Operations Research Perspectives 9 (2022) 100237

Contents lists available at ScienceDirect

Operations Research Perspectives

journal homepage: www.elsevier.com/locate/orp

Optimization of subsurface models with multiple criteria using Lexicase Selection

Yifan He [a], Claus Aranha [b,*], Antony Hallam [c], Romain Chassagne [c]

[a] School of Systems and Information Engineering, University of Tsukuba, Tsukuba, Japan
[b] Faculty of Engineering, Information and Systems, University of Tsukuba, Tsukuba, Japan
[c] Energy, Geoscience, Infrastructure and Society, Heriot-Watt University, United Kingdom

ARTICLE INFO

Keywords:
Seismic History Matching
Lexicase Selection
Multi-objective optimization
Multi-Objective Evolutionary Algorithm

ABSTRACT

Seismic History Matching (SHM) is a key problem in the geosciences community, requiring optimal parameters of a subsurface model that match the observed data from multiple in-situ measurements. Therefore, the SHM problems are usually solved with Multi-Objective Evolutionary Algorithms (MOEAs). This group of algorithms optimize multiple objectives simultaneously, considering the trade-off between objectives. However, SHM requires the solutions that are good on all objectives rather than a trade-off. In this study, we propose a Differential Evolution algorithm using Lexicase Selection to solve the SHM problems. Unlike the MOEAs, this selection method pushes the solutions to perform well on all objectives. We compared this method with two MOEAs, namely Non-dominated Sorting Genetic Algorithm II and Reference Vector-guided Evolutionary Algorithm, on two SHM problems. The results show that this method generates more solutions near the ground truth.

1. Introduction

Optimization problems for subsurface flow processes, are a key problem in the geosciences community, especially the Seismic History Matching (SHM) problem, which we will use here as a case study. It is about the matching of model parameters with data obtained from in-situ measurements. The objective of SHM is to find a model or small set of models which best match the in-situ measurements. This calibration process improves the prediction accuracy of the starting model/s and is a necessity for safe and economically sound development of subsurface energy systems (energy storage, geothermal, Oil and Gas). A small accurate set of final models is needed because the simulation of reservoir fluid flow is computationally expensive.

The SHM data assimilation problem is well-known as being a highly difficult inverse problem to solve [1–7]. The typical data to assimilate in a history matching exercise are wells and seismic data, which are usually represented as time-series and matrices, respectively. How to merge these two rather different attributes in a single objective function is non-trivial problem [8]. One “classical” approach is to calculate the mismatch of wells and seismic maps and then compose a weighted single objective function. These weights are generally chosen by engineering judgement, consequently it is then questionable to sum directly these two objectives together, as they represent different entities and physical measurements.

To circumvent these problems, several authors have employed Multi-Objective Evolutionary Algorithms (MOEAs) [9–19], and Many Objective Evolutionary Algorithms [20]. Typically, these algorithms work simultaneously on several conflicting objectives (i.e. the optimality on each objective cannot be achieved at the same time), looking for the best trade-off set of solutions, which is called the Pareto Front [21]. Therefore, selection methods that consider the trade-offs of multiple objectives at the same time are usually introduced in this group of algorithms. For example, the Reference Vector-guided Evolutionary Algorithm [22] uses a reference vector-guided selection based on the linear combinations of the objectives. Non-dominated Sorting Genetic Algorithm II [21] uses the fast non-dominated sorting that consider a partial order based on all objectives.

However, this trade-off approach is precisely what we propose to investigate in this paper. We postulate that a “trade-off of objectives” is not a good mental model to describe SHM problems, and therefore MOEAs are not well adapted for these problems. This is because in SHM, the different objectives are not strictly trade-offs, but rather different descriptions of a same physics. Eventual differences in objective values are explained by the sparseness and uncertainty of the available data, and not because of some inherent incompatibility of the two objectives (compare this with the more traditional financial

* Corresponding author.
E-mail addresses: he.yifan.xs@alumni.tsukuba.ac.jp (Y. He), caranha@cs.tsukuba.ac.jp (C. Aranha), arh5@hw.ac.uk (A. Hallam), r.l.chassagne@hw.ac.uk (R. Chassagne).

https://doi.org/10.1016/j.orp.2022.100237
Received 10 February 2022; Received in revised form 3 April 2022; Accepted 20 April 2022
Available online 26 April 2022

portfolio optimization problem, where return and risk of an investment are objectives that are normally at odds with each other). Ideally, we are interested in solutions that show high performance in all objective measures, and a solution that is very good in one objective and very bad in others is likely unphysical.

Other problem formulations in the Evolutionary Computation literature are close to this "non trade-off" model, such as multi-task optimization and multi-form optimization. The readers may refer to Gupta's review work [23] for the general background. In particular, we highlight program synthesis, where the goal is to optimize a computer program that can solve a generalized logical task. Each instance of the task is considered a separate objective, and an optimal program must solve as many of these instances as possible [24].

We consider the SHM problem to be more similar to this formulation, in the sense of objective aggregation, so we propose a method for aggregating multiple objectives for the subsurface history matching problem based on *Lexicase Selection* [24]. Lexicase Selection is a method originally proposed for program synthesis tasks. The primary concept, is to filter the solutions based on a shuffled arrangement of all objectives, so that it can drive the solutions to a better quality in all objectives at the same time, while providing enough diversity during the search process. We provide a detailed background of this method in Section 2.3. We introduce a new Differential Evolution algorithm using Lexicase Selection for the SHM problem, which is described in Section 3.

We tested our proposed algorithm on two SHM problems, referred in the following as: (1) TS2N and (2) Volve. TS2N [25] is a simple model containing a single injection and production well pair. This model contains very low trade-offs between the objectives, but it will serve as a calibration for the new implemented method. The Volve model [26], on the other hand, is a real-world case, and significantly more difficult to optimize, with multiple well and seismic objectives.

We compared the proposed algorithm with two well-known MOEAs, NSGA-II [21] and RVEA [22]. We discuss the results from various perspectives, including the distance to the ground truth, the difference on set coverage, the distribution of the non-dominated solutions, as well as the prediction performance. Our experiment shows that this method has the following characteristics:

1. A better optimization performance can be achieved for the SHM problem (Tables 4, 6, 5, and 7 in Section 4.3).
2. The final solution set is concentrated in the center of the Pareto Front, with fewer extreme solutions which would possibly be non-physical (Fig. 3 in Section 4.3).
3. Weighting or merging of independent objectives is not required, greatly simplifying the data assimilation processes (Algorithm 2 in Section 3).

2. Background

In this section, we first provide background of MOEAs and their application in the seismic history matching literature. After that, we point out the limitations of the MOEAs and introduce a different approach, Lexicase Selection, to handle with the multiple objectives in the SHM problems.

Subsurface flow data assimilation problems [27,28] and more precisely, seismic history matching [1–7], developed in this paper, are very challenging to solve and very active area of research in the Geosciences community, which consists on merging predictions with observations. We aim in matching a multi-stream of data: map-based and "point scale" based (respectively, time-lapse seismic and multiple well data), to the reciprocal maps issued by the simulation model. The ultimate goal being to have at disposal a robust and reliable model update, able to inform on field connectivity [29,30], geological features identification [31] or support to decision-making [32], for instance. The complexity lies in the fact that seismic contains uncertainty, such as structural errors and noise, which are very difficult to estimate. For this very reason it is hard to find wells and seismic data in agreement, therefore a multi-objective function approach is sound.

