

Chassagne, Romain; Aranha, Claus

## Article

# A pragmatic investigation of the objective function for subsurface data assimilation problem

Operations Research Perspectives

## Provided in Cooperation with:

Elsevier

*Suggested Citation:* Chassagne, Romain; Aranha, Claus (2020) : A pragmatic investigation of the objective function for subsurface data assimilation problem, Operations Research Perspectives, ISSN 2214-7160, Elsevier, Amsterdam, Vol. 7, pp. 1-15, <https://doi.org/10.1016/j.orp.2020.100143>

This Version is available at:

<https://hdl.handle.net/10419/246414>

### Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

### Terms of use:

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*



<https://creativecommons.org/licenses/by-nc-nd/4.0/>



# A pragmatic investigation of the objective function for subsurface data assimilation problem

Romain Chassagne<sup>\*,a</sup>, Claus Aranha<sup>b</sup>

<sup>a</sup> Energy, Geoscience, Infrastructure and Society, Heriot-Watt University, United Kingdom

<sup>b</sup> Faculty of Engineering, Information and Systems, University of Tsukuba, Tsukuba, Japan



## ABSTRACT

One of the main mechanisms of an optimization problem is the effectiveness and relevance of the objective function. In the context of an optimization problem in the subsurface domain, called seismic history matching, this study proposes to investigate further aspects of assimilating data. We focus on two main characteristics of the objective function: the influence and the sensitivity to the amount of data used in the seismic history matching. We select four metrics to analyse the similarity/dissimilarity measurement used in the matching. The optimization method used to perform the seismic history matching is an auto-adaptive differential evolution algorithm. This study has been carried out on three real datasets. Based on the results and analysis of the seismic history matching experiments, we are able to draw some practical suggestions on what kind of objective function should be established. Despite its simplicity, the Least Square metric performs as well as any other metric. Using all the possible data is safer but it is not compulsory to obtain good history matching results, in some cases using less data leads to the same answer. Using different metrics or more data does not change the computing time.

## 1. Introduction

In the subsurface community, a key problem is how to generate simulation models that can predict the behavior of the fluids flow within the reservoir, using data collected from the field [1–3]. In this context, *History Matching* (HM) is one key component of the subsurface management. HM consists of updating the reservoir model based on an optimization process that matches the simulated outputs with available observed data [4].

To update the reservoir model, the Geoscience Community generally uses all the available field data [5]. Namely, the production well data, which is very spatially localized type of data, and the seismic survey data, which is a more spatially global data. The integration of data from production wells into HM is a well-known process [6,7]. On the other hand, the assimilation of seismic data is not yet well established and still under very active study [8–14].

In this paper, we focus on the integration of four-dimensional (4D) inverted seismic data into the HM, also called *Seismic History Matching* (SHM). 4D seismic data consists of multiple three dimensional (area and depth) seismic surveys at different times. Such seismic data constitute an invaluable source of information on fluid displacement and geology over extensive areas of the reservoir [15]. It is believed that the large quantities of data available in the seismic surveys could be used to generate a much more robust prediction as compared to using only well production data. However, incorporating 4D seismic data into the

reservoir model through an optimization procedure is a challenging task [16–18], and no complete evidence of its benefits to HM has been seen in the literature so far.

This leads us to believe that something is still to be understood concerning the crucial step of integrating the seismic data into the history matching workflow. Therefore in this study we explore in more detail the main characteristics of SHM. Specifically, we focus our analysis on the choice of the objective function and how it affects the performance of the SHM. To simplify the SHM workflow we will consider the seismic data to be already inverted into pressure and saturation and focus on the data assimilation questions, or how we should formulate the objective function of this problem.

Classical formulations of the objective function are maybe not well adapted to inverting seismic data [19]. For example, the traditional least square based mismatch may not be the best representation of the visual difference between two seismic images, as it does not directly takes into account the relationship between values in different pixels of the seismic map image.

Furthermore, traditionally it is believed that a maximum number of different seismic attributes should be used and will guarantee the best assimilation into the history matching process (the more data the better). However, there is no study up until now that verifies this statement. This crucial point is about the amount of data used in the optimization process versus its relevance in the reduction of uncertainty. It raises the question: is it economically relevant, in terms of

\* Corresponding author.

E-mail addresses: [r.l.chassagne@hw.ac.uk](mailto:r.l.chassagne@hw.ac.uk) (R. Chassagne), [caranha@cs.tsukuba.ac.jp](mailto:caranha@cs.tsukuba.ac.jp) (C. Aranha).

computing time, to incorporate as much data as we can?

To better understand these issues, we perform extensive comparative experiments with the goal to better assess the mechanisms triggered by the design of the objective function in the seismic history matching procedure. In these experiments, we use a framework based on Auto-adaptive Evolutionary Computation [20] on three different real data sets to explore the following questions: (1) How different image analysis methods compare to each other, (2) How the amount of data used affects the update of the model and (3) the influence of the complexity of the dataset itself. To guarantee the relevance of these tests, we perform these experiments within a SHM loop with large real world data sets, making the whole exercise very costly in terms of computational time.

## 2. Set up and background

In this work, we investigate different similarity and dissimilarity metrics and different attribute combinations as objective functions for the SHM problem. These objective functions are applied on a self-adaptive Differential Evolution algorithm in a Seismic History Matching framework. In this section we describe the details of each procedure and the algorithm used.

### 2.1. The seismic history matching challenge

The problem of Seismic History Matching (SHM) consists of using data from production wells' history and from seismic volumes or maps to update the parameters of a simulation model describing the subsurface fluids flow. A successful optimization will make the model more reliable for future forecasting, allowing it to be used for decision-making.

The seismic data is obtained from surveys, and is subject to various processing and interpretation steps [21–23]. This results in a map that will be used in the objective function, where it is compared to an equivalent map generated by the simulation model, and both are employed to update the model [24].

In this work, we made the choice of using a simplified history matching workflow and setting in order to study strictly the impact of multiple data (attributes) and the influence the similarity metric on the history matching result. This avoids the results of the experiments to be unnecessary biased by data processing, physical modeling and expensive procedures to achieve them. Indeed, to history match we need to transform seismic attributes into the simulation domain or the other way around. For instance, the SHM can be set up in different domains of the above process: Observed Seismic, Impedance, or Simulation domain (pressure and saturation of fluids). The choice of the comparison domain should depend on the smallest encompassed uncertainty of the data, but this is still an open problem [25] and a matter of debate among geoscientists.

Our workflow uses synthetic seismic maps, which are created directly from a simulation model of real cases (existing fields). We extract pressure and saturation information directly from the cell's attributes in the model, forming a 3D synthetic seismic map. These maps are averaged over the depth axis, we apply a min/max linear scaling to the [0,1] range. The optimization procedure operates on an initial random ensemble, and performs the history matching using only the synthetic seismic model from the reference model. Finally, to compare the different formulations of the objective functions, the final ensemble obtained from SHM using different attributes and metrics are evaluated by comparing the wells' production data (Fig. 2). A summary of this workflow is shown in Fig. 1.

In this workflow, a critical question is how much seismic data to use in the optimization procedure to reach a satisfactory update of the model. A common belief is that "the more data the better". However, some of the data might be not reliable or redundant, and simply slow down the convergence of the optimization. Therefore in this study we

want to question this statement, and analyze the influence of the amount of data used in the process of SHM. Following the choice that we made for performing SHM in the simulation domain, we have direct access to three attributes, namely *Pressure* (P), *water saturation* (W) and *gas saturation* (G) (three phase flow), as well as their different combinations. By changing the number of attributes available to the optimizer, we vary the amount of data inserted into the SHM procedure, and observe how this change influences the update of the models.

Another critical question for the SHM is how to assemble the objective function used for evaluation of the similarities/dissimilarities between seismic maps. The standard formulation for the SHM used in the industry is to use Least Squares (LS) [26], which relies on a point-by-point analysis between the reference and model-generated maps, as the similarity metric, but other formulations are also possible [27]. A correct evaluation of how similar two maps are is a very important point to drive the optimization algorithm towards a minimum misfit and consequently to obtain the best updated models. Therefore, in this study we compare four different formulations of this objective function (including the LS), which are described in Section 2.3.

### 2.2. Differential evolution for history matching

In order to examine the effects of different attributes and similarity measures on the SHM, we use the *Success History Adaptive Differential Evolution* (SHADE [28]) as the optimization method. SHADE is a state-of-art variant of Differential Evolution (DE) [29]. Recently, we have shown that SHADE is effective at updating models in traditional HM [20,30].

In SHADE, we define the initial model ensemble as the *Solution Set*  $X = \{x_1, x_2, \dots, x_N\}$ , where  $N$  is the size of the ensemble. Each  $x_i \in X$  is a real valued vector  $x_i = \{x_{i1}, x_{i2}, \dots, x_{iD}\}$ , where  $D$  is the number of model parameters, and  $x_{ij}$  is a value between 0 and 1.

At each iteration  $t$  of the optimization, new solutions are generated according to the following rules: a mutant vector  $v_{i,t}$  is generated from an existing solution  $x_{i,t}$  by applying Eq. (1):

$$v_{i,t} = x_{r_1,t} + F \cdot (x_{r_2,t} - x_{r_3,t}). \quad (1)$$

The indices  $r_1, r_2, r_3$  are randomly selected from  $[1, N]$  such that they differ from each other as well as  $i$ . The parameter  $F \in [0, 1]$  controls the magnitude of change in the mutant vector. After generating  $v_{i,t}$ , it is crossed with the original solution  $x_{i,t}$  in order to generate a trial vector  $u_{i,t}$  using the "Binomial Crossover" strategy, described in Eq. (2):

$$u_{j,i,G} = \begin{cases} v_{j,i,G} & \text{if } \text{rand}[0, 1] \leq CR \text{ or } j = j_{rand}, \\ x_{j,i,G} & \text{otherwise.} \end{cases} \quad (2)$$

Where  $\text{rand}[0, 1]$  denotes a uniformly selected random number and  $j_{rand}$  is a decision variable index uniformly selected from  $[1, D]$ .  $CR \in [0, 1]$  is the crossover rate. After all of the trial vectors have been generated, a model is generated and simulated for each trial vector. The result of the simulation generated by the trial vector  $u_{i,t}$  is compared with the model generated by the corresponding original vector  $x_{i,t}$ , using the objective function described in Section 2.3. Then the best model is kept in the solution set  $X$ .

The values for the CR and F parameters are adjusted in an adaptive manner by SHADE's scheme of storing former parameter values in a historical memory, and sampling new values near the historically successful ones.

