

# AUTOMATED PARAMETER ESTIMATION AND SENSITIVITY ANALYSIS: IMPLEMENTATION ISSUES (1ST REPORT)

HECTOR KLIE<sup>1</sup>, MIGUEL ARGÁEZ<sup>2</sup>, LETICIA VELÁZQUEZ<sup>2</sup>, CARLOS QUINTERO<sup>2</sup>  
AND MARY F. WHEELER<sup>1</sup>

## CONTENTS

1. Introduction	1
2. Problem Formulation	3
2.1. Parameter Estimation Problem	3
2.2. Objective Function	4
3. Optimization Framework	4
3.1. Simultaneous Perturbation Stochastic Approximation (SPSA)	4
3.2. Newton-Krylov Interior-Point (NKIP)	4
3.3. Global Levenberg-Marquardt (GLM)	6
3.4. Surrogate Model Approach	8
3.5. The SPSA-GLM Approach	8
3.6. Implementation Details	9
4. Conclusions	10
Acknowledgements	11
References	11

**ABSTRACT.** The present report establishes the computational issues that will be considered for the implementation of hybrid optimization approaches oriented to automated parameter estimation problems. The proposed hybrid optimization approaches are based on the coupling of the Simultaneous Perturbation Stochastic Approximation (SPSA) approach (a global and derivative free optimization method) with two local and derivative dependent optimization methods: (1) a globalized Newton-Krylov Interior-Point algorithm (NKIP) and, (2) the global Levenberg-Marquardt (GLM) method. The first coupling will imply the generation of a metamodel or surrogate model that will allow to incorporate derivative information on a simpler representation of the original problem. The second type of coupling assumes that there is some derivative information available but its utilization is postponed until the SPSA algorithm has made sufficient progress towards the solution. At that point, the GLM takes over the computation to speedup the parameter estimation process. This preliminary description will set the ground for the implementation of a Matlab prototypes on these ideas.

**Keywords:** Parameter estimation, inverse problems, global optimization, SPSA, local optimization, interior-point methods, nonlinear programming, Newton method, nonlinear least-squares, Levenberg-Marquardt method, large-scale optimization.

## 1. INTRODUCTION

Finding a global optimal solution in parameter estimation is a challenging task due to the fact that data and models are usually nonlinear, non-smooth and subject to different source of errors. Major DOD environmental examples include

---

*Date:* August 28, 2006

<sup>1</sup> The University of Texas at Austin, TX.

<sup>2</sup> The Universidad de Texas at El Paso, TX.

obtaining subsurface properties from pressure head and saturation observations in contaminant remediation sites, estimating temporal and spatial coefficients for shallow water models from observed bathymetry data, and fish prediction behavior. All these problems require substantial computational resources and efforts to solve coupled nonlinear partial differential equations for different sets of parameter configurations. Moreover, in many cases the number of these parameters may be significantly large since they generally depend on a prescribed level of spatial and temporal resolution.

Most successful approaches in global optimization use some stochastic or heuristic elements that do not depend on derivative information; e.g., genetic algorithms, tabu search, neural networks, and simulated annealing [10]. Despite inherent costs, their success resides in making a broad and systematic exploration of the search domain, although this may vary with the nature of the problem. Therefore, it is useful to obtain several local optimal solutions and compare them before selecting a global one. This has proven to be an effective way to cope with the underlying uncertainty. However, the performance of these algorithms depends significantly on the choice of free parameters and this task tends to be less obvious at large-scale since the number of function evaluations grow dramatically with the number of parameters.

On the other hand, the availability of derivative information such as gradient, Jacobian or Hessian operators allow for achieving higher rates of convergence towards a local optimum. Among several derivative dependent methods, Newton method has been proven to be the most effective. However, the use of derivatives can be prohibitively expensive and even unfeasible to compute. In that sense, nonlinear methods such as Broyden (secant type) and Levenberg-Marquardt (regularization type) have been found to be very effective [7]. In any case, the inclusion of physical bound constraints is a powerful mean to reduce the search space and speedup the computation. In this setting, interior-point methods are perhaps the method of choice to generate iterates within the interior of a feasible region.

