



**HAL**  
open science

## Constrained nonlinear optimization for extreme scenarii evaluation in reservoir characterization.

N. Metla, Frédéric Delbos, Sébastien da Veiga, Delphine Sinoquet

### ► To cite this version:

N. Metla, Frédéric Delbos, Sébastien da Veiga, Delphine Sinoquet. Constrained nonlinear optimization for extreme scenarii evaluation in reservoir characterization.. ECMOR European conference on mathematics of oil recovery, Sep 2010, Oxford, United Kingdom. hal-02284357

**HAL Id: hal-02284357**

**<https://ifp.hal.science/hal-02284357>**

Submitted on 11 Sep 2019

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

## Introduction

The goal of reservoir characterization is the estimation of unknown reservoir parameters (*the history matching problem*), by integrating available data in order to take decisions for production scheme and to predict the oil production of the field in the future (*the forecast problem*). The reservoir parameters could be classified in two classes:

- those related to the geological modeling (spatial distribution of porosity, permeability, faults),
- and those related to the fluid flow modeling (relative permeability curves, productivity index of the wells).

Those parameters could not be directly determined by measurements (or only locally using well logs), this is the reason why this parameter estimation problem is formulated as an inverse problem with forward simulators that compute synthetic measurable data from those parameters. Observed data are well data acquired at production/injection wells (bottom-hole pressure, gas-oil ratio, oil rate) at different calendar times during the production of the field.

The main contribution of this work is the integration of nonlinear optimization methodology to predict the oil production of a field and to give a confidence interval on this prediction. We believe that applying non linear optimization methods will increase accuracy and then give more reliable production forecast than approaches with simplified models of forward operators (linear approximations or response surfaces). The first and second sections of this paper are respectively dedicated to the history matching problem and to the forecast problem. In the third section, we described the optimization methods used to solve both problems. Then, in the last section the previous methodology is applied to a 3D synthetic reservoir application (the PUNQ test case).

### The History Matching Problem

This History Matching Problem is formulated as an inverse problem, where the objective function is the quadratic errors of the model response with respect to the measured response at each well, defined as

$$\begin{aligned} \text{Minimize } f(x) &:= \|d(x) - d_{obs}\|_{[T_0, T_1]}^2 \\ x &\in R^n \end{aligned} \quad (1)$$

under bound constraints,

where  $x = (x^1, \dots, x^n) \in R^n$  are the unknown reservoir parameters,  $d_{obs}$  and  $d(x)$  are respectively the observed and computed data from the initial time  $T_0$  up to the final measured time  $T_1$ . Problem (1) is a nonlinear least-squares optimization problem subject to bound constraints. The associated forward problem, or the computation of  $d(x)$ , is a fluid flow simulator, see Fornel et al. (2007) and Feraille et al. (2003) for further details. Solving the forward problem is often computationally expensive and the derivatives with respect to the parameters are not available. Both problems (forward and inverse) are nonlinear. In the following, the solution of problem (1) is noted  $x_{HM}$ .

### The Forecast Problem

Probabilistic history matching is a standard approach to solve the Forecast Problem. Feraille and Busby (2009) use statistical technique that combines non-parametric response surface modeling with adaptive design method to approximate the objective function. This OF approximation is then used in a Markov Chain Monte Carlo algorithm to compute the posterior distribution of the parameters that respect the production data. Then, reservoir parameters are optimized to maximize the cumulative oil production while minimizing the risk using the response surface model.

In the following we propose to use another approach based on nonlinear optimization. This approach is made up of two step :

- once a suitable reservoir model  $x_{HM}$  is found to satisfy (1), the fluid flow simulator can be run once more to compute simulator responses up to a future time  $T_2$ . Then one can give, as plotted in Figure 1, a first prediction of the future cumulated oil production. This methodology corresponds to the first step of the forecast problem,
- the second step of the forecast problem consists not only in finding one prediction value, but also in computing its uncertainty. In fact, the solution  $x_{HM}$  is only one solution among others which well fit the observed data  $d_{obs}$ . In other words, another reservoir parameters  $\tilde{x}$  different from  $x_{HM}$  may be such that the associated simulated value  $d(\tilde{x})$  matches the observed data  $d_{obs}$  with an equivalent accuracy as  $f(x_{HM})$  on  $[T_0, T_1]$ . Due to the nonlinearity of the forward operator  $d(x)$ , a small perturbation of the model can lead to a much larger perturbation of the simulator responses and then could yield to errors in the first step of the forecast problem.