2.1. Multi-objective evolutionary algorithms

Multi-Objective Evolutionary Algorithms (MOEAs) are frequently used for solving the SHM problems. These methods aggregate the different objectives of the SHM problem through *Pareto-optimality*. In terms of Pareto-optimality, a solution $\mathbf{x}$ dominates (i.e., is better than) another solution $\mathbf{y}$, if and only if they satisfy the following two statements:

1. All the objective values of $\mathbf{x}$ are no worse than those of $\mathbf{y}$.
2. There is at least one objective value of $\mathbf{x}$ is better than that of $\mathbf{y}$.

This definition is equivalent to Eq. (1), which describes a minimization problem with M objectives. This aggregation method focuses on retrieving an optimal set of trade-off solutions (called Pareto Front) under several conflicting objectives.

$$\begin{aligned}\mathbf{x} \prec \mathbf{y} \Leftrightarrow \; & \forall i = 1, \ldots, M, f_i(\mathbf{x}) \leq f_i(\mathbf{y}) \wedge \\ & \exists i = 1, \ldots, M, f_i(\mathbf{x}) < f_i(\mathbf{y})\end{aligned} \tag{1}$$

There are two classical MOEA algorithms that have been frequently cited in SHM literature, and illustrate two different approaches to find a Pareto Front: The Non-dominated Sorting Genetic Algorithm II (NSGA-II) [21] uses a domination approach, while the Reference Vector-guided Evolutionary Algorithm (RVEA) [22] uses a decomposition approach. (see Table 3).

2.1.1. Non-dominated sorting genetic algorithm II

NSGA-II [21] is a traditional genetic algorithm where the *selection operator*, which selects the solutions to keep for subsequent iterations, is modified to take into account the notion of Pareto-optimality. This selection method contains two components, fast non-dominated sorting and crowding distance assignment. Fast non-dominated sorting assigns a rank to each solution depending to their dominance relationship to the rest of the solution set. The tie-breaker for solutions with the same rank is the crowding distance, which makes similar solutions in the objective space less likely to be selected.

Other than the selection method, NSGA-II [21] uses elitism strategy, and generates new solutions by simulated binary crossover and polynomial mutation.

2.1.2. Reference vector-based evolutionary algorithm

RVEA [22] is a decomposition-based MOEA proposed for many-objective optimization problems (MaOPs). MaOPs contain more than four objectives. In these high dimensionality situations, domination-based selection does not provide enough selection pressure for the optimization process. To address this issue, RVEA [22] uses a novel selection method called reference vector-guided selection. It contains four steps: objective value transition, population partition, angle penalized distance (ADP) calculation, and the elitism selection. The objective value transition scales the objective values so that the length of the objective vectors is between 0 and 1. The next step, population partition, divides the whole population into several sub-population based on the cosine similarity between the objective vectors of the individual and the unit reference vectors. After that, ADP of the individuals will be computed as the objective vector length with a penalty related to the angle between the objective vector and the reference vector. The algorithm selects the elites of each sub-population based on ADP.

The remainder of RVEA [22] contains random parent selection, simulated binary crossover and polynomial mutation, as well as an additional step to automatically adjust the reference vectors by generations.

2.2. Multi-objective evolutionary algorithms in seismic history matching literature

The early studies in the SHM literature aggregated the objectives by weighted sum and solved with Single-Objective Evolutionary Algorithms such as Particle Swarm Optimization (PSO) [33,34]. However, many recent studies started to solve the SHM problems using MOEAs.

Schulze applied Strength Pareto Evolutionary Algorithm (SPEA) to solve the SHM problem [9], while Min and Negash modified and applied Multi-Objective GA (MOGA) in three separate works [35–37]. NSGA-II is the most frequently used algorithm in the SHM literature [15–19,38]. Mohamed [11] and Christie [12] did the similar comparison studies between Multi-Objective PSO (MOPSO) and Single-Objective PSO (SOPSO) in two separate works. Hutahaean applied MOPSO in his two studies [13,14]. Ilamah applied MOEA/DD [39]. Hutahaean applied RVEA to solve many-objective SHM problems (4-objectives and 6-objectives) [20].

Based on our review, subsurface researchers have applied various MOEAs to solve the SHM problems, including domination approaches such as SPEA [9], SPEA2 [10], MOPSO [11–14], and NSGA-II [15–19], as well as decomposition approach such as RVEA [20] and MOEA/DD [39]. Most of these studies have indicated that the MOEAs can achieve better results on the SHM problems, compared to using single-objective EAs.

However, we suggest that there are at least three limitations of using MOEAs to solve the SHM problems.

1. Many multi-objective optimization problems assume that the objectives are strict trade-offs in competition with each other. We think, that for SHM problems the trade-off between competing objectives can be considered weak and that many objectives are either localized (do not impact each other) or are complimentary, where an improving fitness in one objective is usually complimentary of other objectives. This notion also fits with the overall desire to find models which best fit all objectives in SHM. The complexity and variability of SHM leads to a unique level of objective competition for each model, making generalization of this rule difficult.
2. Following from 1, it is usually not necessary (or desirable) to achieve the entire Pareto Front in SHM problems. Reservoir engineers are normally not interested in extreme points that hold high misfit on several objectives but low misfit on the rest, as such points usually represent unphysical solutions (data unbalanced solutions).
3. Some components of the MOEAs, such as the crowding distance assignment in NSGA-II [21], aggregate different objective values by the simple addition. This is a questionable practice when different objectives correspond different physical entities, and requires arbitrary tuning of the scaling factors for these objectives.

In the next section, we introduce a different selection method named *Lexicase Selection* [24] that handles the multiple objectives in a different manner from MOEAs. We argue that Lexicase selection is able to overcome these limitations.

2.3. Lexicase selection

The *Lexicase Selection* was first proposed for Genetic Programming (GP) to solve the program synthesis problems [24]. In a program synthesis problem [24], the goal is to find a computer program that passes a set of example input and output. Therefore, in the program synthesis problem [24], solutions are expected to have good performance on all objectives. This shows a difference from Multi-Objective Optimization (MOO), where solutions are allowed to perform sub-optimally on some objectives (i.e., extreme points). We think the SHM problems are more similar to the program synthesis tasks rather than MOOs, since a solution performing well in one objective should not, in principle, always cause it to perform worse in another.

Algorithm 1 Automatic ϵ-Lexicase Selection

Shuffle the objective list into $F = \{f_{r_1}, \dots, f_{r_M}\}$;
Set the candidate pool A as the entire population X;
while $|F| > 0$ and $|A| > 1$ **do**
  Denote the first item in the objective list F as f;
  for x_j in A **do**
    if $f(x_j) > f_{best} + \sigma$ **then**
      Delete x_j from A;
    end if
  end for
  Pop the first item from F;
end while

The basic idea of Lexicase Selection [24] is to filter the population based on each of the objectives in a random order. Each time before the algorithm selects an individual, the order of the objectives will be shuffled. Then the algorithm keeps the best individuals based on the objectives in the previous shuffled order, until there is only one individual left or all the objectives have been used (in this case, the algorithm returns a random individual from the rest individuals).

One disadvantage of this selection scheme is that it can lead to a poor performance on problems with a continuous fitness space, since few individuals share the same elitism unless they are exactly identical. Therefore, in this case, only one objective will be used to select a parent in the continuous fitness space.