Standard optimization software in environmental parameter estimation is mainly based on nonlinear squares solution approaches such as Levenberg- Marquardt method; see e.g. [1]. Despite that this implementation may be robust for a good initial guess due to a regularizing scheme based on the combination of a Tikhonov procedure with a truncated singular value decomposition method, the method may not be affordable at large-scale in the absence of sufficient regularity. This is due to the fact that the method relies on expensive direct factorization methods and on the assumption of derivative information. The focus of the present work is to extend the capabilities of this approach by combining global stochastic optimization and deterministic local optimization methods (i.e., hybrid optimization methods). The former will allow for a better exploration of the parameter space without relying on derivatives, whereas the second approach will allow to speedup the parameter estimation process on a neighborhood of the optimal solution.

The global stochastic optimization is performed by means of the simultaneous perturbation stochastic approximation (SPSA) algorithm [12]. The philosophy of this algorithm is to perform random simultaneous perturbations of all model parameters to generate a descent direction at each iteration. The magnitude of this

perturbation and the step direction is decreased as the iteration proceeds. This allows for narrowing the search on promising optimal solution regions. This algorithm has been successfully used in several application areas, see [9, 11, 12].

We identify two alternatives to generate hybrid optimization algorithms: (1) metamodel-based, which implies to use the sampling generated by SPSA, interpolate the resulting response surface, generate the metamodel (or surrogate model), establish some additional constraints (if necessary) and find the optimal solution in such metamodel; and, (2) gradual gradient-based, which implies to perform a set of SPSA iterations by gradually reducing the degree of stochasticity and switching eventually to a global Levenberg-Marquardt (GLM) type of method.

The present report is structured as follows. We provide a description of the optimization problem to be solved in terms of parameter estimation. We follow with a detailed description of the three main algorithms under study: SPSA, NKIP and GLM. Additionally, we focus on the description of the metamodel and gradual gradient-based approach for combining these methods. Some implementation details are discussed in Section 4. The report ends with some concluding remarks.

## 2. PROBLEM FORMULATION

We consider the nonnegative constrained minimization problem in the form of

$$\begin{aligned} & \text{minimize} && f(x) \\ & \text{subject to} && x \geq 0, \end{aligned} \tag{1}$$

where  $x \in \mathbb{R}^n$  and  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  is a twice continuously differentiable function. The global optimization problem is to find a point  $x^*$  such that  $f(x^*) \leq f(x)$  for all  $x \in \mathbb{R}^n$ .

We are interested in finding global minima of  $f$  that may possess many local minima for the class of constrained minimization problems (1). In order to simplify the presentation of the problem, we look at the unconstrained case. Our goal is to use SPSA as the sampling device to perform a global search of the parameter space and switch to either NKIP or GLM. If necessary, imposition of functional constraints are possible using the NKIP approach.

**2.1. Parameter Estimation Problem.** A general parameter estimation problem can be written as a system of equations

$$G(x) = d, \tag{2}$$

where  $G$  is a nonlinear function that, in most cases, may involve the discrete solution of a set of ordinary or partial differential equations,  $x$  is a vector expressing the model parameters (i.e., problem coefficients), and  $d$  is a set of discrete data observations (i.e., outputs).

In general, the data  $d$  predicted by the function  $G$  does not match with the observed data in the lab or field, say  $d_T$ . This is mainly due to noise or numerical errors. Thus, in parameter estimation, one is interested in finding the vector  $x$  that minimizes the mismatch between  $d$  and  $d_T$  under some norm. However, one of the major challenges of ill-posed problems is that there may be infinite models satisfying the same degree of mismatch even in the absence of noise or numerical errors.