An optimization viewpoint of the forecast problem is to find reservoir models which maximize and minimize profit up to time  $T_2$  and which at the same time sufficiently fits the observed data on  $[T_0, T_1]$ . This idea was first proposed by Guérillot and Roggero (1995). This corresponds to look for upper and lower bounds or for a confidence interval of the forecast. We define the optimistic and the pessimistic forecast scenario as :

- (i) "*Optimistic*" scenario means to identify the parameters that **maximize** the production forecast function (up to  $T_2$ ) with respect to matched data, *i.e.*,

$$\begin{aligned} & \text{Maximize}(\text{forecast function}) \\ & \text{s.t. History Matching constraint,} \end{aligned}$$

- (ii) "*Pessimistic*" scenario means to identify the parameters that **minimize** the production forecast function with respect to matched data, *i.e.*,

$$\begin{aligned} & \text{Minimize}(\text{forecast function}) \\ & \text{s.t. History Matching constraint.} \end{aligned}$$

Based on above definitions, we state the forecast problem as the optimization problem

$$\text{Maximize}_x / \text{Minimize}_x p(x, T_2) \tag{2}$$

$$\text{subject to } f(x) \leq \varepsilon \tag{3}$$

and bound constraints,

where  $\varepsilon$  is an absolute error on the history matching problem and  $f$  is the objective function of the history matching problem already defined in equation (1). The objective function (2) focuses on minimizing or maximizing the total cumulated oil production  $p$  at  $T_2$ . Equation (3) is the nonlinear history matching constraint. Optimistic and the pessimistic forecast scenario for a given value of  $\varepsilon$  are represented in Figure 2. Note that if  $p$  stands here for the cumulated oil production, it can also be defined by any function of simulator responses. Moreover, note also that the simulator responses can correspond to computed data which are different from the ones observed (for instance the computed pressure at the top of the reservoir).

## Optimization Methods

The objective of this section is to explain which optimization techniques are used and why they are efficient to solve the previously described history matching and forecast problems. Both problems, (1) and (2)–(3), correspond to nonlinear optimization problem without derivatives. The optimization methods that we can use to solve those problems can be gathered into two classes:

- the class of standard nonlinear optimization methods, which includes the sequential quadratic programming (SQP) approach. Sequential quadratic programming is one of the most effective methods for solving nonlinearly constrained optimization problems. The approach was first suggested by Wilson (1963) for the special case of convex optimization, then popularized mainly by Biggs (1972), Han (1976), and Powell (1978b,a) for general nonlinear constraints. Gould and Toint (2000) survey the recent development in SQP. The main idea of the SQP approach is to solve the nonlinearly constrained problem using a sequence of quadratic programming (QP) subproblems. In each QP subproblem, the constraints are obtained by linearizing the constraints in the original problem, and the objective function is a quadratic approximation to the Lagrangian function,
- the class of free derivative optimization methods: besides the direct search methods (Nelder and Mead, 1965; Audet and Dennis, 2003; Kolda et al., 2003) and the metaheuristic methods as evolutionary approaches (Hansen and Ostermeier, 1996; Kirkpatrick et al., 1983), a quite popular class of methods based on sequential minimizations of models of the objective function turns out to be very efficient in terms of number of simulations required to converge to a minimum (Schonlau, 1997; Villemonteix, 2008; Conn et al., 2009; Langouët et al., 2010).

While the second class is dedicated to optimization without derivatives, the first class when applied on problems without any available analytical derivatives needs strong computational effort to evaluate them (for instance with a standard finite differences method). However, when applied on least-squares problems this class of method can have a very good converge rate due to the Gauss-Newton approximation of the hessian of the Lagrangian (see below for more details). In the following, we give some details about the first class of method and more particularly about SQP solver. For more details about the second class, free derivative methods, and an application in reservoir engineering see Langouët et al. (2010).

The SQPAL solver is a sequential quadratic programming method suited to nonlinear constrained optimization problems. SQPAL stands for Sequential Quadratic Programming and Augmented Lagrangian: the tangent quadratic problem is solved thanks to an original method based on a combination of an augmented Lagrangian method and active-set method (cf Sinoquet and Delbos (2008)). Quasi-Newton techniques are used to approximate the Hessian of the Lagrangian. Two types of quasi-Newton methods are implemented into the solver:

- the Gauss-Newton method, which is suitable for least-square formulations in which Hessian is not available, while the Jacobian matrix of the forward problem can be computed and stored. The convergence rate is closed to Newton method (quadratic convergence) for weakly nonlinear problems or when the residuals between observed data and modeled data are small.
- the BFGS method (see Bonnans et al. (2003)) which is adapted to applications where second order derivatives of the cost function are not available but where gradient can be estimated. Larger is the number of iterations, better is the Hessian approximation, the given initial approximation being usually the identity matrix.

It is well known that the convergence of SQP methods described so far is ensured only if the initial guess  $x_0$  is in a sufficiently small neighborhood of the optimal solution, which is unknown. To overcome this

difficulty one has to apply a good globalization technique. The main approaches are line search and trust-region methods, see e.g. Nocedal and Wright (1999). Line search and trust-region methods both generate steps with the help of a quadratic model of the objective function, but they use this model in different ways. Line search algorithms uses it to generate a search direction, and then focus their efforts on finding a suitable step length along this direction. Trust-region methods define a region around the current iterate where they trust the model to be an adequate representation of the objective function, and then minimize the model in this trust region to determine the next iterate. The trust region size is updated according to the adequation of the model with the objective function at this new iterate. In the following we focus on two methods: line search and dog-leg. Latter method is a trust-region method with an approximate solution of the trust-region subproblem. Dog-leg method could not be applied to optimization problems subject to nonlinear constraints.

