Multi-Objective Stochastic Optimization by Co-Direct Sequential Simulation for History Matching of Oil Reservoirs

João Daniel Trigo Pereira Carneiro*

under the supervision of Amílcar de Oliveira Soares†

Dep. Mines, IST, Lisbon, Portugal

December 2010

Abstract

In the present work, a modification of the stochastic History Matching method of Mata-Lima [8] is proposed.

The method is modified in order to use a multi-objective optimization based on the estimation of the Pareto front of the generated image set. The additional computational weight imposed by the multi-objective approach on the perturbative (Co-Direct Sequential Simulation) and fluid simulation stages of the process is minimized with the sequential application of image match, Multi-Dimensional Scaling, Kernel-PCA and Clustering techniques in order to reduce the set of secondary images need.

The resulting method is applied to a case study with promising results.

Keywords: History Matching, Multi-Objective Optimization, Inverse Problem, Kernel-PCA, Multi-Dimensional Scaling, Geostatistics

1 Introduction

History Matching of Oil Reservoirs is a much need procedure in the workflow of reservoir analysis. It enables the planning of future exploration considering economic constraints by reducing the uncertainty associated with the knowledge of fundamental petrophysical proprieties (e.g. permeability and porosity).

*carneirus@gmail.com

†Professor at Instituto Superior Técnico: asoares@ist.utl.pt

The subject as received continuous considerable attention since the early works of Jacquard [6] and Chen et al. [2] among others.

The subject is currently well developed, as it can be realized in the extensive review recently preformed by Oliver and Chen [9].

History Matching problems are mathematically categorized as an ill-posed and inverse problems, i.e., one knows the solution but lacks information about the conditions which enabled to arrive to the observed solution. In terms of oil reservoir analysis it means that one uses the observed dynamic data obtained from currently producing wells and from that data one tries to compute - or in History Matching parlance, to obtain a "match" - a petrophysical model (or models) capable of generating the observed dynamic response. Hence, it can also be mathematically seen as an optimization problem.

According to Sarma et al. [11], current approaches to History Matching problems can be classified in four main categories:

- gradient calculation;
- streamline simulation-based algorithms;
- ensemble Kalman filters;
- stochastic algorithms.

In this paper, the focus rests on the last category. Stochastic algorithms are based on the perturbation of the initial image of the property subject to match, and the application of an objective function in order to rank the response obtained from fluid simulation of the perturbed image. The process is iterative and can find several minima in the objective space. The main drawback in this class of algorithms concerns the high cost in terms of fluid simulations needed in order to achieve a reasonable match.

In this paper, a modification of the stochastic algorithm developed by Mata-Lima [8] is proposed in order to use a multi-objective function based on the simple concept of Pareto Front (or Pareto Optimal Set). The multi-objective approach carries a large computational penalty which is minimized using the image clustering method outlined by Scheidt and Caers [12]. These three key concepts are the subject of the following sections.

2 Stochastic History Matching using Co-Direct Sequential Simulation

The stochastic History Matching algorithm of Mata-Lima [8] can be summarized in the following steps:

1. **Initial Seed**: initial generation of an image set of size N of the property subject to "match" (e.g. permeability) by Direct Sequential Simulation [15] conditioned only to the hard data[1] available and to the bi-point spacial continuity information revealed by the same data;

2. **Transfer Function**: the N images are used as input in N finite difference simulation flows in order to obtain time responses of dynamic data of the variable used to perform the match (e.g. WWCT - Well Water Cut). In the general framework of stochastic algorithms, the use of a fluid simulator can be viewed as the simple application of a (usually) non-linear transfer function, i.e., fluid simulation is treated as a black box;

3. **Response Ranking**: the responses of the transfer function are ranked using some objective function - e.g. the modified Hausdorff distance [1] - to the observed historic data of the dynamic match variable (e.g., WWCT). If the obtained match is satisfactory the process stops here;

4. **Perturbation**: a new set of N images is generated by Co-Direct Sequential Simulation, where the image with the best rank from the previous step is used as conditioning information. The process iterates to step 2.