**2.2. Objective Function.** The aforementioned mismatch between numerical observations and real observations can be quantified by an objective function  $f$  defined in a weighted least squares sense

$$f(x) = (d - G(x))^t C_d^{-1} (d - G(x)). \quad (3)$$

Here, the observation covariance matrix  $C_d$  represents the error distribution (or weight) in the data. In the particular case of groundwater problems, these model parameters depend on location (permeability, porosity, rock compressibility, initial or boundary conditions). The vector of measurements  $d_T$  may be obtained at different locations and time intervals. In this case, (3) may be seen as a mismatch instance in a summation of several similar mismatch terms over time.

### 3. OPTIMIZATION FRAMEWORK

**3.1. Simultaneous Perturbation Stochastic Approximation (SPSA).** The SPSA for either equation (1) or (3) is defined by the following recursion for the parameter vector  $x$ :

$$x_{k+1} = x_k - a_k \hat{g}_k(x_k), \quad (4)$$

where  $a_k$  is a positive scalar that monotonically decreases with respect to  $k$ , and  $\hat{g}_k(x_k)$  is a stochastic approximation to the gradient given by a simultaneous perturbation of all elements of  $x_k$ , that is,

$$\hat{g}_k(x_k) = \frac{1}{2c_k} [f(x_k + c_k \Delta_k) - f(x_k - c_k \Delta_k)] \Delta_k^{-1}, \quad (5)$$

where  $c_k$  is also a positive scalar that monotonically decreases with respect to  $k$ ,  $\Delta_k$  is a vector consisting of  $\{-1, 1\}$  values randomly generated with a Bernoulli distribution and  $\Delta_k^{-1}$  stands for the componentwise reciprocal of each of the entries of  $\Delta_k$ . The parameters  $a_k$  and  $c_k$  are chosen to ensure asymptotic convergence of the algorithm; for more details and pointers on SPSA see [12].

The SPSA algorithm has received considerable attention for global optimization problems where it is difficult or impossible to compute first order derivative information associated with the problem. Clearly, as indicated by (5), this algorithm only requires two function evaluations per iteration independently of the parameter space size to generate a stochastic descent direction for (4). The SPSA algorithm turns out to be very attractive for large-scale implementations and for a “black-box” treatment of complex multiphysics models where derivatives with respect to the parameters are a challenging and even unfeasible task.

**3.2. Newton-Krylov Interior-Point (NKIP).** We describe the NKIP algorithm with a linesearch globalization strategy for solving nonnegative constrained minimization problems. This algorithm has been developed for obtaining an optimal solution for large scale and degenerate problems from any initial point. The most expensive part of the algorithm is the need for solving iteratively a least squares problem, but the rest of the operations are based on matrix-vector multiplications.

The Lagrange function associated with problem (1) is given by

$$\ell(x, y, z) = f(x) - x^T z$$

where  $z \geq 0 \in \mathbb{R}^n$  is the Lagrange multipliers associated with the inequality constraints.

In order to promote the global convergence of the Newton interior-point method and keep iterates away from the boundary of the feasible region, the algorithm is based on the perturbed Karush-Kuhn-Tucker (KKT) conditions given by

$$F_\mu(x, z) \equiv \begin{pmatrix} \nabla_x \ell(x, z) \\ XZe - \mu e \\ (x, z) \geq 0, \end{pmatrix} = 0 \quad (6)$$

where  $\nabla_x \ell(x, z) = \nabla f(x) - z$  is the gradient of the Lagrangian function,  $X = \text{diag}(x)$ ,  $Z = \text{diag}(z)$ ,  $e = (1, 1, \dots, 1)^T \in \mathbb{R}^n$ , and the perturbation parameter  $\mu \geq 0$ . For  $\mu = 0$ , these conditions are merely the KKT conditions associated to problem (1).

A point  $(x, z)$  is said to be an interior point if  $(x, z) > 0$ . The methodology of interior-point methods entails to keep this property during the entire procedure except at an optimal solution.

