ECONOMETRIC INSTITUTE

BAYESIAN STOPPING RULES FOR MULTISTART GLOBAL OPTIMIZATION METHODS

C.G.E. BOENDER AND A.H.G. RINNOOY KAN

REPORT 8536/A

ERASMUS UNIVERSITY ROTTERDAM - P.O. BOX 1738 - 3000 DR ROTTERDAM - THE NETHERLANDS

BAYESIAN STOPPING RULES FOR MULTISTART GLOBAL OPTIMIZATION METHODS

C.G.E. Boender*
A.H.G. Rinnooy Kan* **

Abstract

By far the most efficient methods for global optimization are based on starting a local optimization routine from an appropriate subset of uniformly distributed starting points. As the number of local optima is frequently unknown in advance, it is a crucial problem when to stop the sequence of sampling and searching. By viewing a set of observed minima as a sample from a generalized multinomial distribution whose cells correspond to the local optima of the objective function, we obtain the posterior distribution of the number of local optima and of the relative size of their regions of attraction. This information is used to construct sequential Bayesian stopping rules which find the optimal trade off between reliability and computational effort.

1. Introduction.

So far, only a few solution methods have been developed for the unconstrained global optimization problem, which is to find the global optimum x_* (say, the global minimum) of a real valued multimodal objective function f over a compact set S which contains x_* as an interior point [Dixon & Szegö 1975, 1978]. As confirmed in recent computational experiments [Rinnooy Kan & Timmer 1984; Timmer 1984], the currently most efficient way to solve this problem is through methods that perform a local search from each point in an appropriate subset of a random sample drawn from the uniform distribution over S. If the true number of local minima of f is unknown, such methods can of course never provide an absolute guarantee that the globally optimal value has been found: all that can be assured is that the probability of this event rapidly

* Econometric Institute, Erasmus University Rotterdam, The Netherlands
** Department of Industrial Engineering & Operations Research, School of Business Administration, University of California, Berkeley, U.S.A.

approaches 1 as the sample size goes to infinity. Thus, there typically exists a need for stopping rules to determine the sample size which corresponds to the optimal trade off between reliability and computational effort.

In most computational experiments reported in the literature this crucial feature is dealt with in ad hoc fashion, which is strongly inspired by the properties of the test functions that happen to be involved. In practice, one often simply performs a prespecified number of local searches. Here, we shall describe a rigorous Bayesian framework for the development of optimal stopping rules. In the Bayesian approach, the user is asked to express his beliefs about some unknown relevant parameters of f in the form of a prior distribution. Experimental information gathered on f is then used to convert these initial beliefs into a posterior distribution through Bayes' Theorem. This posterior distribution reflects the way in which the initial beliefs are affected by the outcomes of the experiments. A decision whether or not to continue the search can then be taken which is optimal with respect to a loss function based on a termination loss if sampling is stopped before all local minima have been found and an execution loss which expresses the cost of sampling and of performing new local searches. Given the initial beliefs, such a decision incorporates all information derived from the experiments to weigh expected costs and benefits against each other in an optimal fashion.

The construction of our stopping rules is based on a statistical analysis of the Multistart method. In this approach a local search is applied to each point in the uniform sample. A crucial observation about Multistart is that its outcome, in the form of a set of local minima, can be viewed as a sample from a multinomial distribution whose cells correspond to the local minima of f. However, since the details of this correspondence are unknown in advance, the multinomial formula turns out to be inappropriate for statistical inference about the local minima structure of f. A first solution to this problem is described in [Zielinkski 1981] where the probability distribution is derived of the different number of local minima W which will be found in n local searches. Conditioned on a set of observed minima, this result is applied to determine the optimal Bayesian estimator of the true number of local minima with respect to a loss function which imputes a fixed cost if an incorrect decision is made. Also, the optimal estimator of the total relative size of the regions of attraction of the observed minima is derived under a

quadratic loss function. However, the loss functions considered do not involve the cost of sampling and of performing local searches. Hence these results are insufficient to construct stopping rules which provide the optimal trade-off between reliability and computational effort.

In Section 2 of this paper we will first describe how Zielinski's results can be extended through the use of the generalized multinomial distribution. In addition to dealing with the different number of local minima W, this approach also allows us to determine the probability that an application of Multistart to a sample of size n leads to the i-th local minimum being found N_i times (i = i,...,W; $\Sigma_{i=1}^{W} N_i = n$). The generalized multinomial distribution is then used to compute relevant posterior results concerning the minima of f which extend the work in [Zielinski 1981]. In Section 3, these posterior results are applied to construct optimal sequential Bayesian stopping rules with respect to loss functions which not only impute a cost if sampling is stopped too early, but which also take the cost of extending the sample into account. Our experimental results on the standard set of test functions from [Dixon & Szegö 1978] are contained in Section 4. In Section 5, we will show that our results can be extended to deal with the important case where the local minima are generated by computationally superior clustering variants of Multistart that aim to perform only one local search with respect to each local minimum of f.

2. Statistical analysis of Multistart.

In the Multistart method, a local search procedure L is started from each point in a sample that has been generated from the uniform distribution over the feasible region S. We recall that the region of attraction R_{x^*} of a local minimum x^*, given L, is defined as the subset of points in S starting from which L will arrive at x^* [Dixon & Szegö 1975, 1978]. Let k be the number of local minima of the objective function f, and let us denote the relative volume of the i-th region of attraction by θ_i (i = 1,...,k). If these values would be known, then we have several obvious stopping rules for the sample at our disposal. We may terminate Multistart, for example, if the number of different local minima observed is equal to k, or if the total relative volume of the observed regions of attraction exceeds some prespecified value.

In practice, $k, \theta_1, \ldots, \theta_k$ are always unknown. The local minima sampled, however, clearly provide information about the values of these parameters. The crucial observation that enables us to learn about these values is that, since the starting points of Multistart are uniformly distributed over S, the i-th local minimum at each trial has a fixed probability of being found that is equal to the relative volume θ_i of its region of attraction ($i = 1, \ldots, k$). This implies that, given a number of local searches n, the observed minima can be viewed as a sample from a multinomial distribution whose cells correspond to the minima of f: the number of cells is equal to the unknown number k of minima and the cell probabilities are equal to the unknown relative volumes $\theta_1, \ldots, \theta_k$. Hence, if the random variables N_i with realizations n_i ($i = 1, \ldots, k$) are defined as the number of times that the i-th local minimum is found in n local searches, then the probability of the event $\{(N_1, \ldots, N_k) = (n_1, \ldots, n_k)\}$ is given by

$$(1) \qquad p(n_1, \ldots, n_k) = \frac{n!}{\prod_{i=1}^{k} n_i!} \prod_{i=1}^{k} \theta_i^{n_i}.$$

In a statistical approach one would like to obtain information about $k, \theta_1, \ldots, \theta_k$ by substituting observed sample values $n_1, \ldots, n_k$ into (1). This would yield the likelihood function of $k, \theta_1, \ldots, \theta_k$, with respect to which one could compute estimates of these unknowns, such as, for example, the maximum likelihood estimates which correspond to values for $k, \theta_1, \ldots, \theta_k$ that maximize the probability that the observed sample values will occur. Here, however, we encounter a serious difficulty. Since the existence of a local minimum cannot be postulated until it has been observed, it is impossible to decide to which local minimum a θ_i corresponds. For example, if in 8 local searches one minimum is found 5 times, another one twice, and a third minimum once, it is impossible to distinguish between the events $\{(N_1, N_2, N_3) = (5, 2, 1)\}$, $\{(N_1, N_2, N_3) = (1, 2, 5)\}$, $\{(N_1, N_2, N_3, N_4, N_5, N_6) = (0, 5, 0, 1, 2, 0)\}$ etc. Hence, (1) is inappropriate for the computation of the probability of such an outcome.

