradient approximation.

A similar trend is reflected in the results presented in Fig. A-1b when using a perturbation number of 5. The optimal values of Adam-SPSA when using the projection method and logarithmic transformation method are 35.51 MM USD and 33.84 MM USD, respectively. Although these individual values are an improvement to results in A-1a, the percent difference is still around 4.9%, on average. In regard to SD-SPSA, the results of optimal values have also improved to 34.19 MM USD and 33.3 MM USD for the projection and logarithmic transformation methods, respectively. However, this is a difference of 2.7% between the two bound constraint handling techniques.

Fig. A-1c displays the results when using a perturbation number of 10. Again, there is a slight improvement with regards to the optimal values obtained by Adam-SPSA of 35.76 MM USD and 34.00 MM USD when using the projection and logarithmic transformation methods, respectively. Based on these average results, this is a 5.2% difference between the constraint handling techniques. Similarly, the results of SD-SPSA show a 3.8% difference between the results obtained when using the two methods. An optimal value of 34.85 MM USD was obtained when applying the projection method. On the other hand, an optimal value of 33.59 MM USD was obtained when utilizing the logarithmic transformation method.

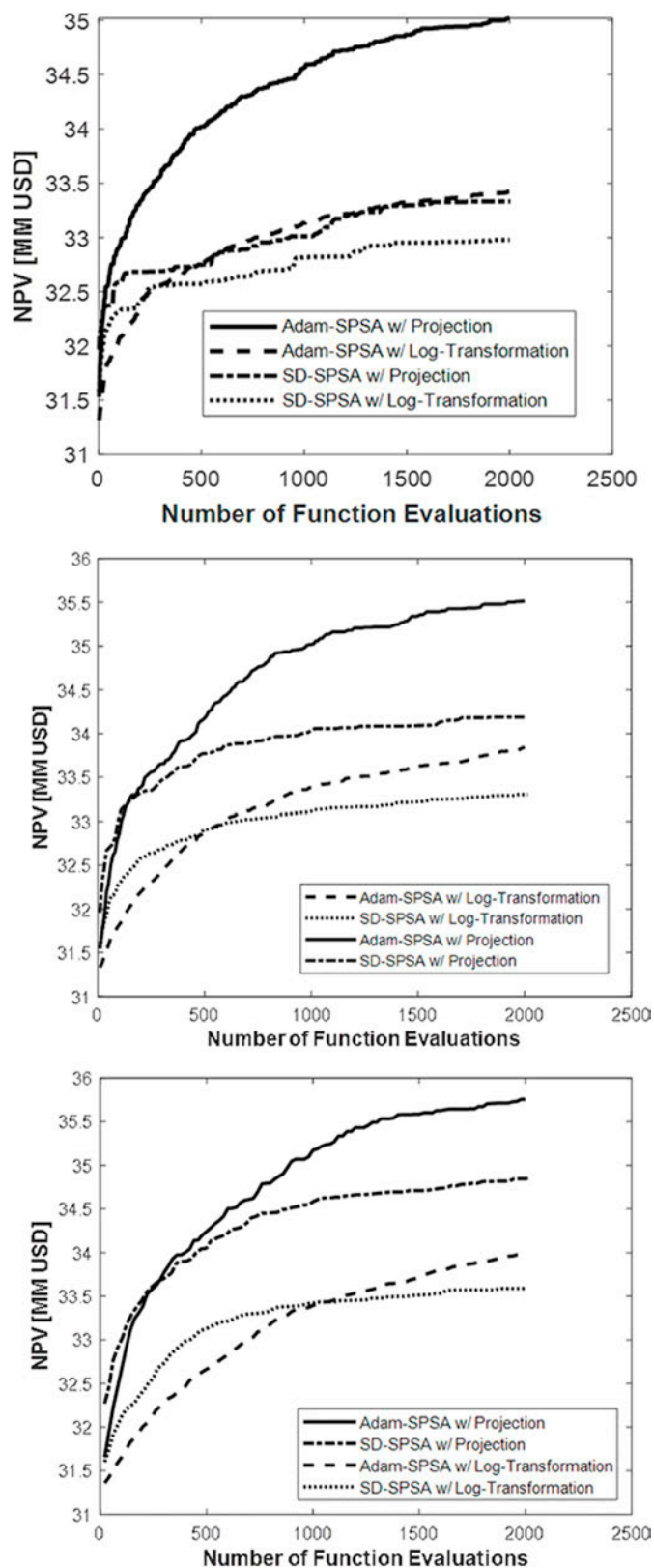


Fig. A-1. Investigation into the effect of bound constraint handling methods for (a) 1, (b) 5 and (c) 10 perturbations for SPSA and Adam-SPSA

References

- Asadollahi, M., Naevdal, G., 2009. Waterflooding optimization using gradient based methods. In: SPE/EAGE Reservoir Characterization & Simulation Conference.
- Asadollahi, M., Naevdal, G., Dadashpour, M., Kleppe, J., 2014. Production optimization using derivative free methods applied to Brugge field case. *J. Pet. Sci. Eng.* 114, 22–37.
- Asheim, H., 1988. Maximization of water sweep efficiency by controlling production and injection rates. In: European Petroleum Conference. Society of Petroleum Engineers, London, United Kingdom, p. 10.
- Ballin, P., Journel, A., Aziz, K., 1992. Prediction of uncertainty in reservoir performance forecast. *J. Can. Pet. Technol.* 31 (4).
- Brouwer, D.R., Jansen, J.D., 2004. Dynamic Optimization of Waterflooding with Smart Wells Using Optimal Control Theory.
- Brouwer, D.R., Naevdal, G., Jansen, J.D., Vefring, E.H., van Kruijsdijk, C.P.J.W., 2004. Improved reservoir management through optimal control and continuous model updating. In: SPE Annual Technical Conference and Exhibition. Society of Petroleum Engineers, Houston, Texas, p. 11.
- Chaudhri, M.M., Phale, H.A., Liu, N., Oliver, D.S., 2009. An improved approach for ensemble-based production optimization. In: SPE Western Regional Meeting. Society of Petroleum Engineers, San Jose, California.
- Chen, B., Fonseca, R.-M., Leeuwenburgh, O., Reynolds, A.C., 2017. Minimizing the Risk in the robust life-cycle production optimization using stochastic simplex approximate gradient. *J. Pet. Sci. Eng.* 153, 331–344.
- Chen, X., Liu, S., Sun, R., Hong, M., 2018. On the Convergence of a Class of Adam-type Algorithms for Non-convex Optimization arXiv preprint arXiv:1808.02941.
- Chen, Y., Oliver, D.S., Zhang, D., 2009. Efficient ensemble-based closed-loop production optimization. *SPE J.* 14 (4), 634–645.
- Dehdari, V., Oliver, D.S., 2012. Sequential quadratic programming for solving constrained production optimization—case study from Brugge field. *SPE J.* 17 (3), 874–884.
- Dehdari, V., Oliver, D.S., Deutsch, C.V., 2012. Comparison of optimization algorithms for reservoir management with constraints—a case study. *J. Pet. Sci. Eng.* 100, 41–49.
- Deutsch, V.C., Begg, S.H., 2001. The Use of Ranking to Reduce the Required Number of Realizations.
- Do, S.T., 2012. Application of Spsa-type Algorithms to Production Optimization. University of Tulsa.
- Do, S.T., Reynolds, A.C., 2013. Theoretical connections between optimization algorithms based on an approximate gradient. *Comput. Geosci.* 17 (6), 959–973.
- Duchi, J., Hazan, E., Singer, Y., 2011. Adaptive subgradient methods for online learning and stochastic optimization. *J. Mach. Learn. Res.* 12 (Jul), 2121–2159.
- Fathi, Z., Ramirez, F.W., 1984. Optimal injection policies for enhanced oil recovery: Part 2—surfactant flooding. *SPE-11285-PA* 24 (3), 333–341.
- Fonseca, R., Leeuwenburgh, O., Della Rossa, E., Van den Hof, P.M., Jansen, J.-D., 2015. Ensemble-based multiobjective optimization of on/off control devices under geological uncertainty. *SPE Reserv. Eval. Eng.* 18 (4), 554–563.
- Fonseca, R., Stordal, A., Leeuwenburgh, O., Van den Hof, P., Jansen, J.D., 2014. Robust ensemble-based multi-objective optimization. In: 14th European Conference on the Mathematics of Oil Recovery (ECMORE XIV), Catania, Sicily.
- Fonseca, R.R.M., Chen, B., Jansen, J.D., Reynolds, A., 2017. A stochastic simplex approximate gradient (StoSAG) for optimization under uncertainty. *Int. J. Numer. Methods Eng.* 109 (13), 1756–1776.
- Foroud, T., Baradaran, A., Seifi, A., 2018. A comparative evaluation of global search algorithms in black box optimization of oil production: a case study on Brugge field. *J. Pet. Sci. Eng.* 167, 131–151.
- Fu, M.C., Hill, D., 1997. Optimization of discrete event systems via simultaneous perturbation stochastic approximation. *IEE Trans.* 29 (3), 233–243.
- Gregor, K., Danihelka, I., Graves, A., Rezende, D.J., Wierstra, D., 2015. Draw: A Recurrent Neural Network for Image Generation arXiv preprint arXiv:1502.04623.
- Hasan, A., Foss, B., 2013. Optimal wells scheduling of a petroleum reservoir. In: 2013 European Control Conference (ECC), pp. 1095–1100.
- Hinton, G., Srivastava, N., Swersky, K., 2012. Neural Networks for Machine Learning Lecture 6a Overview of Mini-Batch Gradient Descent.
- Hou, J., Zhou, K., Zhang, X.-S., Kang, X.-D., Xie, H.J.P.S., 2015. A review of closed-loop reservoir management, 12 (1), 114–128.
- Isebor, O.J., 2009. Constrained Production Optimization with an Emphasis on Derivative-free Methods. Stanford University, Stanford, CA.
- Isebor, O.J., Durllofsky, L.J., Echeverría Ciaurri, D., 2014. A derivative-free methodology with local and global search for the constrained joint optimization of well locations and controls. *Comput. Geosci.* 18 (3), 463–482.
- Jansen, J.D., 2011. Adjoint-based optimization of multi-phase flow through porous media – a review. *Comput. Fluids* 46 (1), 40–51.
- Kingma, D.P., Ba, J., 2014. Adam: A Method for Stochastic Optimization arXiv preprint arXiv:1412.6980.
- Ledig, C., et al., 2017. Photo-realistic single image super-resolution using a generative adversarial network. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 4681–4690.
- Lefebvre, M., 2006. Applied Probability and Statistics, vol. XVIII. Springer-Verlag New York, p. 358.
- Liu, W., Ramirez, W.F., Qi, Y.F., 1993. Optimal control of steamflooding. *SPE-21619-PA* 1 (2), 73–82.
- Lu, R., Forouzanfar, F., Reynolds, A.C., 2017. An efficient adaptive algorithm for robust control optimization using StoSAG. *J. Pet. Sci. Eng.* 159, 314–330.
- McLennan, J., Deutsch, C.V., 2005. Ranking geostatistical realizations by measures of connectivity. In: SPE International Thermal Operations and Heavy Oil Symposium. Society of Petroleum Engineers, Calgary, Alberta, Canada, p. 13.
- Mehos, G.J., Ramirez, W.F., 1989. Use of optimal control theory to optimize carbon dioxide miscible-flooding enhanced oil recovery. *J. Pet. Sci. Eng.* 2 (4), 247–260.
- Nocedal, J., Wright, S., 2006. Numerical optimization, 35. Springer Science, p. 664.
- Peters, L., et al., 2010. Results of the Brugge benchmark study for flooding optimization and history matching. *SPE Reserv. Eval. Eng.* 13 (3), 391–405.
- Ramirez, W.F., Fathi, Z., Cagnol, J.L., 1984. Optimal injection policies for enhanced oil recovery: Part 1 theory and computational strategies. *SPE-11285-PA* 24 (3), 328–332.
- Sadegh, P., 1997. Constrained optimization via stochastic approximation with a simultaneous perturbation gradient approximation. *IFAC Proceedings Volumes* 30 (11), 281–285.
- Sarma, P., Aziz, K., Durllofsky, L.J., 2005. Implementation of adjoint solution for optimal control of smart wells. In: SPE Reservoir Simulation Symposium. Society of Petroleum Engineers, The Woodlands, Texas.
- Sisser, F.S., 1981. Elimination of bounds in optimization problems by transforming variables. *Math. Program.* 20 (1), 110–121.
- Spall, J.C., 1992. Multivariate stochastic approximation using a simultaneous perturbation gradient approximation. *IEEE Trans. Autom. Control* 37 (3), 332–341.
- Spall, J.C., 1998. Implementation of the simultaneous perturbation algorithm for stochastic optimization. *IEEE Trans. Aerosp. Electron. Syst.* 34 (3), 817–823.
- Spall, J.C., Hill, S.D., Stark, D.R., 2006. Theoretical framework for comparing several stochastic optimization approaches. In: Calafiore, G., Dabbene, F. (Eds.), Probabilistic and Randomized Methods for Design under Uncertainty. Springer London, London, pp. 99–117.
- Sudaryanto, B., Yortos, Y.C., 2001. Optimization of displacements in porous media using rate control. In: SPE Annual Technical Conference and Exhibition. Society of Petroleum Engineers, New Orleans, Louisiana, p. 12.
- Virnovsky, G., 1991. Waterflooding Strategy Design Using Optimal Control Theory, IOR 1991-6th European Symposium on Improved Oil Recovery.
- Wang, C., Li, G., Reynolds, A.C., 2009. Production optimization in closed-loop reservoir management. *SPE J.* 14 (3), 506–523.
- Wang, L., Spall, J.C., 2003. Stochastic optimization with inequality constraints using simultaneous perturbations and penalty functions. In: 42nd IEEE International Conference on Decision and Control (IEEE Cat. No.03CH37475), vol. 4, pp. 3808–3813.
- Wu, Y., et al., 2016. Google’s Neural Machine Translation System: Bridging the Gap between Human and Machine Translation arXiv preprint arXiv:1609.08144.
- Yeten, B., Brouwer, D.R., Durllofsky, L.J., Aziz, K., 2004. Decision analysis under uncertainty for smart well deployment. *J. Pet. Sci. Eng.* 44 (1), 175–191.
- Zhao, H., et al., 2013. Maximization of a dynamic quadratic interpolation model for production optimization. *SPE J.* 18 (6), 1012–1025.

4 Efficient techniques for well placement optimization

4.1 An adaptive moment estimation framework for well placement optimization

Arouri, Y. & Sayyafzadeh, M.

Computational Geosciences, 2022.

Statement of Authorship

Title of Paper	An adaptive moment estimation framework for well placement optimization
Publication Status	<input checked="" type="checkbox"/> Published <input type="checkbox"/> Accepted for Publication <input type="checkbox"/> Submitted for Publication <input type="checkbox"/> Unpublished and Unsubmitted work written in manuscript style
Publication Details	Arouri, Y. & Sayyafzadeh, M. An adaptive moment estimation framework for well placement optimization. Computational Geosciences, 2022.

Principal Author

Name of Principal Author (Candidate)	Yazan Arouri		
Contribution to the Paper	Coding, experimentation, analysis of results, writing original manuscript, corresponding author development of algorithm, conceptualization		
Overall percentage (%)	80%		
Certification:	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.		
Signature		Date	10/04/2022

Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

- the candidate's stated contribution to the publication is accurate (as detailed above);
- permission is granted for the candidate to include the publication in the thesis; and
- the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author	Mohammad Sayyafzadeh		
Contribution to the Paper	formulation, supervision, conceptualization, analysis of results, discussions, revisions		
Signature		Date	06/06/2022

Name of Co-Author			
Contribution to the Paper			
Signature		Date	

Please cut and paste additional co-author panels here as required.



An adaptive moment estimation framework for well placement optimization

Yazan Arouri¹ · Mohammad Sayyafzadeh¹

Received: 2 March 2021 / Accepted: 25 January 2022
© The Author(s) 2022

Abstract

In this study, we propose the use of a first-order gradient framework, the adaptive moment estimation (Adam), in conjunction with a stochastic gradient approximation, to well location and trajectory optimization problems. The Adam framework allows the incorporation of additional information from previous gradients to calculate variable-specific progression steps. As a result, this assists the search progression to be adjusted further for each variable and allows a convergence speed-up in problems where the gradients need to be approximated. We argue that under computational budget constraints, local optimization algorithms provide suitable solutions from a heuristic initial guess. Nonlinear constraints are taken into account to ensure the proposed solutions are not in violation of practical field considerations. The performance of the proposed algorithm is compared against steepest descent and generalized pattern search, using two case studies — the placement of four vertical wells and placement of 20 nonconventional (deviated, horizontal and/or slanted) wells. The results indicate that the proposed algorithm consistently outperforms the tested methods in terms computational efficiency and final optimum value. Additional discussions regarding nonconventional parameterization provide insights into simultaneous perturbation gradient approximations.

Keywords Well placement · Gradient-based algorithm · Adaptive moment estimation · Field development optimization

1 Introduction

Field development strategies are a crucial part of the efficient management of fluids in the subsurface. This includes water resources [1], carbon storage in geological formations [2] and hydrocarbon reserves [3–5]. A major component of a field development plan is the configuration of the wells. This may include the number of wells, type (injector or producer) and their trajectories, amongst other considerations. The reservoir response for a set of well configurations is used to assess the suitability for the given objective, such as Net-Present-Value (NPV) or cumulative oil production. Obtaining the reservoir response entails the running of an expensive reservoir simulation for each set

of well configurations. Computational demand increases when geological uncertainty needs to be considered, which is typically taken into account by using a suite of reservoir models. Additional considerations include field constraints, such as inter-well distance and maximum well length. As a result, optimization methods have been employed to efficiently find an optimal well configuration (solution).

Well placement optimization problems are characterized by the highly nonlinear relationship between the input (well locations and reservoir model) and the output (fluid production volumes). This results in nonconvex and multimodal objective function landscapes that require computationally efficient methods to traverse. Consequently, population-based algorithms have been investigated and are commonly applied to well placement optimization problems. These include the application of genetic algorithms (GA) [5–7], particle swarm optimization (PSO) [4, 8, 9] and covariance matrix adaptation evolution strategy [10].

Although these techniques produce promising solutions by exploring the space more globally, they arguably require a prohibitively large number of computationally expensive reservoir simulation calls. In some practical situations, the available computational budget may not allow the

✉ Yazan Arouri
yazan.arouri@adelaide.edu.au

Mohammad Sayyafzadeh
mohammad.sayyafzadeh@adelaide.edu.au

¹ Australian School of Petroleum and Energy Resources,
The University of Adelaide, Adelaide, 5005, SA, Australia

application of these global optimizers to well placement problems. These techniques stochastically explore the search space using a pool of solutions (known as the population). Through learning parameters, solutions which are more promising undergo selection and stochastic manipulation which results in a new population. As such, in the case of heuristic initialization, it is difficult to ensure that the initial solutions based on reservoir engineering considerations are used or exploited adequately. As a result, such solutions may be absent in future generations and replaced by completely different ones. Although these generational mechanisms can be fruitful for global search, they inevitably increase the computational demand.

In these situations, local optimizers may provide suitable solutions in a more computationally efficient manner. Given the nature of local optimization methods, they are more likely to converge to a local optimum near the starting point. Consequently, their performance is heavily affected by the quality (location in the search space) of the initial guess. In an attempt to overcome this limitation, methods such as multi-start [11] and restart techniques [12] have been introduced. As their name suggests, multi-start methods initialize the algorithm from multiple different initial guesses which will (ideally) converge to different optimums. The aim is then to select the most optimal solution from all starting points. On the other hand, restart techniques execute the algorithm for a number of iterations before restarting the algorithmic parameters and using the last solution found as the initial guess. Although these approaches may succeed in delaying the convergence to an optimum, they inevitably reduce the computational efficiency of local approaches. This becomes counter-productive for optimization problems that are computationally constrained as is commonly the case for practical scenarios. By contrast, a less burdensome approach is to utilize prior knowledge to improve the quality of the initial guess. In this undertaking, the aim is to leverage the advantages of local methods to increase the convergence speed to an improved solution. It should be noted that this approach assumes the initial guess is based on sound knowledge, as is the case for the optimization problem under consideration. However, this does not preclude the use of local optimization methods to improve lower quality initial guesses either.

Optimization of well location using local methods has received some attention in the literature. These techniques can be categorized into two groups, gradient-based and pattern-search methods. Pattern-search algorithms have been used for well placement optimization, typically as local optimizers within hybrid approaches [13–16]. These methods provide a gradient-free procedure that relies on the direct search of the solution space. This search occurs through the use of a stencil, which is typically

a collection of points obtained by perturbations of equal size in all directions. The center of the stencil moves when there is an improvement in the objective function value. If no improvement occurs, the stencil size is reduced and the procedure is repeated. However, in high-dimensional optimization problems (e.g., placement of multiple nonconventional — deviated, horizontal and/or slanted — wells) the need to search perturbations in all directions may become a computational limitation.

Gradient-based approaches involve the utilization of a gradient computed through either an adjoint system [17–20] or approximation techniques [21, 22]. The adjoint method has been applied in several pieces of literature regarding vertical well placement optimization [19, 20, 23]. These methods rely on indirect approaches, which derive the well location sensitivity through gradients based on well control. Vlemmix et al. [23] extended this method to nonconventional wells by placing pseudo-sidetracks in grid blocks adjacent to the well path at specified trajectory points. However, this method is limited to only modifying the well along these trajectory points and not its overall location in the reservoir. More recently, Volkov and Bellout [24] use a combined technique for the optimization of nonconventional wells. The method uses the adjoint formulation to find key partial derivative terms that are then approximated using finite difference methods. The practical application of methods which rely on either adjoint formulations or adjoint-based gradients can be limited as they are not readily available in all commercial simulators in well placement optimization contexts. In addition, finite difference-based methods may lose efficiency when scaled to large dimensional problems, as they require one or two function evaluations for the perturbation in each dimension.

On the other hand, approximating the gradient, using simultaneous perturbation, provides an alternative. These methods rely on only objective function values and act as black-box methods. Bangerth et al. [21] applied an integer variant of the Simultaneous Perturbation Stochastic Approximation (SPSA) algorithm to well placement problems, including placing seven vertical wells in a simple two-dimensional model. Leeuwenburgh et al. [25] applied an ensemble method (EnOpt) to optimize the areal locations (x and y-coordinates) of nine vertical wells in a two-dimensional model. Li et al. [26] applied the integer variant of SPSA to the joint optimization of well controls and well placement on three case studies, including the benchmark PUNQ-S3 model. Jesmani et al. [22] applied continuous variants of SPSA to optimize the location of a single nonconventional well in the presence of four pre-existing injection wells. Simultaneous perturbation methods can provide an approximated gradient with only two function evaluations (if central), regardless of the number of decision variables. This makes the algorithms,

which employ simultaneous perturbation gradient approximation, computationally efficient for nonconventional well placement optimization problems.

Another important consideration is the parameterization of the decision variables in the problem formulation. Various parameterizations have been utilized in the literature. The simplest kind is the placement of vertical wells, which requires only two variables for each well location. In such a case, the well locations are defined either by the continuous Cartesian coordinates (x and y) or by the cell indices (i and j) [27]. In nonconventional wells (deviated, horizontal, slanted), the completion trajectory can be defined as two (heel and toe) points in space connected by a straight line. There are two common parameterizations for nonconventional wells: Cartesian coordinates and spherical coordinates. When using spherical coordinates, the completion trajectory is defined by the heel point using Cartesian coordinates (x , y and z), and spherical coordinates to define the length of the well (L), the inclination angle (ϕ) and the azimuth angle (θ) in the horizontal plane [7, 28, 29]. In such parameterization, a well length constraint can be handled through simple bounds. On the other hand, when using Cartesian parameterization, both the heel and toe are defined by Cartesian coordinates [9, 30]. In Sayyafzadeh and Alrashdi [9], the implemented parameterization represents the heel and toe by x and y -coordinates and the z -coordinate is defined as a percentage between the bounding surfaces of the reservoir. The selection of parameterization can be important when approximating gradients using simultaneous perturbation. This is because the decision variables need to have similar sensitivities to the perturbation size in the gradient approximation. If this is not the case, the sensitivity to some parameters will be masked by others.

In practice, field development planning is a multi-disciplinary task, which includes an understanding of suitable well locations based on reservoir engineering judgement. To this end, we argue in situations under computational constraints, local methods can be leveraged to improve on these initial guesses in an efficient manner. These local methods will be able to produce improvements that are in line with considerations accounted for in the initial guess. In this study, we focus on the development of a first-order algorithm based on the adaptive moment estimation (Adam), for application to constrained well placement optimization. The algorithm is a combination of SPSA as a gradient approximation within an adaptive moment estimation framework. It is referred to as Adam-SPSA.

The adaptive moment estimation framework introduces the idea of using dimension-wise step-sizes in gradient-based algorithms. This allows the search direction to be

tailored for each dimension accordingly. In addition, this framework considers the accuracy of the approximated gradient by estimating the first and second order moments. As a result, these estimations aid the algorithm in guiding the search to more promising areas. The adaptive moment estimation (Adam) framework has found significant success in optimization problems in machine learning applications, including Google's translation system [31] and image processing [32, 33]. In these applications, the objective function is considered noisy as it is the summation of a random subset of cost (or loss) functions, from which a gradient is approximated. Although the objective function in well placement optimization using Adam-SPSA is not noisy (even if geological uncertainty is incorporated), stochasticity is still introduced by the random and simultaneous perturbations used in SPSA. This enables the extension of Adam to well placement optimization problems. More recently, olkov and Bellout [3] successfully applied the technique to well control optimization and shows its applicability to such problems.

In this study, we investigate the application of Adam-SPSA to vertical and nonconventional well placement optimization. This includes a two-dimensional visual example to compare the search directions of the adaptive moment estimation framework and the steepest descent framework. Additionally, Adam-SPSA is employed to two well placement optimization problems in the PUNQ-S3 model. The results are compared to the conventional steepest descent SPSA algorithm (SD-SPSA) and a pattern search technique, generalized pattern search (GPS).

The outline of the paper is as follows. The problem statement, including the formulation of the optimization problem, as well as the objective function are presented first. Next, the adaptive moment estimation framework is introduced. Following this, we begin the numerical results with a visualization of the gradient-based methods using a simple two-dimensional problem to reflect the improved search directions of the proposed algorithm. In addition, the numerical results for well placement optimization on two numerical case studies of increasing complexity are presented. The significance and implications of these results are then discussed, followed by concluding remarks.

2 Problem formulation

2.1 Problem statement

The optimization problem involves the minimization of a defined objective function where the well locations are the variables of interest. The well placement optimization

problem can be formulated as a general optimization problem as follows:

$$\begin{aligned} & \min_{\mathbf{x} \in \mathbb{R}^n} f(\mathbf{x}) \\ & \text{subject to:} \\ & c_i(\mathbf{x}) \leq 0 \quad i \in K \\ & c_i(\mathbf{x}) = 0 \quad i \in I, \end{aligned} \tag{1}$$

where, $f(\mathbf{x})$ is the objective function, \mathbf{x} denotes the vector of decision variables, n is the number of decision variables, I and K are sets of indices for equality and inequality constraint functions, $c_i(\mathbf{x})$, respectively. These constraint functions may include linear and nonlinear constraints, as well as bound constraints.

2.2 Objective function

Here, we consider the objective function to be the negative of the NPV for a given lifespan. The objective function is defined as

$$\begin{aligned} f(\mathbf{x}) &= -\text{NPV}(\mathbf{x}) \\ &= \sum_{i=1}^{N_t} \frac{-r_{o,i} Q_{o,i}(\mathbf{x}) + c_{wp,i} Q_{wp,i}(\mathbf{x}) + c_{wi,i} Q_{wi,i}(\mathbf{x})}{(1+b)^{t_i}}, \end{aligned} \tag{2}$$

where $r_{o,i}$, $c_{wp,i}$ and $c_{wi,i}$ are the sale price of oil and costs of water separation and injection, respectively, all of them per unit volume and defined from time t_i to t_{i+1} (there are N_t such intervals), $Q_{o,i}$, $Q_{wp,i}$ and $Q_{wi,i}$ denote the field oil production, field water production and field water injection volumes during the mentioned output interval and b is the discount rate.

3 Methodology

3.1 Gradient approximation

The gradient approximation utilized in this study is SPSA. SPSA was first introduced by Spall [34] for problems whose analytical derivatives are unavailable. SPSA only requires two function evaluations to approximate a gradient using the central-difference method. This allows SPSA to be very efficient for high dimensional problems. The stochastic gradient approximation is:

$$\begin{aligned} \hat{g}_k(\mathbf{x}_k) &= \begin{bmatrix} \frac{f(\mathbf{x}_k + c_k \Delta_k) - f(\mathbf{x}_k - c_k \Delta_k)}{2c_k \Delta_{k1}} \\ \vdots \\ \frac{f(\mathbf{x}_k + c_k \Delta_k) - f(\mathbf{x}_k - c_k \Delta_k)}{2c_k \Delta_{kp}} \end{bmatrix} \\ &= \frac{f(\mathbf{x}_k + c_k \Delta_k) - f(\mathbf{x}_k - c_k \Delta_k)}{2c_k} \\ &\quad \times \left[\Delta_{k1}^{-1}, \Delta_{k2}^{-1}, \dots, \Delta_{kp}^{-1} \right]^T, \end{aligned} \tag{3}$$

where, the mean-zero p -dimensional random perturbation vector, $\Delta_k = [\Delta_{k1}^{-1}, \Delta_{k2}^{-1} \dots \Delta_{kp}^{-1}]^T$, has a user-specified distribution and c_k is a positive scalar. The convergence theory of the SPSA algorithm can be found in Spall [34]. An important consideration for this theory is that Spall [34] recommends the use of the Bernoulli distribution. The selection of the optimal distribution perturbation vector was studied in Sadegh and Spall [35]. The gradient approximation is calculated using SPSA with the parameters following the recommendations given in Spall [36]. These are used to calculate the proposed perturbation size from the following gain sequence:

$$c_k = \frac{c}{(k+1)^\gamma}, \tag{4}$$

where, k is the iteration number, γ is a positive scalar constant and c is the initial perturbation size. Readers are referred to Spall [36] for additional implementation guidelines.

3.2 Adaptive moment estimation framework

Kingma and Ba [37] first introduced the adaptive moment estimation (Adam) framework as a gradient-based optimization algorithm which utilizes first-order information. The framework computes distinct progression steps for each decision variable (dimension) based on estimates of the first and second moments extracted from the gradient approximations [37]. In comparison, the steepest descent framework utilizes the approximated gradient as the progression step and, as such, the same step is taken for all decision variables.

The two pieces of information Adam extracts and utilizes are estimates of the first and second moments. For a random variable, X , the first moment is defined as the mean, or expected value, $E[X]$, about the origin [38]. This random variable, X , is taken to be the gradient approximation. Consequently, the estimation of the first moment is in fact an estimation of the expected value of the gradient. Following this, the definition of the second central moment of the random variable X is the variance about the mean, defined by Lefebvre [38]:

$$\text{Var}(X) = \text{Diag}(E[(X - E[X])^2]), \tag{5}$$

This can be further simplified with algebraic manipulation while assuming the variables of X are independent. Additionally, in the long-run, the mean of the gradients tends to approach zero as a local optimum is approached. This results in the following definition of the uncentred variance:

$$\text{unVar}(X) = \text{Diag}(E[X^2]), \tag{6}$$

Given that the random variable, X , is the gradient approximation, the uncentred variance is the expected value

of the element-wise gradient squared. The Adam framework then uses an exponential moving average to put greater weighting on the most recent approximated gradients when estimating the first and second moments. Details of the convergence theory for the adaptive moment estimation framework is presented in Kingma and Ba [37]. Framework 1 presents the steps of Adam.

Framework 1 - Adaptive Moment Estimation

0. Initialize iteration counter, select initial guess, x_0 , assign nonnegative constant parameters: step-size ω and exponential decay rate for moment estimates β_1 and β_2
 1. Initialize first and second moment vectors, m_0 and v_0 , which are $n \times 1$ column vectors
 2. Approximate stochastic gradient
 3. Update the first moment vector using update rule:
 - (a) $m_k = \beta_1 \times m_{k-1} + (1 - \beta_1) \times \hat{g}_k$, where m_k, m_{k-1} and \hat{g}_k are $n \times 1$ column vectors
 - (b) $\hat{m}_k = \frac{m_k}{1 - \beta_1^k}$, where \hat{m}_k is a $n \times 1$ column vector
 4. Update the second raw moment vector:
 - (a) $v_k = \beta_2 \times v_{k-1} + (1 - \beta_2) \times (\hat{g}_k \odot \hat{g}_k)$, where \odot is the element-wise multiplication and v_k and v_{k-1} are $n \times 1$ column vectors
 - (b) $\hat{v}_k = \frac{v_k}{1 - \beta_2^k}$, where \hat{v}_k is a $n \times 1$ column vector
 5. Update the iterate:
 - (a) $x_{k+1} = x_k - \omega \times \frac{\hat{m}_k}{\sqrt{\hat{v}_k + \epsilon}}$
-

It should be noted that all operations are done in an element-wise manner. The \odot represents the element-wise multiplication of two vectors. Also, the operation on the hyper-parameters (β_1^k, β_2^k) are denoted as β_1 and β_2 raised to the power k (iteration number). Typical values for β_1, β_2 are 0.9 and 0.999, respectively [37]. A small positive value, $\epsilon = 10^{-8}$, is used to avoid the division by zero [37].

In machine learning applications, the objective function is typically in the form of a loss function representing a sum of differences over a set of observations. For example, a common loss function is the summation of the squared differences, $f(x) = \sum_{i=1}^N (y_i - \hat{y}_i)^2$, where N is the number of observations and y_i and \hat{y}_i are the true value and predicted value for observation i , respectively. In many occasions, instead of calculating the gradient as $\nabla_k f(x) = \sum_{i=1}^N \nabla_k (y_i - \hat{y}_i)^2$, a random set of observations is used in each iteration instead of all the observations ($\nabla_k f(x) \approx \sum_{i=1}^n \nabla_k (y_i - \hat{y}_i)^2$), where n is smaller than N . This gives an approximate gradient in each iteration. If these gradient approximations are averaged, the result is in the same direction as the true gradient.

In contrast, the objective function in well placement problems is not of similar form to a loss function and the gradient approximation can be computationally expensive. That is, the objective function, $f(x) = -NPV(x)$, requires reservoir simulations to obtain a objective value. To reduce the computational costs of gradient approximations, we use a simultaneous perturbation approach. It is worth mentioning that, in both cases (subset of objective functions or simultaneous perturbations), the gradient approximation accuracy increases if averaged over a number of approximations. That is, let $g(x) = \nabla_k f(x)$ then the averaged gradient approximation is given by $\hat{g}(x) = \sum_{j=1}^s g_j(x)$, where s is the number of gradient approximations. Additionally, the theory of simultaneous perturbation suggests the approximated gradient is an unbiased estimator of the true gradient within a bias bound [34]. Adam has been shown to work reasonably well in such stochastic optimization problems, and because of the similarities, we use this framework with a simultaneous perturbation to address computational intensity of well placement problems.

3.3 Parameter selection

In this section, we present the rationale for the parameter values selected for perturbation size and step-size for each algorithm. The selection of appropriate parameter values for perturbation size and step-size are pivotal in the convergence of local algorithms. Previous work has investigated this topic for well control optimization problems [3].

3.3.1 Perturbation size

The perturbation size determines the amount of perturbation taken in a direction when calculating the gradient. In this work, we use a central difference to approximate the gradient using SPSA. As such, the perturbation is done in the forward and backward directions from the current solution. The gradient aims to capture the sensitivity of the objective function to changes in the solution. For well placement problems, such as those under consideration, this is guided by the grid block size. Considering this, the numerical simulation may not properly capture the sensitivity if the perturbation is too small. Similarly, a perturbation too large will not be representative of the local landscape. Due to spatially varying property values (e.g., saturations and permeabilities), there is an inherent discontinuity in the landscape. The selection of an appropriate perturbation becomes important as it is one of the remedies for this issue, to some extent.

Consequently, an initial perturbation size (c) value of 0.05 was found to be suitable for a perturbation of

approximately one grid block. Adam-SPSA and SD-SPSA both employ the decaying sequence given by Eq. 4 to update the perturbation size at each iteration. A value of 0.101 was used for γ as recommended by Spall [36]. It is important to note that although Eq. 4 represents a decaying sequence the sensitivities are still captured by the gradient.

Firstly, this work is focused on optimization under practical computational constraints. As such, the decay of the perturbation size is limited for the computational budget considered. In addition, in this work we employ a continuous parameterization for well location for vertical and nonconventional wells. Lastly, the wells are not necessarily at the centre of the grid blocks as the search progresses.

In turn, this rationale was used to guide the selection of an initial stencil size for optimization using GPS. The stencil size in GPS represents the perturbation taken in each direction (decision variable) that forms the search points of the stencil at the current solution. For this reason, a stencil size that represents a perturbation of approximately one grid block is suitable. Consequently, the selected initial stencil size for GPS optimization runs is similar to the perturbation size value for Adam-SPSA and SD-SPSA.

3.3.2 Step-size

The step-size determines the progression of an algorithm from the current solution to the next proposed solution in the search direction. The effective step-size can be thought of as the step-size multiplied by the search direction. In other words, this is the difference between the current solution and the proposed solution. As such, a suitable effective step-size is one where the search does not swing from one bound to another as this may result in overstepping local minimums. On the other hand, an effective step-size that is too small may cause a prohibitive slowdown in convergence. Accordingly, for this work an effective step-size which represented approximately one grid block was used. To do this, the first two iterations of Adam-SPSA were used to find a suitable value for ω (step-size) that gave an effective step-size of one grid block. In Adam-SPSA, the selected ω value is constant throughout the optimization run (i.e. the step-size does not change as the search progresses). In addition, previous studies have shown that Adam-SPSA does not require a backtracking line-search given its adaptive search direction [3].

A similar routine was undertaken to find a similar effective step-size for SD-SPSA. SD-SPSA updates the step-size based on Eq. 7, where a is the initial step-size, A is a stability constant, and α is a positive scalar constant. From Spall [36] we use the recommended values of 0.602 and $0.1 \times \text{maximum number of iterations}$ for α and A , respectively. Additionally, for SD-SPSA a backtracking

line-search is implemented to help guide the search to an improved objective function value. In this work, a cut-back value of 0.5 and a maximum of 5 cut-backs are used, after which the gradient is re-calculated and the line-search is repeated.

$$a_k = \frac{a}{(A + k + 1)^\alpha}, \quad (7)$$

3.3.3 Constraint handling

The consideration of physical field constraints is an important aspect for practical application of optimization in well location problems. In this work, three nonlinear constraints are considered to ensure the proposed solutions do not violate engineering principles. In addition, the decision variables are normalized between 0 (lower bound) and 1 (upper bound) for which simple bound constraints are applied. The first constraint considers a minimum inter-well distance between any pair of wells. When placing three-dimensional nonconventional wells, this constraint considers the minimum distance between any two points on the line segments representing the two wells. The second constraint is a polygon reservoir bound. In this constraint, the reservoir bound is approximated using a polygon shape to ensure any proposed well (both heel and toe for nonconventional wells) lies within this polygon. The third constraint considered relates to a maximum well length for nonconventional wells.

The violation of any of these nonlinear constraints is treated through the projection of an infeasible point onto the feasible domain. This projection is formulated into an optimization problem where a distance metric, the Euclidean distance, with respect to the infeasible point is minimized. That is, the distance between the original violating set of well location/s and a set of proposed well location/s is minimized. This optimization problem is subject to the same constraints as in the original problem. In this work, the constraint handling is implemented through MATLAB's *fmincon* function, which is a nonlinear programming solver. The solver is set to use a sequential quadratic programming (SQP) algorithm to solve this optimization problem. Details can be found on the MathWorks reference manual [39].

3.3.4 Constraint handling in gradient approximation

A sensitivity analysis was undertaken to gain insights into the effect that different types of constraints in gradient approximation had on the algorithm's performance. All parameter values were kept constant for each method. The difference between each method is the constraints considered when approximating the gradient using SPSA. This means once the forward ($x_k + c_k \Delta_k$) and backward

$(x_k - c_k \Delta_k)$ perturbations are proposed they are subjected to the assigned constraints.

Three different combinations were tested. The first considered simple bound constraints to ensure the proposed forward and backward perturbations are within the lower and upper bounds. The second implemented a refined bound constraint which included simple bound constraints and the reservoir polygon constraint. That means if a forward or backward perturbation proposed well locations outside the defined reservoir polygon, they are projected back into the reservoir. This was to ensure the number of wells was consistent throughout the gradient approximation. The third set-up considered all three nonlinear constraints (mentioned in Section 3.3.3) as well as the bound constraints.

The results showed the worst performing combination was the simple bound set-up which considered only bound constraints for the forward and backward perturbation. This could indicate that although the perturbations are within the bounds, they are not corrected enough to result in a useful gradient. On the other hand, including all the nonlinear constraints could be over-correcting the perturbations. This may result in a gradient that is not representative of the local landscape causing it to perform worse than the refined bounds set-up. The best performing combination is the refined bound case containing bound constraints and the reservoir polygon constraint. This insight shows that the constraint handling considered for the perturbations has an effect on the quality of the gradient approximation. It is worth mentioning that only one simultaneous perturbation is used for gradient approximation.

4 Case studies

This section begins with a simple two-dimensional example to illustrate the differences in search direction between Adam-SPSA and SD-SPSA. Next, the results for a case study investigating the placement of four infill vertical

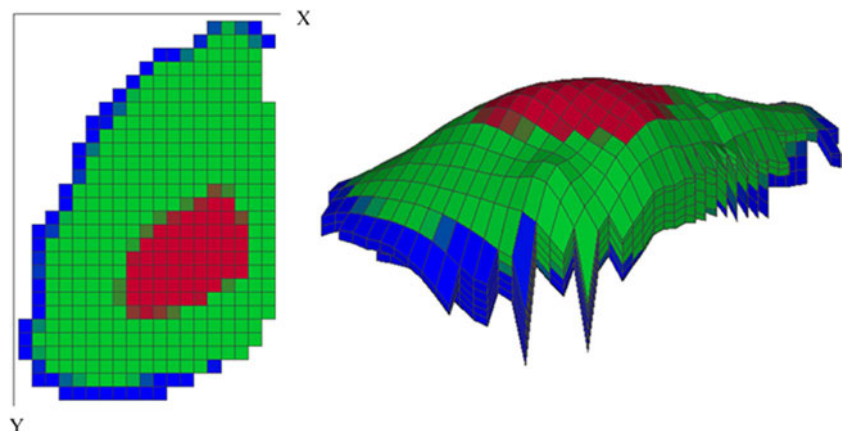
production wells are presented. The third case study considers the placement of 20 nonconventional wells considering three physical field constraints. All case studies use the three-dimensional PUNQ-S3 benchmark model, shown in Fig. 1 [40]. The model is an oil-saturated heterogeneous reservoir with a small gas-cap and strong aquifer support with bottom-drive and side support on two sides. The other two sides are defined as no-flow boundaries. The model is discretized into 19 by 28 by 5 grid blocks, in which 1,761 are active. The areal extent of the model is $17 \times 10^6 m^2$, with varying thicknesses between 20 and 30 meters. The reservoir has a lifespan of 10 years.

4.1 Case study 1 - two-dimensional visual example

In this case study, we visually compare the search directions from Adam-SPSA and SD-SPSA. The optimization problem is the placement of one vertical production well, resulting in 2 decision variables (x- and y- coordinates). Figure 2 shows the search steps undertaken by Adam-SPSA and SD-SPSA for three different initial starting points (represented by the cross). In Fig. 2a, both Adam-SPSA and SD-SPSA are able to reach the closest local minimum to the initial guess. However, Adam-SPSA only requires 4 iterations to do so while SD-SPSA requires 6. Figure 2b shows a different result where only Adam-SPSA is able to reach the closest local minimum after 6 iterations, whilst SD-SPSA converges after 5 iterations to a lower quality solution. It must be noted that SD-SPSA was allowed to continue for 3 additional iterations without any improvement. Similarly, Fig. 2c shows that Adam-SPSA was able to reach the closest local minimum in 3 iterations, whilst SD-SPSA did not converge to the same solution after 3 iterations.

As shown in Fig. 2 the search directions of SD-SPSA are always 45 degrees from the previous solution. This is attributed to the use of the Bernoulli distribution in the gradient (when using only one gradient approximation), which is directly used as the search direction. This problem

Fig. 1 Top view of Layer 1 and three-dimensional representation of the PUNQ-S3 benchmark model showing the ternary saturations. Red is gas, green is oil and blue is water



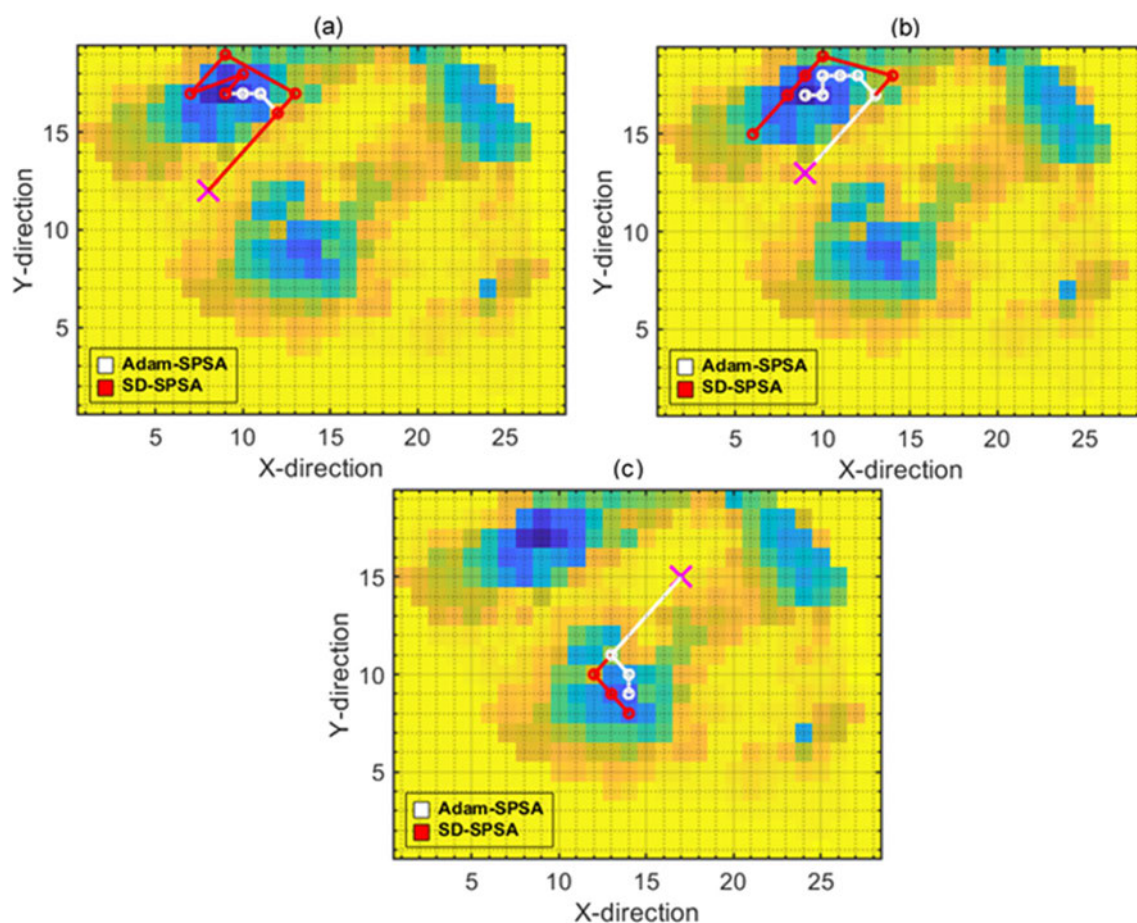


Fig. 2 Visual representations of search steps of optimization of one vertical well using three different initial starting points (pink cross). The numbers inside the box indicates the number of iterations needed

is more severe when the number of dimensions is greater and/or the sensitivity to variables is not similar. Unlike SD-SPSA, Adam-SPSA uses a dimension-wise adjusted search step which gives the algorithm more control and flexibility over movement in both x and y directions.

4.2 Case study 2 - four vertical infill production wells

In this case study, two producers are considered pre-existing. These wells are PROD5 and PROD6 with x and y-coordinates of (2970,1890) and (1890, 4230), respectively, shown in Fig. 3. The objective is to minimize the function given by Eq. 2 by placing four vertical infill producers, resulting in a total of 8 decision variables. The wells are controlled by bottom-hole pressure (BHP) with a pressure of 2900 psi (200 bars) and a maximum liquid rate of 5660 STB/day (900 sm^3 /day). The economic parameters are given in Table 1. Since the number of wells are fixed (i.e., is not a decision variable) and all wells are vertical, there is no impact of drilling and completion costs on NPV calculation. As such, the drilling and completion costs are

to reach final value for each algorithm. Yellow is high objective function value and blue is low objective function value

not considered in this case study, and the associated values in Eq. 2 are equated to zero. The optimization variables were normalized and bounded between 0 and 1. In this case study, a minimum inter-well distance of 300 meters is used. A box constraint was used to represent the reservoir bounds by placing upper and lower bounds on the x and y-coordinates.

As previously discussed, the step-sizes were used to ensure a similar effective step was taken in each algorithm. Adam-SPSA had parameter values of 0.07 and 0.05 for ω and c , respectively. SD-SPSA had parameter values of 0.2 and 0.05 for a and c , respectively. To reduce the effect of the stochastic nature of the SPSA gradient approximation, both Adam-SPSA and SD-SPSA results were averaged over 10 optimization runs from the same reservoir engineering initial guess. Since GPS is a deterministic method, it was only run once. That is, for the same initial guess, GPS will result in the same solution. GPS was implemented using MATLAB's pattern search optimization tool [39]. The polling method used was the *Positive basis 2N* with an initial and maximum mesh size of 0.05. An incomplete polling was employed where the first search direction at

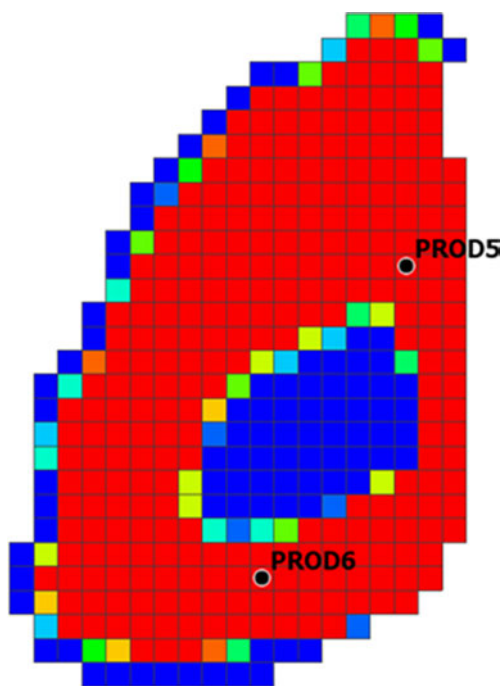


Fig. 3 Initial oil saturation of Layer 1 in two-dimensions with two pre-existing production wells (PROD5 and PROD6). Red represents high oil saturation and blue represented low oil saturations

each iteration is the direction which gave the best solution in the previous iteration. The stopping criterion was set to a maximum number of function evaluations of 100 in this case study. For all optimization runs, the initial guess was taken to be an example of a possible solution based on reservoir engineering judgement, represented by the initial well locations from the PUNQ-S3 benchmark model. These wells are located around the gas/oil contact (GOC) to ensure that the water breakthrough is delayed whilst maximizing oil production.

Figure 4 shows the convergence plot for the results of the well location optimization for case study 2. The results show that SD-SPSA, on average, converged to a final optimum NPV of 1,359 MM USD after 103 function evaluations. In comparison, Adam-SPSA, on average, required only 30 function evaluations to achieve a similar NPV value (1,361 MM USD). This represents up to a 71% decrease in the required number of functional evaluations. Another important consideration is the standard deviation of the results from Adam-SPSA compared to SD-SPSA. The

Table 1 Economic parameters for the two example problems

Parameter	Case study 2	Case study 3
Oil price (r_o)	\$80 USD/bbl	\$40 USD/bbl
Water handling costs (c_{wp})	\$13 USD/bbl	\$5 USD/bbl
Discount rate (b)	0%	0%

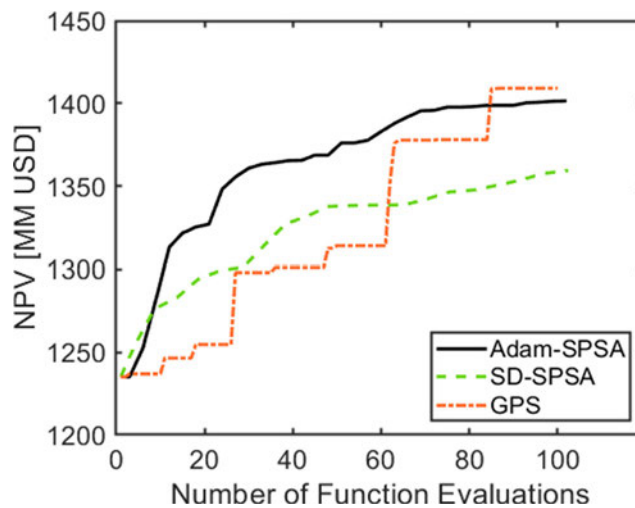


Fig. 4 Convergence plot for the optimization of four vertical infill wells comparing Adam-SPSA (solid line), SD-SPSA (dashed line), and GPS (dash-dotted line). The results of Adam-SPSA and SD-SPSA represent an average of 10 runs

standard deviation of the 10 runs for both these algorithms gives an indication of the effect that the stochastic nature of the gradient approximation has. Although an element of stochasticity can be advantageous, it can become a hindrance if an algorithm is susceptible to it. For case study 2, Adam-SPSA had a standard deviation in the final optimum value of 41.69 MM USD across the 10 runs. In comparison, SD-SPSA had a standard deviation of 51.22 MM USD across the 10 runs. This is up to almost a 19% difference in standard deviation. That is, the optimization runs of SD-SPSA showed a 19% more spread in the final optimum values found when compared to Adam-SPSA.