For a given  $\mu > 0$ , the Newton step  $(\Delta x, \Delta z)$  at the interior point  $(x, z)$  is obtained as the solution of the following nonsymmetric linear system:

$$\begin{pmatrix} \nabla^2 f(x) & -I \\ Z & X \end{pmatrix} \begin{pmatrix} \Delta x \\ \Delta z \end{pmatrix} = \begin{pmatrix} e_d \\ e_c \end{pmatrix}. \quad (7)$$

In the right hand side vector,  $e_d = -\nabla f(x) + z$  represents the gradient of the Lagrange function, and  $e_c = \mu e - XZe$  the perturbed complementarity condition associated to the problem for a positive perturbed parameter  $\mu$ . To guarantee that the updated iterate is an interior point, a steplength  $\alpha \in (0, 1]$  is selected such that  $(x, z) + \alpha(\Delta x, \Delta z) > 0$ .

Under this framework, we state the idea behind the path-following interior-point Newton algorithm for nonnegative constrained minimization.

**Definition 3.1.** (*Path-Following Strategy*). *For  $\mu > 0$ , and working from the interior  $(x, z) > 0$ , apply a linesearch Newton's method to the perturbed KKT conditions until the iterate satisfies a specified proximity to the perturbed complementarity condition and hopefully reduces the objective function. Then decrease  $\mu$ , specify a new proximity, and repeat the process. Under appropriate conditions, an optimal solution will be obtained as  $\mu$  approaches zero.*

In order to implement this strategy, we need to state three primary issues among others: First the centrality region to be followed, second how close we want to approach a centrality region, and third which merit function to be used for obtaining such proximity. These issues are addressed in the papers of Argáez and Tapia [3], and Argáez, Saéñz, and Velázquez [2]).

The interest in larger-scale applications of interior-point methods brought about the idea of performing inexact Newton steps of (7), and coupling it with the globalized path-following strategy. The use of system (7) has some disadvantages: it is nonsymmetric, and the condition of non-singularity in real applications is not satisfied. Therefore, a current challenge is to find such a robust iterative method using this system or a decoupled system. Most of these advances have relied on the use of Krylov subspace methods [5, 6].

In particular, we present an algorithm for obtaining a Newton-Krylov direction. In line with this objective the above linear system is decouple in the following manner: Once the direction  $\Delta x$  is known, then from the second block of equations

we have

$$\Delta z = X^{-1}(\mu e - XZe - Z\Delta x). \quad (8)$$

Then substituting  $\Delta z$  in the first block of equations we have

$$(\nabla^2 f(x) + X^{-1}Z)\Delta x = \mu x^{-1} - \nabla f(x) \quad (9)$$

where  $x^{-1}$  denotes the vector in  $\mathbb{R}^n$  whose components are the reciprocals of the components of  $x$ . Observe that once  $\Delta x$  is known, then  $\Delta z$  is found using simple arithmetic operations. Therefore, the difficulty lies in solving the linear system (9). The direction  $(\Delta x, \Delta z)$  given by (9) and (8) is the same Newton direction of (7) if the linear system (9) is solved exactly.

At this moment we have two possibilities for obtaining an approximate solution of large linear systems of equations. The first possibility is to consider an iterative procedure for the entire system (7), and the second one is to consider an iterative procedure for obtaining an approximate solution of the linear system (9) and then define an inexact Newton direction using (8). We discard the first possibility since the matrix associated with this linear system is highly indefinite, and to find an effective preconditioner for the entire system is hard, therefore, defining a robust iterative method represents a major challenge. Instead we turn our attention to the decoupled procedure.

Even though we still require to solve a linear system, this new system of equations has several advantages over system (7). First of all, the coefficient matrix in system (9) is always symmetric, and its size is half of the size of system (7). Now, since we do not have any control over the positive definiteness of the second order information of the problem,  $\nabla^2 f(x)$ , then by adding a diagonal positive matrix,  $X^{-1}Z$ , we increase the likelihood that the matrix  $\nabla^2 f(x) + X^{-1}Z$  will be positive definite. This fact is of great importance because it yields the implementation of an iterative procedure as the conjugate gradient method for obtaining an approximate solution. Second, we only need to find a good preconditioner for matrices that have high possibilities to be positive definite.