To solve this problem, Cava proposed the *Automatic ϵ-Lexicase Selection* in 2016 [40]. Automatic ϵ-Lexicase Selection [40] differs from the basic Lexicase Selection [24] by introducing an adaptive threshold parameter σ to solve the previous issue. For a minimization problem, the individuals whose fitness value is less than $f_{min} + \sigma$ are considered as the "best" individuals. f_{min} is the minimum objective value in the population. σ is calculated based on median absolute deviation (MAD) in (2), where $med(\cdot)$ takes the median of a set. A detailed procedure of this method is provided in Algorithm 1.

$$\sigma = med\left(\left\{|\, med(f(A)) - f(x_i)\,|\right\}_{i=1,\dots,|A|}\right) \tag{2}$$

In the Automatic ϵ-Lexicase Selection [40], each parent is elite on at least the first objective used to select it. Since each parent is selected by a random order of the objectives, the individuals are pressured to perform well on various combinations of the objectives, which enhances the diversity of the population. However, the disadvantage of this method is also obvious. When the number of objectives is small, there are not enough combinations of objectives to provide the diversity. Thus, the algorithm can become greedy. But since the available computational budget in the SHM problems is usually limited, we believe this greedy characteristic does little harm for this application case.

3. Proposed method: Differential evolution based on automatic ϵ-lexicase selection

In this study, we propose a novel method to solve the SHM problems based on the Differential Evolution (DE) [41] and the Automatic ϵ-Lexicase Selection [40]. DE [41] is a simple yet powerful optimization algorithm especially on continuous domain. Its superiority has been proven in many prior studies [7,42,43].

In most MOEAs, there are some components doing arithmetic operations on the objective values of the different objective functions. For example, the crowding distance in NSGA-II [21] is computed as a Manhattan distance that adds objective values of different objective functions directly. This step is influenced by the different scales of

Algorithm 2 Differential Evolution based on Automatic ϵ-Lexicase Selection

```
Initialize a population X = {x_1, ..., x_N};
Evaluate the fitness of every individuals in X;
while termination criteria are not satisfied do
    Create an empty population X' = {};
    for i in 1 to N do
        Select a parent x_lexicase by Automatic ε-Lexicase Selection;
        Select two parents x_1 and x_2 randomly;
        Reproduce offspring y based on (4);
        Perform polynomial mutation on y to generate y';
        Add y' to X';
    end for
    Replace X with X';
end while
```

the objective functions. Therefore, a proper weight assigning step is necessary. This step is not trivial and usually based on an engineering judgement. However, by using Automatic ϵ-Lexicase Selection [40], there is no need to do arithmetic operations on the objective values of the different objective functions. Therefore, this selection method is not influenced by the different scales of the objectives and there is no need to do weighting.

While there are multiple proposed DE variants, one of the most frequently used versions is DE/rand/1/bin [41]. In this method, three parents x_1, x_2, and x_3 are selected randomly to generate an offspring individual y based on (3). In (3), the subscript j means the jth dimension of a vector. F is a scaling factor to control the mutation strength, and CR controls the binary crossover rate. j_{r_2} is a randomly selected dimension that ensures at least one dimension in the solution is mutated.

$$y_j = \begin{cases} x_{1,j} + F \cdot (x_{2,j} - x_{3,j}), & r_1 \leq CR \text{ or } j = j_{r_2} \\ x_{1,j}, & \text{otherwise} \end{cases} \tag{3}$$

In this study, we propose DE/lexicase/1/bin in Algorithm 2. This method is similar as DE/rand/1/bin [41], however, we replace the first parent with a "good" individual x_l selected by Automatic ϵ-Lexicase Selection [40]. This method performs mutation on the selected individual, by adding a differential vector between two random individuals.

We further modify the original DE algorithm [41] as follows. We omit the survival selection in the original DE procedure, since the Lexicase Selection can provide sufficient selection pressure. We also introduce the polynomial mutation after the differential mutation to prevent premature convergence. We provide the procedure of our proposed method in Algorithm 2.

$$y_j = \begin{cases} x_{l,j} + F \cdot (x_{1,j} - x_{2,j}), & r_1 \leq CR \text{ or } j = j_{r_2} \\ x_{l,j}, & \text{otherwise} \end{cases} \tag{4}$$

As pointed out in Section 2.3, when the number of the objectives in the problem is small, the Lexicase Selection can become greedy. However, for a SHM problem, the available number of iterations is usually small, since it costs several minutes to hours per evaluation. Therefore, the greedy performance may not harm the optimization results (in the SHM problems).

In the remaining sections, we abbreviate our method to Lex-DE.

4. Experiments

To test the proposed Lex-DE algorithm, we prepared two SHM problems, TS2N and Volve. The optimal parameters, also known as ground truth, for both problems are known for us (of course not used by the algorithms). The datasets generated and analyzed during the current study are available in a public repository.[1]

[1] https://github.com/Y1fanHE/lexde-subsurface-model

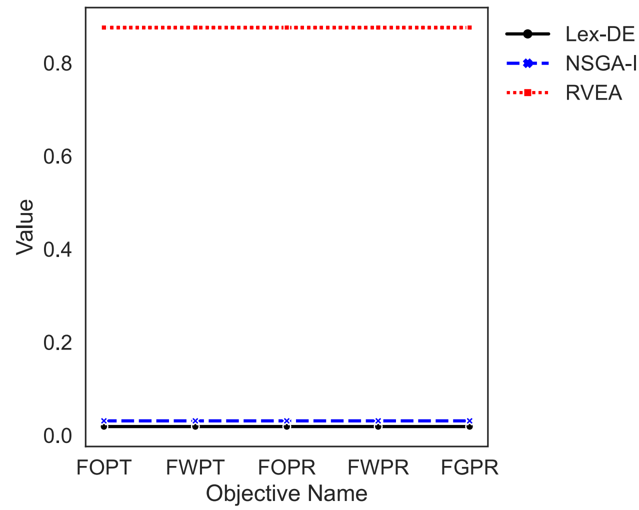


Fig. 1. Parallel objective plots for the best run on the Volve problem. Each line corresponds to the five (scaled) objective values of a non-dominated solution.

Table 1
Details of the objectives in the TS2N problem.

Name	Details
FWPR	Field Water Production Rate
FOPR	Field Oil Production Rate
FWPT	Field Water Production Total $\int_0^t FWPR(t)dt$
FOPT	Field Oil Production Total $\int_0^t FOPR(t)dt$
FGPR	Field Gas Production Rate

Table 2
Details of the objectives in the Volve problem.

Name	Details
seis-mean	Seismic metric using Mean Attribute
seis-spa	Seismic metric using SPA Attribute
P-F-14	Composite well F-14 fitness metric
P-F-12	Composite well F-12 fitness metric
P-F-15C	Composite well F-15C fitness metric

4.1. Test problems

SHM models based on real world datasets, while unique, are broadly similar in their objectives and the challenges they present for optimization. We have selected two test problems which cover different model scales and complexities. Additionally the models are open and available for additional research.

The TS2N model [25] simulates a reservoir located in the Gulf of Mexico with a single production well. The model includes monthly production data for Oil, Gas and Water volumes from 1996 to 1999. Within the model there are five geological layers with uniform properties. It is a real life example, albeit a simple one. It includes five objectives, namely FWPR, FOPR, FWPT, FOPT, and FGPR. The detail of the five objectives is provided in Table 1. The production totals (PT) metrics are time integrals of the production rates (PR). Including them as an objective emphasizes the requirement for the model to match the actual produced volumes of the field rather than just the production rates. This is a key requirement in history matching. Production totals tend towards increasing error overtime which places more weight on matching the total produced volume at the end of the optimization period.