Then, from previous remarks, to solve problem

- (1), which is a least-squares problem, we use the "SQPAL" solver with a Gauss-Newton approximation of the Hessian of the Lagrangian. As this problem includes only bound constraints, we choose a dog-leg trust-region globalization technique (preferred to line-search because of its robustness).
- (2)–(3), since the objective function is not a least-square function, we use the "SQPAL" solver with a BFGS approximation of the Hessian of the Lagrangian. As it is not possible to use the dog-leg method (due to presence of nonlinear constraint), we use line search method as a globalization technique.

For both problems derivatives of the objective function and of simulated data with respect to the parameters  $\frac{\partial d(x)}{\partial x^i}$  are approximated using a finite differences method. In next section results of problem (2)–(3) will be also computed with SQA method (the derivative free optimization method), and then compared to ones from SQPAL method proposed above.

### Application to a 3D synthetic reservoir : the PUNQ test case

The PUNQ test case is a 3D synthetic reservoir model derived from real eld data. It was already used for comparative inversion studies in the European PUNQ project (Floris (2001)) and for validation of constrained modeling and optimization scheme development methods (Roggero (2001)). The top structure of the reservoir is presented in Figure 3. The reservoir is surrounded by an aquifer in the north and the west, and delimited by a fault in the south and the east. A small gas cap is initially present. The geological model is composed of five independent layers. The layers 1, 3, 4 and 5 are assumed to be of good quality, while the layer 2 is of poorer quality. The reservoir is made up of two block units. Its size is  $3\text{km} \times 5\text{km} \times 90\text{m}$ . Four different facies are considered: two sandstones and two clays. In the following we will consider only the two sandstones. We consider a production scheme with 11 wells : 6 producers and 5 injector (water is injected in the aquifer to maintain pressure). The upscaling procedure is used for building a fluid flow simulation 3D-model with  $19 \times 28 \times 5$  grid blocks, and fluid flows in 1761 active cells. The flow simulator uses a black oil model described. The available data are the cumulated oil production, the gas-oil-ratio and the water-cut of 6 producers observed during 10 years. To this synthetic data, 4% of Gaussian relative noise have been added. The unknown reservoir parameters are:

- the permeability of the aquifer
- 4 permeability multipliers
- the gasflood oil residual saturation
- the waterflood oil residual saturation

It ends up with 7 parameters and 228 measurements. The goal is to infer these unknowns and to forecast the cumulated oil production of entire reservoir after 22 years.

The results of the History Matching problem are plotted in Figure 6. The SQPAL Gauss-Newton optimizer with a dog-leg globalization needs 75 simulations to converge. The solution  $x_{HM}$  is reached after 59 simulations and the value of the cost function is  $f(x_{HM}) = 2.679$ . In Figures 5 one can observe the match between measured and computed GOR data (for 3 different wells). Figures 6 shows the adequation between measured and computed cumulated oil production data (in the history matching part).

The cumulated oil production forecast and its upper / lower bound with respect to different  $\varepsilon = \{5, 10, 20, 30\}$  can be observed in Figure 6. Identical results are observed in Figure 7 with a y-scale that represents the forecast error (computing the difference of a curve with the HM forecast curve). In both Figures, one can observe that the larger  $\varepsilon$ , the larger the bounds on the cumulated oil production forecast are. This behavior is logic because when  $\varepsilon$  increase the set of admissible models for the nonlinear constraint (3) increase as well. And minimizing or maximizing on a larger set can only lead to larger (or equal) maximum values and smaller (or equal) minimum values. For  $\varepsilon = 30$  the error on the cumulated oil production forecast after 22 years of production belongs to the interval  $[-0, 2; 0.1]$  in millions of  $m^3$ . From this result one can also deduce that the posterior distribution of the cumulated oil production is not symmetric (this is due to the nonlinear property of the inverse and forward problems). Extreme scenarios obtained without the HM nonlinear constraint ( $\varepsilon = +\infty$ ) are plotted in Figure 8. In this case the confidence interval becomes almost twice larger than the one obtained with  $\varepsilon = 30$  (from 0.3 to 0.6 millions of  $m^3$ ). That means that nonlinear constraint plays crucial role in prediction phase.

Analogous results, not plotted here, are obtained with the derivative free optimization approach : the SQA solver (Sequential Quadratic Approximation). In tables 1–2 the two classes of approaches recommended to solve the forecast problem (2)–(3)(see previous section) are compared : classical SQP-BFGS method with a finite differences method for the computation of derivatives (with a line search as globalisation technique) versus the derivative free algorithm SQA. First column of the tables gives the values of  $\varepsilon$ . Second and third column show respectively the number of simulations needed by SQPAL and SQA to convergence. Next two column give the forecast objective function values, i.e. the final cumulated oil production at time  $T_2$  using both solvers. Finally, in the last column we compute the relative distance between reservoir models obtained by SQPAL and SQA (it consist in computing the  $l_2$ -norm of the difference between both reservoir models divided by the  $l_2$ -norm of  $x_{HM}$  ). one can remark that whether in term of objective function value or in term of relative distance of the computed models, both optimizers give similar results. However, one can see that SQA is able solve the forecast problem about 1.5 times faster than SQPAL (the CPU time needed to get one solution of the forecast problem is proportional to the number of simulations needed to converge).