The solution to this problem is to define mutually exclusive and exhaustive aggregates of the individual events $n_1, \ldots, n_k$ by disregarding the order and omitting the n_i's that are equal to 0. Thus, for example, the aggregate event corresponding to the outcome mentioned above is the (multi)set

$\{N_1, N_2, N_3\} = \{1, 2, 5\}$. If we define the random variable W, with realization w, as the number of different observed minima in n local searches, the probability of these aggregate events turns out to be given by the generalized multinomial distribution [Boender & Rinnooy Kan 1983a; Boender 1984], i.e.

$$(2) \qquad p(\{n_1,\ldots,n_w\}) = \frac{1}{\prod_{j=1}^{n} h_j!} \frac{n!}{\prod_{i=1}^{w} n_i!} \sum_{(g_1,\ldots,g_w)\in S_k[w]} \prod_{i=1}^{n_i} \theta_{g_i},$$

where h_j is the number of n_i's that are equal to j, and $S_k[w]$ is the set of all permutations of w different elements from $\{1,\ldots,k\}$ ($\Sigma_{i=1}^{w} n_i = n$; $n_i > 0$ for $i = 1,\ldots,w$).

However, if we now try to use (2) to obtain the maximum likelihood estimate of the number of local minima k, a second difficulty appears: we find this estimate to be equal to ∞ for all possible outcomes $\{n_1,\ldots,n_w\}$. Hence, this approach does not provide a proper setting for the problem. We therefore adopt the Bayesian approach in which the unknowns $k,\theta_1,\ldots,\theta_k$ are assumed to be themselves random variables $K,\Theta_1,\ldots,\Theta_K$ with realizations $k,\theta_1,\ldots,\theta_k$, for which a prior distribution can be specified. Given the outcome $\{n_1,\ldots,n_w\}$ of a number of local searches, we then use Bayes' Theorem to compute the posterior distribution of $K,\Theta_1,\ldots,\Theta_K$, which incorporates both our prior beliefs and the sample information.

For the number of local minima K we shall first assume each integer of $[1,\infty)$ to be a priori equiprobable. Given $K = k$, we assume the relative sizes of the regions of attraction $\Theta_1,\ldots,\Theta_k$ to follow a uniform distribution on the (k-1)-dimensional unit simplex $I_{k-1} = \{(\theta_1,\ldots,\theta_k) \mid \theta_i \geq 0 \ (i = 1, \ldots, k), \ \Sigma_{i=1}^{k} \theta_i = 1\}$. Hence, our joint prior probability density function is given by

$$(3) \qquad p(k,\theta_1,\ldots,\theta_k) \propto (k-1)!,$$

where $\propto$ denotes proportionality. Other choices for a prior distribution can be accommodated as well (see Section 5).

Theorem 1 [Boender & Rinnooy Kan 1983a].

Given the uniform prior (3) and a Multistart outcome $\{n_1,\ldots,n_w\}$, the posterior density function of the number of local minima K, and the relative volumes of the regions of attraction $\theta_1,\ldots,\theta_k$ for $n \geq w+2$ is given by

$$p(k,\theta_1,\ldots,\theta_k|\{n_1,\ldots,n_w\}) =$$

$$(4) \qquad = \frac{(n-1)!(n-2)!(k-1)!}{w!(w-1)!(n-w-2)! \prod_{i=1}^{w} n_i!} \sum_{(g_1,\ldots,g_w)\in S_k[w]} \prod_{i=1}^{w} \theta_{g_i}^{n_i}.$$

Sketch of proof.

Substitution of the generalized multinomial formula (2) and the assumed prior (3) in Bayes' Theorem yields

$$p(k,\theta_1,\ldots,\theta_k|\{n_1,\ldots,n_w\}) =$$

$$= \frac{p(\{n_1,\ldots,n_w\})\ p(k,\theta_1,\ldots,\theta_k)}{\sum_{m=w}^{\infty} \int_{I_{m-1}} p(\{n_1,\ldots,n_w\})\ p(m,\psi_1,\ldots,\psi_m) \prod_{i=1}^{m} d\psi_i}$$

$$(5) \qquad = \frac{(k-1)! \sum_{(g_1,\ldots,g_w)\in S_k[w]} \prod_{i=1}^{w} \theta_{g_i}^{n_i}}{\sum_{m=w}^{\infty} \int_{I_{m-1}} (m-1)! \sum_{(g_1,\ldots,g_w)\in S_m[w]} \prod_{i=1}^{w} \psi_{g_i}^{n_i} \prod_{i=1}^{m} d\psi_i}.$$

We observe that

$$(6) \qquad \frac{(n+m-1)!}{\prod_{i=1}^{w} n_i!} \prod_{i=1}^{w} \psi_{g_i}^{n_i}$$

is an m-dimensional Dirichlet density function with parameters $n_1+1,\ldots,n_w+1$, $1,\ldots,1$ so that (5) simplifies to

$$p(k,\theta_1,\ldots,\theta_k|\{n_1,\ldots,n_w\}) =$$

$$(7) \qquad = \frac{(k-1)!}{\prod_{i=1}^{w} n_i! \sum_{m=w}^{\infty} \frac{(m-1)!m!}{(m-w)!(n+m-1)!}} \sum_{(g_1,\ldots,g_w)\in S_k[w]} \prod_{i=1}^{w} \theta_{g_i}^{n_i}.$$

The proof is completed by using the equality

$$(8) \qquad \sum_{m=w}^{\infty} \frac{(m-1)!m!}{(m-w)!(n+m-1)!} = \frac{w!(w-1)!(n-w-2)!}{(n-1)!(n-2)!} \qquad (n \geq w+2).$$

The following corollaries are consequences of Theorem 1. The result (15) has previously been published in [Zielinski 1981]; (10), (11), (13) and (15) appeared in [Boender & Rinnooy Kan 1983a]. The other results are published here for the first time; their proofs can be found in Appendix A. The corollaries involve the random variable Ω, with realization ω, which is defined as the total relative volume of the observed regions of attraction. The posterior density of this quantity yields useful information if, for example, one is not interested in local minima with an extremely small region of attraction. Also, since the complement of Ω is equal to the probability that an additional local search will render a new local minimum, it will play a crucial role in our subsequent construction of optimal stopping rules for a sequential sample of local minima.

Corollary 1.1. Posterior density of the number of local minima and of the total volume of the regions of attraction of the observed minima:

$$(9) \qquad p(k,\omega|\{n_1,\ldots,n_w\}) = \begin{cases} \prod_{i=1}^{w} \frac{n-1-i}{n-1+i} & \text{if } k=w,\ \omega=1, \\ \frac{(k-1)!k!(n-1)!(n-2)!}{(n+w-1)!(k-w-1)!(k-w)!(n-w-2)!}\, \omega^{n+w-1}(1-\omega)^{k-w-1} & \\ \text{otherwise } (n \geq w+2). & \end{cases}$$

Corollary 1.2 Marginal posterior probability distribution of the number of local minima, its expected value E, mode M and variance σ^2:

$$(10) \qquad p(k|\{n_1,\ldots,n_w\}) = \frac{(k-1)!k!(n-1)!(n-2)!}{(n+k-1)!(k-w)!w!(w-1)!(n-w-2)!} \qquad (n \geq w+2)$$

$$(11) \qquad E(K|\{n_1,\ldots,n_w\}) = \frac{w(n-1)}{n-w-2} \qquad (n\geq w+3)$$

$$(12) \qquad M(K|\{n_1,\ldots,n_w\}) = \frac{w(n-1)}{n-w} \qquad (n\geq w+1)$$

$$(13) \qquad \sigma^2(K|\{n_1,\ldots,n_w\}) = \frac{w(w+1)(n-1)(n-2)}{(n-w-2)^2(n-w-3)} \qquad (n\geq w+4).$$

Corollary 1.3 Posterior expected value of the relative volume of a region of attraction of a minimum which has been found n_j times:

$$(14) \qquad E(\theta_{n_j}|\{n_1,\ldots,n_w\}) = \frac{(n_j+1)(n+w)}{n(n-1)} \qquad (n\geq w+2)$$