Using this general framework, Mata-Lima [8] explores how the perturbation stage can be improved, either by varying the correlation with the conditioning (or secondary image), and/or by using a linear combination of best ranked images as secondary information. A regionalized version is also explored where the secondary image is a composite of the images which gives the best match according to each well.

This regionalized variant is the one used in the present work.

3 Multi-Objective Approach: Using the Pareto Front

The concept of Pareto Front or Pareto Optimal Set is a widely used concept in multi-objective optimization and it as already been used in multi-criteria approaches to History Matching problems [14].

The first approach considered when dealing with a multi-objective optimization problem usually consists in the linear combination of the objective functions for each optimization variable. The problem is thus reduced to a single-objective one. This approach as inherent difficulties such as the fact that many times the objective variables conflict with each other, i.e., a better fit in one will necessary lead to a worse

[1]hard data sets are those which result from direct observation of petrophysical properties on core samples or from well logging.

fit in the other. Another difficulty avoided by the use of the Pareto front concept is the use of a scaling procedure in the combined multi-objective approach. When badly done, such scaling can lead to very slow convergence [10].

The Pareto Front or Pareto Optimal Set is defined as follows: Consider a multi-objective problem with m parameters (to minimize) and n objectives:

$$\begin{aligned} \text{to minimize} \quad & \vec{y} = f(\vec{x}) = (f_1(x), ..., f_n(x)) \\ \text{where} \quad & \vec{x} = (x_1, ..., x_n) \ \epsilon \ X \\ \text{and} \quad & \vec{y} = (y_1, ..., y_n) \ \epsilon \ Y \end{aligned} \tag{1}$$

The vector $\vec{x}$ is called the decision vector in the parameter space X and $\vec{y}$ a vector in the Y objective space spanned by the individual objective functions. A decision vector $\vec{a} \ \epsilon \ X$ is said to dominate another $\vec{b} \ \epsilon \ X$ (or with the same meaning $\vec{a} \succ \vec{b}$), iff:

$$\begin{cases} \forall_i \epsilon \{1, ..., n\} : f_i(\vec{a}) \leq f_i(\vec{b}) \\ \exists_j \epsilon \{1, ..., n\} : f_j(\vec{a}) < f_j(\vec{b}) \end{cases} \tag{2}$$

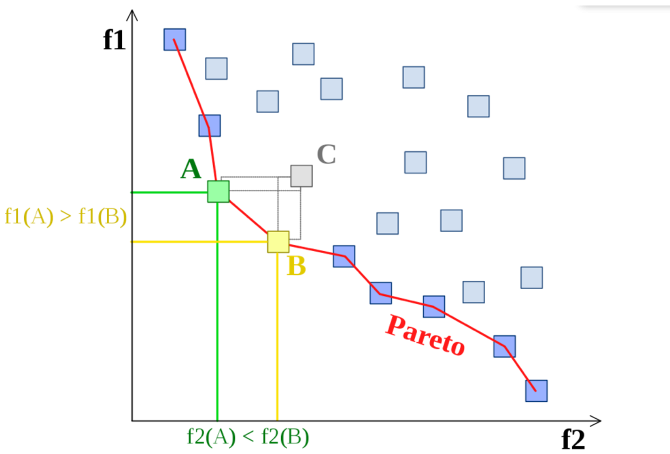


Figure 1: Pareto Front in a two dimensional objective space: point C is dominated by both point A and point B. No point dominates over these last two, hence they are part of the set of optimal choices defining part of the Pareto Front.

The set of undomitaded vectors constitutes a set of optimal solutions on all optimization parameters known as the Pareto Optimal Set or Pareto Front (figure 1).

The use of this concept on the History Matching algorithm previously described (section 2 - regionalized variant) will generated not one but several best images per well. The permutation of all these local best fit images will generate a potentially very large set of composite optimal images entailing a consequentially very large computational weight in the application of the transfer function. As such, these image set must be reduced in order to use just a few representative images.