The TS2N model has 13 model parameters that include horizontal and vertical permeability multipliers for each layer, the oil–water contact depth, porosity and the reservoir compressibility.

To increase the uncertainty of the TS2N model, we added randomized noise to the production rate history and reintegrated the rates to

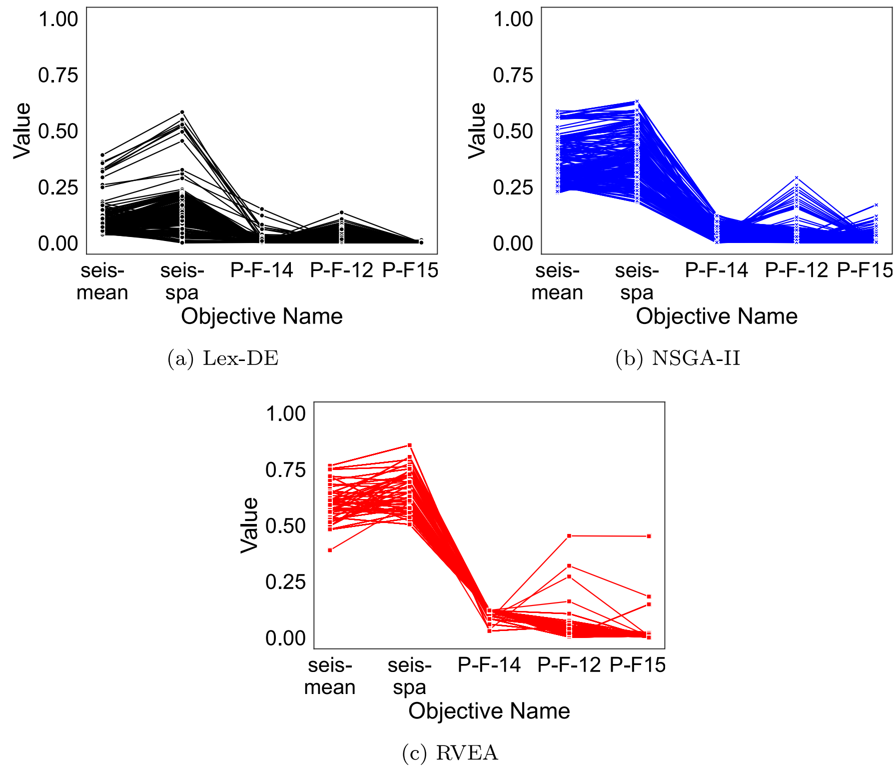


Fig. 2. Parallel objective plots for the best run on the Volve problem. Each line corresponds to the five (scaled) objective values of a non-dominated solution.

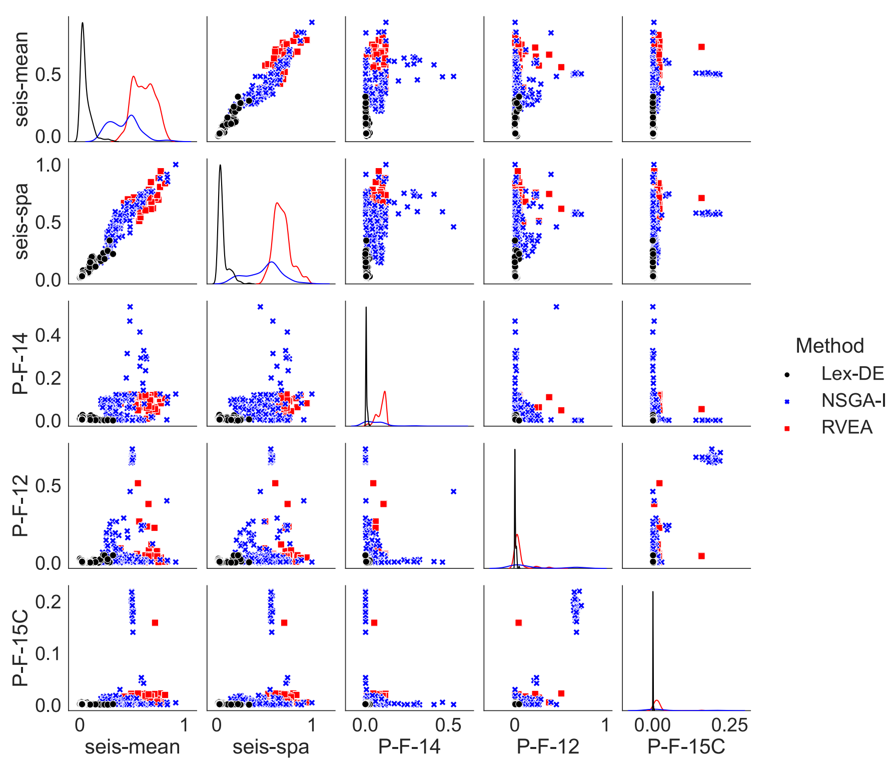


Fig. 3. Scatter plot of the objective function values of non-dominated solutions in the best run on the Volve problem.

create production totals which are slightly different to the truth case. The noise is not correlated between objectives and this creates a small degree of trade-off in the objective function that would otherwise be absent.

The Volve field is an open dataset [26]. This problem also includes five objectives, namely seis-mean, seis-spa, P-F-14, P-F-12, and P-F-15C. The detail of the five objectives are provided in Table 2. The Volve model has 63 model parameters to generate a large search space. The

Table 3
Comparison on the proposed Lex-DE with NSGA-II and RVEA.

Algorithm	Objective aggregation	Variation methods
Lex-DE	Shuffle the priority of the objectives	Differential mutation Binary crossover Polynomial mutation
NSGA-II	Non-dominated sorting Crowding distance	Simulated binary crossover Polynomial mutation
RVEA	Tchebycheff decomposition	

Table 4
Average distance to the ground truth $\bar{d}_{g.t.}$ of three algorithms on the TS2N problem.

Method	Best	Median	Worst	Mean	Std.
Lex-DE	0.157	0.339	0.623	**0.352**	0.170
RVEA	0.684	0.719	0.953	0.793	0.123
NSGA-II	0.256	**0.319**	0.498	0.359	0.099

Table 5
Difference on set coverage $\Delta C(A, B)$ of three algorithms on the TS2N problem.

Methods (A, B)	(Lex-DE, RVEA)	(Lex-DE, NSGA-II)	(RVEA, NSGA-II)
ΔC	100%	100%	−100%

Table 6
Average distance to the ground truth $\bar{d}_{g.t.}$ of three algorithms on the Volve problem.

Method	Best	Median	Worst	Mean	Std.
Lex-DE	2.102	**2.368**	3.076	**2.464**	0.376
RVEA	2.947	3.127	3.222	3.086	0.110
NSGA-II	2.383	2.764	2.840	2.675	0.187

Table 7
Difference on set coverage $\Delta C(A, B)$ of three algorithms on the Volve problem.

Methods (A, B)	(Lex-DE, RVEA)	(Lex-DE, NSGA-II)	(RVEA, NSGA-II)
ΔC	100%	95%	−98%

parameters of the model include the oil–water contact, fault transmissibilities, region and zone permeability and porosity multipliers and aquifer volume. Unlike the TS2N problem, this problem includes data from multiple wells and multiple seismic.

