B44 Adapted Nonlinear Optimization Method for Production Data and 4D Seismic Inversion

D. Sinoquet, Frédéric Delbos

B44

Adapted Nonlinear Optimization Method for Production Data and 4D Seismic Inversion

D. Sinoquet* (IFP) & F. Delbos (IFP)

SUMMARY

Integrated inversion of production history data and 4D seismic data for reservoir model characterization leads to a nonlinear inverse problem that is usually cumbersome to solve : the associated forward problem based, on one hand, on fluid flow simulation in the reservoir for production data modeling, and on the other hand, on a petro-elastic model for 4D time lapse seismic data modeling, is usually computationally time consuming, the number of measurements to be inverted is large (up to 500 000), the number of model parameters to be determined is up to 100. Moreover, all the derivatives of the modeled data with respect to those parameters are usually not available. We propose an optimization method based on a Sequential Quadratic Programming algorithm which uses gradient approximation coupled with a BFGS approximation of the Hessian. In addition, the proposed method allows to handle equality and inequality nonlinear constraints. Some realistic applications are presented to illustrate the efficiency of the method.

1. Introduction

The goal of reservoir characterization is the estimation of the unknown reservoir parameters 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 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 some forward simulators that compute synthetic measurable data from those parameters : production data acquired at production/injection wells (bottom-hole pressure, gas-oil ratio, oil rate), time lapse seismic data (more precisely compressional and shear wave impedances for different seismic campaigns at different calendar times during the production of the field). The associated forward models are on one hand a fluid flow simulator, on the other hand, a petro-elastic model (PEM) based on rock physic Gassmann equations. For further details on this application see Fornel et al.(2007) and Feraille et al.(2003). Solving the forward problem is often CPU time consuming and the derivatives with respect to the parameters are usually not available.

The optimization problem is formulated as the minimization of a least-square objective function composed of two terms, one for the production data mismatch and one for the seismic data mismatch, some weights are introduced to take into account data uncertainties and modeling errors. A natural choice for the optimization method for this kind of problems is the Gauss-Newton algorithm which relies on the computation of the Jacobian matrix. But, in our application, the estimation of the Jacobian matrix is CPU time consuming (the first derivatives not being available, they are estimated by finite differences) and its storage is impossible for large datasets involving seismic data. This is the reason why, in this paper, we propose adapted techniques that furnish approximations of derivatives and a quasi-Newton approach based on classical BFGS approximation of the Hessian of the objective function. The IFP SQPAL optimization package (Delbos et al.(2006)) provides the flexibility to switch from one method to an other depending on the type of applications under study. In a first part, we describe the main features of the SQPAL package useful for the history matching applications and show its efficiency on some benchmark problems in a second part. The last part is dedicated to a 2D synthetic reservoir application with the joint inversion of production data and 4D seismic data.

2. IFP SQPAL package

Optimization takes place in many IFP applications: estimating the parameters of numerical models from experimental data (earth sciences, combustion in engines), design optimization (networks of oil pipelines), optimizing the settings of experimental devices (calibration of engines, catalysis). These optimization problems consist in minimizing a functional that is complex (nonlinearities, noise) and expensive to estimate (solution of a numerical model based on differential systems, experimental measurements), and for which derivatives are often not available, with nonlinear constraints, and sometimes with several objectives among which it is necessary to find the best compromise. IFP has engaged an active research in this field for a number of years and develops its own optimization tools in order to match the needs of its applications as well as possible. 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. Delbos et al.(2006)).

We consider the general constrained optimization problem

$$\min_{x\in\Omega} f(x) \quad \text{subject to } \; c_E(x) = 0, \;\; c_I(x) \leq 0, \tag{1}$$

where a real-valued function $f : \Omega \to R$ is defined on an open set Ω in $\mathbb{R}^n$, c_E and c_I are the vectors of equality and inequality constraint functions, respectively. We further define the feasible set

$$X = \{x \in \Omega : c_E(x) = 0, \;\; c_I(x) \leq 0\}$$

and assume that f, c_E and c_I are differentiable functions. Moreover, c'_E is surjective or onto for all x in the open set Ω. Presently, numerical methods to solve (1) can by gathered into two classes:

- the class of penalty methods, which includes the augmented Lagrangian approaches and the interior point (IP) approaches,
- the class of direct Newtonian methods, which is mainly formed of the sequential quadratic programming (SQP) approach.