4 Reducing the Composite Optimal Image Set

The extraction of a few representative images from the generated composite optimal imaged set is accomplished following the image clustering procedure used by Scheidt and Caers [12]. It entails the following steps:

1. **Image Feature Extraction**: The individual images needs to be simplified in order to perform efficient pattern recognition. For that purpose, all the images are submitted to edge-detection using the Sobel Operator, and the resulting images are then subject to a cut around their median value, resulting in a reduced point-set that preserves the main features of each image. This procedure is common in pattern extraction workflows [4], and can be observed in figure 2 applied to one of the optimal images resulting from the procedure described previously (section 3).

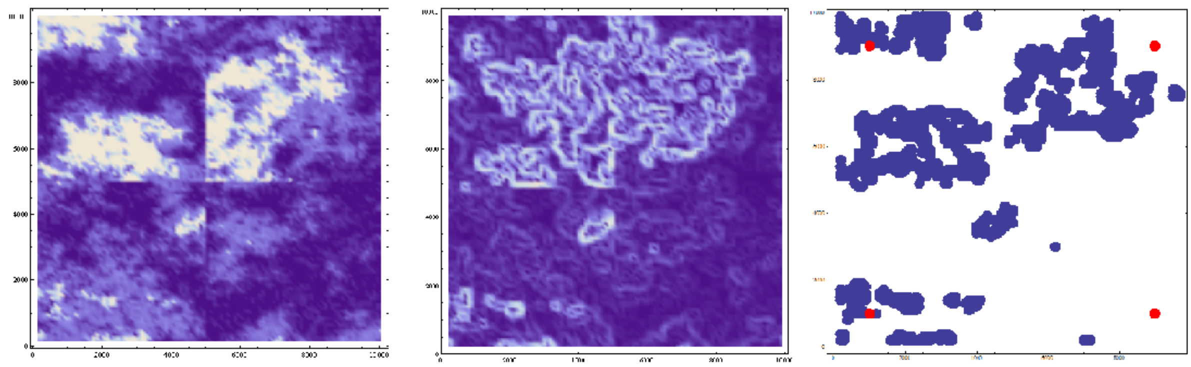


Figure 2: One of the composite optimal images obtained (left) is submitted to edge-detection using the Sobel Operator (center) and finally a cut around its median value is performed (right).

2. **Hausdorff Matrix Distance Calculation**: After the pattern extraction procedure is performed for all composite images, the Modified Hausdorff Distance is calculated between all of them. This metric is quite useful and again frequently used in pattern recognition workflows [4]. It is defined as

follows; Let $A = \{a_1, ..., a_n\}$ and $B = \{b_1, ..., b_n\}$ be two finite point-sets, the Hausdorff distance between A and B is given by:

$$H(A,B) = \max(h(A,B), h(B,H)) \qquad \text{where} \tag{3}$$

$$h(A,B) = \max_{a \epsilon A} \min_{b \epsilon B} \|a - b\| \tag{4}$$

where $h(A,B)$ is the direct Hausdorff distance from set A to set B for a given norm $\|.\|$. The Modified Hausdorff Distance is defined as follows [4, 7]:

$$h_{mod}(A,B) = \frac{1}{|A|} \sum_{a \epsilon A} \min \|a - b\| \tag{5}$$

This modification, calculating the average of the point-to-point distances, reduces the impact of outliers, thus being more appropriate for pattern recognition purposes.

3. **Generation of an Euclidian Space by Multi-Dimensional Scaling**: Once the distance matrix between all composite images is available, it is used as input to a statistical procedure for dimensionality reduction known as Multi-Dimensional Scaling [3, 5]. This procedure uses as input a distance matrix (or any dissimilarity matrix) and builds a N dimensional Euclidean projection, i.e., it provides a mapping on a N-dimensional euclidean space of the original data.

