Nonlinear Parameter Estimation by Global Optimization — Efficiency and Reliability

T. CSENDES

1. Introduction

Our original task [8] was to solve parameter estimation problems having very complex nonlinear objective functions with a relatively small number of parameters. Their evaluation was quite expensive: about 100 times as much CPU time is needed as to the standard test functions. The objective functions usually turned out to have a large number of local minima in the region of interest. Although we can compute these functions, at times we do not even know their explicit form. Thus, determination of the exact, analytical derivatives is impossible in such cases, and we are forced to use non-derivative techniques.

The literature on global optimization [5] suggested that the method of Boender et al. [2] was the most promising for our purposes. Although a later version of this algorithm [11] seemed to be more efficient, we did not implement this modification because it was less reliable.

In this paper we discuss the relationship between the nonlinear least squares problem and global optimization, and we deal with the efficiency and reliablity of the above global optimization method using a quasi-Newton procedure and a random walk direct search technique.

2. Nonlinear parameter estimation and global optimization

The nonlinear parameter estimation problem is usually given as

$$\min_{x} F(x) \tag{1}$$

where $F(x): \mathbf{R}^n \to \mathbf{R}$,

$$F(x) = \left(\sum_{i=1}^{m} (\bar{f}_i - f_i(x))^2 \right)^{1/2}$$

$\bar{f}_i \in \mathbf{R}$; $f_i(x): \mathbf{R}^n \to \mathbf{R}$, $i = 1, 2, \ldots, m$; $m > 0$ integer; $x \in S \subseteq \mathbf{R}^n$, where the region of interest S is a compact set. S is in most cases a hypercube $a_i \leq x_i \leq b_i$; $a_i, b_i \in \mathbf{R}$ $i = 1, 2, \ldots, n$. Thus the objective function $F(x)$ is of least squares type.

In solving (1), it is often supposed that $F(x)$ is unimodal (it has only one local minimum) or that a suitable starting point is at hand for the iterative solving algorithm [6]. Since we have found many practical nonlinear parameter estimation problems whose objective functions were not unimodal, we examine the relationship between the nonlinear least squares problem (1) and the global optimization problem:

Consider a compact set S in $\mathbf{R}^n$ and a not necessarily unimodal function $G(x)$: $\mathbf{R}^n \rightarrow \mathbf{R}$. The problem is to find a global minimizer $x^* \in S$ such that $G(x) \geq G(x^*)$ for all $x \in S$.

S is usually given by simple bounds on the parameters of $G(x)$:

$$a_i \leqq x_i \leqq b_i, \quad a_i, b_i \in \mathbf{R}, \quad i = 1, 2, ..., n$$

We found that the structure of $F(x)$ guarantees only the non-negativity of $F(x)$. More exactly:

Proposition. *For every non-negative real function* $G(x)$: $\mathbf{R}^n \rightarrow \mathbf{R}$, *positive integer* m *and real numbers* $\bar{f}_i \in \mathbf{R}$, $i = 1, 2, ..., m$, *there are real functions* $f_i(x)$, $i = 1, 2, ...$ *..., m such that*

$$F(x) = \left(\sum_{i=1}^{m} (\bar{f}_i - f_i(x))^2 \right)^{1/2}$$

and $G(x) = F(x)$ *for every* $x \in \mathbf{R}^n$.

For example, let $g_i(x)$, $i = 1, 2, ..., m$ be real, non-negative functions such that

$$G(x)^2 = \sum_{i=1}^{m} g_i(x)$$

There exist such functions $g_i(x)$, since $g_i(x) = G(x)^2/m$ is suitable, for instance. Then, let

$$f_i(x) = g_i(x)^{1/2} + \bar{f}_i.$$

Note that the functions $g_i(x)$ can be almost freely chosen, and in this way we can ensure further desirable properties of the functions $f_i(x)$. For example, when $G(x)$ is continuous, then all $f_i(x)$ can be continuous, too. On the other hand, for all sets of functions $f_i(x)$, $i = 1, 2, ..., m$ there obviously exists a real function $f(i, x)$: $\mathbf{R}^{n+1} \rightarrow \mathbf{R}$, so that $f_i(x) = f(i, x)$ for all $i = 1, 2, ..., m$; and $f(i, x)$ is even continuous in the variable i.

According to the Proposition, the objective function of a nonlinear parameter estimation problem can be any non-negative real function. Thus, a nonlinear parameter estimation problem can have an arbitrary large number (or even a continuum) of local minima. The structure of $F(x)$, i.e. the least squares form, results only in the non-negativity of $F(x)$, and not in any further regularity.