Often, actual algorithms combine elements of the two classes, but their main features make them belonging to one of them. The choice of the class of algorithms strongly depends on the features of the optimization problem to solve. The key issue is to balance the time spent in the simulator (to evaluate the functions defining the nonlinear optimization problem) and in the optimization procedure (to solve the linear systems or the quadratic programs). In the seismic reflection tomography application Delbos et al.(2006) argue that the SQP approach is the best fitted. Generally, this is particularly true for applications where the forward modeling is CPU time consuming and where the number of iterations with a Newton-like algorithm is less smaller than the one generated with IP algorithms. This type of conditions are widely encountered at IFP, and particularly in inverse problems issued from earth sciences. This is our main motivation for developing the SQPAL solver, which implements an SQP-like algorithm.

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(1978)a, Powell(1978)b for general nonlinear constraints. Gould & 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.

SQPAL is a software developed for the general nonlinear optimization problem (1). 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 (Bonnans et al.(2006), Nocedal & Wright(1999)) which is adapted to applications where second order derivatives of the cost function are not available but where gradient can be estimated (see algorithm 1). Larger is the number of iterations, better is the Hessian approximation, the given initial approximation being usually the identity matrix.

To be practical, an SQP method must be able to converge from remote starting points. Two class of globalization methods can be used at this point: line-search or trust regions. For BFGS method, a classic line-search based on Armijo rule is implemented. The BFGS Hessian approximations remain positive definite if the initial approximation is positiv definite (see algorithm 1). The globalization of Gauss-Newton approach is based on a trust-region method, a globalization by line-search being delicate: the Gauss-Newton matrix may be ill-conditioned (as for many inverse problems) leading to Gauss-Newton directions closed to be orthogonal to the gradient direction. Two kinds of trust-region methods have been implemented (see algorithm 2 for the general framework of trust region methods):

- the well-known Levenberg-Marquardt method: a l^2 penalty term is added to the Gauss-Newton approximation in order to limit the size of the perturbation:

$$H_k = J_k^T J_k + \lambda_k I \tag{2}$$

with J_k being the Jacobian matrix of the modeled data $d(x)$ in the least-square formulation $f(x) = \|d(x) - d^{obs}\|^2$.

- and the Dog-Leg method which combines the steepest direction and the Gauss-Newton direction if the perturbation computed by Gauss-Newton method $p_k^{GN} = -(J_k^T J_k)^{-1} J_k^T (d(x_k) - d^{obs})$ is such that $\|p_k^{GN}\| > \Delta_k$, i.e.:

$$p_k = p_k^{SD} + (\tau - 1)(p_k^{GN} - p_k^{SD})$$

with $p_k^{SD} = -\frac{\nabla f_k^T \nabla f_k}{\nabla f_k^T J_k^T J_k \nabla f_k} \nabla f_k$, perturbation along the steepest descent direction.

Algorithm 1. BFGS method for unconstrained problem

---

data: Starting point x_0, convergence tolerance $\epsilon > 0$, initial Hessian approximation H_0, $c_1 > 0$
$k = 0$
while $\|\nabla f_k\| > \epsilon$ **do**
compute search direction

$$p_k = -H_k^{-1} \nabla f_k \tag{3}$$

set $x_{k+1} = x_k + \alpha_k p_k$
where α_k is computed from a line search procedure to satisfy Armijo's condition

$$f(x_k + \alpha_k p_k) \leq f(x_k) + c_1 \alpha \nabla f_k^T p_k \tag{4}$$

define $s_k = x_{k+1} - x_k$ and $y_k = \nabla f_{k+1} - \nabla f_k$;
compute H_{k+1} thanks to BFGS formula:

$$H_{k+1} = H_k - \frac{H_k s_k s_k^T H_k}{s_k^T H_k s_k} + \frac{y_k y_k^T}{y_k^T s_k} \tag{5}$$

end

For constrained optimization problems, as previously said, the QP at each SQP iterate becomes a quadratic problem under linearized constraints which is solved thanks to an original method based on the Augmented Lagrangian method (Delbos et al.(2006)). The constrained QP is transformed in