4. **Kernel-Principal Component Analysis and Clustering**: On the Multi-Dimensional generated euclidean space, a technique known as kernel-Principal Component Analysis proposed by Scholkpof et al. [13] is used. This technique is a non-linear extension of Principal Component Analysis in which the feature space is mapped by a non-linear function Φ, of which usually only the inner - product matrix between all points (the kernel matrix) is known. The use of a non-linear feature space ensures that every possible cluster, which might not be linearly separable in the original space, is separable in the non-linear feature space. As such, the projection of the input data points in the feature space where separability between clusters is enhanced is ideal for the use of non-hierarquichal clustering techniques such as the K-means algorithm. That is exactly what it is done in the described methodology, being the number of final clusters pre-determined by the number of flow-simulation sets computationally feasible. The point (i.e. image) extracted is the nearest neighbor of the cluster centroid.

When the composite image reduction process is completed, the method of Mata-Lima [8] continues to the perturbation stage with the images selected by the explained procedure has secondary information.

5 Case Study

The described method was applied to a five-spot (four production wells and one central water injection well) with the following overall parameters:

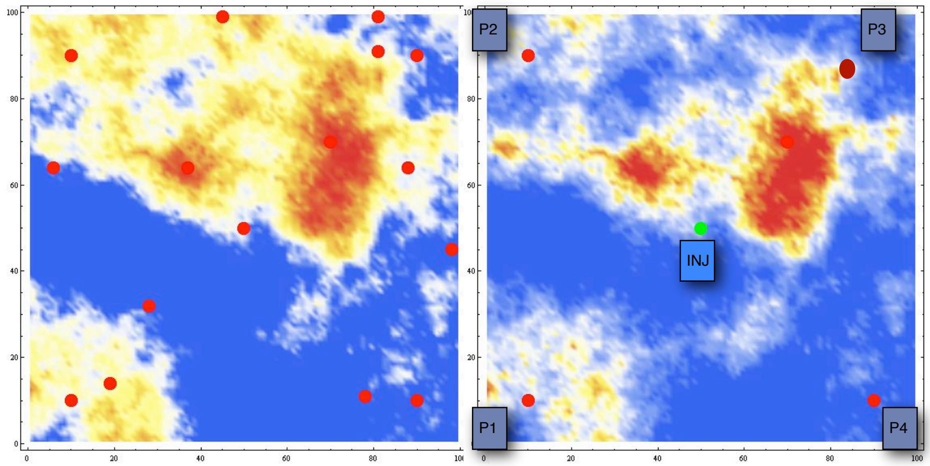
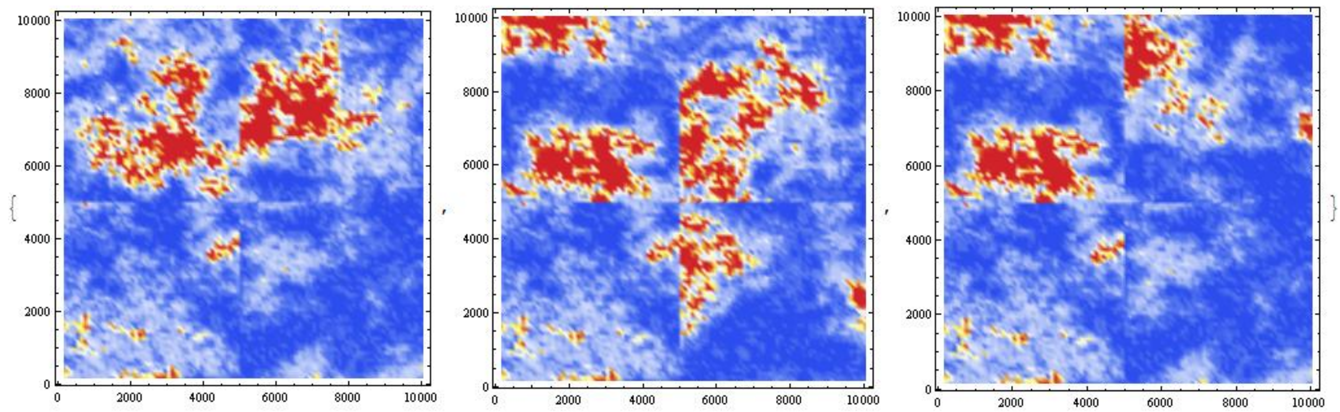
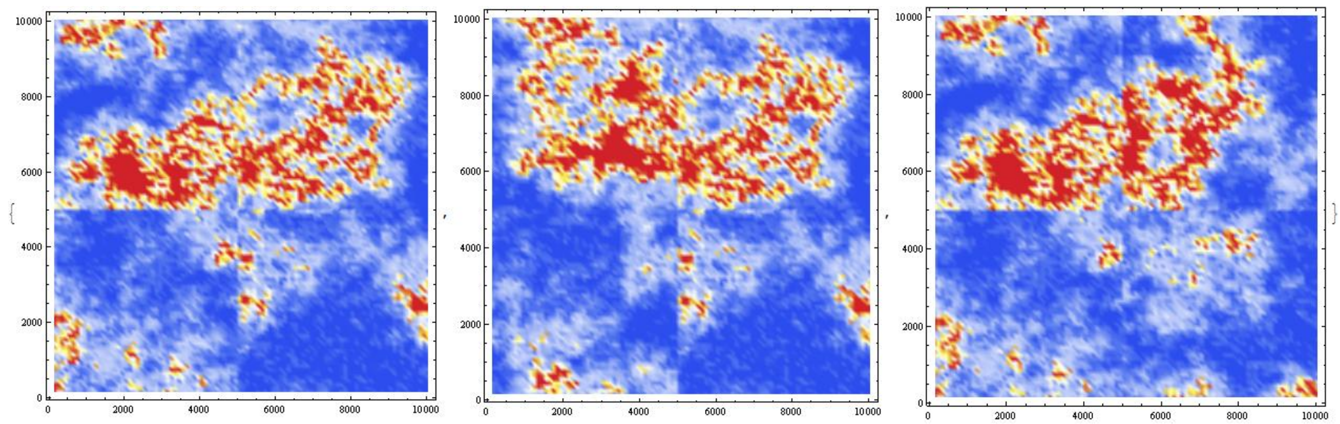