Since the global minimum of a well-posed global optimization problem is finite (e.g. $G(x^*) \in \mathbf{R}$), every such problem can be transformed with $G'(x) = = G(x) - G(x^*)$ and $S' = S$ to a problem having a non-negative objective function and the same structure of local minima. Thus, loosely speaking, every global optimization problem can be written in the form of a nonlinear parameter estimation problem with any m, and $\bar{f}_i$ values fixed in advance. This confirms the use of a global optimization algorithm to solve problems such as (1).

3. Implementations

We discuss here an algorithm to solve the global optimization problem defined in the previous section. In most cases, the result of a global optimization algorithm is only an approximation of the global optimum, though the precision of the modern sophisticated nonlinear optimization methods approaches that of the given computer.

The global optimization method of Boender et al. [2] has been implemented in two versions. These have the same structure, the only difference between them being the local search procedure (an algorithm to find a local minimizer) used: a quasi-Newton procedure with the DFP update formula [6] and a random walk direct search method UNIRANDI [9, 12]. In the following these algorithms are denoted by A and B, respectively. Both are derivative-free, i.e. they do not use the partial derivatives of the objective functions. Evaluation of the latter would be difficult or even impossible in the case of our original problem [8]. UNIRANDI proved to be robust but inefficient, whereas the quasi-Newton method was rather sensitive to the initial points but more accurate [3]. The global optimization method and UNIRANDI were implemented by using solely [2] and [12].

The global optimization algorithm discussed in this paper can be described concisely as follows:

Step 1. Draw N points with uniform distribution in S, and add them to the current sample C. Construct the transformed sample T by taking the γ percent of the points in C with the lowest function values.

Step 2. Apply the clustering procedure to T. If all points of T can be assigned to a cluster, go to Step 4.

Step 3. Apply the local search procedure to the points in T not yet clustered. Repeat Step 3 until every point has been assigned to a cluster. If a new local minimizer has been found, go to Step 1.

Step 4. Determine the smallest local minimum value found, and stop.

The local search procedure mentioned here is either the quasi-Newton method or UNIRANDI. We chose the single linkage clustering procedure as being the more promising of the two discussed in [2]. The aim of this procedure is to recognize those sample points starting from which the local search would possibly result in an already found local minimizer. Clusters are grown around seed points (local minimizers or such points of the local search procedure from which an already known local minimum was reached). A distance $d(\underline{x}, \underline{x}')$ is defined [2] for the clustering between two points $\underline{x}$ and $\underline{x}'$ in the neighbourhood of a local minimizer $\underline{x}^*$ by

$$d(\underline{x}, \underline{x}') = ((\underline{x} - \underline{x}')^T H(\underline{x}^*)(\underline{x} - \underline{x}'))^{1/2}.$$

The quasi-Newton method of algorithm A gives a good approximation to the Hessian $H(\underline{x}^*)$ of the objective function. In the case of UNIRANDI the identity matrix replaces $H(\underline{x}^*)$ (cf. [2]). A new point x is added to a cluster if there is a point x' in this cluster for which

$$d(\underline{x}, \underline{x}') \leq \left[\frac{\Gamma\left(1 + \frac{1}{2}n\right) |H(\underline{x}^*)|^{1/2} \dot{m}(S)}{\pi^{n/2}} (1 - \alpha^{1/(N'-1)}) \right]^{1/n}$$

where $|H(\underline{x}^*)|$ denotes the determinant of $H(\underline{x}^*)$, $m(S)$ is a measure of the set S, N' is the total number of sampled points, and $0 < \alpha < 1$ is a parameter of the clustering procedure [2].

The two most important changes in the original algorithm are as follows:

1. We do not use a steepest descent step to transform the current sample. Its efficiency was examined in the early phase of the implementation, and it turned out to be omittable.

2. The parameters of the objective function are scaled [6] by the global optimization subroutine with the transformations

$$x_i' = \frac{2x_i - a_i - b_i}{b_i - a_i} \quad i = 1, 2, \ldots, n.$$

This can be done, of course, without using the explicit form of the objective function. The scaling does not have much effect on the efficiency of algorithms A and B in the case of the test functions. On the other hand, it is indispensable for the solution of practical problems.