Algorithm 2. Trust-Region method for unconstrained problem

data: Given a maximal trust region radius $\bar{\Delta}$, the initial Trust Region radius $\Delta_0 \in (0, \bar{\Delta})$, $\eta \in [0, 1/4)$

for $k = 1, 2, \ldots$ **do**

obtain the perturbation p_k by solving (approximately) the Tangent Quadratic Problem

$$\min_p m_k(p) = f_k + \nabla f_k^T p + 1/2 p^T H_k p \text{ s.t. } \|p\| \leq \Delta_k \tag{6}$$

compute ρ_k from

$$\rho_k = \frac{f(x_k) - f(x_k + p_k)}{m_k(0) - m_k(p_k)} \tag{7}$$

if $\rho_k < 1/4$ *(the predicted reduction is bad)* **then**

$\Delta_{k+1} = \frac{1}{4}\|p_k\|$

else

if $\rho_k > 3/4$ *and* $\|p_k\| = \Delta_k$ *(the predicted reduction is good and the step size is equal to the trust region)* **then**

$\Delta_{k+1} = \min(2\Delta_k, \bar{\Delta})$

else

$\Delta_{k+1} = \Delta_k$

end

end

end

a sequence of bound constrained quadratic problems via the Lagrangian Augmented method. Each bound constrained problem being solved by a classical combination of gradient projection algorithm, active set method and conjugate gradient algorithm. The difficulty linked to the augmentation parameter determination is based on a precise theoretical study Delbos & Gilbert(2005), which has led us to design a suitable and effective heuristic (Delbos et al.(2006)).

Concerning the globalization of BFGS method by line-search, a l_1 exact penalty function is used as merit function in order to measure the progress of the new iterate according to the minimization of the objective function but also to the respect of the constraints. Up to now, the Gauss-Newton methods coupled with trust region are generalized only to linear constraints. To overcome the difficulty associated with linearized constraint in-feasibility, we follow the elastic mode idea proposed by Gill et al.(2005) in SNOPT implementation.

3. Reservoir application: inversion of monitor seismic datasets

The presented reservoir application is 2D synthetic case : a reservoir cross section made up of two block units. Its size is 3240m in x-direction and 90m in z-direction. Four different facies are considered: two sandstones and two clays. In the following we will consider only the two sandstones (facies 3 and 4 on Figure 1). A black oil model is considered, three wells are used, two producers and one injector (water is injected in the aquifer to maintain pressure). Two synthetic monitor surveys are available after 181 and 547 days (see Figure 3 for P impedance data of base seismic dataset). The objective function is thus composed of 3 different parts associated with the base seismic and 2 different monitor seismic surveys: the mismatch between data modeled by forward operator and P impedances for base seismic, the mismatch of ΔIp (differences of Ip between base seismic and monitor seismic for the two calendar times t =181 days and 547 days).

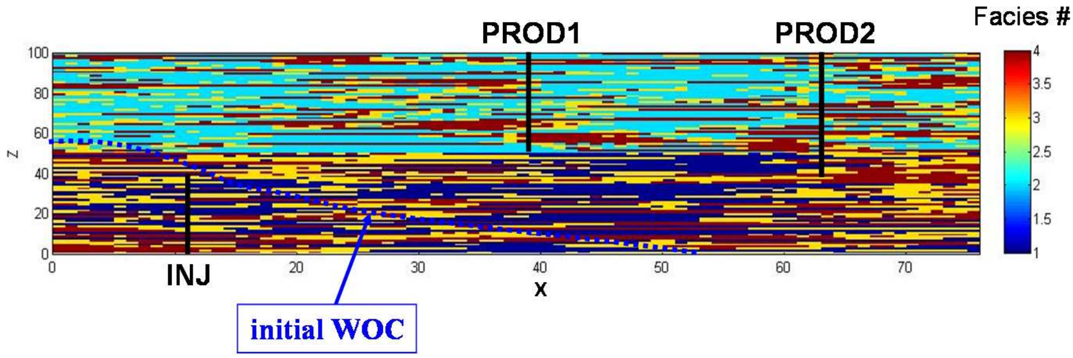

Figure 1: Facies description of the cross section of the considered reservoir : the facies 1 and 2 are associated with clays, 3 and 4 with sandstones. There are two production wells and one injector well.