4.2. Experimental methods

We compared Lex-DE with two MOEAs, namely RVEA [22] and NSGA-II [21]. Before the formal experiments, we tuned several important parameters but set the rest based upon experience in prior studies [21,22] with the exception of population size. For each algorithm, we run five repetitions. We set the total number of evaluations as 1500 for all three algorithms on TS2N, and 2000 on Volve. The population size is set as 20 (75 generations on TS2N and 100 generations on Volve). For Lex-DE, we set the scaling vector $F = 0.5$ and the crossover rate $CR = 0.5$ without tuning. The mutation rate p_m is set to $1/n$ (n is the dimension of the problem). For NSGA-II, the crossover rate p_c is set to 0.9 (tuned from {0.6, 0.7, 0.8, 0.9, 1.0}). The mutation rate p_m is set to $1/n$ (n is the dimension of the problem). For RVEA, the crossover rate and mutation rate are the same as NSGA-II. We generate 15 weight vectors based on *das-dennis method* [44]. We set the rest parameters, $\alpha = 2.0$ and $f_r = 0.1$, based on the original RVEA paper [22].

We include two numerical metrics of performance, average distance to the ground truth and the difference on set coverage. Before computing the metrics, we scaled the objective values into $[0, 1]$ based on the non-dominated solutions over all evaluations in the five repetitions of the three algorithms.

- **Average distance to the ground truth** ($\bar{d}_{g.t.}$). This metric shows the scaled Euclidean distance between non-dominated solution set A and the ground truth x^* in the parameter space.

$$\bar{d}_{g.t.} = \frac{\Sigma_{x \in A} \parallel x - x^* \parallel}{\mid A \mid} \tag{5}$$

- **Difference on set coverage** ($\Delta C(A, B)$). Let A and B be two non-dominated sets, the set coverage $C(A, B)$ is defined as the percentage of the solutions in B that are dominated by at least one solution in A. When computing this set coverage, the non-dominated solution set of every methods is the non-dominated set of union of the solutions from five repetitions. The difference on set coverage is computed as in (6). A positive value of $\Delta C(A, B)$ shows that A is better than B considering all the objectives.

$$\Delta C(A, B) = C(A, B) - C(B, A) \tag{6}$$

$$C(A, B) = \frac{|\{u \in B | \exists v \in A : v \prec u\}|}{|B|} \tag{7}$$

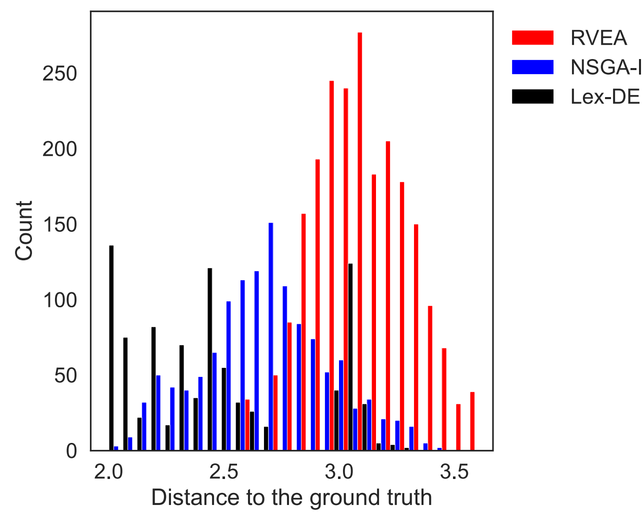


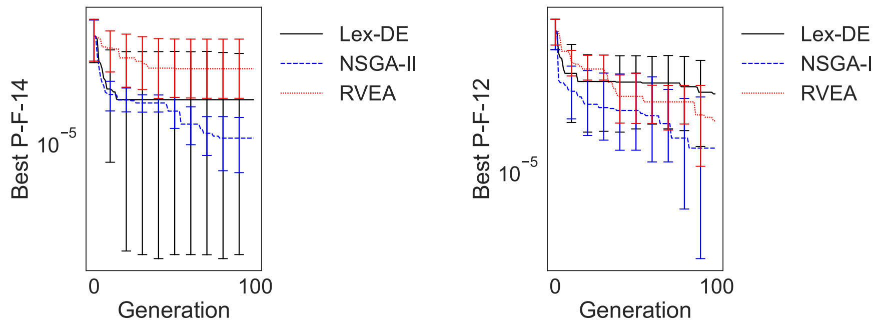
Fig. 4. Frequency of scaled Euclidean distance between non-dominated solutions to the ground truth in the parameter space.

4.3. Experimental results

Tables 4 and 6 provide the average distance to the ground truth $\bar{d}_{g.t.}$ on the TS2N and Volve problems. On both problems, the solution set found by Lex-DE is closest on average to the ground truth. Tables 5 and 7 shows the difference on set coverage on the two problems. On both problems, ΔC(Lex-DE, RVEA) and ΔC(Lex-DE, NSGA-II) are positive, indicating that a high proportion of solutions by Lex-DE dominates the solution sets of both RVEA and NSGA-II.

Figs. 7 and 8 shows the change in best objective values by generation on the TS2N and Volve problem, indicating the convergence of the algorithms on these problems. The objective values are scaled into $[0, 1]$ based on all the solutions found in five repetition of three algorithms. These figure shows that Lex-DE and NSGA-II approach the best solutions faster than RVEA in TS2N, and that on Volve Lex-DE finds better solutions faster on seis-mean, seis-spa and P-F-14, while NSGA-II finds better solutions faster on P-F-15C.

Figs. 1 and 2 provide the parallel coordinate graph of the TS2N and Volve problems, respectively. These figures show the non-dominated solutions over all evaluations in the run with best average distance to the ground truth. In the graph, a non-dominated solution is represented by five points connected by a line. The five points show the five objective values (scaled into $[0, 1]$) of this solution. We find that on the TS2N problem (Fig. 1), there is only one non-dominated solution for



(a) P-F-14 (b) P-F-12

Fig. 5. Best prediction fitness by generations (y-axis in log scale) on the Volve problem.

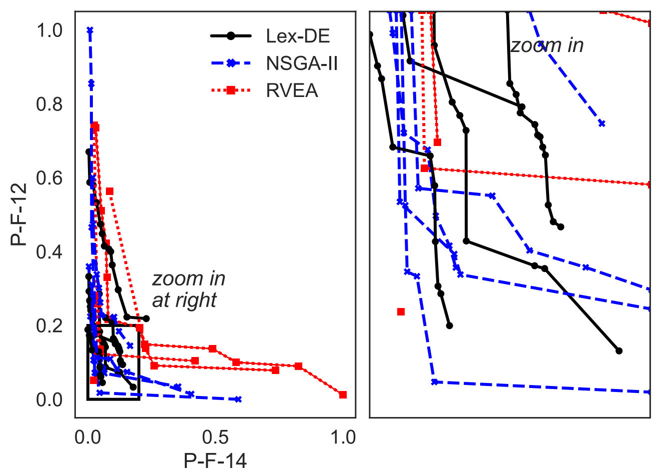


Fig. 6. Scatter plot of the non-dominated solutions in the prediction period on the Volve problem. The right plot is the zoom-in of the rectangle area in the left graph.

each algorithm. This shows the lack of trade-off between the objectives in TS2N.