The result of the implementation was a FORTRAN subroutine of just over 400 program lines, occupying 44 kilobytes of core space (without the local search routines). It serves to solve global optimization problems of up to 15 parameters. The program documents its progress, and when the problem is solved it makes a list of local minima with increasing function values.

4. Efficiency

The numerical tests were carried out on a ROBOTRON R55M computer. The programs were coded to use single precision arithmetic (with 7.2 decimal digits). The standard time unit $\left(1000 \text{ evaluations of the S5 function at } \underline{x}^T = (4.0, 4.0, 4.0, 4.0)^T\right)$ was measured ten times. The average of these was 2.00 seconds with a standard deviation of 0.15. We used the usual test functions whose detailed description can be found in [5], [4] and [7]. With these functions, mostly the efficiency of a global optimization algorithm can be measured. Wherever possible, the results from the original papers are included in our tables. These data differ slightly from those in [5] and [2]. Algorithms A and B were applied to each test function ten times. The parameters of the procedures were chosen so that they were able to find the global minimum each time.

We found that the computational effort (CPU time and number of function evaluations) was proportional to the required precision of the estimation of local minima. Thus, when different global optimization methods are compared, their accuracy should also be taken into account. First of all, the exact global minimum values should be determined. Table 1 gives the accurate global minimum and global minimizer values for every test function. These data are in good agreement with the results of Price [10]. It should be mentioned that slightly different numbers can be obtained with another computer precision. Certain global minimizer values in Table 1 are given by four or five decimal digits only, since for these test functions the same global minimum value can be achieved with somewhat different minimizers.

We subsequently determined the precision of the results obtained with the global

Table 1: *Global minimum and global minimizers of the test functions*

Test function	$F(\underline{x}^*)$	x_1^*	x_2^*	x_3^*	x_4^*	x_5^*	x_6^*
S5	-10.153206	3.99995	4.00014	4.00011	4.00016		
S7	-10.402947	4.00061	4.00072	3.99945	3.99958		
S10	-10.536416	4.00075	4.00061	3.99967	3.99948		
H3	-3.8627815	0.1146	0.5557	0.8525			
H6	-3.3223667	0.201536	0.149909	0.476906	0.275239	0.311593	0.657353
GP	2.9996490	0.000068	-1.0001				
RCOS	0.39788723	-3.1416	12.275				
	0.39788723	3.1416	2.2750				
	0.39788723	9.425	2.4750				
SHCB	-1.0316286	0.0899	-0.7126				
	-1.0316286	-0.0899	0.7126				
RB	0.0	1.0	1.0				

optimization methods cited in [2] wherever this was possible. The number of significant digits was defined by

$$-\log \frac{F(\underline{x}') - F(\underline{x}^*)}{F(\underline{x}^*)} \tag{2}$$

where $\underline{x}^*$ is a global minimizer of the given test function $F(\underline{x})$, and $\underline{x}'$ is its estimate. In the particular case of the Rosenbrock function (RB) where the global minimum is zero, the following expression was used for this:

$$-\log F(\underline{x}').$$

The numbers of significant digits are listed in Table 2 for every test function. We tried to tune the procedures A and B so that they achieve a similar accuracy (2) on the various test functions. The reliability and the accuracy of our method can be tuned almost independently.

The numbers of function evaluations required by the global optimization methods to solve test functions are given in Table 3. Since those local minimization

Table 2: *Number of significant digits in the global minimum*

Method	Test function								
	S5	S7	S10	H3	H6	GP	RCOS	SHCB	RB
Branin	—	—	—	—	—	—	—	—	—
Törn*	3.0	2.9	3.2	4.3	2.6	3.9	4.5	—	—
Price*	6.2	5.3	5.8	6.4	6.0	3.9	6.3	—	—
De Biase	2.9	3.4	2.0	4.7	4.7	4.8	—	6.4	—
Boender	—	—	—	—	—	—	—	—	—
A*	7.0	6.8	6.7	6.8	6.9	4.3	7.2	7.1	10.1
B*	4.9	4.0	5.2	4.3	4.3	4.4	4.1	4.5	4.7

* These methods do not use the partial derivatives of the objective function.