The unknown parameters of the inversion problem are the mean porosity of the two units and some parameters which control the spatial variations of the porosity (gradual deformations of Gaussian stochastic models of porosity, see Hu et al.(2001)). It ends up with 10 parameters and 70000 measurements. This is a small test case, a real case is usually larger: especially the number of measurements, up to $1M$ (Berthet et al.(2007)).

An important point is that the gradient of the objective function is not available in the forward simulators, the derivatives are thus computed thanks to finite differences. Then, a key point is the choice of the perturbation size: we have adopted an adaptive step depending on the size of the trust region. If the size of the trust region is small, the step size is reduced. A too small step size leads to difficulties, some numerical instabilities in the forward simulator being observed. Even in BFGS method with line-search globalization, a trust region is updated thanks to the estimation of ρ in (algorithm 2), the trust region radius being used to compute an adapted perturbation step size for gradient computation as explained before. Approximations of cost function may be used to estimate numerical gradients (linear and quadratic approximations), simulations already performed at previous iterations are re-used in order to reduce the number of simulations to be done at each iteration: finite differences require np simulations at each iterate (np being the number of optimized parameters).

The applied methodology is composed of two steps: the fist one consists in fitting the base seismic data (P impedance and time shift at the reference time) and then in a second step in adding the fitting of the differences between seismic data associated with the different calendar times and the base seismic data. The results for P impedance of the base seismic are presented in Figure 3 for initial and solution models. The misfits are reduced, especially in the upper part of the reservoir. The 2 curves in Figure 2 illustrate the decrease of the cost function for the two steps : fitting the base seismic and fitting the two seismic surveys. We succeed to reduce the cost function by 30% and 39%. A comparison of SQPAL methods is done for a simplified version of the first step: 4 parameters of gradual deformations are tuned to fit only the base seismic. The results are presented in Figure 4. We observed the decreasing objective function obtained with the three different methods. BFGS method with line-search globalization, Gauss-Newton approach with Levenberg-Marquardt penalization and Gauss-Newton Dog-Leg method. We observed similar decrease for the three methods, even if the two Gauss-Newton approaches are more efficient in the first iterations than the BFGS method. This is not surprising: the initial Hessian approximation being the identity matrix, several iterations are necessary to build a more accurate Hessian approximation. Despite this remark, BFGS approach is preferred for large dataset, the Jacobian matrix of the Gauss-Newton approach being too large to be estimated and stored.

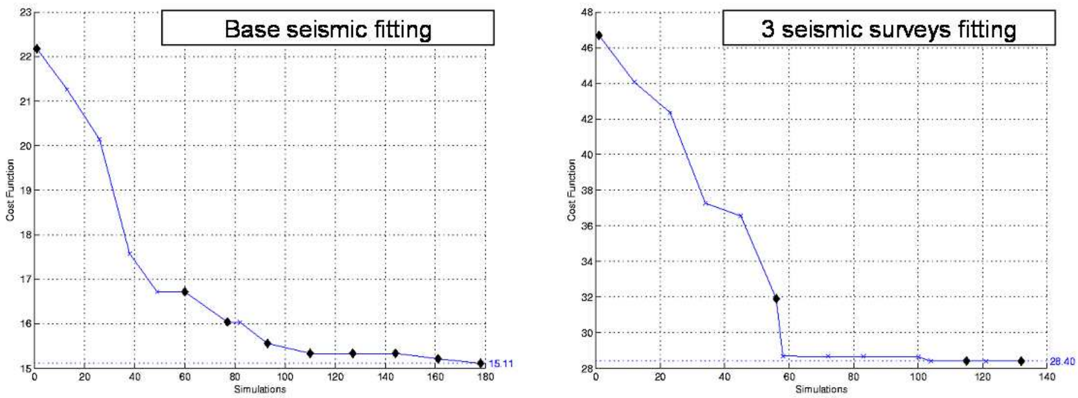

Figure 2: The cost function versus the simulation numbers is displayed for the two steps of the application: the first step corresponds to the fitting of base seismic (left) and the second step the fitting of the two monitor seismic surveys (right). These results have been obtained with BFGS option of SQPAL.