The hyper-parameters of SHADE used in this work were selected using the Sequential Model-based Algorithm Configuration method (SMAC [31]). A suite of 30 optimization benchmark functions from the CEC2014 benchmark set was used in this process. The hyper-parameters used are: Population Size 8, Initial DE mutation F: 0.856, Initial DE crossover CR: 0.244, selection strategy: current-to-pbest/1, p: 0.017, archive rate: 0.996, memory size: 9. More details about the choice of hyper-parameters, including a discussion of the sensitivity of these

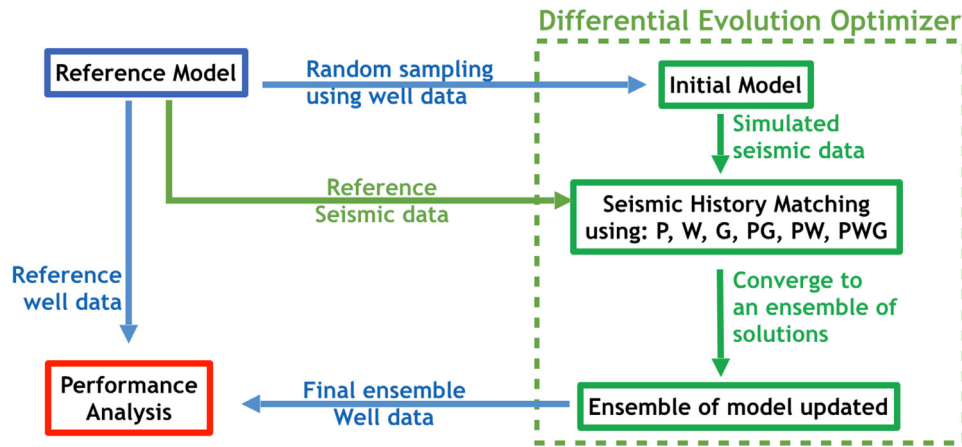


Fig. 1. Workflow of the seismic history matching method presented and used in this work.

parameters on the history matching problem can be found in Aranha et al. [20]

2.3. Objective functions, metrics and attributes

Within seismic history matching, an objective function is defined through the number of attributes used and the metric utilized to measure similarity/ dissimilarity of the seismic maps. The optimization algorithm generates sets of model parameter vectors  $x_i$ , which are compared with the reference model parameter vector  $ref$  following the objective function  $F(x_i)$  described in Eq. (3).

$$F(x_i) = similarity \left( \sum^{attribs} map(x_i), \sum^{attribs} map(ref) \right). \tag{3}$$

In this equation,  $map(param)$  is a function that generates the seismic map of a certain attribute for a given set of model parameters  $param$ . So the optimization process compares the seismic map generated by the individual solution  $x_i$ , with the seismic map generated by the reference parameter ( $ref$ ).

$Attribs$  is the subset of attributes chosen from (P, W, G, PW, PG, WG, PWG), as described in Section 2.1. The seismic map for each attribute is scaled to a [0,1] range and added together, so that the attributes have the same weight.

Finally,  $similarity$  is the similarity function used to compare the seismic maps. The choice of similarity function is another question that

we explore in this study. We choose four well established image similarity metrics [32] described in the formulations below. For each formulation,  $P = \{p_1, p_2, \dots, p_N\}$  is a seismic map obtained from the simulation of a solution  $x_i$ , and  $R = \{r_1, r_2, \dots, r_N\}$  is the seismic map obtained by using the reference model,  $ref$ .

**Least Squares (LS):** This metric is the way traditionally used to calculate the misfit in both the production well and seismic cases. The misfit between two maps is calculated as

$$LS(P, R) = 1/M \sum_{y=0}^M (p_y - r_y)^2. \tag{4}$$

where  $M$  is the total number of pixels.

**Pearson Correlation (PC):** This metric takes into account not only the difference between individual cells in the seismic map, but also their relative distance from the total distribution of values. The misfit is calculated as

$$PC(P, R) = 1 - \sum_{y=0}^M ((p_y - \mu_P)/\sigma_P) * ((r_y - \mu_R)/\sigma_R), \tag{5}$$

where  $\mu(k)$  and  $\sigma(k)$  are the mean and standard deviation of a seismic map, respectively.

**Kendall Tau (KT):** This metric uses the number of concordant and discordant cells in the two maps. Let  $i, j, i \neq j$  be two indexes of a

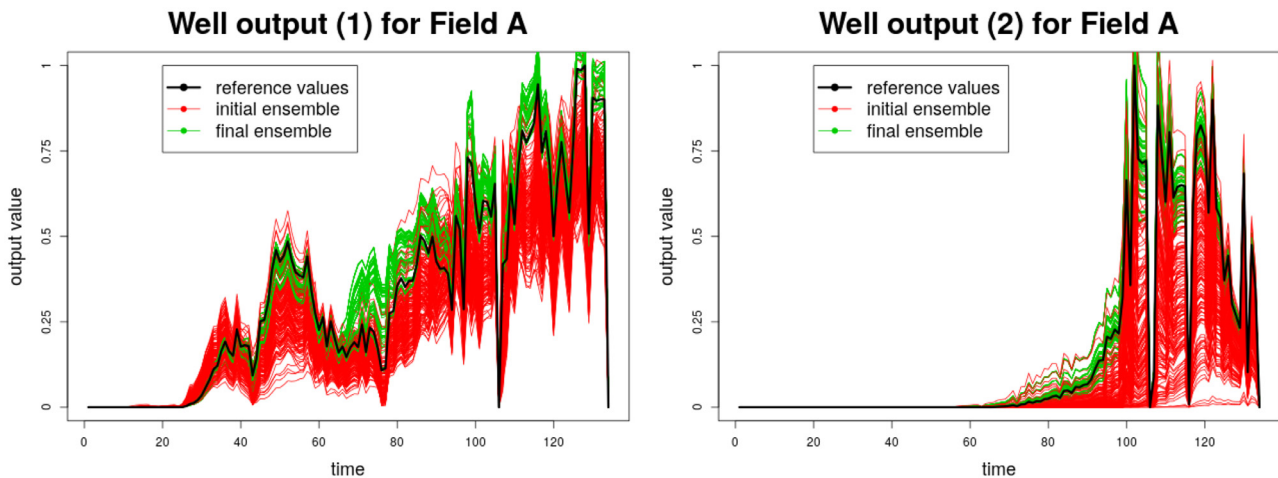


Fig. 2. Traditional HM performance evaluation is using well production data. Left and right figures are two examples of production well outputs for field A. The initial and final (optimized) set of models are represented by the red and green lines, respectively, while the reference data is represented by the black line. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

seismic map with size  $M$ .  $i$  and  $j$  are a concordant pair if  $sign(p_i - p_j) = sign(r_i - r_j)$ , and a discordant pair otherwise.  $N_c$  is the number of concordant pairs, and  $N_d$  is the number of discordant pairs ( $N_c + N_d = M*(M - 1)/2$ ) The Kendall Tau similarity is then given by:

$$KT(P, R) = 1 - |(N_c - N_d)/(M*(M - 1)/2)| \tag{6}$$

**Minimum Ratio (MR):** This metrics focuses on the ratio between each pair of cells in both maps. It is similar to the LS in that the value of the ratio at each cell is independent from the values in the rest of the seismic image. It is given by

$$MR(P, R) = 1 - \frac{1}{M} \sum_{y=0}^M \min(p_y/r_y, r_y/p_y). \tag{7}$$

In the equation above, if either  $p_y$  or  $r_y$  is 0, then the result of the minimum function is 0. If both are 0, then the result of the minimum function is 1.

It is worth noting the comparative characteristics of the metrics above. MR and LS are local, so results from a cell are independent of the value of others cells, while PC and KT are global, meaning that the result of a cell depends on quantities from the entire map. Also, LS and PC depend on the raw difference of values between the maps, while KT and MR are based on more indirect comparisons.

### 3. Experiments and results

#### 3.1. Experiment design and methodology

In this study the effects of the choice of *seismic attributes* and *comparison metrics* within the SHM loop are analyzed. Following the methodology described in Fig. 1, the SHM process for different combinations of attributes and metrics is performed, and the final ensemble of models is compared against the original reference model. We use the SHADE optimization algorithm and its hyperparameters, described in more details in Section 2.2. The stopping criteria is 2000 evaluations for the Teal South and Field A, and 1000 evaluations for field B. The low number of evaluations reflects the complexity of the simulations and the number of repetitions required by the analysis, described below.

In these experiments, we use three real world datasets, namely, Teal South [33], Field A and Field B. Field A consists of stacked reservoirs made up of the tertiary age turbidites, with respect to an optimisation complexity one can have a look at Fursov [34], which is a very similar simulation model. Field B is a complex reservoir comprising heavily compartmentalised with faults cross-cutting the turbidite sand depositional axes, a very similar field (in terms of complexity) is described in Obdegwu [8]. The number of adjustable parameters for the HM process in each field is, respectively, 13, 52 and 8. These parameters have been selected after a sensitivity analysis, which used the same approach as in [8] (Latin hypercube and tornado plots based on well production attributes). The model parameters for Teal South are Porosity multiplier, Water Oil Contact position, Rock Compressibility, Permeability values in the X direction (5 layers) and Permeability values in the Z direction (5 layers) for Teal South. The parameters for Field A are Permeability Multiplier per layer (35 layers), Permeability values in the Z direction, Pore volume Multiplier per layer (8 layers) and Transmissibility Multiplier per layer (8 layers). The parameters for Field B are Net to Gross Value, permeability in the X, Y and Z directions, transmissibility in the X, Y and Z directions, and Porosity Multiplier (showed later in Figs. 7–10).

By their very essence, these datasets (models and well production data) are highly difficult to work with. Some of these difficulties include the representation of the fluid flow interacting with the rocks and the necessary mathematical framework needed to solve the flow equations

based on mass balance and Darcy’s law. Moreover, the level of grid details necessary to describe the complex geology can become quite high. For this reason, we upscaled (rebuilding the grid structure to a coarser mesh and flow based upscaling) all models [35] in order to save some computational time and make our study not too expensive, we applied the same procedure as in [8]. Nevertheless, one run of the simulation model takes from three minutes on average for the Teal South dataset to one hour on average for the Field A dataset. This means that one complete optimization exercise can take from one to six weeks. Additionally these three datasets are not equally complex in regards to the physical processes encompassed (geology, rock mechanisms, multiphase flow), Teal south is simpler as compared to Field A and Field B.

Overall we performed a total of 84 complete optimization exercises, including three data sets, seven different attributes, four selected metrics and 20 repetitions per combination for the statistical analysis. This took us several months of calculation and analysis of the produced data. To reduce this computational burden, a pre-analysis of the results concerning only the Teal South field was conducted, where we observed that the performance of the metrics KT and MR was much lower compared to LS and PC (see Section 3.2). Accordingly, in the more costly Field A and Field B we do not perform the complete analysis of these less promising metrics, and focus only on the higher performing ones.