## Conclusion

The History Matching and Forecast problems are both formulated as a nonlinear optimisation problem. This is as an innovative methodology in reservoir modelling. We emphasize that while the History Matching problem involves quite simple mathematical concepts (minimization of a least-squares function subject only to bound constraints), the forecast problem is complex : it is not anymore a least-squares problem and it includes a nonlinear inequality constraint. The SQPAL and SQA optimizers has given promising results on the presented 3D reservoir problem : results obtained with both solvers are similar. However SQA is 1.5 times faster than SQPAL. This is a very nice result since taking account nonlinear constraints in derivative free optimization is still a research subject in progress. Next step of this work will be to compare these results to the one obtained with a probabilistic History Matching approach.



## References

- Audet, C. and Dennis, J.E. [2003] Analysis of generalized pattern searches. *SIAM J. Optim.*, **13**, 889–903.
- Biggs, M.C. [1972] Constrained minimization using recursive equality quadratic programming. In: F.A. Lootsma, e. (Ed.) *Numerical Methods for Nonlinear Optimization*. Academic Press, 411–428.
- Bonnans, J.F., Gilbert, J.C., Lemarechal, C. and Sagastizabal, C.A. [2003] *Numerical Optimization: Theoretical And Practical Aspects*. Springer-Verlag.
- Conn, A.R., Scheinberg, K. and Vicente, L.N. [2009] *Introduction to Derivative-Free Optimization*. Society for Industrial and Applied Mathematics (MPS-SIAM Series on Optimization), Philadelphia, PA, USA.
- Feraille, M. and Busby, D. [2009] Uncertainty management on a reservoir workflow. *International Petroleum Technology Conference*.
- Feraille, M., Roggero, F., Manceau, E., Hu, L.Y., Zabalza-Mezghani, I. and Reis, L.C. [2003] Application of advanced history matching techniques to an integrated field case study. *SPE Annual technical conference and exhibition*.
- Floris [2001] *Methods for quantifying the uncertainty of production forecasts - a comparative study*, vol. 7. published in *J.Petroleum Geoscience*.
- Fornel, A., Noetinger, B. and Roggero, F. [2007] Using time domain seismic attributes for history matching. *IEA - International workshop and symposium of the International Energy Agency*.
- Gould, N.I.M. and Toint, P.L. [2000] *System Modelling and Optimization*, Kluwer Academic, chap. SQP method for large scale nonlinear programming. 149–178.
- Guérrillot, D. and Roggero, F. [1995] Matching the future for the evaluation of extreme reservoir development scenarios. *8th European Symposium on Improved Oil Recovery*.
- Han, S.P. [1976] Superlinearly convergent variable metric algorithms for general nonlinear programming problems. *Mathematical Programming*, **11**, 263–282.
- Hansen, N. and Ostermeier, A. [1996] Adapting arbitrary normal mutation distributions in evolution strategies: The covariance matrix adaptation. In *Proceedings of the 1996 IEEE International Conference on Evolutionary Computation*, p. 312–317.
- Kirkpatrick, S., Gelatt, C.D. and Vecchi, M.P. [1983] Optimization by simulated annealing. *Science*, **220**, 671–680.
- Kolda, T.G., Lewis, R.M. and Torczon, V. [2003] Optimization by direct search: New perspectives on some classical and modern methods. *SIAM Review*, **45**, No. 3, 385–482.
- Langouët, H., Delbos, F., Sinoquet, D. and Da Veiga, S. [2010] A derivative free optimization method for reservoir characterization inverse problem. *ECMOR European Conference on the Mathematics of Oil Recovery, 12th, Oxford, UK, 6-9 september 2010*.
- Nelder, J.A. and Mead, R. [1965] A simplex method for function minimization. *Comput. J.*, **7**, 308–313.
- Nocedal, J. and Wright, S. [1999] *Numerical Optimization*. Springer.
- Powell, M. [1978a] The convergence of variable metric methods for nonlinearly constrained optimization calculations. *Nonlinear Programming 3*, Academic Press, New York, 27–63.
- Powell, M. [1978b] A fast algorithm for nonlinearly constrained optimization calculations. *Numerical Analysis Dundee 1977*, no. 630 in *Lecture Notes in Mathematics*, Springer-Verlag, Berlin, 144–157.
- Roggero [2001] Constraining stochastic reservoir models to dynamic data: an integrated approach. Rueil, France, OAPEC seminar.
- Schonlau, M. [1997] *Computer Experiments and Global Optimizations*. Ph.D. thesis, the University of Waterloo.
- Sinoquet, D. and Delbos, F. [2008] Adapted nonlinear optimization method for production data and 4d seismic inversion. *ECMOR European Conference on the Mathematics of Oil Recovery, 11th, Bergen, Norway, 8-11 september 2008*.
- Villemonteix, J. [2008] *Optimisation De Fonctions Coûteuses: Modèles Gaussiens Pour Une Utilisation Efficace Du Budget D'évaluations : Théorie Et Pratique Industrielle*. Ph.D. thesis, Université Paris-Sud Xi Faculté Des Sciences D'Orsay.
- Wilson, R. [1963] *A simplicial algorithm for concave programming*. Ph.D. thesis, Graduate School of Business Administration, Harvard University, Cambridge, MA, USA.