Corollary 1.4 Posterior expected value and variance of the total volume of the observed regions of attraction:

$$(15) \qquad E(\Omega|\{n_1,\ldots,n_w\}) = \frac{(n-w-1)(n+w)}{n(n-1)} \qquad (n\geq w+2)$$

$$(16) \qquad \sigma^2(\Omega|\{n_1,\ldots,n_w\}) = \frac{2(n+w)(n-w-1)w(w+1)}{(n-1)^2n^2(n+1)} \qquad (n\geq w+2)$$

Several remarks about the above results are appropriate at this point. First of all, we observe that the posterior probability that all local minima (including the global minimum) have been discovered, is given by $\Pi_{i=1}^{w}((n-1-i)/(n-1+i))$ (cf. (9)). Secondly, we note that the corollaries (with the exception of (14)) do not involve the values of the n_i, but only the sample size n and the number of different local minima discovered w. Thirdly, since the prior for K is improper, the posterior results are only defined for $n \geq w + i$ ($i = 1, 2, 3$ or 4); this creates no problem in practical applications. Finally, we recall that the posterior expected values can be shown to be optimal Bayesian estimators under a quadratic loss function.

By way of illustration, the posterior probability density of $k,\theta_1,\ldots,\theta_k$ has been computed for the case that $k = 3$, given the Multistart outcomes $\{N_1, N_2, N_3\} = \{3, 3, 3\}$ (Figure 1a) and $\{N_1, N_2, N_3\} = \{7, 1, 1\}$ (Figure 1b). Note that since we do not know to which θ_i an n_i corresponds, these figures are symmetrical around the centre of gravity of I_2. In Figures 1c and 1d we depict the posterior probability that all local minima have been discovered, and the posterior expectation of the total volume of the observed regions of attraction. Both are given as a function of n and w, where n and w range from 1 to 25.

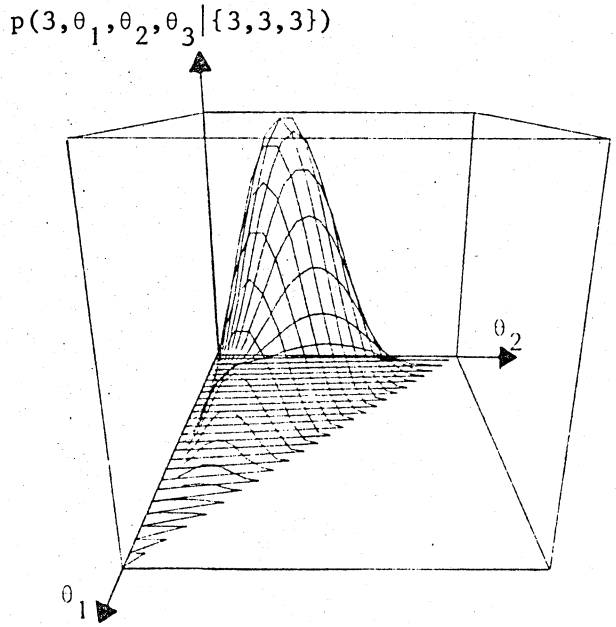


Figure 1a

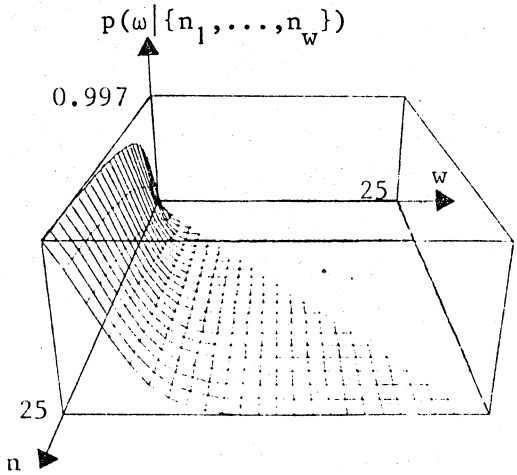


Figure 1b

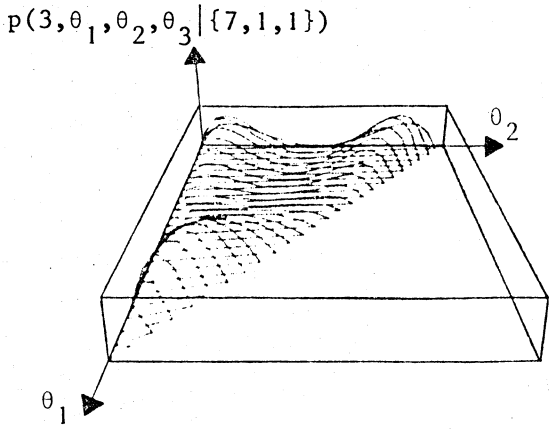


Figure 1c

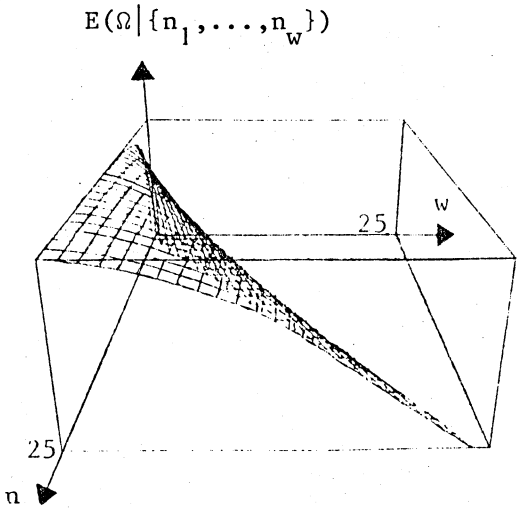


Figure 1d

3. Stopping rules.

We recall from Section 2 that if in n local searches w different local minima have been discovered, the Bayesian estimate of the unknown number of local minima is equal to

$$(17) \qquad w \frac{n-1}{n-w-2}.$$

For most (n,w) pairs (17) will yield a real valued estimate, although the true number of local minima is evidently an integer. It is easily verified that the optimal integer Bayesian estimate under a quadratic loss function is a round-off of the real valued estimate. Therefore an appealing simple stopping criterion is to terminate if

$$(18) \qquad w \frac{n-1}{n-w-2} - \tfrac{1}{2} \leq w,$$

i.e. to stop if the optimal integer Bayesian estimate of the unknown number of minima is equal to the number of distinct local minima observed.

If the stopping criterion (18) is satisfied the estimated number of unobserved minima is equal to 0. This may cause the algorithm to run for an extremely long time if the objective function has many minima with very small regions of attraction. If one is not willing to continue the search for all minima in these situations, an appropriate stopping criterion may be to terminate the algorithm if the total relative volume of the observed regions of attraction exceeds a prescribed value t ($0 < t < 1$), i.e.: stop if

$$(19) \qquad \frac{(n-w-1)(n+w)}{n(n-1)} \geq t.$$

It is preferable, however, to seek sequential stopping rules which also take into account the cost of sampling. To do so, we have to impute a termination loss which attaches costs to the deviation of an estimated unknown quantity from its true value, as well as an execution loss corresponding to the cost of further experiments. We can then try to construct the rule that minimizes expected posterior loss [De Groot 1970].

We will consider 4 loss structures: for each, the execution loss is assumed to

be proportional to the number of additional sample points.