From the optimization exercises, we extract the following data: Both the mean seismic misfit of the final ensemble model, and the well production misfit; the proportion of model parameter values found on the final ensemble and the differential seismic maps between a model ensemble and the reference seismic data. In the following subsections, we describe and discuss the implications of each set of results.

#### 3.2. Comparison of metrics

A well performing metric will be able to efficiently measure differences and then drive better the optimization process towards better candidates, as compared to other metrics less performing. Therefore, a good convergence reflects the ability of a metric to perform well, to better estimate similarity/dissimilarity between patterns. Overall smallest misfit means better metric. In the case of Teal South, according to Table 1 the lowest final mean misfit is achieved using the Least Squares (LS) and Pearson Correlation (PC) metrics. In both metrics we obtained a final mean misfit of 1.41. Even across different attributes these two metrics are very comparable in terms of final mean misfit value and standard deviation (See Figs. 3 and 4). Minimum Ratio (MR) is slightly under performing as compared to LS and PC but still within the same range of amplitude. On the other hand, Kendall Tau (KT) is definitely not performing well, for many attributes the final mean misfit is one order of magnitude above the other metrics.

In the case of Field A, according to Table 2 the lowest final mean misfit is achieved using the Pearson Correlation (PC), but overall LS seems to narrow down the uncertainty of the final realizations as compared to PC. As noted in the case of Teal South, here too LS and PC deliver rather comparable results, in the same order of magnitude (see

**Table 1**  
Final mean and deviation of misfits for the Teal South field, for the different methods and attributes under consideration. Blue and Red highlights the highest and lowest value of the misfit, respectively.

Method	P	W	G	PG	PW	WG	PWG
LS	27.41	3.57	9.94	7.23	3.57	<b>1.41</b>	1.55
(StDev)	10.42	4.88	11.02	8.32	3.15	3.17	1.95
PC	7.35	3.22	8.15	8.02	2.96	<b>1.41</b>	1.85
(StDev)	4.91	2.07	9.91	11.76	2.52	1.59	2.29
KT	37.76	<b>17.04</b>	31.97	37.95	48.74	21.49	27.65
(StDev)	19.44	11.17	14.56	19.14	30.76	12.82	12.12
MR	32.41	<b>1.89</b>	43.39	9.34	1.98	3.92	3.49
(StDev)	11.16	1.75	25.47	7.18	2.02	4.92	5.17



### Misfit comparison of Methods (Teal South)

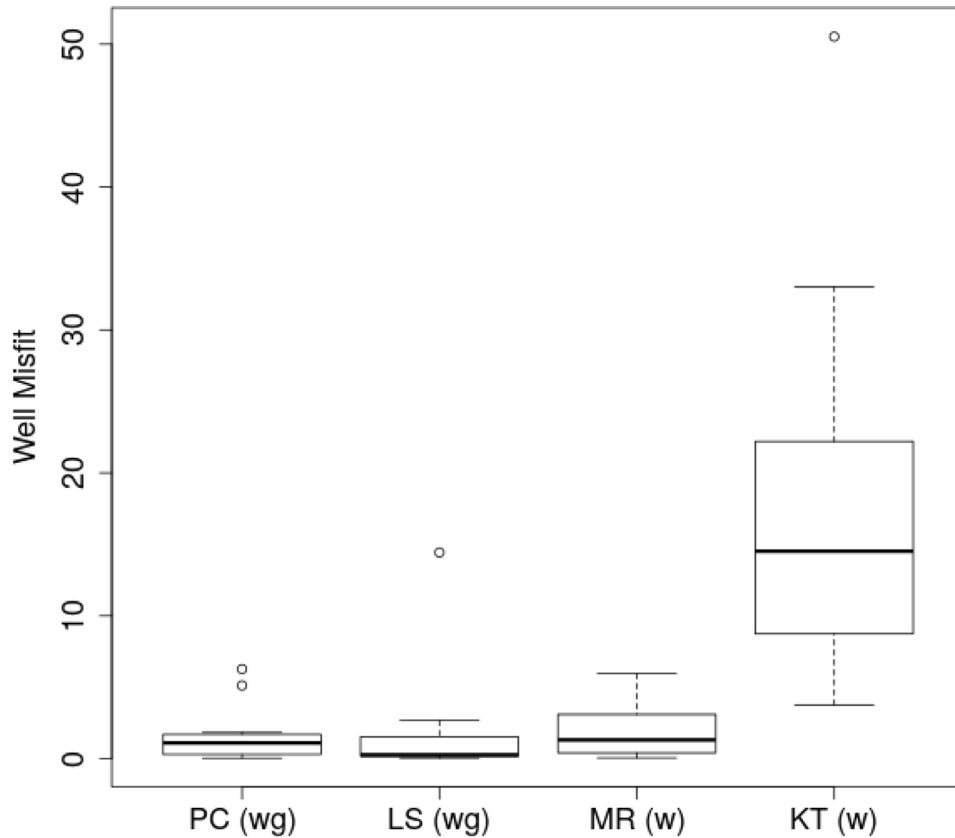


Fig. 3. Misfit values after seismic history matching for the Teal South field, using the four chosen metrics (PC, LS, MR, KT).

Fig. 4).

In the case of field B, according to Table 3 the lowest final mean misfit is obtained with Pearson Correlation. In that case the results are not comparable across these two metrics, but not drastically different either; as the uncertainty in the final ensemble of realizations seems to be larger with LS than PC, larger mean and standard deviation.

Among these experiments the LS metric, despite its simplicity, seems to deliver satisfactory results overall as compared to more sophisticated metrics. Taking into account its simplicity to implement, it

could be considered a fair and reliable choice in a context of simple patterns (see Figs. 11–13) to identify.

#### 3.3. Comparison of attributes

Taking a first look at the results while going across the experiments (metrics and attributes, see Tables 1–3) we note that the best realization, the one with the lowest misfit, is systematically the one considering three attributes (except with the KT metric where the attribute

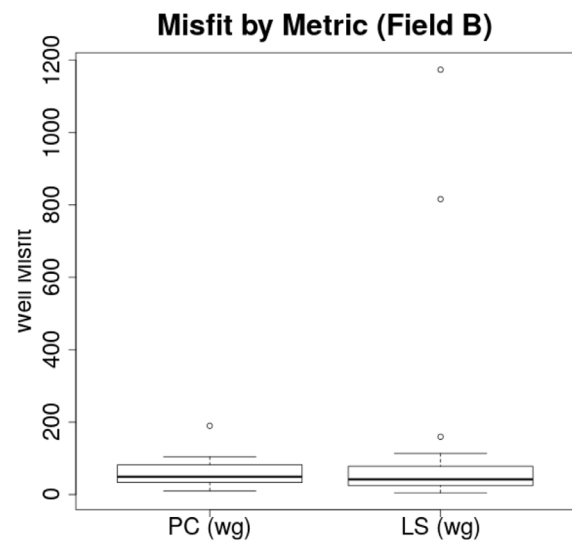
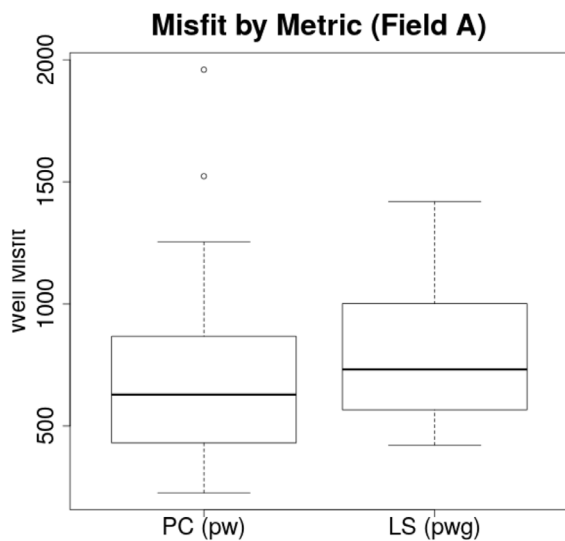


Fig. 4. Misfit values after seismic history matching for field A (left) and field B (right), using two metrics, PC and LS.

**Table 2**

Final mean and deviation of misfits for Field A, for the different methods and attributes under consideration. Blue and Red highlights the highest and lowest value of the misfit, respectively.

Method	P	W	G	PG	PW	WG	PWG
LS	1394.09	1040.61	1004.26	993.30	<b>773.31</b>	823.41	788.30
(StDev)	386.46	340.55	330.92	438.93	272.99	284.92	425.28
PC	1430.13	993.06	1112.45	891.63	836.79	795.18	<b>744.87</b>
(StDev)	432.59	276.98	386.56	532.99	290.91	390.57	439.65

**Table 3**

Final mean and deviation of misfits for Field B, for the different methods and attributes under consideration. Blue and Red highlights the highest and lowest value of the misfit, respectively.

Method	P	W	G	PG	PW	WG	PWG
LC	1110.16	399.84	439.65	530.46	255.27	<b>142.27</b>	258.39
(StDev)	635.53	469.58	678.04	755.08	393.11	299.62	408.15
PC	831.69	342.32	104.42	163.16	156.89	<b>58.22</b>	258.39
(StDev)	667.72	448.59	169.21	313.18	287.32	41.79	385.95

W gives the lowest misfit. We will come back to this point in the rest of the study). As we are dealing with stochastic methods and consequently multiple realizations, a more comprehensive comparison should look at combination of the mean of the misfits given the corresponding standard deviation. The mean misfits and their standard deviations for different attributes across different metrics are shown in Table 1 (For Teal South), Table 2 (For Field A) and Table 3 (For Field B).

The comparison of the lowest and highest mean misfit for each case leads the reader to notice that using more data (all three attributes at once, PWG) does not imply a better misfit, on average.

To investigate the validity of this observation, we perform the following statistical analysis: first we conduct an Analysis of Variance (ANOVA), grouped by metric, to test whether the attribute selection has a statistically significant influence on the mean misfit. After this, we prepare a 95% confidence interval of the difference between the misfit using all attributes (PWG) and every other combination of attributes, using Tukey's HSD correction for multiple comparisons [36]. This second test indicates, what sets of attributes (if any) show statistically significant different mean misfits from using all data.

The ANOVA test showed that, in all cases, some selection of attributes had a clear influence in the mean misfit (F statistic over 1000 in each case). An examination of the confidence intervals let us know which attributes are responsible for this difference. In the case of a simple dataset like Teal South (described in Section 3.1), the Fig. 5 indicates that on average, lowest misfits can be reached while using only one or two attributes. When it comes to analyze the results with more challenging real datasets (Fig. 6), the results are even clearer in this regard: Field A can achieve similar results to PWG with only two attributes, and in Field B even one attribute can be enough.