On TS2N, the non-dominated solution of Lex-DE is slightly better than that of NSGA-II, but much better than that of RVEA. On the Volve problem, more than one non-dominated solutions have been found for all three methods. The parallel objectives plot in Fig. 2 shows that Lex-DE generates solutions with clearly better values than the other methods for the objectives seis-mean and seis-spa, and somewhat better values on P-F-14, P-F-12, and P-F-15. This shows that the proposed method finds solutions that are generally good across all objectives, including objectives of different nature (seismic and well) when compared to the other two MOEAs.

To better understand these results, Fig. 3 provides a scatter plot of the solutions in Fig. 2. This figure shows how the non-dominated solutions of the Lex-DE algorithm are distributed in a central area in the non-dominated front. However, for the other two MOEAs (especially the NSGA-II), their solution sets are spread over a larger area, including some "extreme points" that perform sub-optimally on some objectives.

5. Discussion

In Section 4.3, we find that the performance of Lex-DE and NSGA-II on the simple TS2N problem are close. However, on the harder Volve problem, Lex-DE has a better results compared to NSGA-II and RVEA. In this section, we provide further discussion based on the Volve results.

Table 8
Difference on set coverage $\Delta C(A, B)$ in prediction of three algorithms on the Volve problem.

Methods (A, B)	(Lex-DE, RVEA)	(Lex-DE, NSGA-II)	(RVEA, NSGA-II)
ΔC	−45%	−28%	−23%

5.1. Distribution of distance to the ground truth

Lex-DE uses the Lexicase Selection. Therefore, the solutions are pressured to perform "good" on all objectives. This feature brings two main benefits.

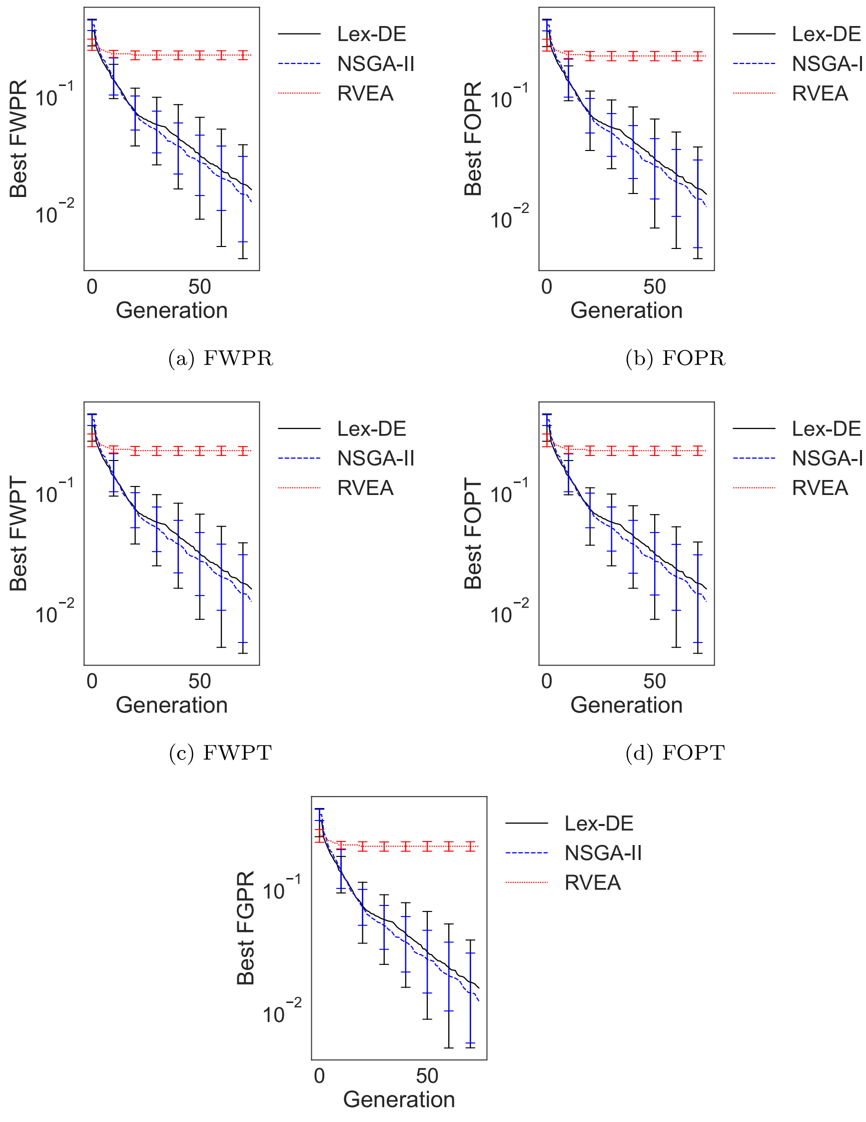
1. For the final solution set, Lexicase Selection centralizes the solutions in a small area. To illustrate this point, in Fig. 4, we show the histogram of the scaled distance between every non-dominated solution and the ground truth in the parameter space in all repetitions. We clearly find that Lex-DE generates more solutions than the other two MOEAs in the area that is close to the ground truth (distance less than 2.2). In real-world engineering tasks, the final set of models when using Lex-DE as an optimization tool is closer to the truth and contains less nonphysical results.
2. For the evolutionary process itself, a centralized solution set usually holds a stronger exploitation, and thus the algorithm can converge to a better result in a shorter amount of time. This is advantageous for SHM problems where evaluations a usually limited due to computation cost.

However, the disadvantages is also obvious. Focusing on a specific area may lead to many solutions within a local optimum. For example, in Fig. 4, there are solutions of Lex-DE distributed between 3.0 and 3.2. We suspect this run falls within a local optimum, and thus results are worse than other runs.

5.2. Performance in the prediction period

In the real world, the subsurface model is used to do forecasting on the field production. We perform prediction based on every individuals generated during the optimization (2000 individuals per run). We only compute the misfit on P-F-12 and P-F-14, since the other three objectives are not available in the prediction period. The difference on set coverage based on the fitness in prediction is provided in Table 8. Regarding the set coverage in prediction and the evolution of the prediction fitness (Fig. 5), NSGA-II performs the best. We consider the following three possible reasons.

- Our Lex-DE performs much better on the two seismic related objectives. However, they are not available in the prediction. In the optimization period (Fig. 2), NSGA-II also generated several solutions that are good in P-F-14 and P-F-12. Therefore, it is not strange for NSGA-II to have a good prediction result.