1. The termination loss is equal to a fixed constant if sampling is stopped before all minima have been discovered and 0 otherwise [Boender & Zielinski 1982; Boender & Rinnooy Kan 1983a]:

$$L_1 = \begin{cases} n + c_1 & \text{if } K > W \\ n & \text{if } K = W \end{cases} \tag{20}$$

Given $\{n_1,\ldots,n_w\}$ the posterior loss is equal to (cf. (9))

$$E(L_1|\{n_1,\ldots,n_w\}) = c_1\left(1 - \prod_{i=1}^{w} \frac{n-1-i}{n-1+i}\right) + n. \tag{21}$$

2. The termination loss is proportional to the number of unobserved minima:

$$L_2 = c_2(K - W) + n, \tag{22}$$

The posterior loss is equal to (cf. (11))

$$E(L_2|\{n_1,\ldots,n_w\}) = c_2 \frac{w(w+1)}{n-w-2} + n. \tag{23}$$

3. The termination loss is proportional to the fraction of unobserved minima:

$$L_3 = c_3\left(\frac{K - W}{K}\right) + n. \tag{24}$$

From (8) and (10), the posterior loss can be seen to be equal to

$$E(L_3|\{n_1,\ldots,n_w\}) = c_3 \frac{w}{n-1} + n. \tag{25}$$

4. The termination loss is proportional to the total relative volume of the unobserved regions of attraction:

$$L_4 = c_4(1 - \Omega) + n. \tag{26}$$

From (15), the posterior loss is equal to

$$E(L_4|\{n_1,\ldots,n_w\}) = c_4 \frac{w(w+1)}{n(n-1)} + n. \tag{27}$$

Since the posterior losses, given a sample outcome $\{n_1,\ldots,n_w\}$, turn out to depend only on the number of local searches n, and the number of observed local minima w, we will in this section use the notation (n,w) to describe the outcome $\{n_1,\ldots,n_w\}$ from now on.

We note that for the first loss structure (20) the termination loss can only decrease if ultimately all local minima are found, whereas for the second loss function (22) the termination loss is reduced by a fixed constant each time an additional local minimum is observed. The third loss function (24) attempts to involve the global minimum in the analysis directly: if the true number of local minima is equal to k and w different minima have been observed, then, loosely speaking and in absence of further information about f, the chance that the global minimum is not among the observed ones equals $(k - w)/k$. The fourth loss structure (26), which is based on the total relative volume of the unobserved regions of attraction, reflects the case in which one is less interested in finding local minima with extremely small regions of attraction.

Given a sample of size n, the posterior loss after $n' > n$ observations is a random variable $E(L_j|(n',W))$. Our purpose, then, is to find for each loss function the stopping rule that minimizes the expected value of the sequence of posterior losses $\{E(L_j|(n',W)\}_{n'=n+1}^{\infty}$. Given a current pair (n,w) only two relevant outcomes can occur as the result of an additional local search: either a yet unobserved local minimum is found ($(n,w) \to (n+1,w+1)$), or not ($(n,w) \to (n+1,w)$). The posterior probability that the next search will not result in the discovery of an unobserved minimum is equal to the posterior expected volume of the observed regions of attraction (cf. (15)). Hence the sequence of posterior losses satisfies the recurrence relation:

$$(28)\qquad E(E(L_j|(n+1,W))|(n,w)) =$$

$$= E(\Omega|(n,w))E(L_j|(n+1,w)) + (1 - E(\Omega|(n,w))E(L_j|(n+1,w+1))$$

$$= \frac{(n-w-1)(n+w)}{n(n-1)} E(L_j|(n+1,w)) + \frac{w(w+1)}{n(n-1)} E(L_j|(n+1,w+1)).$$

Thus, given the current pair (n,w), we can compute the conditional expected

value of the posterior loss of n+1 observations according to (28). Therefore, an appealing one-step stopping rule is to terminate if this conditional expected posterior loss of n+1 observations $E(E(L_j|(n+1,W))|(n,w))$ is greater than the current posterior loss $E(L_j|(n,w))$. Observe that this decision is based on the assumption that the search is immediately stopped once the next observation would have been performed. In contrast, optimal stopping rules are obtained by comparing the current posterior loss with the expected posterior loss of another observation under the assumption that the best strategy is adopted thereafter [De Groot 1970]. For such an approach to be feasible we have to find a value n_j^* such that for all pairs (n,w) with $n \geq n_j^*$ the sequence of posterior losses is a submartingale, i.e. the expected posterior loss for all (n,w) pairs with $n \geq n_j^*$ never decreases if another observation is performed. We then know for all pairs (n,w) with $n = n_j^*$ that the optimal decision is to stop, and we can compute the corresponding posterior losses according to the formulae (21), (23), (25) or (27) for $E(L_j|(n,w))$. Given this result, we can start working backwards form $n = n_j^*$. Since we know that for all (n,w) pairs with $n = n_j^*$ the optimal decision is to stop, we have that for the (n,w) pairs with $n = n_j^* - 1$ the above one-step rule is the optimal one. Thus comparing the expected posterior loss of n+1 observations (28), and the current posterior loss (21), (23), (25) or (27), we can determine for each pair (n,w) with $n = n_j^* - 1$ if the optimal decision is to continue or not. Then we proceed to stage $n_j^* - 2$. Again we use (28) to compute the expected posterior loss of continuation. Now, however, if the optimal decision in (n+1,w) or (n+1,w+1) is to continue once more, we substitute the previously computed corresponding expected posterior loss of continuation in (28). By working backwards in this way to the first stage we can compute for each (n,w) pair if the optimal decision is to continue, or not. Note, given c_j, that a single table will summarize the optimal stopping policy for all possible objective functions.

Theorem 2 contains the relevant results for the four loss structures under consideration. The proofs are contained in Appendix B.

Theorem 2.

Submartingale results ($n \geq w + 2$)

(29) $E(E(L_1|(n+1,W))|(n,w)) \geq E(L_1|(n,w))$ for $n > n_1^* = c_1 + 1 - \sqrt{(4c_1+1)}$

(30) $E(E(L_2|(n+1,W))|(n,w)) = E(L_2|(n,w)) - c_2(1-E(\Omega|(n,w))) + 1$

(31) $E(E(L_3|(n+1,W))|(n,w)) \geq E(L_3|(n,w))$ for $n > n_3^* = \frac{c_3}{4}$

(32) $E(E(L_4|(n+1,W))|(n,w)) \geq E(L_4|(n,w))$ for $n > n_4^* = \frac{c_4}{3}$.

Theorem 2 indicates that for the first, third and fourth loss function a value n_j^* exists such that the sequence of local searches should be stopped if $n \geq n_j^*$. Thus, for these loss functions we can compute optimal stopping rules by backwards recursion. In Figure 2 the results are depicted for $c_j = 1000$. If the current (n,w) pair is within the appropriate convex region the optimal decision is to continue: as soon as the boundary of the region is reached no more local searches should be performed.

The theorem shows further that for the second loss function no such value n_2^* exists, so that for this loss function it is impossible to determine an optimal sequential stopping rule. However, the theorem does imply that the expected posterior loss after one additional local search is lower than the current posterior loss if the posterior probability that the next local search will render a yet unidentified local minimum is greater than $1/c_2$ (cf. (15)). Hence a proper one-step stopping rule is to continue the search as long as this probability is greater that $1/c_2$, and to stop otherwise. We depicted this rule in Figure 2 for the case $c_2 = 1000$: as soon as the dotted line is crossed the search should be terminated.