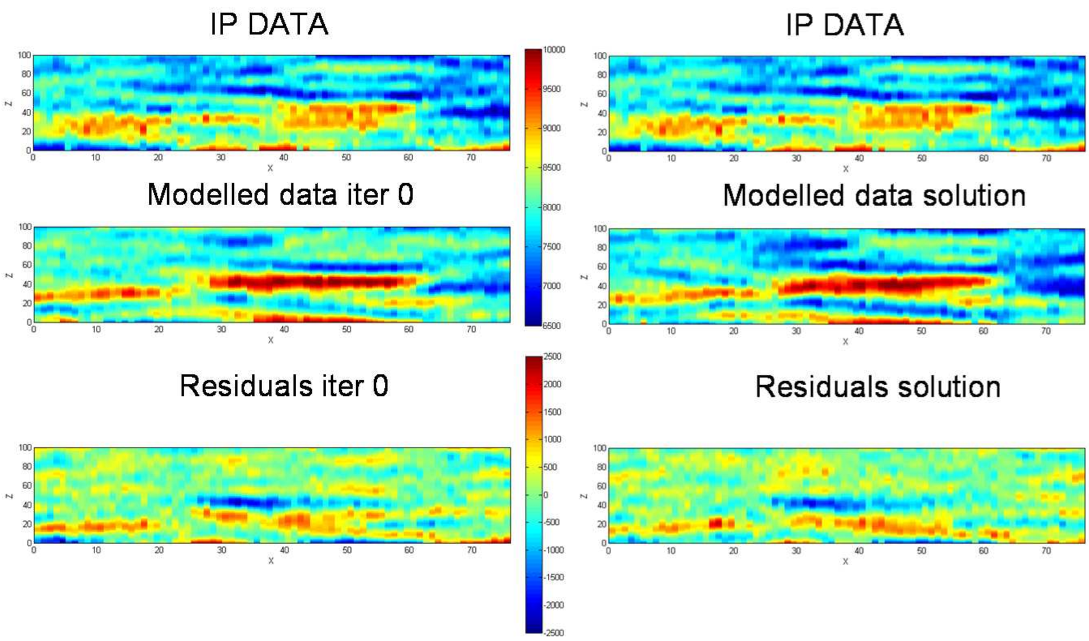

Figure 3: Results obtained by SQPAL BFGS option - P impedance of base seismic dataset (top), modeled P impedance for initial (middle left) and solution parameters (middle right) and the P impedance residuals for initial (bottom left) and solution parameters (bottom right).

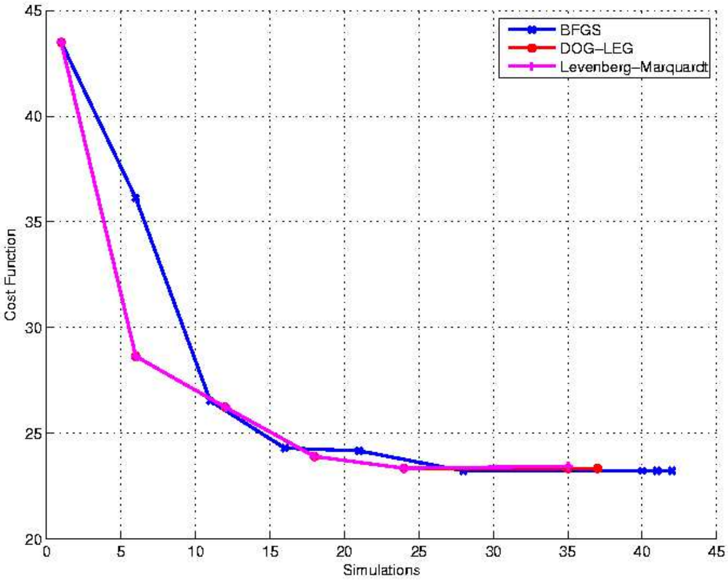

Figure 4: Cost function versus simulation numbers for 3 optimizations performed, respectively, with BFGS method with line search globalization (in blue), with Levenberg-Marquardt Gauss-Newton method (in pink) and with Dog-Leg trust region method (in red). Each tag on the curves indicates one nonlinear iteration. This application is the fitting of the base seismic with 4 parameters.