### 3.4. Model parameter distribution

As an inverse problem, the goal of the HM procedure is to update the input parameters of the initial set of models, proposing at the end of the run a new range of variation for the selected input parameters. In this part these three figures: Figs. 7–10 will be used to support this study. Also, it is to be noted that we draw a black line on each of these plots to place the "true" synthetic answer, which helps identify the correctness of the history matching result.

Fig. 7 shows the initial (red) and final (blue) distributions for a model parameter (*poro*) for different combinations of similarity metric (rows) and seismic attributes used (columns). Whatever the attributes used, there is a clear convergence of the final range towards the value 0.31. This convergence is especially strong for the LS, PC and MR

metrics. Nevertheless, when using only attribute P (leftmost column of Fig. 7), the result is a somewhat wider spread of values between 0.29 and 0.33. This means that the convergence was not yet completed, when compared to the results obtained with other seismic attribute combinations, for which the range of values are more narrow. To some extent, we also see this lack of convergence in the case of using the seismic parameter G in isolation (third column from the left in the same Figure).

A similar trend can be observed for the parameter P5, shown in Fig. 8. In this case, a better convergence is also observed when more seismic attributes are used. Still, while using only one attribute contains indeed more uncertainty, it also can lead to a reasonable updated range for the input parameter, in the case of attribute W (second column from the left in Fig. 8). It is clear that the seismic attributes have to be selected wisely, otherwise convergence is not possible, for example compare columns 1 and 3 from the left in Fig. 8, which correspond to attributes P and G.

These two figures also give a very clear image of the comparative performance among the similarity metrics. The KT metric, represented in the bottom row in both Figs. 7 and 8 produced mixed results compared to LS for the first Figure, and in the second Figure is not able to decide for a final range of the parameter at all, whatever the number of attributes used. This could be explained by the fact that KT does not use absolute values only whether the value changes or not. For problem where values do not matter it is not an issue, but for this study, where it matters, it is preferable to use absolute value, like LS does. As for the MR metric, in Fig. 7 it obtains similar results as LS, except that it fails to reduce the uncertainty significantly when using only the P and G attributes, succeeding only with W. For the P5 parameter (Fig. 8), MR is not providing a robust convergence for most attribute combinations.

For the PC metric (second row in both figures), the final ranges of the two input parameters are very comparable with the results obtained by LS. Overall, the LS and PC metrics seem to produce the most reliable results, by reducing the initial uncertainty of the input parameters to a satisfactory new range. Despite using only one attribute, both metrics appear to converge properly as compared to the other metrics. This point will be discussed further in the Discussion section.

Extending the same analysis to the cases of Field A and Field B (Figs. 9 and 10), a first general comment would be that because of the added complexity of these fields, we do not see such a drastic reduction of the uncertainty in the final range of the selected input parameters as compared with the Teal South case. The parameters selected for inclusion in these figures associated with the PC and LS metrics for fields A and B offer a fair illustration of what it looks like to deal with more physical complexity, and how difficult the optimization problem becomes.

Fig. 9 shows a comprehensive illustration of the results of using additional information in a HM process. Rows a) and b) show the distribution of values for parameter M21, using LS and PC as similarity measures. Rows c) and d) shows the same information, for the model parameter M37.

In the bottom two rows (Parameter M37) we see that even only one seismic attribute, if chosen correctly, will deliver a satisfactory reduction of the initial uncertainty. In this case, attribute G (third column) shows this property. For two attributes, the best result is obtained in the case where G is associated, namely WG and PG (columns five and six, respectively). Finally, using all three attributes (PWG) did not show any significant enhancement compared with the two-attribute cases.

In the top two rows (Parameter M21) shows an inconclusive result for both PC and LS. Usually when the values are widely spread in the final model ensemble, it indicates that the considered input parameter is not sensitive. Indeed, using one or two seismic attributes does not change the final spread of the value distribution. Nevertheless, LS seems to narrow down the suggested range very slightly in comparison with PC when using all three seismic attributes.

We observe similar patterns when looking at the model parameters

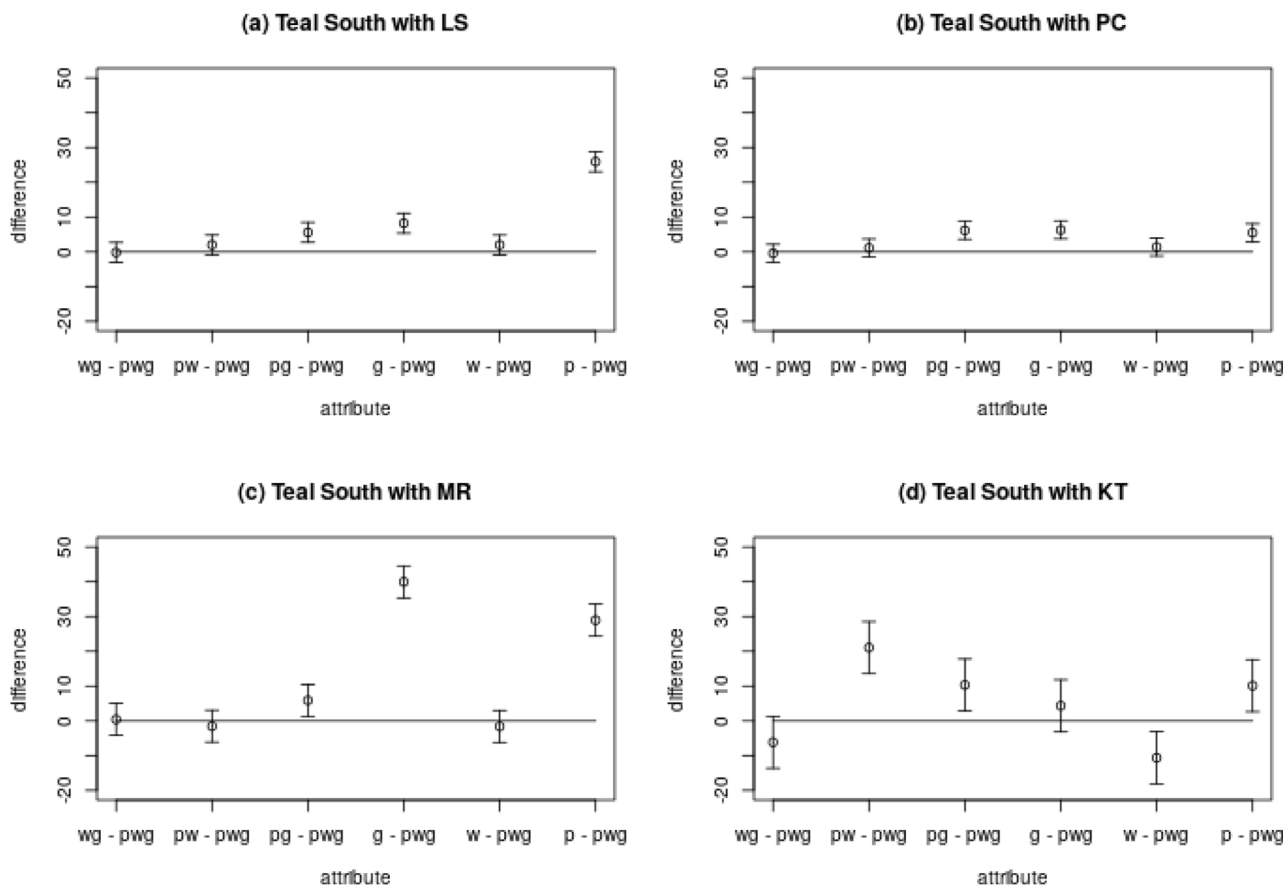


Fig. 5. Mean Misfit Difference (MMD) between using PWG as the seismic attribute set, and every other combination of attribute (Teal South Field). The Y axis is the MMD, and the X axis is each attribute combination. The bars are 95% confidence interval of the MMD with a Tukey’s HSD correction for multiple comparisons. Attribute sets whose bars are above the horizontal line are worse than PWG, sets below the line are better, and sets that touch the line are inconclusive.

for Field B (Fig. 10). For the parameter Y in rows a) and b), different seismic attributes deliver a similar updates model parameter range, at around 0 and 3.2. From these results, we see that using only one seismic attribute is an option to envisage. Looking at the second input parameter for the same field (poro, rows c) and d)), we can formulate roughly the same observations. The PC metric leads to a wrong convergence when using attributes WG as compared to the LS (column 6), which converges to the correct final range when compared to the range reached by other attribute cases (between 0.35 and 3.14). In the other attributes, LS and PC are generally comparable metrics in the case of Field B.

In terms of computing time there is no sharp differences in using one metric or another. Moreover, counter intuitively there is no difference in the optimization procedure time involving one or more attributes. However, the time to pre-process the data for each attribute before the optimization has a non-negligible cost, including experienced engineering interpretation and processing. Therefore, there is an interest in using as few seismic attributes as possible.

### 3.5. Qualitative analysis of the misfit

Previously a quantitative analysis of the performance of SHM have been investigated on the three datasets, a final look at this problem could be from the pattern perspective as the objective function is evaluating patterns see Figs. 11–13. This section can be seen as a global qualitative comparison of the fields, although we cannot strictly speaking compare the results across fields. By comparing the differences of two maps (in LS and PWG case) at different time, before and after SHM, two main kinds of results emerge. In the case of Teal South

and field B the remaining patterns after the SHM process are very little whereas in the case of the Field A they are significant for pressure and water. Teal South is physically simple (geology and fluid flow), but Field A and B are much more complex, then, apparently the complexity/non linearity of a model is not a limit to how good a history matching can be achieved in a reasonable time. One explanation for the field A results could be that the optimization procedure has been terminated before convergence, but it is not the case here as we checked that aspect. The most probable explanation comes from the setting of the problem; the choice of the inputs parameters, the range of variation of each, could not allow the optimizer to converge with all the attributes. Also it is interesting to note that SHM on Field A is performed in a higher dimension space (number of input parameters) than for Field B. This leads to another argument in favor of a well-thought and well-parameterized SHM should be carefully prepared as compared to the use of all possible data we have access to.

## 4. Discussion

It has been noticed that using pressure only does not end up with good history matching results, for the three tested fields. A physical explanation could be that because pressure rapidly stabilizes everywhere it makes the different fits hard to conceal, whereas for the other attributes the changes are more local and thus less challenging. On the other hand, using more than one attribute does not guarantee the correct uncertainty reduction, see for instance WG as compared to PWG (Fig. 9-b) and this is apparently not dependent on the metric. The use of only one attribute has been shown to be useful as it obtains comparable – sometimes strictly as good as – results as using the three attributes.