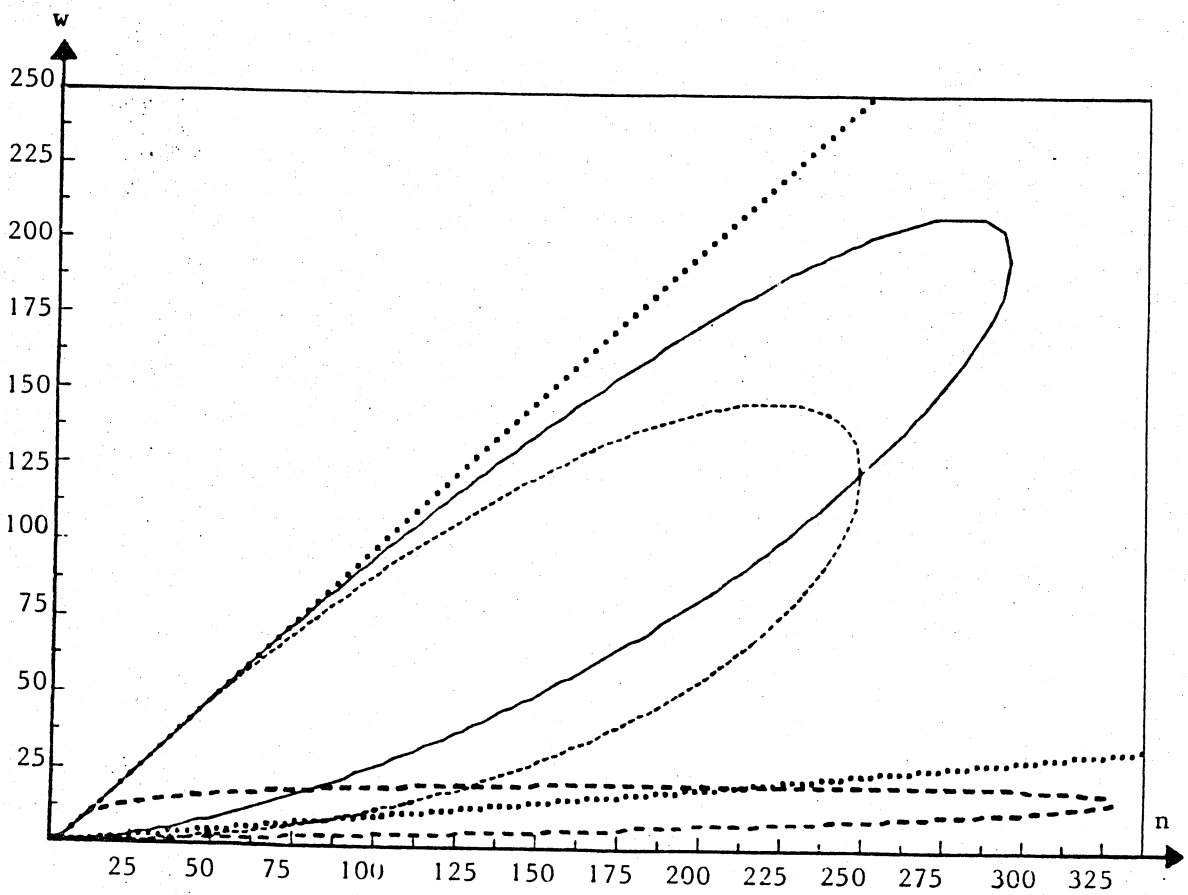


Figure 2

STOPPING RULES

- Optimal rule 1 (c_1 =1000)
- One-step rule 2 (c_2 =100)
- Optimal rule 3 (c_3 =1000)
- Optimal rule 4 (c_4 =1000)

4. Test results.

In this section, rules 1 to 4 are applied to the standard set of test functions for global optimization [Dixon & Szegö 1978]. The number of local minima of these test functions are listed in Table 1. In our experiments we used the original Multistart algorithm, although the analysis carries over to more superior clustering variants (cf. Section 5). In Figure 3 we have depicted for each test function the number of local minima which were discovered during the search. Also we indicated at which sample size the rules 1, 2, 3 and 4 prescribed termination. Finally, at each point during the search Figure 3 shows the current optimal estimate of the number of local minima, as well as the optimal estimate of the total relative size of the observed regions of attraction (for reasons of comparability the latter estimate is multiplied by the true number of local minima).

Table 1.

Test functions [Dixon & Szegö 1978].

	Name	k
A	Goldstein & Price	4
B	Branin	3
C	Hartman 3	3
D	Hartman 6	2
E	Shekel 5	5
F	Shekel 7	7
G	Shekel 10	10

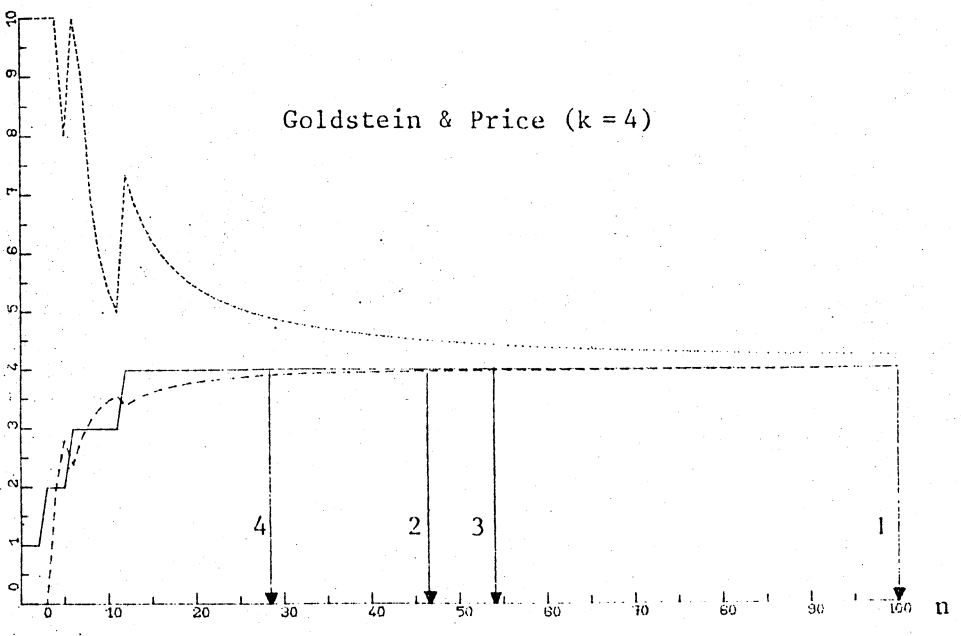


Figure 3

—— Observed number of local minima: w.

..... Posterior expected number of minima: $E(K|\{n_1,\ldots,n_w\}) = \frac{w(n-1)}{n-w-2}$

---- Posterior expected total relative volume of the regions of attraction of the observed minima multiplied by the true number of local minima: $E(\Omega|\{n_1,\ldots,n_w\}) * k = \frac{(n-w-1)(n+w)}{n(n-1)} * k$

Stopping rule 1. The execution loss is proportional to the number of sample points n, and the termination loss is equal to 0 iff all minima are discovered, and equal to 0 otherwise:

$$L_1 = n + \begin{cases} 1000 & \text{if } K > W \\ 0 & \text{if } K = W. \end{cases}$$

Stopping rule 2. The termination loss is proportional to the number of unobserved minima:

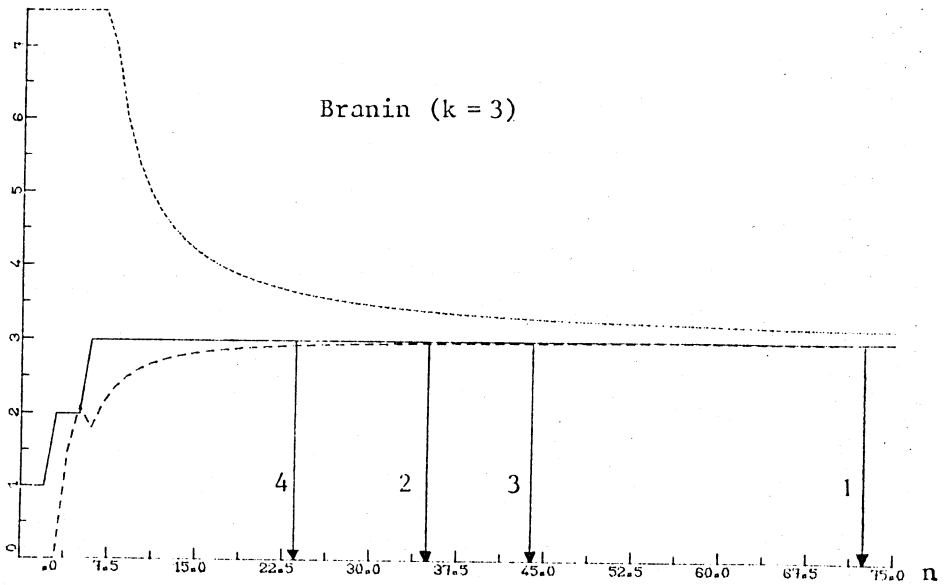
$$L_2 = 100 * (K - W) + n.$$

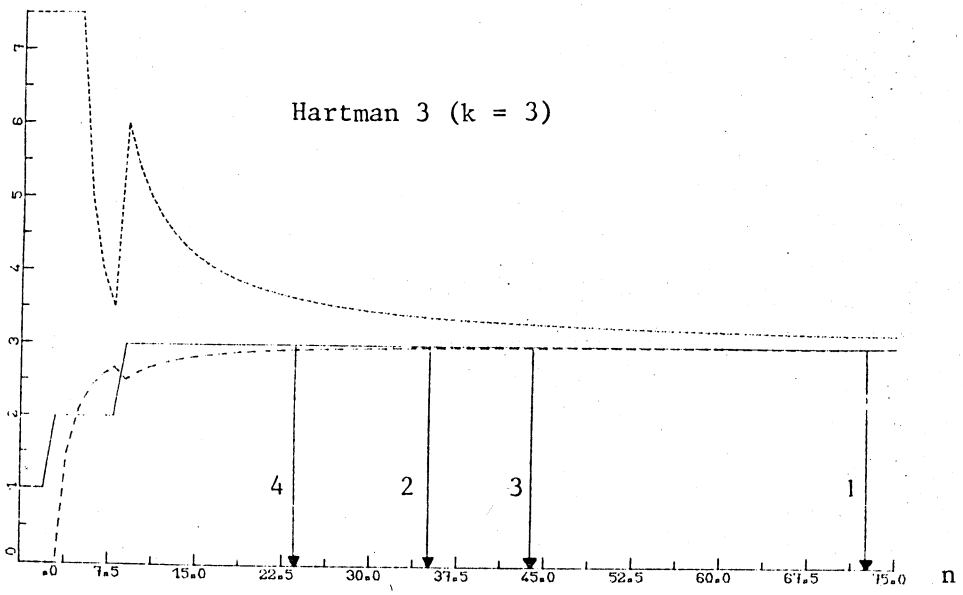
Stopping rule 3. The termination loss is proportional to the fraction of unobserved minima:

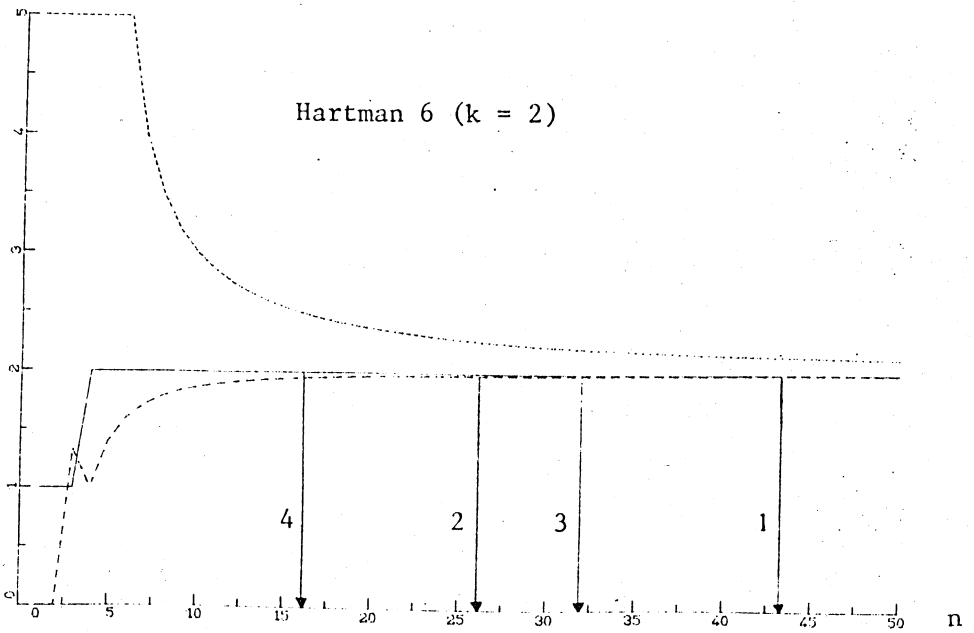
$$L_3 = 1000 * (\frac{K - W}{K}) + n.$$

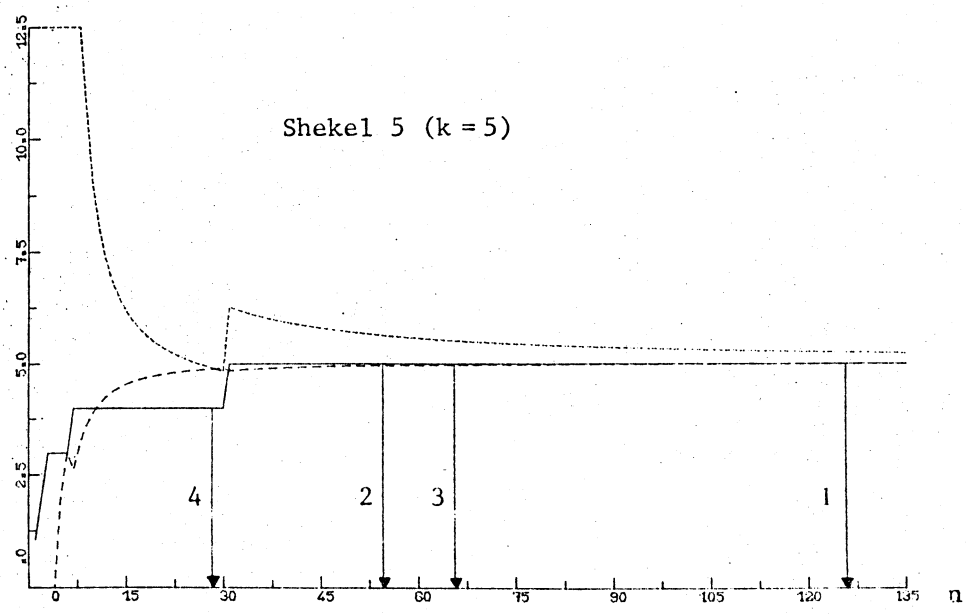
Stopping rule 4. The termination loss is proportional to the total relative volume of the unobserved regions of attraction:

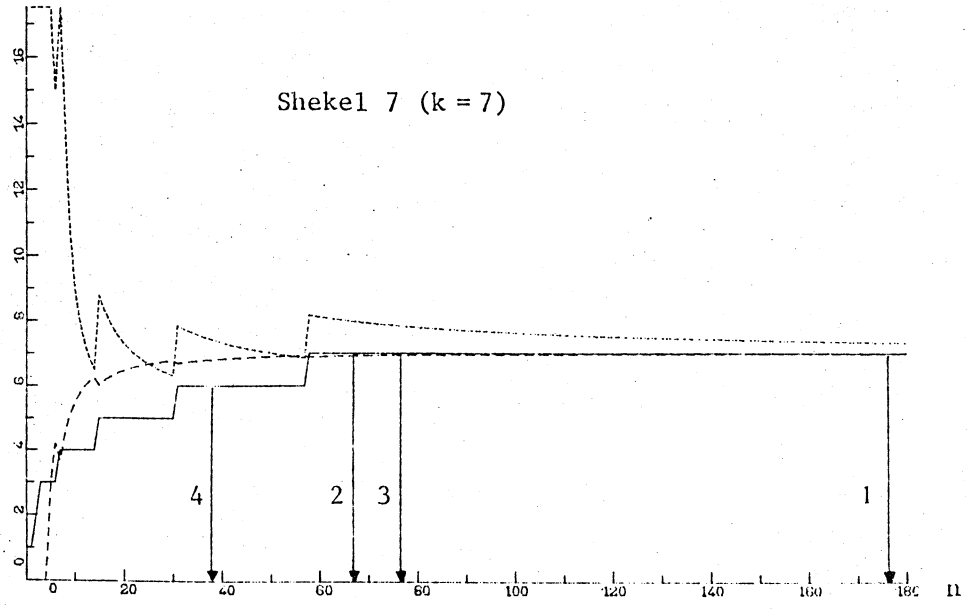
$$L_4 = 1000 * (1 - \Omega) + n.$$


Branin (k = 3)
4
2
3
1
n


Hartman 3 (k = 3)
4
2
3
1
n


Hartman 6 (k = 2)
4
2
3
1
n


Shekel 5 (k = 5)
12,5
10,0
7,5
5,0
2,5
0,0
4
2
3
1
0
15
30
45
60
75
90
105
120
135
n