The average results of Adam-SPSA and the result of GPS are competitive with regards to the final optimum value. The GPS optimization run results in a final optimum NPV value of 1,409 MM USD after 100 function evaluations. On the other hand, Adam-SPSA, on average, results in a final optimum value of 1,401 MM USD. The convergence plot of GPS exhibits a characteristic pattern of a local optimizer with improvements occurring after searching the local landscape. Given the stencil-based search of GPS, it is not uncommon to see large increases in NPV after one function evaluation. An example of this is shown in Fig. 4 where a significant jump in NPV occurs at 61 function evaluations.

Although the one GPS run outperforms the average runs of Adam-SPSA in terms of final NPV, it should be mentioned that three of the 10 runs of Adam-SPSA outperform GPS. An example is shown in Fig. 5, which shows the best run of Adam-SPSA against the GPS run. These results show that Adam-SPSA outperforms GPS in terms of convergence speed and final optimum value. For

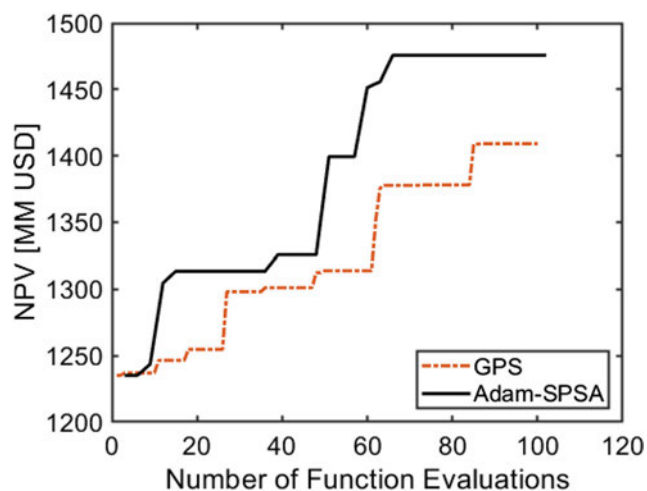


Fig. 5 Convergence plot for the best run (out of 10) of Adam-SPSA (solid line) and GPS (dash-dotted line) run for the optimization of four vertical infill wells

example, GPS converges to an optimum value of 1,409 MM USD after 85 function evaluations, whilst Adam-SPSA converges to similar value (1,399 MM USD) after only 51 function evaluations. This translates to a 40% improvement in number of function evaluations to achieve a similar NPV value. This Adam-SPSA optimization run converges to a final optimum value of 1,476 MM USD after 66 function evaluations. On the other hand, after 66 function evaluations, GPS reaches an NPV value of 1,378 MM USD, which is a 7.11% reduction compared to Adam-SPSA. The convergence plots shown in Fig. 5 are characteristic of local optimizers. The staggered pattern represents the algorithms traversing the local landscape in search of an improvement in objective function value.

Figure 6 shows the final oil saturation maps for the best optimized solutions of all three algorithms. The major difference is seen when comparing the locations of the wells relative to the gas cap, where the wells were initially located (shown in top left map). Given this work investigates the use of local optimization algorithms for well placement problems, it is expected that the final solutions will be relatively close (with respect to well locations) to the initial guess. The aim is to improve on an initial guess that is based on reservoir engineering judgement to maximize the objective function. This is shown by the well locations from the solutions of the three algorithms. The wells are slightly moved away from the gas cap to produce more oil from the northern and western flanks of the reservoir. However, this needs to be balanced as a move too far north or westerly would result in a large water influx from the strong aquifer on either side. From the final oil saturations shown, the Adam-SPSA solution produces a solution that best balances

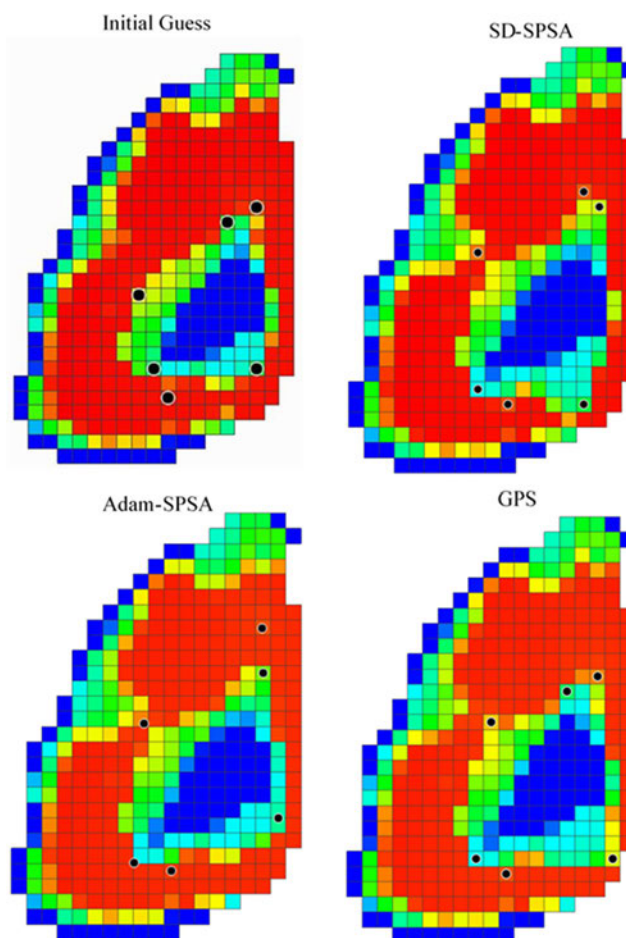


Fig. 6 Oil saturation maps of Layer 1 at the end of the production time for the initial guess (top left) and the three best solutions from SD-SPSA (top right), Adam-SPSA (bottom left) and GPS (bottom right). Red represents high oil saturation and blue represented low oil saturations. Wells are represented by solid black circles

this as shown by the drainage of oil saturation in the western side of the reservoir model.

4.3 Case study 3 - 20 nonconventional wells

The third case study investigates the placement of 20 non-conventional wells in the PUNQ-S3 model. Consequently, this problem has a total of 120 decision variables. The optimization variables were normalized and bounded between 0 and 1. In this paper, we use a parameterization similar to Sayyafzadeh and Alrashdi [9]. Each nonconventional well is defined with 6 variables. The heel and toe are defined using x and y Cartesian coordinates, while the z -coordinate is defined as a percentage between the top and bottom layer. For a proposed set of x - and y -coordinates, the corresponding z -coordinates of the top and bottom layers are found using an interpolation surface. Then the proposed z -coordinate is calculated using the proposed percentage

(i.e. the decision variable) between these two z-coordinates. This helps ensure that the full perforation length of the well intersects the reservoir model in the z-direction. The wells are defined as fully perforated straight lines between the heel and toe. Lastly, this parameterization allows the nonconventional wells to be defined as vertical, slanted or horizontal.

The objective is to place 20 nonconventional wells. As in case study 2, the wells are controlled by BHP with a pressure of 2900 psi (200 bars) and a maximum liquid rate of 5660 STB/day (900 sm^3/day). The economic parameters used in this case study are given in Table 1. As previously mentioned, three nonlinear constraints are considered in this case study. This includes a maximum well length of 500 meters, an inter-well constraint (in three-dimensions) of 200 meters and a reservoir polygon bound (Fig. 7) constraint.

The GPS employed is the same as that employed in case study 2, with an initial stencil size of 0.05. Adam-SPSA had parameter values of 0.1 and 0.05 for ω and c , respectively. SD-SPSA had parameter values of 0.25 and 0.05 for a and c , respectively. In this case study, an initial guess was used based on possible reservoir engineering considerations. The wells were placed in a manner that would take advantage of the added contact with the reservoir when using nonconventional wells. However, the effect of

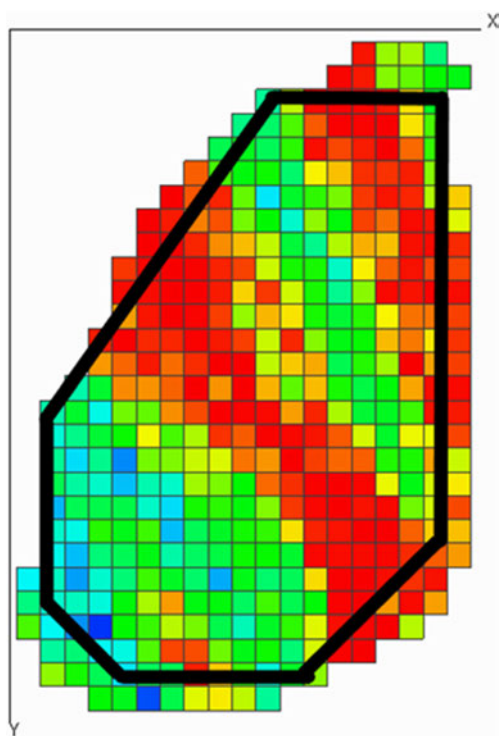


Fig. 7 X-directional permeability of Layer 1 showing the reservoir polygon boundary (solid black lines) used to define the piecewise linear polynomials for boundary constraint. Red represents a high permeability value and green represents a low permeability

the strong aquifer needed to be considered to ensure the water breakthrough did not force wells to be shut-in. The stopping criterion was set to a maximum number of function evaluation of 200.

Figure 8 compares the performance of Adam-SPSA to GPS and SD-SPSA for case study 3. Similar to case study 2, Adam-SPSA and SD-SPSA were run with the same initial guess using 10 different seeds. On average, SD-SPSA required 203 function evaluations to converge to a final optimum NPV value of 719.9 MM USD. In comparison, Adam-SPSA required 87 function evaluations to reach a similar NPV value (721.7 MM USD). This translates to a convergence speed-up of up to 57% to reach the SD-SPSA final optimum value.

An interesting insight shown in Fig. 8 is the behaviour of Adam-SPSA and SD-SPSA in early iterations. Between 0 and 50 function evaluations SD-SPSA, on average, seems to be performing slightly better than Adam-SPSA. However, after this initial period, the convergence speed of Adam-SPSA increases, while SD-SPSA slows significantly, resulting in a noticeable difference in the average final optimums. This initial difference can be attributed to the nature in which the search directions are calculated by Adam-SPSA, where a running exponential average is used to estimate the first and second moments. As such, in order to improve the search direction, Adam-SPSA requires first-order information from a number of iterations. Once this is achieved, Adam-SPSA is able to perform significantly better based on this extracted information. This is a noticeable difference compared to case study 2 where only 8 dimensions were considered where Adam-SPSA was able to update the dimension-wise search directions more efficiently.

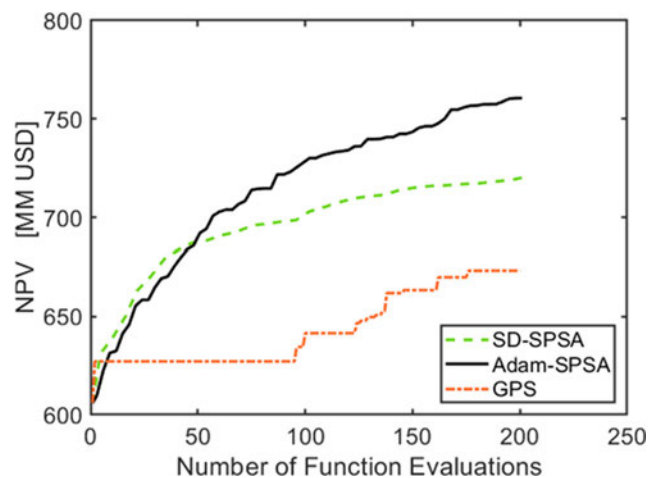
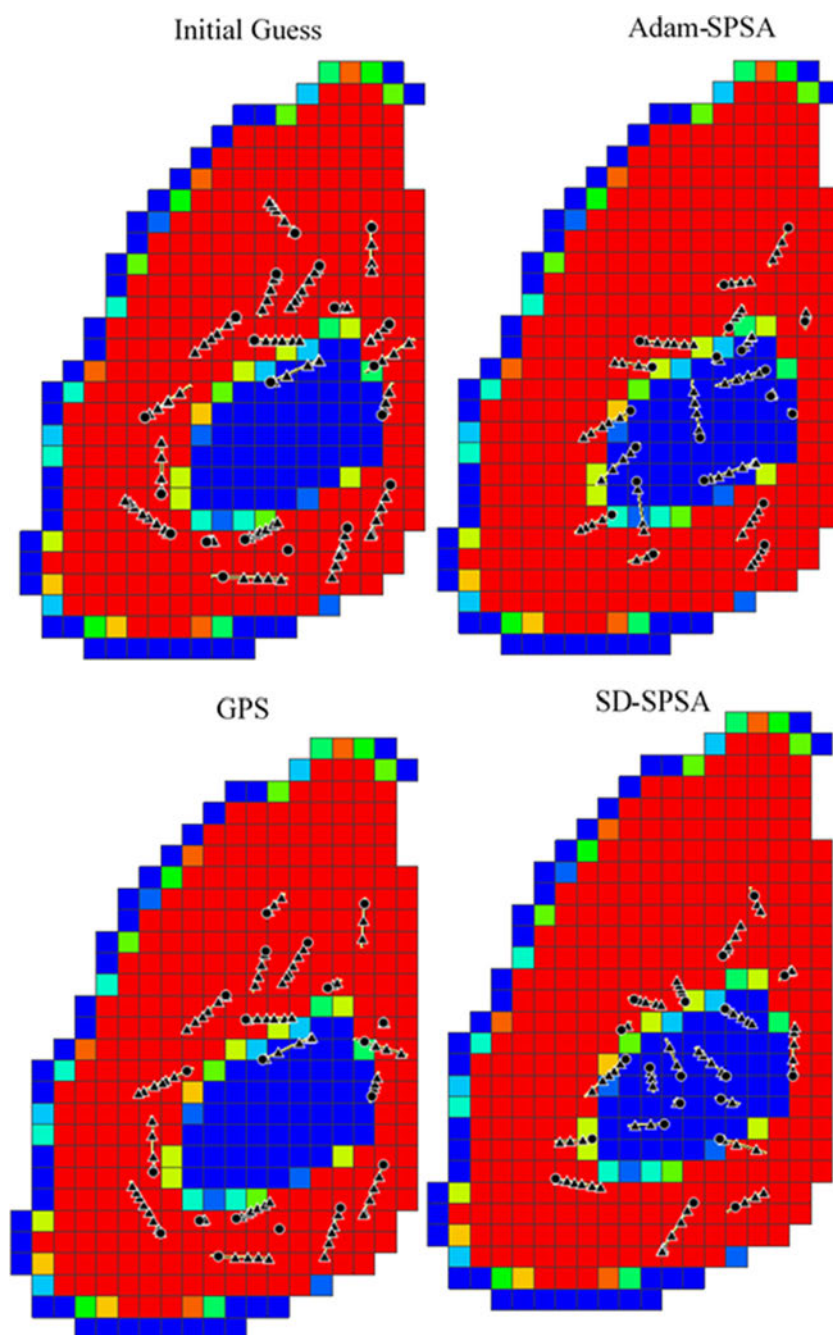


Fig. 8 Convergence plot showing the results of Adam-SPA (solid line), SD-SPSA (dashed-dotted line) and GPS (dashed line) for the placement of 20 nonconventional wells

The large difference in performance of the best run of SD-SPSA and its average over 10 runs is the stochastic nature of the gradient calculation. As such, the standard deviation can give insights into the spread of optimization results. For case study 3, Adam-SPSA has a standard deviation in the final optimum value of 15.90 MM USD. On the other hand, SD-SPSA has a standard deviation over 10 runs of 26.38 MM USD, which represents up to 40% more spread in final optimum values. From a practical standpoint, Adam-SPSA results in a more consistent solution with respect to the final optimum NPV.

The comparison of the GPS and Adam-SPSA results is a reflection of the “curse of dimensionality” that pattern search methods are susceptible to. For a 120-dimensional problem, as in case study 3, it is computationally inefficient to traverse the local landscape using a stencil-based approach. Even when employing an incomplete polling approach, this search technique may result in prolonged periods of no improvement in function value, as shown in Fig. 8. The outperformance of GPS by Adam-SPSA (and SD-SPSA) gives insight into the advantage that a stochastic element (i.e., in gradient approximation) can

Fig. 9 The well locations for initial guess (top left) and initial oil saturation maps with the optimal well locations found by Adam-SPSA (top right), SD-SPSA (bottom right), and GPS (bottom left) in two-dimensions. Circles represent well head location and triangles represent a connection. Red represents high oil saturation and blue represents low oil saturation



have, in addition to the simultaneous perturbation. Although these local methods are still susceptible to the same dimensionality curse, having a stochastic element gives these methods an opportunity to traverse the landscape more efficiently. This results in a more computationally efficient search and hence final optimum value. It is worth mentioning that there is potential for these advantages to extend to larger models containing finer grid cells. However, in such cases careful consideration would need to be given when selecting appropriate step-sizes and perturbation sizes. As a step-size and perturbation size too small may have an affect on the convergence of the algorithm.

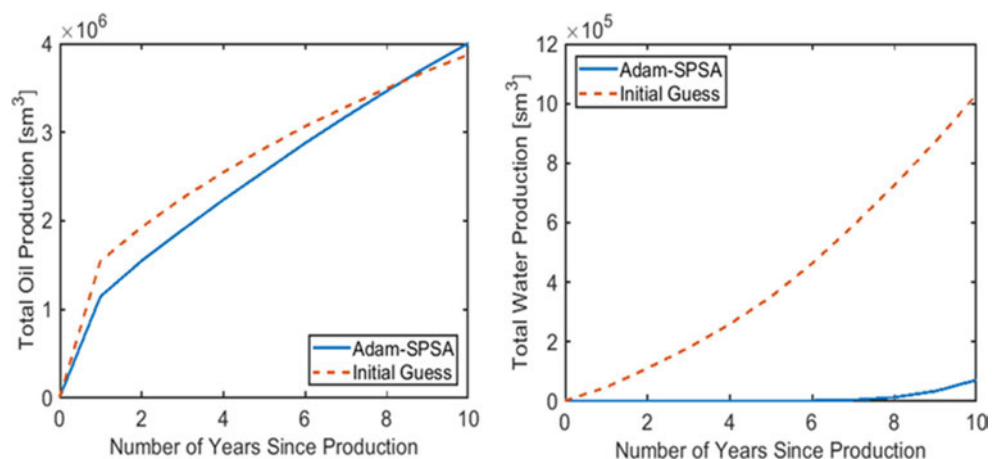
The well locations of the initial guess as well as selected optimization results for the three algorithms are presented in Fig. 9. The first observation is the similarity of the well locations from the initial guess to each of the algorithm results. This is what is expected with the use of local methods, which exploit the initial guess and progress to the closest minimum. The well locations exhibited in Fig. 9 show that placing wells closer to the gas cap assists in lowering the water production due to aquifer encroachment. However, placing too many wells in (or too close to) the gas cap, as in the SD-SPSA solution, may undermine the oil production total and as such reduce the objective function (NPV) value. Overall, the solution found by Adam-SPSA is able to balance these two aspects of water production and oil production to improve NPV over the initial guess. This gives an indication that local methods, both gradient-based and pattern-search (although to a lesser degree) type algorithms can be used to improve on initial guesses for well location optimization. Furthermore, although the well locations are similar to the initial guess, it reflects the ability for local methods to handle nonconventional well trajectory placement. In computationally constrained practical scenarios, these methods may provide suitable solutions within the budget.

Another important aspect of practical application for nonconventional wells is the ability for algorithms to

navigate the local landscape in the presence of physical field constraints. As mentioned earlier, three nonlinear constraints were considered: a maximum well length, a minimum three-dimensional inter-well distance and a reservoir polygon bound. The local methods were still able to improve on the initial guess whilst operating with the nonlinear constraint handling employed. The constraint handling technique is successful in ensuring the proposed solutions are not violating the physical constraints.

Additional insight can be obtained when reflecting on the oil and water production totals, presented in Fig. 10, for the Adam-SPSA solution presented in Fig. 9. These totals give an indication of the reasons for the differences in the objective function value, NPV, obtained from Adam-SPSA compared to the initial guess. The placement of the wells plays a pivotal role in the volume of fluids produced from a field. As previously mentioned, this solution places more wells closer to the gas cap, with a number of wells directly on top of the dome structure. By the same token, this means that the wells in this solution are further away from the strong aquifers that are present on the outer boundaries. After 10 years of production, the Adam-SPSA solution and the initial guess have similar oil production totals of $4.00 \times 10^6 \text{sm}^3$ and $3.87 \times 10^6 \text{sm}^3$, respectively. However, there are more substantial differences between the water production totals. Specifically, the Adam-SPSA solution results in a water production total of $7.02 \times 10^4 \text{sm}^3$ compared to $1.03 \times 10^6 \text{sm}^3$ produced in the initial guess, which is more than a magnitude difference. Given there are no associated costs with gas handling, the key driver for the difference in NPV between the Adam-SPSA solution and the initial guess is the significant reduction in water production. This gives additional insight to possible field development considerations with regards to water handling costs. Although not considered in this study, further analysis of these results in practical applications may lead to the consideration of the gas production. This could be limited through additional constraints or the introduction

Fig. 10 Oil (left) and water (right) production totals for the best Adam-SPSA (solid lines) solution compared to the initial guess (dashed line)



of gas handling costs into the objective function. In such a scenario, it is suitable to expect the solutions from the optimization will look different.

5 Discussion

In practical scenarios under computational constraint the use of optimization algorithms may be limited by a number of expensive reservoir simulation calls. As such, the employment of local optimization algorithms to these problems, such as well location optimization, may provide suitable improvements to proposed solutions. Often in practice an in-depth understanding of the reservoir dynamics is available. This may culminate itself through multiple possible development plans or potential infill well locations. This provides an ideal starting point for local optimization techniques to exploit and fine-tune these development plans to improve returns and highlight potential bypassed opportunities. Previous applications of local optimization algorithms have shown promise for vertical well placement problems. A limited number of studies have also shown applications for local optimization algorithms for nonconventional well placement.

In this study, we compare a derivative-free local algorithm, GPS, and two gradient-based algorithms, Adam-SPSA and SD-SPSA (with some form of stochasticity) for well placement optimization under a limited computational budget. For problems of relatively low dimensions, as in case study 2 (8 dimensions), the performance of the algorithms was competitive as the local landscape could be exploited efficiently. However, as the dimension increases to more realistic problems, as in case study 3 (120 dimensions), the simultaneous perturbation stochastic nature of SPSA offers some advantages. The element of randomness introduced by SPSA, especially when using one approximation, increases the probability of the search covering the local landscape more efficiently.

In addition, by simultaneously perturbing all the decision variables at once the progression may occur in all dimensions. On the other hand, the deterministic nature of pattern search methods, such as GPS, makes this type of progression less likely. In addition, pattern search methods employ perturbations to each dimension individually, which is computationally inefficient for problems with a large number of decision variables. The decrease in performance of GPS from case study 2 to case study 3 reflects this inefficiency and susceptibility to the dimensionality curse. This further becomes an issue under computational constraint where an incomplete polling is more favourable, which may result in more promising directions being overlooked.

Although Adam-SPSA and SD-SPSA both employ SPSA as the gradient approximation, the first-order information extracted and utilized differs. The steepest descent framework employs the gradient as the search direction to progress the search in each dimension. However, the step-size taken is identical across all dimensions. A backtracking line-search was employed for SD-SPSA to improve the search progression by adjusting the step-size. On the other hand, Adam-SPSA, utilizes an adaptive framework that allows dimension-wise steps. This is done by estimating the first and second order moments of the gradients, which represents the accuracy of the gradient in each direction. The search direction in Adam-SPSA is the ratio of the first moment (mean) to the second moment (variance). This can be thought of as a signal-to-noise ratio (SNR). A smaller SNR indicates that there is substantial uncertainty as to whether the estimated first moment corresponds to the true gradient. As a result, the effective step in such a direction should be small. On the other hand, a large SNR indicates that the estimated first moment gives a better approximation of the true gradient as the estimated second moment would be relatively low. Consequently, this allows Adam-SPSA to adaptively change the search direction for each decision variable to improve the search. This is shown through the results for the two case studies involving both the placement of vertical wells (case study 2) and nonconventional wells (case study 3) where Adam-SPSA outperformed SD-SPSA in both convergence speed and final optimum value.

Another important consideration for gradient approximations which use simultaneous perturbation is the parameterization. A gradient approximation attempts to find the sensitivity of the objective function to each decision variable. In simultaneous perturbation methods, this is done for all decision variables often using the same perturbation size (namely, c in SPSA). However, if the decision variables do not have the same sensitivity, this will result in the less sensitive decision variables being masked by more sensitive decision variables. As a result, the gradient approximation will not be representative of the landscape.

In our nonconventional well placement studies, we trialed both spherical coordinates and Cartesian coordinates to define the well trajectory. Given the different types of variables in spherical parameterization, the sensitivity of the objective function to each will be varied. Our results, which are not shown, found that when using spherical parameterization for gradient approximations that use simultaneous perturbation were susceptible to this and performed very poorly. To further investigate this difference in performance, we calculated the gradients using finite difference for both parameterizations for 15 nonconventional wells. Figure 11 presents the histogram showing the (finite-difference) gradient values across all

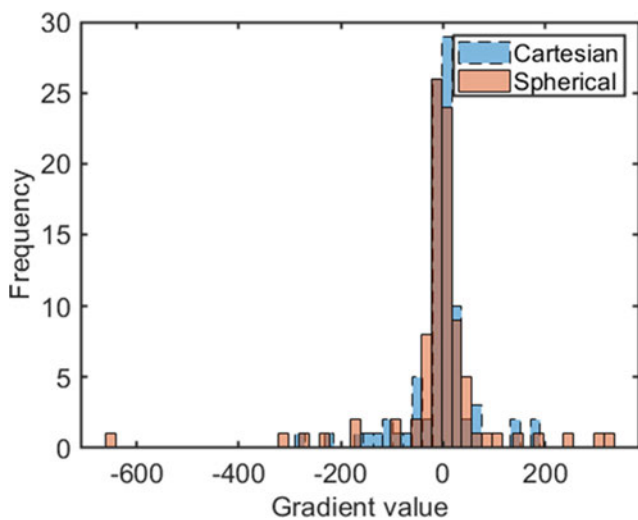


Fig. 11 Histogram of gradient values calculated using finite difference for all decision variables using a perturbation size (c) of 0.1 comparing Cartesian and spherical parameterization

90 decision variables. The mean absolute deviation for the Cartesian parameterization histogram is 36.44, whilst it is 51.83 for the spherical parameterization histogram. This indicates that the gradient values when using spherical coordinates exhibit more spread when using the same perturbation size. This additional sensitivity can be attributed to the dependence of the location of the toe to the location of the heel in spherical parameterization. When a coordinate of the heel is perturbed, the physical location of the toe will also move even though the values for the decision variables related to the toe (ϕ , θ , L) are unchanged. This gives insight into the effect that different types of parameterization can have on gradient approximations.

Although the performance of simultaneous perturbation gradients improve when utilizing a Cartesian parameterization, an additional nonlinear constraint for well length needed to be considered. In addition, it is more difficult to put constraints on the dogleg severity when using this parameterization. This could be the reason why spherical parameterization is popular in the literature as population-based methods do not require a gradient approximation.

The practicality of optimized solutions is dependent on their feasibility when physical constraints are present. This further complicates the optimization problem, especially for nonconventional wells where we considered three practical field constraints. These included a minimum inter-well distance, a maximum well length and a reservoir polygon bound. These constraints were useful in obtaining final optimum solutions that were plausible and in line with general engineering knowledge (e.g. no intersecting wells). Similar constraint handling techniques applied in this work could be extended to deal with practical constraints surrounding a minimum distance between a well and a fault.

In this case, the fault would need to be represented by a plane in 3-dimensional space.

Furthermore, the role these constraints play when approximating gradients and their effects on algorithm performance was investigated. Although not extensive, the preliminary results indicated that there is an ideal amount of (forward and backward) perturbation correction needed during gradient approximation. The correction needs to keep the integrity of the underlying direction, yet ensure the perturbations fulfil certain constraints. There is potential to extend this investigation further by comprehensively studying the available constraint handling techniques and their effects on gradient approximation and algorithm performance.

Also, the consideration of geological uncertainty in gradient-based methods is an important extension. One may extend the proposed method by simply using the expected NPV across a number of realizations as the objective function. Other more efficient, but less accurate, approaches may approximate a gradient using an average of gradient approximations which utilize stochastically selected realizations. This requires additional research to investigate the implications of such techniques on convergence.

6 Concluding remarks

This study compared the proposed algorithm (Adam-SPSA) with the conventional steepest descent SPSA (SD-SPSA) and a derivative-free pattern search algorithm (GPS) for constrained well placement optimization. The results presented showed the successful application of local optimization algorithms to well placement optimization, including vertical wells and nonconventional wells. These algorithms leveraged local exploitation to improve the objective function value from an initial guess. The gradient-based methods (Adam-SPSA and SD-SPSA) performed effectively in both the low-dimensional and high-dimensional case studies by significantly improving on the initial objective function value. However, the derivative-free method (GPS) was not competitive in the high-dimensional case. This can be attributed to the simultaneous perturbation used in the gradient-based methods, which allows for the progression in all dimensions during one step. When comparing the gradient-based methods, Adam-SPSA consistently outperformed SD-SPSA in both case studies investigated. This can be ascribed to the adaptive nature of the search direction, which incorporates estimates of the first and second moment. This allows the search direction to be calculated in a dimension-wise manner leading to more suitable progression steps for each decision variable. Lastly, the decision variables must have a similar sensitivity to the objective

function for the parameterization employed. For nonconventional wells, the Cartesian parameterization showed lower sensitivity compared to the spherical parameterization.

Acknowledgements The authors would like to thank Steve Begg for his support and fruitful discussion. The authors would also like to thank David Echeverría Ciaurri for his help with the code relating to the inter-well constraint. The authors are also grateful to Rock Flow Dynamics for providing licenses for tNavigator. This work was supported with supercomputing resources provided by the Phoenix HPC service at the University of Adelaide.

Funding Open Access funding enabled and organized by CAUL and its Member Institutions.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Ahmad, A., El-Shafie, A., Razali, S.F.M., Mohamad, Z.S.: Reservoir optimization in water resources: a review. *Water Resour. Manag.* **28**(11), 3391–3405 (2014)
- Cameron, D.A., Durlofsky, L.J.: Optimization of well placement, CO₂ injection rates, and brine cycling for geological carbon sequestration. *Int. J. Greenhouse Gas Control* **10**, 100–112 (2012). <https://doi.org/10.1016/j.ijggc.2012.06.003>, <https://www.sciencedirect.com/science/article/pii/S1750583612001296>
- Aroui, Y., Sayyafzadeh, M.: An accelerated gradient algorithm for well control optimization. *J. Pet. Sci. Eng.* **190**, 106872 (2020). <https://doi.org/10.1016/j.petrol.2019.106872>, <http://www.sciencedirect.com/science/article/pii/S0920410519312884>
- Alrashdi, Z., Sayyafzadeh, M.: (+) evolution strategy algorithm in well placement, trajectory, control and joint optimisation. *J. Pet. Sci. Eng.* **177**, 1042–1058 (2019). <https://doi.org/10.1016/j.petrol.2019.02.047>, <http://www.sciencedirect.com/science/article/pii/S0920410519301846>
- Sayyafzadeh, M.: Reducing the computation time of well placement optimisation problems using self-adaptive metamodelling. *J. Pet. Sci. Eng.* **151**, 143–158 (2017)
- Bittencourt, A.C., Horne, R.N.: Reservoir development and design optimization. In: SPE Annual Technical Conference and Exhibition. Society of Petroleum Engineers (1997). <https://doi.org/10.2118/38895-MS>
- Yeten, B., Durlofsky, L.J., Aziz, K.: Optimization of nonconventional well type, location, and trajectory. *SPE J.* **8**(03), 200–210 (2003)
- Onwunalu, J.E., Durlofsky, L.J.: Application of a particle swarm optimization algorithm for determining optimum well location and type. *Comput. Geosci.* **14**(1), 183–198 (2010). <https://doi.org/10.1007/s10596-009-9142-1>
- Sayyafzadeh, M., Alrashdi, Z.: Well controls and placement optimisation using response-fed and judgement-aided parameterisation: Olympus optimisation challenge. *Comput. Geosci.* <https://doi.org/10.1007/s10596-019-09891-y> (2019)
- Bouzarkouna, Z., Ding, D.Y., Auger, A.: Well placement optimization with the covariance matrix adaptation evolution strategy and meta-models. *Comput. Geosci.* **16**(1), 75–92 (2012)
- Martí, R.: Multi-start Methods, pp. 355–368. Springer (2003)
- O'Donoghue, B., Candàs, E.: Adaptive restart for accelerated gradient schemes. *Found. Comput. Math.* **15**(3), 715–732 (2015). <https://doi.org/10.1007/s10208-013-9150-3>
- Bellout, M.C., EcheverríaCiaurri, D., Durlofsky, L.J., Foss, B., Kleppe, J.: Joint optimization of oil well placement and controls. *Comput. Geosci.* **16**(4), 1061–1079 (2012). <https://doi.org/10.1007/s10596-012-9303-5>
- Isebor, O.J., EcheverríaCiaurri, D., Durlofsky, L.J.: Generalized field-development optimization with derivative-free procedures. *SPE J.* 891–908. <https://doi.org/10.2118/163631-PA> (2014)
- Wang, H., Ciaurri, D.E., Durlofsky, L.J.: Use of retrospective optimization for placement of oil wells under uncertainty. In: Proceedings of the 2010 Winter Simulation Conference, pp. 1750–1757 (2010)
- Wang, H., Echeverría-Ciaurri, D., Durlofsky, L., Cominelli, A.: Optimal well placement under uncertainty using a retrospective optimization framework. *SPE J.* **17**(01), 112–121 (2012). <https://doi.org/10.2118/141950-PA>
- Jansen, J.D.: Adjoint-based optimization of multi-phase flow through porous media - a review. *Comput. Fluids* **46**(1), 40–51 (2011). <https://doi.org/10.1016/j.compfluid.2010.09.039>, <http://www.sciencedirect.com/science/article/pii/S0045793010002677>
- Sarma, P., Chen, W.H.: Efficient well placement optimization with gradient-based algorithms and adjoint models. In: Intelligent Energy Conference and Exhibition (2008)
- Wang, C., Li, G., Reynolds, A.C.: Optimal well placement for production optimization. In: Eastern Regional Meeting, p. 5. Society of Petroleum Engineers (2007). <https://doi.org/10.2118/111154-MS>
- Zandvliet, M., Handels, M., van Essen, G., Brouwer, R., Jansen, J.-D.: Adjoint-based well-placement optimization under production constraints. *J. Pet. Sci. Eng.* <https://doi.org/10.2118/105797-PA> (2008)
- Bangerth, W., Klie, H., Wheeler, M.F., Stoffa, P.L., Sen, M.K.: On optimization algorithms for the reservoir oil well placement problem. *Comput. Geosci.* **10**(3), 303–319 (2006). <https://doi.org/10.1007/s10596-006-9025-7>
- Jesmani, M., Jafarpour, B., Bellout, M.C., Hanea, R.G., Foss, B.: Application of simultaneous perturbation stochastic approximation to well placement optimization under uncertainty. In: ECMOR XV-15th European Conference on the Mathematics of Oil Recovery, pp. cp-494–00133. European Association of Geoscientists & Engineers (2016)
- Vlemmix, S., Joosten, G., GerardJP., Brouwer, R., Jansen, J.-D.: Adjoint-based well trajectory optimization. In: EUROPEC/EAGE Conference and Exhibition. Society of Petroleum Engineers (2009)
- Volkov, O., Bellout, M.C.: Gradient-based constrained well placement optimization. *J. Pet. Sci. Eng.* **171**, 1052–1066 (2018). <https://doi.org/10.1016/j.petrol.2018.08.033>, <http://www.sciencedirect.com/science/article/pii/S0920410518306995>

25. Leeuwenburgh, O., Egberts, P.J., PAbbink, O.A.: Ensemble methods for reservoir life-cycle optimization and well placement. In: SPE/DGS Saudi Arabia Section Technical Symposium and Exhibition, p. 8. Society of Petroleum Engineers (2010). <https://doi.org/10.2118/136916-MS>
26. Li, L., Jafarpour, B., Mohammad-Khaninezhad, M.R.: A simultaneous perturbation stochastic approximation algorithm for coupled well placement and control optimization under geologic uncertainty. *Comput. Geosci.* **17**(1), 167–188 (2013). <https://doi.org/10.1007/s10596-012-9323-1>
27. Guyaguler, B., Horne, R.N.: Uncertainty assessment of well placement optimization. In: SPE Annual Technical Conference and Exhibition. Society of Petroleum Engineers (2001)
28. Ding, Y.: Optimization of well placement using evolutionary methods. In: Europec/EAGE Conference and Exhibition. Society of Petroleum Engineers (2008)
29. Forouzanfar, F., Reynolds, A.C.: Well-placement optimization using a derivative-free method. *J. Pet. Sci. Eng.* **109**, 96–116 (2013). <https://doi.org/10.1016/j.petrol.2013.07.009>, <http://www.sciencedirect.com/science/article/pii/S0920410513001952>
30. Emerick, A.A., Silva, B., Almeida, L.F., Szwarcman, D., Pacheco, M.A.C., Vellasco, M.M.B.R.: Well placement optimization using a genetic algorithm with nonlinear constraints. In: SPE Reservoir Simulation Symposium. Society of Petroleum Engineers (2009)
31. Wu, Y., Schuster, M., Chen, Z., Le, Q.V., Norouzi, M., Macherey, W., Krikun, M., Cao, Y., Gao, Q., Macherey, K.: Google's neural machine translation system: Bridging the gap between human and machine translation. arXiv:1609.08144 (2016)
32. Gregor, K., Danihelka, I., Graves, A., Rezende, D.J., Wierstra, D.: Draw: A recurrent neural network for image generation. arXiv:1502.04623 (2015)
33. Ledig, C., Theis, L., Huszár, F., Caballero, J., Cunningham, A., Acosta, A., Aitken, A., Tejani, A., Totz, J., Wang, Z.: Photo-realistic single image super-resolution using a generative adversarial network. In: Proceedings of the IEEE conference on Computer Vision and Pattern Recognition, pp. 4681–4690 (2017)
34. Spall, J.C.: Multivariate stochastic approximation using a simultaneous perturbation gradient approximation. *IEEE Trans. Autom. Control* **37**(3), 332–341 (1992)
35. Sadegh, P., Spall, J.C.: Optimal random perturbations for stochastic approximation using a simultaneous perturbation gradient approximation. *IEEE Trans. Autom. Control.* **43**(10), 1480–1484 (1998)
36. Spall, J.C.: Implementation of the simultaneous perturbation algorithm for stochastic optimization. *IEEE Trans. Aerospace. Electron. Syst.* **34**(3), 817–823 (1998)
37. Kingma, D.P., Ba, J.: Adam: A method for stochastic optimization. arXiv:1412.6980 (2014)
38. Lefebvre, M. *Applied Probability and Statistics*, 1st edn. Springer, New York (2006)
39. MathWorks. Matlab (2018)
40. Floris, F.J.T., Bush, M.D., Cuypers, M., Roggero, F., Syversveen, A.-R.: Methods for quantifying the uncertainty of production forecasts: a comparative study. *Pet. Geosci.* **7**(S), S87–S96 (2001). <https://doi.org/10.1144/petgeo.7.S.S87>, <https://www.earthdoc.org/content/journals/10.1144/petgeo.7.S.S87>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

4.2 A study of simulation-based surrogates in well placement optimization for hydrocarbon production

Aroui, Y.; Echeverría Ciaurri, D. & Sayyafzadeh, M.

Journal of Petroleum Science and Engineering, 2022.

Statement of Authorship

Title of Paper	A study of simulation-based surrogates in well-placement optimization for hydrocarbon production
Publication Status	<input checked="" type="checkbox"/> Published <input type="checkbox"/> Accepted for Publication <input type="checkbox"/> Submitted for Publication <input type="checkbox"/> Unpublished and Unsubmitted work written in manuscript style
Publication Details	Arouri, Y.; Echeverría Ciaurri, D & Sayyafzadeh, M. A study of simulation-based surrogates in well-placement optimization for hydrocarbon production. Journal of Petroleum Science and Engineering, Published

Principal Author

Name of Principal Author (Candidate)	Yazan Arouri		
Contribution to the Paper	Coding, experimentation, analysis of results, original manuscript		
Overall percentage (%)	60%		
Certification:	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.		
Signature		Date	21/05/2022

Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

- the candidate's stated contribution to the publication is accurate (as detailed above);
- permission is granted for the candidate to include the publication in the thesis; and
- the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author	David Echeverría Ciaurri		
Contribution to the Paper	Coding, analysis of results, original manuscript, revisions		
Signature		Date	25/05/2022

Name of Co-Author	Mohammad Sayyafzadeh		
Contribution to the Paper	Discussions, analysis of results, revisions writing, supervision, methodology		
Signature		Date	06/06/2022

Please cut and paste additional co-author panels here as required.



A study of simulation-based surrogates in well-placement optimization for hydrocarbon production

Yazan Arouri ^{a,*}, David Echeverría Ciaurri ^{b,1}, Mohammad Sayyafzadeh ^a

^a Australian School of Petroleum and Energy Resources, The University of Adelaide, Adelaide, 5005, SA, Australia

^b Urbasa Energetica, New York, 10024, NY, USA

ARTICLE INFO

Keywords:

Surrogate-based optimization
Well placement
Reservoir flow simulation
Derivative-free optimization
Hydrocarbon production

ABSTRACT

An essential part of hydrocarbon field-development planning is the determination of well-drilling locations within a reservoir, which can be formulated as an optimization problem. Lending to the nonlinear nature and potentially high-dimensional search spaces, this problem is computationally intensive as it entails time-demanding reservoir flow simulations. To make matters worse, most practical decision-making scenarios in field development are subject to constraints on available computing resources (this also includes commercial-software licenses). In this paper, we study a surrogate-based well-placement optimization approach especially devised for these challenging scenarios. In this approach, surrogates are constructed either analytically (using a number of previous simulation runs) or by simplifying the underlying physics of the problem. The surrogates are iteratively corrected in a derivative-free, noninvasive (regarding the simulator) manner, through manifold mapping, a multi-fidelity technique. This is the first application of manifold mapping to well-placement optimization problems. We considered two optimization problems that involve benchmark models with different degrees of complexity. Multiple implementations of the methodology are compared with a single-level optimization method (i.e., that does not utilize any surrogate explicitly). The effect of surrogate accuracy and its impact on the quality of the optimized solutions is explored. We also discuss limitations of the analytical surrogates in the presence of complex optimization spaces and physical phenomena. It is shown here that physics-based surrogates are more suitable choices in these situations and still accelerate computations noticeably with respect to the single-level method. The promising findings in this study reflect how the optimization of well placement in hydrocarbon fields under (stringent) constraints on computing resources could be approached.

1. Introduction

Once a prospective natural resource is discovered and appraised to contain economic volumes of hydrocarbons (oil and/or gas), the subsequent step involves designing the most sustainable and economic extraction procedure, also known as a field development plan (FDP). The FDP comprises the specific details and decisions for aspects such as the location and trajectory of the wells, the type of wells (e.g., producer or injector), the number of wells and the drilling schedule, amongst other considerations. An essential part of field development planning is to conduct what if analysis and study the effect that different FDPs have on the recovery of the hydrocarbons. This entails performing full physics reservoir simulation to solve a set of nonlinear partial differential equations and find the state variables (pressures and saturations) throughout a reservoir's lifetime in response to any proposed

FDP. The output of the simulation provides a forecast of the produced fluid volumes. These volumes can then be used to calculate economic metrics such as net present value (NPV). The nonlinear governing equations render reservoir simulations that are computationally demanding, especially for systems with complex geological settings and/or fluid properties each run may take hours or even days to execute.

Given the highly nonlinear relationship between the input variables (development strategy and parameters of the system) and the output of reservoir simulation (produced fluid volumes), computer aided FDP optimization (see, e.g., Echeverría Ciaurri et al. 2012, Isebor et al. 2014b, Echeverría Ciaurri et al. 2011b, Alrashdi and Sayyafzadeh 2019, Arouri and Sayyafzadeh 2020, Arouri et al. 2022b) has garnered significant attention recently. The decisions of interest in FDP (e.g., well location and trajectory) are approached through the formulation of an

* Corresponding author.

E-mail addresses: yazan.arouri@adelaide.edu.au (Y. Arouri), david.echeverria@urbasa-energetica.com (D. Echeverría Ciaurri), mohammad.sayyafzadeh@adelaide.edu.au (M. Sayyafzadeh).

¹ Now at Saudi Aramco.

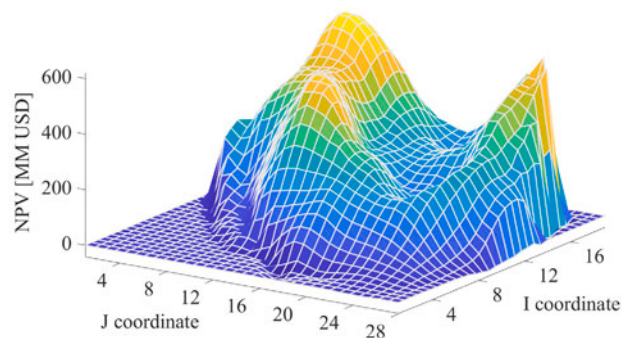


Fig. 1. NPV associated with the optimization of the areal location of one vertical well for a given reservoir model.

optimization problem where the values for certain variables are set with the aim of, without the loss of generality, minimizing (or maximizing) a predefined objective function (or more than one, if multi objective criteria are considered), while meeting a number of constraints.

In this paper we study the well placement component of FDP optimization, which, due to the heterogeneity of rock properties and their impact on fluid flow, is nontrivial and yields multimodal and nonconvex objective function landscapes. Fig. 1 illustrates the NPV associated with the optimization of the areal positioning of just one vertical well for a given reservoir model. The number of peaks in the cost function is expected to increase with the number of optimization variables. Hence, optimization problems that include the selection of well locations are often addressed using global search techniques. The global approaches typically applied to these problems are derivative free optimization algorithms (see, e.g., Kramer et al. 2011, Conn et al. 2009). Examples of global, derivative free methods used in well placement optimization are genetic algorithms (GAs; see, e.g., Emerick et al. 2009, Guyaguler and Horne 2001, Sayyafzadeh 2017), evolution and covariance matrix adaptation evolution strategies (see, e.g., Sayyafzadeh and Alrashdi 2019, Bouzarkouna et al. 2012, Ding 2008), and particle swarm optimization (see, e.g., Ding et al. 2014, Onwunali and Durlofsky 2010, Arouri et al. 2022b). Although the application of global search algorithms to field development studies is promising, they arguably rely on a prohibitive number of computationally expensive reservoir simulations. Indeed, global exploration suffers from the curse of dimensionality: note that, for example, a box in a space of n dimensions has 2^n vertices and that a thorough search within that box may generally entail evaluating more points than these vertices.

Solutions in practice are often required in a relatively short time frame – few days, which typically translates to a computing budget of a few hundred simulations. Instances of such scenarios may be given when a solution has to be computed, for example, in one week with only one software license for a simulation model that takes approximately 30 min (this yields a computational budget of 300–400 simulations). In another case, a solution may need to be determined within a weekend for simulations that, provided the required software licenses are available, clock in at roughly 15 min when parallelized on multiple cores/nodes (for a computing budget of about 200 simulations).

Local optimization of well placement may provide, in some situations, acceptable solutions readily. Gradient based methods, with the gradient computed precisely through adjoint procedures (see, e.g., Jansen 2011, Sarma and Chen 2008, Wang et al. 2007, Zandvliet et al. 2008) or approximated (see, e.g., Arouri et al. 2022a, Bangerth et al. 2006, Jesmani et al. 2016), can be alternatives in these situations. Note that adjoint based methods are simulator invasive and for that

reason very often not applicable when commercial software is used. Numerical gradients may need certain amount of tuning with respect to the perturbation size. Additionally, in some simulators this size cannot be (for well placement optimization) smaller than one grid block. In these circumstances one could resort to simplex gradients (see, e.g., Wang et al. 2012) or to pattern search algorithms (see, e.g., Bellout et al. 2012). Local methods can incorporate mechanisms that, to some extent, aim at avoiding convergence to solutions that might not be satisfactory regarding cost function value. Examples of these mechanisms are heuristics for the initial guess in the optimization, gradient estimation using a random direction in simultaneous perturbation stochastic approximation (see, e.g., Spall 1992) and the use of varying stencils in mesh adaptive direct search (see, e.g., Audet and Dennis 2006) or of large stencils in initial iterations of pattern search methods. Hybridization of global and local procedures (see, e.g., Arouri et al. 2022b, Alghareeb 2015, Bittencourt and Horne 1997, Isebor et al. 2014a, Nwankwor et al. 2013) represents a tradeoff between the benefits (and sometimes the drawbacks) of the two types of approaches. In this paper we focus on acceleration of local, simulator noninvasive methods for optimal well placement.

The performance of all algorithms can be improved by means of surrogates, also known as proxies or meta models. We can distinguish two major families of surrogates. In the first instance, surrogates can be constructed using analytical techniques based on previous simulation results or measurements. In essence, these surrogates are constructed by generating a set of data points that are used to “train” and calibrate the model, which is then used to predict. Most of these techniques perform interpolation or approximation in some form or another. Examples of these analytical surrogates and of their use include polynomial approximation (see, e.g., Peng and Gupta 2004, Yeten et al. 2005, Zubarev 2009), kriging (see, e.g., Peng and Gupta 2004, Yeten et al. 2005, 2003), thin plate splines (see, e.g., Li and Friedmann 2005), and artificial neural networks (ANNs; see, e.g., Yeten et al. 2003, Cullick et al. 2006, Sayyafzadeh 2017).

The second type of surrogates are based on simplifications of scientific laws (and/or of their respective numerical solution) present in the full order models. These approximate models, which will be referred to here as physics based surrogates, frequently require domain knowledge and, for comparable computational cost, are normally more accurate than the first type of surrogates mentioned (or, equivalently, they require a smaller computing effort for similar precision). On the other hand, surrogates built upon analytical techniques are of more straightforward implementation (this easier construction may be translated to time savings as well). Examples of physics based surrogates in reservoir fluid flow are decline curve analysis (see, e.g., Li and Horne 2005), capacitance resistance models (see, e.g., Sayarpour et al. 2009, Yousef et al. 2006), and the use of the Buckley Leverett equation (see, e.g., Buckley and Leverett 1941), of streamlines (see, e.g., Batycky et al. 1997, Datta Gupta and King 2007), of single phase flow (see, e.g., Wolfsteiner et al. 2000) and of component lumping for compositional simulation (see, e.g., Alavian et al. 2014).

There are surrogates that could be considered a combination of the two types just introduced. These surrogates are computed using analytic procedures, such as linearization and projection from high dimensional to low dimensional spaces, to reduce the number of degrees of freedom related to the state variables. The implementation of these models also requires some domain knowledge. In these approaches the surrogates refer to the simulator itself and not only to the cost and constraints functions, and in that manner they may capture model trends that could be difficult to detect through variation of one or few scalar quantities. Examples of techniques used for these surrogates are proper orthogonal decomposition (see, e.g., Alghareeb 2015, Cardoso et al. 2009, Jansen and Durlofsky 2017, van Doren et al. 2006), missing point estimation (see, e.g., Cardoso et al. 2009), trajectory piecewise linearization (see, e.g., Cardoso et al. 2009, Jansen and Durlofsky 2017) and the discrete empirical interpolation method

(see, e.g., [Alghareeb 2015](#), [Suwartadi et al. 2015](#)). Since linearization is a cornerstone of these methods, we can expect that if applied to the optimization of well placement, due to the significant nonlinearities often present in these problems, the accuracy of the surrogates obtained may not be high. However, as proposed in [Alghareeb \(2015\)](#), this type of surrogates, when iteratively corrected, can still be used to accelerate local search.

Surrogates that aim at the approximation of the entire search space, if constructed using previous simulation runs, may demand excessive computational cost (these surrogates, as global exploration, can be expected to suffer from the curse of dimensionality). In the applications addressed in this work, the computational budget, which could be just a few hundred simulation runs, precludes targeting accurate and global surrogates. Nevertheless, a surrogate based on a relatively small number of data points can still capture some global trends and detect promising regions in the search space, which may be later explored locally through refinement of the surrogate. That is our approach, which consists of first performing optimization based on an initial surrogate that is later progressively corrected as the search advances. The initial optimization includes some amount of global exploration, consistent with the computing budget available. The surrogate correction is also paired with local, efficient optimization of the improved surrogate. This iterative correction and optimization of the surrogate is accomplished through manifold mapping (see, e.g., [Echeverría Ciaurri 2007](#), [Echeverría and Hemker 2005](#)), a multi level optimization technique with solid theoretical foundations that has been successfully applied in multiple disciplines, including reservoir engineering (see [Echeverría Ciaurri and Wagenaar 2016](#), [Echeverría Ciaurri et al. 2021](#) for examples in well control optimization). This work is the first application of manifold mapping to well placement optimization problems.

Optimization of well placement aided by surrogates built on simulation runs has received noticeable attention in the last two decades. For the reasons already mentioned regarding nonlinearities in the problem, the majority of the methods resort to global exploration. In [Yeten et al. \(2003\)](#) and in [Guyaguler et al. \(2000\)](#) a GA was combined with ANNs and, in the latter work, also with kriging. In both cases the global search method was hybridized with a local optimization procedure. The approach presented in [Artus et al. \(2006\)](#) incorporates specific treatment of geological uncertainty and also relies on a GA but in this case the (statistical) proxy is constructed using k means clustering. Although all these GA based techniques yield solutions in a relatively efficient manner, due to the global nature of the optimization approach and surrogate used, it is unclear how performance will scale in problems with more variables (curse of dimensionality). In [Bouzarkouna et al. \(2012\)](#) they coupled a covariance matrix adaptation evolution strategy, a local optimization algorithm, with quadratic surrogates based upon k nearest points. However, this surrogate based approach still requires a few thousand simulation runs for optimizing two unilateral wells (12 optimization variables). Multiple analytic surrogates (i.e., kriging, thin plate splines, and ANNs) are considered in [Zubarev \(2009\)](#); nonetheless the problems solved in that work (optimization of areal location of at most two vertical wells) are somewhat too simple to draw meaningful conclusions. In all these works on surrogate based well placement optimization, the surrogate, although iteratively improved, only refers to the cost function and not to other richer simulation output that may be leveraged to obtain more precise approximations. The methodology presented in [Alghareeb \(2015\)](#) relies on local optimization enhanced with some global exploration and resorts to surrogates for the simulator itself. These surrogates are of the hybrid type discussed above and require knowledge of the particular system of equations and implementation of its numerical solving (which in certain cases can be far from trivial). In our work, we opt for ease of implementation and flexibility with respect to the simulator used at the expense of some additional computing effort (still within the budget given).