4*

Table 3: *Number of function evaluations*

Method	Test function								
	S5	S7	S10	H3	H6	GP	RCOS	SHCB	RB
Branin	5500	5020	4860	—	—	—	—	—	—
Törn*	3649	3606	3874	2584	3447	2499	1558	—	—
Price*	3800	4900	4400	2400	7600	2500	1800	—	—
De Biase	620	788	1160	732	807	378	597	717	—
Boender	567	624	755	235	462	398	235	—	—
A*	990	1767	2396	216	1446	436	330	233	410
B*	1083	1974	2689	697	2610	386	464	267	1524

* These methods do not use the partial derivatives of the objective function.

procedures that are not allowed to use the partial derivatives of the objective function are usually less efficient than the others, the efficiencies of algorithms A and B should be compared only with those of the similar non-derivative methods. The methods known to be non-derivative are marked by asterisks in Tables 2—4. The numbers in these Tables are results of a single sample run for each of the first four methods, the average of four independent runs for the method of Boender et al [2], and the averages of ten runs for algorithms A and B. Table 3 indicates that the procedure of Boender et al. works best of all, and the non-derivative methods of Törn [12] and Price [10] are less efficient than A and B.

Table 4 contains the numbers of standard time units required. As concerns these data, algorithms A and B seem to be definitely quicker than the other non-derivative ones, and procedure A is about as rapid as that of Boender et al. [2]. From the user's point of view Table 3 is more important, since in practical cases the evaluation of the objective function is more expensive than that of the standard test functions. Therefore, Table 4 is informative as to the overhead costs.

To summerize our numerical experience, we can state that these two non-derivative versions of the global optimization method of Boender et al. work definitely better than the other non-derivative procedures. The efficiency of implementation A

Table 4: *Numbers of standard time units*

Method	Test function								
	S5	S7	S10	H3	H6	GP	RCOS	SHCB	RB
Branin	9.0	8.5	9.5	—	—	—	—	—	—
Törn*	10.0	12.4	14.4	8.0	15.6	4.1	3.7	—	—
Price*	13.9	20.0	19.7	7.5	47.5	2.8	4.4	—	—
De Biase	26.1	23.0	33.7	17.6	23.1	16.8	15.2	23.2	—
Boender	3.5	4.5	7.0	1.7	4.3	1.5	1.0	—	—
A*	3.0	4.9	7.0	1.2	4.2	1.3	1.4	1.2	1.0
B*	3.5	6.0	8.8	1.9	14.2	1.5	1.6	1.3	1.5

* These methods do not use the partial derivatives of the objective function.

approaches that of the original one. The discussed global optimization method with a non-derivative quasi-Newton procedure can be highly recommended for the solution of smooth global optimization problems when calculation of the partial derivatives is inconvenient or impossible. The same global optimization method, together with the direct search method UNIRANDI, can be an efficient tool for locating the global minimum of non-smooth or non-differentiable objective functions.

5. Reliability

Almost all global optimization methods use only local information, i.e. the values of the objective function and its first and second derivatives at certain points. It is easy to show that, for the solution of this problem in general, there is no algorithm that uses only such local information at a finite number of points. For this reason, the research efforts on global optimization are concetrated mainly on evaluating increasingly reliable and efficient heuristics.

The size of the region of attraction [1] of the global minimum (the points of continuous curves in $\mathbf{R}^n$ that end in a global minimizer, and along which the objective function decreases strictly monotonously) characterizes the difficulty of a given problem. From this point of view, the most frequently used test problems [5] are rather easy to solve, and mostly the efficiency of an algorithm can be tested with them.

A new global optimization test problem is proposed below for comparing algorithms in terms of reliability and for testing the degree of difficulty of global optimization problems that can be solved with them.

The suggested n-dimensional test function is very simple:

$$F(\underline{x}) = \sum_{i=1}^{n} f_i(x_i) \tag{3}$$

where for every $i = 1, 2, ..., n$:

$$f_i(x_i) = x_i^6 \left(\sin(1/x_i) + 2\right)$$

if $x_i \neq 0$, and

$$f_i(0) = 0.$$

If $x_i \neq 0$, the gradient and Hessian of $F(\underline{x})$, respectively, are

$$g_i(\underline{x}) = 6x_i^5 \left(\sin(1/x_i) + 2\right) - x_i^4 \cos(1/x_i) \tag{4}$$

and

$$H_{i,j}(\underline{x}) = 0 \quad (i \neq j)$$

$$H_{i,i}(\underline{x}) = 30x_i^4 \left(\sin(1/x_i) + 2\right) - 10x_i^3 \cos(1/x_i) - x_i^2 \sin(1/x_i) \tag{5}$$

$i, j = 1, 2, ..., n$. Otherwise, $g_i(x)$ and $H_{i,j}(\underline{x})$ are zeros, $i, j = 1, 2, ..., n$. The gradient and the Hessian are continuous everywhere in $\mathbf{R}^n$. Since

$$\sum_{i=1}^{n} x_i^6 \leq F(\underline{x}) \leq 3 \sum_{i=1}^{n} x_i^6 \tag{6}$$

the global minimum of $F(\underline{x})$ on $\mathbf{R}^n$ is zero, and this value is reached only in the origin.