Figure 3: Synthetic porosity (left) and permeability (right) images of the five-spot test case. Over-imposed on theses images are the sample grid of hard data (right) and well locations (right).

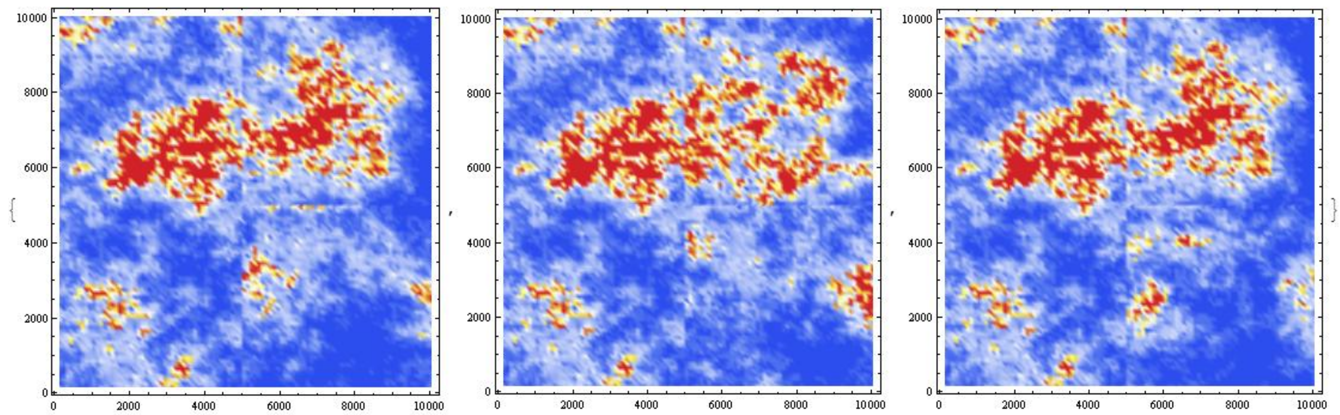
- initial permeability image set of thirty images generated by Direct Sequential Simulation;
- three secondary images generated in each iteration;
- dimension three used in the image clustering procedure;
- kernel-PCA used with a Gaussian Kernel;
- each secondary image is correlated at 0.8 in the perturbation by Co-Direct Sequential Simulation stage. Twenty images are generated per secondary image used;
- two dimensional multi-objective space spanned by WWCT (Well Water Cut) and WBHP (Well Bottom Hole Pressure) match variables. The scalar objective function used is the Modified Hausdorff Distance;
- two iterations performed.