As noted earlier, the approach followed here for the optimization of well location is based on local search with some amount of global

exploration (consistent with computing resources that allow around few hundred simulation runs in total). The optimization method is of derivative free (simulator noninvasive) nature and is combined with a surrogate that is initially constructed either upon a number of previous simulation runs or by simplifying the underlying physics. The surrogate is later corrected locally in every iteration to improve the current best solution. Uncertainty, which is outside the scope of this work, poses additional challenges and can be handled, for example, by means of the retrospective optimization framework presented in [Wang et al. \(2012\)](#) or using a rank based subset selection technique introduced in [Arouri et al. \(2022a\)](#) or as in [Sayyafzadeh and Alrashdi \(2019\)](#) for the OLYMPUS challenge ([Fonseca et al., 2020](#)). As we will see later in the paper, we have found evidences within optimal well location of possible limitations of the use of surrogates built on a relatively small number of simulation runs. To the best of our knowledge, these limitations have not been identified in FDP for other similar methodologies published. The use of physics based surrogates may be, as also shown in one example in this research, an alternative to efficiently optimize well locations in practical scenarios.

To begin, the definition of the optimization problem for well placement is presented. Next, the different optimization methodologies utilized, including manifold mapping, a multi level approach for iterative correction of the surrogate, are described. Following this, the results for two case studies are presented and discussed. Finally, we end the paper with some concluding remarks and future work directions.

2. Problem definition

In this section we frame the optimization problem mathematically and introduce the objective function that is used in the remainder of the paper.

2.1. Problem statement

The optimization problem considered is aimed at minimizing a defined objective function, where the decision variables \mathbf{x} refer to well locations. In this study, we parameterize the well locations as continuous areal Cartesian coordinates (x and y) since the (commercial) simulator used allows the positioning of wells anywhere in each discretization grid block. Parameterizations based on cell indices (i and j) are often used when wells can only be located at the center of the blocks. We formulate the well placement optimization problem using a single objective function as follows:

$$\min_{\mathbf{x} \in \mathbb{R}^n} f(\mathbf{x}) \quad \text{subject to} \quad \mathbf{c}(\mathbf{x}) \leq 0, \quad (1)$$

where $f(\mathbf{x})$ is the objective function, \mathbf{x} denotes the vector of decision variables, n is twice the number of wells whose location is optimized and $\mathbf{c}(\mathbf{x}) \in \mathbb{R}^m$ represents the problem constraints.

Here, we consider the objective function to be the negative of the NPV for a given lifespan. Note that in the remainder of the paper NPV will be plotted in the figures. NPV is an economic metric that, for the problems studied in this paper, takes into account oil price, produced water handling expenses and water injection costs. The objective function is defined as

$$\begin{aligned} f(\mathbf{x}) &= -\text{NPV}(\mathbf{x}) \\ &= \sum_{i=1}^{N_t} \frac{-r_{o,i} Q_{o,i}(\mathbf{x}) + c_{wp,i} Q_{wp,i}(\mathbf{x}) + c_{wi,i} Q_{wi,i}(\mathbf{x})}{(1+b)^{t_i}}, \end{aligned} \quad (2)$$

where $r_{o,i}$, $c_{wp,i}$ and $c_{wi,i}$ are the sale price of oil and costs of water separation and injection, respectively, all of them per unit volume and defined from time (in years) t_i to t_{i+1} (there are N_t such intervals), $Q_{o,i}$, $Q_{wp,i}$ and $Q_{wi,i}$ denote the field oil production, field water production and field water injection volumes during the mentioned output interval and b is the annual discount rate. The production and injection volumes are obtained from the reservoir simulation for a defined set

of optimization variables (well locations) and for a specified reservoir model. Since the number of wells to be drilled (and their respective type) and the drilling sequence are not subjected to optimization in these problems, the cost and timing associated with drilling future wells is not included in the NPV (see, e.g., Isebor et al. 2014b for a definition of NPV that incorporates this cost in the context of generalized field development, a broader set of optimization problems). Other possible objective functions may include cumulative oil production or, in the presence of uncertainty, risk attitude considerations (see, e.g., Yeten et al. 2004).

The constraint functions, in the case studies presented in this paper, incorporate bounds for the drilling locations and for the minimum distance between any two wells. The bounds for the drilling locations prevent the wells from being located in certain (undesirable) regions.

3. Optimization methodology

As indicated earlier, the optimization methodology studied in this research relies on local, simulator noninvasive, surrogate based methods in order to obtain solutions rapidly, with implementations that can be applied, in principle, to any simulation software. Next, we will provide more details on the basic components of the methodology.

3.1. Pattern search optimization

Pattern search optimization represents a family of derivative free methods that have solid theoretical foundations (see, e.g., Torczon 1997, Kolda et al. 2003) and deliver fairly efficient performance. The gradient free nature of these methods makes them suitable to be used with any simulator. In essence, pattern search unconstrained optimization proceeds in each iteration as follows. A stencil, which is a collection of points in the (n dimensional) search space, is centered at the best solution found so far. The stencil is required to satisfy certain properties (see Kolda et al. 2003), in particular it needs to have at least $n + 1$ points. A popular stencil is the compass, that is, a collection of points obtained via positive and negative perturbations of equal size and in all coordinate directions in the search space (that yields $2n$ points, besides the stencil center). If the objective function value of a stencil point improves the best solution, the stencil can be centered at that point for a new iteration (the stencil size could also be increased although usually it is not modified). Otherwise, the stencil size is reduced and the new smaller stencil is considered in the next iteration. Convergence theory requires that the stencil size progressively tends towards zero (see Kolda et al. 2003). The stencil size being smaller than a given threshold (that minimum size can be related to actual application tolerances such as resolution in the well drilling coordinates) can be used as stopping criterion in the optimal search process (this criterion is often combined with a budget for the number of simulation runs). Since the points in the stencil can be evaluated in parallel, pattern search methods are amenable to distributed computing implementations that can be very efficient in terms of wall clock time. Alternatively, when computing resources are scarce (this includes licenses for commercial software) the stencil points can be polled in opportunistic fashion, i.e., the basic iteration is finished as soon as a new point improving the current best solution is found. Pattern search methods may also incorporate some degree of global exploration, for example, by means of a relatively large initial stencil size or, as in mesh adaptive direct search (see, e.g., Audet and Dennis 2006), through random variation of the stencil in each iteration.

The pattern search method applied in this study is Hooke Jeeves direct search (HJDS; see, e.g., Hooke and Jeeves 1961). HJDS has been already applied to FDP, concretely to the well control optimization component, and can be an attractive alternative when computing resources are modest (see Echeverría Ciaurri et al. 2011a). In HJDS, the basic pattern search iteration described above, which is referred

to as the exploratory move in HJDS lexicon, is followed by a so called pattern move. The stencil points in the exploratory move are evaluated opportunistically. However, upon success in this evaluation, the poll in the new stencil does not start from the first direction in the stencil but from the next direction after the last one checked. This continues until all search directions have been considered. As in the basic pattern search iteration, if no improvement occurs within the exploratory move, the stencil size is reduced. Otherwise, the pattern move is undertaken, which is an aggressive step in the previous overall successful exploratory direction. If this pattern move does not improve the current best solution, a new exploratory move centered at this point is performed (reduction of the stencil size follows when this move is unsuccessful).

3.2. Surrogate models

In this work the surrogates do not approximate the objective function $f(\mathbf{x})$ but a vectorial model response $\mathbf{m}(\mathbf{x}) \in \mathbb{R}^p$ related to it:

$$f(\mathbf{x}) = U(\mathbf{m}(\mathbf{x})), \quad (3)$$

where $U: \mathbb{R}^p \rightarrow \mathbb{R}$ denotes a functional that connects objective function and model response. While in the physics based surrogate used in this work we approximate $\mathbf{m}(\mathbf{x})$ directly, for the analytical surrogates we construct an independent surrogate for each of the vector components $m_i(\mathbf{x})$. For the sake of clarity in the notation, in the remainder of the paper we will refer to any of these components as $m(\mathbf{x})$. The rationale behind targeting a vector in the approximation is that surrogate corrections may be then improved due to a larger use of information than if only a scalar is considered. Specifically, the model response represents a times series of NPV contributions for different intervals and the functional is the negative of the sum of all the vector components. The number of time intervals N_t are the same for the surrogate and the full physics simulation model. Note that other magnitudes such as field and well injection/production rates may be selected as model response.

We reiterate that our applications are constrained by computational budgets that are common in practice, frequently of few hundred simulations. These budgets include the construction of the analytical surrogates employed. In order to make economic use of simulation runs required in the surrogate generation stage we will leverage design of experiments (DoE; see, e.g., Santner et al. 2003, Koehler and Owen 1996). The DoE techniques tested are central composite design (CCD; see, e.g., Khuri and Cornell 1996) and Latin hypercube sampling (LHS; see, e.g., McKay et al. 1979). The DoE procedures used aim at performing some global exploration of the search space so that the surrogate obtained captures overall trends to better drive the optimization. Once a promising region is identified, local correction of the surrogate may facilitate exploitation within the optimal search. Next, we will briefly describe the surrogate methods applied.

3.2.1. Polynomial approximation

Polynomial approximation of the function $m(\mathbf{x})$ can be formulated as the fitting of N_s polynomial basis functions $p_j(\mathbf{x})$

$$m(\mathbf{x}) \approx \tilde{m}(\mathbf{x}) = \sum_{j=1}^{N_s} \beta_j p_j(\mathbf{x}), \quad (4)$$

where $\boldsymbol{\beta} = [\beta_1 \beta_2 \dots \beta_{N_s}]^T$ is a vector of parameters determined by means of a regression problem built upon N evaluations of the function $m(\mathbf{x})$. It is important to note that N_s increases exponentially if all polynomials up to a given order are considered. This, within the applications targeted in this work, can lead to a large number of evaluations of $m(\mathbf{x})$ even for relative small search spaces. For example, if a full quadratic (cubic) is used, N_s would be between 66 (286) and 231 (1771) for values of n between 10 and 20. Since these (or larger) values of n are common in real life problems, for example, when the

drilling areal locations of 5 to 10 vertical wells are optimized, full cubic expansions may lead to prohibitive implementations. We note that quadratic approximation, the one tested in this work, is suggested as a companion of CCD (see [Khuri and Cornell 1996](#)). The computational cost associated with the evaluation of a polynomial surrogate is negligible when compared with reservoir fluid flow simulation.

3.2.2. Kriging

Kriging (see, e.g., [Journal and Huijbregts 1981](#), [Kleijnen 2009](#)) is an interpolation technique based on Gaussian processes (see, e.g., [Rasmussen and Williams 2006](#)) where the function of interest $m(\mathbf{x})$ is assumed to be composed of a regression model M (to remove possible trends) and a stochastic function Z (residuals):

$$m(\mathbf{x}) \approx \tilde{m}(\mathbf{x}) = M(\mathbf{x}) + Z(\mathbf{x}) = \mathbf{g}(\mathbf{x})^T \boldsymbol{\beta} + Z(\mathbf{x}), \quad (5)$$

where $\mathbf{g}(\mathbf{x}) \in \mathbb{R}^{N_s}$ is a vector of known functions and $\boldsymbol{\beta} \in \mathbb{R}^{N_s}$ are the unknown regression parameters. The stochastic function Z represents a Gaussian process with mean zero and covariance matrix

$$\Sigma_{\mathbf{xw}} = \sigma^2 R(\boldsymbol{\theta}, \mathbf{x}, \mathbf{w}), \quad (6)$$

where R is a correlation model with an unknown vector of parameters $\boldsymbol{\theta}$. In this work the following linear correlation model was considered:

$$R(\boldsymbol{\theta}, \mathbf{x}, \mathbf{w}) = \prod_{i=1}^n \max\{0, 1 - \theta_i |x_i - w_i|\}. \quad (7)$$

The kriging surrogate was obtained by means of the toolbox Design and Analysis of Computer Experiments (DACE) in MATLAB (see [Nielsen et al. 2002](#)).

3.2.3. Two dimensional reservoir model approximation

As a physics based surrogate we employ a two dimensional reservoir model approximation that relies on local grid coarsening (LGC). LGC reduces the number of grid blocks in the model with the aim of decreasing the run time associated with the solution of the underlying system of partial differential equations. The particular implementation in this study is based on a coarsened discretization where the cell dimensions in the horizontal directions are equal to the respective average over each direction whilst the cell dimension in the vertical direction is the sum. With regards to the rock properties, such as permeability and porosity, for each coarsened grid block we take the value of the pore volume weighted average of the corresponding original scale cells. Note that initial conditions for the differential equations have to be treated as well. Initial pressure is also averaged and, as far as the initial saturation is concerned, the centermost block of the selected original cells is taken as representative. Therefore, in a model, for example, with nine layers, the fifth layer will be representative of the coarsened grid blocks. In this study, an LGC implementation was available in the commercial reservoir simulator used (tNavigator). This LGC implementation arguably preserves areal heterogeneity in the model, which has significant impact on vertical well placement. Since solution methods are, in general, at a minimum of linear computational complexity with respect to the number of grid blocks, we can expect that the implementation will provide speed up factors at least on the order of the reduction in the number of cells. LGC was used as it is readily accessible and simple to implement. More sophisticated upscaling techniques may be used and would be expected to produce more accurate output relative to the full physics simulation.

3.2.4. Surrogate correction through manifold mapping

The solution of (1) is initially approximated using the surrogate $\tilde{m}(\mathbf{x})$. Thereafter, a sequence of corrections of this surrogate, together with the corresponding optimization runs, are performed using the manifold mapping technique. Manifold mapping (MM; see, e.g., [Echeverría Ciaurri 2007](#)) is a multi level optimization technique that belongs

to a larger family of methods named space mapping (SM; see, e.g., [Bandler et al. 2004](#), [Koziel et al. 2006](#)). In MM (and also in SM) the optimization problem of interest is formulated in terms of the so called fine model. The fine model is assumed to be satisfactorily accurate but computationally expensive. The coarse model, on the other hand, is of much faster evaluation but less precise. With the notation introduced earlier in this section, $\mathbf{m}(\mathbf{x})$ represents the fine model and $\tilde{m}(\mathbf{x})$ the coarse model.

The k th iteration in MM is based on the following corrected model

$$\tilde{m}_k(\mathbf{x}) = \mathbf{m}(\mathbf{x}_k) + S_k (\tilde{m}(\mathbf{x}) - \tilde{m}(\mathbf{x}_k)), \quad (8)$$

where \mathbf{x}_k is the current solution of the optimization problem and S_k is a $p \times p$ matrix that aggregates all components of the response in the correction and is constructed using previous evaluations of the fine and coarse models. Specifically, we define S_k for $k \geq 1$ as

$$S_k = \Delta \mathbf{m}_k \Delta \tilde{m}_k^+, \quad (9)$$

where $\Delta \mathbf{m}_k$ and $\Delta \tilde{m}_k$ are the $p \times \min\{k, n\}$ matrices defined as

$$\Delta \mathbf{m}_k = [\mathbf{m}(\mathbf{x}_k) - \mathbf{m}(\mathbf{x}_{k-1}) \quad \cdots \quad \mathbf{m}(\mathbf{x}_k) - \mathbf{m}(\mathbf{x}_{\max\{k-n, 0\}})], \quad (10)$$

$$\Delta \tilde{m}_k = [\tilde{m}(\mathbf{x}_k) - \tilde{m}(\mathbf{x}_{k-1}) \quad \cdots \quad \tilde{m}(\mathbf{x}_k) - \tilde{m}(\mathbf{x}_{\max\{k-n, 0\}})], \quad (11)$$

with $^+$ being the pseudo inverse operator (see [Golub and Van Loan, 2013](#)).

The next solution \mathbf{x}_{k+1} is obtained by solving the optimization problem formulated for the corrected model $\tilde{m}_k(\mathbf{x})$

$$\mathbf{x}_{k+1} = \min_{\mathbf{x} \in \mathbb{R}^n} U(\tilde{m}_k(\mathbf{x})) \quad \text{subject to} \quad \mathbf{c}(\mathbf{x}) \leq 0. \quad (12)$$

The MM iteration is usually started at $k = 0$. The starting point \mathbf{x}_0 , if not given, could be the optimal solution based on the coarse model $\tilde{m}(\mathbf{x})$ and S_0 is generally taken as the identity matrix. When the coarse model has negligible computational cost, the choice of the initial guess is less relevant, especially if the cost function is not expected to present strong nonlinearities (as is often the case for a coarse model built with quadratic approximation).

In this optimization one can use the same methods as in the optimization based on the original surrogate $\tilde{m}(\mathbf{x})$ since the additional computational cost associated with the correction is in most cases negligible. The correction in (8) can be interpreted as a translation and a rotation of the coarse model response and, upon convergence, yields first order consistency with the fine model response. That is, not only both responses but also their Jacobians coincide in the limit (see [Echeverría Ciaurri 2007](#)). MM and SM are supported by rigorous convergence theory for least squares optimization (see, e.g., [Echeverría Ciaurri 2007](#)) and have been successfully applied to other objective functions such as NPV (see, e.g., [Echeverría Ciaurri and Wagenaar 2016](#), [Echeverría Ciaurri et al. 2021](#)), the one considered in this work. The MM iteration can be stopped using criteria related to the best solution found so far (i.e., when the value of the objective function is acceptable), to the variation between changes in the solution (i.e., when the difference between latest updates for the objective function and/or solution is small) and/or to the number of function evaluations performed (i.e., when the computing budget is exhausted). Different enhancements of the basic MM algorithm have been proposed. For example, in [Hemker and Echeverría \(2007\)](#) the MM iteration includes a trust region strategy to avoid searching regions where the corrected surrogate is not sufficiently accurate, and in [Echeverría \(2007\)](#) the method is extended to deal with constraint functions $\mathbf{c}(\mathbf{x})$ that are expensive to evaluate. We would like to add that MM and SM usually perform satisfactorily with surrogates $\tilde{m}(\mathbf{x})$ that (at first sight) may appear to be inaccurate for practical applications (see, e.g., [Echeverría Ciaurri 2007](#), [Bandler et al. 2004](#), for the use of equivalent circuits as coarse models in optimization problems where the fine models are determined through solution of partial differential equations). The interested reader is referred to [Echeverría Ciaurri \(2007\)](#) and [Echeverría \(2007\)](#) for examples of pseudo codes for MM.

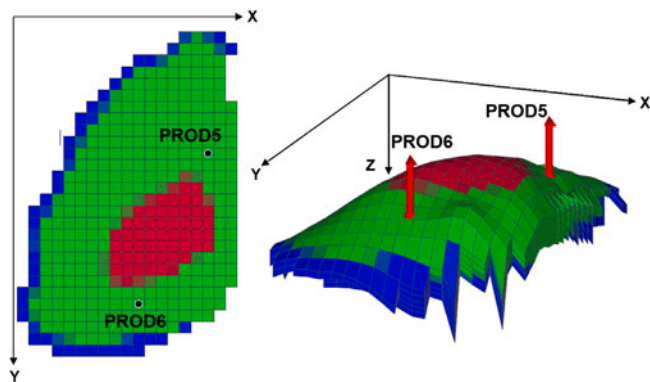


Fig. 2. Top view of the first layer and a three-dimensional representation of the PUNQ-S3 model showing the two preexisting production wells, PROD5 and PROD6, and initial saturation distribution (red, green and blue indicate gas, oil and water, respectively). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

3.3. Constraint handling

General constraints can be handled in most optimization techniques (that includes pattern search algorithms) by means of the filter method, which is essentially an add on (see, e.g., Fletcher and Leyffer 2002, Echeverría Ciaurri et al. 2011a). However, the nonlinear constraints in the problem solved in this work, that is, bounds for the minimum distance between any two wells, unlike, for example, limits for field water production and injection rates, do not require evaluation of computationally expensive functions. For that reason, they can be treated in a somewhat more specific manner than through the filter method. In pattern search optimization we resort to a projection operator to determine feasible alternatives to stencil points that violate constraints for well location and for inter well distance. The projection of an infeasible point onto the feasible domain is formulated as an optimization problem where a distance metric (Euclidean) with respect to the infeasible point is minimized subject to the same constraints as in the original problem. In this work, when used within HJDS the inter well constraint handling is implemented using Sequential Quadratic Programming (see, e.g. Nocedal and Wright 2006) and in MM optimization via an interior point algorithm (see, e.g., Byrd et al. 1999).

4. Results and discussion

In this section we present the results for the application of the aforementioned methods to two realistic, but synthetic, well location optimization problems. The first case investigates the problem of finding the optimal location of four additional production wells in the PUNQ S3 model. The second case involves the placement of five production wells in the presence of five preexisting injection wells for a model based on the Brugge benchmark. We used commercial software (tNavigator) to represent the fine model (i.e., reservoir fluid flow equations) in both cases.

4.1. Example Problem 1: PUNQ S3

The three dimensional PUNQ S3 model (see, e.g., Floris et al. 2001) represents a heterogeneous reservoir with a gas cap and a strong aquifer, which provides bottom water drive and pressure support. The model is discretized into a $17 \times 28 \times 5$ grid, of which 1761 are active cells (i.e., cells that are relevant regarding fluid flow). The reservoir covers an area of approximately 17 km^2 with layer thickness varying between 20 and 30 m. In this example problem we have considered the preexistence of two (fully penetrating, vertical) production

Table 1

Economic parameters for the two example problems.

Parameter	Problem 1	Problem 2
Oil price (r_o)	\$80 USD/bbl	\$40 USD/bbl
Water handling costs (c_{wp})	\$13 USD/bbl	\$5 USD/bbl
Water injection costs (c_{wi})	–	\$5 USD/bbl
Discount rate (b)	0%	2.34%

wells, namely PROD5 and PROD6, which are located close to the gas cap. In Fig. 2 we show a cross section of the top layer alongside a three dimensional representation of the PUNQ S3 model and initial saturation distribution together with the two preexisting wells (red, green and blue indicate gas, oil and water, respectively).

This problem has a total of eight decision variables for the areal location corresponding to the four (fully penetrating, vertical) production wells. Note that in both problems studied in this work the decision variables are normalized and bounded between 0 and 1. The production wells are operated by bottom hole pressure (BHP) with settings of 200 bars (approximately 2900 psi) and a maximum liquid rate of $900 \text{ sm}^3/\text{day}$ (approximately 5660 STB/day). In addition, the production wells are operated in a reactive manner under an economic water cut limit of 85% (i.e., if the water cut of a production well exceeds this limit it is shut in). The distance between wells has to be larger than 300 m. The optimization time frame is equal to 10 years and the economic parameters considered in the objective function are summarized in Table 1.

4.1.1. Surrogate accuracy

In order to assess the accuracy of the surrogates for this example problem, we utilized 1000 random points in the feasible solution space and computed the corresponding responses of the fine and coarse models. As previously mentioned, the model response represents a time series of NPV contributions for different intervals. The two coarse models investigated in this section are based on kriging and quadratic approximation, using 50 sample points obtained with CCD in both cases. The means of the responses associated with the fine and coarse models over the 1000 random points are shown in Fig. 3. In this figure we also include a plot with the root mean square error (RMSE) calculated for the coarse model responses with respect to the fine model responses. The RMSE is calculated using:

$$RMSE = \sqrt{\frac{\sum_{i=1}^N (m(x_i) - \tilde{m}(x_i))^2}{N}}, \quad (13)$$

where, $m(x_i)$ is the reservoir simulation output, $\tilde{m}(x_i)$ is the surrogate output, x_i denotes each of the sample points used, and N is the number of sample points.

The average of the NPV during the 10 year time frame (negative of the objective function) for the fine model responses is 956.5 MM USD whilst for the kriging and quadratic surrogates is 895.8 MM USD and 1073.3 MM USD, respectively. This translates to an average error of 60.7 MM USD and -116.8 MM USD for those approximations, respectively (i.e., a relative error of 6.4% and 12.2%). These results are summarized in Table 2. Although there are some differences between the average fine and coarse responses, especially in the early time intervals, the coarse responses are able to reproduce the general (decreasing) trend seen in the fine model. The coarse models considered may appear relatively inaccurate for modeling purposes but, as can be seen next, yield satisfactory performance when used in optimization.

4.1.2. Optimization results

The optimization based on surrogates and on their correction through MM was solved using an interior point algorithm with stopping criterion given by a threshold equal to 10^{-9} for changes in solution and function value. MM was terminated when a maximum number of 350 reservoir fluid flow simulations, or equivalently, evaluations of the

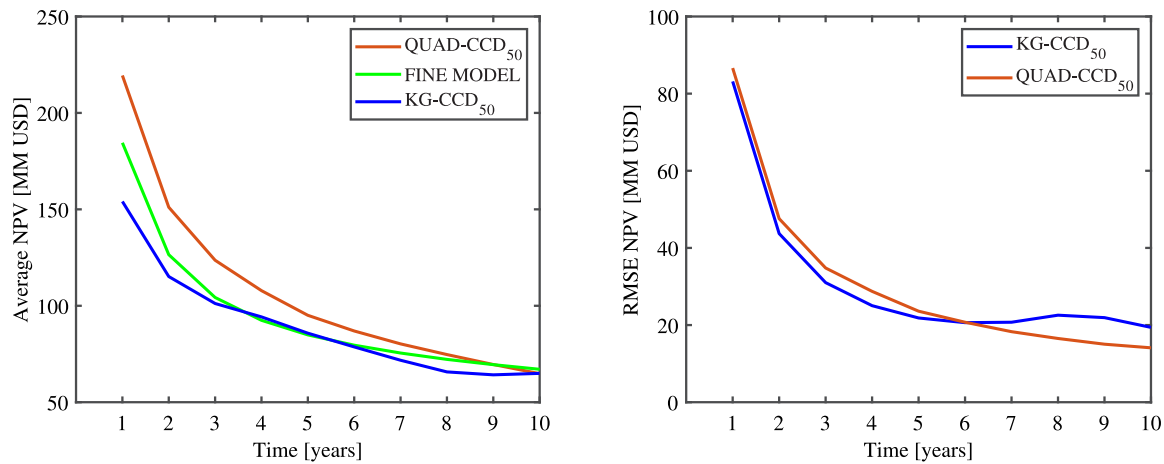


Fig. 3. Left: Average of responses of fine and coarse models for 1000 random points in the feasible solution space. The two coarse models are based on quadratic approximation and kriging, using CCD with 50 sample points. All responses represent time series of NPV contributions during a 10-year production period. Right: RMSE of the two coarse-model responses with respect to the fine-model responses. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

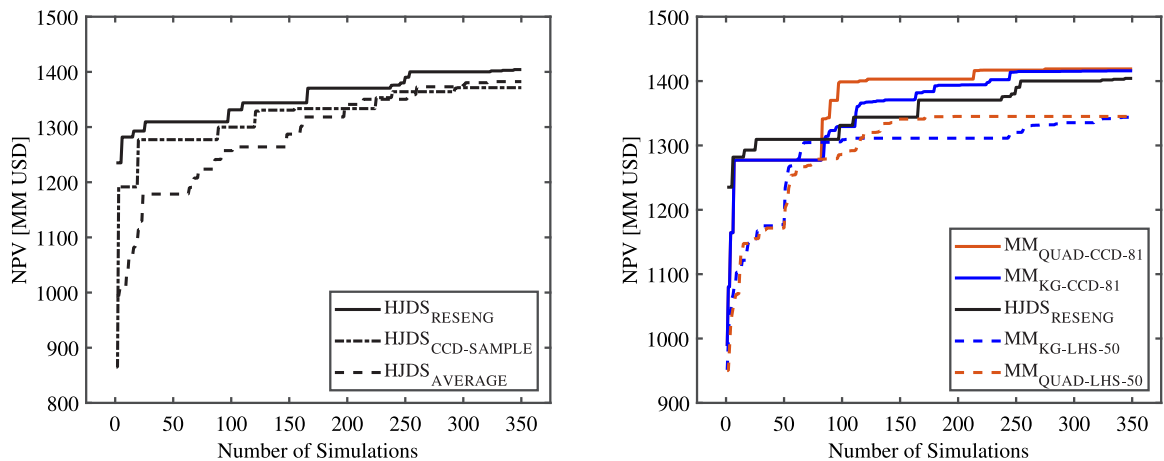


Fig. 4. Left: Optimization results of HJDS for the first example problem using different techniques for the choice of the initial guess: reservoir-engineering considerations, sample in CCD with highest NPV and random selection (in this case, the average of 10 points is shown). Right: Optimization results comparing the performance of HJDS with multiple surrogate-based methods. The surrogates rely on quadratic approximation and kriging, are built on CCD and LHS and are corrected through MM. The experimental-design techniques CCD and LHS have 81 and 50 samples points, respectively. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 2
Comparison of the accuracy for different surrogates used in Example Problem 1.

Model	Average NPV	Relative error
Fine-model (full-physics simulation)	956.5 MM USD	-
Kriging	895.8 MM USD	6.4%
Quadratic approximation	1073.3 MM USD	12.2%

objective function $f(x)$, was reached. Note that in this example the computational cost associated with the surrogates and corresponding correction via MM is negligible when compared with a reservoir simulation. The samples sizes for the experimental design techniques CCD and LHS were 81 and 50, respectively. Tests of LHS with 50 and 81 sample points on a similar optimization problem were inconclusive regarding the choice of the sample size. We opted in this case for 50 sample points in order to provide results for additional algorithmic configurations. Due to the stochastic nature of LHS, the results reported for this technique were determined with the average of 5 optimization runs using different LHS samples.

The surrogate based methods will be compared with HJDS (this latter applied directly to the objective function). The initial stencil size for HJDS was 0.25 (the decision variables are normalized between

0 and 1) and the maximum number of function evaluations (simulation calls) allowed was 350. Given the local optimizer nature of HJDS, two initial guesses were considered. The first initial guess was obtained through reservoir engineering considerations, in particular, by distributing the new (production) wells around the gas oil contact (GOC). This distribution aims at some degree of synchronization of the water breakthrough from the strong aquifer so that oil production below the GOC is maximized. The sample point in CCD with the highest NPV was taken as the second initial guess.

The results of HJDS starting from these two initial guesses are presented in Fig. 4. The initial guess based on reservoir engineering considerations has an NPV equal to 1235 MM USD, which, as one may expect, is a relatively high value. The highest NPV for all the 81 sample points in CCD is 1277 MM USD. However, the selection of the point with the best objective function value within an experimental design procedure requires additional computing cost (81 simulations in the case of CCD studied here). In practical situations under computational constraints, this may not be an efficient method for selecting an initial guess. The initial guess based on reservoir engineering considerations resulted in a higher final NPV (1404 MM USD) than the second initial guess (1371 MM USD). Additional analysis was undertaken by performing the optimization run using 10 random initial guesses. The average

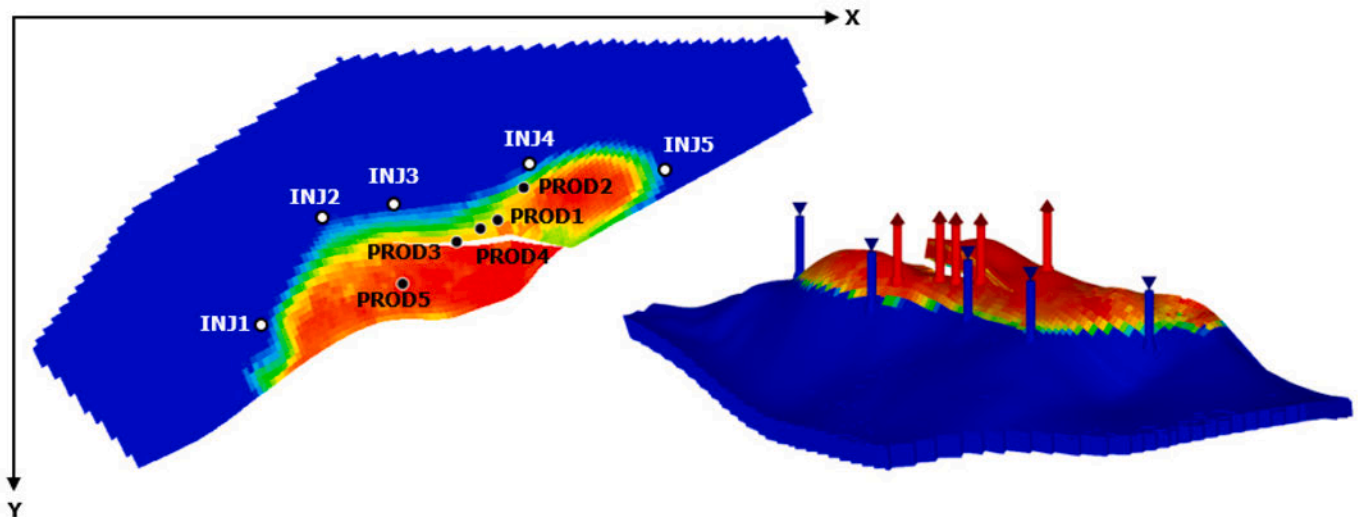


Fig. 5. Top-view and three-dimensional representation for realization 104 selected from the Brugge benchmark and considered in this study. The plots illustrate the average initial oil saturation over the nine reservoir layers (red is high and blue is low oil saturation). The five preexisting injection wells are also included in both representations, together with the five production wells that correspond to the initial guess used for HJDS as stand-alone solver and for the first iteration of MM with LGC (that is, $k = 0$ in MM). The injection and production wells are shown in the three-dimensional representation as blue and red pillars, respectively. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

evolution of the NPV for these 10 runs is shown in Fig. 4. While the optimized NPVs in a number of the 10 runs are larger than the NPV obtained in the run that utilizes reservoir engineering considerations, it still outperforms the average result of the 10 runs. This outperformance is in regard to the final NPV as well as the number of required function evaluations (simulation calls) within the optimization. The optimized NPV in one of the runs was only 1315 MM USD and the standard deviation of the optimized NPVs of the 10 runs is equal to 49.57 MM USD. Although reservoir engineering judgement does not always guarantee selection of an initial guess that yields satisfactory results in terms of optimized objective function, we can expect that in many situations this would be case. In the rest of the paper reservoir engineer considerations will be used to determine initial guesses for HJDS.

Fig. 4 also illustrates the comparison of surrogate based methods with HJDS. Please note that the presented NPVs have been computed using the fine model (i.e., reservoir simulator) The simulations needed to construct each surrogate are included in each plot. The optimization runs with the two surrogates built upon LHS and corrected via MM yield similar NPVs after 350 function evaluations. The optimized NPVs for quadratic approximation and kriging are 1345 MM USD and 1344 MM USD, respectively (as indicated above, the LHS results are an average of 5 runs). In comparison, the surrogates based on CCD outperformed those based on LHS regarding final optimized value and most intermediate NPVs during the run. For example, quadratic approximation and kriging, both using CCD within the MM framework, reach an NPV equal to 1400 MM USD after 115 and 225 function evaluations, respectively. On the other hand, HJDS requires 254 function evaluations to achieve this objective function value. This translates to a 55% (12%) reduction in computational cost with respect to HJDS when a surrogate based on quadratic approximation (kriging) is combined with MM. The final values for the optimization runs using CCD with quadratic approximation and kriging are 1419 MM USD and 1416 MM USD, respectively (remember that the optimized NPV for HJDS with an initial guess determined through reservoir engineering considerations was 1404 MM USD).

We also tested a surrogate based on generalized barycentric coordinates (see, e.g. Floater 2015). This type of surrogate relies strongly on linearity assumptions and has been observed to deal successfully with well control optimization (see, e.g., Echeverría Ciaurri and Wagenaar 2016). However, that surrogate did not perform satisfactorily in this case example, possibly due to the strong nonlinear character of the optimization problem.

4.2. Example problem 2: Brugge

The second example problem utilizes the synthetic Brugge benchmark (see, e.g., Peters et al. 2010). This benchmark was constructed as an analogue of the typical reservoir found in the North Sea. The three dimensional model is discretized by means of a $139 \times 48 \times 9$ grid, of which 44,550 cells are active, and describes an oil water system where geological uncertainty is quantified through 104 realizations. The areal dimensions of the field are approximately 10 by 3 km and each grid block in the oil bearing zone is around 120 by 100 by 7 m. In our optimization problem we consider geological realization 104, which was arbitrarily selected, and five (fully penetrating, vertical) initial peripheral injectors. We illustrate in Fig. 5 a top view and a three dimensional representation of the selected Brugge realization showing the five preexisting injectors and the average initial oil saturation across the nine reservoir layers (red and blue indicate oil and water, respectively).

In this problem we seek to determine the optimal areal locations of five (fully penetrating, vertical) production wells (i.e., there are 10 decision variables) during a production time frame of 30 years. As a reminder, the decision variables are normalized and bounded between 0 and 1. The injection wells were controlled by BHP, which was set to 180 bars (approximately 2610 psi), subject to a maximum water injection rate of $636 \text{ m}^3/\text{day}$ (4000 STB/day). The production wells were also controlled by BHP, which was set to 50 bars (approximately 725 psi), with a maximum liquid production rate of $477 \text{ m}^3/\text{day}$ (3000 STB/day). Additionally, the distance between any two wells cannot be smaller than 200 m. The economic parameters in the optimization are summarized in Table 1.

4.2.1. Surrogate accuracy

We utilized 100 random points in the feasible solution space and computed the responses of the fine and three different coarse models in order to investigate the accuracy of the surrogates in this second problem. Fig. 6 shows the responses averaged over the 100 points. The two analytical surrogates, kriging and quadratic approximation, were built using CDD with 100 sample points. We also considered LGC as the basis for a physics based surrogate. Note that this surrogate does not need fine model simulations for its construction and that its run time is around 10 times faster than the fine model response (remember

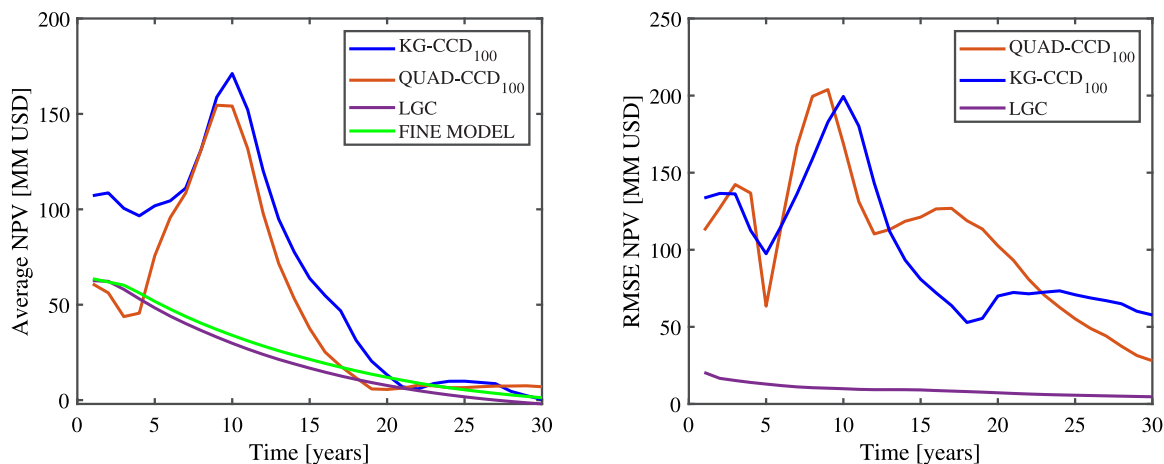


Fig. 6. Left: Average of responses of fine and coarse models for 100 random points in the feasible solution space. The three coarse models are based on quadratic approximation and kriging, using CCD with 100 sample points, and on LGC. All responses represent time series of NPV contributions during a 30-year production period. Right: RMSE for the three coarse-model responses with respect to the fine-model responses. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

that the analytical surrogates, once they are built, have a negligible computational cost when compared to the fine model). The difference in quality between the analytical and the physics based approximations is apparent in the figure. More importantly, the trend in the response of the coarse model based on LGC is similar to the trend in the fine model response. This behavior may be crucial for guiding the optimization effectively (as can be seen later in the results).

As in the previous example problem, we also illustrate the RMSE that corresponds to each of the three surrogates relative to the fine model response (Fig. 6). The RMSEs associated with the analytical surrogates present clear oscillations. This behavior may be indicative that the number of sample points in the design of experiments is not large enough to sample the search space adequately. The average of the NPV during the 30 year time frame for the fine model responses is 757 MM USD whilst it is 1933.6 MM USD, 1455.4 MM USD and 646.9 MM USD for kriging, quadratic approximation and LGC, respectively. The relative error in the NPV during the 30 year time frame averaged over the 100 random points for the surrogates based on LGC, quadratic approximation and kriging is 14.5%, 92.3% and 155.4%, respectively. These results are summarized in Table 3.

The relative error observed for the analytical surrogates in this problem is noticeably larger than in the PUNQ S3 model. The manifest discrepancy shown in Fig. 6 anticipates that the performance of these surrogates in the optimization problem will be unsatisfactory. The search space has now only two additional dimensions with respect to the first problem and the number of sample points in the second case has just doubled. Although the accuracy of analytical surrogates, in general, can be improved with an increase in the number of sample points used, the results obtained may indicate that this number still needs to grow further (remember the existing relationship between global exploration and the curse of dimensionality). Since the number of points sampled in this problem is already equal to 100, it is unclear how effective analytical surrogates can be in practical problems, which very likely may have more than 10 optimization variables and, not infrequently, computing budgets of few hundred simulations. The performance deterioration (from the first to the second problem) of the analytical surrogates tested can be also attributed to the difference in complexity of the problems and of the underlying physical phenomena modeled. Although, as mentioned above, the second problem has slightly more optimization variables, it also presents additional changes with respect to the first problem such as a different geological heterogeneity and the use of secondary recovery through water injection wells. Physics based surrogates, in principle, are inherently more suited to model these phenomena more precisely (usually at a higher computational cost).

Table 3

Comparison of the accuracy for different surrogates used in Example Problem 2.

Model	Average NPV	Relative error
Fine-model (full-physics simulation)	757.0 MM USD	–
Kriging	1933.6 MM USD	155.4%
Quadratic approximation	1455.4 MM USD	92.3%
Local grid-coarsening (LGC)	646.9 MM USD	14.5%

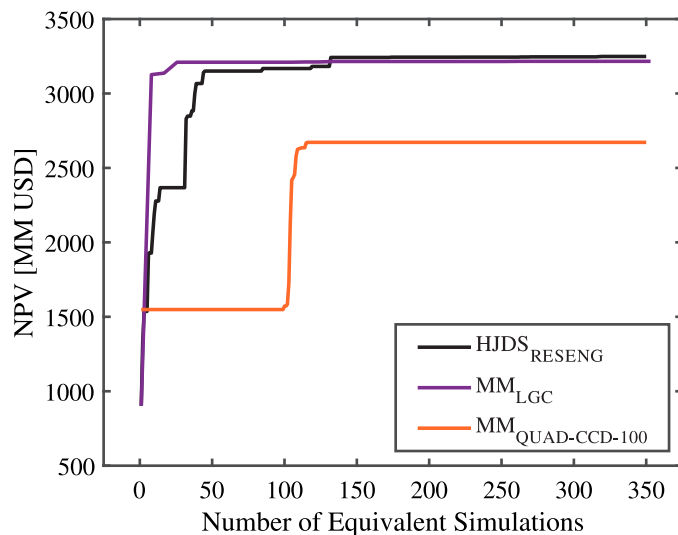


Fig. 7. Optimization results comparing the performance of HJDS with two surrogate-based methods for the second example problem. One surrogate relies on quadratic approximation that is built on CCD using 100 sample points and is corrected through MM. The second surrogate relies on LGC and is also corrected by means of MM.

4.2.2. Optimization results

The optimization based on the quadratic approximation surrogate and its iterative correction was handled, as in the previous problem, with an interior point algorithm. The stopping criterion used was again determined by means of a threshold equal to 10^{-9} for changes in solution and function value. The sample size for CCD was 100 sample points. Since LGC does not have negligible computational cost when compared to the fine model response and the use of numerical gradients may be potentially problematic (e.g., selection of perturbation size in the numerical derivative), the optimization runs based on that surrogate and its corrections through MM were solved using HJDS, a

derivative free method. In this case the initial stencil size for HJDS was equal to 0.25 (remember that the decision variables are normalized between 0 and 1) and the stopping criterion was a maximum of 50 function evaluations, that is, 50 executions of the LGC surrogate. Note that the LGC surrogate did not require any simulations to construct. However, the computational cost required to run this surrogate is accounted for in the results. To this end, equivalent function evaluations (simulation calls) will be used: each 10 evaluations of LGC response are counted as one evaluation of the fine model response.

The initial guess for the first iteration in MM with LGC as coarse model, that is $k = 0$ in MM, was determined through reservoir engineering considerations because the results obtained in the first example problem with this type of starting point were satisfactory. These considerations aim now to achieve pressure support at the production wells through a low elevation on the dome structure and their relative proximity to the existing injectors (see Fig. 5). For any subsequent iteration in MM, the initial guess in (12) was x_k . The MM algorithm was terminated when a maximum number of 350 equivalent simulations was reached. This criterion, as we will see later, may not be practical here because it is translated to a large number of simulations without significant change in the objective function. Nevertheless, it provides some information regarding asymptotic convergence that can be useful in a comparative study as this one.

The surrogate based methods will be once more compared with HJDS. The initial stencil size was taken equal to 0.25 and the starting point was the same one taken in the first iteration for MM with LGC and illustrated in Fig. 5. The maximum allowable number of function evaluations (simulation calls) was 350.

The comparison of the performance between HJDS and the surrogate based methods is represented in Fig. 7. Please note that the presented NPVs have been determined using the fine model (i.e., reservoir simulator). The simulations needed to build the quadratic surrogate are included in the plot. The NPV found with HJDS after the 350 equivalent simulations was 3249 MM USD. MM based on LGC reached a comparable value, 3216 MM USD, also after 350 equivalent simulations. The optimization solved at the first iteration of MM combined with LGC, i.e., $k = 0$ in MM, yielded an NPV equal to 3126 MM USD and required 8 equivalent simulation calls. The MM iteration is able to rapidly increase that value further afterward. We would like to stress that in spite of LGC being a relatively coarse surrogate (it relies heavily on averages of reservoir properties), the optimization results, what matters in the end, are acceptable. On the other hand, the NPV obtained with MM based on quadratic approximation after 350 equivalent simulations was only 2672 MM USD. While MM provides certain improvement in this latter case as well, the solution obtained has clearly lower objective function value than the other two runs. This is in line with the results shown in Fig. 6, which may indicate the inability of the quadratic approximation to capture general trends in the search space for this problem.

Fig. 8 shows the well locations of the optimized solutions for HJDS (left) and LGC using MM (right). The similarities in the solutions give an indication that there is no disadvantage when using LGC with MM. Interestingly, in both solutions a well (PROD3) is placed to target the southern region while two wells (PROD1 and PROD5) are also located at the top of the anticline structure (see Fig. 5). Note that these wells adhere to the inter well distance constraint. The major difference between the solutions is that in LGC a well (PROD2) is placed at the dome structure on the northern side of the fault while in HJDS the well (PROD2) is closer to the injectors.

We reiterate that the stopping criterion for MM is not practical and that it was selected to make better comparisons between the methods. In real applications this criterion should be adjusted in accordance with the available computing resources and include additional metrics (e.g., based on changes in the solution and objective function). The initial guess determined by means of reservoir engineering considerations, unlike in the first example problem, does not improve the sampling

performed by the DoE technique. We can expect that in real life scenarios the use of experience and domain knowledge for determining a baseline or initial guess is, in general, beneficial. However, there may be situations where systematic analysis, such as the one performed by DoE, can be advantageous.

Although the solution eventually determined by HJDS has slightly higher NPV than the one found with MM combined with LGC, during the first part of the run the surrogate based approach shows better performance. For example, the optimization based on MM with LGC reaches an NPV equal to 3210 MM USD after only 26 equivalent function evaluations. In order to achieve a similar objective function value, HJDS required approximately 132 equivalent function evaluations. This translates to roughly an 80% reduction in computational cost when MM is combined with LGC regarding HJDS.

We would like to add that optimization tests were also undertaken for quadratic approximation using a sample size of 50 both for CCD and LHS. Given the low performance observed regarding final optimized NPV and for succinctness, we opt for not reporting these tests here. One may increase the number of simulation runs used in DoE but, judging from the results analyzed above, it is unclear if the surrogate constructed upon these runs would be competitive with respect to HJDS or MM based on LGC.

4.3. Discussion

The first optimization problem showcases the potential applicability of analytical surrogates to well placement problems and FDP. The results when using analytical surrogates, namely polynomial approximation and kriging, for a case involving a few optimization variables and a relatively simple physical phenomenon (e.g., reservoir geology without complex features and primary hydrocarbon recovery) give an indication that these surrogates are able to approximate the general trends seen in the optimization space and yield reasonably precise solutions. The first problem also showed us that, when the analytical surrogates exhibit acceptable accuracy, local improvements and corrections may improve performance. In particular, we employed corrections based on MM, which aimed at first order consistency of the coarse model response with the fine model response.

In the second example problem we saw that a surrogate based on polynomial approximation and iteratively corrected via MM clearly underperformed stand alone HJDS and LGC combined also with MM. This case was somewhat more involved than the first optimization problem: there was one more well (areal) location to optimize, the reservoir geology was to a certain extent realistic and production was accomplished through water injection, that is, a secondary recovery mechanism. It may be argued that the quality of an analytical surrogate, in principle, improves with additional sample points. This strategy, besides defeating its purpose in time constrained scenarios, may deteriorate performance in the optimization when compared to methods that avoid a stage to construct a surrogate by performing a number of simulation runs. Although alternative analytical surrogates could be considered in this problem, we can expect that the observed underperformance may very likely reappear in many other practical problems that have more than just five (vertical) wells to optimize, and, as a consequence, much more complex search spaces. This can be linked to the curse of dimensionality as its effects take a more dominant role when the number of dimensions of a space grows. The question remains how one can explore (relatively) large optimization spaces effectively and efficiently when no information about the behavior of a noticeably nonlinear objective function in these spaces is available.

The analytical surrogates in this study were built upon few simulation runs (one hundred at most) in an attempt to model practical scenarios where solutions are required in a very short time frame. Moreover, in some real life situations, not only rapid turnaround is required but also, to exacerbate the problem, computational resources, this includes licenses for commercial simulators, are scarce or have

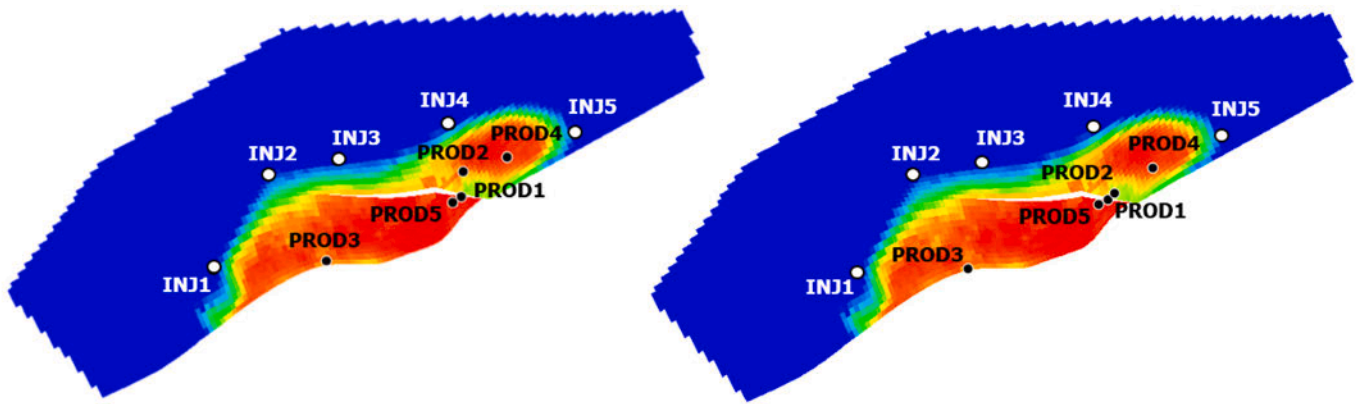


Fig. 8. Left: Well locations of the optimized solution found with HJDS. Right: Well locations of the optimized solution found when using LGC in combination with MM. Both figures show the average initial oil saturation over the nine reservoir layers (red is high and blue is low oil saturation). The producers are represented by black circles and the injectors are represented by white circles. Note: wells adhere to the minimum inter-well constraint. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

to be shared. These constraints preclude the use in these scenarios of surrogates that leverage immense sets of previous simulation runs. This is the case, in particular, of approximations based on the recently popularized ANNs. The evaluation of the training sets required by most ANNs may entail, in the context of well placement and field development, execution of a prohibitively large number of simulations (many more than the numbers considered in this work). For example, in the work described in De Paola et al. (2020), the network training for a field development problem regarding only two additional new wells needs around 10,000 simulations and the network constructed yields relatively high error for reservoir engineering relevant magnitudes. Nevertheless, in environments where resources are bountiful, ANNs may be a promising option to approach surrogate based optimization. We would like to add that some other problems in FDP optimization are less sensitive to adverse consequences of the curse of dimensionality because the objective functions addressed are very often relatively smooth. In that case, the construction of sufficiently precise surrogates may require moderate and practical numbers of simulation runs. A conspicuous example of one such problem is the optimization of well controls, which has already been addressed successfully by means of similar analytical surrogates to the ones applied here (see, e.g., Echeverría Ciaurri et al. 2021).

Physics based surrogates, as exemplified in the second optimization problem in this study, incorporate domain knowledge and problem physics, and could be an alternative in some real world scenarios. The results from that problem show the potential upside of utilizing a two dimensional reservoir model approximation as physics based surrogate for well placement. Physics based surrogates are expected to be more precise than analytical surrogates and typically yield better optimization outcome. However, physics based surrogates also pose challenges. The implementations of such surrogates are frequently more elaborate and require a more in depth understanding of the underlying physics and associated numerical computation. As a consequence, there is a trade off that needs to be considered when applying surrogate based optimization (to, in this case, well placement and field development problems). On the one hand, we have the analytical surrogates, many of them of straightforward construction but sometimes unacceptably imprecise. Note that success in this optimization context is measured in being able to yield a satisfactory solution to the problem. On the other hand, there are physics based surrogates, nontrivial, more time consuming (both in development and run time) and usually providing higher levels of accuracy.