Shekel 7 (k = 7)
16
14
12
10
8
6
4
2
0
4
2
3
1
0
20
40
60
80
100
120
140
160
180
n

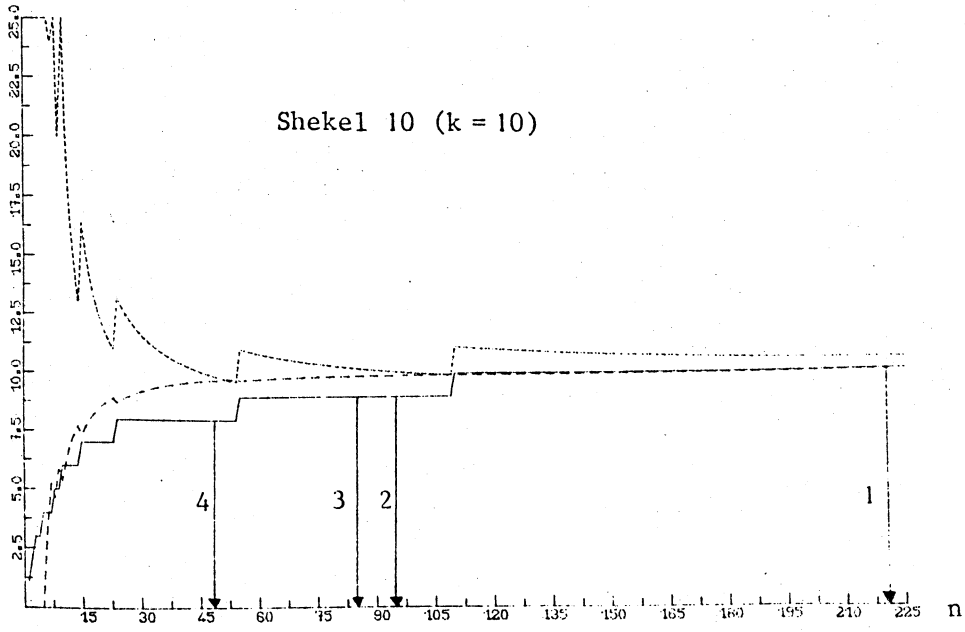

Shekel 10 (k = 10)
25,0
22,5
20,0
17,5
15,0
12,5
10,0
7,5
5,0
2,5
4
3
2
1
15
30
45
60
75
90
105
120
135
150
165
180
195
210
225
n

Figure 3 shows that for each test function the sample size approximately has to be doubled to find a new local minimum. The table also shows that the Bayesian estimates behave as one would hope: the expected value of K converges to the true number of local minima around the time at which all local minima are found, and the expected value of Ω becomes very close to 1 at that point. The convergence appears less rapid for the simpler test problems Branin and Hartman 3 & 6. In interpreting this phenomenon it should be kept in mind that for the simpler test problems the initial prior for K on $[1, \infty)$ may be less appropriate. However (as is always the case in applied Bayesian analysis) the effect of that initial prior drops off, so that for problems where the total number of experiments is not quite miniscule (as would typically be the case in larger practical problems), the behaviour of the Bayesian estimates is as satisfactory as could be achieved from any reasonable prior.

Figure 3 shows interesting differences between the four stopping rules. As expected, rule 1, concentrating on finding all minima, requires the most local searches. Rule 4, which is based on the volume of the observed regions of attraction stops first, and rule 3, whose underlying loss structure depends on the fraction of unobserved minima, is intermediate between these two for all test functions. The one-step rule 2, which is based on a fixed loss reduction if a new minimum is obtained and which prescribes to stop if the probability of this event is too small, turns out to be surprisingly similar to rule 3. However, Figure 2 indicates that for objective functions with many local minima fundamentally different decisions will be taken, since then rules 1, 3 and 4 ultimately will tell the user to stop, whereas the one-step rule, which does not take sampling cost considerations into account, may continue the search forever. On the whole, the stopping rules are successful in coping with the difficulties for which some of the larger dimensional test problems are renowned. An exception occurs on the Shekel-10 function, where rule 1 is the only rule which allows the search to continue long enough to find the local minimum. The price that we pay for this reliability is that the procedure then continues twice as long as necessary.

The choice of an appropriate rule in practice may require some careful thought about the true objectives of the problem solver and about the costs involved. We regard that as an advantage rather than a disadvantage of this approach. Any stopping rule chosen reflects implicit estimates of the costs and benefits involved; it is much better to make these as explicit as possible.

5. Concluding remarks.

We conclude our paper with some remarks on the assumptions underlying our stopping rules and on their domain of applicability.

First of all, the rules that we have developed are based on the uniform prior distribution for the number of local minima k, and the volumes of the regions of attraction $\theta_1,\ldots,\theta_k$. There is a vast statistical literature to support this choice of the uniform prior for cases where the user knows little or nothing. However, if the user has specific prior information about the values of $k,\theta_1,\ldots,\theta_k$ the uniform prior is inappropriate. We can then resort to posterior results derived for other priors [Berbee et al. 1985; Boender & Rinnooy Kan 1983b; Boender 1984]. They require a somewhat larger computational effort and seem less useful in the context of global optimization.

Our second remark concerns the probability distribution of the points from which the local search routine is started. The Multistart method invariably chooses its next starting point from the uniform distribution over S. However, the previous sections reveal that our stopping rules do not crucially depend on this assumption: any other distribution can be accommodated as well. Of course, the probability θ_i of a region of attraction is then equal to its new measure.

Thirdly, in spite of the reliability of Multistart, the method is lacking in efficiency. The inefficiency of Multistart stems from the fact that each local minimum, particularly the ones with a large region of attraction, will generally be found several times. For reasons of efficiency the local search procedure should be performed no more than once in each region of attraction. Computationally successful adaptations of Multistart in that direction are provided by clustering methods [Becker & Lago 1970; Törn 1978; Boender et al. 1982; Timmer 1984]. These methods also generate points in S according to the uniform distribution. Now, however, only a prespecified fraction q containing the points with the lowest function values are retained in the sample. Let f_q be the largest function value in the reduced sample, and define $R_q \subset S$ as the set of all points in S whose function value does not exceed f_q. R_q will consist of a number of disjoint components that together contain all the points from the reduced sample: the reduced sample points that are contained

in one component of R_q form a cluster. Ideally, the clusters should be in 1-1 correspondence with the regions of attraction whose intersection with R_q is nonempty. Then one local search from the best point in each cluster will suffice to find the set of local minima with function value smaller than f_q, which obviously includes the global minimum.

In the global optimization algorithms Single Linkage en Mode Analysis [Timmer 1984] clusters are efficiently identified by exploiting the fact that the points in the reduced sample are uniformly distributed over R_q. For these methods the above ideal situation is virtually achieved: it can be proved that the probability that a local search is started in the region of attraction of a local minimum already observed decreases to 0 with increasing sample size n. Furthermore, for sufficiently large n the probability that no local search is started in a component of R_q containing at least one point from the reduced sample tends to 0 as well.