iteration 0 (DSS): composite representative optimal images obtained from a total of 72.



iteration 1 (Co-DSS): composite representative optimal images obtained from a total of 160.



iteration 2 (Co-DSS): composite representative optimal images obtained from a total of 240.

Figure 4: Resulting representative optimal images for each iteration.

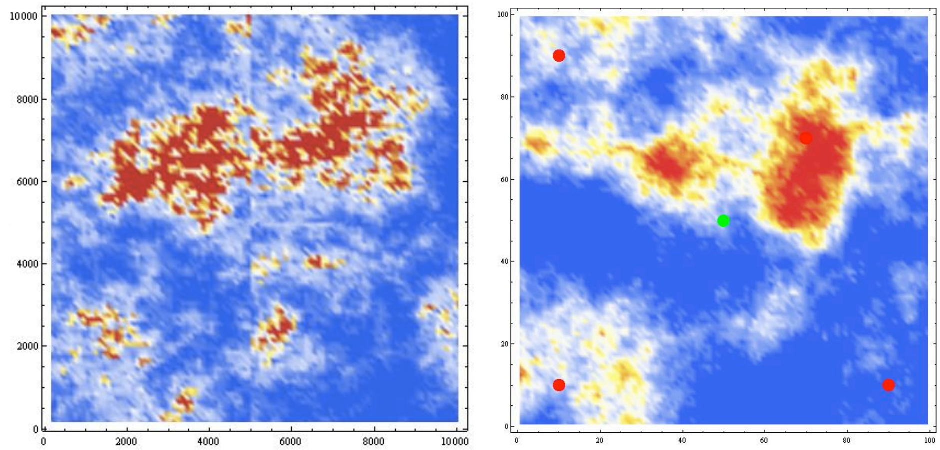


Figure 5: final match: one of the images obtained with iteration 2 (left) and the synthetic image of permeability (right).

The composite images resulting from the clustering process are sequentially presented in figure 4 and one of the final images of iteration 2 is presented in figure 5 along with the synthetic permeability image. As it can be seen, a very good match can be obtained as early as the second iteration, even with parameters that may not be ideal. The composite final images reveal a remarkable increase in spatial continuity, being the artifact introduced by the regionalized well areas progressively reduced.

The number of composite images per iteration increases considerably while simultaneously the similarities (by visual inspection) between the representative images tend to lessen. This fact can be considered natural if the process is convergent, as the images should be progressively more alike.Hence, the final match (figure 5) can be considered quite good.