In view of this trade off, we may opt to employ a general strategy to (attempt to) predict the suitability of using an analytical surrogate (built on a limited number of simulations). If the surrogate is not deemed satisfactorily precise, one could resort to incorporating domain

knowledge and experience through some physics based surrogate. A metric that gives insight into the complexity of the model or of the problem and that is determined running few simulations (consistent with the computational budget) could be used to estimate the suitability of an analytical surrogate. Another procedure to validate the surrogate accuracy may include a training and testing setup.

5. Concluding remarks

In this paper we studied the application of surrogate based optimization to the well placement problem within hydrocarbon field development planning (FDP). The problem was contextualized by considering practical scenarios under tight computational budget (of at most few hundred reservoir flow simulations). The employed methodology is noninvasive regarding the simulator (i.e., derivative free) and is combined with a surrogate to allow for efficient optimization. Two main categories of surrogates, analytical and physics based, have been investigated. The analytical surrogates included quadratic approximation and kriging, both of them relying on two types of design of experiments: central composite design and Latin hypercube sampling. The physics based surrogate used is a two dimensional reservoir approximation obtained through local grid coarsening (LGC). The surrogate in all cases is iteratively and locally corrected with manifold mapping (MM), a multi level optimization technique, to improve the current best solution. The results from this work show the first application of MM to well placement optimization problems.

The methodology was tested on two example problems: (1) the placement of four production wells in the presence of two other preexisting production wells in the PUNQ S3 reservoir model; and (2) the placement of five production wells in the presence of five water injection wells in a somewhat more realistic model based on the Brugge benchmark. The results of the proposed methodology were compared against those obtained by Hooke Jeeves direct search (HJDS), a derivative free optimization method without any explicit surrogate. It is worth mentioning that HJDS had performed successfully in other FDP problems and is especially suited to scenarios with modest computing resources. For the first example problem, the two analytical surrogates, quadratic approximation and kriging, were compared with HJDS. The use of quadratic approximation and kriging translated to a reduction in computing time of 55% and 12%, respectively, with respect to HJDS to obtain an acceptable solution. The second example problem, with a more complex search space due to the recovery mechanism considered (i.e., waterflooding) and increased reservoir heterogeneity, demonstrated the potential limitations of analytical surrogates for well placement problems in time constrained decision making scenarios. Quadratic approximation, which provided satisfactory results in the

first example, clearly underperformed HJDS. The optimization using a physics based surrogate constructed upon LGC resulted in a reduction of 80% in computational cost compared with HJDS to reach a reasonably adequate solution.

In addition, the investigations into surrogate accuracy (for both problems) gave insight into the importance of capturing the general trends of the fine model especially when used within an optimization framework. The results show that when the fidelity of the surrogate is reasonably high, MM was able to bring improvement. This enables MM to obtain solutions of similar accuracy to those determined by means of a single level technique and with significant computational savings. More specifically, the accuracy of the analytical surrogates in the first example problem was sufficient to provide improvements with the use of MM. However, this was not the case for the second example problem, where the analytical surrogates were unable to help in the optimization. This may be associated with their inability to capture the general trends seen in the fine model response. The performance in the optimization clearly improved when a physics based surrogate was used as an alternative.

All in all, the findings in this study give valuable insights into the applicability and shortcomings of analytical surrogates and the tradeoff associated with alternatives, such as physics based surrogates. The problems that can be found in real life well placement applications are harder than those analyzed here, for example, they require positioning of many more wells, hundreds in large hydrocarbon fields, so we can expect that the issues identified here will emerge, magnified in some cases, in those significantly larger optimization spaces.

The research in this study could be continued in multiple directions. The surrogates used in the two examples were built on aggregate model responses, in particular, time series of NPV contributions corresponding to the production of the entire hydrocarbon field. We could expect that a more granular approach to surrogate modeling, for example, based on responses for individual well injection and production rates, may improve the quality not only of the overall approximation but also of the MM correction due to the larger amount of information involved. Other physics based surrogates, such as the ones indicated in the Introduction, can be considered. The methodology may be enhanced by incorporating ways to deal with uncertainty, for instance, through the retrospective optimization framework introduced in Wang et al. (2012) or the use of the adaptive rank based subset selection technique presented in Arouri et al. (2022a). Typically, when incorporating geological uncertainty, the objective function becomes the expected value of NPV (for a risk neutral decision maker). Possible implementations could include an analytical surrogate that approximates expected NPV as a function of time or a physics based surrogate built for each realization separately. Further tests have to be performed for more complex reservoir models and optimization problems. In this regard, addressing generalized field development, that is, extending the formulation of the well placement problem with other decision variables relevant in FDP, such as the number of wells, their type, respective controls and drilling times, may be especially practical. Finally, as indicated above, a metric that, consistent with the computational resources available, aims to predict the suitability of a given analytical surrogate will certainly be very useful in order to help decide if dedicating time to develop an alternative physics based surrogate is a worthy enterprise.

CRediT authorship contribution statement

Yazan Arouri: Conceptualization, Methodology, Software, Formal analysis, Investigation, Data curation, Writing original draft. **David Echeverría Ciaurri:** Conceptualization, Methodology, Investigation, Software, Validation, Writing review & editing. **Mohammad Sayyafzadeh:** Resources, Validation, Methodology, Writing review & editing, Supervision.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The authors would like to thank Rock Flow Dynamics for providing the academic licenses for tNavigator. This research did not receive any specific grant from funding agencies in the public, commercial, or not for profit sectors.

References

- Alavian, S., Whitson, C., Martinsen, S., 2014. Global component lumping for EOS calculations. In: SPE Annual Technical Conference and Exhibition. <http://dx.doi.org/10.2118/170912-MS>.
- Alghareeb, Z.M., 2015. Optimal Reservoir Management using Adaptive Reduced-Order Models (Ph.D. thesis). Massachusetts Institute of Technology.
- Alrashdi, Z., Sayyafzadeh, M., 2019. $(\mu + \lambda)$ Evolution strategy algorithm in well placement, trajectory, control and joint optimisation. *J. Pet. Sci. Eng.* 177, 1042–1058. <http://dx.doi.org/10.1016/j.petrol.2019.02.047>, URL: <http://www.sciencedirect.com/science/article/pii/S0920410519301846>.
- Arouri, Y., Lake, L.W., Sayyafzadeh, M., 2022b. Bilevel optimization of well placement and control settings assisted by capacitance-resistance models. *SPE J.* 1–20. <http://dx.doi.org/10.2118/210562-PA>.
- Arouri, Y., Sayyafzadeh, M., 2020. An accelerated gradient algorithm for well control optimization. *J. Pet. Sci. Eng.* 190, 106872. <http://dx.doi.org/10.1016/j.petrol.2019.106872>, URL: <http://www.sciencedirect.com/science/article/pii/S0920410519312884>.
- Arouri, Y., Sayyafzadeh, M., 2022. An adaptive moment estimation framework for well placement optimization. *Comput. Geosci.* <http://dx.doi.org/10.1007/s10596-022-10135-9>.
- Arouri, Y., Sayyafzadeh, M., Begg, S., 2022a. Adaptive rank-based selection of geological realizations for optimum field development planning. *SPE J.* 1–18. <http://dx.doi.org/10.2118/209584-PA>.
- Artus, V., Durlafsky, L.J., Onwunali, J., Aziz, K., 2006. Optimization of nonconventional wells under uncertainty using statistical proxies. *Comput. Geosci.* 10 (4), 389–404. <http://dx.doi.org/10.1007/s10596-006-9031-9>.
- Audet, C., Dennis, J.E., 2006. Mesh adaptive direct search algorithms for constrained optimization. *SIAM J. Optim.* 17 (1), 188–217.
- Bandler, J.W., Cheng, Q.S., Dakrouy, S.A., Mohamed, A.S., Bakr, M.H., Madsen, K., Søndergaard, J., 2004. Space mapping: The state of the art. *IEEE Trans. Microw. Theory Tech.* 52 (1), 337–361.
- Bangerth, W., Klie, H., Wheeler, M.F., Stoffa, P.L., Sen, M.K., 2006. On optimization algorithms for the reservoir oil well placement problem. *Comput. Geosci.* 10 (3), 303–319. <http://dx.doi.org/10.1007/s10596-006-9025-7>.
- Batycky, R.P., Blunt, M.J., Thiele, M.R., 1997. A 3D field-scale streamline-based reservoir simulator. *SPE Reserv. Eng.* 12 (04), 246–254. <http://dx.doi.org/10.2118/36726-PA>.
- Bellout, M.C., Echeverría Ciaurri, D., Durlafsky, L.J., Foss, B., Kleppe, J., 2012. Joint optimization of oil well placement and controls. *Comput. Geosci.* 16 (4), 1061–1079.
- Bittencourt, A.C., Horne, R.N., 1997. Reservoir development and design optimization. In: SPE Annual Technical Conference and Exhibition. <http://dx.doi.org/10.2118/38895-MS>.
- Bouzarkouna, Z., Ding, D.Y., Auger, A., 2012. Well placement optimization with the covariance matrix adaptation evolution strategy and meta-models. *Comput. Geosci.* 16 (1), 75–92.
- Buckley, S., Leverett, M., 1941. Mechanism of fluid displacement in sands. *Trans. AIME* 146, <http://dx.doi.org/10.2118/942107-G>.
- Byrd, R.H., Hribar, M.E., Nocedal, J., 1999. An interior point algorithm for large-scale nonlinear programming. *SIAM J. Optim.* 9 (4), 877–900.
- Cardoso, M.A., Durlafsky, L.J., Sarma, P., 2009. Development and application of reduced-order modeling procedures for subsurface flow simulation. *Internat. J. Numer. Methods Engrg.* 77 (9), 1322–1350. <http://dx.doi.org/10.1002/nme.2453>, URL: <https://onlinelibrary.wiley.com/doi/abs/10.1002/nme.2453>, arXiv:<https://onlinelibrary.wiley.com/doi/pdf/10.1002/nme.2453>.
- Conn, A.R., Scheinberg, K., Vicente, L.N., 2009. Introduction to Derivative-Free Optimization. Society for Industrial and Applied Mathematics.
- Cullick, A.S., Johnson, W.D., Shi, G., 2006. Improved and more-rapid history matching with a nonlinear proxy and global optimization. In: SPE Annual Technical Conference and Exhibition. p. 13. <http://dx.doi.org/10.2118/101933-MS>.
- Datta-Gupta, A., King, M., 2007. Streamline Simulation: Theory and Practice. SPE TextBook Series, Society of Petroleum Engineers.

- De Paola, G., Ibanez-Llano, C., Rios, J., Kollias, G., 2020. Reinforcement learning for field development policy optimization. In: SPE Annual Technical Conference and Exhibition.
- Ding, Y., 2008. Optimization of well placement using evolutionary methods. In: EAGE Conference and Exhibition. European Association of Geoscientists & Engineers, <http://dx.doi.org/10.2118/113525-MS>.
- Ding, S., Jiang, H., Li, J., Tang, G., 2014. Optimization of well placement by combination of a modified particle swarm optimization algorithm and quality map method. *Comput. Geosci.* 18 (5), 747–762. <http://dx.doi.org/10.1007/s10596-014-9422-2>.
- van Doren, J.F.M., Markovinić, R., Jansen, J.-D., 2006. Reduced-order optimal control of water flooding using proper orthogonal decomposition. *Comput. Geosci.* 10 (1), 137–158. <http://dx.doi.org/10.1007/s10596-005-9014-2>.
- Echeverría, D., 2007. Two new variants of the manifold-mapping technique. *COMPEL - Int. J. Comput. Math. Electr. Electron. Eng.* 26 (2), 334–344.
- Echeverría, D., Hemker, P.W., 2005. Space mapping and defect correction. *Comput. Methods Appl. Math.* 5 (2), 107–136.
- Echeverría Ciaurri, D., 2007. Multi-Level Optimization: Space Mapping and Manifold Mapping (Ph.D. thesis). University of Amsterdam.
- Echeverría Ciaurri, D., Conn, A.R., Mello, U.T., Onwunali, J.E., 2012. Integrating mathematical optimization and decision making in intelligent fields. In: Paper SPE 149780 Presented at the SPE Intelligent Energy Conference and Exhibition, Utrecht, the Netherlands, 27–29 March.
- Echeverría Ciaurri, D., Isebor, O.J., Durlafsky, L.J., 2011a. Application of derivative-free methodologies to generally constrained oil production optimisation problems. *Int. J. Math. Modelling Numer. Optim.* 2 (2), 134–161.
- Echeverría Ciaurri, D., Moreno Beltrán, G.A., Camacho Navarro, J., Prada Mejía, J.A., 2021. Improved well-control progressive optimization with generalized barycentric coordinates and manifold mapping. *SPE Reserv. Eval. Eng.* 24 (04), 940–951. <http://dx.doi.org/10.2118/206718-PA>.
- Echeverría Ciaurri, D., Mukerji, T., Durlafsky, L.J., 2011b. Derivative-free optimization for oil field operations. In: Yang, X.S., Koziel, S. (Eds.), *Computational Optimization and Applications in Engineering and Industry*. Springer, pp. 19–55.
- Echeverría Ciaurri, D., Wagenaar, C., 2016. Manifold-mapping optimization applied to oil field operations. In: Paper P92 Presented at the 15th European Conference on the Mathematics of Oil Recovery, Amsterdam, the Netherlands, 29 August – 1 September.
- Emerick, A.A., Silva, E., Messer, B., Almeida, L.F., Szwarcman, D., Pacheco, M.A.C., Vellasco, M.M.B.R., 2009. Well placement optimization using a genetic algorithm with nonlinear constraints. In: SPE Reservoir Simulation Symposium. Society of Petroleum Engineers, <http://dx.doi.org/10.2118/118808-MS>.
- Fletcher, R., Leyffer, S., 2002. Nonlinear programming without a penalty function. *Math. Program.* 91 (2), 239–269.
- Floater, M., 2015. Generalized barycentric coordinates and applications. *Acta Numer.* 24, 161–214.
- Floris, F.J.T., Bush, M.D., Cuypers, M., Roggero, F., Syversveen, A.-R., 2001. Methods for quantifying the uncertainty of production forecasts: A comparative study. *Pet. Geosci.* 7 (S), S87–S96. <http://dx.doi.org/10.1144/petgeo.7.S.S87>, URL: <https://www.earthdoc.org/content/journals/10.1144/petgeo.7.S.S87>.
- Fonseca, R.M., Rossa, E.D., Emerick, A.A., Hanea, R.G., Jansen, J.D., 2020. Introduction to the special issue: Overview of OLYMPUS optimization benchmark challenge. *Comput. Geosci.* 24 (6), 1933–1941. <http://dx.doi.org/10.1007/s10596-020-10003-4>.
- Golub, G.H., Van Loan, C.F., 2013. *Matrix Computations*, fourth ed. The Johns Hopkins University Press.
- Guyaguler, B., Horne, R.N., 2001. Uncertainty assessment of well placement optimization. In: SPE Annual Technical Conference and Exhibition. Society of Petroleum Engineers, <http://dx.doi.org/10.2118/71625-MS>.
- Guyaguler, B., Horne, R.N., Rogers, L., Rosenzweig, J.J., 2000. Optimization of well placement in a Gulf of Mexico waterflooding project. In: SPE Annual Technical Conference and Exhibition. Society of Petroleum Engineers, <http://dx.doi.org/10.2118/63221-MS>.
- Hemker, P.W., Echeverría, D., 2007. A trust-region strategy for manifold mapping optimization. *J. Comput. Phys.* 224 (1), 464–475.
- Hooke, R., Jeeves, T.A., 1961. Direct search solution of numerical and statistical problems. *J. ACM* 8 (2), 212–229.
- Isebor, O.J., Durlafsky, L.J., Echeverría Ciaurri, D., 2014a. A derivative-free methodology with local and global search for the constrained joint optimization of well locations and controls. *Comput. Geosci.* 18 (3), 463–482. <http://dx.doi.org/10.1007/s10596-013-9383-x>.
- Isebor, O.J., Echeverría Ciaurri, D., Durlafsky, L.J., 2014b. Generalized field-development optimization with derivative-free procedures. *SPE J.* 19 (5), 891–908.
- Jansen, J.D., 2011. Adjoint-based optimization of multi-phase flow through porous media – A review. *Comput. & Fluids* 46 (1), 40–51. <http://dx.doi.org/10.1016/j.compfluid.2010.09.039>, URL: <http://www.sciencedirect.com/science/article/pii/S0045793010002677>.
- Jansen, J.D., Durlafsky, L.J., 2017. Use of reduced-order models in well control optimization. *Opt. Eng.* 18 (1), 105–132. <http://dx.doi.org/10.1007/s11081-016-9313-6>.
- Jesmani, M., Jafarpour, B., Bellout, M.C., Hanea, R.G., Foss, B., 2016. Application of simultaneous perturbation stochastic approximation to well placement optimization under uncertainty. In: ECMOR XV-15th European Conference on the Mathematics of Oil Recovery. European Association of Geoscientists & Engineers, pp. cp-494–00133.
- Journel, A.G., Huijbregts, C.J., 1981. *Mining Geostatistics*. Academic Press, London.
- Khuri, A.I., Cornell, J.A., 1996. *Response Surfaces: Designs and Analyses*. CRC Press.
- Kleijnen, J.P.C., 2009. Kriging metamodeling in simulation: A review. *European J. Oper. Res.* 192 (3), 707–716. <http://dx.doi.org/10.1016/j.ejor.2007.10.013>, URL: <http://www.sciencedirect.com/science/article/pii/S0377221707010090>.
- Koehler, J.R., Owen, A.B., 1996. *Computer experiments*. In: Ghosh, S., Rao, C.R. (Eds.), *Handbook of Statistics*. Elsevier Science B.V., pp. 261–308.
- Kolda, T.G., Lewis, R.M., Torczon, V., 2003. Optimization by direct search: New perspectives on some classical and modern methods. *SIAM Rev.* 45 (3), 385–482.
- Koziel, S., Bandler, J.W., Madsen, K., 2006. A space mapping framework for engineering optimization: Theory and implementation. *IEEE Trans. Microw. Theory Tech.* 54 (10), 3721–3730.
- Kramer, O., Echeverría Ciaurri, D., Koziel, S., 2011. Derivative-free optimization. In: Koziel, S., Yang, X.S. (Eds.), *Computational Optimization, Methods and Algorithms*. Springer, pp. 61–83.
- Li, B., Friedmann, F., 2005. Novel multiple resolutions design of experiment and response surface methodology for uncertainty analysis of reservoir simulation forecasts. In: SPE Reservoir Simulation Symposium. p. 16. <http://dx.doi.org/10.2118/92853-MS>.
- Li, K., Horne, R., 2005. An analytical model for production decline-curve analysis in naturally fractured reservoirs. *SPE Reserv. Eval. Eng.* 8, <http://dx.doi.org/10.2118/83470-PA>.
- McKay, M.D., Beckman, R.J., Conover, W.J., 1979. A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics* 21 (2), 239–245, URL: <http://www.jstor.org/stable/1268522>.
- Nielsen, H., Lophaven, S., Søndergaard, J., 2002. DACE – A MATLAB Kriging Toolbox. Informatics and Mathematical Modelling, Technical University of Denmark, DTU.
- Nocedal, J., Wright, S., 2006. *Numerical Optimization*, second ed. In: Springer Science, vol. 35, Springer, p. 664.
- Nwankwor, E., Nagar, A.K., Reid, D.C., 2013. Hybrid differential evolution and particle swarm optimization for optimal well placement. *Comput. Geosci.* 17 (2), 249–268. <http://dx.doi.org/10.1007/s10596-012-9328-9>.
- Onwunali, J.E., Durlafsky, L.J., 2010. Application of a particle swarm optimization algorithm for determining optimum well location and type. *Comput. Geosci.* 14 (1), 183–198. <http://dx.doi.org/10.1007/s10596-009-9142-1>.
- Peng, C.Y., Gupta, R., 2004. Experimental design and analysis methods in multiple deterministic modelling for quantifying hydrocarbon in-place probability distribution curve. In: SPE Asia Pacific Conference on Integrated Modelling for Asset Management. p. 16. <http://dx.doi.org/10.2118/87002-MS>.
- Peters, E., Arts, R.J., Brower, G.K., Geel, C.R., Cullick, S., Lorentzen, R.J., Chen, Y., Dunlop, K.N.B., Vossepoel, F.C., Xu, R., Sarma, P., Alhuthali, A.H., Reynolds, A.C., 2010. Results of the brugge benchmark study for flooding optimization and history matching. *SPE Reserv. Eval. Eng.* 13 (03), 391–405.
- Rasmussen, C.E., Williams, C.K.I., 2006. *Gaussian Processes for Machine Learning*. MIT Press.
- Santner, T.J., Williams, B., Notz, W., 2003. *The Design and Analysis of Computer Experiments*. Springer.
- Sarma, P., Chen, W.H., 2008. Efficient well placement optimization with gradient-based algorithms and adjoint models. In: Intelligent Energy Conference and Exhibition. Society of Petroleum Engineers, <http://dx.doi.org/10.2118/112257-MS>.
- Sayarpour, M., Zuluaga, E., Kabir, C.S., Lake, L.W., 2009. The use of capacitance-resistance models for rapid estimation of waterflood performance and optimization. *J. Pet. Sci. Eng.* 69 (3), 227–238. <http://dx.doi.org/10.1016/j.petrol.2009.09.006>, URL: <http://www.sciencedirect.com/science/article/pii/S0920410509002046>.
- Sayyafzadeh, M., 2017. Reducing the computation time of well placement optimisation problems using self-adaptive metamodeling. *J. Pet. Sci. Eng.* 151, 143–158.
- Sayyafzadeh, M., Alrashdi, Z., 2019. Well controls and placement optimisation using response-fed and judgement-aided parameterisation: Olympus optimisation challenge. *Comput. Geosci.* <http://dx.doi.org/10.1007/s10596-019-09891-y>.
- Spall, J.C., 1992. Multivariate stochastic approximation using a simultaneous perturbation gradient approximation. *IEEE Trans. Automat. Control* 37 (3), 332–341.
- Suwardati, E., Krogstad, S., Foss, B., 2015. Adjoint-based surrogate optimization of oil reservoir water flooding. *Opt. Eng.* 16 (2), 441–481. <http://dx.doi.org/10.1007/s11081-014-9268-4>.
- Torczon, V., 1997. On the convergence of pattern search algorithms. *SIAM J. Optim.* 7 (1), 1–25.
- Wang, H., Echeverría Ciaurri, D., Durlafsky, L.J., Cominelli, A., 2012. Optimal well placement under uncertainty using a retrospective optimization framework. *SPE J.* 17 (1), 112–121.
- Wang, C., Li, G., Reynolds, A.C., 2007. Optimal well placement for production optimization. In: Eastern Regional Meeting. Society of Petroleum Engineers, p. 5. <http://dx.doi.org/10.2118/111154-MS>.
- Wolfsteiner, C., Durlafsky, L.J., Aziz, K., 2000. Approximate model for productivity of nonconventional wells in heterogeneous reservoirs. *SPE J.* 5 (02), 218–226. <http://dx.doi.org/10.2118/62812-PA>.

- Yeten, B., Brouwer, D.R., Durlofsky, L.J., Aziz, K., 2004. Decision analysis under uncertainty for smart well deployment. *J. Pet. Sci. Eng.* 44 (1), 175–191. <http://dx.doi.org/10.1016/j.petrol.2004.09.002>, URL: <http://www.sciencedirect.com/science/article/pii/S0920410504001299>.
- Yeten, B., Castellini, A., Guyaguler, B., Chen, W.H., 2005. A comparison study on experimental design and response surface methodologies. In: *SPE Reservoir Simulation Symposium*. p. 15. <http://dx.doi.org/10.2118/93347-MS>.
- Yeten, B., Durlofsky, L.J., Aziz, K., 2003. Optimization of nonconventional well type, location, and trajectory. *SPE J.* 8 (03), 200–210.
- Yousef, A.A., Gentil, P.H., Jensen, J.L., Lake, L.W., 2006. A capacitance model to infer interwell connectivity from production and injection rate fluctuations. *SPE Reserv. Eval. Eng.* 9 (06), 630–646. <http://dx.doi.org/10.2118/95322-PA>.
- Zandvliet, M., Handels, M., van Essen, G., Brouwer, R., Jansen, J.-D., 2008. Adjoint-based well-placement optimization under production constraints. *SPE J.* <http://dx.doi.org/10.2118/105797-PA>.
- Zubarev, D.I., 2009. Pros and cons of applying proxy-models as a substitute for full reservoir simulations. In: *SPE Annual Technical Conference and Exhibition*. Society of Petroleum Engineers.

5 Bilevel optimization of well placement and well control settings assisted by capacitance-resistance models


Arouri, Y.; Lake, L. W. & Sayyafzadeh, M.

SPE Journal, 2022.

Statement of Authorship

Title of Paper	Bilevel optimization of well placement and control settings assisted by capacitance resistance models
Publication Status	<input checked="" type="checkbox"/> Published <input type="checkbox"/> Accepted for Publication <input type="checkbox"/> Submitted for Publication <input type="checkbox"/> Unpublished and Unsubmitted work written in manuscript style
Publication Details	Arouri, Y.; Lake, L. W. & Sayyafzadeh, M. Bilevel optimization of well placement and control settings assisted by capacitance resistance models. SPE Journal, Published.


Principal Author


Name of Principal Author (Candidate)	Yazan Arouri
Contribution to the Paper	Coding, experimentation, analysis of results, original manuscript
Overall percentage (%)	70%
Certification:	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party. I do not intend to constrain its inclusion in this thesis. I am the primary author of this paper.
Signature	
Date	21/05/2022

Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

- i. the candidate's stated contribution to the publication is accurate (as detailed above);
- ii. permission is granted for the candidate to include the publication in the thesis; and
- iii. the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author	Larry W. Lake
Contribution to the Paper	Discussion, analysis of results, revisions
Signature	
Date	May 20, 2020

Name of Co-Author	Mohammad Sayyafzadeh
Contribution to the Paper	Supervision, revisions, discussions, editing
Signature	
Date	06/06/2022

Please cut and paste additional co-author panels here as required.

Bilevel Optimization of Well Placement and Control Settings Assisted by Capacitance-Resistance Models

Yazan Arouri^{1*}, Larry W. Lake², and Mohammad Sayyafzadeh¹

¹Australian School of Petroleum and Energy Resources, The University of Adelaide

²Hildebrand Department of Petroleum and Geosystems Engineering, The University of Texas at Austin

Summary

Well control and well placement optimization have typically been considered as separate problems. More recently, there have been a number of works which have shown improved results when these two problems are considered in a joint manner. However, this joint optimization problem, whether in a sequential or simultaneous manner, is more computationally demanding. In light of this, we propose the use of capacitance-resistance models (CRMs) to assist the computational demand of the joint optimization of well controls and well placement. Specifically, we use a bilevel (or nested) approach, where the outer loop is the well placement problem and the inner loop is the well control problem assisted by CRMs. The well placement problem is solved using particle swarm optimization (PSO), and the well control problem is solved using Adam-simultaneous perturbation stochastic approximation (SPSA). The proposed approach is compared with the conventional implementation using only high fidelity full-physics simulations on two reservoir models of varying complexity. We also investigate the accuracy of the CRMs during the optimization procedure. The proposed approach resulted in solutions for the joint optimization problems with objective function values of up to 21.8% higher than the conventional approach and up to a 99.6% decrease in the number of required reservoir simulations.

Introduction

A field development plan outlines the specific strategy that aims at maximizing the economic returns from a hydrocarbon reserve over the course of its production lifetime. Typical details of a field development plan include well location, trajectory, drilling schedule, platform location (if offshore), well type and well control settings [injection/production rates and/or bottomhole pressures (BHPs)], among other considerations. All these aspects have an influence on the recovery of hydrocarbon volumes and hence on the monetary costs and returns. High fidelity full-physics reservoir simulations are typically used to perform what-if analyses to assess the suitability of any proposed field development plan.

Owing to the complex nonlinear relationships between the development decisions and the hydrocarbon production volumes, the application of optimization techniques to field development planning has been an active arena of research over the past few decades. Traditionally, because of the differences in ruggedness of the objective function landscapes, well placement and well control optimization have been considered as separate problems and solved—for the most part—using different techniques.

Well placement optimization involves varying the well locations and/or trajectories subject to physical field constraints, such as reservoir boundaries and interwell constraints, to maximize an objective. Well placement optimization is typically characterized by a highly nonlinear relationship between the well locations/trajectories and the produced fluid volumes. This results in a rough, highly multimodal, nonconvex objective function.

Consequently, derivative-free optimization algorithms are applied to this type of problem in an effort to traverse the solution space in search of a global solution—albeit in a relatively low dimensional space. Examples of derivative-free optimization algorithms applied to well placement include PSO (Onwunalu and Durlafsky 2010; Arouri et al. 2022), genetic algorithms (GA) (Emerick et al. 2009; Sayyafzadeh 2017), evolutionary strategies (Sayyafzadeh and Alrashdi 2019; Alrashdi and Sayyafzadeh 2019), differential evolution (Ding 2008), covariance matrix adaptation-evolutionary strategy (Bouzarkouna et al. 2012), and simulated annealing (Beckner and Song 1995). Such techniques use objective function values to direct the search by, either deterministically or stochastically, updating the algorithmic parameters to propose a new population at each generation.

Given these optimization algorithms typically require a large number of function evaluations (reservoir simulation calls for convergence), gradient-based algorithms have been shown to provide a suitable alternative for an improved local solution—particularly in computationally constrained scenarios (Arouri and Sayyafzadeh 2020b). In well placement optimization, the well control settings (i.e., injection/production rates and/or BHPs) are set to predefined values. Common control strategies include fixed settings (e.g., maximum injection rates and minimum producer BHP) or a reactive control strategy—where a production well is shut in once becoming uneconomic.

In contrast, well control optimization problems aim at maximizing a defined objective by assigning values for control settings (injection/production rates and/or BHP of each well) for a fixed set of well locations. This type of optimization problem presents a less rugged (i.e., smoother) objective function landscape, characterized by long flat valleys, which are influenced by the piecewise, temporal, continuous variables (Do and Reynolds 2013; Zhao et al. 2013; Fonseca et al. 2014).

Given their computational efficiency in high dimensional settings, gradient-based algorithms have been favored for such problems. Gradient-based algorithms applied to well control optimization include SPSA (Wang et al. 2009; Foroud et al. 2018; Arouri and Sayyafzadeh 2020a), ensemble-based techniques (Chen et al. 2009; Chaudhri et al. 2009; Leeuwenburgh et al. 2010; Perrone and Rossa 2015; Oguntola and Lorentzen 2021), simplex approximations (Fonseca et al. 2017; Lu et al. 2017b; Liu and Reynolds 2020; Silva et al. 2020), and adjoint-based formulations (Sarma et al. 2005; Jansen 2011; Volkov and Bellout 2017). These algorithms rely on either an

*Corresponding author; email: yazan.arouri@adelaide.edu.au

approximation or a calculation of a gradient to determine the search direction. Derivative-free optimization algorithms, such as Hooke-Jeeves direct search and generalized pattern search, have also been applied to well control optimization (Echeverria Ciaurri et al. 2010).

More recently, research has shown that although the well placement and well control problems may present different objective function landscape architectures, the consideration of the two as separate problems may lead to suboptimal solutions (Bellout et al. 2012; Sayyafzadeh and Alrashdi 2019). As one would expect, there is an intrinsic interdependency between well locations and the associated control settings which leads to the expectation that their joint consideration would lead to better solutions. However, this would come at the expense of a significant number of computationally demanding reservoir simulations.

The implementation of joint optimization has been varied in the literature. A common implementation is the simultaneous optimization of well locations and well controls. As the name suggests, in this implementation the decision variable vector is composed of the variables relating to well location and well control, which are optimized concurrently. This increases the complexity of the optimization, as the dimensionality of the problem increases significantly and ruggedness arising from the well placement inclusion remains.

Another method is the sequential approach, in which either a well placement or well control optimization problem is solved followed by the other. The second problem is solved using the optimal solution found in the preceding problem. For example, if well placement optimization is solved first, the optimal well locations found are then used for well control optimization (or vice versa). In this approach, the two problems can be solved sequentially once or in an iterative manner until no additional improvements occur. An advantage of the sequential approach is the ability to apply problem-specific algorithms for each decision variable type.

There is still debate in the research community about which implementation (simultaneous or sequential) is preferred. Several works argue that simultaneous optimization produces higher quality solutions as it incorporates all the decision variables into one optimization problem (Bellout et al. 2012; Isebor et al. 2014; Forouzanfar et al. 2016; Sayyafzadeh and Alrashdi 2019). On the other hand, there is also research that shows the superiority of a (typically, iterative) sequential approach over a simultaneous one (Humphries et al. 2014; Humphries and Haynes 2015; Wang et al. 2016; Lu et al. 2017a; Lu and Reynolds 2020). The development and implementation of optimization techniques for the solution of the joint optimization problem include gradient-based algorithms (Li and Jafarpour 2012; Li et al. 2013; Lu et al. 2017b; Tanaka et al. 2018; Salehian et al. 2021), derivative-free algorithms (Awotunde 2014; Forouzanfar et al. 2016; Wang et al. 2016; Tanaka et al. 2018; Sayyafzadeh and Alrashdi 2019; Alrashdi and Sayyafzadeh 2019), or a combination of these (Bellout et al. 2012; Isebor et al. 2014; Humphries et al. 2014; Humphries and Haynes 2015; Lu and Reynolds 2020).

Given the nature of the joint optimization problem, hybrid algorithms attempt to take advantage of local and global capabilities to traverse the search space more efficiently than single-level optimization techniques. Bellout et al. (2012) used pattern search methods, including Hooke-Jeeves direct search and generalized pattern search for well placement, in combination with sequential quadratic programming (using adjoint gradients) for well control optimization. They implemented joint optimization in a nested fashion, where each proposed well location underwent an inner loop of well control optimization. Isebor et al. (2014), Humphries and Haynes (2015), and Humphries et al. (2014) implemented a combination of PSO and pattern search methods (either mesh-adaptive direct search or generalized pattern search) for both the simultaneous and sequential optimization. Isebor et al. (2014) and Humphries et al. (2014) considered only vertical wells while Humphries and Haynes (2015) studied the placement of 3D wells. Lu and Reynolds (2020) implemented a sequential approach using GA with mixed encodings to optimize well drilling paths, drilling sequences, and well types and used stochastic simplex approximate gradient (or StoSAG) for well control optimization.

The computational costs of high fidelity (full-physics) reservoir simulations can be reduced with the effective use of surrogate models, also known as proxy models or metamodels. Surrogate models can be categorized into those based on response surface modeling and those which are based on reduced-order modeling. In the first category, a model is trained and calibrated using a number of previously simulated high fidelity measurements. The implementation of these surrogates tends to be straightforward. Examples of these surrogates and their implementation include artificial neural networks (Yeten et al. 2003; Cullick et al. 2006; Sayyafzadeh 2017), polynomial approximation (Peng and Gupta 2004; Yeten et al. 2005; Zubarev et al. 2009), thin plate splines (Li and Friedmann 2005), and kriging (Yeten et al. 2003, 2005; Peng and Gupta 2004).

The second category is based on simplifications of the underlying scientific laws or the numerical computations. These surrogates tend to require an understanding of the governing equations and the underlying physics for proper implementation. Some examples of reduced-physics surrogates include streamline simulations (Batycky et al. 1997; Datta-Gupta and King 2007), component lumping for compositional simulation (Alavian et al. 2014), single-phase flow (Christian et al. 2000), and CRMs (Yousef et al. 2006; Weber et al. 2009; Sayarpour et al. 2009; Jafroodi and Zhang 2011; Hong et al. 2017). Another type of reduced-order surrogates is the reduction of the system of equations in the numerical computation from high to low dimensional. These surrogates include proper orthogonal decomposition (Van Doren et al. 2006; Cardoso et al. 2009; Alghareeb 2015; Jansen and Durlofsky 2017), missing point estimation (Cardoso et al. 2009), trajectory piecewise linearization (Cardoso et al. 2009; Jansen and Durlofsky 2017), and the discrete empirical interpolation method (Alghareeb 2015; Suwartadi et al. 2015).

There have been numerous applications of surrogate models within well placement, well control, and joint optimization to improve the computational efficiency of solving these problems. Suwartadi et al. (2015), Alghareeb (2015), Van Doren et al. (2006), and Jansen and Durlofsky (2017) investigated the use of proper orthogonal decomposition and its variants for surrogate-based well control optimization. The proper orthogonal decomposition-based models are optimized by adjoint methods (Van Doren et al. 2006; Suwartadi et al. 2015; Jansen and Durlofsky 2017) or direct search methods (Van Doren et al. 2006; Jansen and Durlofsky 2017). Zhao et al. (2013) used approximated gradients from ensemble-based techniques and SPSA to optimize a quadratic interpolation model for well controls. Wen et al. (2014) used streamline simulations alongside time-of-flight measurements to optimize well controls in mature fields. Echeverria Ciaurri and Wagenaar (2016) applied a corrective surrogate treatment (manifold mapping) in conjunction with surrogates based on generalized barycentric coordinates and quadratic approximation for well controls. Using CRMs, Weber et al. (2009) optimized the well controls for 197 producers using real historical field data. Hong et al. (2017) and Jafroodi and Zhang (2011) also used CRMs; however, their works applied ensemble-based optimization techniques to CRM for well control optimization.

Surrogates have also been implemented for well placement optimization in various studies. Pan and Horne (1998) used least squares and kriging to build multivariate interpolation models for the optimization of well location, well patterns, and operations scheduling. Yeten et al. (2003) and Guyaguler et al. (2000) used an artificial neural network in conjunction with GA to improve the computational efficiency of well placement optimization. Similarly, Sayyafzadeh (2017) combined GA with a self-adaptive artificial neural network to reduce the computational time of well placement optimization. Artus et al. (2006) used a statistical proxy to incorporate uncertainty in the well placement problem within a GA framework. Zubarev et al. (2009) investigated and compared multiple analytical surrogate techniques including kriging, artificial neural networks, quadratic approximation, and thin plate splines for the placement of two vertical wells. In Bouzarkouna et al. (2012), metamodels were coupled with covariance matrix adaptation-evolutionary strategy to optimize two wells that were represented in three dimensions.

Given the recent interest in joint optimization, there are a limited number of works (Møyner et al. 2014; Aliyev and Durlofsky 2017) which investigate the use of surrogates for such problems. However, the potential benefits for such a computationally intensive optimization task can be significant. More recently, De Brito and Durlofsky (2020) implemented a nested approach for the joint optimization using a set of surrogate treatments. The inner loop of well control optimization is divided further into two subproblems each with a different objective function based on different surrogates.

In our work, we focus on using CRMs for the efficient joint optimization. The robustness of CRMs comes from the fact that they only require historical field data of injection and production rates and, when available, BHP from producers. In addition, their straightforward implementation and ability to provide accurate production forecasts for the fraction of the computational cost of a high fidelity grid-based reservoir simulation make them a potentially powerful tool for joint optimization. Our work develops a bilevel (nested) approach for joint optimization using the strengths of CRMs for lower level decision variable (well control settings) optimization within a PSO framework where the upper decision variables (well placement) are optimized.

Following the Introduction, we state the joint optimization problem as well as the objective function in Problem Formulation. Method provides an overview of the optimization method, including a detailed description of CRMs for well control optimization. Experimental Problems presents the 2D and 3D models used for the experimental studies. The applicability of the proposed approach is demonstrated and discussed in Results and Discussion. This is followed by the Concluding Remarks.

Problem Formulation

This section constructs the problem for joint optimization. We first present the joint optimization as a general optimization problem followed by the definition of the objective function used for this work.

Problem Statement. The optimization problem involves the minimization of a defined objective function where the well locations and well control settings are the variables of interest. The joint optimization problem can be formulated as a general optimization problem as follows:

$$\begin{aligned} & \min_{\mathbf{x} \in \mathbb{R}^{n_x}, \mathbf{u} \in \mathbb{R}^{n_u}} f(\mathbf{x}, \mathbf{u}) \\ & \text{s. t. :} \\ & c_i(\mathbf{x}, \mathbf{u}) \leq 0 \quad i \in K \\ & c_i(\mathbf{x}, \mathbf{u}) = 0 \quad i \in I, \end{aligned} \quad (1)$$

where $f(\mathbf{x}, \mathbf{u})$ is the objective function, \mathbf{x} denotes the vector of well location decision variables, \mathbf{u} denotes the vector of well control settings, n_x is the number of well placement decision variables, n_u is the number of well control decision variables, I and K are sets of indices for equality and inequality constraint functions, $c_i(\mathbf{x}, \mathbf{u})$, respectively. These constraint functions may include linear and nonlinear constraints, as well as bound constraints. The constraint functions, in the case studies presented in this paper, incorporate bounds for the drilling locations and for the minimum distance between any two wells. The bounds for the drilling locations ensure the wells are located in active gridblock cells. The well control settings (\mathbf{u}) have lower (\mathbf{u}_{lb}) and upper limits (\mathbf{u}_{ub}) based on operational constraints.

Objective Function. The objective function is the negative of the net present value (NPV) for a given reservoir lifespan. The objective function is defined as:

$$f(\mathbf{x}, \mathbf{u}) = -\text{NPV}(\mathbf{x}, \mathbf{u}) = \sum_{i=1}^{N_t} \frac{-r_{o,i} Q_{o,i}(\mathbf{x}, \mathbf{u}) + c_{wp,i} Q_{wp,i}(\mathbf{x}, \mathbf{u}) + c_{wi,i} Q_{wi,i}(\mathbf{x}, \mathbf{u})}{(1+b)^i}, \quad (2)$$

where $r_{o,i}$, $c_{wp,i}$, and $c_{wi,i}$ are the price of oil and costs of water separation and injection, respectively, all of them per unit volume and defined from time t_i to t_{i+1} (there are N_t such intervals); $Q_{o,i}$, $Q_{wp,i}$, and $Q_{wi,i}$ denote the field oil production, field water production, and field water injection volumes during the output interval, and b is the discount factor.

In addition, although not considered in this work, one may incorporate geological uncertainty in the objective function through the use of multiple geological realizations (Arouri et al. 2022).

Method

In this section, we first present the overview of the bilevel optimization. Then, the individual components of the optimization method, including a global derivative-free optimization algorithm, PSO, and a gradient-based algorithm (Adam-SPSA) are outlined. We also describe the mathematical equations and analytical solutions that make up CRM, specifically the producer-based CRM (CRMP). Finally, we present the implementation of CRM in the bilevel (nested) approach.

Bilevel Optimization. To solve the joint optimization problem, we apply a bilevel (or nested) approach similar to the one proposed in Bellout et al. (2012). The outer loop is the well placement optimization and is solved using PSO, while the inner loop is the well control optimization and is solved using Adam-SPSA. While PSO and Adam-SPSA are used within the proposed CRM bilevel approach, any optimization algorithm(s) may be used. In other words, for every proposed solution (i.e., well location) from PSO, a well control optimization is undertaken by Adam-SPSA.

The algorithm begins with the initialization of the PSO parameters (ω , c_1 , and c_2), a randomly generated swarm (population), and the generation counter k . The particles are proposed solutions, \mathbf{x}_j , for well locations only (i.e., the decision variables in PSO are the well locations). The feasibility of each proposed well location is checked, and, if a solution is infeasible, it is repaired as detailed in Arouri and Sayyafzadeh (2022). Next, for each particle (well location) in the swarm, a well control optimization is undertaken using Adam-SPSA to find optimal well control settings, \mathbf{u}_j^* , for the associated well location \mathbf{x}_j (i.e., $\mathbf{u}_j^* = \text{argmin}_{\mathbf{u}_j \in \mathbb{R}^{n_u}} f(\mathbf{x}_j, \mathbf{u}_j)$). This is what makes joint optimization a computationally expensive problem to solve. In the remainder of this work, for the sake of brevity, we define \mathbf{u}_j^* as \mathbf{u}_i^* .

The objective function values, $f(\mathbf{x}_j, \mathbf{u}_i^*)$, are used in PSO to propose the new particles for the following generation. This procedure is repeated until a stopping criterion is met (number of generations in our work). **Fig. 1** shows the approach in a flow chart.

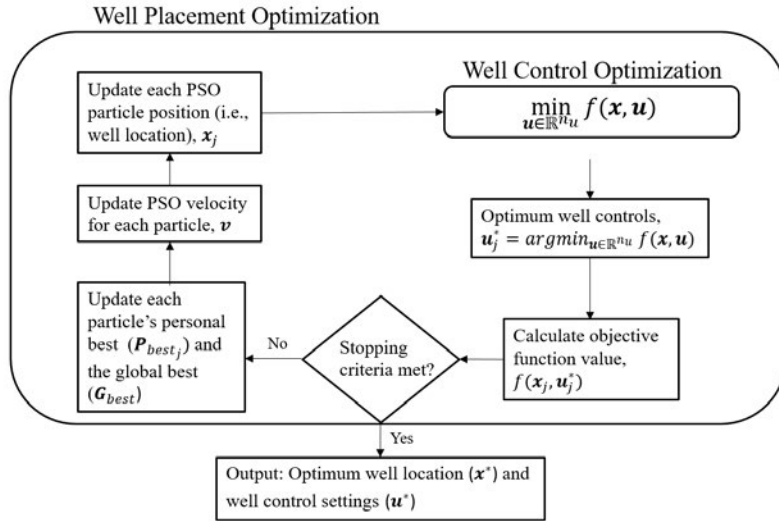


Fig. 1—Flow chart presenting the bilevel approach for the joint optimization of well control settings and well locations.

PSO. PSO is a stochastic, global, derivative-free optimization algorithm that takes inspiration from the behaviors of biological swarms of animals (Kennedy and Eberhart 1995). The stochastic nature of the algorithm allows it to traverse the multimodal landscape of well placement problems. As previously mentioned, PSO has been applied to field development optimization problems and has shown to be successful in Onwunalu and Durlofsky (2010), Foroud et al. (2018), Alrashdi and Sayyafzadeh (2019), and Arouri et al. (2022).

PSO begins with a swarm (population) of (typically) randomly generated particles (solutions). The fitness value of each particle is evaluated using the objective function, $f(\mathbf{x}_j, \mathbf{u}_j^*)$. If there is an improvement in fitness value of the best global (swarm) solution and each particle's best solution, then these are updated. Next, the position of each particle, \mathbf{x} , is updated based on its velocity vector, \mathbf{v} , as follows:

$$\mathbf{x}_{\text{new}} = \mathbf{x} + \mathbf{v} \quad (3)$$

The velocity vector is calculated based on a particle's personal best and the neighborhood best. There are different possible neighborhood topologies, including star, ring, and global (Engelbrecht 2013; Liu et al. 2016). This study uses a global topology, so each particle is influenced by the global (swarm) best found in previous generations. Hence, the velocity equation for particle \mathbf{x} is updated as follows:

$$\mathbf{v} = \omega \mathbf{v}_{\text{old}} + c_1 r_1 (\mathbf{P}_{\text{best}} - \mathbf{x}) + c_2 r_2 (\mathbf{G}_{\text{best}} - \mathbf{x}), \quad (4)$$

where ω , c_1 , and c_2 are the algorithmic parameters known as the inertial, cognitive, and social parameters, respectively. r_1 and r_2 are random vectors between 0 and 1, and \mathbf{P}_{best} and \mathbf{G}_{best} are the personal best for particle \mathbf{x} from previous generations and the global best of the swarm, respectively. The algorithmic parameters play a role in the balance between exploration and exploitation of PSO. For our work, we use values of 0.721 for ω and 1.193 for c_1 and c_2 , based on Onwunalu and Durlofsky (2010) and a population size of 20.

Adam-SPSA. Adam-SPSA was first introduced for well control optimization problems in Arouri and Sayyafzadeh (2020a) and has also been implemented for well placement optimization in Arouri and Sayyafzadeh (2022). It uses an adaptive moment estimation framework in conjunction with SPSA as the gradient approximation technique. It gives superior results to the typical steepest descent framework that is usually implemented for well control optimization.

The framework finds its success in the stochasticity and the incorporation of additional information from the approximated gradients in previous iterations to calculate variable-specific search directions. The search direction is a vector composed of an estimated exponential moving average of previous gradients that is normalized elementwise by an estimated uncentered variance. This allows the search direction to be adjusted for each control variable, resulting in convergence speedups (Arouri and Sayyafzadeh 2020a). The search direction can be thought of as a signal-to-noise ratio, where a small signal-to-noise value represents a lower quality search direction and hence greater uncertainty in its representation of the true gradient (Kingma and Ba 2014).

The gradient approximation uses the SPSA. Spall (1992) introduced this gradient approximation for problems where an analytical derivative could not be calculated. The advantage of SPSA is that it only requires two function evaluations to approximate a gradient using the central-difference method. This enables SPSA to be computationally efficient for high dimensional problems. The stochastic gradient approximation is:

$$\begin{aligned} \hat{\mathbf{g}}_k(\mathbf{x}_k) &= \begin{bmatrix} \frac{f(\mathbf{x}_k + c_k \Delta_k) - f(\mathbf{x}_k - c_k \Delta_k)}{2c_k \Delta_{k1}} \\ \vdots \\ \frac{f(\mathbf{x}_k + c_k \Delta_k) - f(\mathbf{x}_k - c_k \Delta_k)}{2c_k \Delta_{kp}} \end{bmatrix} \\ &= \frac{f(\mathbf{x}_k + c_k \Delta_k) - f(\mathbf{x}_k - c_k \Delta_k)}{2c_k} \times [\Delta_{k1}^{-1}, \Delta_{k2}^{-1}, \dots, \Delta_{kp}^{-1}]^T, \end{aligned} \quad (5)$$

where the mean-zero p -dimensional random perturbation vector, $\Delta_k = [\Delta_{k1}^{-1}, \Delta_{k2}^{-1}, \dots, \Delta_{kp}^{-1}]^T$, has a user-specified distribution, and c_k is a positive scalar. The gradient approximation is calculated using SPSA with the parameters following the recommendations given in Spall (1998). These are used to calculate the proposed perturbation size from the following gain sequence:

$$c_k = \frac{c}{(k+1)^\gamma}, \quad (6)$$

where k is the iteration number, γ is a positive scalar constant, and c is the initial perturbation size. Readers are referred to Spall (1998) for additional implementation guidelines.

In this paper, Adam-SPSA with an average of five stochastic gradient approximations is used to solve the well control optimization problem (i.e., the inner loop of the bilevel approach). The step size and perturbation size were taken to be 0.01. For conciseness, the details of Adam-SPSA are not presented. Interested readers are directed to Arouri and Sayyafzadeh (2020a) for details of Adam-SPSA and applications to well control optimization.

CRM. Albertoni and Lake (2003) first introduced the idea of using multivariate linear regression to obtain interwell connectivity information to predict total fluid production rates. Later, Yousef et al. (2006) advanced and formalized the idea by incorporating material balance equations with the linear productivity model (or the deliverability equation) into, what is now known as, CRMs. As such, CRMs encompass the compressibility (capacitance) as well as the transmissibility (resistance) of the reservoir, which could then be used to extract reservoir and fluid properties between injectors and producers. This was formalized into two model parameters: λ (or f ; interwell connectivity) and τ (time constant), which represent the interwell connectivity between an injector-producer pair and a direct measure of the dissipation of pressure between an injector and producer pair, respectively.

In Sayarpour et al. (2009), superposition in time was used to analytically solve the CRM equation for three different control volumes. The control volumes reflect multiple levels of investigation from a field scale, with the CRM-tank, to a producer-level with CRMP, and to individual injector-producer pairs with CRM injector-producer.

The original CRM proposed by Yousef et al. (2006) only predicts the total production rate, so fractional flow models are required to obtain separate fluid-phase flow rates. Multiple works investigated this by combining CRM with a Buckley-Leverett-based fractional flow equation (Sayarpour et al. 2011), the Koval model (Cao et al. 2015), and a fully coupled two-phase fractional flow model (Cao et al. 2014). An extensive literature review on CRMs and their applications is found in Holanda et al. (2018).

Governing Equations. The governing equations for CRM can be derived from the material balance equation for a reservoir tank control volume in conjunction with the linear productivity equations to give a first-order ordinary differential equation (Yousef et al. 2006):

$$\frac{dq_{prod}(t)}{dt} + \frac{1}{\tau} q_{prod}(t) = \frac{1}{\tau} q_{inj}(t) - J \frac{dp_{wf}}{dt}, \quad (7)$$

where $q_{prod}(t)$ and $q_{inj}(t)$ are the total production and injection rates, p_{wf} is the bottomhole flowing pressure in the producer, J is the productivity index (PI), and τ is the time constant defined as $\tau = \frac{c_i V_p}{J}$ for compressibility (c_i) and pore volume (V_p ; V_p).

Liang et al. (2007) extended the governing equations of CRM for a producer-based control volume (CRMP; see Fig. 2) to the following:

$$\frac{dq_{prod,j}(t)}{dt} + \frac{1}{\tau_j} q_{prod,j}(t) = \frac{1}{\tau_j} \sum_{i=1}^{N_{inj}} \lambda_{ij} q_{inj,i}(t) - J_j \frac{dp_{wf,j}}{dt}, \quad (8)$$

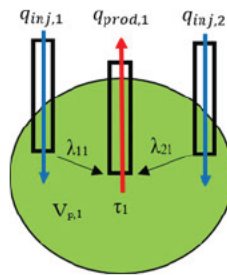


Fig. 2—An example of a control volume for a CRMP showing the interwell connectivity between Producer 1 and Injector 1 (λ_{11}) and Producer 1 and Injector 2 (λ_{21}). Producer 1 has a movable PV of $V_{p,1}$ and a time constant of τ_1 .

where the subscripts j and i represent the producer and injector number, respectively, and λ_{ij} is the interwell connectivity between producer j and injector i .

Using superposition in time, Sayarpour et al. (2009) derived the general form solution of Eq. 8 for piecewise constant injection rates (see Fig. 3) and linear variations in producer BHPs to be:

$$q_{prod,j}(t_n) = q_{prod,j}(t_0) \left(e^{-\frac{t_n - t_0}{\tau_j}} \right) + \sum_{k=1}^n \left\{ e^{-\frac{t_n - t_0}{\tau_j}} \left(1 - e^{-\frac{\Delta t_k}{\tau_j}} \right) \left[\sum_{i=1}^{N_{inj}} \lambda_{ij} q_{inj,i}(t) - J_j \frac{\Delta p_{wf,j}^k}{\Delta t_k} \right] \right\}. \quad (9)$$

This solution assumes that all the interwell connectivities (λ_{ij}), the time constants (τ_j), and PIs (J_j) are time-independent (i.e., constant for all time intervals Δt_k).