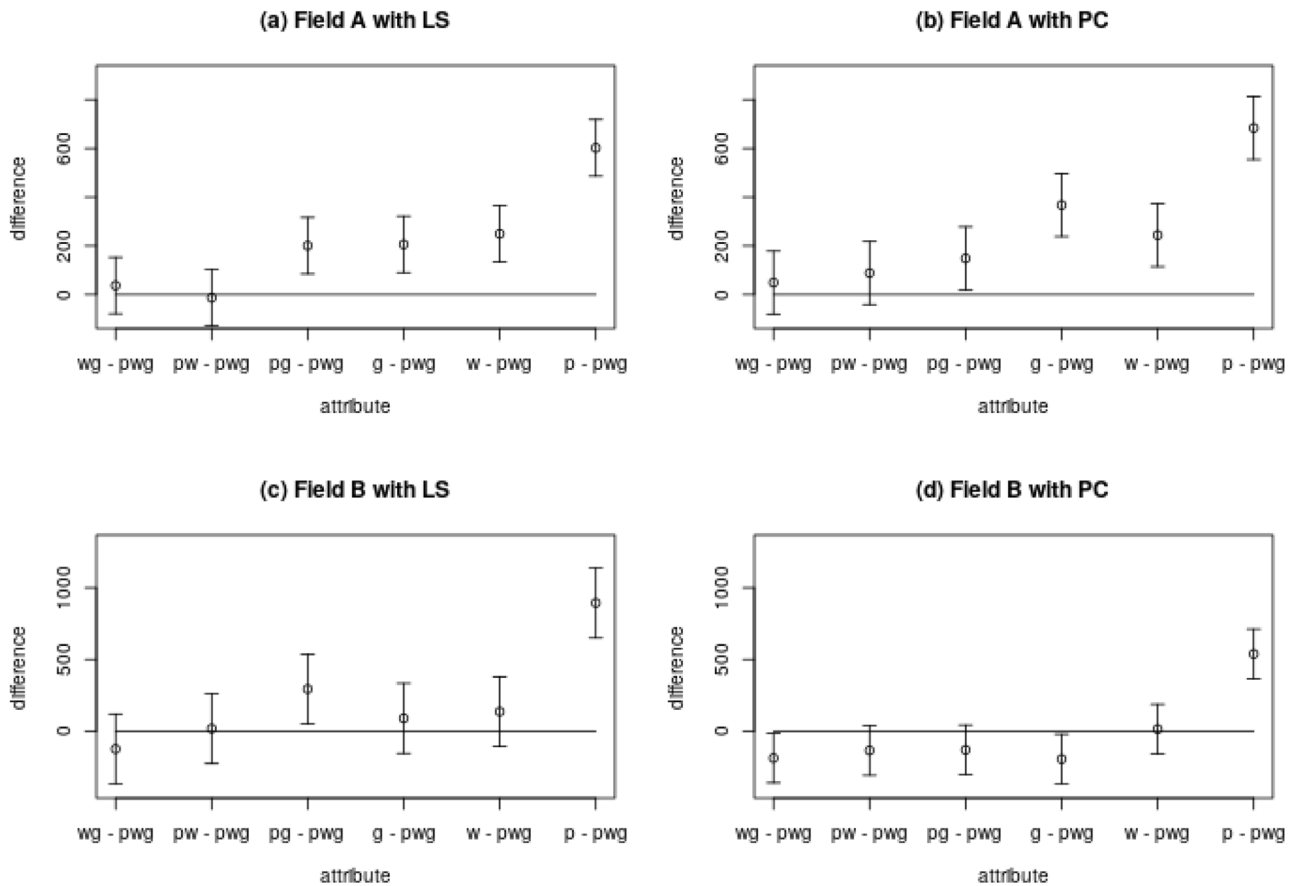


Fig. 6. Mean Misfit Difference (MMD) between using PWG as the seismic attribute set, and every other combination of attribute (Fields A and B). The Y axis is the MMD, and the X axis is each attribute combination. The bars are 95% confidence interval of the MMD with a Tukey’s HSD correction for multiple comparisons. Attribute sets whose bars are above the horizontal line are worse than PWG, sets below the line are better, and sets that touch the line are inconclusive.

Within the workflow the use of only one attribute could play the role of a proxy or a quick look for different scenario of parameters to optimize before using many attributes. By then we could save time in the processing of the data and in the setting up of the seismic history-matching problem [37].

Another point to make is by using an attribute that did not converge, will not affect the overall convergence if other converging attributes are incorporated. In such way that, adding more data is no harm to the global convergence, if at least one other useful attribute is integrated into the objective function. This leads us to the question of the choice of useful attributes and data, how to determine the convergence is something that could be achieved while looking at the physical meaning of the data. Being able to establish a clear link between the input parameters to be optimized and the attribute to use could deliver faster and more efficient SHM. Finding such a relation is out of the scope of this study as the setting of the experiments would have to be different and a new analysis to be designed.

One aspect of this study is that it has been carried out in the simulation domain, as we wanted to simplify the framework of the study and focus only on the mechanisms involved in the Objective function. While the use of real observed seismic rises many processing questions, two of them are rather interesting to point out here. The use of the least square metric, which is a point-to-point metric is overall performing very well in all the cases we considered. Nonetheless we looked at seismic maps, which are an average in depth, a point-to-point metric could be less adapted while using vertical section of the seismic data [38]. Low seismic resolution being one of the characteristics of the vertical direction, a spatial pattern correlation could be more suited. In other words if the use of the full three dimensions of the seismic data is

thought, then further experiments regarding the metrics accuracy should be carried out. The second point to make is about noise, indeed these extensive experiments have been carried out under the noise free assumption. Whereas in practice the seismic data cannot be free of noise. Nevertheless adding relevant noise to the synthetic seismic would have been itself a non trivial question among the community of geophysicists [39,40]. However, the DE family of optimization methods used in this study has been shown to be quite robust to noise [41], which leaves us with the interesting problem of the effect of noise on the metrics (LS, PC, KS and MR) as a future direction of this work.

Also to be noted is that given one metric the number of attributes used is not affecting in any manner the global computing time. In every considered case the convergence is reached in the same amount of iterations, for the three different fields and the seven different attributes.

## 5. Conclusion

In the energy industry no study has been conducted about the comparative efficiency of the objective function (attributes and metrics used) during the Seismic History Matching process. Yet it is one of the key drivers to achieve a good update of the model.

Regarding the metric used to compare two seismic images, our study shows that when it comes to choosing a similarity metric for seismic history matching, “the simpler the better”. The traditional Least Squares metric gave enough information to perform a good history matching, compared to other, more elaborate metrics (Kendall Tau, Minimum Ratio, Pearson Correlation), which gave no significant advantages on the overall SHM performance, in the context of this study.

# Distribution of parameter 'Poro' before/after optimization

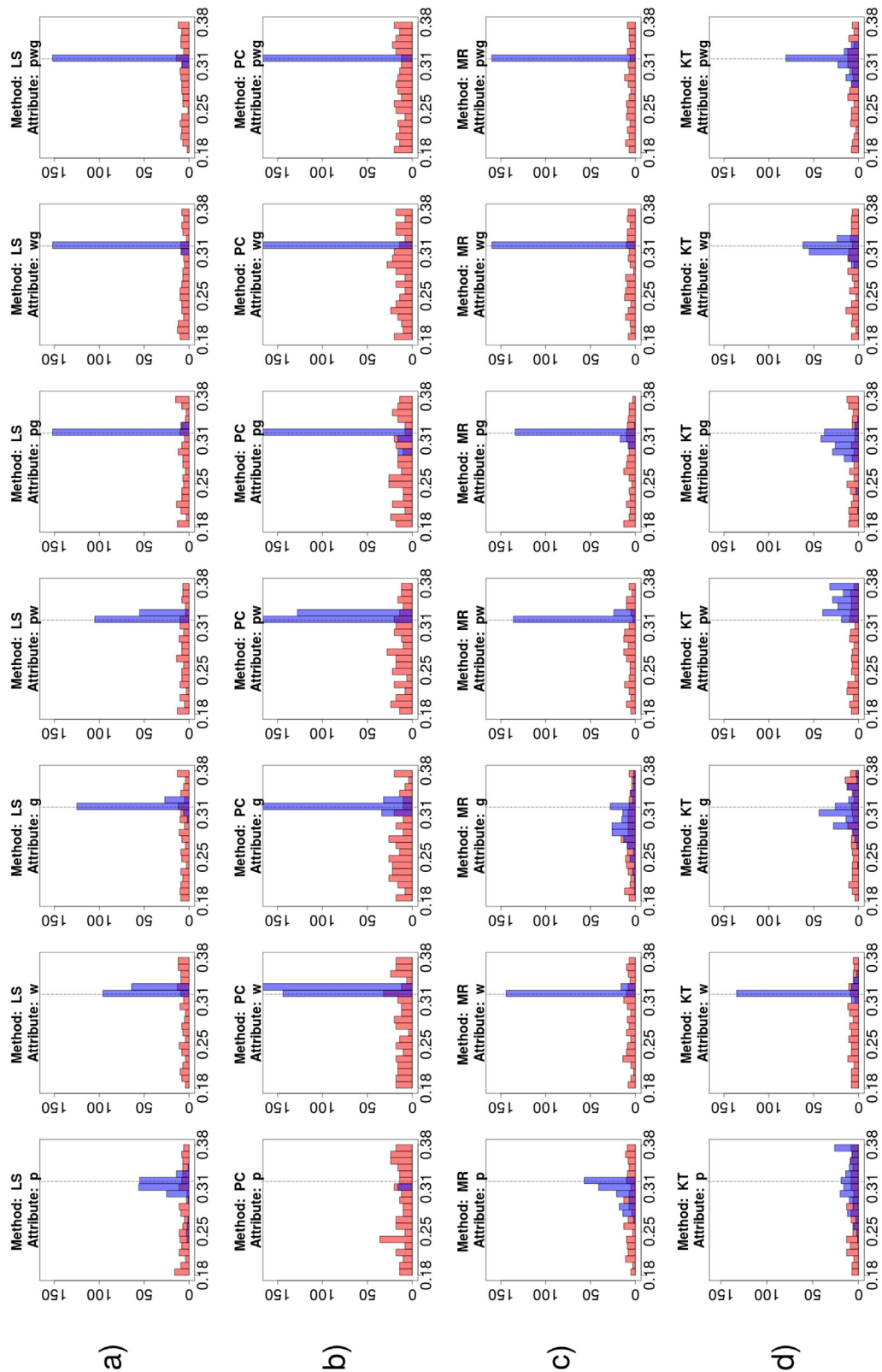


Fig. 7. Different Parameter distributions before (in red) and after (in blue) optimization, for the seven attributes applied to the Teal South field using the same parameter, *poro*, for the seven attributes and the four chosen metrics in a), b), c) and d) respectively. The dotted line indicates the true value of the parameter. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

That said, although each metric has different characteristics (such as calculating individual differences, overall ratios, sensitivity to scaling, etc), they do not explicitly consider the spatial structure of the seismic image, treating it as a numerical vector. With this in mind, it would be

interesting to explore metrics that explicitly focus on the spatial structure of the image [42], its spectral qualities [43] or its graph qualities [44].