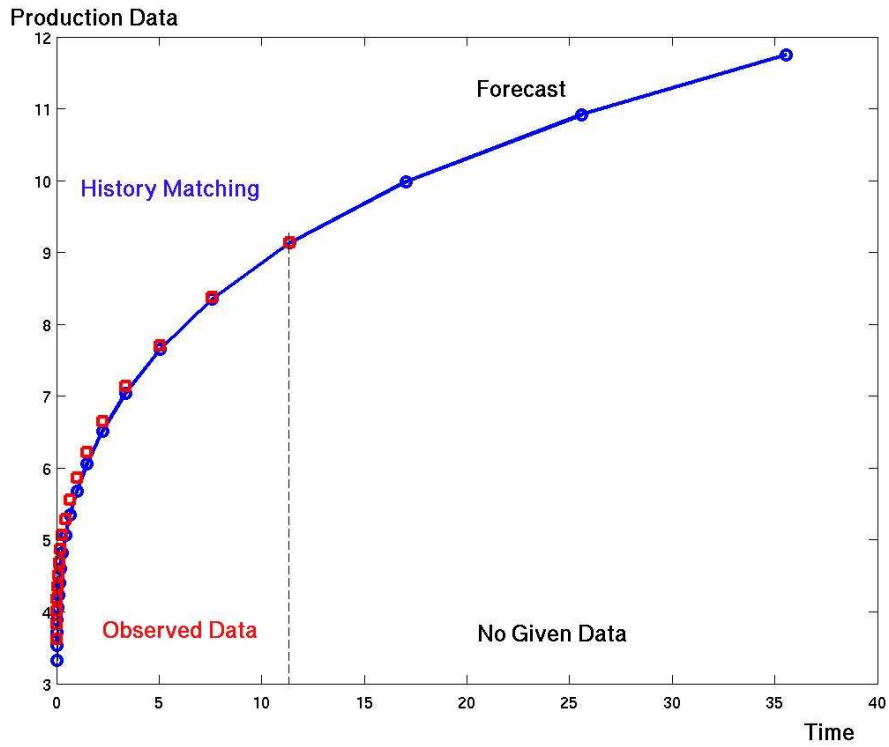


Figure 1: Step 1 of the forecast problem : red point – observed data ; blue line (from time  $T_1$  to time  $T_2$ ) – simulated data before (from time  $T_0$  to time  $T_2$ )

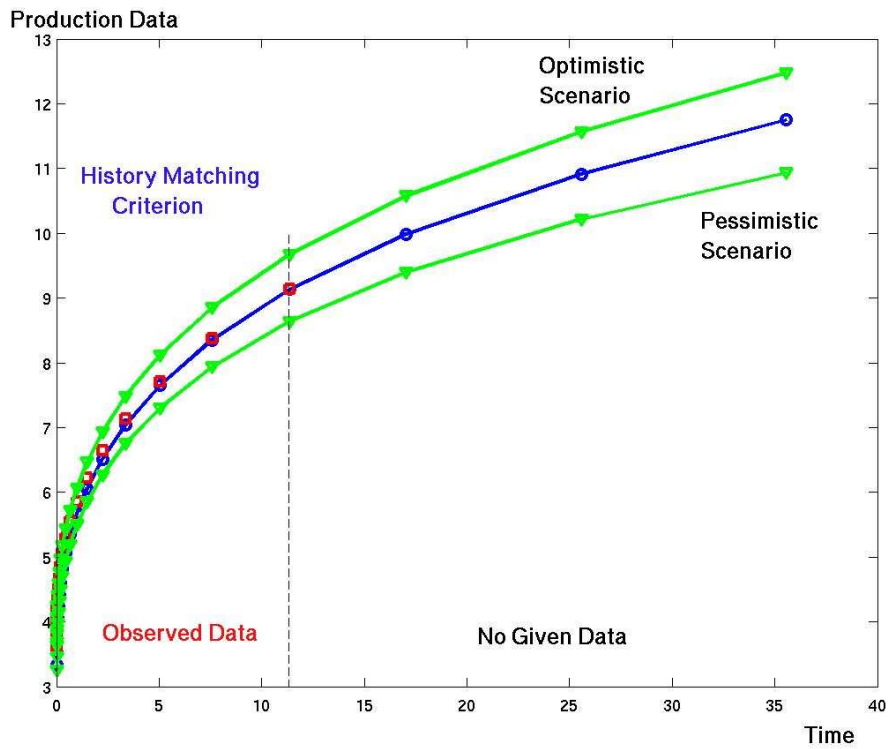


Figure 2: Step 2 of the forecast problem : optimistic / pessimistic scenarios (green lines) give a upper and lower bound of the forecast