If the producer BHPs are kept constant, the solution of Eq. 8 for the total production equation for CRMP is (Sayarpour 2009):

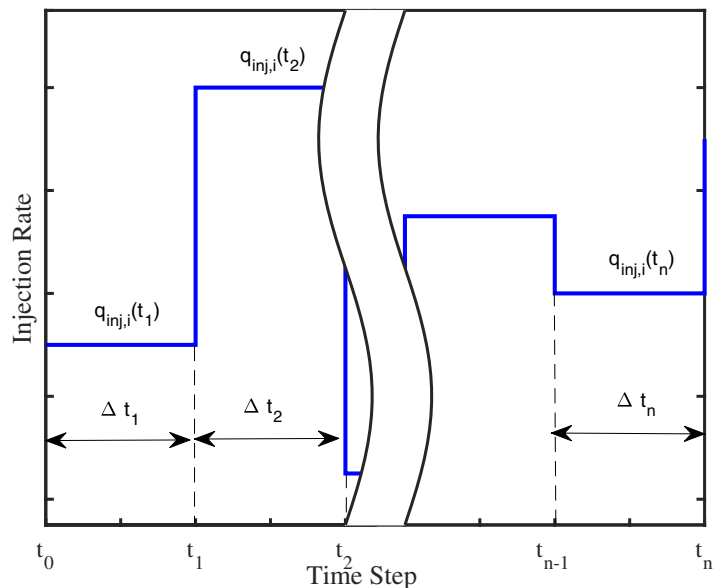


Fig. 3—An example of piecewise constant injection rates ($q_{inj,i}$) for each timestep interval Δt_k used to tune the CRMP.

$$q_{prod,j}(t_n) = q_{prod,j}(t_0) \left(e^{-\frac{t_n - t_0}{\tau_j}} \right) + \left(1 - e^{-\frac{\Delta t_k}{\tau_j}} \right) \left[\sum_{i=1}^{N_{inj}} \lambda_{ij} q_{inj,i}(t) \right]. \quad (10)$$

To obtain oil and water production rates, we use the Koval model method used in Cao et al. (2015). To do this the water cut at producer j at timestep t_n , $f_{w,j}(t_n)$, is given from the following:

$$f_{w,j}(t_n) = \begin{cases} 0, & \text{if } t_{D_j} < \frac{1}{K_{valj}} \\ \frac{K_{valj} - \sqrt{K_{valj} t_{D_j}}}{K_{valj} - 1}, & \text{if } \frac{1}{K_{valj}} \leq t_{D_j} \leq K_{valj} \\ 1, & \text{if } t_{D_j} > K_{valj}, \end{cases} \quad (11)$$

where K_{valj} is the Koval factor for producer j , t_{D_j} is the dimensionless time or the fraction of cumulative water injected into the movable PV of producer j and is defined as

$$t_{D_j} = \frac{\sum_{k=1}^n \left[\left(\sum_{i=1}^{N_{inj}} \lambda_{ij} q_{inj,i}(t_n) \right) \Delta t_n \right]}{V_{pj}}. \quad (12)$$

This gives the summation of the total water injection contribution from all injectors to producer j during time interval Δt_n . This assumes the water injection rate of an injector is constant over Δt_n (as is the case for piecewise constant injection rates). Once the water cut is known, the oil and water production for producer j can be calculated, respectively, as follows:

$$q_{w_j}(t_n) = q_j(t_n) \times f_{w,j}(t_n), \quad (13)$$

$$q_{o_j}(t_n) = q_j(t_n) \times (1 - f_{w,j}(t_n)). \quad (14)$$

Parameter Tuning. A least squares analysis is performed to tune the parameters of CRM based on simulation data (i.e., injection/production rates and BHPs). That is, only one high fidelity reservoir simulation is needed to tune (or “train”) the CRM. The model parameters for CRMP when used in conjunction with the Koval model are the interwell connectivity between each injector-producer pair (λ_{ij}), the time constant for each producer (τ_j), the Koval factor for each producer (K_{valj}), the movable PV for each producer (V_{pj}), and the PI for each producer (J_j). If the producer BHPs are unavailable or are kept constant, then the PI is not a model parameter. In this work, we define \mathbf{p} as the vector which consists of the CRMP parameters to be tuned. The minimization problem is formulated as follows:

$$\begin{aligned}
& \min_{\mathbf{p}} z \\
& \text{s.t.} \\
& \tau_j \geq 0, \\
& \lambda_{ij} \geq 0, \\
& K_{valj} \geq 1, \\
& \sum_{j=1}^{N_{prod}} V_{p_j} \leq V_{p_{Field}}, \\
& \sum_{j=1}^{N_{prod}} \lambda_{ij} \begin{cases} \leq 1, & \text{if injection loss exists} \\ = 1, & \text{if no injection loss occurs} \end{cases}
\end{aligned}$$

In this work, we define z as

$$z = \sum_{k=1}^n \sum_{j=1}^{N_{prod}} \left[\left(q_j(t_k) - q_j^{obs}(t_k) \right)^2 + \left(q_{o,j}(t_k) - q_{o,j}^{obs}(t_k) \right)^2 \right],$$

where n is the number of timesteps, N_{prod} is the number of producers, the superscript *obs* represents the ‘‘observed’’ data (i.e., simulation output). In our work, we use a sequential quadratic programming algorithm in the nonlinear optimization package in MATLAB® to solve this problem (Nocedal and Wright 2006).

Once the CRMP parameters (\mathbf{p}) are found, they can be used in Eq. 9 or Eq. 10 to predict production, depending on what information is available (i.e., injection/productions rates and/or producer BHPs). Although the well locations are not explicitly used when building a CRMP, the tuned CRMP is applicable for the well placements that were used to build it. In other words, if the well locations change, then a new CRMP will need to be built. However, as will be discussed later, this is still a fraction of the computational cost of a single high fidelity grid-based reservoir simulation. It should be noted that the assumptions made when deriving CRMs make them less applicable under certain conditions. This includes reservoirs with a gas cap, time-varying CRM parameters due to shut in of wells, and a constant number of wells. These assumptions are met in this study as the studied models are undergoing waterflooding by vertical wells in undersaturated reservoirs. The interested reader is referred to the extensive literature review in Holanda et al. (2018) for alternative CRM developments.

Bilevel Optimization Assisted by CRMs. To reduce the computational intensity of joint optimization, we propose the use of CRMPs in the inner loop for well control optimization. Specifically, when undertaking the well control optimization rather than using a high fidelity full-physics reservoir simulation ($f(\mathbf{x}_j, \mathbf{u}_i)$) to evaluate the proposed well control strategy, \mathbf{u}_i , a CRMP, $f'(\mathbf{x}_j, \mathbf{u}_i, \mathbf{p}_j^*)$, is used. \mathbf{p}_j^* denotes the tuned CRMP parameters for \mathbf{x}_j well placement solution.

To do this, for each particle (well location) in the swarm, a CRMP is built using the method described in subsection Parameter Tuning. The well controls (i.e., injection rates and producer BHPs) used to tune CRMP are randomly generated for each particle. Once the CRMP is built, the quality is checked to ensure it is above a coefficient of determination value (R^2), which is calculated by

$$R^2 = 1 - \frac{\sum_n (d_{t_n}^{obs} - d_{t_n}^{CRMP})^2}{\sum_n (d_{t_n}^{obs} - \bar{d}^{obs})^2}, \quad (15)$$

where $d_{t_n}^{obs}$ is the ‘‘observed’’ production data at time t_n from the simulation output, $d_{t_n}^{CRMP}$ is the predicted production data at time t_n from the CRMP, and \bar{d}^{obs} is the mean of $d_{t_n}^{obs}$ over all timesteps t_n (Hong et al. 2017).

If the CRMP does not meet the user-defined threshold R^2 value, then the tuning procedure is repeated using the same well locations but a different random sample of injection rates and producer BHPs. Each tuning procedure that is undertaken will require one high fidelity simulation to obtain ‘‘observation’’ data. This is repeated for a predefined maximum number (Max_{CRMP}) of times after which the best fit is taken if the threshold value is still not reached.

Next, well control optimization is undertaken for each particle (well location) using the associated tuned CRMP (to obtain production rates). Adam-SPSA is the algorithm used to optimize the CRMP with the decision variables being the well control settings (\mathbf{u}_i).

The optimized solution (\mathbf{u}_i^*) from the well control optimization with the associated well locations (\mathbf{x}_j) is re-evaluated using the high fidelity reservoir simulation. This objective function value ($f(\mathbf{x}_j, \mathbf{u}_i^*)$) is used to update the personal best (P_{best}) and global best (G_{best}) in PSO. The velocities are then updated using Eq. 4 followed by updating the particle positions using Eq. 3. This procedure is repeated until a stopping criterion is met. A pseudocode is presented in **Algorithm 1**.

Experimental Problems

This section presents the two reservoir models used for the numerical experiments. The first reservoir model is a 2D heterogeneous reservoir undergoing waterflooding. The second reservoir model is a more complex 3D benchmark case, known as the egg model, which is also undergoing water-flooding.

Model 1: 2D Example. The 2D model is a channelized reservoir developed with two production wells and four injection wells (Isebor et al. 2014). The model is discretized into a 40×40×1 grid, with each gridblock having a size of 50 ft in the x - and y -directions. **Fig. 4** shows the x -directional permeability field.

In our work, the decision variables for well placements were the x - and y -coordinates of the 6 wells, resulting in 12 decision variables. The decision variables for the well control strategies were the injection rates for the four injection wells and the BHPs for the production wells. The well control settings were divided into 50 control steps which were 30 days each over the reservoir lifespan of 1,500 days. This resulted in a total of 300 decision variables for the well control optimization problem. The injection wells were operated with minimum and maximum rates of 0 and 630 STB/D, respectively. On the other hand, the producers were operated by BHP with lower and upper

Algorithm 1—Bilevel optimization using CRMPs

```
1: Initialize a random swarm of particles; algorithmic parameters ( $\omega, c_1, c_2$ );  $k = 0$ 
2: while Termination condition(s) not met do
3:   for Each particle  $j$  do
4:     Check feasibility of particle
5:     if Infeasible then
6:       Repair
7:       Build CRMP for particle (well location)  $\mathbf{x}_j$  using “observation” data from high fidelity simulations
8:     if  $R^2 < \text{threshold value}$  AND  $\text{Max}_{\text{CRMP}} < N_{\text{CRMP}}$  then
9:       Obtain “observation” data from high fidelity simulation for new random injection rates and/or producer BHPs
10:      Repeat CRMP build using random injection rates and/or producer BHPs
11:       $N_{\text{CRMP}} = N_{\text{CRMP}} + 1$ 
12:     else
13:       Use CRMP with highest  $R^2$  value
14:       Perform well control optimization for particle  $\mathbf{x}_j$  using Adam-SPSA with CRMP to find  $\mathbf{u}_i^*$ 
15:       Re-evaluate optimized well control settings and well location using high fidelity reservoir simulation
16:     if  $f(\mathbf{x}_j, \mathbf{u}_i^*) < f(P_{\text{best}})$  then
17:        $P_{\text{best}j} \leftarrow \mathbf{x}_j$ 
18:     if  $f(\mathbf{x}_j, \mathbf{u}_i^*) < f(G_{\text{best}})$  then
19:        $G_{\text{best}} \leftarrow \mathbf{x}_j$ 
20:   Update velocities of each particle
21:   Update the positions of each particle
22:    $k = k + 1$ 
```

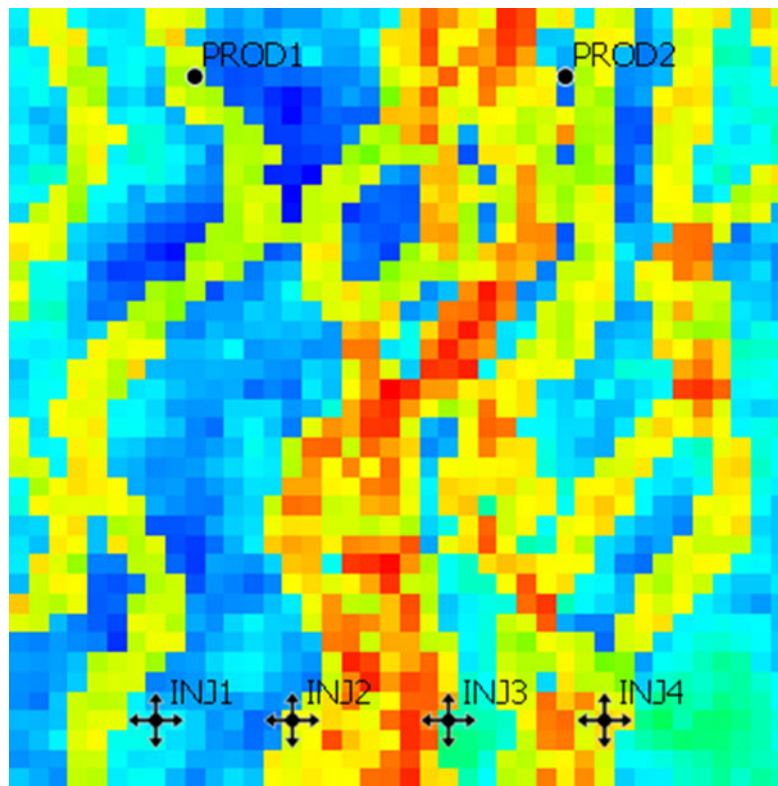


Fig. 4—A 2D heterogeneous channelized reservoir model (Model 1) with two production wells (circles) and four injection wells (crosses) showing x-directional permeability field. Red represents high permeability values, and blue represents low permeability values.

limits of 1,000 and 2,500 psi, respectively. A minimum interwell distance of 600 ft is used for this problem. The economic parameters and their values can be found in **Table 1**.

Model 2: 3D Example. The egg model (see **Fig. 5**) is a synthetic 3D reservoir model introduced into the reservoir optimization research community by Jansen et al. (2014). The model was introduced with an ensemble of 100 permeability realizations to incorporate the subsurface uncertainty in reservoir description. For the purpose of our work, we utilized Realization 1 in a deterministic manner.

Parameter	Value
Oil price, c_o	50 USD/STB
Water handling costs, c_{wp}	8 USD/STB
Water injection costs, c_{wi}	2 USD/STB
Discount rate, b	8%

Table 1—Economic parameters for experimental problems.

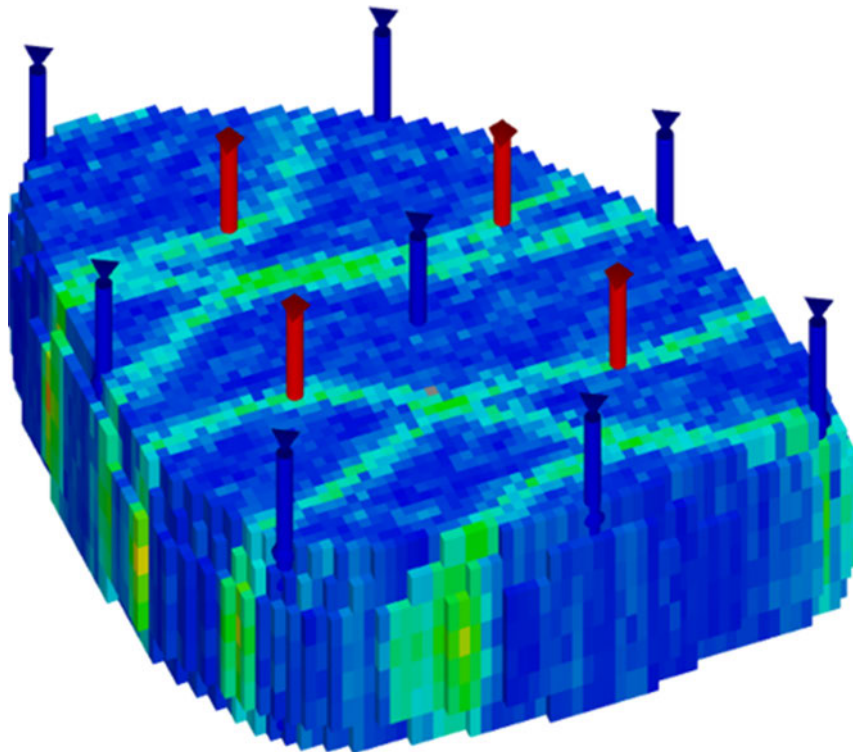


Fig. 5—A 3D benchmark case, the egg model, with eight injection wells (blue) and four production wells (red), showing the x -directional permeability field. Blue represents lower permeability values, and green represents higher permeability values.

The model is discretized into a $60 \times 60 \times 7$ grid, or 25,200 gridblocks, with 18,553 being active (i.e., having nonzero property values). The permeability fields are characterized with high permeability streaks which meander, representing river patterns seen in typical fluvial deposition environments. Given there are no gas cap and no aquifer present, the reservoir requires secondary recovery through water-flooding. This is accomplished with 12 vertical wells, of which 8 are injections wells and 4 are production wells.

In this work, the well placement strategies defined the x - and y -coordinates for the 12 vertical wells, resulting in a total of 24 decision variables. The well control strategies were defined as the injection rates for the eight water injection wells over the life span of the reservoir. These control settings were piecewise constant between each timestep, which are at 30-day intervals over a 10-year period. This results in a total of 960 decision variables. The lower and upper limit for the injection rates were 0 and 377 STB/D (60 std m^3 /D), respectively. The BHP of the production wells was kept constant at 5,730 psi (395 bar).

Additionally, we considered two physical field constraints to ensure only practical solutions were proposed. The first constraint is a minimum interwell distance of 164 ft (50 m), which represents six gridblocks. The second constraint is a reservoir boundary. To ensure wells were placed only in active cells, a projection-type repair mechanism was used. The reservoir boundary was represented by a polygon defined using piecewise linear functions. If a well is placed outside the defined polygon, the well is projected back onto a location on the boundary with the shortest distance to the violating well location. Further details of the constraint handling techniques can be found in Arouri and Sayyafzadeh (2022). The economic parameters and their values can be found in **Table 1**.

Results and Discussion

In this section, we present the experimental results for each model. This includes the results for examples of the built CRMPs, an example of well control optimization, and the results for the joint optimization for well placement and well controls.

Experimental Results: Model 1. Evaluating CRM. The well locations presented in Fig. 4 were used to build the CRMP. Random producer BHPs (see Fig. 6a for an example for PROD1) and injection rates (see Fig. 7a) between the lower and upper limits were generated and used. Since only the BHPs of the producers are used, the CRMP parameters (i.e., \mathbf{p}) were the interwell connectivities (λ_{ij}) between each producer and each injector, time constant for each producer (τ_j), the Koval factor for each producer (K_{val_j}), the movable PV for each producer (V_{p_j}), and the PI for each producer (J_j). This parameter tuning problem is solved using the procedure described in subsection Parameter Tuning.

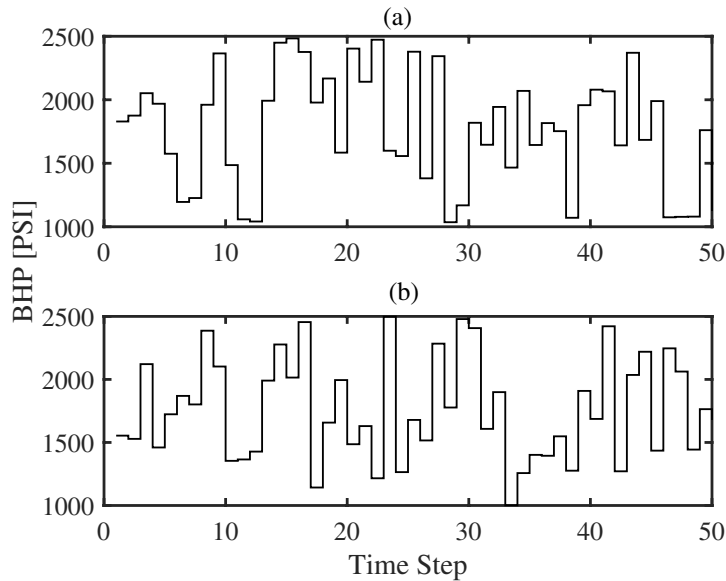


Fig. 6—The (a) training sample and (b) validation sample of random BHPs for PROD1 used to tune the CRMP parameters for the 2D model.

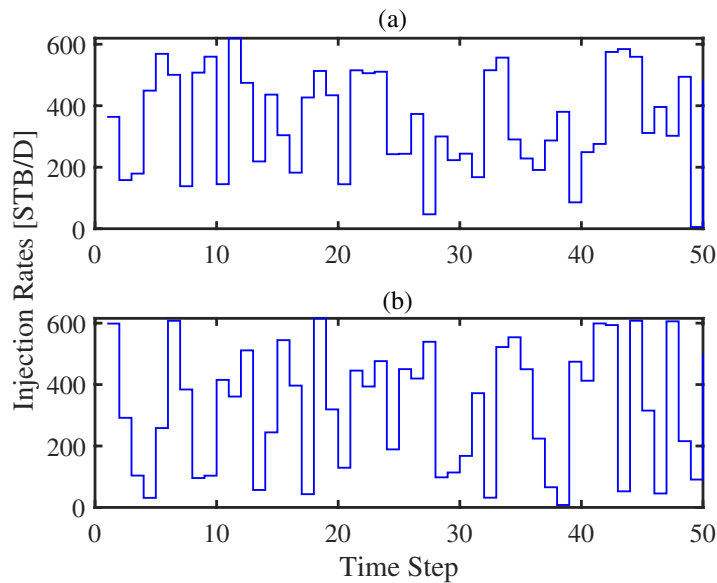


Fig. 7—The (a) training sample and (b) validation sample of random injection rates for INJ1 used to tune and test the CRMP for 2D model.

The output of the CRMP (using tuned parameters) is compared with the full-physics simulation in Fig. 8. The coefficient of determination values (Eq. 15) are 0.9859 and 0.9841 for oil and total production rates, respectively. Similar values of 0.9481 and 0.9618 were obtained for PROD2 for oil and total production rates, respectively. More importantly, as the objective function for the joint optimization was NPV, the difference between the CRMP and the full physics could potentially change the ranking of solutions in PSO. However, the absolute average percent difference between the NPV value for CRMP and full physics for this tuned model is only 2.96%. Although this value will change for each CRMP, it gives a promising indication of the quality that CRMP can provide.

The run-time computational savings of using a CRMP was significant. To run the grid-based reservoir simulation, it took around 3 seconds, while the CRMP took less than 0.01 seconds. This represents a reduction of more than 2 factors of 10. This becomes pivotal in making joint optimization more computationally efficient through the use of CRMPs.

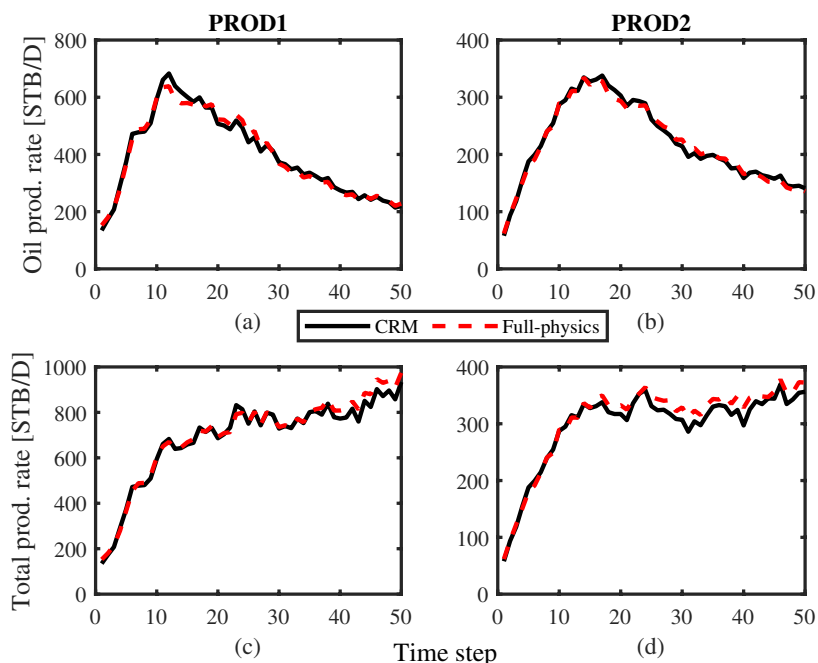


Fig. 8—The output from the tuned CRMP for the two production wells compared with the output from the full-physics simulation. **Top:** Oil production rate plots against timestep for PROD1 (left) and PROD2 (right). **Bottom:** Total production rate against timestep for PROD1 (left) and PROD2 (right).

To gain insight into the ability of the CRMP to forecast production for random well control settings, we ran the CRMP and full-physics simulation for an ensemble of 100 different random samples of injection rates and producer BHPs as validation samples. As an example, one of the validation samples of random INJ1 rates and PROD1 BHPs used to predict oil and total production rates is in **Figs. 6b and 7b**. The associated oil and total production rate plots are in **Fig. 9**. The coefficient of determination values (R^2) for oil and total production rates for PROD1 are 0.9745 and 0.9815, respectively. For PROD2, the values are 0.9809 and 0.9228 for oil and total production rates, respectively. The absolute average percent difference in NPV between the CRMP and the full-physics simulation for the random sample of well control strategies is only 1.89%. This reflects a very promising trade-off between the fidelity of a model and its associated runtime. In other words, for this example, with a reduction in computational costs of more than 2 factors of 10 there is only 1.89% difference in NPV.

For the ensemble of validation samples, the R^2 values for the oil production rates for PROD1 and PROD2 are 0.9620 and 0.9681, with standard deviation values of 0.0140 and 0.0170, respectively. Similarly, the R^2 values for the total production rates of the ensemble of validation samples are 0.9761 and 0.9402 with standard deviations of 0.0048 and 0.0228 for PROD1 and PROD2, respectively. This shows that CRMPs are able to produce sufficiently large R^2 values for random sets of water injection rates and BHPs.

Well Control Optimization. We tested the applicability of CRMP as a surrogate for high fidelity grid-based reservoir simulations in well control optimization as a proof of concept. The well locations are as depicted in **Fig. 4**. The decision variables were normalized between 0 and 1. The optimization algorithm used was Adam-SPSA. The stopping criterion for the optimization is a maximum number of function evaluations (full-physics reservoir simulations) of 250. The algorithmic parameters, specifically step size and perturbation size, were used to ensure the solutions were not swinging from the lower to upper bounds and vice versa.

When using the CRMP-based approach, only one reservoir simulation is required to build the model. Then, it can be used for the entire optimization process since the well locations do not change. To test the robustness of the CRMP, five different random initial guesses were used. **Table 2** presents the differences in the NPV between the CRMP and full-physics simulation for each of the five initial guesses. The absolute differences in NPV across the five different initial guesses are very promising, ranging from as little as 0.37% to a maximum of only 1.19%.

		Seed 1	Seed 2	Seed 3	Seed 4	Seed 5
Initial guess (million USD)	CRMP	33.16 (−1.19%)	33.74 (−0.18%)	33.32 (−0.92%)	32.38 (−0.37%)	32.85 (−0.42%)
Initial guess (million USD)	Full-physics	33.56	33.8	33.63	32.5	32.71
CRMP optimum solution (million USD)	CRMP	36.92 (+7.20%)	36.92 (+7.43%)	36.89 (+7.32%)	36.89 (+7.37%)	36.88 (+7.52%)
CRMP optimum solution (million USD)	Full-physics	34.44	34.37	34.37	34.36	34.3

Table 2—Comparison of results for well control optimization of Model 1.

The results of the optimization using both the CRMP and full-physics approaches are also presented in **Table 2**. The optimum solutions found by the CRMP-based optimization were also re-evaluated using the full-physics simulation. This was done to ensure that a solution

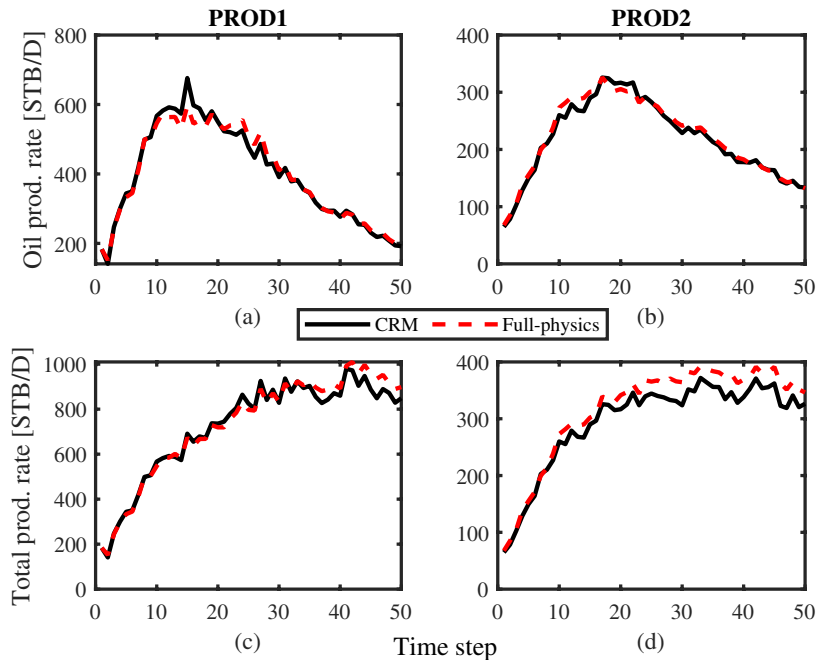


Fig. 9—The CRMP output of the validation sample for the two production wells compared with the output from the high fidelity grid-based numerical simulation using a random sample of well control settings. Top: Oil production rate plots against timestep for PROD1 (left) and PROD2 (right). Bottom: Total production rate against timestep for PROD1 (left) and PROD2 (right).

which represents an increase in CRMP NPV is also an increase in NPV calculated by the full-physics simulation. As shown, this is the case for all the initial guesses.

However, there is around a 7% difference in the NPV values. This increase in percent difference can be attributed to the well control settings of the optimal solutions found for the optimization. The optimal solutions are “bang-bang” control settings, which, as found in Zandvliet et al. (2007), can be the case for some waterflooding problems—especially for smaller models. This means the injection rates (or injection BHPs) and producer BHPs are at (or close to) their upper and lower limits, respectively. Typically, this can be expected to occur when water injection and handling costs are relatively low, which induces early water breakthrough in return for higher oil production rates initially. Consequently, the water cut of the producers approaches 100% (or close to it) relatively early in the reservoir’s lifespan. This has been shown in Cao (2014) to affect the quality of the CRMP when combined with the Koval model.

Joint Optimization. The results comparing the joint optimization of Model 1 assisted by CRMPs (**Algorithm 1**) and the approach using only full-physics simulations are shown in **Fig. 10**. To allow for a clearer comparison, the results are plotted on a semilog (x -axis) plot. Both implementations had a stopping criterion of 50 PSO iterations and the well control optimization (inner loop) of the full-physics implementation had a maximum of 100 function evaluations (for each particle). For the full-physics implementation, this results in a maximum number of function evaluations of 100,000. For Model 1, an R^2 threshold value (i.e., Line 8 in **Algorithm 1**) of 0.95 and a value of 5 for Max_{CRMP} are used.

The left plot in **Fig. 10** shows the results for the joint optimization when using three different seeds. Both plots in **Fig. 10** display the best solution found in each generation of PSO. As such, the plots do not cross the x -axis. This was done to ensure the results were robust to the stochastic nature of the optimization algorithms used. For each seed, the CRMP-based optimization and the full-physics optimization begin from the same random population. From the results, it is evident that there is a significant reduction in computational cost of joint optimization when using CRMPs as surrogates.

Furthermore, this does not come at the expense of the quality of the solution found. The best solution (using Seed 3) found by the CRMP-based method is 49.42 million USD after only 1,854 full-physics simulations. On the other hand, the best solution found by the full-physics implementation is 48.68 million USD after 100,000 full-physics simulations. This represents a reduction of computational cost (in terms of the number of required full-physics simulations) of more than two orders of magnitude with an improvement in NPV of more than 1%. With respect to NPV, even the worst performing CRMP-based run (Seed 1) obtained an optimum NPV of 48.08 million USD, which is competitive with the NPV of the best run of the full-physics-based approach.

The right plot in **Fig. 10** shows the average of the three runs for the two implementations. On average, the CRMP-based approach obtained an NPV value of 48.76 million USD after 2,126 full-physics simulations. On the other hand, on average, the full-physics approach obtained an NPV value of 47.37 million USD after 100,000 full-physics simulations. From a computational efficiency standpoint, the CRMP-based approach required, on average, only 459 full-physics simulations to obtain the optimal NPV (47.37 MM USD) found by the full-physics approach. This translates to a reduction in the number of required high fidelity grid-based simulations by up to 3 factors of 10.

Fig. 11 shows the well locations for the optimal solutions found by the CRMP-based (left) and the full-physics (right) approaches. There are a number of similarities between the well locations of the two solutions. For instance, in both solutions, the injection wells are placed on the boundary of the reservoir model. This delays the water breakthrough at the producers, which is in line with typical reservoir engineering judgment. Also, the solutions both place a production well (specifically, PROD1) in a similar location within the southern section of the reservoir model, in a high permeability streak. This shows that the joint optimization using CRMP proposed well locations which are in line with reservoir engineering concepts and are feasible for the given field constraint (an interwell distance of 600 ft).

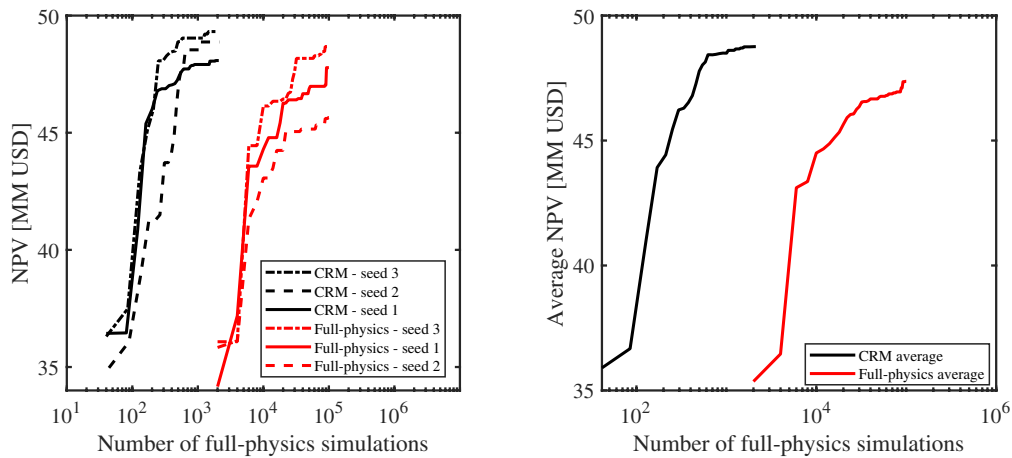


Fig. 10—Left: Convergence plots for the joint optimization of Model 1 comparing CRMP and full-physics-based approaches for the three seeds. Right: Average of the three seeds for joint optimization of Model 1 using CRMP and full-physics simulations.

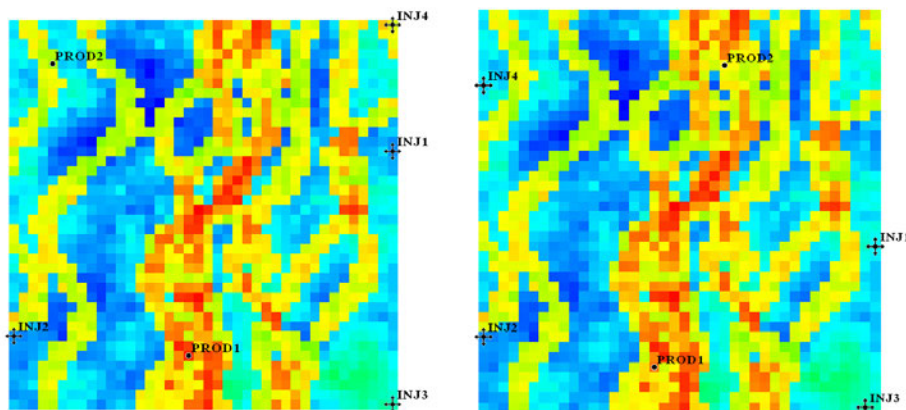


Fig. 11—Left: Optimal solution found by CRMP-based joint optimization showing associated well locations. Right: Optimal solution found by full-physics joint optimization showing associated well locations.

The significant improvement in computational efficiency and resilience of the CRMP-based approach to obtain improvements in NPV is a reflection of its practicality for its given fidelity. The fidelity of any surrogate is expected to be lower than the full-physics simulation; however, its functionality (in this case for joint optimization) is more pertinent than the absolute difference. In other words, the attractiveness of a surrogate is a balance between its fidelity and computational efficiency. If a surrogate is very computationally efficient (i.e., a significantly lower runtime than the full-physics model) but its fidelity is too low, then its usefulness decreases. However, the results presented here show that the fidelity of the CRMP is sufficient to take advantage of the reduction in computational costs without the loss of accuracy (which would be reflected in lower optimal NPV).

Fig. 12 provides insight into the fidelity of the CRMPs during joint optimization for each of the three seeds. The y -axis shows the absolute average percent difference between $f(\mathbf{x}_j, \mathbf{u}_i^*)$ and $f(\mathbf{x}_j, \mathbf{u}_i^*, \mathbf{p}_j^*)$ (see Line 11 in **Algorithm 1**). The difference shown is the average absolute difference across the particles in the swarm (a swarm size of 20 is used). Although there are varying values for the 25th and 75th percentiles, the median values are relatively similar. Seeds 1, 2, and 3 have median values of 2.5, 5.1, and 6.7%, respectively. While there are outliers to these averages (e.g., a maximum of 57% for Seed 3 is not shown) which can be associated to the stochastic nature of the algorithm, these are relatively low values given the computational savings achieved.

Experimental Results: Model 2. Evaluating CRM. To build an example CRMP for Model 2, the well locations displayed in **Fig. 5** were used. Additionally, for illustration, the training sample of random water injection rates for INJ1 is presented in **Fig. 13a**. For Model 2, the BHPs of the producers were assumed to be constant and as such Eq. 10 was used to tune the CRMP parameters. This means there was a total of 44 model parameters which were the interwell connectivities (λ_{ij}), the time constant for each producer (τ_j), the Koval factor for each producer (K_{valj}), and the movable PV (V_{pj}) for each producer.

The output of the CRMP (using the tuned parameters) and the full-physics simulation is compared in **Fig. 14**. For conciseness, only the oil production rates of PROD1 are shown. The coefficient of determination (R^2) for each of the producers is 0.9787, 0.9236, 0.9789, and 0.9645 for PROD1, PROD2, PROD3, and PROD4, respectively. These high values, in addition to the visual inspection of **Fig. 14**, provide promising results for the application of CRMP for a 3D model. Furthermore, the average absolute percent difference in the resulting NPV is only 0.38%. Although this is only one development strategy (well control and locations), it provides an indication to the potential reduction of computational cost that can be achieved without consequential reduction in solution quality.

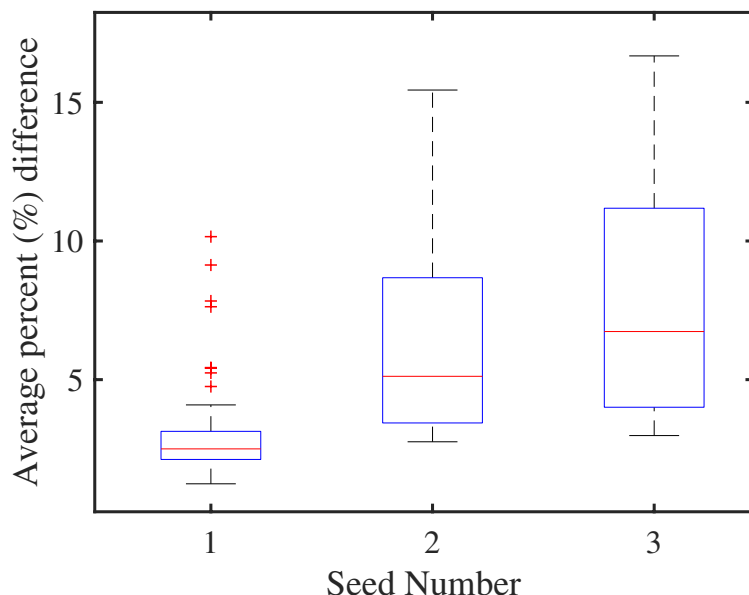


Fig. 12—Box plots presenting the average of the absolute percent difference between $f(x_j, u_i^*)$ and $f(x_j, u_i^*, p_j^*)$ for each of the three seeds. The red line within each box represents the median, and the bottom and top lines of each box represent the 25th and 75th percentiles, respectively. The lines above and below the boxes represent the most extreme data points which are not considered outliers. The outliers are plotted individually using the cross symbol in a red color.

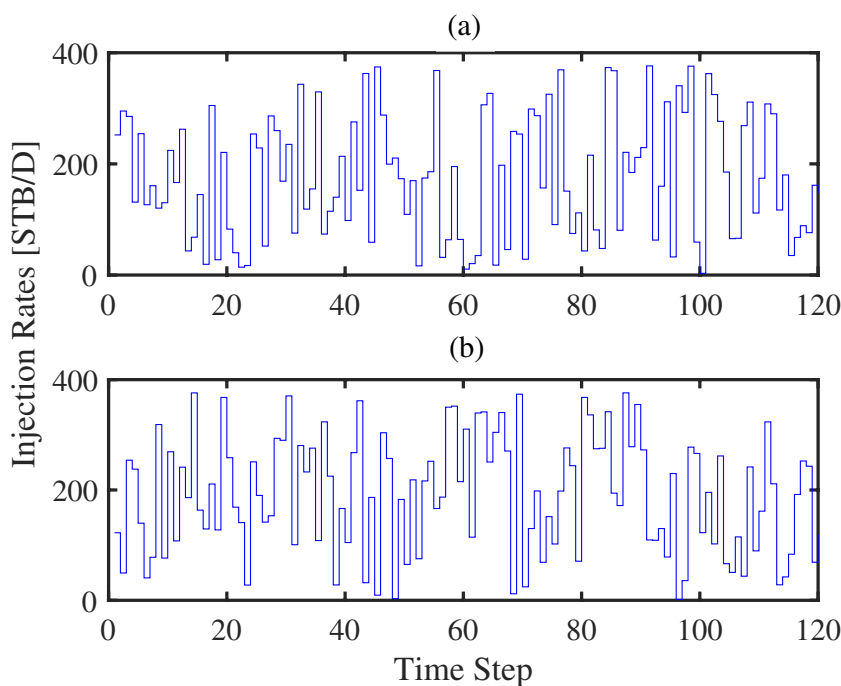


Fig. 13—The (a) training sample and (b) validation sample of random water injection rates for INJ1 used to tune and then test the CRMP of Model 2.

As Model 2 is a 3D model with a larger number of wells than Model 1, the runtime is significantly more. The simulation runtime of Model 2 was around 3 minutes. In comparison, the runtime of the CRMP was less than 0.01 seconds. This represents a computational saving of more than 4 factors of 10 with only an absolute average percent difference of 0.38% in NPV. Although the percent difference may not be expected to be this low throughout the joint optimization, the potential reduction in computation is substantial.

Similar to what was done for Model 1, an ensemble of 100 validation samples of random water injection rates was generated to validate the CRMP's ability to forecast production rates for comparison to that of the full-physics simulation. An example of the validation sample of random water injection rates for INJ1 is shown in **Fig. 13b**. The resulting oil production rates are presented in **Fig. 14b** for PROD1. The associated values for the coefficient of determination (R^2) for the production wells are 0.9269, 0.8493, 0.9652, and 0.8420 for PROD1, PROD2, PROD3, and PROD4, respectively.

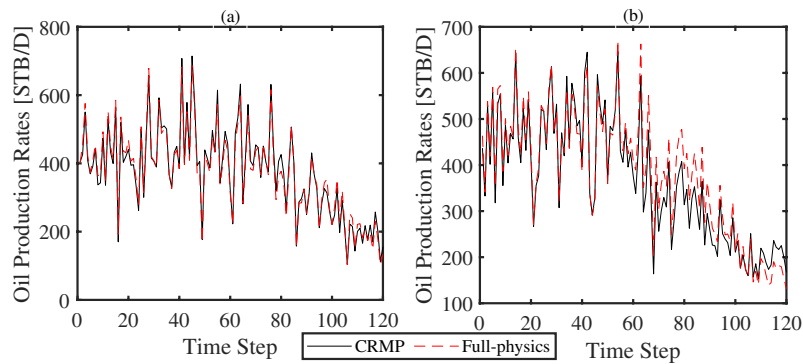


Fig. 14—Oil production rates of PROD1 comparing CRMP and full-physics simulations when using a training sample (left) and a validation sample (right).

The average R^2 values across the ensemble of 100 validation samples produced by the CRMP for the oil production rates for PROD1, PROD2, PROD3, and PROD4 are 0.9618, 0.8579, 0.9691, and 0.9057, respectively. The associated standard deviation of these values across the 100 validation samples points are 0.0178, 0.0336, 0.0092, and 0.0370 for PROD1, PROD2, PROD3, and PROD4, respectively. Again, this reflects sufficiently large R^2 values for predicting oil production rates for random water injection rates. However, more importantly for this study, is how the CRMP will perform when combined in a bilevel approach as proposed here. This will be investigated in the following section in more detail when the proposed approach is applied to the joint optimization of Model 2.

Joint Optimization. The results of the joint optimization of Model 2 using the CRMP-based method and the full-physics simulation approach are presented in Fig. 15. Similar to the results presented in Fig. 10, a semilog (x -axis) plot is used to allow a clear comparison. As before, both plots in Fig. 10 display the best solution found in each generation of PSO. As such, the plots do not cross the x -axis. Given the runtime (around 3 minutes or more) of the full-physics simulation for Model 2, the stopping criterion for this problem was 25 PSO iterations. The inner well control optimization loop had a maximum of 100 function evaluations for the full-physics approach given the significantly longer runtime compared with Model 1. For the full-physics implementation, this results in a maximum number of function evaluations of 50,000. For Model 2, a R^2 threshold value (i.e., Line 8 in Algorithm 1) of 0.90 and a value of 10 Max_{CRMP} were used.

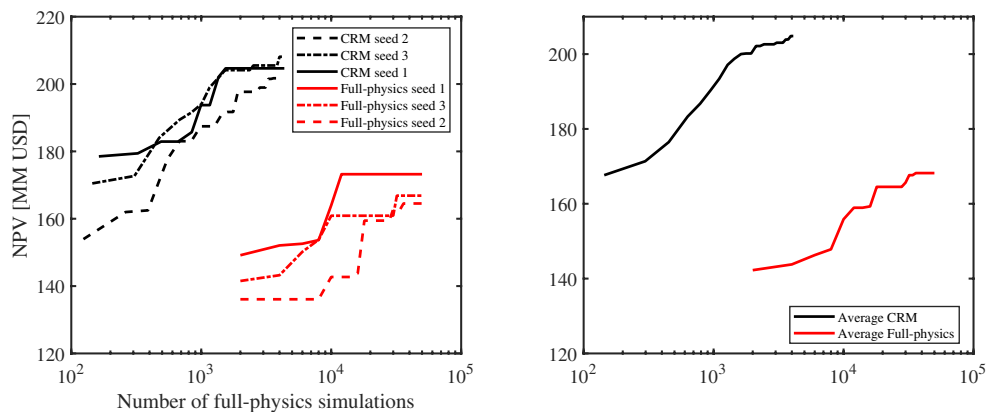


Fig. 15—Left: Convergence plots for the joint optimization of Model 2 comparing CRMP and full-physics-based approaches for the three seeds. Right: Average of the three seeds for joint optimization of Model 2 using CRMP and full-physics simulations.

The plot on the left shows the results for each of the three random seeds for both the CRMP-based optimization and the optimization using only full-physics simulations. The CRMP-based optimization significantly outperforms the full-physics optimization for all three seeds. Although marginal, Seed 3 resulted in the best NPV for the CRMP-based optimization compared with Seeds 1 and 2. Seed 3 resulted in an NPV of 208.1 million USD after 4,175 full-physics simulations. In comparison, the best result (Seed 1) for the full-physics simulations is an NPV of 173.2 MM USD after 12,000 full-physics simulations. This translates to a 16.77% improvement in NPV.

Similar significant improvements are seen from a computational savings standpoint. For example, the CRMP-based optimization obtained an NPV of 178.5 million USD after only 168 full-physics simulations. This is almost 2 factors of 10 decrease in computational intensity without a decrease in the quality of the solution found.

Similar performance of CRMP-based optimization is seen when looking at the average results of the three seeds in the right plot of Fig. 15. On average, the CRMP-based optimization found an optimal solution of 204.8 million USD after 4,103 full-physics simulations. On the other hand, the full-physics optimizations found, on average, an optimal solution with an NPV of 168.2 million USD after 36,000 full-physics simulations. This translates to an improvement of 21.8% in NPV using a magnitude less of full-physics simulations.

From a computational efficiency standpoint, the CRMP-based optimization required, on average, only 144 full-physics simulations to find a solution with an NPV of 167.6 million USD. On the other hand, the full-physics optimization required, on average, 36,000 full-physics simulations. This translates to a reduction of more than 2 factors of 10 in computational costs. In other words, if a runtime of 3 minutes is taken for each full-physics simulation then a reduction of 35,856 full-physics simulations translates to 90 hours or almost 4

days of computation time. This is significant in considering the need for continuous access to software licenses, which may not be available or are shared across multiple teams and departments.

One may attempt to improve the full-physics approach by reducing the maximum number of allowed function evaluations in the inner loop of well control optimization. However, it would be expected that the quality of the optimal solution obtained would also decrease as the algorithm may not converge to a suitable solution within the limited number of function evaluations. From previous studies, namely in Arouri and Sayyafzadeh (2020a, 2022), it was found that Adam-SPSA is able to accelerate to an improved solution within the first 250 function evaluations. As such, this was used as the basis for the maximum number of function evaluations in this work. This further highlights the advantages of using CRMP to assist the joint optimization by significantly reducing the number of full-physics simulations without compromising the quality of the solution. One should note that in the scenario where an unlimited amount of resources is available (i.e., time, human, and software resources), it is expected that using the full-physics simulation would ultimately result in a higher NPV.

Fig. 16 shows the well locations of the optimal solutions found by the CRMP-based approach (left) and the full-physics approach (right). There are a couple of similarities between the two. For example, the placements of PROD2, PROD3, and PROD4 are similar. PROD1 in both solutions is placed closer to the injection wells. A sensitivity was undertaken to manually move PROD1 away from the injectors to see if any improvements in NPV results. However, the solution found by the proposed approach was still the highest NPV. Apart from INJECT8 in the CRMP-based solution, the injection wells in both solutions are on the eastern flank of the model. This

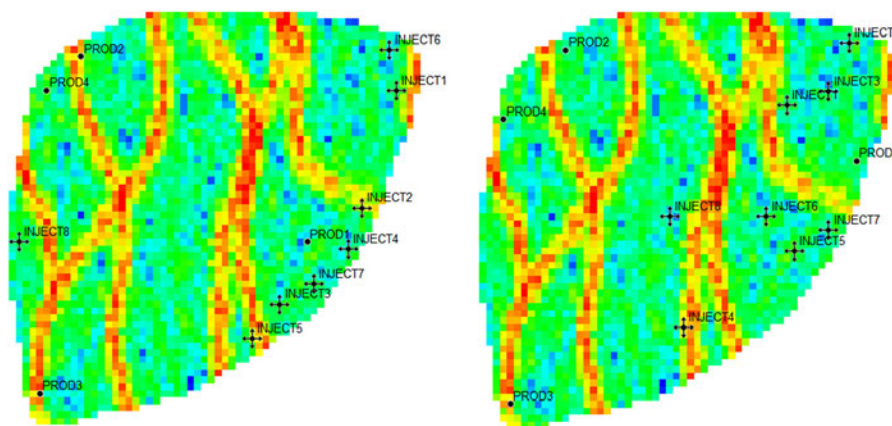


Fig. 16—Left: Optimal well placements found by the proposed CRMP-based optimization method. Right: Optimal well locations found by the full-physics approach. Producers represented by black circles, and injectors represented by cross symbols. Red indicates high permeability values, and blue indicates low permeability values.

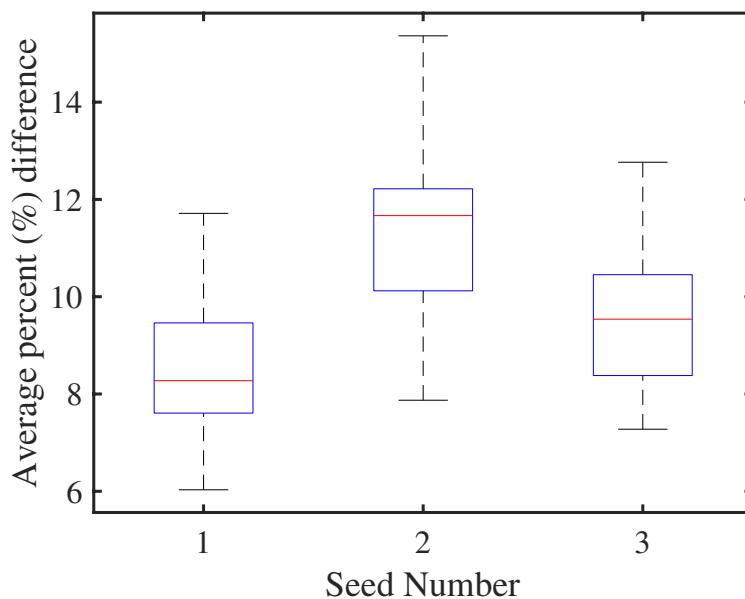


Fig. 17—Box plots presenting the average of the absolute percent difference between $f(x_j, u_i^*)$ and $f(x_j, u_i^*, p_j^*)$ for each of the three seeds. The red line within each box represents the median, and the bottom and top lines of each box represent the 25th and 75th percentiles, respectively. The lines above and below the boxes represent the most extreme data points which are not considered outliers. The outliers are plotted individually using the cross symbol in a red color.

represents waterflooding direction from the eastern side to the western side where the production wells are located. In such a well placement scenario, the injection rates would be pivotal in attempting to ensure a high sweep efficiency.