5. Conclusions

IFP SQPAL package has given promising results on the presented 2D reservoir characterization problem: the results obtained with BFGS and Gauss-Newton options are similar but the BFGS method needs only the estimation of the gradient of the objective function whereas the Gauss-Newton method requires the estimation and the storage of the Jacobian matrix which is an obstacle for application on large datasets. A special care has been put on the choice of the perturbation step size for numerical gradient computation with finite differences. A study of surrogate optimization techniques as Derivate Free Optimization approach proposed by Conn et al.(2000) and Powell(2000) is in progress in order to limit the number of expensive simulations.

Acknowledgments

The authors would like to thank Frédéric Roggero and Elodie Tillier for their advice for the reservoir characterization application and for providing the test case.

References

Berthet, P., Cap, J., Ding, D. Y., Lerat, O., Roggero, F., & Screiber, P.-E., 2007. Matching of production history and 4d seismic data: application to the girassol field, offshore angola, in SPE Annual technical conference and exhibition, *p. SPE 109929.*

Biggs, M. C., 1972. Constrained minimization using recursive equality quadratic programming, pp. 411–428, Academic Press, London, New York, F.A. Lootsma, ed.

Bonnans, J. F., Gilbert, J. C., Lemaréchal, C., & Sagastizábal, C. A., 2006. Numerical Optimization–Theoretical and Practical Aspects, Springer, Berlin.

Conn, A. R., Gould, N. I. M., & Toint, P. L., 2000. Trust-region methods, MPS-SIAM Series in Optimization.

Delbos, F. & Gilbert, J. C., 2005. Global linear convergence of an augmented lagrangian algorithm for solving convex quadratic optimization problems, J. of Convex Analysis, **12***, 45–69.*

Delbos, F., Gilbert, J. C., Glowinski, R., & Sinoquet, D., 2006. Constrained optimization in seismic reflection tomography: a gauss-newton augmented lagrangian approach, Geophys. J. Int., **164***, 670–684.*

Feraille, M., Roggero, F., Manceau, E., Hu, L. Y., Zabalza-Mezghani, I., & Reis, L. C., 2003. Application of advanced history matching techniques to an integrated field case study, in SPE Annual technical conference and exhibition*, 5-8 October, Denver, Colorado.*

Fornel, A., Noetinger, B., & Roggero, F., 2007. Using time domain seismic attributes for history matching, in IEA - International workshop and symposium of the International Energy Agency*, 4-7 September, Vedbaek, Denmark.*

Gill, P. E., Murray, W., & Saunders, M. A., 2005. Snopt: An sqp algorithm for large-scale constrained optimization, SIAM Review, **47***, 99–131.*

Gould, N. I. M. & Toint, P. L., 2000. SQP method for large scale nonlinear programming, pp. 149–178, Kluwer Academic, Boston, MA.

Han, S. P., 1976. Superlinearly convergent variable metric algorithms for general nonlinear programming problems, Math. Program., **11***, 263–282.*

Hu, L., Blanc, G., & Noetinger, B., 2001. Gradual deformation and iterative calibration of sequential stochastic simulations, Mathematical Geology, **33, No. 4**.

Nocedal, J. & Wright, S. J., 1999. Numerical optimization, Springer, New York.

Powell, M. J. D., 1978. A fast algorithm for nonlinearly constrained optimization calculations, pp. 144–157, Springer-Verlag, Berlin.

Powell, M. J. D., 1978. The convergence of variable metric methods for nonlinearly constrained optimization calculations, vol. 3, pp. 27–63, Special Interest Group Math. Programming, University of Wisconsin, Madison, WI, 1977, New York.

Powell, M. J. D., 2000. Developments of newuoa for unconstrained minimization without derivatives, Tech. Rep. Report No. DAMTP 2007/NA05, Technical Report, Department of Applied Mathematics and Theoretical Physics, University of Cambridge, England.

Wilson, R. B., 1963. A simplicial algorithm for concave programming, Ph.D. thesis, Graduate School of Business Administration, Harvard University, Cambridge, MA, USA.