Since the case where we have a non positive definite matrix is still possible, then we use the conjugate gradient algorithm to solve the linear system

$$H\Delta x = b,$$

where  $H = \nabla^2 f(x) + X^{-1}Z$  and  $b = \mu x^{-1} - \nabla f(x)$ , and we consider the Steihaug's approach to find an approximate solution.

**Remark 3.1.** *In the event  $H$  is not positive semidefinite on  $\mathcal{N}(A)$ , other Krylov subspace methods (GMREs, bi-conjugate gradient, etc) can be implemented.*

**3.3. Global Levenberg-Marquardt (GLM).** Velázquez, Phillips, Tapia and Zhang [13] proposed a global Levenberg-Marquardt algorithm, based on the concept of selective minimization introduced in Zhang, Tapia and Velázquez [14], for approximating global minima of zero or small residual nonlinear least-squares problems. Under these conditions, the algorithm was locally attracted to minima with sufficiently low function values, but often repelled from local minima with high function values. We are proposing to take advantage of this property that could be a useful tool for constructing or improving global minimization techniques in parameter estimation problems. Our intent is to use this algorithm coupled with SPSA, and do the refinement stage (if necessary) by using NKIP.

We rewrite the weighted nonlinear least-squares problem (3) as

$$\min f(x) = \frac{1}{2} \sum_{i=1}^m c_i^{-1} r_i^2(x), \quad (10)$$

where the errors or residuals  $r(x) = [r_1(x), r_2(x), \dots, r_m(x)]^T \in \mathbb{R}^m$  for some  $m > n$ , and  $r_i(x) = G(x, t_i) - d_i, i = 1, 2, \dots, m$ . The gradient of  $f(x)$  is given by  $\nabla f(x) = J(x)^T C^{-1} r(x)$  where  $J(x) \in \mathbb{R}^{m \times n}$  is the Jacobian matrix  $[J(x)]_{i,j} = \frac{\partial r_i(x)}{\partial x_j}, i = 1, \dots, m, j = 1, \dots, n$ .

A problem for which  $r(x^*) = 0$ , and hence  $f(x^*) = 0$ , is called a zero-residual problem. A problem is called a small-residual problem if these quantities are relatively small. (Obviously, smallness is a problem-dependent concept.)

The classical multi-start strategy for solving nonlinear least-squares problems combines random sampling with a local minimization algorithm. Each randomly sampled point is a starting point from which one seeks a local minimizer via a local minimization algorithm. At the end, the local minimizer with lowest function value is taken as an approximation to the global minimizer. A variable steplength approach, such as line search or a trust region technique, is implemented in order to guarantee that at each iteration the function value is decreased sufficiently and enhance convergence to a local minimizer (see, for example, Björck [4], and Dennis and Schnabel [7]).

On the contrary, we propose to implement the GLM algorithm without enforcing descent. In this strategy, a unit steplength is always used regardless whether the function value is decreased or not. This feature prevents the algorithm from getting trapped easily at local minima, while still letting the algorithm converge to global minima or local minima of very small function values. As a result, the algorithm will more likely locate a global minimum instead of a local minimum of large function values. In our computational experience, we have indeed observed that iterates were often repelled from undesirable local minima which classical multi-start would force iterates to converge to, but still attracted to global minima or those local minima with close to zero function values.

We now state an algorithmic framework based on the selective minimization approach:

**Algorithm GLM:**

1. Set upper bounds for the number of random initial points  $l_{max}$  and iterations  $k_{max}$  allowed.
2. Choose a tolerance  $\epsilon > 0$ , a parameter  $\tau > 0$ , and set  $l = 0$ .
3. **While** (  $f(x_k) < \epsilon$  or  $l > l_{max}$  ) **do**
  - 3.1 Select a random starting point  $x_0$  and set  $k = 0$ .
  - 3.2 Increment  $l = l + 1$ .
  - 3.3 **While** (  $\|\nabla f(x_k)\| \leq \epsilon$  or  $k > k_{max}$  ) **do**
    - 3.3.1 Set  $\mu_k = \tau \|r(x_k)\|$ .
    - 3.3.2 Compute  $s_k = -(J(x_k)^T J(x_k) + \mu_k I)^{-1} \nabla f(x_k)$ .
    - 3.3.3 Set  $x_{k+1} = x_k + s_k$  and increment  $k$ .