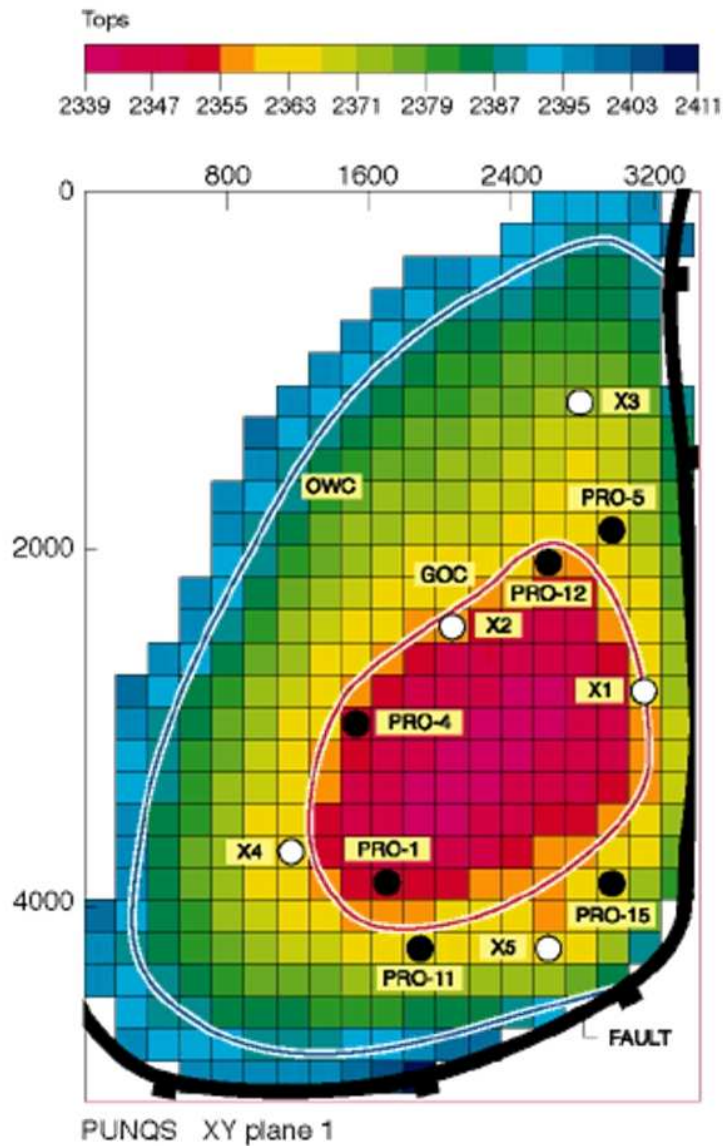


Figure 3: PUNQ

Table 1: SQPAL vs. SQA on the maximization problem (2)–(3)

# Simulations Forecast Obj. Func. Values

Case of Study	SQPAL	SQA	SQPAL	SQA	$\frac{\ x_{SQPAL} - x_{SQA}\ _2}{\ x_{HM}\ _2}$
“max, $\varepsilon = 5$ ”	126	51	4.22009	4.22255	0.009232
“max, $\varepsilon = 10$ ”	100	72	4.25268	4.25299	0.014125
“max, $\varepsilon = 20$ ”	102	74	4.26783	4.27149	0.01854
“max, $\varepsilon = 30$ ”	120	86	4.27568	4.2748	0.029325
“max, $\varepsilon = +\infty$ ”	81	78	4.39805	4.45779	0.228918