6 Conclusions and Future Work

Among the proposed modifications to the History Matching procedure of Mata-Lima [8] and considering the results of the case study, the following conclusions can be stated:

1. the procedure is successfully adapted to multi-objective optimization by use of the Pareto Front without an excessive computational weight due to the image clustering procedure applied [12];

2. the resulting composite images show a reestablishment of spatial continuity, which can be seen as an indicator of convergence;

3. the match is quite good at a very early stage of the procedure (iteration 2)

The proposed procedure was only tested on the presented case study and as such, it needs extensive testing and fine-tuning in each of its steps, validation against other History Matching methods and testing on real reservoirs. In the execution of the presented test case, several issues worthy of further scrutiny were noticed and are stated:

1. in the application of the vectorial objective function and following calculation of the Pareto Optimal Set, every previous iteration should be considered. It was observed that a few optimal points of previous iterations remained so, hence the corresponding images should be preserved;

2. the number of composite optimal images increases considerably with each passing iteration and simultaneously the similarity between representative images also increases, hence it can be expected that the distance values contained in the calculated Hausdorff matrix will show a significant decrease. Hence, a matrix norm could be used as measure of convergence.

3. the dimension used in the image clustering procedure needs to be evaluated in a systematic way. For instance, in the paper of Scheidt and Caers [12], a four dimensional euclidean space was used after correlation analysis on the Multi-Dimensional Scaling stage. This was not done in the presented test case;

4. the regionalized version of the Mata-Lima [8] procedure enables the use of local correlations per well zone of influence. Hence the perturbation stage of the procedure can be fine tuned according to convergence information made available after each iteration.

References

[1] Caeiro, M. H., Soares, A., Santos, J., Carvalho, A., and Guerreiro, L. (2010). Geostatistical History Matching with Direct Transformation of Images. Application to a Middle East Reservoir. In *Rio Oil & Gas Expo and Conference 2010*, pages 1–10.

[2] Chen, W., Gavalas, G., Seinfeld, J., and M.L., W. (1974). A New Algorithm for Automatic History Matching. *SPE Journal*.

[3] Cox, T. F. and Cox, M. A. A. (2001). *Multidimensional Scaling*. Chapman & Hall/CRC, New York.

[4] Dubuisson, M.-P. and Jain, A. (1994). A modified Hausdorff distance for object matching. *Proceedings of 12th International Conference on Pattern Recognition*, pages 566–568.

[5] Härdle and W. Simar, L. (2003). *Applied Multivariate Statistical Analysis.* Humboldt-Universität zu Berlin, Berlin.

[6] Jacquard, P. (1965). Permeability Distribution from Field Pressure Data. *SPE Journal*, 5:281–294.

[7] Jesorsky, O., Kirchberg, K. J., and Frischholz, R. W. (2001). Robust Face Detection Using the Hausdorff Distance. In *In Proc. Third International Conference on Audio- and Video-based Biometric Person Authentication*, number June, pages 90–95, Halmstad, Sweden. Springer.

[8] Mata-Lima, H. (2008). Reservoir characterization with iterative direct sequential co-simulation: Integrating fluid dynamic data into stochastic model. *Journal of Petroleum Science and Engineering*, 62(3-4):59–72.

[9] Oliver, D. S. and Chen, Y. (2010). Recent progress on reservoir history matching: a review. *Computational Geosciences*, (February).

[10] Oliver, D. S., Reynolds, A. C., and Ning, L. (2008). *Inverse Theory for Petroleum Reservoir Characterization and History Matching.* Cambridge University Press.

[11] Sarma, P., Durlofsky, L., Aziz, K., and Chen, W. (2007). A New Approach to Automatic History Matching Using Kernel PCA. *Proceedings of SPE Reservoir Simulation Symposium*, (1974).

[12] Scheidt, C. and Caers, J. (2008). Representing Spatial Uncertainty Using Distances and Kernels. *Mathematical Geology*, pages 397–419.

[13] Scholkpof, B., Smola, A., and Muller, K.-R. (2006). Kernel Principal Component Analysis. In *Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Conference*, volume 1.

[14] Schulze-Riegert, R. and Ghedan, S. (2007). Modern Techniques for History Matching. In *9th International Form On Reservoir Simulation*, Abu Dhabi, United Arab Emirates.

[15] Soares, A. (2001). Direct Sequential Simulation and Cosimulation. *Mathematical Geology*, 33(8):911–926.