**End**

**End**

Some observations for implementing this algorithm are:

- (1) Algorithm GLM is attracted to stationary points with a zero residual level. In particular, it is attracted to global minimizers whenever

$$\min f(x) \leq \epsilon \text{ for sufficiently small } \epsilon.$$

- (2) Algorithm GLM is not attracted to maximizers and saddle points.
- (3) Algorithm GLM compared favorably with other global optimization methods (Classical Multistart Damped Levenberg-Marquardt Method, Simulated Annealing, and Tunneling Methods) for solving high dimensional least-squares problems.
- (4) In general, the set  $\{x \in \mathbb{R}^n : \nabla f(x) = 0\}$  changes with  $C^{-1}$ ; however for a zero-residual problem the set of global minimizers of the weighted least-squares problem is invariant with respect to nonzero weights.
- (5) The proper choice of weights for a zero-residual least squares problem can make the problem easier to solve by local techniques.

**3.4. Surrogate Model Approach.** The parameter estimation is first carried out by means of the SPSA algorithm. This process may be performed by starting with different initial guesses. This not only augments the chances for finding a global optimal solution, it also allows for a rich sampling of the parameter space. Moreover, the search performed by the SPSA algorithm guides the sampling toward promising regions containing a global solution (“hot spots”). For small to moderate size problems (i.e., less than a one hundred state variables) thousands of computations may be affordable in a few hours. Below, we indicate other possibilities for decreasing the computational overburden.

Based on the mapping between parameters and the objective function provided by SPSA samples, we proceed to interpolate the response surface and construct a surrogate representation of the original model in a section of the original search space. This can be performed in different ways, e.g., kriging, radial basis functions, regression analysis, or using artificial neural networks. In our particular case, we are interested in finding an analytical representation of the original function of quadratic or cubic order. Hence, this provides us with an approximate smooth representation of the objective function within the region of the most promising optimal solutions explored by SPSA.

Once the surrogate model is constructed, we can proceed to solve a problem of the nature of (1) with the NKIP algorithm. Derivatives are readily available from the analytical representation of the surrogate model. This provides means to study the objective function sensitivity with respect to parameter estimates in the vicinity of explored regions. The NKIP approach allows for further refining the solution yield by SPSA. Moreover, points evaluated in the surrogate model are validated against the original model (e.g., groundwater flow simulator, integral equation for wave propagation, etc). Figure 1 illustrates how the surrogate model establishes the bridge between the SPSA and the NKIP algorithms. The reader interested in the construction of surrogate models is referred to [8].

**3.5. The SPSA-GLM Approach.** This approach can be employed when derivative information is available or can be efficiently estimated from the SPSA sampling process. The idea is to postpone the construction and use of derivative operators during the global search process in order to ameliorate the computational cost of the optimization. As described above, the SPSA is a very efficient algorithm that requires at most two function evaluations per iteration. Once the SPSA suffers