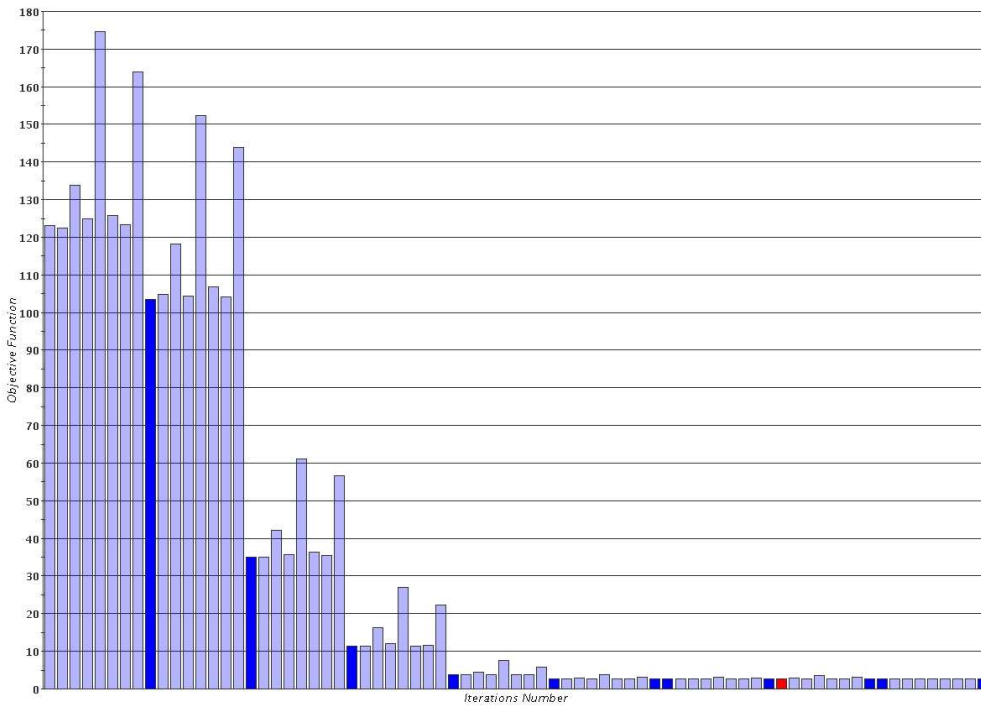


Figure 4: History Matching problem : cost function versus simulation numbers obtained with the SQPAL Gauss-Newton dog-leg optimizer; sky blue bars corresponds to the computation of finite difference gradients; dark blue bars stands for dog-leg iterations; the red bar is the simulation where minimal value of the cost function is reached

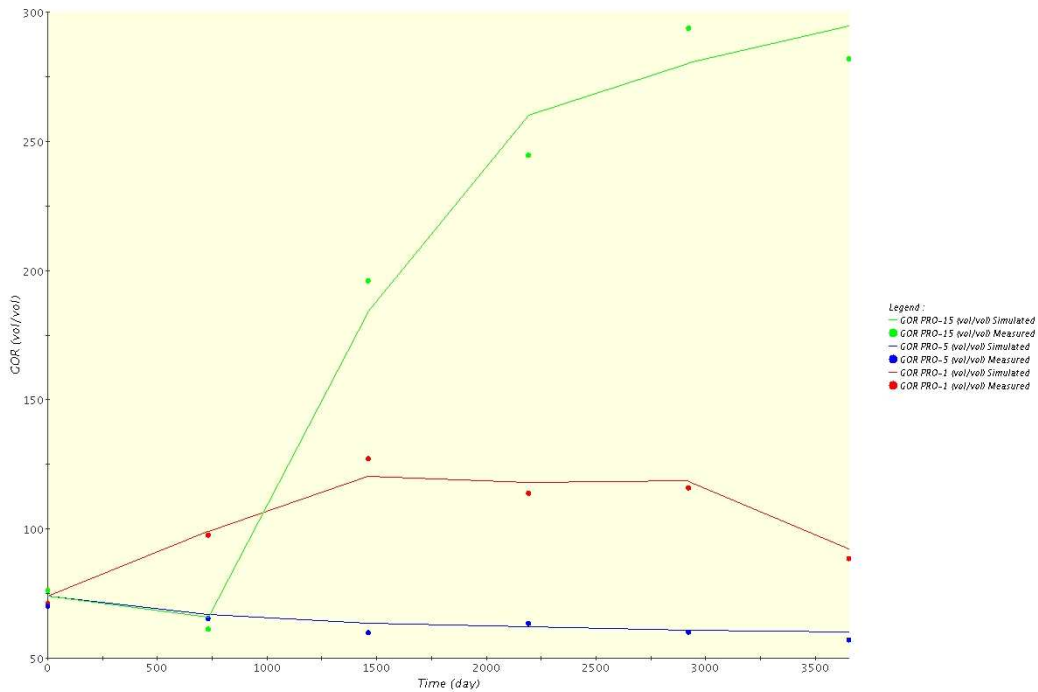


Figure 5: GOR (gas oil ratio) observed from the 3 producers (PRO-1, PRO-5, PRO-15) during 10 years (dots lines), simulated GOR using the solution of the History Matching problem (solid lines)

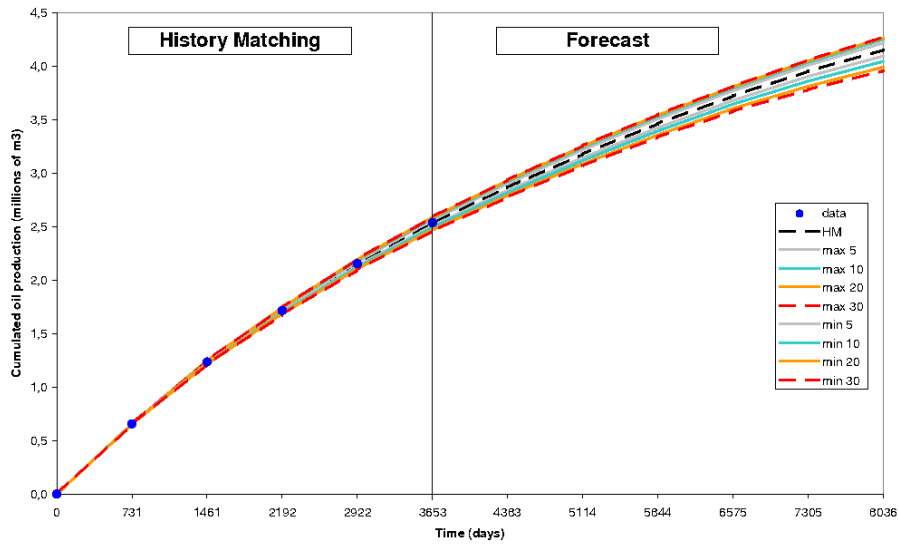


Figure 6: Cumulated Oil Production: blue dots represent observed data during 10 years, dashed black line stands for production forecast computed from the HM solution model, other lines shows max / min of production forecast for various  $\epsilon = 5, 10, 20, 30$