Theorem. *The function $F(\underline{x})$ has a countable infinity of local minima and maxima. All these extrema are in the hypercube*

$$-1 \leqq x_i \leqq 1 \quad i = 1, 2, \ldots, n. \tag{7}$$

Proof. First consider the case when $n = 1$. Supposing that the first derivative is equal to zero and $x \neq 0$, it holds that

$$6x \left(\sin(1/x) + 2 \right) = \cos(1/x). \tag{8}$$

The right side of this equation varies from -1 to 1, while the left side is between $6x$ and $18x$. Hence, the first derivative can be zero only in the interval $(-1/6, 1/6)$. The right side of equation (8) takes the values -1 and 1 in each interval of

$$\left[(1/2k\pi), 1/(2(k+1)\pi) \right) \quad k = \pm 3, \pm 4, \ldots \tag{9}$$

whereas in the same interval

$$-1 < 6x \left(\sin(1/x) + 2 \right) < 1. \tag{10}$$

This proves that the function $F(x)$ has at least one local minimum and one local maximum in every interval of (9), since the first derivative is a continuous function. These extrema are diverse, because they are all inside the intervals. Thus, there is at least a countable infinity of local minima and maxima in (7).

It can easily be seen that the first and the second derivatives of $F(x)$ can not be zeros in the same place. Since the second derivative is continuous, each local extremum is associated with an open interval of **R**, in which it is the only local extremum. Consequently, there cannot be a continuum of local extrema of $F(x)$ in **R**.

For any positive integer n the same proof holds, by using the fact that $F(\underline{x})$ is separable.   □

Thus, the unconstrained problem has the same set of local minima as the problem with the bounds (7). The global minimizer is non-isolated, in the sense that it is an accumulation point of local minimizers [1] (and it is the only one). The region of attraction of the global minimum is obviously of zero measure. The most important property of $F(\underline{x})$ is that the smaller the local minimum, the smaller the measure of the region of attraction relating to this local minimum. This feature can be used to asses the degree of difficulty of global optimization problems that can be solved via the given method.

The local minimizers of the one-dimensional version of the test function can be ordered according to the magnitude of the function value. The serial number N_x of the local minimizer x can be calculated using the equation

$$N_x = 2 \lfloor |1/x|/2\pi \rfloor - 1 + \left(\operatorname{sgn}(x) - 1 \right)/2 \tag{11}$$

where $\lfloor \cdot \rfloor$ denotes the largest integer not greater than the argument, and *sgn* stands for the signum function. In the one-dimensional case the size of the region of attraction A_x of local minimizer x can be well estimated by using equation (8), provided that the absolute value of x is small. The left side of this equation is then close to zero, and A_x is approximately equal to the distance between the two zeros of the right side of equation (8) that are adjacent to x:

$$A_x \approx \frac{2}{\dfrac{1}{x_2} - \pi^2}. \tag{12}$$

The numerical form of $F(\underline{x})$ obviously differs from the analytical one, especially near the origin. Thus, it is important to code this test function very carefully. Our version was written in FORTRAN and run on the mainframe R55M, by using single precision arithmetic.

The proposed test function can be computed quickly: in the one-dimensional case 1000 evaluations of $F(x)$ need 0.306 ± 0.006 (SD) standard time units [5]. When $n = 4$, the corresponding figure is 0.829 ± 0.001 (SD). Accordingly, the computation of even the four-dimensional version requires somewhat less computational effort than that of the S5 function [5]. The numerical form of $F(\underline{x})$ is zero in the hypercube $x_i \in (-1.0 \ 10^{-13}, 1.0 \ 10^{-13})$, $i = 1, 2, ..., n$. In spite of this, there are more than one million local minima whose regions of attraction contain at least 100 points that can be represented by using single precision.

The algorithm A was tested by running it independently ten times on the one- and four-dimensional versions of this test problem with the bounds (7). The parameters of the algorithm were set so that the estimate of the global minimum was as close to zero as possible, and they were different from those used in the previous section. From the point of view of this reliability test the type of the local search procedure is indifferent.