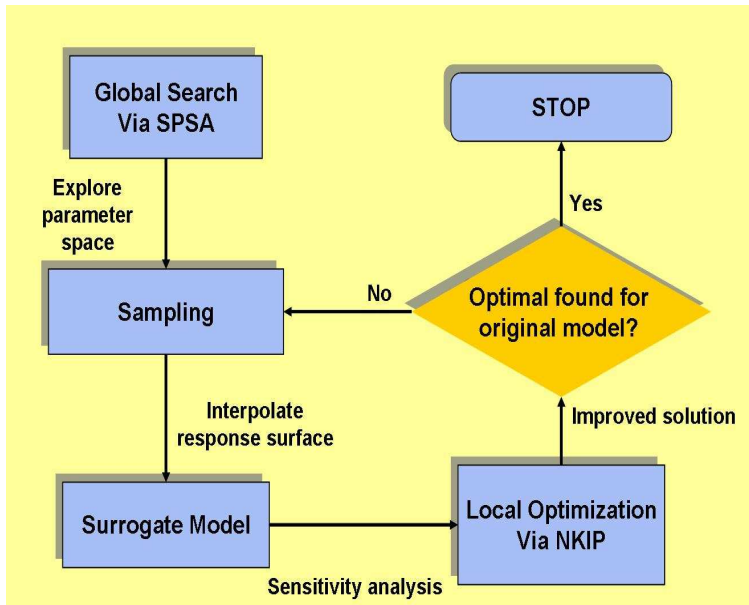


FIGURE 1. The surrogate model approach.

from slow convergence or even stagnation, the process is switched to the GLM method. The idea is that the SPSA may be trapped in a local minima, so inclusion of derivatives may speed up the convergence rate. Note that the GLM is not a strict local optimization algorithm, so there is still a chance that the solution could be improved substantially. One may think of a parameter space characterized by few noisy valleys with many surrounding bumps.

Figure 2 shows how this aforementioned process is tied together. The figure suggests to rely on the SPSA algorithm as much as possible before using the GLM approach. The challenge here is to determine when is computationally convenient to switch to the GLM.

**3.6. Implementation Details.** There are several aspects that have not been fully discussed in the previous sections and that are critical for a robust and efficient implementation of the procedures described above. In particular, we identify five important ones:

- scaling
- parameterization,
- regularization,
- preconditioning, and
- stopping criteria.

Scaling seeks to remove the significant difference in magnitude among dependent variables and independent variables. This may be solved by either changing the units in the problem (e.g., meters to miles, Pascals to psi) or by changing the variable distribution (e.g., work in terms of the log of the variable instead of its original distribution). Parameterization allows one to decrease the parameter space

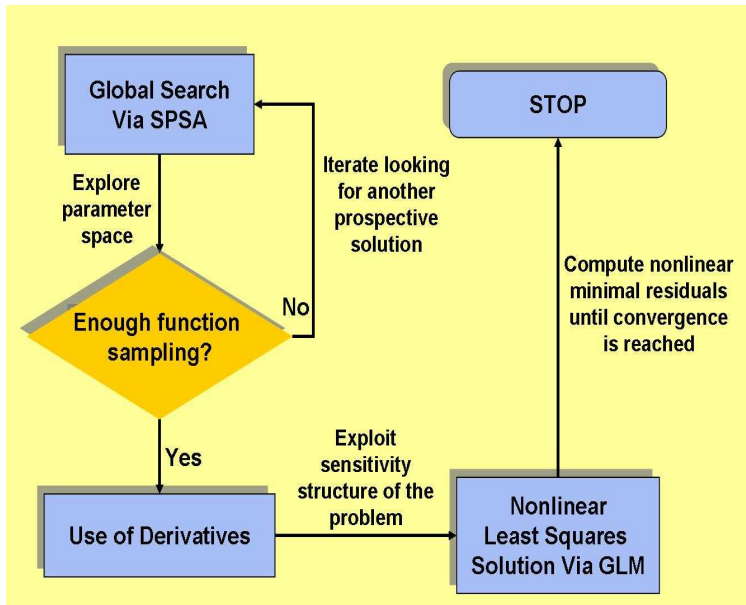


FIGURE 2. The SPSA-GLM approach.

in terms of a new basis. We consider that methods such as multiscale, model reduction and image compression are very effective to reduce the computational cost of the optimization (see e.g., [11]). Regularization is the process of smoothing the parameter space in order to balance the simultaneous decrease of the errors and residuals associate with the objective function. Regularization is mostly achieved by constraining the problem or avoiding the contamination of the solution with numerical errors due to the ill-posedness. Preconditioning focus in accelerating the convergence of iterative methods (mainly Krylov based, such as the conjugate gradient of GMRES). Finally, stopping criteria are conditions to ensure the satisfactory termination of the algorithms. Obviously, the coupling of global and local algorithms implies a careful definition of stopping criteria.