The addition of more data to the optimization process does not

# Distribution of parameter 'P 5' before/after optimization

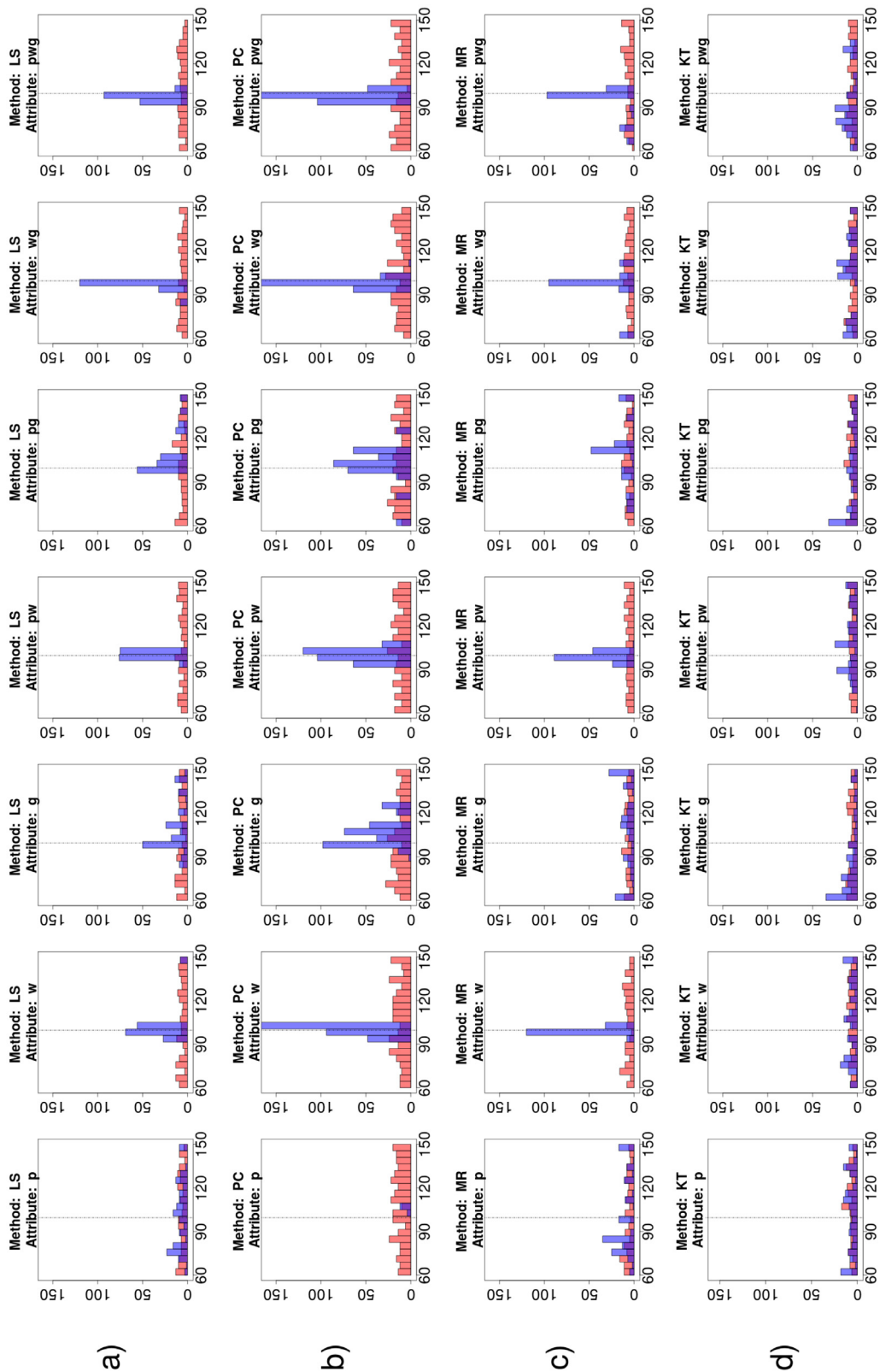


Fig. 8. Different parameter distributions before (in red) and after (in blue) optimization, for the seven attributes applied to the Teal South field using the same parameter, P5 for the seven attributes and the four chosen metrics in a), b), c) and d) respectively. The dotted line indicates the true value of the parameter. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

mean necessarily a better result but will guarantee a safer reduction of the uncertainty. We can consider that the addition of more data is in a way counter balancing a bad convergence from other attributes. The results also showed that in some cases only one attribute can lead to a

proper update the model. This aspect could put forward the redundancy in the data, which is also a data integration question and opening another area of processing this redundancy prior to history match. For that reason, a careful selection of only a few attributes/information

Distribution of parameters in Field A

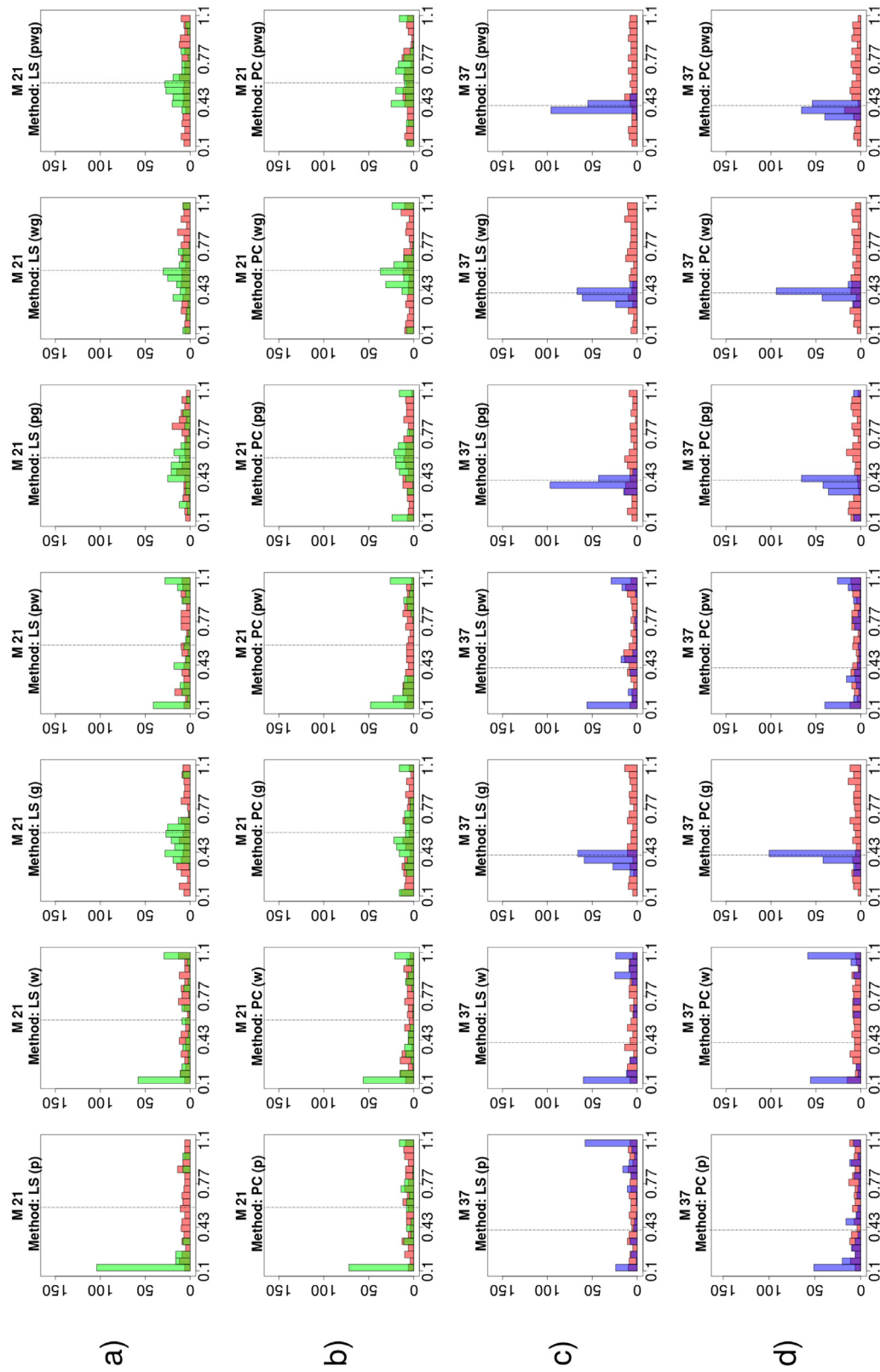


Fig. 9. Different Parameter distributions for parameter M21 and M37, before and after optimization, across metrics and seismic attributes. Parameter values before optimization are in red, and after optimization are in blue and green for Layer 21 and Layer 37 attributes respectively. The dotted line indicates the true value of the parameter. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

could be enough, and it could save time in the processing of the data and the whole workflow. How to identify what data is really useful is still to be investigated, and how to characterize seismic data usefulness would be an interesting continuation of the present study. On that note,

it has to be said also that more input parameters (issued by a sensitivity analysis) does not guarantee a successful HM, it is appearing to be better to select few meaningful model parameters.