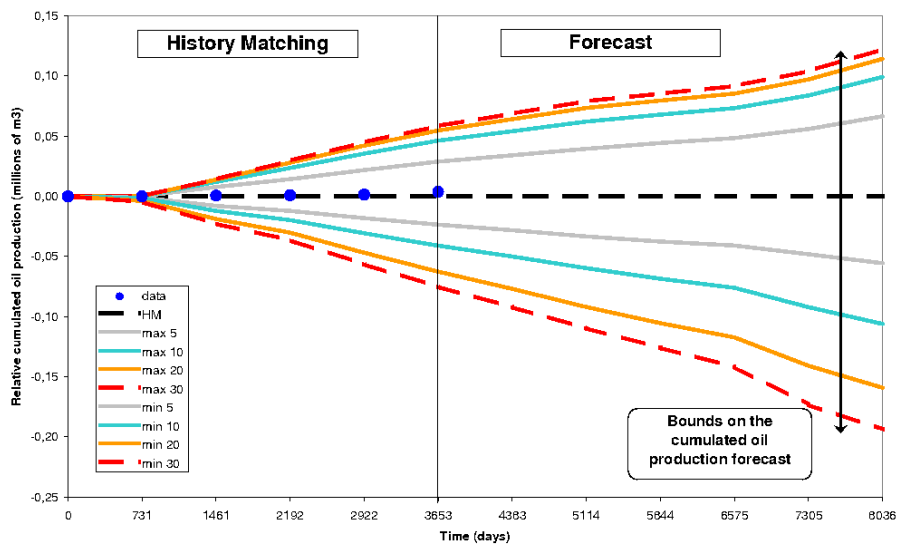


Figure 7: Relative error of cumulated oil production forecast

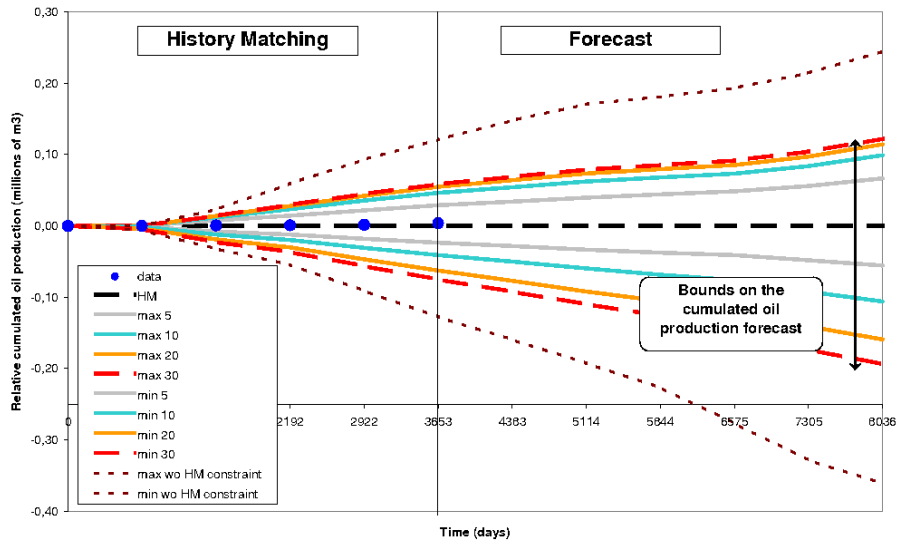


Figure 8: This plot is similar to Figure 7 except for the dashed braun line which corresponds to max / min forecast production without the nonlinear inequality constraint

Table 2: SQPAL vs. SQA on the minimization problem (2)–(3)

# Simulations      Forecast Obj. Func. Values

Case of Study	SQPAL	SQA	SQPAL	SQA	$\frac{\ x_{SQPAL} - x_{SQA}\ _2}{\ x_{HM}\ _2}$
“min, $\varepsilon = 5$ ”	81	65	4.09791	4.09856	0.020384
“min, $\varepsilon = 10$ ”	94	77	4.04742	4.04545	0.026749
“min, $\varepsilon = 20$ ”	149	65	3.9944	3.99438	0.033698
“min, $\varepsilon = 30$ ”	65	70	3.95969	3.9663	0.011092
“min, $\varepsilon = +\infty$ ”	162	63	3.7925	3.78991	0.066514