In the one-dimensional case, the smallest minimum found was $0.523449 \ 10^{-52}$ in $0.193281 \ 10^{-8}$. This was the 164,687,623rd local minimizer in the sequence discussed above, and the size of its region of attraction A_x was $0.23472 \ 10^{-16}$ according to (12). The worst estimate of the global minimum was $0.319144 \ 10^{-23}$ in $-0.119009 \ 10^{-3}$; this was the 2,673rd local minimizer, with $A_x = 0.88989 \ 10^{-7}$. The average run consumed 33.5 standard time units and 22,137 function evaluations. In the four-dimensional case, the best and the worst estimates of the global minimum were $0.272099 \ 10^{-8}$ and $0.598347 \ 10^{-6}$, respectively. The average run consumed 46.1 standard time units and 22,020 function evaluations.

In conclusion, the results of this reliability test have shown that the studied global optimization method can be tuned to solve most practical problems with satisfactory reliability.

Abstract

In this paper we first show that the objective function of a least squares type nonlinear parameter estimation problem can be any nonnegative real function, and therefore this class of problems corresponds to global optimization. Two non-derivative implementations of a global optimization method are presented, with nine standard test functions applied to measure their efficiency. A new nonlinear test problem is then presented for testing the reliability of global optimization algorithms. This test function has a countable infinity of local minima and only one global minimizer. The region of attraction of the global minimum is of zero measure. The results of efficiency and reliability tests are given.

Key words. Global optimization, nonlinear parameter estimation, sum of squares, least squares, test problem.

T. CSENDES
KALMÁR LABORATORY OF CYBERNETICS
JÓZSEF ATTILA UNIVERSITY
ÁRPÁD TÉR 2.
SZEGED
HUNGARY

References

[1] F. ARCHETTI, G. P. SZEGŐ: Global optimization algorithms, in: Nonlinear Optimization — Theory and Algorithms (Dixon, L. C. W., Spedicato, E., Szegő, G. P., eds., Birkhäuser, Boston, 1980), 429—469.

[2] C. G. E. BOENDER, A. H. G. RINNOOY KAN, G. T. TIMMER, L. STOUGIE: A stochastic method for global optimization, *Math. Programming,* 22 (1982), 125—140.

[3] T. CSENDES, B. DARÓCZY, Z. HANTOS: Nonlinear parameter estimation by global optimization: comparison of local search methods in respiratory system modelling, in: System Modelling and Optimization (Thoma, M., Wyner, A., eds., Springer-Verlag, Berlin, 1986), 188—192.

[4] L. DE BIASE, F. FRONTINI: A stochastic method for global optimization: its structure and numerical performance, in: Towards Global Optimisation 2. (Dixon, L. C. W., Szegő, G. P., eds., North-Holland, Amsterdam, 1978), 85—102.

[5] L. C. W. DIXON, G. P. SZEGŐ: The global optimisation problem: an introduction, in: Towards Global Optimisation 2. (Dixon, L. C. W., Szegő, G. P., eds., North-Holland, Amsterdam, 1978), 1—15.

[6] P. E. GILL, W. MURRAY, M. H. WRIGHT: Practical Optimization (Academic Press, London, 1981).

[7] J. GOMULKA: Deterministic vs probabilistic approaches to global optimisation, in: Towards Global Optimisation 2. (Dixon, L. C. W., Szegő, G. P., eds., North-Holland, Amsterdam, 1978), 19—29.

[8] Z. HANTOS, B. DARÓCZY, B. SUKI, G. GALGÓCZY, T. CSENDES: Forced oscillatory impedance of the respiratory system at low frequencies. J. Appl. Physiol. 60 (1986), 123—132.

[9] T. JÄRVI: A random search optimizer with an application to a maxmin problem, *Publications of the Institute for Applied Mathematics,* University of Turku, No. 3, 1973.

[10] W. L. PRICE: A controlled random search procedure for global optimisation, in: Towards Global Optimisation 2. (Dixon, L. C. W., Szegő, G. P., eds., North-Holland, Amsterdam, 1978), 71—84.

[11] G. T. TIMMER: Global optimization: a stochastic approach, (Ph. D. Thesis, *Erasmus University,* Rotterdam, 1984).

[12] A. A. TÖRN: A search-clustering approach to global optimization, in: Towards Global Optimisation 2. (Dixon, L. C. W., Szegő, G. P., eds., North-Holland, Amsterdam, 1978), 49—62.