We also found that adding more data does not change the

Distribution of parameters in Field B

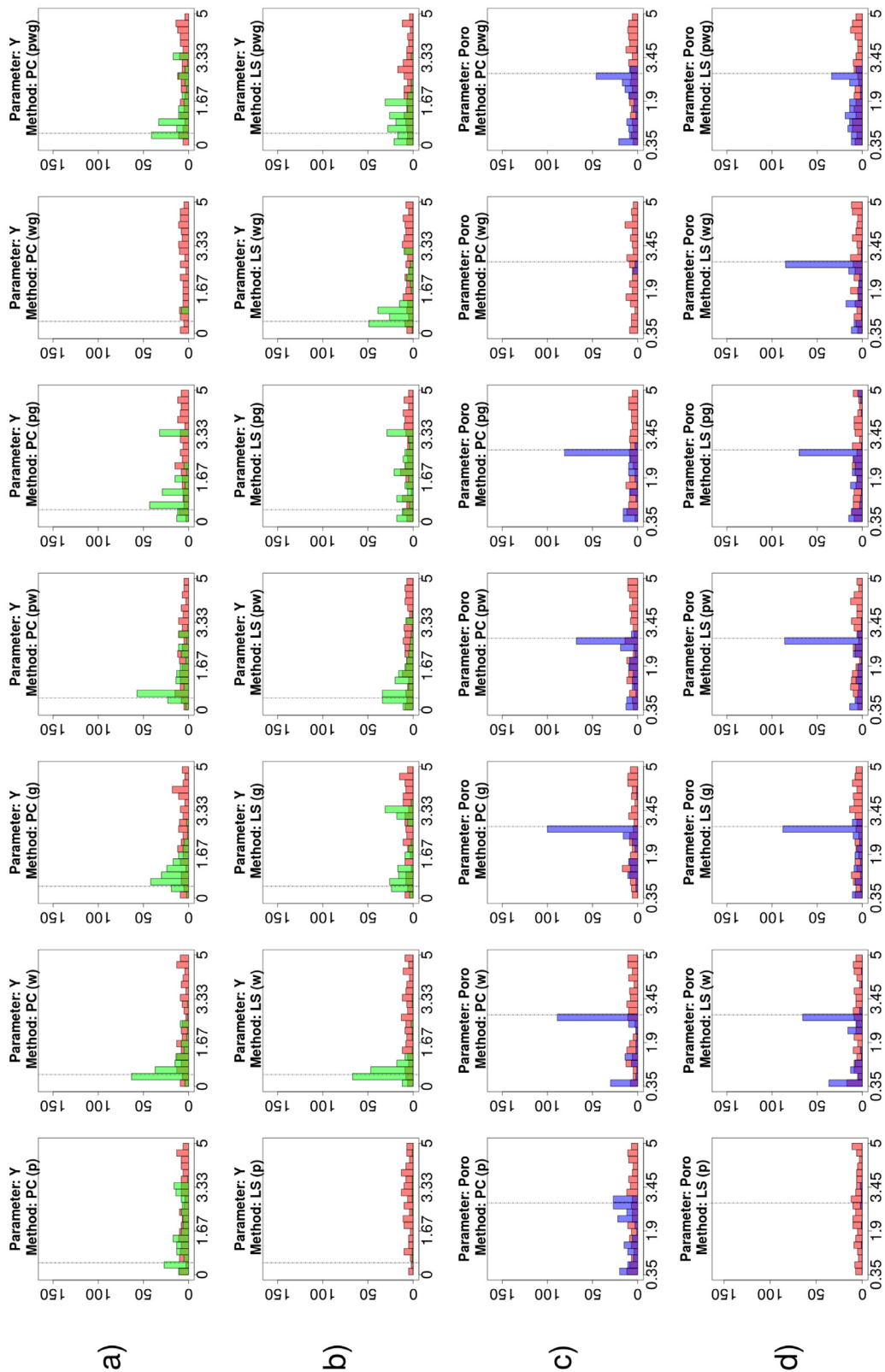


Fig. 10. Different Parameter distributions for parameter Y and *poro* before and after optimization, across metrics and seismic attributes. Parameter values before optimization are in red, and after optimization are in blue and green for Y and *poro* attributes respectively. The dotted line indicates the true value of the parameter. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

computing time given our current setting and optimization algorithm used.

In a larger context of data assimilation and black-box optimization problem, one important opening message from this study could be

about the relevance of the data to assimilate, instead of assimilating all possible data; a prior step of selection should be incorporated before to run an optimization. Such a physics-based selection in a prior step is a new research by itself, but it should lead to more specific and efficient



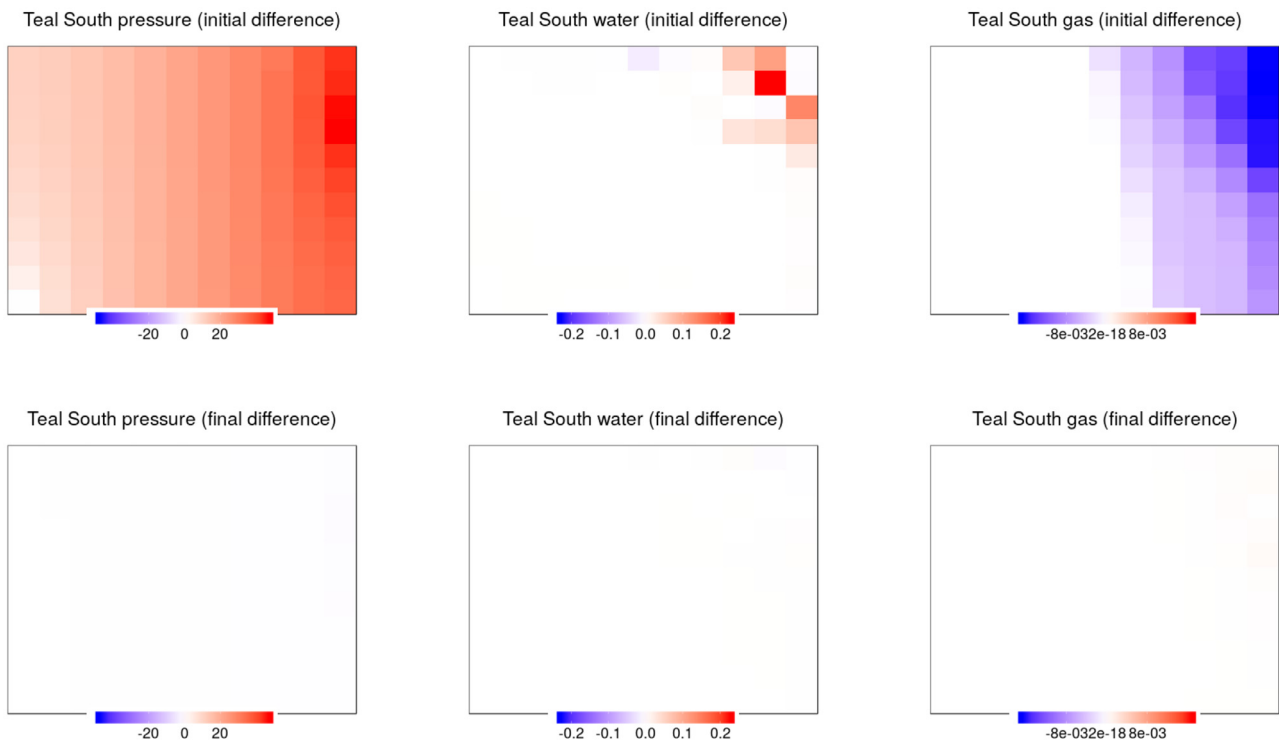


Fig. 11. Initial and Final seismic images for the Teal South field. Darker colors signify a larger difference from the reference model.

data assimilation workflow, for interest of time and computing cost.

**Declaration of Competing Interest**

The authors declare that they do not have any financial or non-financial conflict of interests.

**CRediT authorship contribution statement**

**Romain Chassagne:** Data curation, Validation, Resources, Writing - review & editing. **Claus Aranha:** Investigation, Software, Validation, Visualization, Writing - review & editing.

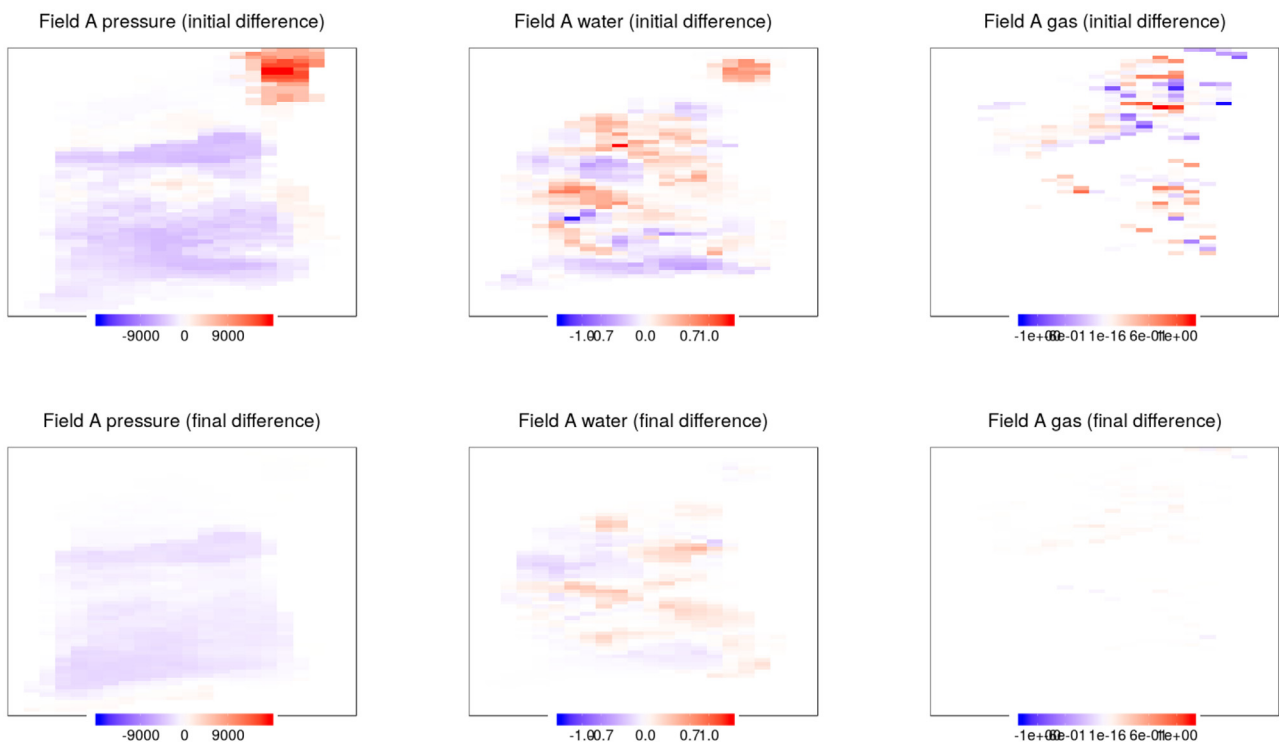


Fig. 12. Initial and Final seismic images for the Field A. Darker colors signify a larger difference from the reference model.