(a) FWPR (b) FOPR (c) FWPT (d) FOPT (e) FGPR

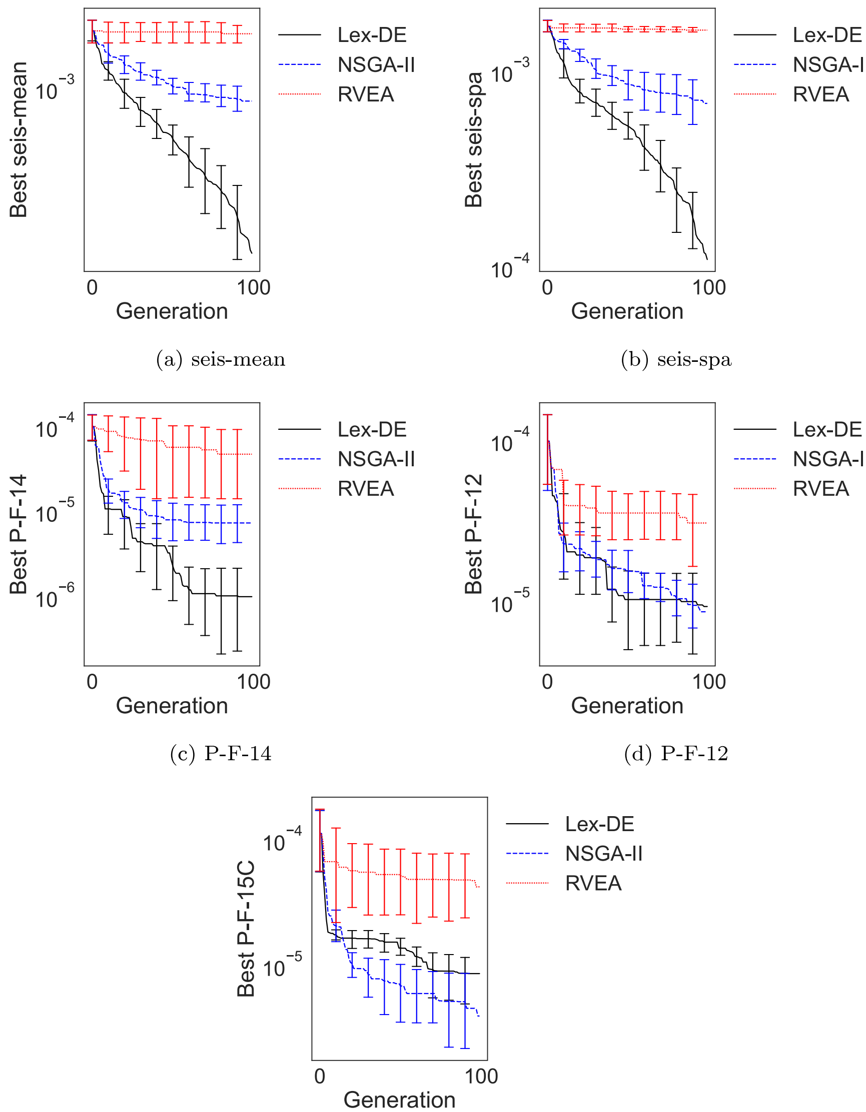
Fig. 7. Best fitness by generations (y-axis in log scale) on the TS2N problem.

- Fig. 6 shows the non-dominated solutions in the prediction period of three algorithms in all five runs, as well as a zoom-in graph of the center part of the non-dominated front. The solutions in the same run are connected by a line. In most runs, NSGA-II and RVEA generate several solutions that perform well on some objectives, but sub-optimally on the others. These "sub-optimal" solutions may not be dominated by any of the solutions of Lex-DE, but they are less useful in the SHM problem.
- Our method may overfit the problem during the optimization phase. Though all three algorithms do not contain any explicit way to overcome overfitting, NSGA-II and RVEA tend to maintain a more diverse solution set. This may lead to the better results in the prediction period. To enhance the diversity of Lex-DE without generating sub-optimal solutions, we can use other strong global mutation methods or restart strategies, and keep the Lexicase Selection.

6. Conclusions

In this study, we introduced the Lexicase Selection method [24,40] and proposed the Differential Evolution based on Automatic ϵ-Lexicase Selection algorithm to solve the history matching problems. We compared the proposed algorithm with two other literature methods, the NSGA-II [21] and the RVEA [22], on two real-world examples [25,26]. The results have shown the superiority of the proposed method with better optimization results (i.e., positive difference on set coverage and smaller average distance to the ground truth) and a more centralized solution set. What is more, we found that this centralized set usually provides more solutions close to the ground truth in the parameter space. For an engineering problem, this feature generates a final ensemble of models which better characterize the true model and parameter uncertainty.

Despite the above advantages, this method sometimes falls into local optimum. What is more, the prediction performance (difference on set coverage in prediction) of Lex-DE is not as good as in the optimization phase. This may caused by the following reasons: (1) some objectives are not available in the prediction period; (2) the set coverage may be affected by the extreme points; (3) our Lex-DE may get overfitting in the optimization phase. In the future, we are going to increase the diversity of Lex-DE by applying strong global mutation methods. This can solve the local optimum and the overfitting issue, without generating sub-optimal solutions.



(a) seis-mean

(b) seis-spa

(c) P-F-14

(d) P-F-12

(e) P-F-15C

Fig. 8. Best fitness by generations (y-axis in log scale) on the Volve problem.

In addition, the main outcome of this paper is developed from questioning whether we should model the history matching problem as Multi-Objective Optimization. In the recent optimization literature, Multi-Form Optimization [23] has been proposed to reconcile multiple alternate formulations of a single target task of interest. Part of our future research will consider the history matching problem as a Multi-Form Optimization task.

CRediT authorship contribution statement

Yifan He: Methodology, Software, Investigation, Visualization, Writing – original draft, Writing – review & editing. **Claus Aranha:** Conceptualization, Supervision, Writing – original draft, Writing – review & editing. **Antony Hallam:** Methodology, Software, Investigation, Writing – original draft, Writing – review & editing. **Romain Chassagne:** Conceptualization, Supervision, Writing – original draft, Writing – review & editing, Funding acquisition.

Declaration of competing interest

One or more of the authors of this paper have disclosed potential or pertinent conflicts of interest, which may include receipt of payment, either direct or indirect, institutional support, or association with an entity in the biomedical field which may be perceived to have potential conflict of interest with this work. For full disclosure statements refer to https://doi.org/10.1016/j.orp.2022.100237.Tony Hallam, Romain Chassagne reports financial support was provided by Edinburgh Time-Lapse Project (ETLP) Phase VII. Romain Chassagne reports financial support was provided by Japanese Society for Promotion of Science (JSPS), fellowship L19544. YH reports financial support was provided by Japanese Society for Promotion of Science (JSPS), grant 17K12690

Acknowledgments

AH and RC would like to thank the sponsors of Edinburgh Time-Lapse Project (ETLP) Phase VII (AkerBP, BP, CGG, Chevron/Ithaca Energy, CNOOC, Equinor, ConocoPhillips, ENI, Petrobras, Norsar, Woodside, Taqa, Halliburton, ExxonMobil, OMV and Shell) for funding this work and Equinor for provision of the public data set. RC acknowledges the Japanese Society for Promotion of Science (JSPS) for the fellowship L19544 which initiated this work. YH acknowledges the Japanese Society for Promotion of Science (JSPS) for support under grant 17K12690 We also acknowledge Schlumberger for the use of their Eclipse software and the Python open source community.