Fig. 17 shows the absolute average percent difference between $f(\mathbf{x}_j, \mathbf{u}_i^*)$ and $f(\mathbf{x}_j, \mathbf{u}_i^*, \mathbf{p}_i^*)$ for each of the three seeds. The median values are 8.3, 11.7, and 9.5% for Seeds 1, 2, and 3, respectively. The maximum absolute average percent difference in all three seeds is only 15.4% (Seed 2) and a minimum of 6.0% (Seed 1). As one would expect, for a more larger complex model, the median percent differences are larger than those seen in Model 1. These results coupled with the results presented in Fig. 15 show that the fidelity of the CRMP is sufficiently high enough to reap the significant computational benefits when used in a joint optimization problem.

Concluding Remarks

We proposed a CRMP-assisted approach for the joint optimization of well controls and well locations for general field development problems. The proposed method uses a bilevel implementation with the outer loop being the well placement problem solved by PSO and the inner loop being the well control problem solved by Adam-SPSA. In past work, only full-physics simulations have been used to evaluate the proposed general field development plans; however, this results in an intractable computational cost. To ameliorate this issue, we proposed and investigated the use of (producer-based) CRMs as a computationally efficient alternative. Specifically, the CRMPs were used within the inner loop of the optimization method which consisted of the well control problem. For each proposed well location, a CRMP was built and thereafter optimized to find the optimal well control settings.

The proposed technique was tested on two example problems of varying complexity. The first problem was a heterogeneous 2D reservoir model with four injection wells and two production wells. The second problem used the 3D egg model populated with eight injection wells and four production wells. For each model, we investigated the fidelity of the CRMP before implementing the proposed approach for joint optimization.

The results of the proposed method were compared against an implementation which used only full-physics simulations (as has typically been done in previous literature). To investigate the accuracy of CRMPs, we tuned the models using a sample of random well control settings (injection rates and/or BHPs). The tuned CRMP parameters were then used to validate the accuracy of the CRMP using a validation sample of well control settings. For both models, the validation resulted in coefficient of determination values above 0.92.

Promising results were also obtained when the proposed CRMP-based approach was used for the joint optimization of well controls and well locations. When implemented for Model 1, the CRMP-based approach resulted, on average, in computational savings of up to 3 factors of 10 in the number of required full-physics simulations. More importantly, the computational reduction did not come at the cost of the optimal solution value (NPV).

Similar positive results were obtained when the CRMP-based approach was used for the joint optimization of Model 2. On average, the proposed method found solutions of up to 21.8% higher than the optimal found by the full-physics approach. In addition, on average, the computational savings were nearly 3 factors of 10. For the runtime of 3 minutes (on average) for Model 2, this translates to a runtime savings of almost 4 days. This is significant when considering the potential cost savings associated with software licensing and the likes.

We also investigated the fidelity of the CRMP when used within joint optimization. The fidelity of the CRMPs for Model 1 varied between 1.2 and 5% (apart from the first two iterations), while for Model 2 it varied between 8 and 15%. When considered in conjunction with the results from the joint optimization this indicates that these levels of fidelity are sufficient to encourage the use of CRMP without compromising the performance with respect to NPV.

The work in this study can be furthered through a number of avenues. The incorporation of geological uncertainty is an important aspect of general field development. To incorporate geological uncertainty in the CRMP-based approach, one will simply need to build CRMPs for each realization at each well location. We would expect the computational savings to be similar, if not greater than, those found for the deterministic case as the computational cost increases proportionally with the number of realizations considered. Another avenue of potential research could investigate the use of efficient surrogates for the outer well placement loop. It would be intriguing to investigate the interplay between the inner and outer loop, including the required fidelity of each to reap the computational benefits without compromising the quality of the solution. Additionally, investigations into combinations of different optimization algorithms for the outer and inner loops could add value to literature.

Acknowledgments

The authors would like to thank Rock Flow Dynamics for providing the academic licenses for tNavigator. This work was supported with supercomputing resources provided by the Phoenix HPC service at the University of Adelaide. Yazan would also like to acknowledge the funding and support received from Fulbright Australia under the Fulbright Future Scholarship to undertake research at The University of Texas at Austin.

References

- Alavian, S. A., Whitson, C. H., and Martinsen, S. O. 2014. Global Component Lumping for EOS Calculations. Paper presented at the *SPE Annual Technical Conference and Exhibition*, Amsterdam, The Netherlands, 27–29 October. SPE-170912-MS. <https://doi.org/doi:10.2118/170912-MS>.
- Albertoni, A. and Lake, L. W. 2003. Inferring Interwell Connectivity Only From Well-Rate Fluctuations in Waterfloods. *SPE Res Eval & Eng* 6 (1): 6–16. SPE-83381-PA. <https://doi.org/doi:10.2118/83381-PA>.
- Alghareeb, Z. M. 2015. *Optimal Reservoir Management Using Adaptive Reduced-Order Models*. PhD dissertation, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA.
- Aliyev, E. and Durlafsky, L. J. 2017. Multilevel Field Development Optimization Under Uncertainty Using a Sequence of Upscaled Models. *Math Geosci* 49 (3): 307–339. <https://doi.org/doi:10.1007/s11004-016-9643-0>.
- Alrashdi, Z. and Sayyafzadeh, M. 2019. M+ Λ Evolution Strategy Algorithm in Well Placement, Trajectory, Control and Joint Optimisation. *J Pet Sci Eng* 177: 1042–1058. <https://doi.org/doi:10.1016/j.petrol.2019.02.047>.
- Arouri, Y. and Sayyafzadeh, M. 2020a. An Accelerated Gradient Algorithm for Well Control Optimization. *J Pet Sci Eng* 190: 106872. <https://doi.org/doi:10.1016/j.petrol.2019.106872>.
- Arouri, Y. and Sayyafzadeh, M. 2020b. Adaptive Moment Estimation Framework for Well Placement Optimization. In *ECMOR XVII*, 1–15. The Netherlands: European Association of Geoscientists & Engineers.
- Arouri, Y. and Sayyafzadeh, M. 2022. An Adaptive Moment Estimation Framework for Well Placement Optimization. *Comput Geosci* (open access; available online 19 February 2022). <https://doi.org/doi:10.1007/s10596-022-10135-9>.
- Arouri, Y., Sayyafzadeh, M., and Begg, S. 2022. Adaptive Rank-Based Selection of Geological Realizations for Optimum Field Development Planning. *SPE J*. SPE-209584-PA (in press; posted 7 March 2022). <https://doi.org/doi:10.2118/209584-PA>.

- Artus, V., Durllofsky, L. J., Onwunali, J. et al. 2006. Optimization of Nonconventional Wells under Uncertainty Using Statistical Proxies. *Comput Geosci* **10** (4): 389–404. <https://doi.org/doi:10.1007/s10596-006-9031-9>.
- Awotunde, A. A. 2014. On the Joint Optimization of Well Placement and Control. Paper presented at the *SPE Saudi Arabia Section Annual Technical Symposium and Exhibition*, Al-Khobar, Saudi Arabia, 21–24 April. SPE-172206-MS. <https://doi.org/doi:10.2118/172206-MS>.
- Batycky, R. P., Blunt, M. J., and Thiele, M. R. 1997. A 3D Field-Scale Streamline-Based Reservoir Simulator. *SPE Res Eng* **12** (4): 246–254. SPE-36726-PA. <https://doi.org/doi:10.2118/36726-PA>.
- Beckner, B. L. and Song, X. 1995. Field Development Planning Using Simulated Annealing - Optimal Economic Well Scheduling and Placement. Paper presented at the *SPE Annual Technical Conference and Exhibition*, Dallas, Texas, USA, 22–25 October. SPE-30650-MS. <https://doi.org/doi:10.2118/30650-MS>.
- Bellout, M. C., Echeverría Ciaurri, D., Durllofsky, L. J. et al. 2012. Joint Optimization of Oil Well Placement and Controls. *Comput Geosci* **16** (4): 1061–1079. <https://doi.org/doi:10.1007/s10596-012-9303-5>.
- Bouzarkouna, Z., Ding, D. Y., and Auger, A. 2012. Well Placement Optimization with the Covariance Matrix Adaptation Evolution Strategy and Meta-Models. *Comput Geosci* **16** (1): 75–92. <https://doi.org/doi:10.1007/s10596-011-9254-2>.
- Cao, F. 2014. *Development of a Two-Phase Flow Coupled Capacitance Resistance Model*. PhD dissertation, The University of Texas at Austin, Austin, Texas.
- Cao, F., Luo, H., and Lake, L. W. 2014. Development of a Fully Coupled Two-Phase Flow Based Capacitance Resistance Model CRM. Paper presented at the *SPE Improved Oil Recovery Symposium*, Tulsa, Oklahoma, USA, 12–16 April. SPE-169485-MS. <https://doi.org/doi:10.2118/169485-MS>.
- Cao, F., Luo, H., and Lake, L. W. 2015. Oil-Rate Forecast by Inferring Fractional-Flow Models From Field Data With Koval Method Combined With the Capacitance/Resistance Model. *SPE Res Eval & Eng* **18** (4): 534–553. SPE-173315-PA. <https://doi.org/doi:10.2118/173315-PA>.
- Cardoso, M. A., Durllofsky, L. J., and Sarma, P. 2009. Development and Application of Reduced-Order Modeling Procedures for Subsurface Flow Simulation. *Int J Numer Methods Eng* **77** (9): 1322–1350. <https://doi.org/doi:10.1002/nme.2453>.
- Chaudhri, M. M., Phale, H. A., Liu, N. et al. 2009. An Improved Approach for Ensemble-Based Production Optimization. Paper presented at the *SPE Western Regional Meeting*, San Jose, California, USA, 24–26 March. SPE-121305-MS. <https://doi.org/doi:10.2118/121305-MS>.
- Chen, Y., Oliver, D. S., and Zhang, D. 2009. Efficient Ensemble-Based Closed-Loop Production Optimization. *SPE J.* **14** (4): 634–645. SPE-112873-PA. <https://doi.org/doi:10.2118/112873-PA>.
- Christian, W., Durllofsky, L. J., and Aziz, K. 2000. Approximate Model for Productivity of Nonconventional Wells in Heterogeneous Reservoirs. *SPE J.* **5** (2): 218–226. SPE-62812-PA. <https://doi.org/doi:10.2118/62812-PA>.
- Cullick, A. S., Johnson, W. D., and Shi, G. 2006. Improved and More Rapid History Matching With a Nonlinear Proxy and Global Optimization. Paper presented at the *SPE Annual Technical Conference and Exhibition*, San Antonio, Texas, USA, 24–27 September. SPE-101933-MS. <https://doi.org/doi:10.2118/101933-MS>.
- Datta-Gupta, A. and King, M. 2007. *Streamline Simulation Theory and Practice*. Richardson, Texas, USA: Society of Petroleum Engineers.
- De Brito, D. U. and Durllofsky, L. J. 2020. Field Development Optimization Using a Sequence of Surrogate Treatments. *Comput Geosci* **25** (1): 35–65. <https://doi.org/doi:10.1007/s10596-020-09985-y>.
- Ding, Y. 2008. Optimization of Well Placement Using Evolutionary Methods. Paper presented at the *Europec/EAGE Conference and Exhibition*, Rome, Italy, 9–12 June. SPE-113525-MS. <https://doi.org/doi:10.2118/113525-MS>.
- Do, S. T. and Reynolds, A. C. 2013. Theoretical Connections between Optimization Algorithms Based on an Approximate Gradient. *Comput Geosci* **17** (6): 959–973. <https://doi.org/doi:10.1007/s10596-013-9368-9>.
- Echeverría Ciaurri, D. and Wagenaar, C. 2016. Manifold-Mapping Optimization Applied to Oil Field Operations. In *ECMOR XV - 15th European Conference on the Mathematics of Oil Recovery*. The Netherlands: European Association of Geoscientists & Engineers. <https://doi.org/doi:10.3997/2214-4609.201601827>.
- Echeverría Ciaurri, D., Isebor, O. J., and Durllofsky, L. J. 2010. Application of Derivative-Free Methodologies to Generally Constrained Oil Production Optimization Problems. *Procedia Comput Sci* **1** (1): 1301–1310. <https://doi.org/doi:10.1016/j.procs.2010.04.145>.
- Emerick, A. A., Silva, E., Messer, B. et al. 2009. Well Placement Optimization Using a Genetic Algorithm With Nonlinear Constraints. Paper presented at the *SPE Reservoir Simulation Symposium*, The Woodlands, Texas, USA, 2–4 February. SPE-118808-MS. <https://doi.org/doi:10.2118/118808-MS>.
- Engelbrecht, A. P. 2013. Particle Swarm Optimization: Global Best or Local Best? In *2013 BRICS Congress on Computational Intelligence and 11th Brazilian Congress on Computational Intelligence*. Piscataway, New Jersey: IEEE Xplore.
- Fonseca, R. M., Stordal, A. S., Leeuwenburgh, O. et al. 2014. Robust Ensemble-Based Multi-Objective Optimization. In *ECMOR XIV - 14th European Conference on the Mathematics of Oil Recovery*. The Netherlands: European Association of Geoscientists & Engineers. <https://doi.org/doi:10.3997/2214-4609.20141895>.
- Fonseca, R. R.-M., Chen, B., Jansen, J. D. et al. 2017. A Stochastic Simplex Approximate Gradient (StoSAG) for Optimization under Uncertainty. *Int J Numer Methods Eng* **109** (13): 1756–1776. <https://doi.org/doi:10.1002/nme.5342>.
- Foroud, T., Baradaran, A., and Seifi, A. 2018. A Comparative Evaluation of Global Search Algorithms in Black Box Optimization of Oil Production: A Case Study on Brugge Field. *J Pet Sci Eng* **167**: 131–151. <https://doi.org/doi:10.1016/j.petrol.2018.03.028>.
- Forouzanfar, F., Poquioma, W. E., and Reynolds, A. C. 2016. Simultaneous and Sequential Estimation of Optimal Placement and Controls of Wells With a Covariance Matrix Adaptation Algorithm. *SPE J.* **21** (2): 501–521. SPE-173256-PA. <https://doi.org/doi:10.2118/173256-PA>.
- Guyaguler, B., Horne, R. N., Rogers, L. et al. 2000. Optimization of Well Placement in a Gulf of Mexico Waterflooding Project. Paper presented at the *SPE Annual Technical Conference and Exhibition*, Dallas, Texas, USA, 1–4 October. SPE-63221-MS. <https://doi.org/doi:10.2118/63221-MS>.
- Holanda, R. W. D., Gildin, E., Jensen, J. L. et al. 2018. A State-of-the-Art Literature Review on Capacitance Resistance Models for Reservoir Characterization and Performance Forecasting. *Energies (Basel)* **11** (12): 3368. <https://doi.org/doi:10.3390/en11123368>.
- Hong, A. J., Bratvold, R. B., and Nævdal, G. 2017. Robust Production Optimization with Capacitance-Resistance Model as Precursor. *Comput Geosci* **21** (5–6): 1423–1442. <https://doi.org/doi:10.1007/s10596-017-9666-8>.
- Humphries, T. D. and Haynes, R. D. 2015. Joint Optimization of Well Placement and Control for Nonconventional Well Types. *J Pet Sci Eng* **126**: 242–253. <https://doi.org/doi:10.1016/j.petrol.2014.12.016>.
- Humphries, T. D., Haynes, R. D., and James, L. A. 2014. Simultaneous and Sequential Approaches to Joint Optimization of Well Placement and Control. *Comput Geosci* **18** (3–4): 433–448. <https://doi.org/doi:10.1007/s10596-013-9375-x>.
- Isebor, O. J., Durllofsky, L. J., and Echeverría Ciaurri, D. 2014. A Derivative-Free Methodology with Local and Global Search for the Constrained Joint Optimization of Well Locations and Controls. *Comput Geosci* **18** (3–4): 463–482. <https://doi.org/doi:10.1007/s10596-013-9383-x>.
- Jafroodi, N. and Zhang, D. 2011. New Method for Reservoir Characterization and Optimization Using CRM–EnOpt Approach. *J Pet Sci Eng* **77** (2): 155–171. <https://doi.org/doi:10.1016/j.petrol.2011.02.011>.
- Jansen, J. D. 2011. Adjoint-Based Optimization of Multi-Phase Flow through Porous Media – A Review. *Comput Fluids* **46** (1): 40–51. <https://doi.org/doi:10.1016/j.compfluid.2010.09.039>.

- Jansen, J. D. and Durlafsky, L. J. 2017. Use of Reduced-Order Models in Well Control Optimization. *Optim Eng* **18** (1): 105–132. <https://doi.org/doi:10.1007/s11081-016-9313-6>.
- Jansen, J. D., Fonseca, R. M., Kahrobaei, S. et al. 2014. The Egg Model - a Geological Ensemble for Reservoir Simulation. *Geosci Data J* **1** (2): 192–195. <https://doi.org/doi:10.1002/gdj3.21>.
- Kennedy, J. and Eberhart, R. 1995. Particle Swarm Optimization. In *ICNN'95 - International Conference on Neural Networks*. Piscataway, New Jersey: IEEE. <https://doi.org/doi:10.1109/ICNN.1995.488968>.
- Kingma, D. P. and Ba, J. 2014. Adam: A Method for Stochastic Optimization. arXiv:1412.6980. <https://arxiv.org/abs/1412.6980>(preprint; last revised 30 January 2017).
- Leeuwenburgh, O., Egberts, P. J. P., and Abbink, O. A. 2010. Ensemble Methods for Reservoir Life-Cycle Optimization and Well Placement. Paper presented at the *SPE/DGS Saudi Arabia Section Technical Symposium and Exhibition*, Al-Khobar, Saudi Arabia, 4–7 April. SPE-136916-MS. <https://doi.org/doi:10.2118/136916-MS>.
- Li, B. and Friedmann, F. 2005. Novel Multiple Resolutions Design of Experiment/Response Surface Methodology for Uncertainty Analysis of Reservoir Simulation Forecasts. Paper presented at the *SPE Reservoir Simulation Symposium*, The Woodlands, Texas, USA, 31 January–2 February. SPE-92853-MS. <https://doi.org/doi:10.2118/92853-MS>.
- Li, L. and Jafarpour, B. 2012. A Variable-Control Well Placement Optimization for Improved Reservoir Development. *Comput Geosci* **16** (4): 871–889. <https://doi.org/doi:10.1007/s10596-012-9292-4>.
- Li, L., Jafarpour, B., and Mohammad-Khaninezhad, M. R. 2013. A Simultaneous Perturbation Stochastic Approximation Algorithm for Coupled Well Placement and Control Optimization under Geologic Uncertainty. *Comput Geosci* **17** (1): 167–188. <https://doi.org/doi:10.1007/s10596-012-9323-1>.
- Liu, Q., Wei, W., Yuan, H. et al. 2016. Topology Selection for Particle Swarm Optimization. *Inf Sci* **363**: 154–173. <https://doi.org/doi:10.1016/j.ins.2016.04.050>.
- Liu, Z. and Reynolds, A. C. 2020. A Sequential-Quadratic-Programming-Filter Algorithm with A Modified Stochastic Gradient for Robust Life-Cycle Optimization Problems with Nonlinear State Constraints. *SPE J* **25** (4): 1938–1963. SPE-193925-PA. <https://doi.org/doi:10.2118/193925-PA>.
- Liang, X., Lake, L. W., Edgar, T. F. et al. 2007. Optimization of Oil Based on a Capacitance Model of Production and Injection Rates. Paper presented at the *Hydrocarbon Economics and Evaluation Symposium*, Dallas, Texas, USA, 1–3 April. SPE-107713-MS. <https://doi.org/doi:10.2118/107713-MS>.
- Lu, R., Forouzanfar, F., and Reynolds, A. C. 2017a. Bi-Objective Optimization of Well Placement and Controls Using StoSAG. Paper presented at the *SPE Reservoir Simulation Conference*, Montgomery, Texas, USA, 20–22 February. SPE-182705-MS. <https://doi.org/doi:10.2118/182705-MS>.
- Lu, R., Forouzanfar, F., and Reynolds, A. C. 2017b. An Efficient Adaptive Algorithm for Robust Control Optimization Using StoSAG. *J Pet Sci Eng* **159**: 314–330. <https://doi.org/doi:10.1016/j.petrol.2017.09.002>.
- Lu, R. and Reynolds, A. C. 2020. Joint Optimization of Well Locations, Types, Drilling Order, and Controls Given a Set of Potential Drilling Paths. *SPE J* **25** (3): 1285–1306. SPE-193885-PA. <https://doi.org/doi:10.2118/193885-PA>.
- Møyner, O., Krogstad, S., and Lie, K.-A. 2014. The Application of Flow Diagnostics for Reservoir Management. *SPE J* **20** (2): 306–323. SPE-171557-PA. <https://doi.org/doi:10.2118/171557-PA>.
- Nocedal, J. and Wright, S. 2006. *Numerical Optimization*, second edition, Vol. 35. Berlin/Heidelberg, Germany: Springer.
- Oguntola, M. B. and Lorentzen, R. J. 2021. Ensemble-Based Constrained Optimization Using an Exterior Penalty Method. *J Pet Sci Eng* **207**: 109165. <https://doi.org/doi:10.1016/j.petrol.2021.109165>.
- Onwunali, J. E. and Durlafsky, L. J. 2010. Application of a Particle Swarm Optimization Algorithm for Determining Optimum Well Location and Type. *Comput Geosci* **14** (1): 183–198. <https://doi.org/doi:10.1007/s10596-009-9142-1>.
- Pan, Y. and Horne, R. N. 1998. Improved Methods for Multivariate Optimization of Field Development Scheduling and Well Placement Design. Paper presented at the *SPE Annual Technical Conference and Exhibition*, New Orleans, Louisiana, USA, 27–30 September. SPE-49055-MS. <https://doi.org/doi:10.2118/49055-MS>.
- Peng, C. Y. and Gupta, R. 2004. Experimental Design and Analysis Methods in Multiple Deterministic Modelling for Quantifying Hydrocarbon In-Place Probability Distribution Curve. Paper presented at the *SPE Asia Pacific Conference on Integrated Modelling for Asset Management*, Kuala Lumpur, Malaysia, 29–30 March. SPE-87002-MS. <https://doi.org/doi:10.2118/87002-MS>.
- Perrone, A. and Rossa, E. D. 2015. Optimizing Reservoir Life-Cycle Production Under Uncertainty: A Robust Ensemble-Based Methodology. Paper presented at the *SPE Reservoir Characterisation and Simulation Conference and Exhibition*, Abu Dhabi, UAE, 14–16 September. SPE-175570-MS. <https://doi.org/doi:10.2118/175570-MS>.
- Salehian, M., Sefat, M. H., and Muradov, K. 2021. Robust Integrated Optimization of Well Placement and Control under Field Production Constraints. *J Pet Sci Eng* **205**: 108926. <https://doi.org/doi:10.1016/j.petrol.2021.108926>.
- Sarma, P., Aziz, K., and Durlafsky, L. J. 2005. Implementation of Adjoint Solution for Optimal Control of Smart Wells. Paper presented at the *SPE Reservoir Simulation Symposium*, The Woodlands, Texas, USA, 31 January–2 February. SPE-92864-MS. <https://doi.org/doi:10.2118/92864-MS>.
- Sayarpour, M. 2009. *Development and Application of Capacitance-Resistive Models to Water/CO₂ Floods*. PhD dissertation, The University of Texas at Austin, Austin, Texas.
- Sayarpour, M., Kabir, C. S., Sepehrmoori, K. et al. 2011. Probabilistic History Matching with the Capacitance–Resistance Model in Waterfloods: A Precursor to Numerical Modeling. *J Pet Sci Eng* **78** (1): 96–108. <https://doi.org/doi:10.1016/j.petrol.2011.05.005>.
- Sayarpour, M., Zuluaga, E., Kabir, C. S. et al. 2009. The Use of Capacitance–Resistance Models for Rapid Estimation of Waterflood Performance and Optimization. *J Pet Sci Eng* **69** (3–4): 227–238. <https://doi.org/doi:10.1016/j.petrol.2009.09.006>.
- Sayyafzadeh, M. 2017. Reducing the Computation Time of Well Placement Optimisation Problems Using Self-Adaptive Metamodelling. *J Pet Sci Eng* **151**: 143–158. <https://doi.org/doi:10.1016/j.petrol.2016.12.015>.
- Sayyafzadeh, M. and Alrashdi, Z. 2019. Well Controls and Placement Optimisation Using Response-Fed and Judgement-Aided Parameterisation: Olympus Optimisation Challenge. *Comput Geosci* **24** (6): 2001–2025. <https://doi.org/doi:10.1007/s10596-019-09891-y>.
- Silva, V. L. S., Cardoso, M. A., Oliveira, D. F. B. et al. 2020. Stochastic Optimization Strategies Applied to the OLYMPUS Benchmark. *Comput Geosci* **24** (6): 1943–1958. <https://doi.org/doi:10.1007/s10596-019-09854-3>.
- Spall, J. C. 1992. Multivariate Stochastic Approximation Using a Simultaneous Perturbation Gradient Approximation. *IEEE Trans Automat Contr* **37** (3): 332–341. <https://doi.org/doi:10.1109/9.119632>.
- Spall, J. C. 1998. Implementation of the Simultaneous Perturbation Algorithm for Stochastic Optimization. *IEEE Trans Aerosp Electron Syst* **34** (3): 817–823. <https://doi.org/doi:10.1109/7.705889>.
- Suwartadi, E., Krogstad, S., and Foss, B. 2015. Adjoint-Based Surrogate Optimization of Oil Reservoir Water Flooding. *Optim Eng* **16** (2): 441–481. <https://doi.org/doi:10.1007/s11081-014-9268-4>.
- Tanaka, S., Wang, Z., Dehghani, K. et al. 2018. Large Scale Field Development Optimization Using High Performance Parallel Simulation and Cloud Computing Technology. Paper presented at the *SPE Annual Technical Conference and Exhibition*, Dallas, Texas, USA, 24–26 September. SPE-191728-MS. <https://doi.org/doi:10.2118/191728-MS>.

- Van Doren, J. F. M., Markovinović, R., and Jansen, J.-D. 2006. Reduced-Order Optimal Control of Water Flooding Using Proper Orthogonal Decomposition. *Comput Geosci* **10** (1): 137–158. <https://doi.org/doi:10.1007/s10596-005-9014-2>.
- Volkov, O. and Bellout, M. C. 2017. Gradient-Based Production Optimization with Simulation-Based Economic Constraints. *Comput Geosci* **21** (5–6): 1385–1402. <https://doi.org/doi:10.1007/s10596-017-9634-3>.
- Wang, C., Li, G., and Reynolds, A. C. 2009. Production Optimization in Closed-Loop Reservoir Management. *SPE J.* **14** (3): 506–523. SPE-109805-PA. <https://doi.org/doi:10.2118/109805-PA>.
- Wang, X., Haynes, R. D., and Feng, Q. 2016. A Multilevel Coordinate Search Algorithm for Well Placement, Control and Joint Optimization. *Comput Chem Eng* **95**: 75–96. <https://doi.org/doi:10.1016/j.compchemeng.2016.09.006>.
- Weber, D., Edgar, T. F., Lake, L. W. et al. 2009. Improvements in Capacitance-Resistive Modeling and Optimization of Large Scale Reservoirs. Paper presented at the *SPE Western Regional Meeting*, San Jose, California, USA, 24–26 March. SPE-121299-MS. <https://doi.org/doi:10.2118/121299-MS>.
- Wen, T., Thiele, M. R., Ciaurri, D. E. et al. 2014. Waterflood Management Using Two-Stage Optimization with Streamline Simulation. *Comput Geosci* **18** (3–4): 483–504. <https://doi.org/doi:10.1007/s10596-014-9404-4>.
- Yeten, B., Castellini, A., Guyaguler, B. et al. 2005. A Comparison Study on Experimental Design and Response Surface Methodologies. Paper presented at the *SPE Reservoir Simulation Symposium*, The Woodlands, Texas, USA, 31 January–2 February. SPE-93347-MS. <https://doi.org/doi:10.2118/93347-MS>.
- Yeten, B., Durlafsky, L. J., and Aziz, K. 2003. Optimization of Nonconventional Well Type, Location, and Trajectory. *SPE J.* **8** (3): 200–210. SPE-86880-PA. <https://doi.org/doi:10.2118/86880-PA>.
- Yousef, A. A., Gentil, P. H., Jensen, J. L. et al. 2006. A Capacitance Model To Infer Interwell Connectivity From Production and Injection Rate Fluctuations. *SPE Res Eval & Eng* **9** (6): 630–646. SPE-95322-PA. <https://doi.org/doi:10.2118/95322-PA>.
- Zandvliet, M. J., Bosgra, O. H., Jansen, J. D. et al. 2007. Bang-Bang Control and Singular Arcs in Reservoir Flooding. *J Pet Sci Eng* **58** (1–2): 186–200. <https://doi.org/doi:10.1016/j.petrol.2006.12.008>.
- Zhao, H., Chen, C., Do, S. et al. 2013. Maximization of a Dynamic Quadratic Interpolation Model for Production Optimization. *SPE J.* **18** (6): 1012–1025. SPE-141317-PA. <https://doi.org/doi:10.2118/141317-PA>.
- Zubarev, D. et al. 2009. Pros and Cons of Applying Proxy-Models as a Substitute for Full Reservoir Simulations. Paper presented at the *SPE Annual Technical Conference and Exhibition*, New Orleans, Louisiana, USA, 4–7 October. SPE-124815-MS. <https://doi.org/doi:10.2118/124815-MS>.

6 Adaptive rank-based selection of geological realizations for optimum field development planning

Arouri, Y.; Sayyafzadeh, M. & Begg, S.

SPE Journal, 2022.

Statement of Authorship

Title of Paper	Adaptive rank-based selection of geological realizations for optimum field development planning
Publication Status	<input checked="" type="checkbox"/> Published <input type="checkbox"/> Accepted for Publication <input type="checkbox"/> Submitted for Publication <input type="checkbox"/> Unpublished and Unsubmitted work written in manuscript style
Publication Details	Arouri, Y.; Sayyafzadeh, M. & Begg, S. Adaptive rank-based selection of geological realizations for optimum field development planning. SPE Journal, 2022.

Principal Author

Name of Principal Author (Candidate)	Yazan Arouri
Contribution to the Paper	Coding, experimentation, literature review, analysis, original manuscript
Overall percentage (%)	70%
Certification:	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.
Signature	Date 10/04/2022

Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

- the candidate's stated contribution to the publication is accurate (as detailed above);
- permission is granted for the candidate to include the publication in the thesis; and
- the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author	Mohammad Sayyafzadeh
Contribution to the Paper	Supervision, original idea, discussion, analysis of results writing, mathematical formulation
Signature	Date 06/06/2022

Name of Co-Author	Steve Begg
Contribution to the Paper	Writing, supervision, editing
Signature	23/5/22

Please cut and paste additional co-author p

Adaptive Rank-Based Selection of Geological Realizations for Optimum Field Development Planning

Yazan Arouri^{1*}, Mohammad Sayyafzadeh^{1*}, and Steve Begg¹

¹Australian School of Petroleum and Energy Resources, The University of Adelaide

Summary

Uncertainties are present in many decision-making processes. In field development planning, these uncertainties, typically represented by a set of geological realizations, need to be propagated in response to any proposed alternative (solution). Incorporation of the full set of realizations results in the oil and gas field development optimization problem—where either an algorithm iteratively tries to find the best solution from all the possible alternatives or the best solution must be selected from a set of predefined engineering judgment-driven development scenarios (i.e., set of either well control or well placement settings)—becoming computationally demanding. As such, realization subset selection techniques are required to reduce the computational overhead. We first introduce a reformulation of the subset selection problem to one that aims at ensuring consistent ranking of alternatives between those obtained by the full set and the selected subset. We argue that this should be the ultimate goal of any subset selection technique in such problems. In addition, we also propose a technique which selects a subset that minimizes the difference between the rankings obtained by the full set and subset, for a small batch of alternatives. The key idea, which we investigate thoroughly, is that there is a positive association between the goodness (in terms of ranking alternatives) of the subset selected using a small batch of alternatives and its fidelity in ranking other alternatives. Unlike previous methods, this technique does not depend on selecting subjective (static) properties to perform the subset selection nor does it rely only on flow-response vectors of a base-case scenario. In this work, the proposed technique is assessed using well placement and well control development alternatives to determine the applicability within field development planning. Additionally, the proposed subset selection technique is implemented in an adaptive scheme to solve a well placement optimization problem. The results are promising as the proposed technique consistently selects subsets that are able to rank development alternatives in a similar manner to the full set regardless of the type of development strategy (well control settings or well placement). Furthermore, the implementation of the proposed technique in an adaptive scheme is able to reduce the computational costs, on average, by a factor close to 9 without compromising the solution found for well placement optimization.

Introduction

Decision-making is essentially the design and ranking of alternatives, based on their value to the decision-maker (Bratvold and Begg 2010). Typically, a predetermined (discrete) number of alternatives are designed/proposed (through engineering judgment), assessed, and then ranked. The purpose of modeling in engineering studies, such as well placement optimization, is usually to make predictions that inform decisions by allowing the evaluation of each alternative and subsequently ranking them. It is worth mentioning that most algorithms used in population-based optimization frameworks perform in a similar fashion to decision-making processes where, in each iteration, the algorithm ranks and selects solutions (alternatives) based on their fitness value.

An assessment of the uncertainty associated with a model's prediction is vital because it impacts the rankings (and hence the decisions) that are made. An accurate recognition of the uncertainty is important, as it can enable decision-makers (or algorithms) to make different and better decisions, ones that have an increased chance of achieving the desired outcome. In the context of uncertainty, accuracy means that all possible outcomes have been identified and that the probability of each is consistent with our state of information and the extent to which our models are an approximation to reality. If we are unsure of the values of the input parameters, which describe the system and its conditions, or if the model is a crude approximation of the reality, these must be reflected and propagated in the uncertainty of the predictions. It is also important that the model inputs (and the structure of the model itself) are not biased (Welsh and Begg 2016).

When the uncertainty has been assessed, decision-makers (or algorithms) are tasked with ranking and selecting the most value-creating alternative. To do so, for each alternative, the decision-makers (or algorithms) construct a single value measure—the decision metric—that incorporates the full spectrum of possible outcomes, their probabilities, and the decision-maker's attitude to risk. Common decision metrics include expected value for a risk-neutral decision-maker (as assumed in this work) or expected utility for risk-averse or risk-seeking decision-making. Understanding the (correctly) propagated uncertainty can not only increase the chance of getting desired outcomes, but it can also be exploited to create additional value (Begg et al. 2002) by building-in flexibility (future decisions contingent upon the resolution of uncertainty) to mitigate downside risk and capture upside opportunity. In this study, our focus is on the impact of uncertainty on ranking of the alternatives (either within a decision-making context or a population-based optimization framework).

In field development planning, reservoir flow simulations provide the key input for decision-making when evaluating alternative development strategies against the objectives. This allows the investigation of system responses to various alternatives and carrying out what-if sensitivities to find the optimum solution (alternative/development strategy). If the input parameters describing the system and its conditions (i.e., geological model, fluid properties, development strategy, etc.) are considered to be known with reasonable certainty, then the output can be considered to be fully determined—assuming modeling errors are negligible. However, this is not the case, and normally, the geological models have a great deal of uncertainty. Hence, an ensemble of geological realizations is typically used to represent the current state of information and is used to propagate uncertainty in the prediction of fluid production volumes (because of the nonlinearity of flow equations, Monte Carlo simulation is used).

*Corresponding author; email: yazan.arouri@adelaide.edu.au; mohammad.sayyafzadeh@adelaide.edu.au

Depending on the complexity of the subsurface and available information, the ensemble may be composed of 100s to 1,000s of realizations. As such, when evaluating the fitness (against the objectives) of a development strategy, it must be assessed against the full ensemble. This results in a cumulative distribution function (CDF) of the fitness metric—this could be an economic measure, such as net present value (NPV), but field production measures may also be used. The CDF (or its associated probability density function), corresponding to the fitness metric, serves as a mechanism to reflect the uncertainty needed for a robust decision-making process. As aforementioned, this then needs to be represented by a single value (i.e., the decision metric) to allow for the ranking of each development strategy. Such techniques have been used within optimization frameworks to enable robust optimization of field developments (Yeten et al. 2003; Artus et al. 2006; Onwunali and Durlafsky 2010; Bouzarkouna et al. 2012; Sayyafzadeh and Alrashdi 2019).

However, the significant computational burden of evaluating (and subsequent ranking of) each development strategy against the full set of realizations becomes intractable. As a result, scenario reduction techniques are used for the selection of representative realizations from the full set. Once selected, this subset can then be used to perform field development studies at a significantly lower computational cost. The used subset selection technique must not alter the ranking of the solutions. In essence, this means that whether a full suite or the selected subset of realizations is used, the ranking of the solutions, based on the decision metric, should be consistent to ensure reliable and robust decisions.

Related Work. Previous literature has investigated multiple techniques for the selection of a representative subset of realizations from the full suite. We have categorized these techniques into four classifications. It includes techniques that are based on (i) random selection, (ii) static property methods, (iii) flow-response vector-based methods, and (iv) a combination of (ii) and (iii).

Because of its simplicity, the random subset selection technique is widely implemented. This technique involves randomly selecting a predefined number of realizations from the full ensemble to form a subset. Lorentzen et al. (2009) used the random selection technique within a closed-loop reservoir management framework for the Brugge benchmark model (Peters et al. 2010). Li et al. (2013) used the random selection technique at each iteration of the optimization procedure. The authors explained that this would ensure each realization would be used throughout the optimization process and hence improve the results. Jesmani et al. (2016, 2020) applied a similar technique for robust well placement optimization using a randomly selected subset of a predefined size at each iteration.

The approximated CDF of the output resulting from the randomly selected subset may not be representative of the full set CDF, because the relationship between input and output is highly nonlinear, and more importantly, the similarity between the subset CDF and full set CDF is solution (development plan) dependent. This may lead to an unintended shift in the decision. Although changing the selected subset at each iteration may overcome the solution dependency, it comes at the cost of changing the shape of the objective function landscape of the optimization problem. Hence, this can delay convergence, as it acts as a noisy objective function. Furthermore, the subset size to be chosen and the frequency of the subset selection procedure should occur are not known a priori.

The second category of subset selection techniques attempts to approximate the CDF of the input by collating geological realizations based on a static measure. A static measure is an easily computed prior feature, such as a geological property (e.g., porosity, permeability, connected PV, etc.), which is taken as a representative property of the full geological model. The static measure-based method consists in sorting the full set of realizations according to the static measure, in an ascending/descending order. Once the realizations are in order, the predefined number of subset realizations is selected to represent the low, medium, and high measure values.

Deutsch and Begg (2001) proposed selecting a predefined number of realizations that are equally spaced to preserve the quantiles of the full set. Other studies investigated the use of a static measure that is a combination of various properties for steam-assisted gravity drainage (McLennan and Deutsch 2005), as well as using connected hydrocarbon volume as a ranking measure (Li et al. 2012). More recently, Rahim and Li (2015) introduced an optimization-based subset selection technique employing probability distribution distances. The aim of this technique is to find the optimal subset that has similar statistical distribution characteristics to the full set.

Although these techniques have been shown to result in positive results for computational reductions, they do not consider the importance of ensuring a consistent ranking between the subset and full set for decision-making. In addition, as discussed earlier, there is an inherently nonlinear relationship between the geological model inputs and the reservoir simulation outputs. Therefore, successfully approximating the CDF of the input does not necessarily result in an approximated CDF of the output of similar quality. As such, the final decision made based on the subset may be inconsistent with the decision based on the full set.

The third category of subset selection techniques aims at approximating the CDF of the flow-response output for a base case field development plan. Once the CDF is approximated, it is then assumed that this CDF is representative of the CDFs from all other possible field development plan flow-response outputs.

The usage of this method first appeared in the reservoir engineering community by Ballin et al. (1992). In this work, faster alternatives to full-physics simulators were used to run the full ensemble of realizations (for a base case). Then, the realizations were ranked based on these flow-response outputs for selection. The selected realizations are then used to undergo full-physics simulation. A similar approach is used in Mishra et al. (2002) and Odai and Ogbe (2011), but in these works, a streamline simulator is used to obtain flow-response outputs for the full set of realizations. Chen et al. (2012) used flow-response outputs from an initial guess of well control optimization for ranking and selection of a subset of realization.

Another subset selection technique that also attempts to approximate the CDF of a base case flow response is the distance-based clustering method. This is accomplished by computing a distance (dissimilarity) matrix for the full set of realizations. The matrix is then transformed into the Euclidean space using multidimensional spacing. Because of the nonlinear nature of the points produced by multidimensional spacing, kernel methods are used to map the data points to a high-dimensional feature space. It is assumed that the feature space has a better linear variation than the multidimensional spacing Euclidean space. As such, techniques such as principal component analysis and k-means clustering can be applied in the resulting feature space. Consequently, based on an a priori defined number of centroids, the full set of realizations are grouped within clusters based on their distance to the centroid. These representative centroids are taken as the realizations in the subset. Scheidt and Caers (2009) and Alpak et al. (2010) both applied the clustering method to large full sets of realizations using flow-based responses from streamline simulations.

Although suitable results can be obtained through flow-response-based techniques, it is crucial to accept that each unique field development plan will have a unique associated CDF. As such, it is a fragile extension to assume an approximated CDF of the output from a base case flow response is representative of the full set output CDF for any development plan (solution). Duly, it is foreseeable that the rankings of the solutions when using the subset may show inconsistencies to the ranking based on the full set.

More recently, there has been an introduction of approaches that use a combination of these categories to improve the subset selection process. These approaches typically incorporate static measures (second category) and flow-response vectors (third category) in the subset selection method. Wang et al. (2012) combine flow-response output and static measures through the projection into a 2D space for all realizations. A clustering technique is then used to select a predefined number of realizations. Haghghat Sefat et al. (2016) used this

approach for well control optimization by using the area between well water cut curves as the dissimilarity measure. Shirangi and Durlfolsky (2016) also introduced a subset selection method based on the combination of static measures and flow-response output. The method is based on a weighted combination between the static measure and flow-response from selected well configurations. Salehian et al. (2021) applied the technique introduced in Wang et al. (2012) to the joint optimization of well control and well placement. In their work, the static measure is permeability and the flow-response output is cumulative oil production for the initial well configuration.

The main objective of any subset selection work should be to support the decision-making process (or algorithm in population-based optimization frameworks) by providing information to assist with the ranking of alternatives. However, previous subset selection techniques, by design, do not necessarily ensure that consistent decisions are made as they are not aimed directly at that. To this end, we propose a reformulation of the subset selection technique to one which aims at ensuring consistent ranking to the full set when a subset of realizations is selected and used. Furthermore, we propose a subset selection technique that ensures a strong association between the solution ranking based on the subset and the full set using a small batch of solutions (development strategies). This study investigates the performance across different subset sizes and training set sizes for well control and well placement development strategies. The proposed technique is then implemented and tested within a population-based algorithm for a well placement optimization problem under uncertainty.

The outline of this paper is as follows. We begin by detailing the key mathematical concepts for the reformulation of the subset selection problem. Following this, we present the proposed subset selection technique methodology for the reformulated subset selection problem. We also present the implementation of the proposed method in an optimization under uncertainty problem. Next, the experimental setup, including a description of the benchmark model, (economic) fitness metric, and decision metric, is introduced. Then, we present the experimental results assessing the applicability of the proposed method for well control and well placement development strategies. We also present the performance of the proposed method when implemented for a well placement optimization problem. Finally, we discuss the implications of these results for subset selection and conclude the paper with a summary of the key findings.

Mathematical Formulation for Subset Selection

For a given geological model, m , and field development strategy, s , the solution of the set of partial differential equations that govern the fluid flow through porous media, represented by g , gives the forecasts, p . This is expressed mathematically with the following relationship:

$$p = g(m, s). \quad (1)$$

The forecasts are used in combination with other (economic) parameters, v , to calculate a (economic) fitness metric, y , using h as the (economic) fitness function, expressed as follows:

$$y = h(g(m, s), v). \quad (2)$$

To incorporate subsurface uncertainty, an ensemble of geological models is used. As such, the geological model is considered a random vector, M . Consequently, for a given field development strategy, the associated (economic) fitness metric will also be represented by a random variable/vector, Y :

$$Y(s) = h(g(M, s), v). \quad (3)$$

N samples are taken from M to represent the suite of realizations (models) and each realization is denoted by m_i . The CDF corresponding to $Y(s)$ is obtained/approximated using a set of realizations which is composed of a large ensemble (with the size N) of models m_i , referred to as the full set of realizations in this study. The CDF corresponding to s is denoted as $F_{Y(s)}$. The computational overhead required to generate $F_{Y(s)}$, for each development plan, s_j , using the full set of realizations, will become computationally expensive. Consequently, a number of realizations, n , are intended to be selected as a representative subset, from the full set with the size of N —where $n \ll N$. $F_{Y(s)}$ corresponding to any s_j can be approximated using the output of the subset. Consequently, the aim for any subset selection technique is to select a subset of realizations, denoted by X_{ss} , such that the resulting CDF of the output for any development strategy $F'_{Y(s)}(X_{ss})$ is approximately representative of the CDF of the output when evaluated using the full set for the same development strategies, $F_{Y(s)}$

$$F'_{Y(s)}(X_{ss}) \approx F_{Y(s)}, j \forall 1 : J, \quad (4)$$

where J is the total number of all the possible development strategies (solutions) that are either proposed based on engineering judgment and/or from an algorithm. Here, we define X_{ss} as a binary vector [i.e., (x_1, x_2, \dots, x_N)] representing the selected combination of realizations in the subset. For example, using a full set size (N) of 5 and $X_{ss} = (1, 0, 0, 1, 0)$, this means that the first and fourth realizations are selected to be in the subset. Because of the complex nonlinearities between the input—geological model and development strategy—and the output—production forecasts, there is no closed-form solution for such a problem. Subsequently, it is not the aim of this work to solve this problem. A more relaxed and approachable aim is to approximate a decision metric (e.g., expected value), defined as l , of $Y(s)$ using a subset. This can be expressed mathematically as follows:

$$\begin{aligned} \text{Let } u &= l(Y(s)), \\ u'_{Y(s)}(X_{ss}) &\approx u_{Y(s)}, j \forall 1 : J \end{aligned} \quad (5)$$

$u'_{Y(s)}(X_{ss})$ is the outcome corresponding to s_j approximated by X_{ss} subset of realizations, whereas $u_{Y(s)}$ is the outcome obtained by the full set of realizations. Herein, for brevity, we refer to $u_{Y(s)}$ and $u'_{Y(s)}(X_{ss})$ as $u(s_j)$ and $u'(s_j, X_{ss})$, respectively.

These are the metric values used to rank solutions (development strategies). Therefore, the approximation in Eq. 5 is still an intermediate goal, and we argue that the ultimate goal of a subset selection technique should be to maintain a consistent ranking of the solutions. To this end, we propose the reformulation of the subset selection problem to one which aims at producing similar rankings of solutions between the subset and the full set. Specifically, we argue the aim should be to find a subset that produces similar ranks for the solutions (development strategies) measured by u and u' . This will help ensure that a consistent decision is made regardless of whether the subset or the full set is used. As such, with this ultimate aim, there is no requirement concerning the quality of CDF approximation nor the

approximation of the decision metric values, as long as the ranks of the solutions do not change. This is a more relaxed approximation problem.

Methodology

In this section, we present the proposed methodology for selecting a representative subset from a full set of geological realizations. Next, we present the methodology for the usage of the proposed subset selection method within optimization under uncertainty.

Rank Correlation Method. The reformulation of the subset selection problem can be expressed mathematically. The difference, $d(X_{SS})$, between the rankings of the full set and the subset can be defined as follows:

$$d(X_{SS}) = \| r - r'(X_{SS}) \|, \quad (6)$$

where r and $r'(X_{SS})$ are the rank vectors for J solutions obtained by the full set and the X_{SS} subset, respectively. These two vectors, r and $r'(X_{SS})$, rank the output in vectors, u and u' , which are calculated for J solutions using the decision metric, l , in Eq. 5. This can be expressed as an optimization problem:

$$\begin{aligned} & \min_{X_{SS} \in \mathbb{Z}_2^N} d(X_{SS}) \\ & \text{subject to : } \sum_{i=1}^N X_{SS}(i) = n \end{aligned} \quad (7)$$

This problem is equivalent to maximizing Spearman's rank correlation coefficient:

$$R(X_{SS}) = 1 - \left(\frac{6 d(X_{SS})}{J^2 - J} \right), \quad (8)$$

where J , as before, is the total number of development strategies (solutions) being ranked. However, in reality, it would be pointless to evaluate all J development strategies (solutions) with the full set. As such, we present the following.

First, we select a batch (with the size of J_T) of development strategies from J , determined by the available computing resources. Accordingly, the difference in Eq. 6 becomes the following:

$$d_T(X_{SS}) = \| r_T - r'_T(X_{SS}) \|, \quad (9)$$

where r_T and $r'_T(X_{SS})$ are the rank vectors for the selected batch of J_T solutions obtained by the full set and the X_{SS} subset, respectively. The effect that the value of J_T has on the proposed method's ability to select representative subsets is investigated in later sections. Next, the fitness metric (e.g., NPV) is calculated for each development strategy (i.e., s_1, \dots, s_{J_T}) against each realization (N) in the full set. This will result in a matrix of size $J_T \times N$, which we will call the fitness metric matrix (FMM). Following this, the decision metric (e.g., expected value of NPV) for each development strategy is calculated to obtain a singular value that can be used to rank them. This will produce a vector of size $J_T \times 1$, where each entry represents the decision metric value for a development strategy. The decision metric values are then ranked to produce the vector r_T .

Then, a subset size, n , is defined and all possible combinations of n realizations from N are enumerated. As this can become a computationally expensive task for large full set and subset sizes, an optimization-based approach may be used. Next, for each X_{SS} subset (i.e., combination of realizations), the corresponding columns in the FMM are extracted for each development strategy. As such, a smaller FMM for the subset, of size $J_T \times n$, will result. Again, the decision metric value for each development strategy is calculated using the subset FMM which results in a vector of size $J_T \times 1$, corresponding to the subset. Similar to the full set, the decision metric values (corresponding to the J_T development strategies) are ranked to produce the vector $r'_T(X_{SS})$ corresponding to X_{SS} .

Finally, a distance, as shown in Eq. 9, can be calculated between the ranked subset, $r'_T(X_{SS})$, and full set, r_T , vectors for that specific subset. The combination of realizations which produces the minimum d_T (i.e., maximum Spearman's rank correlation coefficient) is selected as the representative subset, X_{SS}^* , which can be found by an exhaustive search (if the number of possible combinations is not very large) or by an optimization process. For implementation and presentation purposes, we use Spearman's rank correlation coefficient to obtain and relay results. **Fig. 1** presents the flow chart for the proposed method.

Optimization under Uncertainty. The use of population-based algorithms has been a favored approach to solve reservoir engineering optimization problems, such as well placement (Yeten et al. 2003; Onwunali and Durlofsky 2010; Bouzarkouna et al. 2012; Sayyafzadeh 2017; Alrashdi and Sayyafzadeh 2019; Sayyafzadeh and Alrashdi 2019). Population-based algorithms rely on accurately ranking the proposed solutions in each iteration (generation). These rankings are vital for comparing the proposed solutions in the current iteration (generation) with other solutions in the current and past iterations. If, hypothetically, these rankings are inaccurate, it is possible that the information from higher quality solutions is not retained and is lost. Although the focus of this work is not the comparison of different algorithms, it should be noted that for high-dimensional well control optimization problems, gradient-based algorithms are more favorable [see, e.g., Chen et al. (2009), Fonseca et al. (2017), and Arouri and Sayyafzadeh (2020a)]. Given the proposed method works best when ranking is needed, application of the proposed method to gradient-based algorithms requires further investigation.

Implementation of Rank Correlation Method. In our study, we focus on the use of particle swarm optimization (PSO) as the population-based algorithm (Kennedy and Eberhart 1995). However, the subsequent implementation of our proposed subset selection method applies to any population-based algorithm given the presence of ranking/comparisons of solutions. PSO begins with the initialization of a population of solutions (particles) and the algorithmic parameters (namely, ω , c_1 , and c_2). Next, during each iteration, the proposed solutions in the population are evaluated using the defined objective/fitness function. Next, a comparison is made between each proposed solution and its previous personal best, stored in s_{pj} , as well as with the global best solution found so far, s_g , based on the evaluated fitness values. If an approximate fitness value, u' , is used (which is the case when a subset of realizations is used), instead of the exact fitness value, u (when the full set is used), it is essential to ensure the comparisons/ranking will be made correctly. In other words, the difference between the approximated and exact fitness values is not important as long as the result of rankings/comparisons is in

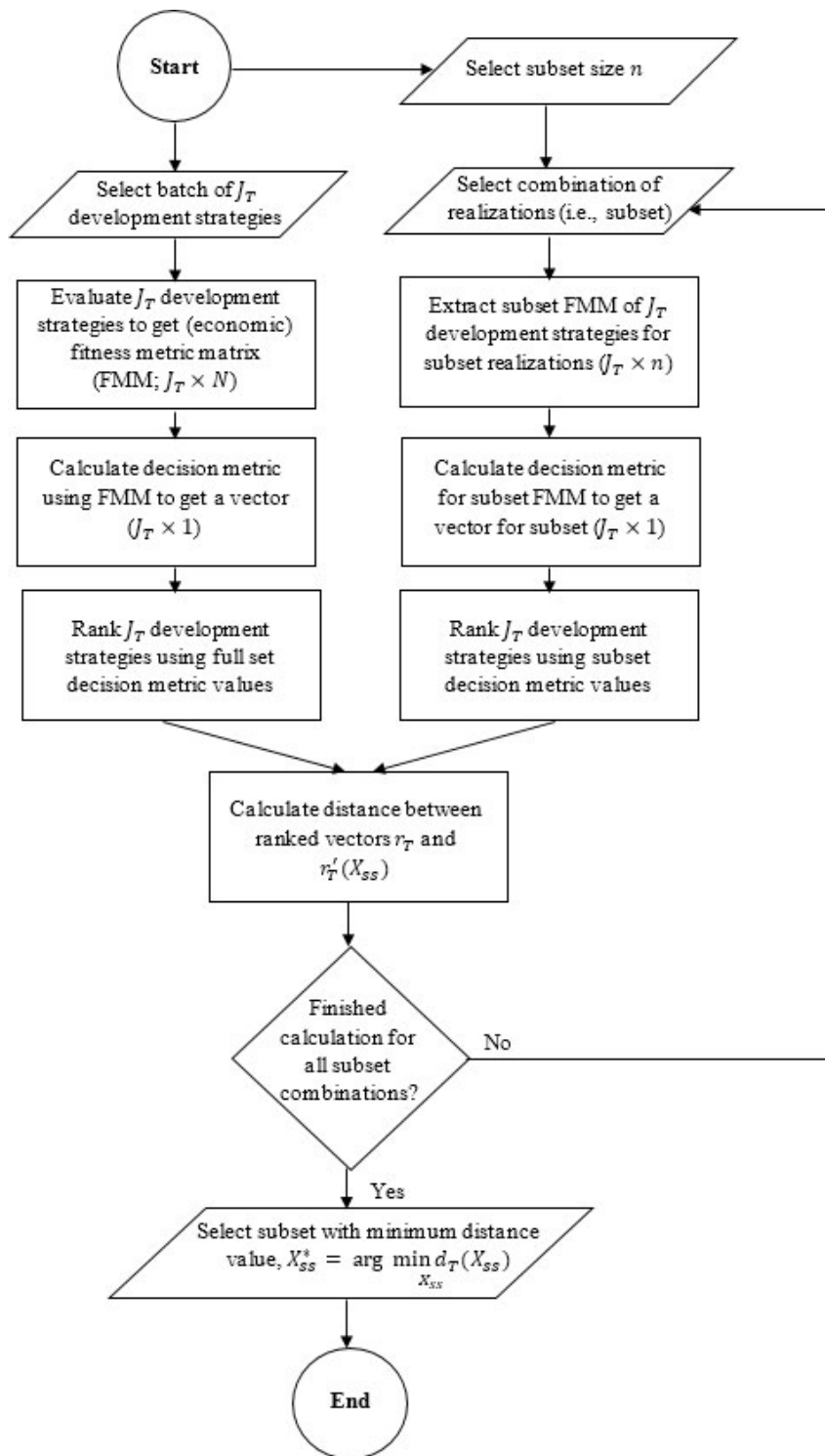


Fig. 1—Flow chart of the proposed method for subset selection.

agreement. Hence, we emphasize the difference in ranking for the selection of a subset which would allow the reduction in computational costs of optimization under uncertainty problems.