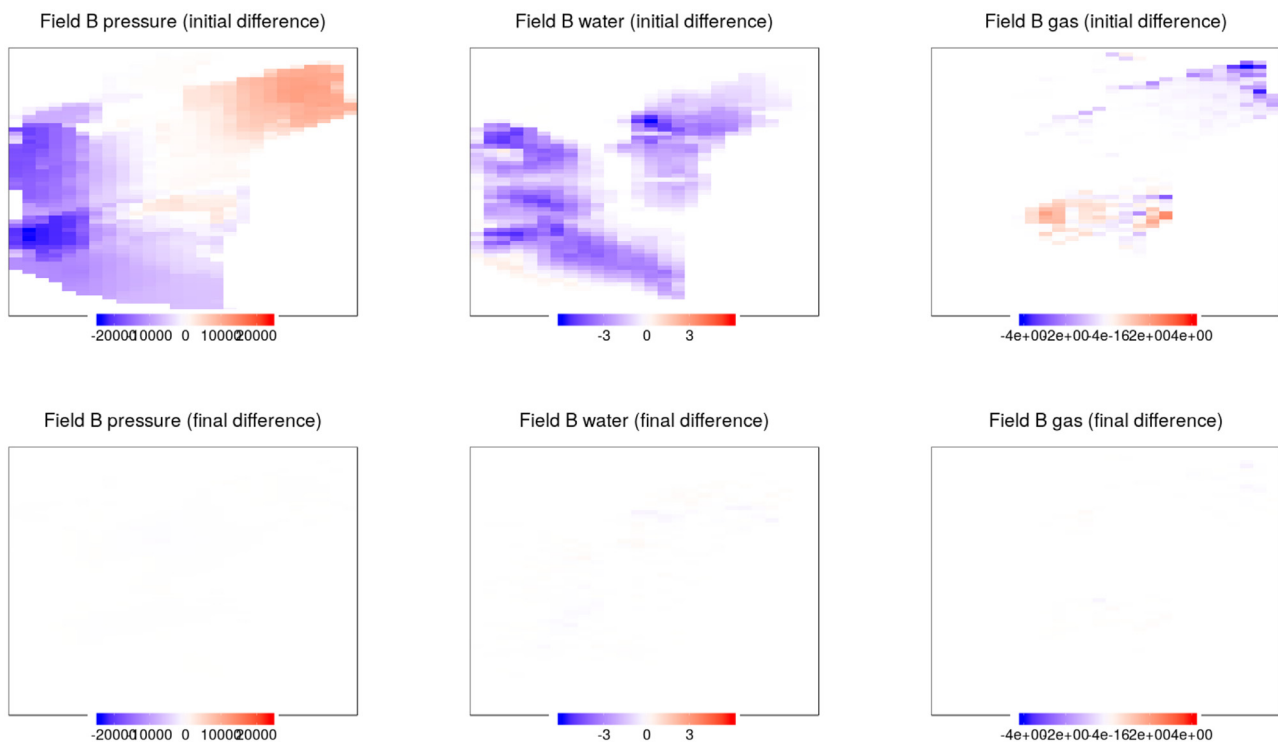


Fig. 13. Initial and Final seismic images for Field B. Darker colors signify a larger difference from the reference model.

## Acknowledgments

RC acknowledge support from the sponsors of the Edinburgh Time-Lapse Project phase VI (BG, BP, CGG, Chevron, ConocoPhillips, ENI, ExxonMobil, Hess, Halliburton, Maersk, Nexen, Norsar, OMV, Petrobras, Shell, Equinor, Taqa) for supporting this research and the Japanese Society for Promotion of Science for the fellowship L19544 during the correction of the present manuscript. We thank Schlumberger for the use of their Eclipse software.

## References

- [1] Multiphase flow in porous media. In: Adler PM, editor. Springer Science Business Media; 1995.
- [2] Ravalec ML, Veiga SD, Derfoul R, Enchéry G, Gervais V, Roggero F. Integrating data of different types and different supports into reservoir models. *Oil and Gas Science and Technology* 2012;67(5):823–39. Doi:10.2516/ogst/2012024
- [3] Ringrose P, Bentley M. Reservoir model types. 2014.
- [4] Parish RG, Calderbank VJ, Watkins AJ, Muggerridge. Effective history matching: the application of advanced software technique to the history-matching process. SPE symposium on reservoir simulation. 1993.
- [5] Oliver DS, Reynolds AC, Liu N. Inverse theory for petroleum reservoir characterization and history matching. Cambridge; 2008.
- [6] Verga F, Cancelliere M, Viberti D. Improved application of assisted history matching techniques. *Journal of Petroleum Science and Engineering* 2013;109:327–47. Doi:10.1016/j.petrol.2013.04.021
- [7] Chakra NCC, Saraf DN. History matching of petroleum reservoirs employing adaptive genetic algorithm. *Journal of Petroleum Exploration and Production Technology* 2015. Doi:10.1006/s13202-015-0216-4
- [8] Obidegwu D, Chassagne R, Macbeth C. Seismic assisted history matching using binary maps. *Journal of Natural Gas Science and Engineering* 2017;42:69–84. Doi:10.1016/j.jngse.2017.03.001
- [9] Chassagne R, Obidegwu D, Dambrine J, Macbeth C. Seismic history matching, a metric study. *Computers and Geosciences* 2016;96:159–72. Doi:10.1016/j.cageo.2016.08.013
- [10] Luo X, Bhakta T, Jakobsen M, Naevdal G. An ensemble 4d seismic history matching framework with wavelet multiresolution analysis – a 3d benchmark case study. 15th European conference on the mathematics of oil recovery. 2016.
- [11] Yin Z, MacBeth C, Chassagne R, Vazquez O. Evaluation of inter-well connectivity using well fluctuations and 4d seismic data. *Journal of Petroleum Science and Engineering* 2016;145:533–47. Doi:10.1016/j.petrol.2016.06.021
- [12] Davolio A, Schiozer DJ. Joining two probabilistic approaches – 4d seismic data and simulation models. 77th EAGE conference and exhibition. 2015.
- [13] Joosten GJ, Altintas A, Essen GV, Doren JV, Gelderblom P, van den Hoek P, et al. Reservoir model maturation and assisted history matching based on production and 4d seismic data. SPE Annual Technical Conference and Exhibition. 2014. Doi:10.2118/170604-MS
- [14] Tolstukhin E, Hu LY, Sudan HH. Geologically consistent seismic history matching workflow for ekofisk chalk reservoir. 14th European Conference on the Mathematics of Oil Recovery. 2014. Doi:10.3997/2214-4609.20141781
- [15] Lumley DE, Behrens RA. Practical issues of 4d seismic reservoir monitoring: what an engineer needs to know. *SPE Reserv Eval Eng* 1998;1(6):528–38.
- [16] Derfoul R, Veiga SD, Gout C, Guyader CL, Tillier E. Image processing tools for better incorporation of 4d seismic data into reservoir models. *Journal of Computational and Applied Mathematics* 2013;240:111–22. Doi:10.1016/j.cam.2012.08.022
- [17] Tillier RDE, Da Veiga S. Appropriate formulation of the objective function for the history matching of seismic attributes. *Comput Geosci* 2013;51:64–73. <https://doi.org/10.1016/j.cageo.2012.07.031>.
- [18] Sagitov I, Stephen KD. Assisted seismic history matching in different domains: what seismic data should we compare? EAGE annual conference and exhibition. 2012.
- [19] Gosselin O, Aanonsen S, Aavastamark I, Cominelli A, Gonard R, Kolasinski M, et al. History matching using time-lapse seismic (huts). Proceedings of the SPE annual technical conference and exhibition. Denver, Colorado, USA. 2013. p. SPE84464.
- [20] Aranha C, Tanabe R, Chassagne R, Fukunaga A. Optimization of oil reservoir models using tuned evolutionary algorithms and adaptive differential evolution. IEEE conference on evolutionary computation. 2015.
- [21] MacBeth C, Floricich M, Soldo J. Going quantitative with 4d seismic analysis. *Geophys Prospect* 2011;54(3):303–17.
- [22] Landa JL, Kumar D. Joint inversion of 4d seismic and production data. SPE annual technical conference and exhibition. 2011.
- [23] Yilmaz O. Seismic data analysis: processing, inversion, and interpretation of seismic data. Investigation in Geophysics 2. Society of Exploration Geophysicists; 2001.
- [24] Lygren M, Fagervik K, Valen TS, Hetlelid A, Berge G, Dahl GV, et al. A method for performing history matching of reservoir flow models using 4d seismic data. *Pet Geosci* 2003;9:85–90.
- [25] Roggero F, Lerat O, Ding DY, Berthet P, Bordenave C, Lefeuvre, et al. History matching of production and 4d seismic data: Application to the girassol field, offshore angola. *Oil and Gas Science and Technology* 2012;67(2):237–62. Doi:10.2516/ogst/2011148
- [26] Sedighi-Dehkordi F, Stephen K. Faster convergence in seismic history matching by dividing and conquering the unknowns. *SPE J* 2010;15(4).
- [27] Mata-Lima H. Evaluation of the objective functions to improve production history matching performance based on fluid flow behavior in reservoirs. *J Petrol Sci Eng* 2011;78(1):42–53.
- [28] Tanabe R, Fukunaga A. Success-history based parameter adaptation for differential evolution. IEEE conference on evolutionary computation. 2013. p. 71–8.
- [29] Storn R, Price K. Differential evolution – a simple and efficient adaptive scheme for global optimization over continuous spaces. Tech. Rep.. International Computer Science Institute, Berkeley, CA; 1995.
- [30] Zhang Q, Chassagne R, MacBeth C. Seismic history matching uncertainty with weighted objective functions. ECMOR XVI-16th European conference on the mathematics of oil recovery. 2018.

- [31] Hutter F, Hoos HH, Leyton-Brown K. Sequential model-based optimization for general algorithm configuration. Proc. of LION-5. 2011. p. 507–23.
- [32] Goshtasby AA. Image registration. Springer; 2012.
- [33] Abdollahzadeh A. Adaptive algorithm for history matching and uncertainty quantification. 2014. Ph.D. thesis.
- [34] Fursov I. Quantitative application of 4d seismic data for updating thin-reservoir models. 2015. Ph.D. thesis.
- [35] Preux C, Ravalec ML, Enchery G. Seleting an appropriate upscaled reservoir model based on connectivity analysis. Oil Gas Sci Technol 2016;71(5):60.
- [36] Montgomery DC. Design and analysis of experiments. eight ed Wiley; 2013.
- [37] Analysis and calibration of 4D seismic data prior to 4D seismic inversion and history matching – Norne field case. In: Alfonso CMM, Oliver DS, editors. 79th EAGE conference and exhibition, Paris. 2017.
- [38] Wu J, Mukerji T, Journel AG. Establishing spatial pattern correlation for interpreting time lapse seismic amplitudes. SEG technical program. 2004. p. 2235–8.
- [39] Insights and methods for 4D reservoir monitoring and characterization. Society of Exploration Geophysicists Distinguished Instructor Series; EAGE; 2005.
- [40] Legijt J, Jenkins G. A 4d noise model used in probabilistic seismic inversion. SEG technical program expanded abstracts. 2013. p. 3262–6.
- [41] Luukka P, Lampinen J. Differential evolution classifier in noisy settings and with interacting variables. Appl Soft Comput 2011;11(1):891–9.
- [42] Cressie N. Statistics for spatial data. J. Wiley & Sons; 1993.
- [43] Kalenova D., Toivanen P., Bochko V.. Spectral image distortion map. vol. 1. 2004, p. 668–671 Vol.1. 10.1109/ICPR.2004.1334257.
- [44] Taha AA, Hanbury A. An efficient algorithm for calculating the exact hausdorff distance. Trans Pattern Anal Mach Intell 2015;37:2153–63.