References

[1] Mitchell P, Chassagne R. 4D assisted seismic history matching using a differential evolution algorithm at the harding South field. In: 81st EAGE conference and exhibition 2019. European Association of Geoscientists & Engineers; 2019.
[2] Chassagne R, Aranha C. A pragmatic investigation of the objective function for subsurface data assimilation problem. Oper Res Perspect 2020;7:100143.
[3] Obidegwu D, Chassagne R, MacBeth C. Seismic assisted history matching using binary maps. J Nat Gas Sci Eng 2017;42:69–84.
[4] Oliver DS, Fossum K, Bhakta T, Sandø I, Nævdal G, Lorentzen RJ. 4D seismic history matching. J Pet Sci Eng 2021;207:109119.
[5] Zhang Q, Chassagne R, MacBeth C. Seismic history matching uncertainty with weighted objective functions. In: ECMOR XVI - 16th european conference on the mathematics of oil recovery. EAGE Publications BV; 2018.
[6] Stephen KD, Soldo J, Macbeth C, Christie MA. Multiple model seismic and production history matching: A case study. SPE J 2006;11(04):418–30.
[7] Aranha C, Tanabe R, Chassagne R, Fukunaga A. Optimization of oil reservoir models using tuned evolutionary algorithms and adaptive differential evolution. In: 2015 IEEE congress on evolutionary computation (CEC). IEEE; 2015.
[8] Zhang Q, Chassagne R, MacBeth C. 4D seismic and production history matching, a combined formulation using Hausdorff and Fréchet metric. In: SPE europec featured at 81st EAGE conference and exhibition. SPE; 2019.
[9] Schulze-Riegert RW, Krosche M, Fahimuddin A, Ghedan SG. Multi-objective optimization with application to model validation and uncertainty quantification. In: SPE middle east oil and gas show and conference. SPE; 2007.
[10] Verga F, Cancelliere M, Viberti D. Improved application of assisted history matching techniques. J Pet Sci Eng 2013;109:327–47.
[11] Mohamed L, Christie M, Demyanov V. History matching and uncertainty quantification: Multiobjective particle swarm optimisation approach. In: SPE EUROPEC/EAGE annual conference and exhibition. SPE; 2011.
[12] Christie M, Eydinov D, Demyanov V, Talbot J, Arnold D, Shelkov V. Use of multi-objective algorithms in history matching of a real field. In: SPE reservoir simulation symposium. SPE; 2013.
[13] Hutahaean JJ, Demyanow V, Christie MA. Impact of model parameterisation and objective choices on assisted history matching and reservoir forecasting. In: SPE/IATMI Asia Pacific oil & gas conference and exhibition. SPE; 2015.
[14] Hutahaean J, Demyanov V, Christie MA. On optimal selection of objective grouping for multiobjective history matching. SPE J 2017;22(04):1296–312.
[15] Sayyafzadeh M, Haghighi M, Carter JN. Regularization in history matching using multi-objective genetic algorithm and Bayesian framework. In: SPE Europec/EAGE annual conference. SPE; 2012.
[16] Kanfar MS, Clarkson CR. Reconciling flowback and production data: A novel history matching approach for liquid rich shale wells. J Nat Gas Sci Eng 2016;33:1134–48.
[17] Park H-Y, Datta-Gupta A, King MJ. Handling conflicting multiple objectives using Pareto-based evolutionary algorithm during history matching of reservoir performance. J Pet Sci Eng 2015;125:48–66.
[18] Kim J, Kang J, Park C, Park Y, Park J, Lim S. Multi-objective history matching with a proxy model for the characterization of production performances at the shale gas reservoir. Energies 2017;10(4):579.
[19] Zhang Z, Jung HY, Datta-Gupta A, Delshad M. History matching and optimal design of chemically enhanced oil recovery using multi-objective optimization. In: SPE reservoir simulation conference. SPE; 2019.
[20] Hutahaean J, Demyanov V, Christie M. Many-objective optimization algorithm applied to history matching. In: 2016 IEEE symposium series on computational intelligence (SSCI). IEEE; 2016.
[21] Deb K, Pratap A, Agarwal S, Meyarivan T. A fast and elitist multiobjective genetic algorithm: NSGA-II. IEEE Trans Evol Comput 2002;6(2):182–97.
[22] Cheng R, Jin Y, Olhofer M, Sendhoff B. A reference vector guided evolutionary algorithm for many-objective optimization. IEEE Trans Evol Comput 2016;20(5):773–91.
[23] Gupta A, Ong Y-S. Back to the roots: Multi-X evolutionary computation. Cogn Comput 2019;11(1):1–17.
[24] Helmuth T, Spector L, Matheson J. Solving uncompromising problems with lexicase selection. IEEE Trans Evol Comput 2015;19(5):630–43.
[25] Christie M, MacBeth C, Subbey S. Multiple history-matched models for teal South. Lead Edge 2002;21(3):286–9.
[26] Hallam T, MacBeth C, Chassagne R, Amini H. 4D seismic study of the volve field - an open subsurface-dataset. First Break 2020;38(2):59–70.
[27] Musuuza JL, Gustafsson D, Pimentel R, Crochemore L, Pechlivanidis I. Impact of satellite and in situ data assimilation on hydrological predictions. Remote Sens 2020;12(5):811.
[28] Essouayed E, Verardo E, Pryet A, Chassagne RL, Atteia O. An iterative strategy for contaminant source localisation using GLMA optimization and data worth on two synthetic 2D aquifers. J Contam Hydrol 2020;228:103554.
[29] Corte G, Chassagne R, MacBeth C. Seismic history matching in the pressure and saturation domain for reservoir connectivity assessment. In: 82nd EAGE annual conference & exhibition. European Association of Geoscientists & Engineers; 2021.
[30] Yin Z, MacBeth C, Chassagne R. Joint interpretation of interwell connectivity by integrating 4D seismic with injection and production fluctuations. In: 77th EAGE conference & exhibition incorporating SPE EUROPEC 2015. SPE; 2015.
[31] Rahon D, Edoa PF, Masmoudi M. Identification of geological shapes in reservoir engineering by history matching production data. SPE Reserv Eval Eng 1999;2(05):470–7.
[32] Cavalcante CCB, Maschio C, Santos AA, Schiozer D, Rocha A. History matching through dynamic decision-making. In: Conejero JA, editor. PLoS One 2017;12(6):e0178507.
[33] Onwunalu JE, Durlofsky LJ. Application of a particle swarm optimization algorithm for determining optimum well location and type. Comput Geosci 2009;14(1):183–98.
[34] Rwechungura R, Dadashpour M, Kleppe J. Application of particle swarm optimization for parameter estimation integrating production and time lapse seismic data. In: SPE offshore europe oil and gas conference and exhibition. SPE; 2011.
[35] Min B, Kang JM, Chung S, Park C, Jang I. Pareto-based multi-objective history matching with respect to individual production performance in a heterogeneous reservoir. J Pet Sci Eng 2014;122:551–66.
[36] Min B, Kang JM, Lee H, Jo S, Park C, Jang I. Development of a robust multi-objective history matching for reliable well-based production forecasts. Energy Explor Exploit 2016;34(6):795–809.
[37] Negash BM, Ayoub MA, Jufar SR, Robert AJ. History matching using proxy modeling and multiobjective optimizations. In: ICIPEG 2016. Springer Singapore; 2017, p. 3–16.
[38] Han YM, Park C, Kang JM. Estimation of future production performance based on multi-objective history matching in a waterflooding project. In: SPE EUROPEC/EAGE annual conference and exhibition. SPE; 2010.
[39] Ilamah O. A multiobjective dominance and decomposition algorithm for reservoir model history matching. Petroleum 2019;5(4):352–66.
[40] Cava WL, Spector L, Danai K. Epsilon-lexicase selection for regression. In: Proceedings of the genetic and evolutionary computation conference 2016. ACM; 2016.
[41] Price KV. Differential evolution. In: Handbook of optimization. Springer Berlin Heidelberg; 2013, p. 187–214.
[42] Mirzabozorg A, Nghiem L, Chen Z, Yang C. Differential evolution for assisted history matching process: SAGD case study. In: SPE heavy oil conference-Canada. SPE; 2013.
[43] Hajizadeh Y, Christie M, Demyanov V. History matching with differential evolution approach; A look at new search strategies. In: SPE EUROPEC/EAGE annual conference and exhibition. SPE; 2010.
[44] Das I, Dennis JE. Normal-boundary intersection: A new method for generating the Pareto surface in nonlinear multicriteria optimization problems. SIAM J Optim 1998;8(3):631–57.