The implementation of our proposed subset selection method allows the use of an adaptive approach with respect to the realizations in the subset as well as the subset size. This is done by continuously assessing the performance of the selected subset throughout the optimization process to ensure the rankings are similar to those obtained by the full set, and if its fidelity drops, it will be substituted with a higher fidelity subset by reperforming the subset selection. More specifically, a prediction matrix, A_p , is used to store the fitness value for the best solution found in each generation, denoted as $u(s_{gk})$, that is evaluated using the full set of realizations. This is equivalent to the FMM in Fig. 1 and allows the calculation of a prediction rank correlation coefficient, R_p , for the current selected subset. If the calculated R_p is below a predefined prediction threshold (i.e., below par performance in ranking), R_{pt} , then a subset selection process is initiated. The effect of the threshold value R_{pt} will be investigated in later sections. Once subset selection is initiated, the A_p matrix is reset (i.e., set to

Algorithm 1 —Rank correlation method with PSO.

```

1: Initialize a random population of particles; algorithmic parameters ( $\omega, c_1, c_2$ ); subset selection parameters ( $R_{pt}, R_{tt}, A_p, A_t$ );  $k = 0$ ; initial subset ( $X_{ss}^*$ )
2: while Termination condition/s not met do
3:   for Each particle  $j$  do
4:     Check feasibility of particle
5:     if Infeasible then
6:       Repair
7:     Evaluate fitness of particle  $s_j$  using selected  $X_{ss}^*$  subset
8:     if  $u'(s_j, X_{ss}^*) < u'(s_{pj}, X_{ss}^*)$  then
9:        $s_{pj} \leftarrow s_j$ 
10:    if  $u'(s_j, X_{ss}^*) < u'(s_g, X_{ss}^*)$  then
11:       $s_g \leftarrow s_j$ 
12:    if  $u'(s_j, X_{ss}^*) < u'(s_{gk}, X_{ss}^*)$  then
13:       $s_{gk} \leftarrow s_j$ 
14:     $A_p = [A_p; u(s_{gk})]$ 
15:    if  $\dim(A_p, 1) \geq 5$  then
16:      Calculate rank correlation coefficient of current subset ( $R_p$ ) using  $A_p$ 
17:      if  $R_p < R_{pt}$  then
18:         $n = 1$ 
19:         $A_t = [A_t; A_p]$ 
20:         $A_p = []$ 
21:        while  $R_t < R_{tt}$  do
22:           $n = n + 1$ 
23:          Find  $X_{ss}^{new}$  subset which maximizes  $R_t$  using either exhaustive search or optimization-based search
24:          if  $X_{ss}^{new} \neq X_{ss}^*$  then
25:             $X_{ss}^{new} \leftarrow X_{ss}^*$ 
26:            Re-evaluate  $u'(s_{pj}, X_{ss}^*)$  and  $u'(s_g, X_{ss}^*)$  using new subset
27:            Update velocities of each particle
28:            Update the positions of each particle
29:             $k = k + 1$ 

```

an empty matrix). As such, the implementation of the proposed subset selection technique is adaptive with respect to the subset (i.e., the realizations in the subset may change) through the use of feedback from the optimization process.

Once the subset selection process is initiated, the procedure follows that described in **Fig. 1**, where the J_T development strategies are the solutions (i.e., population) proposed by the optimization algorithm and are then evaluated by the full set to obtain the FMM. The FMM in **Fig. 1** is denoted as the training matrix, A_t , in this implementation. This matrix comprises all the fitness values that were previously in A_p and were used to calculate R_p . Next, starting at a subset size (n) of 1, A_t is used to find the subset that maximizes the training rank correlation coefficient, R_t . If this is above the predefined training threshold, R_{tt} , then the selected subset is used in subsequent generations until the subset process is initiated again. Otherwise, the subset size is incremented by 1 and the procedure is repeated until the training threshold is met. Either an exhaustive search or an optimization-based approach can be used to find the subset that maximizes R_t .

It is important to note that if a new subset is selected, the current personal best solutions, s_{pj} , and the current global best solution, s_g , must be reevaluated using the newly selected subset. This is to ensure the ranking and selection of solutions are in comparable terms. The pseudocode of the described algorithm is presented in Algorithm 1.

Experimental Setup

In this section, we present the experimental setup used to test the proposed subset selection technique. We begin by introducing the benchmark model, the Egg model, which was used for all experimental problems presented here. Next, we present the (economic) fitness metric and the decision metric that are used to rank the development strategies.

The Egg Model. The benchmark model was introduced into the reservoir optimization community for the application of optimization algorithms to field development studies (Jansen et al. 2014). It is a synthetic 3D reservoir model consisting of 18,553 active grid cells, as shown in **Fig. 2**. The model is a two-phase (oil and water) model, with no aquifer (pressure) support. The geological uncertainty of the model is represented by 100 realizations (full set) of the permeability field. A sample of six x -directional permeability realizations from the full set ensemble is shown in **Fig. 2**.

Two types of development strategies that are typically studied in the reservoir optimization community were used to investigate the proposed method. The first type of development strategy considered is well placement solutions, which consist of the areal locations for a predefined number of wells. In this work, the well placement strategies define the x and y coordinates for 12 (8 injectors and 4 producers) vertical wells, resulting in a total of 24 decision variables. The second type of development strategy consists of the injection rates for eight water injection wells over the life span of the reservoir. These control settings are piecewise constant between each timestep, which are taken to be at 90-day intervals over a 10-year period. This results in a total of 320 decision variables. For example, the well control settings for one injection well are defined as $(q_{wi,1}^1, \dots, q_{wi,N_t}^1)$, where, for example, $q_{wi,1}^1$ represents the water injection rate for well number 1 during the first timestep and N_t is the total number of timestep intervals.

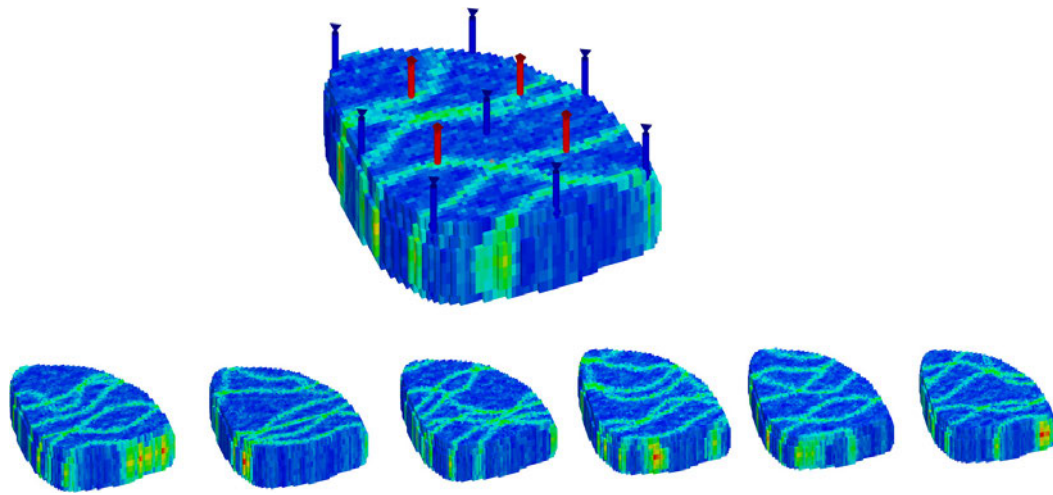


Fig. 2—Top: The Egg model displaying injection wells (blue) and producers (red). **Bottom:** Six realizations selected from the full set of 100 realizations showing the variability in the x-directional permeability across the ensemble. Blue color represents lower permeability values, and green color represents higher permeability values.

Fitness Metric. The (economic) fitness metric used in this study is the NPV of a given development strategy, s , for a given geological model, m . Other metrics, such as cumulative production totals, may also be used as the fitness metric. The values of the fitness metric for each development strategy against each realization form the FMM (see Fig. 1). For example, if there are 200 development strategies and 100 realizations in the full set, this will result in an FMM of size 200×100 . The NPV is defined as follows:

$$y = h(g(m, s), v) = \text{NPV}(m, s) = \sum_{i=1}^{N_t} \frac{c_{o,i} Q_{o,i}(m, s) - c_{wp,i} Q_{wp,i}(m, s) - c_{wi,i} Q_{wi,i}(m, s)}{(1 + b)^t}, \quad (10)$$

where $c_{o,i}$, $c_{wp,i}$, and $c_{wi,i}$ are the price of oil and costs of water separation and injection, respectively, all of them per unit volume and defined from time t_i to t_{i+1} (there are N_t such intervals), $Q_{o,i}$, $Q_{wp,i}$, and $Q_{wi,i}$ denote the field oil production, field water production, and field water injection volumes during the mentioned output interval, and b is the discount rate. The field production and injection volumes are obtained from the reservoir simulation for a defined set of decision variables (development strategy; i.e., well locations or well control settings) and for a specified geological model. To calculate the NPV value for a specific geological realization model and development strategy, a reservoir life of 10 years is used and the economic parameters considered are summarized in Table 1.

Parameter	Value
Oil price, c_o	20 USD/bbl
Water handling costs, c_{wp}	3 USD/bbl
Water injection costs, c_{wi}	0.8 USD/bbl
Discount rate, b	0%

Table 1—Economic parameters for the calculation of (economic) fitness metric.

Decision Metric. The decision metric is applied to the FMM to condense the information into a vector of decision values for each development strategy by which the decision-maker can rank the alternatives. The decision metric used in this study is the expected value of the NPV, assuming a risk-neutral decision-maker. An expected utility may be used for risk-averse or risk-seeking decision-makers. The expected value is used to efficiently incorporate the geological uncertainty exhibited by the ensemble of realizations. The expected value of NPV can be formally stated for the full set of realizations, N , and development strategy (solution), s_j :

$$u(s_j) = \frac{\sum_{i=1}^N \text{NPV}(m_i, s_j)}{N}, \quad (11)$$

where u represents the decision metric output for the full set of realizations. Similarly, the expected value NPV for an X_{ss} subset of n realizations and development strategy (solution), s_j , is defined as:

$$u'(s_j, X_{ss}) = \frac{\sum_{i=1}^N X_{ss}(i) \times \text{NPV}(m_i, s_j)}{n}, \quad (12)$$

where u' represents the decision metric output for the subset of realizations and X_{SS} is the binary vector (1s and 0s) detailing the selected realizations. Various decision metrics have been used in the literature, including expected value (Fonseca et al. 2015, 2017) and risk-adjusted expected values (Yeten et al. 2003; Artus et al. 2006; Tilke et al. 2010). The selection of the decision metric is based upon the objectives of the decision-maker.

Experimental Results

In this section, we begin with the assessment of the rank correlation method for selecting representative realizations for well control and well placement development strategies. Following this, we present the results for a well placement optimization problem under uncertainty using the rank correlation method.

Assessment of Rank Correlation Method. To investigate and assess the ability of the proposed method to select a representative subset, we use training and prediction sets. To do this, we assume a total number of development strategies (solutions), J , exist. These solutions can either be proposed through engineering judgment or an optimization algorithm. For the purpose of this assessment, they were generated randomly. Next, as per Fig. 1, a batch with a size of J_T solutions is selected (from J) and is designated as the training set. The remaining solutions (i.e., $J - J_T$) are known as the prediction set, denoted by J_P . As such, for a specific value of J_T , the split of J can be represented as J_T/J_P .

For a specified subset size, n , the procedure described in Fig. 1 is used to find a subset that minimizes the distance (Eq. 9) between the full set (r_T) and subset ($r'_T(X_{SS})$) rankings using the development strategies in the training set (J_T). Once a subset (X_{SS}^*) is found, it can be used to calculate the predicted rank correlation coefficient for the prediction set (J_P). To assess the risk associated with the proposed method, the splitting procedure was repeated 10 times for every specified subset size. In other words, the strategies placed into J_T and J_P were randomly shuffled 10 times.

The inherent assumption is made that the rank correlation coefficient calculated for a subset of development strategies (i.e., based on J_T) is representative of the rank correlation coefficient for the total number of development strategies (J). The effect of the value of J_T on the performance of the rank combination method will be investigated in greater detail later in the study (i.e., to see if the conclusions/observations hold regardless of the size of J_T). This will allow us to extend the observations to situations where a training/prediction set is not computationally justifiable.

In our assessment, we use random development strategies to correspond with possible solutions proposed by global stochastic search algorithms, such as PSO. The assessment of the proposed method considered multiple conditions. The first condition was the total number of development strategies, J . Two values were studied, 200 and 400. For all results presented, the full set ensemble was defined to have a total of 100 realizations.

In addition, the effect of the proportion of J_T and J_P , for a given J , was also explored. Two splits between J_T and J_P were studied: a 10/90% split and 30/70%. The assessment will be applied to development strategies focussed on well control settings (for fixed well locations) and well placement strategies (with fixed well control settings) separately.

Lastly, for the purpose of initially assessing the proposed method, we use an exhaustive search to select the subset to eliminate any error or noise that may enter when an optimization-based method is used.

Well Control Settings. Fig. 3 shows the results for the two different splits of the 200 well control setting strategies (i.e., water injection rates for each well at 90-day intervals over a 10-year period). Each box plot summarizes the statistics for the predicted rank correlation coefficients corresponding to the most representative subsets found in the 10 random shuffles for different subset sizes. A general trend is seen, regardless of percentage split, that increasing the subset size improves the statistics of the predicted rank correlation coefficient. First, as the subset size increases, the median predicted rank correlation coefficient increases until plateauing when going from a subset size of 4–5. For example, using 30/70 split (left plot), a median of 0.957 is observed for a subset size of 1 and increases to 0.993 for a subset size of 5. This means the proposed method is able to select a subset based on the training rank correlation coefficient that produces a predicted rank correlation coefficient value, on average, of 0.993 (where perfectly associated is represented by a value of 1) for well control strategies.

Additionally, the interquartile range (IQR, i.e., $Q_3 - Q_1$; a measure of spread or variability) typically decreases as the subset size increases. For example, for the 30/70 split, a subset size of 1 has an IQR of 0.0235, while a subset size of 5 has a value of 0.00198. This is an order of magnitude decrease in the IQR. This indicates that for smaller subset sizes, there is a greater sensitivity to the way J_T development strategies are selected. Nonetheless, an interesting observation is that even using a single realization as a subset, a relatively high predicted rank correlation coefficient is achieved. For example, using a subset size of 1 gives minimum values (i.e., worst case of the 10 runs) for predicted rank correlation coefficient of 0.911 and 0.892 for splits of 30/70 and 10/90, respectively. Consequently, this gives the decision-maker greater ability to shift the balance between computational savings (by using a smaller subset) and ranking accuracy.

Fig. 4 presents the results of subset selection when using 400 well control setting strategies. Similar observations can be made to those displayed in Fig. 3, such as the improvement of the median predicted rank correlation coefficient with an increase in subset size. For example, the median value (over 10 subset selection shuffles) increases from 0.941 for a subset size of 1 (using a split of 30/70) to a value of 0.994 for a subset size of 5. In a similar manner, there is a noticeable decrease in the IQR, regardless of split, when there is an increase in subset size. For example, for a 30/70 split, the IQR for a subset size of 1 is 0.0259; in comparison, it is 0.0009 for a subset size of 5. This is consistent with results in Fig. 3 and shows that the method in which the J_T are selected is not important, especially for larger subset sizes.

An intriguing result is that the number of random development strategies (J) does not have a substantial effect on the ability of the rank correlation method to select subsets that produce rankings of the development strategies similar to the full set. This is evident when comparing the results between 200 and 400 random development strategies, with similar trends being observed. This gives an indication that this method can be expected to perform relatively consistently for various numbers of well control strategies, as would be the case in decision-making contexts or optimization frameworks.

Well Placement Strategies. Fig. 5 presents the results for 200 well placement strategies (i.e., x and y coordinates for 12 vertical wells) using two different splits for J_T / J_P . Similar to the well control setting results, there is a general trend of an increase in the predicted rank correlation coefficient as the subset size increases. However, the predicted rank correlation values are generally lower than the well control results. More specifically, for example, when using a subset size of 1 with a 30/70 split, the maximum value (out of the 10 selection runs) is 0.80. Similarly, using a 10/90 split, the maximum predicted rank correlation coefficient is 0.739.

This gives an indication that the increased nonlinearities of well placement strategies when compared with well control settings render these types of development strategies more difficult to rank accurately with a smaller subset. In other words, a larger number of

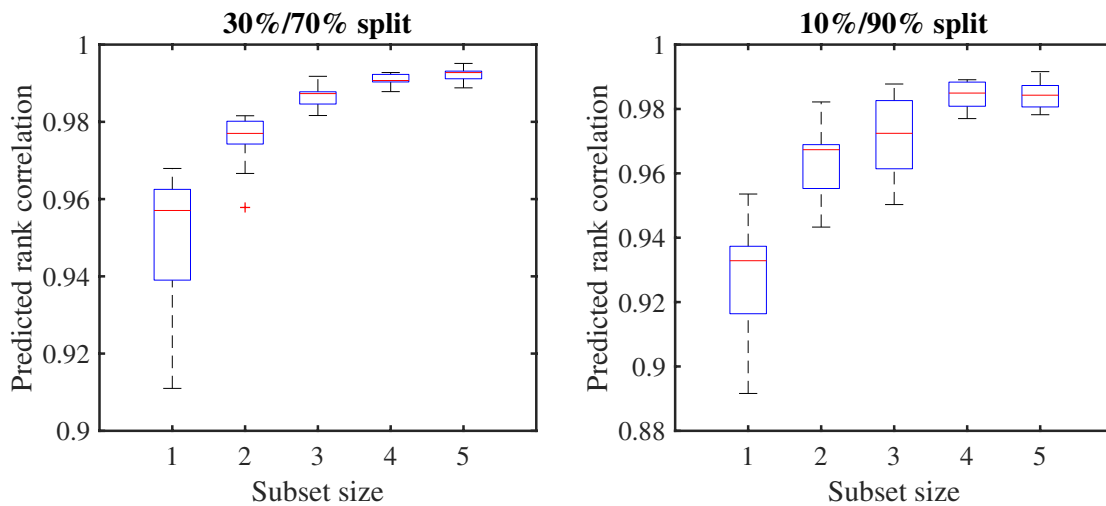


Fig. 3—Box plots for $J = 200$ random well control settings using different subset set sizes. Left: Using a J_T / J_P split of 30–70%. Right: Using a J_T / J_P split of 10–90%. The red line within each box represents the median, and the bottom and top lines of each box represent the 25th and 75th percentiles, respectively. The lines above and below the boxes represent the most extreme data points that are not considered outliers. The outliers are plotted individually using the cross symbol in red color.

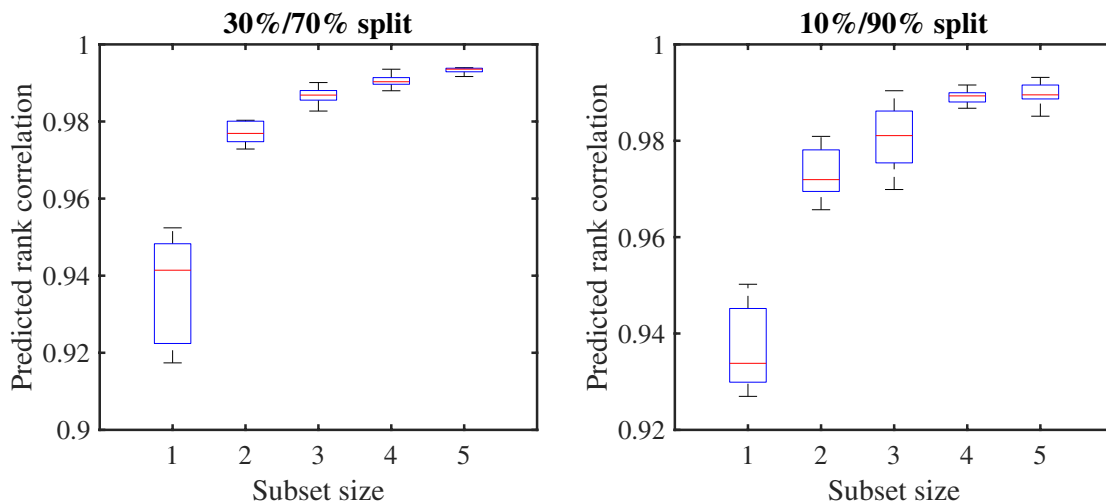


Fig. 4—Box plots for $J = 400$ random well control settings using different subset set sizes. Left: Using a J_T / J_P split of 30–70%. Right: Using a J_T / J_P split of 10–90%. The red line within each box represents the median, and the bottom and top lines of each box represent the 25th and 75th percentiles, respectively. The lines above and below the boxes represent the most extreme data points that are not considered outliers. The outliers are plotted individually using the cross symbol in red color.

realizations are required to obtain a more representative subset when ranking well placement strategies. This is reflected in the significant improvement in results as the subset size increases to 5. For example, the median predicted rank correlation for 30/70 and 10/90 splits are 0.911 and 0.906, respectively.

Unlike for well control strategies, there is not an obvious trend seen when looking at the IQR for either split. This may indicate that the nonlinearity of well placement development strategies overrides any possible effects that subset size may cause when selecting J_T strategies. It should be noted that the IQR values, even for well placement strategies, are still relatively low. For example, using a subset size of 5 with a 30/70 split, the IQR/median ratio is 0.0244/0.911, or 2.68%.

Fig. 6 summarizes the results for subset selection with a total of 400 random well placement strategies using two different splits against different subset sizes. Similar to the results presented in **Fig. 5**, overall, the predicted rank correlation coefficient values are lower than the results using 400 random well control strategies (see **Fig. 4**). Again, this can be attributed to the intrinsic differences in flow response of alternate well placement strategies compared with the differences between alternate well control settings. As such, finding a subset of size of 1 or 2 to represent a full set of 100 realizations may be unrealistic and counter productive. For example, when using a 30/70 split and a subset size of 1, the maximum predicted rank correlation coefficient (out of the 10 selection runs) is 0.754 and a median of 0.712. As the subset size increases to 5, the maximum rank correlation coefficient increases to 0.947 and a median of 0.914. A similar increase is seen for the 10/90 split going from a maximum predicted rank correlation of 0.723 (and a median of 0.682) for a subset size of 1 to a maximum value of 0.927 (and a median of 0.904) for a subset size of 5.

These results give insight into the difficulty of selecting a representative subset, especially for development strategies detailing well locations. It also indicates the importance of selecting representative subsets, rather than employing a random selection technique. Although randomly selecting a subset of realizations may result in the opportunity of a high-rank correlation coefficient, it is vital to

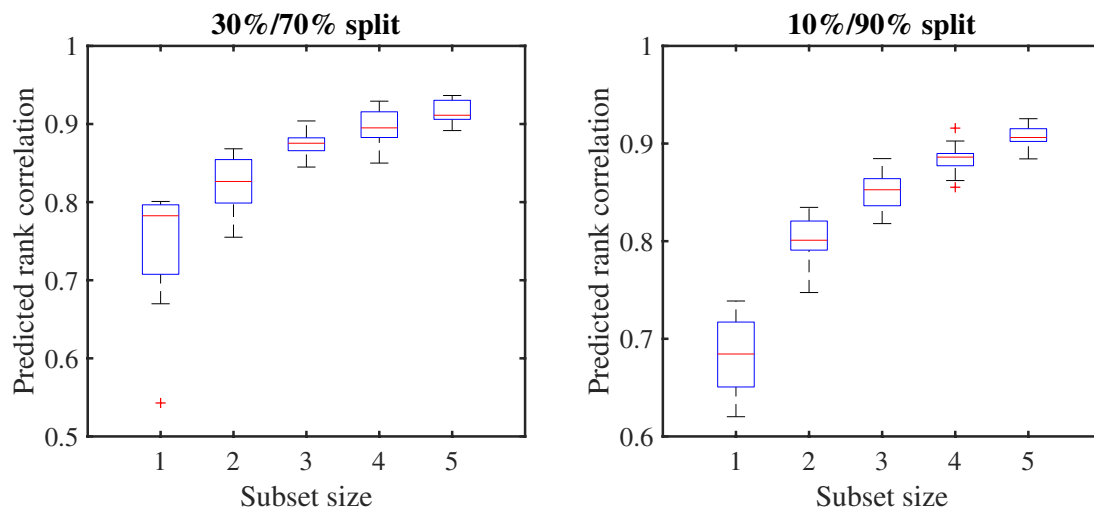


Fig. 5—Box plots for $J = 200$ random well placement strategies using different subset set sizes. Left: Using a J_T / J_P split of 30–70%. Right: Using a J_T / J_P split of 10–90%. The red line within each box represents the median, and the bottom and top lines of each box represent the 25th and 75th percentiles, respectively. The lines above and below the boxes represent the most extreme data points that are not considered outliers. The outliers are plotted individually using the cross symbol in red color.

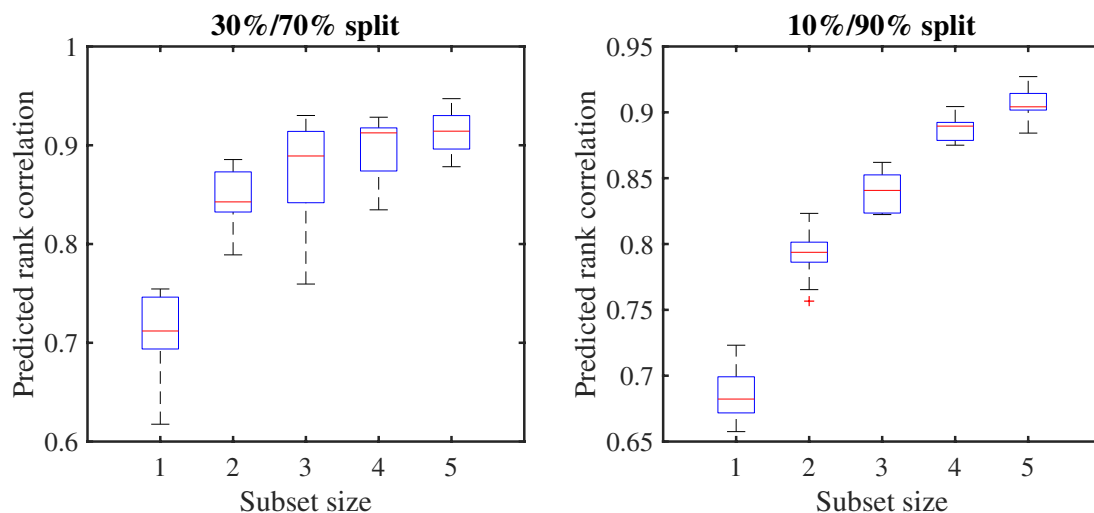


Fig. 6—Box plots for $J = 400$ random well placement strategies using different subset set sizes. Left: Using a J_T / J_P split of 30–70%. Right: Using a J_T / J_P split of 10–90%. The red line within each box represents the median, and the bottom and top lines of each box represent the 25th and 75th percentiles, respectively. The lines above and below the boxes represent the most extreme data points that are not considered outliers. The outliers are plotted individually using the cross symbol in red color.

consider the associated risk of selecting a subset with a low-rank value. This is exhibited in the range of the predicted rank correlation coefficients.

Analysis has shown that the proposed method is consistently able to select a subset that results in the predicted rank correlation coefficient to be in the top quartile of the range. However, for conciseness, these results are not presented. For example, for a 10/90 split using $J = 400$ and $n = 1$, the range in predicted rank correlation coefficient for one of the 10 runs is [0.242, 0.770] with Q1 and Q3 values of 0.446 and 0.603, respectively. For this run, the proposed method, using only 10 % of J development strategies as J_T , selects a subset with a predicted rank correlation coefficient value of 0.705.

If development strategies are ranked (whether in decision-making context or optimization framework) in a manner that is not consistent with the full set of realizations, this may lead to the selection of a poor alternative.

It is also worth mentioning the importance of testing different combinations for each subset size. We found that although a realization may be selected in a smaller subset size (e.g., $n = 1$), it does not necessarily mean it will be selected for larger subset sizes. This was seen when using well placement or well control development strategies.

Realization Subset Selection for Well Placement Optimization. In this section, we compare the results of well placement optimization when using the proposed subset selection technique with the results using the full set of realizations. We first define the problem by describing the decision variables, fixed well control settings, and field constraints considered. Additionally, we present the experimental investigation undertaken into various subset selection schemes for the implementation of the rank correlation method.

Problem Definition. In this problem, we consider well placement optimization, in which the areal locations (x and y coordinates) of 12 vertical wells (8 injectors and 4 producers) that maximize the decision metric (i.e., expected NPV) are sought. We continue with the use of the Egg model (Fig. 2) for this optimization problem. Note that the decision variables are normalized and bounded between 0 and 1. The production wells are operated by bottomhole pressure with settings of 395 bars. The injection wells are controlled by a surface rate of $60 \text{ sm}^3/\text{d}$ and a maximum bottomhole pressure of 405 bars.

Additionally, two field constraints were considered to ensure practical solutions were proposed. First, an interwell distance constraint of 24 m (approximately three grid blocks) was used. Second, a projection-type repair mechanism was used to project the proposed well locations outside the reservoir bounds back into the reservoir. The reservoir boundary was approximated using piecewise linear functions to represent a reservoir polygon. If a proposed solution places a well outside this polygon, the well is projected back onto the location on the boundary with the shortest distance to the violating well location [details of the constraint handling can be found in Arouri and Sayyafzadeh (2020b)]. Table 1 presents the economic parameters used for calculating u and u' (i.e., expected NPV).

The optimization using the proposed subset selection method (see Algorithm 1) was undertaken using various values for R_{pt} . This threshold value, in essence, gives an indication of the quality of the subset and its ability to rank well placement strategies. In turn, this determines if the subset selection process should be initiated or not (see line 17, Algorithm 1). Four values of R_{pt} were investigated: 0.8, 0.85, 0.9, and 0.95. These values were selected based on the results presented in the previous section. Similarly, a value of 0.99 was chosen for R_t as the previous results showed that when a subset with a training rank correlation coefficient of 0.99 was selected, the predicted rank correlation coefficient (of that subset) was above 0.92 in 93% of the runs. This threshold value is important as it determines which subset size (and hence subset) is selected based on the training rank correlation coefficient in the selection process (see line 21, Algorithm 1). This gives the subset size an adaptive nature and hence can play a role in producing an efficient optimization run (compared with a static subset size). The initial subset was chosen for all three seeds using a training threshold of 0.99, which had a size of three realizations, and was used for all subset optimization runs. It should be noted as the initial populations for different, each seed had a different initial subset. This selection was based on the first generation of optimization (i.e., using 50 random solutions representing J_T), where all the solutions are evaluated using the full set (which requires 5,000 reservoir simulations), creating the FMM, and these evaluations are accounted for in the convergence plot of the runs with the subsets.

In comparison, the full set optimization used all 100 realizations to calculate the exact decision metric, u , at each generation. Unsurprisingly, this entails a large computational burden, hence the importance of using a subset. For the full set optimization, a maximum of 75,000 reservoir simulations was allowed. On the other hand, given the computational burden, a maximum of 25,000 reservoir simulations was allowed for the optimization runs using a subset. The values for the PSO algorithmic parameters, namely, ω , c_1 and c_2 , were taken as 0.721, 1.193, and 1.193 based on Onwunalu and Durlofsky (2010) using a population size of 50. In addition, to remove the effect of the stochastic nature of PSO, full set and subset runs used the same seeds.

Optimization Results. Fig. 7 presents the results comparing subset optimization using various subset selection schemes (R_{pt} values) against the full set optimization. To account for the randomness of PSO, the experiments were repeated using three different seeds. The presented expected NPV values in this figure and any subsequent figures are obtained by evaluating the development scenario using the full set [i.e., $u(s_g)$]. This allows the comparison of optimum development scenarios found in the subset and full set optimizations.

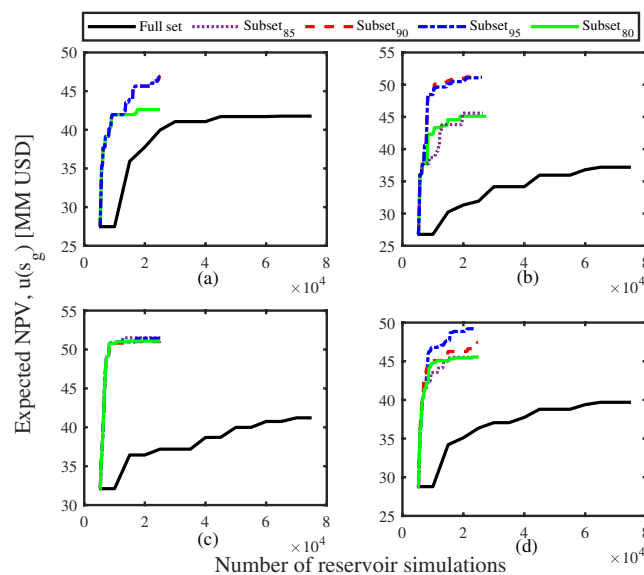


Fig. 7—Well placement optimization results comparing subset optimization using the proposed method to the full set optimization for (a) Seed 1, (b) Seed 2, (c) Seed 3, and (d) average. Various training thresholds (R_t) were used to understand the effect on the progress of the optimization.

First, in all three seeds, the subset optimization runs are able to outperform the full set optimization with respect to computational efficiency and objective function value, regardless of the threshold value. More specifically for Seed 1 (see Fig. 7a), the full set optimization converges to an objective function value of 41.71 MM USD after 45,000 reservoir simulations. In comparison, all subset optimization runs, regardless of R_{pt} value, obtain a similar objective function value (41.64 MM USD) after only 9,000 reservoir simulations. This translates to a reduction in computational expense by a factor of 5 to achieve the same objective function value. Assuming each reservoir simulation takes 10 minutes (for a modest reservoir model) to execute, that is a savings of up to 6,000 hours when optimizing using subsets selected by the proposed method.

Similar comparisons can be made from the results of Seed 2 (Fig. 7b) and Seed 3 (Fig. 7c). For example, in the Seed 2 runs, the full set optimization obtains an NPV value of 37.19 MM USD after 75,000 reservoir simulations. In comparison, the best subset optimization

run (using an R_{pt} of 0.90) obtained an NPV value of 51.21 MM USD after only 20,910 reservoir simulations. Comparably, using Seed 3, the full set optimization obtained an NPV value of 41.21 MM USD after 75,000 reservoir simulations, while the best subset run (using an R_{pt} of 0.85) converged to an NPV value of 51.5 MM USD after 25,100 reservoir simulations.

Second, the results in Fig. 7 allow the comparison of the different subset (R_{pt} values) runs. Note that the subset runs using an R_{pt} of 0.8 and 0.85 result in the same convergence plot and hence are overlapping in Fig. 7a. As a reminder, the R_{pt} value determines whether or not a subset reselection procedure is undertaken. Consequently, it is possible that an initial subset performs above this threshold throughout the optimization procedure. An example of this is reflected in the subset optimization runs using R_{pt} values of 0.80 and 0.85 in Fig. 7a. This means that the optimization occurred using the initial subset selected throughout the process (i.e., a static subset). However, this is still more computationally efficient than using the full set, as these runs converge to an expected NPV of 42.61 MM USD after 17,500 reservoir simulations using Seed 1. This translates to up to a 61% reduction in computational costs with up to a 2% increase in expected NPV. Similar computational savings are seen for seeds 2 and 3 in Figs. 7b and 7c.

The optimization runs using higher R_{pt} values (namely, 0.90 and 0.95) typically outperform the other subset selection runs and the full set run significantly, except in Seed 3 where the performances are similar. For example, in Fig. 7a, the run using a value of 0.95 for R_{pt} converged to an expected NPV value of 46.78 MM USD after 24,510 reservoir simulations, while the run using a value of 0.90 performs in a comparable manner as it converged to an expected NPV of 46.92 MM USD after 24,710 reservoir simulations. These results illustrate up to a 13% increase in expected NPV with a computational cost reduction of up to 52% compared with the full set run. Furthermore, when compared with the subset runs using R_{pt} values of 0.80 and 0.85, these results represent up to a 10% increase in expected NPV.

Fig. 7d presents the average of all three seeds. On average, it is clear the subset optimization runs all outperform the full set optimization. For example, on average, the full set optimization converged to an NPV value of 36.69 MM USD, while the best performing subset run ($R_{pt} = 0.95$), converged to an average of 49.21 MM USD. This represents, on average, an improvement of 24% in NPV value between the best subset run and the full set run. It is also evident from Fig. 7d that the subset runs are, as expected, much more computationally efficient, where a reduction of computational costs is close to a factor of 9.

Evolution of Rank Correlation Coefficient. Fig. 8 shows the subset optimization runs (left y-axis) alongside the evolution of the predicted rank correlation coefficient (right y-axis) for Seed 1. For succinctness, we focus on the results of Seed 1 as they are representative of the two other seeds. This figure gives insights into the effect of various R_{pt} values on the search progression. As previously mentioned, the optimization runs using values of 0.80 and 0.85 do not undergo subset selection throughout the optimization. This is reflected in the fact that the predicted rank correlation coefficient (i.e., R_p) for these runs does not fall below the threshold values (0.85 or 0.80, respectively). As the optimization progresses, the predicted rank correlation coefficient for these runs reaches a maximum value of 0.962 and then stabilizes around 0.90. This stabilization is reflected in the optimization results with the convergence to an expected NPV of 42.61 MM USD.

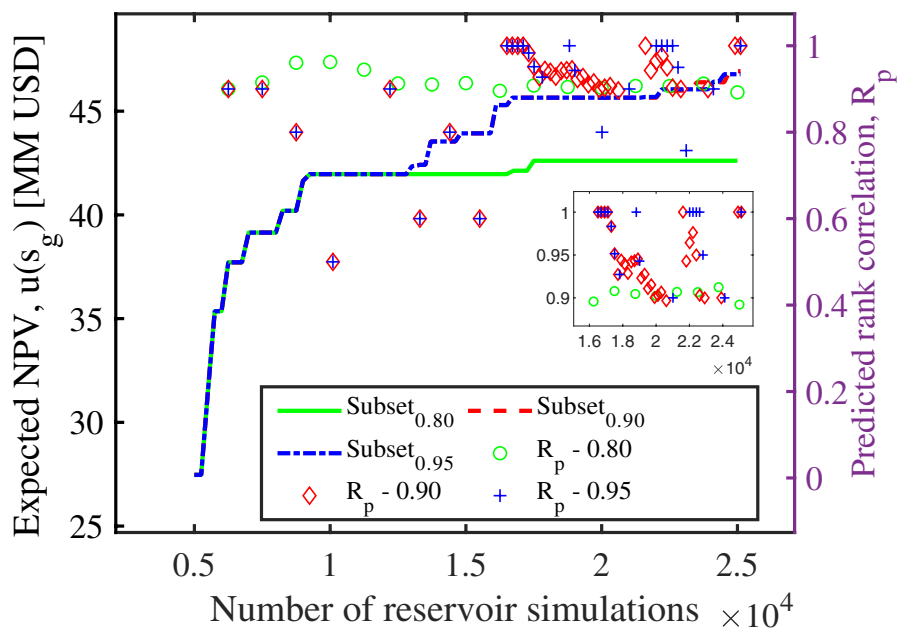


Fig. 8—Well placement optimization results comparing the evolution of the predicted rank correlation coefficient (R_p) for each threshold value (R_{pt}). Left y-axis is the expected NPV (MM USD) and the right y-axis is the predicted rank correlation coefficient. Inset figure is a zoom-in showing differences in the predicted rank correlation coefficient in the later stages of the optimization process.

In comparison, the other two subset runs have higher thresholds and as such undergo the subset selection procedure multiple times in the early stages of the optimization. This is reflected by the discontinuous rank correlation coefficient values. In the early stages, the subset selection results in the same subset as the initial subset, and hence the three convergence plots overlap. The first divergence between the subset runs occurs around 10,000 reservoir simulations which represents the selection of a different subset for the two runs ($R_{pt} = 0.9$ and 0.95). It is also important to note, that the proposed adaptive scheme allowed the reduction in the subset size from 3 to 2 (i.e., a reduction of 33% in computations per generation). A number of subset selections are undertaken (represented by singular predicted rank correlation values) until a subset with a predicted rank correlation coefficient of 1 is selected around 17,000 reservoir simulations (for both runs), shown in detail in the inset of Fig. 8. Soon after, a divergence between the 0.90 and 0.95 runs occurs as indicated by a predicted rank correlation value of 0.927. That is, the threshold is met for the 0.90 run; however, it is not for the 0.95 run and a subset selection procedure

is undertaken resulting in a new subset being selected. In subsequent generations, the differences in the predicted rank correlation values are shown in the inset of Fig. 8. In the later stages of these runs, the subset selection results in predicted rank correlation coefficients of 1. As the optimization progresses, these predicted rank correlation coefficient values gradually decrease, shown by continuous values, compared with earlier stages that saw discontinuous values for the higher threshold runs.

Discussion

Decision-making processes rely on ranking a number of alternatives (solutions) using decision metrics with the aim of selecting the most value-creating option for the given objectives. In field development planning, this process is complicated with the incorporation of geological uncertainty as it increases the computational overhead required. As such, it is crucial to find subset selection techniques that reduce the required computation and yet maintain the rankings of possible development strategies (solutions). This helps ensure robust decisions can be made under uncertainty. We argue that a reformulation of the subset selection problem is needed to propose a technique that allows for robust decision-making. This entails refocusing the aim of subset selection to ensure the realizations are able to rank (and hence select) alternatives (development strategies in the case of field development) in a similar manner to the full set.

Assessment of Underlying Assumption. The overarching distinction between previous subset selection techniques and the proposed method is the underlying assumption. Methods utilizing static geological properties, such as the static ranking method, assume that the CDF of a geological static property is representative of the CDF of the output (e.g., production forecasts or NPV). Other methods use flow-based vector/s from a selected (few) base case development strategy (strategies). The underlying assumption in these methods is that the CDF of the output for a base case development strategy is representative of the CDF of the output for any other development strategy.

In the proposed method, the underlying assumption is that there is a positive relationship between the rank correlation coefficient based on J_T development strategies and the rank correlation coefficient for the other strategies in J . Although this assumption cannot be directly assessed, a number of indicators can be examined. Fig. 9 shows the minimum predicted rank correlation coefficient for a specific training rank correlation coefficient. The results are presented for a number of test cases of full set size (N), subset size (n), and split of J . A minimum predicted rank correlation coefficient represents the worst-case scenario for the associated training rank correlation coefficients. For example, selecting a subset with a training rank correlation value of 0.9, one could expect a minimum predicted rank correlation coefficient of 0.74 for the tested cases or even a minimum as high as 0.86.

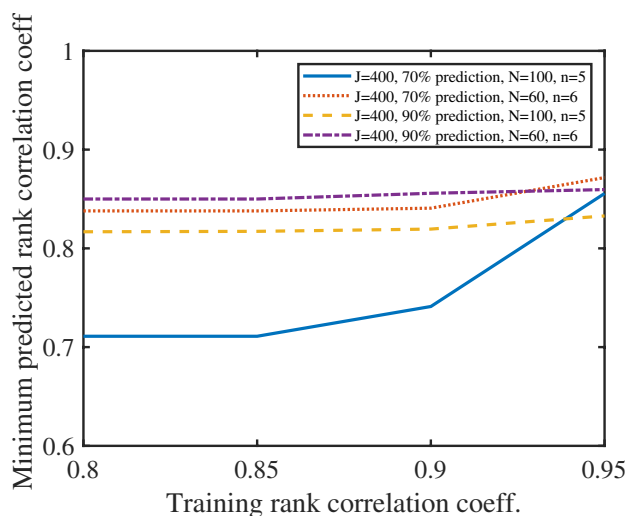


Fig. 9—Minimum rank correlation coefficient of predicted values for corresponding subsets at various training rank correlation coefficient threshold values.

The trend in Fig. 9 shows that as the training rank correlation coefficient increases, the minimum associated rank correlation coefficient also increases. This indicates a positive association between the training rank correlation coefficient and the prediction value for subsets. Such an association cannot be made using a random subset selection method as it is just as likely to select a subset with a low-rank correlation coefficient as it is to select a subset with a high-rank correlation coefficient. Although the results in Fig. 9 are not expansive, they give insight into the validity of the assumption for the proposed technique.

Effect of Development Strategy. The comparison of the results between the two types of development strategies (well placement and well control settings) brings to light some important findings. Because of the relatively smoother objective function landscape of well control development strategies (Arouri and Sayyafzadeh 2020a), the subsets of realizations tend to have rank correlation coefficients close to unity, even for smaller subset sizes. This indicates that individual realizations inherently have a relatively close association to the full set (and one another) with regards to flow-response for different well control development strategies. In turn, this may indicate that the parameters that were deemed uncertain when generating the realizations do not have a significant impact across different well-operating conditions. In comparison, the well placement development strategies have a rougher landscape because of the highly nonlinear relationship between well locations and flow response. These findings are corroborated by Sayyafzadeh and Alrashdi (2019) and Shirangi and Durlofsky (2016) who found similar differences when selecting subsets for well control optimization compared with well placement optimization. Consequently, larger subset sizes are required to ensure the subset has a strong positive association with the full set.

However, as larger subset sizes are selected or larger full set sizes are used, the computational overhead required for exhaustive search using the proposed method becomes unjustifiable. To this end, the use of an integer-based population optimization algorithm will alleviate these computational costs without compromising in finding the subset which maximizes the training rank correlation coefficient. Two

population optimization algorithms were tested for this purpose: genetic algorithm (GA) and PSO. The underlying evolutionary formulations of each algorithm are consistent with the common implementation. Details of GA and PSO can be found in Srinivas and Patnaik (1994) and Kennedy and Eberhart (1995), respectively. The integer GA implemented used a single-point crossover, a single gene mutation, and a random tournament selection scheme. The PSO algorithm implemented a global (star) topology.

The optimization problem is formulated to minimize the distance between ranked vectors for the full set, r_T , and subset, $r'_T(X_{ss})$. The decision variables are the ordinal numbers that represent the realizations in the full set. For example, using a subset size of 5 and a full set size of 100, one possible set of decision variables would be (1, 44, 23, 95). Once a new population is generated, the individuals are rounded to the nearest integer values. Next, to reduce redundant objective function evaluations, newly generated individuals are compared with each other as well as with previous generations. If a combination of integers (i.e., a subset) has been evaluated previously, a new set of integers is generated. This would allow the innards of PSO/GA to work as per usual without much change. Preliminary testing showed that for a subset size of 5, both algorithms were able to find the same subset that maximized the training rank correlation coefficient as the exhaustive search in a limited number of function evaluations.

As shown in Figs. 5 and 6, there is clearly a need for a larger subset size to improve the predicted rank correlation coefficient when using well placement strategies. This will provide further insight into the general trend seen in other results—that an increase in subset size will generally improve the predicted rank correlation coefficient. As such, integer GA was applied to well placement development strategies using $J = 400$ with a split of 30/70. The results are presented in Fig. 10 and are a summary of the 10 shuffles of J_T as was done before. The results show that it is possible to obtain high predicted rank correlation coefficients for well placement strategies. For example, to obtain a predicted rank correlation coefficient of 0.96, a subset size (n) of 10 is required. A subset size of 10 represents a 90% reduction from the full set size of 100 realizations. As one would expect, as the subset size approaches the full set size, the predicted rank correlation coefficient approaches a value of unity (i.e., perfect association). In addition, similar to previously presented results (using exhaustive search), there is minimal sensitivity to the way J_T development strategies are selected are represented by the small IQRs.

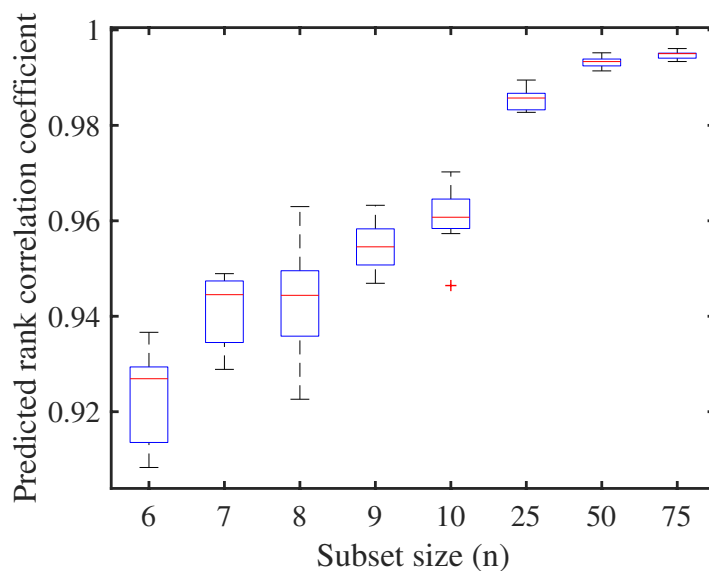


Fig. 10—Box plots for $J = 400$ random well placement strategies using different subset set sizes and a J_T / J_P split of 30–70%. These results are obtained using an integer-based GA using a population size of 100, a crossover probability (P_c) of 0.9, mutation probability (P_m) of 0.05, and tournament size of 4. The red line within each box represents the median, and the bottom and top lines of each box represent the 25th and 75th percentiles, respectively. The lines above and below the boxes represent the most extreme data points that are not considered outliers. The outliers are plotted individually using the cross symbol in red color.

Traversing the Objective Function Landscape. An interesting observation from the optimization results is the possible effect that selecting a subset has on the ability of PSO (and possibly any population-based algorithm for that matter) to traverse the objective function landscape. When the results in Fig. 7 are plotted against iteration number (for brevity not shown) rather than the number of reservoir simulations, there is a noticeable difference in convergence speeds between the subset and the full set. This gives an indication that there may be more at play than solely a computational difference because of the substantial reduction in the number of realizations used. One explanation may be that when using the full set (of 100 realizations), the landscape is more complex, with a larger number of local minimums, which results in a more rugged landscape. This makes it more difficult for PSO to traverse the landscape in search of the global minimum and hence slow down convergence when compared with the subset runs.

In addition, PSO and other population-based algorithms are characterized by the dichotomy of exploration and exploitation. A balance between these two aspects plays a significant role in the success of such algorithms. During the exploratory stages of optimization (specifically in the early iterations), the aim is to visit new regions in the search space. As such, it is important to traverse the landscape in a computationally efficient manner (i.e., using a low number of iterations). When using the full set, the increased complications and ruggedness of the landscape may cause a slow-down in the exploratory stage. This may be because of the increased number of local minima which act as speed bumps for the algorithm. On the other hand, when a subset is used, the landscape ruggedness is reduced and a lower number of local minima are present which allows faster exploration to take place.

In the early stages of the optimization, the importance lies in searching potential regions in the search space. As the optimization progresses, the population begins to converge, and the importance intrinsically shifts to searching previously visited successful regions, the selection of the subset is still crucial. The adaptive nature of the subset (discussed in detail in the next section) allows the selection of a

subset that closely mimics the major local minimums (i.e., has a high-rank correlation coefficient). This becomes pertinent during the exploitation of the search space as the populations converge.

Adaptive Nature of Subset Selection. The optimization results provide insight into the importance of an adaptive scheme for subset selection when implemented within an optimization framework. An adaptive approach uses feedback from the optimization process to determine whether or not a new subset selection process should be undertaken. This allows the realizations within a subset as well as the subset size to be updated. A possible drawback of using a subset of realizations for decision-making is the possibility of selecting a nonrepresentative subset (i.e., a subset of realizations that does not allow for consistent and robust decision-making, when compared with the full set). However, the proposed technique is able to ensure consistent ranking (and hence selection) of the proposed development strategies within an optimization framework. This is evident by the fact that all subset selection schemes (i.e., various R_{pt} values) were able to achieve an objective function value [evaluated by the full set - $u(s_g)$] similar to that of the full set with significant computational savings. This means that the optimization algorithm (i.e., decision-maker) is able to decide (i.e., rank and select) at each generation the development strategies that are consistent with the full set.

Another interesting insight is the effect that the frequency of subset selection can have on the optimization progress. First, it is important to note that the frequency of an adaptive subset selection can distinguish between subset optimization runs; however, it is clear all subset runs outperform the full set optimization. In the proposed technique, the frequency of subset selection is not static (e.g., every five generations); rather, it is based upon feedback from the optimization process (i.e., using a threshold value). This means that a subset selection process (and hence possible subsequent change of subset) only occurs if the optimization process requires it. This in effect is a cost-saving measure in itself, as redundant subset selections may unnecessarily increase the number of evaluations of the full set.