The aforementioned details will be considered at different degree of complexity in our implementations. In most cases, we may follow conventional rules employed in most well-established optimization software. Numerical comparisons will be designed to test standard benchmark cases and numerical software. Nevertheless, a list of major recommendations will be provided for the sake of improving the proposed implementations. Most of our developments will be in Matlab to facilitate the technology transfer to DOD users.

#### 4. CONCLUSIONS

The present report outlines the main theoretical and computational issues for the implementation of a new family of optimization algorithms. The development of these algorithms is motivated by the need to explore more efficient, robust ways to perform parameter estimation. Although we have not stressed other computational aspects such as parameterization, regularization and preconditioning (see section

3.6), they would be somehow considered in the underlying implementation in its more conventional forms. They deserved a separate study that should complement the aforementioned ideas.

The resulting implementation should provide further insights to DOD users for performing parameter estimation process in a routine and large-scale basis. The combination of global and local optimization methods as suggested by the authors is oriented to exploit the best of both approaches.

#### ACKNOWLEDGEMENTS

The authors thanks the grant DOD-PET Project UR-EQM-06-03 for the support of this work.

#### REFERENCES

- [1] *PEST: Model-Independent Parameter Estimation, User Manual: 5th Edition*. WaterMark Numerical Computing, 2005.
- [2] M. Argáez, R. Sáenz, and L. Velázquez. A trust-region interior-point algorithm for nonnegative constrained minimization. Technical report, Department of Mathematics, The University of Texas at El Paso, 2003.
- [3] Miguel Argáez and R.A. Tapia. On the global convergence of a modified augmented lagrangian linesearch interior-point newton method for nonlinear programming. *J. Optim. Theory Appl.*, 114:1–25, 2002.
- [4] A. Björck. *Numerical Methods for Least Squares Methods*. SIAM, Philadelphia, 1996.
- [5] P. N. Brown and Y. Saad. Hybrid Krylov methods for nonlinear systems of equations. *SIAM J. Sci. Statist. Comput.*, 11:450–481, 1990.
- [6] P. N. Brown and Y. Saad. Convergence theory of nonlinear Newton-Krylov algorithms. *SIAM J. Optim.*, 4:297–330, 1994.
- [7] J. E. Dennis and R. B. Schnabel. *Numerical Methods for Unconstrained Optimization and Nonlinear Equations*. Prentice-Hall, Englewood Cliffs, New Jersey, 1983.
- [8] A.J. Keane and P.B. Nair. *Computational Approaches for Aerospace Design: The Pursuit of Excellence*. Wiley, England, 2005.
- [9] H. Klie, W. Bangerth, M. F. Wheeler, M. Parashar, and V. Matossian. Parallel well location optimization using stochastic algorithms on the grid computational framework. In *9th European Conference on the Mathematics of Oil Recovery (ECMOR)*. EAGE, Aug. 30 - Sept 2, 2004.
- [10] P.M. Pardalos and M.G.D. Resende, editors. *Handbook of Applied Optimization*. Oxford University Press, 2002.
- [11] A. Rodriguez, S.G. Thomas, H. Klie, and M. F. Wheeler. A multiscale and metamodel simulation-based method for history matching. In *10th European Conference on the Mathematics of Oil Recovery (ECMOR)*. EAGE, Sept 4 - Sept 7, 2006.
- [12] J. C. Spall. *Introduction to stochastic search and optimization: Estimation, simulation and control*. John Wiley & Sons, Inc., New Jersey, 2003.
- [13] L. Velázquez, G.N. Phillips, R.A. Tapia, and Y. Zhang. Selective search for global optimization of zero or small residual least-squares problems: a numerical study. *Computational Optimization and Applications: An international Journal*, 20(3):299–315, 2001.
- [14] Y. Zhang, R.A. Tapia, and L. Velázquez. On convergence of minimization methods: Attraction, repulsion and selection. *Journal of Optimization Theory and Applications*, 107(3):529–546, 2000.