The frequency of subset (re)selection within an optimization process can either be a hindrance or an advantage. As the subset is changed, the intrinsic objective function landscape that is being explored is changed. When using a static subset (i.e., one subset) throughout the whole optimization process, it is only this subset that is being optimized. In other words, given the multimodal nature of field development objective landscapes, the search most likely converges to a local minimum/maximum of that specific subset. This local minimum/maximum may not be represented in the full set landscape and as such be suboptimal. An example of this is displayed when using a R_{pt} value of 0.85 (or 0.80) in Fig. 7a, where the initial subset was used for the whole optimization process. Similar results would be expected for subset selection techniques, such as static ranking and clustering techniques, which use a single subset throughout the optimization process. On the other hand, undergoing reselection too frequently may introduce additional unnecessary noise into the objective function. In such a scenario, the algorithm may be unable to converge to any local minimum/maximum. This would decrease the computational efficiency dramatically. Although randomness may be useful to escape local minima/maxima, excess stochasticity may become counter productive when under computational constraints—as is usually the case for decision-making scenarios.

The proposed adaptive technique, based on rank correlation coefficient, is not susceptible to these issues when suitable R_{pt} values are used. In actuality, the subset optimization runs which undergo reselection (i.e., have higher R_{pt} values) are able to find solutions with higher objective function values. This may be because as reselection occurs the more promising regions (i.e., local maxima/minima), which are consistently present in the selected subset realizations, are accentuated. This is aided by the use of feedback from the optimization, formalized in the R_p value, which enables the algorithm to delay convergence by reselecting a subset of realizations as it approaches local subset optima. This is seen prominently in Fig. 8, where in later stages, a plateau arises but is soon overcome after a successful reselection. This feedback also ensures that the selection process is undertaken once the current subset is unable to rank the proposed development strategies in a manner similar to the full set.

Compared with other subset selection methods, another advantage of the proposed technique is the fact that it is problem independent. It does not rely on selecting a static property as is required for static ranking and clustering methods, nor does it rely on base case scenarios for flow-response vectors. Additionally, it is applicable to any type of development strategy (i.e., well placement or well control settings), without the need to change any weighting factors, simulate different base case scenarios, or reselect static properties. It should be noted that the proposed technique does require an initial computational investment to select an initial subset; however, this is not unusual for subset selection techniques that incorporate some form of flow-response vectors. If computational budgets are under stringent constraints, then it is reasonable to use proxy flow simulations, such as streamline simulations (Batycky et al. 1997) or capacitance-resistance models (Sayarpour et al. 2009) for the selection of the initial subset.

Lastly, the proposed technique works for any decision metric (i.e., I) that may be used within a decision-making process. In addition to using expected NPV as the decision metric, we also assessed the proposed method using the expected utility for a risk-averse decision-maker. The insights obtained from these experiments are akin to those presented and as such, for conciseness, are not presented in this work.

Conclusions

In this work, we began by reformulating the subset selection problem to a more relaxed approximation with the aim of selecting a subset that can rank (and hence select) development strategies in a consistent manner with the full set. We also proposed a subset selection method to solve this problem that is based on finding the combination of realizations that maximizes the rank correlation coefficient for a selected number of solutions (development strategies).

Furthermore, a realignment of the aim to focus on the rankings of development strategies ensures a computational reduction without compromising the robustness of the decision. This work assessed the ability of the proposed technique to select representative subsets for various numbers of random development strategies and subset sizes. The results showed that the rank correlation subset selection technique is able to consistently select a subset that has a strong positive association with the full set of realizations, regardless of development strategy type (well placement or well control). The proposed method is able to select subsets that produce rank correlation coefficients of up to 0.993 for well control settings using a subset size of only 5 (for a full set size of 100). Furthermore, the proposed method is able to select subsets for well placement strategies with rank correlation coefficients of up to 0.947 for a subset size of 5. We found that smaller subsets are able to successfully rank well control development strategies, while well placement development strategies require larger subset sizes to obtain similar rank correlation coefficients. This is owing to the differences in flow response of geological models to different well placement strategies because of the underlying increased nonlinearities. We also found that the way in which (J_T) development strategies (from J) are selected to be ranked has little to no bearing on the method's ability to select a representative subset. For situations with larger subset sizes or larger full set sizes, we have shown the applicability of integer-based optimization algorithms (e.g., integer GA) for finding the best subset for the training set.

Additionally, we applied the proposed technique within a population-based algorithm for the well placement optimization problem. This problem studied the placement of 12 wells (8 injectors and 4 producers) under geological uncertainty for the Egg benchmark. Implementing the proposed subset selection technique resulted, on average, in a computational reduction of close to a factor of 9 when compared with the full set optimization. Moreover, subset optimization runs that underwent subset reselection during the optimization procedure saw improvements in expected NPV of up to 24%, on average, compared with the final expected NPV found by the full set optimization. The adaptive nature of the proposed technique allows the update of realizations in the subset, as well as the subset size, to take advantage of potential additional computational savings without compromising on the achieved objective function value. Our results also showed that the frequency of subset reselection during an optimization procedure will affect the convergence rate. As reselecting a subset too often could introduce unnecessary noise into the objective function, an infrequent reselection may result in the premature convergence to a subset local optima.

This research can take multiple different directions in the future. One direction would be to expand the application of the proposed technique to problems involving joint optimization of well placement and well control. The combination of these decision variables could mean a greater difference in flow response between realizations and as such result in more local optima. This would mean selection of suitable representative realizations would be more pertinent. It would also be interesting to compare the proposed technique with other subset selection techniques which use static properties, flow-response vectors, or a combination of both. Finally, the adaptation of the proposed technique for gradient-based methods would be of interest to the research community.

Nomenclature

- A_p = prediction matrix
- A_t = training matrix
- b = annual discount rate
- c_1 = cognitive coefficient
- c_2 = social coefficient
- c_o = oil price per unit volume from time t_i to t_{i+1} , USD/STB
- c_{wi} = water injection costs per unit volume from time t_i to t_{i+1} , USD/STB
- c_{wp} = water separation costs per unit volume from time t_i to t_{i+1} , USD/STB
- $d(X_{ss})$ = difference between the rankings of the full set and the subset when J development strategies are used
- $d_T(X_{ss})$ = difference between the rank vectors of the full set and the subset when J_T solutions are used
- $F_{Y(s)}$ = CDF corresponding to the evaluation of development strategy using N (full set) realizations
- $F'_{Y(s)}(X_{ss})$ = CDF corresponding to the evaluation of development strategy using X_{ss} subset
- g = governing equations for fluid flow
- J = total number of all possible development strategies (solutions)
- J_T = batch of selected development strategies
- l = decision metric function
- m = geological model
- M = random vector representing all geological models
- n = number of geological realizations in subset
- N = number of geological realizations in full set
- N_i = timestep intervals
- p = production forecasts
- Q_o = field oil production volume from time t_i to t_{i+1} , STB
- Q_{wi} = field water injection volume from time t_i to t_{i+1} , STB
- Q_{wp} = field water production volume from time t_i to t_{i+1} , STB
- r = rank vector for J solutions obtained using the full set
- r' = rank vector for J solutions obtained using the X_{ss} subset
- r_T = rank vector for batch of J_T solutions obtained using the full set
- $r'_T(X_{ss})$ = rank vector for batch of J_T solutions obtained using X_{ss} subset
- R_p = calculated predicted rank correlation coefficient
- R_t = calculated training rank correlation coefficient
- R_{pt} = threshold value for predicted rank correlation coefficient
- R_{tt} = threshold value for training rank correlation coefficient
- $R(X_{ss})$ = Spearman's rank correlation coefficient when using X_{ss} subset
- s = field development strategy (well control settings or well locations/placements)
- s_g = global best solution
- s_j = proposed solution (development strategy) in Algorithm 1
- s_{gk} = best solution found in generation k
- s_{pj} = particle's personal best solution
- t = time since start of production, years
- u = decision metric value
- $u_{Y(s)}$ = decision metric obtained by evaluating the development strategy using N (full set) realizations
- $u'_{Y(s)}(X_{ss})$ = decision metric obtained by evaluating the development strategy using the X_{ss} subset
- v = economic parameters
- X_{ss} = binary vectors of 0's and 1's defining which realizations (from the full set) are selected as the subset
- y = fitness (economic) metric
- Y = random variable/vector representing an uncertain fitness metric for given field development strategy s
- ω = inertia weight

Acknowledgements

The authors would like to thank Rock Flow Dynamics for providing licenses for tNavigator. This work was supported with supercomputing resources provided by the Phoenix HPC service at The University of Adelaide. This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

References

- Alpak, F. O., Barton, M. D., and Caers, J. 2010. A Flow-Based Pattern Recognition Algorithm for Rapid Quantification of Geologic Uncertainty. *Comput Geosci* **14** (4): 603–621. <https://doi.org/10.1007/s10596-009-9175-5>.
- Alrashdi, Z. and Sayyafzadeh, M. 2019. +1 Evolution Strategy Algorithm in Well Placement, Trajectory, Control and Joint Optimisation. *J Pet Sci Eng* **177**: 1042–1058.
- Arouri, Y. and Sayyafzadeh, M. 2020a. An Accelerated Gradient Algorithm for Well Control Optimization. *J Pet Sci Eng* **190**: 106872. <https://doi.org/10.1016/j.petrol.2019.106872>.
- Arouri, Y. and Sayyafzadeh, M. 2020b. Adaptive Moment Estimation Framework for Well Placement Optimization. In *ECMOR XVII*, 1–15. Houten, The Netherlands: European Association of Geoscientists & Engineers. <https://doi.org/10.3997/2214-4609.202035165>.
- Artus, V., Durlafsky, L. J., Onwunali, J. et al. 2006. Optimization of Nonconventional Wells under Uncertainty Using Statistical Proxies. *Comput Geosci* **10** (4): 389–404. <https://doi.org/10.1007/s10596-006-9031-9>.
- Ballin, P. R., Journel, A. G., and Aziz, K. 1992. Prediction Of Uncertainty In Reservoir Performance Forecast. *J Can Pet Technol* **31** (4). PETSOC-92-04-05. <https://doi.org/10.2118/92-04-05>.
- Batycky, R. P., Blunt, M. J., and Thiele, M. R. 1997. A 3D Field-Scale Streamline-Based Reservoir Simulator. *SPE Res Eng* **12** (4): 246–254. SPE-36726-PA. <https://doi.org/10.2118/36726-PA>.
- Begg, S., Bratvold, R., and Campbell, J. 2002. The Value of Flexibility in Managing Uncertainty in Oil and Gas Investments. Paper presented at the SPE Annual Technical Conference and Exhibition, San Antonio, Texas, USA, 29 September–2 October. SPE-77586-MS. <https://doi.org/10.2118/77586-MS>.
- Bouzarkouna, Z., Ding, D. Y., and Auger, A. 2012. Well Placement Optimization with the Covariance Matrix Adaptation Evolution Strategy and Meta-Models. *Comput Geosci* **16** (1): 75–92. <https://doi.org/10.1007/s10596-011-9254-2>.
- Bratvold, R. and Begg, S. 2010. *Making Good Decisions*. Richardson, Texas: Society of Petroleum Engineers.
- Chen, C., Li, G., and Reynolds, A. C. 2012. Robust Constrained Optimization of Short- and Long-Term Net Present Value for Closed-Loop Reservoir Management. *SPE J.* **17** (3): 849–864. SPE-141314-PA. <https://doi.org/10.2118/141314-PA>.
- Chen, Y., Oliver, D. S., and Zhang, D. 2009. Efficient Ensemble-Based Closed-Loop Production Optimization. *SPE J.* **14** (4): 634–645. SPE-112873-PA. <https://doi.org/10.2118/112873-PA>.
- Deutsch, C. V. and Begg, S. 2001. The Use of Ranking to Reduce the Required Number of Realizations. Report, Centre for Computational Geostatistics (CCG), Edmonton, Alberta, Canada.
- Fonseca, R. R., Chen, B., Jansen, J. D. et al. 2017. A Stochastic Simplex Approximate Gradient (StoSAG) for Optimization under Uncertainty. *Int. J. Numer. Meth. Engng* **109** (13): 1756–1776. <https://doi.org/10.1002/nme.5342>.
- Fonseca, R., Leeuwenburgh, O., Della Rossa, E. et al. 2015. Ensemble-Based Multiobjective Optimization of On/Off Control Devices Under Geological Uncertainty. *SPE Res Eval & Eng* **18** (4): 554–563. SPE-173268-PA. <https://doi.org/10.2118/173268-PA>.
- Haghighat Sefat, M., Elsheikh, A. H., Muradov, K. M. et al. 2016. Reservoir Uncertainty Tolerant, Proactive Control of Intelligent Wells. *Comput Geosci* **20** (3): 655–676. <https://doi.org/10.1007/s10596-015-9513-8>.
- Jansen, J. D., Fonseca, R. M., Kahrobaei, S. et al. 2014. The Egg Model - A Geological Ensemble for Reservoir Simulation. *Geosci. Data J* **1** (2): 192–195. <https://doi.org/10.1002/gdj3.21>.
- Jesmani, M., Jafarpour, B., Bellout, M. et al. 2016. Application of Simultaneous Perturbation Stochastic Approximation to Well Placement Optimization under Uncertainty. In *ECMOR XV - 15th European Conference on the Mathematics of Oil Recovery*, cp-49400133. Houten, The Netherlands: European Association of Geoscientists & Engineers. <https://doi.org/10.3997/2214-4609.201601873>.
- Jesmani, M., Jafarpour, B., Bellout, M. C. et al. 2020. A Reduced Random Sampling Strategy for Fast Robust Well Placement Optimization. *J Pet Sci Eng* **184**: 106414. <https://doi.org/10.1016/j.petrol.2019.106414>.
- Li, S., Deutsch, C. V., and Si, J. 2012. Ranking Geostatistical Reservoir Models with Modified Connected Hydrocarbon Volume. Paper presented at the Ninth International Geostatistics Congress, Oslo, Norway, 11–15 June.
- Lorentzen, R. J., Shafieirad, A., and Naevdal, G. 2009. Closed Loop Reservoir Management Using the Ensemble Kalman Filter and Sequential Quadratic Programming. Paper presented at the SPE Reservoir Simulation Symposium, The Woodlands, Texas, USA, 2–4 February. SPE-119101-MS. <https://doi.org/10.2118/119101-MS>.
- McLennan, J. and Deutsch, C. V. 2005. Ranking Geostatistical Realizations by Measures of Connectivity. Paper presented at the SPE International Thermal Operations and Heavy Oil Symposium, Calgary, Alberta, Canada, 1–3 November. SPE-98168-MS. <https://doi.org/10.2118/98168-MS>.
- Kennedy, J. and Eberhart, R. 1995. Particle Swarm Optimization. In *Proceedings of ICNN'95-International Conference on Neural Networks*, Vol. 4, 1942–1948. Piscataway, New Jersey, USA: IEEE. <https://doi.org/10.1109/ICNN.1995.488968>.
- Li, L., Jafarpour, B., and Mohammad-Khaninezhad, M. R. 2013. A Simultaneous Perturbation Stochastic Approximation Algorithm for Coupled Well Placement and Control Optimization under Geologic Uncertainty. *Comput Geosci* **17** (1): 167–188. <https://doi.org/10.1007/s10596-012-9323-1>.
- Mishra, S., Choudhary, M., and Datta-Gupta, A. 2002. A Novel Approach for Reservoir Forecasting Under Uncertainty. *SPE Res Eval & Eng* **5** (1): 42–48. SPE-75353-PA. <https://doi.org/10.2118/75353-PA>.
- Odai, L. and Ogbe, D. O. 2011. An Approach for Ranking Realizations to Characterize Reservoirs for Fluid Flow Simulation. Paper presented at the Nigeria Annual International Conference and Exhibition, Abuja, Nigeria, 30 July–3 August. SPE-150738-MS. <https://doi.org/10.2118/150738-MS>.
- Onwunali, J. E. and Durlafsky, L. J. 2010. Application of a Particle Swarm Optimization Algorithm for Determining Optimum Well Location and Type. *Comput Geosci* **14** (1): 183–198. <https://doi.org/10.1007/s10596-009-9142-1>.
- Peters, L., Arts, R., Brouwer, G. et al. 2010. Results of the Brugge Benchmark Study for Flooding Optimization and History Matching. *SPE Res Eval & Eng* **13** (3): 391–405. SPE-119094-PA. <https://doi.org/10.2118/119094-PA>.
- Rahim, S. and Li, Z. 2015. Reservoir Geological Uncertainty Reduction: An Optimization-Based Method Using Multiple Static Measures. *Math Geosci* **47** (4): 373–396. <https://doi.org/10.1007/s11004-014-9575-5>.
- Salehian, M., Sefat, M. H., and Muradov, K. 2021. Robust Integrated Optimization of Well Placement and Control under Field Production Constraints. *J Pet Sci Eng* **205**: 108926. <https://doi.org/10.1016/j.petrol.2021.108926>.
- Sayarpour, M., Zuluaga, E., Kabir, C. S. et al. 2009. The Use of Capacitance–Resistance Models for Rapid Estimation of Waterflood Performance and Optimization. *J Pet Sci Eng* **69** (3–4): 227–238. <https://doi.org/10.1016/j.petrol.2009.09.006>.
- Sayyafzadeh, M. 2017. Reducing the Computation Time of Well Placement Optimisation Problems Using Self-Adaptive Metamodelling. *J Pet Sci Eng* **151**: 143–158. <https://doi.org/10.1016/j.petrol.2016.12.015>.

- Sayyafzadeh, M. and Alrashdi, Z. 2019. Well Controls and Placement Optimisation Using Response-Fed and Judgement-Aided Parameterisation: Olympus Optimisation Challenge. *Comput Geosci* **24** (6): 2001–2025. <https://doi.org/10.1007/s10596-019-09891-y>.
- Scheidt, C. and Caers, J. 2009. Uncertainty Quantification in Reservoir Performance Using Distances and Kernel Methods—Application to a West Africa Deepwater Turbidite Reservoir. *SPE J.* **14** (4): 680–692. SPE-118740-PA. <https://doi.org/10.2118/118740-PA>.
- Shirangi, M. G. and Durlofsky, L. J. 2016. A General Method to Select Representative Models for Decision Making and Optimization under Uncertainty. *Comput & Geosci* **96**: 109–123. <https://doi.org/10.1016/j.cageo.2016.08.002>.
- Srinivas, M. and Patnaik, L. M. 1994. Genetic Algorithms: A Survey. *Computer* **27** (6): 17–26. <https://doi.org/10.1109/2.294849>.
- Tilke, P. G., Banerjee, R., Halabe, V. B. et al. 2010. Optimizing Well Placement Planning in the Presence of Subsurface Uncertainty and Operational Risk Tolerance. Paper presented at the 12th European Conference on the Mathematics of Oil Recovery, cp-16300078, European Association of Geoscientists & Engineers, Houten, The Netherlands, <https://doi.org/10.3997/2214-4609.20144997>.
- Wang, H., Echeverría-Ciaurri, D., Durlofsky, L. et al. 2012. Optimal Well Placement Under Uncertainty Using a Retrospective Optimization Framework. *SPE J.* **17** (1): 112–121. SPE-141950-PA. <https://doi.org/10.2118/141950-PA>.
- Welsh, M. and Begg, S. 2016. What Have We Learned? Insights from a Decade of Bias Research. *APPEA J* **56** (1): 435. <https://doi.org/10.1071/AJ15032>.
- Yeten, B., Durlofsky, L. J., and Aziz, K. 2003. Optimization of Nonconventional Well Type, Location, and Trajectory. *SPE J.* **8** (3): 200–210. SPE-86880-PA. <https://doi.org/10.2118/86880-PA>.

7 Summary, Conclusions and Future Work

This thesis tackled different aspects of field development planning optimization through the development and implementation of computationally efficient techniques. The thesis explored well control, well placement and joint optimization as well as subset selection for optimization under geological uncertainty. For each of these different problems, a technique is proposed and studied with the aim of reducing the computational cost of optimization without compromising the quality of the solution.

The thesis began by developing a novel gradient-based algorithm, Adam-SPSA, based on the popular optimization framework, Adam, which has seen significant success in machine learning applications. The newly developed algorithm was first implemented for the high-dimensional well control optimization problem with comparisons to its popular counterpart, the original formulation for SPSA using the steepest descent framework. The comparison was based on two experimental problems: a two-dimensional heterogeneous model and a three-dimensional benchmark model (Brugge). The problems were high-dimensional, consisting of 800 and 1,800 decision variables for the two and three-dimensional models, respectively. This study also included a detailed examination into the effect of two common bound constraint handling techniques, logarithmic transformation and projection, on the algorithms convergence rate.

The success of Adam-SPSA led to the implementation of the algorithm for well placement optimization for computationally constrained and practical scenarios. Here, the assumptions were made that a suitable reservoir engineering judgment-based initial guess is available, and a stringent computational budget is in effect. The aim was to use the algorithm's fast convergence rate to find an improved solution without the excessive computational burden of running a large number of reservoir simulations. The problems also took into account physical field constraints including a reservoir boundary approximated by piecewise linear functions and inter-well distance constraints. Additionally, the study also included an in-depth investigation on the effect of constraint handling techniques, as well as three-dimensional well parameterization, on the gradient approximation.

The thesis then investigated another approach for improving the computational efficiency of well placement optimization by investigating the use of a surrogate-treatment technique, known as manifold mapping. The technique is a surrogate-based optimization, in which a surrogate is iteratively adjusted to improve its accuracy. The objective function using the surrogate is then optimized instead of using high fidelity reservoir simulations. This was the first application of manifold mapping to well placement optimization. A number of different surrogates were combined with manifold mapping to understand the effect of each method's inherent accuracy on the ability of manifold mapping to improve the result of the optimization. Analytical surrogates, such as kriging and quadratic approximation, were used, whilst a local grid coarsening surrogate was used as a physics-based (or reduced-order model) surrogate. The first experimental problem implemented manifold mapping with kriging and quadratic approximation for the placement of four production wells in the presence of two pre-existing production wells in the PUNQ-S3 benchmark model. The second problem optimized the placement of five production wells for a model undergoing water-flooding with five injection wells.

As an intuitive evolution, joint optimization of well control and well placement was investigated next. To reduce the significant computational costs of joint optimization, the thesis developed and implemented a bilevel approach taking advantage of capacitance resistance models (CRMs). The bilevel approach is comprised of an outer loop for well placement optimization using particle swarm optimization and an inner loop of well control optimization using Adam-SPSA and aided by CRMs. The advantages of using CRM as a physics-based surrogate for well control optimization lie in the absence of a requirement for a geological model and the need for only one high fidelity full-physics reservoir simulation to tune the parameters. The proposed framework was compared to the typical approach of solving joint optimization problems using only full physics simulations for both the inner and outer loops. The experimental problems included a simple two-dimensional model with two production wells and two injection wells and a three-dimensional problem with four production wells and eight injection wells.

The next chapter of the thesis tackled the subset selection problem to improve the computational efficiency of optimization under uncertainty. To do this, a reformulation of the subset selection problem

was presented which argued that any proposed technique should ultimately aim at ensuring consistent ranking of alternatives when evaluated using the subset and the full set of realizations. In addition, a framework was proposed which was based on minimizing the rank correlation coefficient between a selected subset and the full set. The proposed technique is adaptive, which means the number of realizations is not constant throughout the optimization process, rather it varies depending on feedback from the search. Furthermore, the technique does not rely on selecting a static measure or the output from a base case scenario. The ability of the proposed technique to ensure consistent ranking of the alternatives was compared to the static ranking and clustering techniques. The proposed technique was applied to the well placement optimization of four producers and eight injectors and compared to the full set approach.

The investigations undertaken into the development and implementation of computationally efficient techniques for field development optimization allow the following conclusions to be drawn:

1. The developed gradient-based algorithm, Adam-SPSA, presents a computationally efficient option for well control optimization. Its noticeably faster convergence, of up to 91% when compared to the steepest descent framework, results from its use of approximated first and second moments to calculate dimension-wise search directions. This makes the algorithm applicable for high-dimensional problems such as well control optimization.
2. The choice of constraint handling technique for bound constrained optimization problems has an effect on the convergence rate for gradient-based algorithms. The commonly-used logarithmic transformation poses challenges for problems whose optimal solution lies on either of the bounds. As the search approaches either bound, a larger perturbation size is required to take a reasonable step in the gradient direction. In addition, gradient-based algorithms are more likely to get trapped at the boundaries since a large step-size would be required to move away from the boundaries. On the other hand, for the problems tested, the projection method didn't face the same risks of entrapment at the boundary.
3. In practical scenarios under strict computational budgets, Adam-SPSA provides a suitable alternative for well placement optimization. When compared to a local derivative-free

optimization algorithm, generalized pattern search (GPS), and steepest descent SPSA, Adam-SPSA resulted in improvements in convergence speeds of up to 57%. This advantage was especially evident for the high-dimensional problems, where the derivative-free algorithm struggled to be computationally efficient.

4. When using gradient-based methods for well placement optimization careful consideration needs to be given to the type of parameterization used for three-dimensional wells. For the problems tested, gradient values calculated when using spherical coordinates exhibited additional sensitivity. This is due to the dependence between the toe of the well to the heel of the well in spherical parameterization. In addition, when applying constraint handling techniques with a gradient-based algorithm requires special attention to be given to the effect of constraint handling on the gradient approximation. Simple constraint handling, such as bound constraints, may not provide the required correction to the gradient perturbation to result in a useful approximation. On the other hand, incorporation of too many constraint handling techniques may over-correct the perturbations, leading to a gradient approximation that may not be representative of the local landscape.
5. Manifold mapping provides a computationally efficient technique for surrogate-based well placement optimization. Analytical surrogates combined with manifold mapping are suitable for models undergoing primary depletion, with computational savings of up to 55% without loss of solution quality compared to a local derivative-free algorithm. For more complex recovery methods, such as secondary recovery in highly heterogeneous reservoir, physics-based methods combined with manifold mapping are more suitable, providing up to 80% reduction in the computational cost compared to a local derivative-free algorithm.
6. The developed bilevel approach combining PSO for well placement optimization in the outer loop and Adam-SPSA with CRM for well control optimization in the inner loop was found to be very computationally efficient. The proposed technique resulted in computational savings of up to 99% in the number of required reservoir simulations and improvements in the objective function values of up to 22% when compared to the conventional approach of using only full-physics simulations. The fidelity of producer-based CRM (CRMP) when used in joint

optimization showed variations from 1.2% to 15%, in some cases. However, this fidelity is sufficient to provide an advantage with the significant reduction in required full physics simulations without the loss of solution quality.

7. The reformulation of the subset selection problem to one which focuses on consistent ranking of the alternatives is an important consideration for optimization under uncertainty. The proposed adaptive rank-based approach ensures that even if a subset of realizations is used the alternatives (i.e., field development plans) are ranked in a similar fashion to the full set. The results showed that the frequency of subset selection affected the ability of the algorithm to converge. Reselection of a new subset too often would introduce unnecessary noise into the objective function, but on the other hand, infrequent reselection may result in the premature convergence to a subset local optima.
8. The implementation of the adaptive rank-based subset selection technique for well placement optimization resulted in improvements of up to 24% in objective function value (NPV), on average, compared the optimization using the full set. More importantly, using the proposed subset selection technique resulted in a computational reduction of close to a factor of 9 when compared to the optimization using the full set. This showed that a significant reduction of reservoir simulations is possible without the loss in the quality of solution.

This research answered important questions by developing and presenting computationally efficient techniques for field development optimization problems. However, this research can be extended in a number of ways.

1. An extension of the Adam framework presented in this work by utilizing other gradient approximations such as EnOpt or StoSAG for well control optimization and comparing the efficiency of each implementation. Furthermore, a comprehensive comparison of the proposed Adam-SPSA with other optimization frameworks, including conjugate gradients, or even higher-order methods which use approximations of the Hessian, would be an opportunity to further provide evidence of the algorithm's efficiency.

2. An in-depth investigation into the effect of constraint handling techniques on gradient approximations is missing in the literature. This could provide important insight for well placement and well control problems, where physical constraints are considered.
3. The proposed adaptive rank-based technique for subset realization selection can be extended to joint optimization problems, where the consideration of different types of decision variables may make other subset selection techniques less efficient.
4. An interesting research item would be the extension of the surrogate-assisted bi-level optimization approach proposed in this thesis to include proxy models for the well placement optimization problem, too. The interplay and interdependency of the surrogates for the well control and well placement levels would be fascinating. This could be implemented by using a surrogate for the well placement problem for a local region of the objective function space as the particles converge.
5. In addition, it would be also a fascinating piece of research to compare the proposed bi-level approach using capacitance-resistance models (CRMs) with other proxy models of varying degrees of complexity. The alternative methods that could be used for comparison include quadratic approximation, kriging, artificial neural networks, fuzzy logic, amongst others.
6. The example problems studied in this thesis involved conventional reservoir models undergoing water-flooding. Applications of the proposed optimization techniques to unconventional reservoirs, or reservoirs undergoing enhanced oil recovery processes such as chemical injection, gas injection or steam-assisted gravity drainage. In addition, the optimization techniques presented in this thesis could also be applied to carbon sequestration problems.

Bibliography

- AL DOSSARY, M. A. & NASRABADI, H. 2016. Well placement optimization using imperialist competitive algorithm. *Journal of Petroleum Science and Engineering*, 147, 237-248.
- ALGHAREEB, Z. M. 2015. *Optimal reservoir management using adaptive reduced-order models*. Massachusetts Institute of Technology.
- ALGHAREEB, Z. M., WALTON, S. P. & WILLIAMS, J. R. Well Placement Optimization Under Constraints Using Modified Cuckoo Search. SPE Saudi Arabia Section Technical Symposium and Exhibition, 2014. SPE-172841-MS.
- ALIYEV, E. & DURLOFSKY, L. J. 2017. Multilevel Field Development Optimization Under Uncertainty Using a Sequence of Upscaled Models. *Mathematical Geosciences*, 49, 307-339.
- ALPAK, F. O., BARTON, M. D. & CAERS, J. 2010. A flow-based pattern recognition algorithm for rapid quantification of geologic uncertainty. *Computational Geosciences*, 14, 603-621.
- ALRASHDI, Z. & SAYYAFZADEH, M. 2019. $(\mu+\lambda)$ Evolution strategy algorithm in well placement, trajectory, control and joint optimisation. *Journal of Petroleum Science and Engineering*, 177, 1042-1058.
- ARTUS, V., DURLOFSKY, L. J., ONWUNALU, J. & AZIZ, K. 2006. Optimization of nonconventional wells under uncertainty using statistical proxies. *Computational Geosciences*, 10, 389-404.
- ASADOLLAHI, M. & NAEVDAL, G. Waterflooding optimization using gradient based methods. SPE/EAGE Reservoir Characterization & Simulation Conference, 2009.
- ASADOLLAHI, M., NÆVDAL, G., DADASHPOUR, M. & KLEPPE, J. 2014. Production optimization using derivative free methods applied to Brugge field case. *Journal of Petroleum Science and Engineering*, 114, 22-37.
- ASHEIM, H. 1988. Maximization of Water Sweep Efficiency by Controlling Production and Injection Rates. *European Petroleum Conference*. London, United Kingdom: Society of Petroleum Engineers.
- AWOTUNDE, A. A. 2014. On The Joint Optimization of Well Placement and Control. *SPE Saudi Arabia Section Technical Symposium and Exhibition*. Al-Khobar, Saudi Arabia: Society of Petroleum Engineers.
- BADRU, O. & KABIR, C. S. Well Placement Optimization in Field Development. SPE Annual Technical Conference and Exhibition, 2003. SPE-84191-MS.
- BALLIN, P., JOURNAL, A. & AZIZ, K. 1992. Prediction of uncertainty in reservoir performance forecast. *Journal of Canadian Petroleum Technology*, 31.
- BANGERTH, W., KLIE, H., WHEELER, M. F., STOFFA, P. L. & SEN, M. K. 2006. On optimization algorithms for the reservoir oil well placement problem. *Computational Geosciences*, 10, 303-319.
- BELLMAN, R. 1966. Dynamic programming. *Science*, 153, 34-37.
- 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, 1061-1079.
- BITTENCOURT, A. C. & HORNE, R. N. 1997. Reservoir Development and Design Optimization. *SPE Annual Technical Conference and Exhibition*. San Antonio, Texas: Society of Petroleum Engineers.
- BOUZARKOUNA, Z., DING, D. Y. & AUGER, A. 2012. Well placement optimization with the covariance matrix adaptation evolution strategy and meta-models. *Computational Geosciences*, 16, 75-92.
- BROUWER, D. R. & JANSEN, J. D. 2004. Dynamic Optimization of Waterflooding With Smart Wells Using Optimal Control Theory.
- BROUWER, D. R., NÆVDAL, G., JANSEN, J. D., VEFRING, E. H. & VAN KRUIJSDIJK, C. P. J. W. 2004. Improved Reservoir Management Through Optimal Control and Continuous Model Updating. *SPE Annual Technical Conference and Exhibition*. Houston, Texas: Society of Petroleum Engineers.

- CARDOSO, M. A., DURLOFSKY, L. J. & SARMA, P. 2009. Development and application of reduced-order modeling procedures for subsurface flow simulation. *International Journal for Numerical Methods in Engineering*, 77, 1322-1350.
- CARVAJAL, G., MAUCEC, M. & CULLICK, S. 2018. Chapter Seven - Smart Wells and Techniques for Reservoir Monitoring. In: CARVAJAL, G., MAUCEC, M. & CULLICK, S. (eds.) *Intelligent Digital Oil and Gas Fields*. Boston: Gulf Professional Publishing.
- CHAUDHRI, M. M., PHALE, H. A., LIU, N. & OLIVER, D. S. 2009. An Improved Approach for Ensemble-Based Production Optimization. *SPE Western Regional Meeting*. San Jose, California: Society of Petroleum Engineers.
- CHEN, B. & REYNOLDS, A. C. 2017. Optimal control of ICV's and well operating conditions for the water-alternating-gas injection process. *Journal of Petroleum Science and Engineering*, 149, 623-640.
- CHEN, C., LI, G. & REYNOLDS, A. C. C. 2012. Robust Constrained Optimization of Short- and Long-Term Net Present Value for Closed-Loop Reservoir Management. *SPE Journal*, 17, 849-864.
- CHEN, H., FENG, Q., ZHANG, X., WANG, S., MA, Z., ZHOU, W. & LIU, C. 2018. A meta-optimized hybrid global and local algorithm for well placement optimization. *Computers & Chemical Engineering*, 117, 209-220.
- CHEN, Y., OLIVER, D. S. & ZHANG, D. 2009. Efficient ensemble-based closed-loop production optimization. *SPE Journal*, 14, 634-645.
- CIAURRI, D. E., ISEBOR, O. J. & DURLOFSKY, L. J. 2010. Application of derivative-free methodologies to generally constrained oil production optimization problems. *Procedia Computer Science*, 1, 1301-1310.
- DE BRITO, D. U. & DURLOFSKY, L. J. 2021. Field development optimization using a sequence of surrogate treatments. *Computational Geosciences*, 25, 35-65.
- DEHDARI, V. & OLIVER, D. S. 2012. Sequential Quadratic Programming for Solving Constrained Production Optimization--Case Study From Brugge Field. *SPE journal*, 17, 874-884.
- DEHDARI, V., OLIVER, D. S. & DEUTSCH, C. V. 2012. Comparison of optimization algorithms for reservoir management with constraints—a case study. *Journal of Petroleum Science and Engineering*, 100, 41-49.
- DEUTSCH, C. V. & BEGG, S. H. 2001. The use of ranking to reduce the required number of realizations. *Annual Report 3*. Edmonton: Centre for Computational Geostatistics (CCG).
- DING, S., JIANG, H., LI, J. & TANG, G. 2014. Optimization of well placement by combination of a modified particle swarm optimization algorithm and quality map method. *Computational Geosciences*, 18, 747-762.
- DING, Y. 2008. Optimization of Well Placement Using Evolutionary Methods. *Europec/EAGE Conference and Exhibition*. Rome, Italy: Society of Petroleum Engineers.
- DO, S. T. & REYNOLDS, A. C. 2013. Theoretical connections between optimization algorithms based on an approximate gradient. *Computational Geosciences*, 17, 959-973.
- ECHEVERRIA CIAURRI, D. & WAGENAAR, C. 2016. Manifold-mapping Optimization Applied to Oil Field Operations.
- EMERICK, A. A., SILVA, E., MESSER, B., ALMEIDA, L. F., SZWARCMAN, D., PACHECO, M. A. C. & VELLASCO, M. M. B. R. 2009. Well Placement Optimization Using a Genetic Algorithm With Nonlinear Constraints. *SPE Reservoir Simulation Symposium*. The Woodlands, Texas: Society of Petroleum Engineers.
- FATHI, Z. & RAMIREZ, F. W. 1984. Optimal Injection Policies for Enhanced Oil Recovery: Part 2-Surfactant Flooding. *Society of Petroleum Engineers Journal*, 24, 333-341.
- FONSECA, R., LEEUWENBURGH, O., DELLA ROSSA, E., VAN DEN HOF, P. M. & JANSEN, J.-D. 2015. Ensemble-based multiobjective optimization of on/off control devices under geological uncertainty. *SPE Reservoir Evaluation & Engineering*, 18, 554-563.
- FONSECA, R., LEEUWENBURGH, O., VAN DEN HOF, P. & JANSEN, J. 2014a. Ensemble-based hierarchical multi-objective production optimization of smart wells. *Computational Geosciences*, 18, 449-461.
- FONSECA, R., STORDAL, A., LEEUWENBURGH, O., VAN DEN HOF, P. & JANSEN, J. D. Robust ensemble-based multi-objective optimization. 14th European conference on the Mathematics of Oil Recovery (ECMORE XIV), 8-11 September 2014b Catania, Sicily

- FONSECA, R. R. M., CHEN, B., JANSEN, J. D. & REYNOLDS, A. 2017. A stochastic simplex approximate gradient (StoSAG) for optimization under uncertainty. *International Journal for Numerical Methods in Engineering*, 109, 1756-1776.
- FOROUD, T., BARADARAN, A. & SEIFI, A. 2018. A comparative evaluation of global search algorithms in black box optimization of oil production: A case study on Brugge field. *Journal of Petroleum Science and Engineering*, 167, 131-151.
- FOROUZANFAR, F., POQUIOMA, W. E. & REYNOLDS, A. C. 2016. Simultaneous and Sequential Estimation of Optimal Placement and Controls of Wells With a Covariance Matrix Adaptation Algorithm.
- FOROUZANFAR, F. & REYNOLDS, A. C. 2013. Well-placement optimization using a derivative-free method. *Journal of Petroleum Science and Engineering*, 109, 96-116.
- FOROUZANFAR, F., ROSSA, E. D., RUSSO, R. & REYNOLDS, A. C. 2013. Life-cycle production optimization of an oil field with an adjoint-based gradient approach. *Journal of Petroleum Science and Engineering*, 112, 351-358.
- GOLZARI, A., HAGHIGHAT SEFAT, M. & JAMSHIDI, S. 2015. Development of an adaptive surrogate model for production optimization. *Journal of Petroleum Science and Engineering*, 133, 677-688.
- GÜYAGÜLER, B. & HORNE, R. N. 2004. Uncertainty Assessment of Well-Placement Optimization. *SPE Reservoir Evaluation & Engineering*, 7, 24-32.
- GÜYAGÜLER, B., HORNE, R. N., ROGERS, L. & ROSENZWEIG, J. J. 2002. Optimization of well placement in a Gulf of Mexico waterflooding project. *SPE Reservoir Evaluation & Engineering*, 5, 229-236.
- HAGHIGHAT SEFAT, M., ELSHEIKH, A. H., MURADOV, K. M. & DAVIES, D. R. 2016. Reservoir uncertainty tolerant, proactive control of intelligent wells. *Computational Geosciences*, 20, 655-676.
- HANSEN, N. 2006. The CMA evolution strategy: a comparing review. *Towards a new evolutionary computation*, 75-102.
- HONG, A. J., BRATVOLD, R. B. & NÆVDAL, G. 2017. Robust production optimization with capacitance-resistance model as precursor. *Computational Geosciences*, 21, 1423-1442.
- HOOKE, R. & JEEVES, T. A. 1961. Direct Search Solution of Numerical and Statistical Problems. *Journal of the ACM (JACM)*, 8, 212-229.
- HUMPHRIES, T. D. & HAYNES, R. D. 2015. Joint optimization of well placement and control for nonconventional well types. *Journal of Petroleum Science and Engineering*, 126, 242-253.
- HUMPHRIES, T. D., HAYNES, R. D. & JAMES, L. A. 2014. Simultaneous and sequential approaches to joint optimization of well placement and control. *Computational Geosciences*, 18, 433-448.
- ISEBOR, O. J. 2009. *Constrained production optimization with an emphasis on derivative-free methods*. Stanford University Stanford, CA.
- ISEBOR, O. J., DURLOFSKY, L. J. & ECHEVERRÍA CIAURRI, D. 2014. A derivative-free methodology with local and global search for the constrained joint optimization of well locations and controls. *Computational Geosciences*, 18, 463-482.
- JAFROODI, N. & ZHANG, D. 2011. New method for reservoir characterization and optimization using CRM–EnOpt approach. *Journal of Petroleum Science and Engineering*, 77, 155-171.
- JANSEN, J. D. 2011. Adjoint-based optimization of multi-phase flow through porous media – A review. *Computers & Fluids*, 46, 40-51.
- JANSEN, J. D. & DURLOFSKY, L. J. 2017. Use of reduced-order models in well control optimization. *Optimization and Engineering*, 18, 105-132.
- JESMANI, M., BELLOUT, M. C., HANEA, R. & FOSS, B. 2016a. Well placement optimization subject to realistic field development constraints. *Computational Geosciences*, 20, 1185-1209.
- JESMANI, M., JAFARPOUR, B., BELLOUT, M., HANEA, R. & FOSS, B. Application of simultaneous perturbation stochastic approximation to well placement optimization under uncertainty. ECMOR XV-15th European Conference on the Mathematics of Oil Recovery, 2016b. European Association of Geoscientists & Engineers, cp-494-00133.
- JESMANI, M., JAFARPOUR, B., BELLOUT, M. C. & FOSS, B. 2020. A reduced random sampling strategy for fast robust well placement optimization. *Journal of Petroleum Science and Engineering*, 184, 106414.

- KENNEDY, J. & EBERHART, R. Particle swarm optimization. Proceedings of ICNN'95-International Conference on Neural Networks, 1995. IEEE, 1942-1948.
- KWON, S., PARK, G., JANG, Y., CHO, J., CHU, M.-G. & MIN, B. 2021. Determination of oil well placement using convolutional neural network coupled with robust optimization under geological uncertainty. *Journal of Petroleum Science and Engineering*, 201, 108118.
- LEEUEWENBURGH, O., EGBERTS, P. J. P. & ABBINK, O. A. 2010. Ensemble Methods for Reservoir Life-Cycle Optimization and Well Placement. *SPE/DGS Saudi Arabia Section Technical Symposium and Exhibition*. Al-Khobar, Saudi Arabia: Society of Petroleum Engineers.
- LI, L. & JAFARPOUR, B. 2012. A variable-control well placement optimization for improved reservoir development. *Computational Geosciences*, 16, 871-889.
- LI, L., JAFARPOUR, B. & MOHAMMAD-KHANINEZHAD, M. R. 2013. A simultaneous perturbation stochastic approximation algorithm for coupled well placement and control optimization under geologic uncertainty. *Computational Geosciences*, 17, 167-188.
- LI, S., DEUTSCH, C. V. & SI, J. 2012. Ranking Geostatistical Reservoir Models with Modified Connected Hydrocarbon Volume. *Ninth International Geostatistics Congress*. Oslo, Norway.
- LIU, W., RAMIREZ, W. F. & QI, Y. F. 1993. Optimal Control of Steamflooding. *SPE Advanced Technology Series*, 1, 73-82.
- LIU, Z. & REYNOLDS, A. C. 2020. A Sequential-Quadratic-Programming-Filter Algorithm with a Modified Stochastic Gradient for Robust Life-Cycle Optimization Problems with Nonlinear State Constraints. *SPE Journal*, 25, 1938-1963.
- LORENTZEN, R. J., SHAFIEIRAD, A. & NAEVDAL, G. 2009. Closed Loop Reservoir Management Using the Ensemble Kalman Filter and Sequential Quadratic Programming. *SPE Reservoir Simulation Symposium*. The Woodlands, Texas: Society of Petroleum Engineers.
- LU, R., FOROUZANFAR, F. & REYNOLDS, A. C. 2017a. Bi-Objective Optimization of Well Placement and Controls Using StoSAG. *SPE Reservoir Simulation Conference*. Montgomery, Texas, USA: Society of Petroleum Engineers.
- LU, R., FOROUZANFAR, F. & REYNOLDS, A. C. 2017b. An efficient adaptive algorithm for robust control optimization using StoSAG. *Journal of Petroleum Science and Engineering*, 159, 314-330.
- LU, R. & REYNOLDS, A. C. 2020. Joint Optimization of Well Locations, Types, Drilling Order, and Controls Given a Set of Potential Drilling Paths. *SPE Journal*, 25, 1285-1306.
- MCLENNAN, J. & DEUTSCH, C. V. 2005. Ranking Geostatistical Realizations by Measures of Connectivity. *SPE International Thermal Operations and Heavy Oil Symposium*. Calgary, Alberta, Canada: Society of Petroleum Engineers.
- MEHOS, G. J. & RAMIREZ, W. F. 1989. Use of optimal control theory to optimize carbon dioxide miscible-flooding enhanced oil recovery. *Journal of Petroleum Science and Engineering*, 2, 247-260.
- MISHRA, S., CHOUDHARY, M. K. & DATTA-GUPTA, A. 2002. A Novel Approach for Reservoir Forecasting Under Uncertainty. *SPE Reservoir Evaluation & Engineering*, 5, 42-48.
- MONTES, G., BARTOLOME, P. & UDIAS, A. L. 2001. The Use of Genetic Algorithms in Well Placement Optimization. *SPE Latin American and Caribbean Petroleum Engineering Conference*. Buenos Aires, Argentina: Society of Petroleum Engineers.
- MØYNER, O., KROGSTAD, S. & LIE, K.-A. 2014. The Application of Flow Diagnostics for Reservoir Management. *SPE Journal*, 20, 306-323.
- NADERI, M. & KHAMEHCHI, E. 2017. Well placement optimization using metaheuristic bat algorithm. *Journal of Petroleum Science and Engineering*, 150, 348-354.
- NELDER, J. A. & MEAD, R. 1965. A Simplex Method for Function Minimization. *The Computer Journal*, 7, 308-313.
- NOCEDAL, J. & WRIGHT, S. 2006. *Numerical optimization*.
- NORRENA, K. P. & DEUTSCH, C. V. Automatic Determination of Well Placement Subject to Geostatistical and Economic Constraints. *SPE International Thermal Operations and Heavy Oil Symposium and International Horizontal Well Technology Conference*, 2002. SPE-78996-MS.
- NWANKWOR, E., NAGAR, A. K. & REID, D. C. 2013. Hybrid differential evolution and particle swarm optimization for optimal well placement. *Computational Geosciences*, 17, 249-268.

- ODAI, L. & OGBE, D. O. An Approach for Ranking Realizations to Characterize Reservoirs for Fluid Flow Simulation. Nigeria Annual International Conference and Exhibition, 2011. OnePetro.
- ONWUNALU, J. E. & DURLOFSKY, L. J. 2010. Application of a particle swarm optimization algorithm for determining optimum well location and type. *Computational Geosciences*, 14, 183-198.
- OZDOGAN, U. & HORNE, R. N. 2006. Optimization of Well Placement Under Time-Dependent Uncertainty.
- RAHIM, S. & LI, Z. 2015. Reservoir geological uncertainty reduction: an optimization-based method using multiple static measures. *Mathematical Geosciences*, 47, 373-396.
- RAMIREZ, W. F., FATHI, Z. & CAGNOL, J. L. 1984. Optimal Injection Policies for Enhanced Oil Recovery: Part 1 Theory and Computational Strategies. *Society of Petroleum Engineers Journal*, 24, 328-332.
- SALEHIAN, M., SEFAT, M. H. & MURADOV, K. 2021. Robust integrated optimization of well placement and control under field production constraints. *Journal of Petroleum Science and Engineering*, 205, 108926.
- SARMA, P., AZIZ, K. & DURLOFSKY, L. J. 2005. Implementation of Adjoint Solution for Optimal Control of Smart Wells. *SPE Reservoir Simulation Symposium*. The Woodlands, Texas: Society of Petroleum Engineers.
- SARMA, P. & CHEN, W. H. 2008. Efficient Well Placement Optimization with Gradient-based Algorithms and Adjoint Models. *Intelligent Energy Conference and Exhibition*. Amsterdam, The Netherlands: Society of Petroleum Engineers.
- SAYYAFZADEH, M. 2017. Reducing the computation time of well placement optimisation problems using self-adaptive metamodelling. *Journal of Petroleum Science and Engineering*, 151, 143-158.
- SAYYAFZADEH, M. & ALRASHDI, Z. 2019. Well controls and placement optimisation using response-fed and judgement-aided parameterisation: Olympus optimisation challenge. *Computational Geosciences*.
- SCHEIDT, C. & CAERS, J. 2008. Representing Spatial Uncertainty Using Distances and Kernels. *Mathematical Geosciences*, 41, 397.
- SCHEIDT, C. & CAERS, J. 2009. Uncertainty Quantification in Reservoir Performance Using Distances and Kernel Methods--Application to a West Africa Deepwater Turbidite Reservoir.
- SHIRANGI, M. G. & DURLOFSKY, L. J. 2016. A general method to select representative models for decision making and optimization under uncertainty. *Computers & Geosciences*, 96, 109-123.
- SILVA, V. L. S., CARDOSO, M. A., OLIVEIRA, D. F. B. & DE MORAES, R. J. 2020. Stochastic optimization strategies applied to the OLYMPUS benchmark. *Computational Geosciences*, 24, 1943-1958.
- SRINIVAS, M. & PATNAIK, L. M. 1994. Genetic algorithms: a survey. *Computer*, 27, 17-26.
- SUDARYANTO, B. & YORTSOS, Y. C. 2001. Optimization of Displacements in Porous Media Using Rate Control. *SPE Annual Technical Conference and Exhibition*. New Orleans, Louisiana: Society of Petroleum Engineers.
- SUWARTADI, E., KROGSTAD, S. & FOSS, B. 2012. Nonlinear output constraints handling for production optimization of oil reservoirs. *Computational Geosciences*, 16, 499-517.
- SUWARTADI, E., KROGSTAD, S. & FOSS, B. 2015. Adjoint-based surrogate optimization of oil reservoir water flooding. *Optimization and Engineering*, 16, 441-481.
- TORCZON, V. 1997. On the convergence of pattern search algorithms. *SIAM Journal on optimization*, 7, 1-25.
- TUPAC, Y. J., ALMEIDA, L. F. & VELLASCO, M. M. B. R. 2007. Evolutionary Optimization of Oil Field Development. *Digital Energy Conference and Exhibition*. Houston, Texas, U.S.A.: Society of Petroleum Engineers.
- VAN ESSEN, G., ZANDVLIET, M., VAN DEN HOF, P., BOSGRA, O. & JANSEN, J.-D. 2009. Robust Waterflooding Optimization of Multiple Geological Scenarios. *SPE Journal*, 14, 202-210.
- VLEMMIX, S., JOOSTEN, G. J. P., BROUWER, R. & JANSEN, J.-D. 2009. Adjoint-Based Well Trajectory Optimization. *EUROPEC/EAGE Conference and Exhibition*. Amsterdam, The Netherlands: Society of Petroleum Engineers.

- VOLKOV, O. & BELLOUT, M. C. 2018. Gradient-based constrained well placement optimization. *Journal of Petroleum Science and Engineering*, 171, 1052-1066.
- WANG, C., LI, G. & REYNOLDS, A. C. 2007. Optimal Well Placement for Production Optimization. *Eastern Regional Meeting*. Lexington, Kentucky USA: Society of Petroleum Engineers.
- WANG, C., LI, G. & REYNOLDS, A. C. 2009. Production optimization in closed-loop reservoir management. *SPE journal*, 14, 506-523.
- WANG, H., ECHEVERRÍA-CIAURRI, D., DURLOFSKY, L. & COMINELLI, A. 2012. Optimal Well Placement Under Uncertainty Using a Retrospective Optimization Framework. *SPE Journal*, 17, 112-121.
- WANG, X., HAYNES, R. D. & FENG, Q. 2016. A multilevel coordinate search algorithm for well placement, control and joint optimization. *Computers & Chemical Engineering*, 95, 75-96.
- WANG, X., HAYNES, R. D., HE, Y. & FENG, Q. 2019. Well control optimization using derivative-free algorithms and a multiscale approach. *Computers & Chemical Engineering*, 123, 12-33.
- WEN, T., THIELE, M. R., CIAURRI, D. E., AZIZ, K. & YE, Y. 2014. Waterflood management using two-stage optimization with streamline simulation. *Computational Geosciences*, 18, 483-504.
- YETEN, B., BROUWER, D. R., DURLOFSKY, L. J. & AZIZ, K. 2004. Decision analysis under uncertainty for smart well deployment. *Journal of Petroleum Science and Engineering*, 44, 175-191.
- YETEN, B., DURLOFSKY, L. J. & AZIZ, K. 2003. Optimization of nonconventional well type, location, and trajectory. *SPE Journal*, 8, 200-210.
- ZANDVLIET, M., HANDELS, M., VAN ESSEN, G., BROUWER, R. & JANSEN, J.-D. 2008. Adjoint-Based Well-Placement Optimization Under Production Constraints.
- ZHANG, K., LI, G., REYNOLDS, A. C., YAO, J. & ZHANG, L. 2010. Optimal well placement using an adjoint gradient. *Journal of Petroleum Science and Engineering*, 73, 220-226.
- ZHAO, H., CHEN, C., DO, S., OLIVEIRA, D., LI, G. & REYNOLDS, A. 2013. Maximization of a Dynamic Quadratic Interpolation Model for Production Optimization. *SPE Journal*, 18, 1012-1025.
- ZUBAREV, D. I. Pros and cons of applying proxy-models as a substitute for full reservoir simulations. SPE Annual Technical Conference and Exhibition, 2009. Society of Petroleum Engineers.