Fortunately our stopping rules immediately carry over to these computationally superior clustering variants of Multistart. Our second remark implies that they can be properly applied provided that the number of trials is taken equal to the number of points qn in the reduced sample, the number of local minima is taken equal to the number of local minima whose function value is not greater than f_q and the probabilities θ_i are taken to be equal to the relative volume of the intersections of the regions of attraction with R_q. In applying these rules, we only have to assume that each local minimum with function value smaller than f_q whose region of attraction does contain at least one point from the reduced sample is actually found, i.e. that the methods identify the same local minima that would be found by performing a local search from each of the qn points in the reduced sample. As we already pointed out, the Multistart variants guarantee for n sufficiently large that a local search is started from each component of R_q containing at least one point from the reduced sample. A component, however, may contain several local minima, so that the methods may fail to locate a local minimum although its region of attraction contains points from the reduced sample. This deficiency obviously does affect the performance of our stopping rules. It is dealt with by methods called Multi Level Single Linkage and Multi Level Mode Analysis [Timmer 1984] that make explicit use of the function values of the sample points to distinguish between regions of attraction that are contained in one component.

For these methods it can be proved that both the probability that a local search procedure is started unnecessarily and the probability that a local search is not started although a yet undiscovered minimum would be found approach 0 with increasing sample size. Hence, our stopping rules can be applied without any problems to these methods as well. The computational superiority of these Multi Level methods is documented in [Timmer 1984; Rinnooy Kan & Timmer 1985].

Finally, it is obvious from our experiments that there is no single rule that could ever be called the best one. The advantage of our approach is that it allows the user to translate his own preferences with respect to the trade-off between reliability and computational effort in a precise fashion, and given this translation, to arrive at an optimal and yet computationally simple procedure to terminate the search for local minima.

Acknowledgements

The authors wish to thank R. Zielinski for fruitful discussions and useful comments. The research of the second author was partially supported by a NATO Senior Scientist Fellowship.

APPENDIX A

Proof of the Corollaries of Theorem 1.

The proof of (9) is based on the property of the Dirichlet distribution that if $\theta_1,\ldots,\theta_k$ follow a Dirichlet distribution with parameters $\alpha_1,\ldots,\alpha_k$ and $r_1,\ldots,r_\ell$ are integers, then $\Sigma_{i=1}^{r_1}\theta_i,\ldots,\Sigma_{i=r_\ell}^{k}\theta_i$ are also Dirichlet distributed with parameters $\Sigma_{i=1}^{r_1}\alpha_i,\ \ldots,\ \Sigma_{i=r_\ell}^{k}\alpha_i$ (cf. e.g. [Wilks 1962]). Thus, if $\theta_1,\ldots,\theta_k$ are Dirichlet distributed with density function

$$(33)\qquad \frac{(n+k-1)!}{\Pi_{i=1}^{w} n_i!}\ \Pi_{i=1}^{w}\ \theta_{g_i}^{n_i}\ \Pi_{i=w+1}^{k}\ \theta_{g_i}^{0},$$

where $(g_1,\ldots,g_k)$ is an arbitrary permutation of the set $\{1,\ldots,k\}$, then the density of $\omega = \Sigma_{i=1}^{w}\ \theta_{g_i}$ is equal to

$$(34)\qquad \frac{(n+w-1)!(k-w-1)!}{(n+k-1)!}\ \omega^{n+w-1}(1-\omega)^{k-w-1}.$$

Application of this result to the posterior density $p(k,\theta_1,\ldots,\theta_k \mid \{n_1,\ldots,n_w\})$ (4) yields

$$(35)\qquad p(k,\omega)|\{n_1,\ldots,n_w\}) = \frac{(n-1)!(n-2)!(k-1)!}{w!(w-1)!(n-w-2)!}\ x$$

$$x\ \frac{1}{(n+k-1)!}\ \sum_{(g_1,\ldots,g_w)\ \in\ S_k[w]}\ \frac{(n+k-1)!}{(n+w-1)!(k-w-1)!}\ \omega^{n+w-1}(1-\omega)^{k-w-1}$$

$$= \frac{(n-1)!(n-2)!(k-1)!}{w!(w-1)!(n-w-2)!}\ \frac{1}{(n+k-1)!}\ \frac{k!}{(k-w)!}\ \frac{(n+k-1)!}{(n+w-1)!(k-w-1)!}\ \omega^{n+w-1}(1-\omega)^{k-w-1}$$

which is equal to (9) in Corollary 1.1.

The marginal posterior density of k (10) then follows from (9) by integrating out ω, using the formula of the Dirichlet density. Next

$$E(K|\{n_1,\ldots,n_w\}) = \Sigma_{m=w}^{\infty}\ mp(m|\{n_1,\ldots,n_w\})$$

$$(36)\qquad = \frac{(n-1)!(n-2)!}{w!(w-1)!(n-w-2)!}\ \Sigma_{m=w}^{\infty}\ \frac{m!m!}{(n+w-1)!(m-w)!}.$$

Furthermore we have for $n \geq w + 3$

$$(37)\qquad \Sigma_{m=w}^{\infty} \frac{m!m!}{(n+m-1)!(m-w)!} =$$

$$= \Sigma_{m=0}^{\infty} \frac{(m+w)!}{m!(n-2)!} \frac{(m+w)!(n-2)!}{(m+w+n-1)!}$$

$$= \Sigma_{m=0}^{\infty} \frac{(m+w)!}{m!(n-2)!} \int_0^1 t^{m+w}(1-t)^{n-2}dt$$

$$= \frac{w!}{(n-2)!} \int_0^1 t^{w}(1-t)^{n-2} \Sigma_{m=0}^{\infty} \frac{(m+w)!}{m!w!} t^{m}dt$$

$$= \frac{w!}{(n-2)!} \int_0^1 t^{w}(1-t)^{n-w-3}dt$$

$$= \frac{w!w!(n-w-3)!}{(n-2)!(n-2)!}$$

which proves (11). To prove (12) we consider

$$(38)\qquad h(k) = \frac{p(k+1|\{n_1,\ldots,n_w\})}{p(k|\{n_1,\ldots,n_w\})} = \frac{(k+1)k}{(n+k)(k-w+1)}.$$

It can be shown that the unique root of $h(k) = 1$ is $k^* = n(w-1)/(n-w)$, $h(k) > 1$ for $k < k^*$, and $h(k) < 1$ for $k > k^*$, so that the marginal posterior of k attains its maximum at $k^* + 1$, from which (12) follows.
Next

$$(39)\qquad \sigma^2(K|\{n_1,\ldots,n_w\}) = \Sigma_{m=w}^{\infty}(m-E(K|\{n_1,\ldots,n_w\}))^2\, p(m|\{n_1,\ldots,n_w\})$$

$$= \Sigma_{m=w}^{\infty} m^2 p(m|\{n_1,\ldots,n_w\}) - (E(K|\{n_1,\ldots,n_w\}))^2.$$

We note that

$$(40)\qquad \Sigma_{m=w}^{\infty} m^2 p(m|\{n_1,\ldots,n_w\}) = \frac{(n-1)!(n-2)!}{w!(w-1)!(n-w-2)!} \Sigma_{m=w}^{\infty} m^2 \frac{(m-1)!m!}{(n+w-1)!(m-w)!}$$

$$= \frac{(n-1)!(n-2)!}{w!(w-1)!(n-w-2)!} \Sigma_{m=w}^{\infty} \frac{(m+1)!m! - m!m!}{(n+w-1)!(m-w)!}$$

$$= \frac{(n-1)!(n-2)!}{w!(w-1)!(n-w-2)!} \Sigma_{m=w}^{\infty} \frac{(m+1)!m!}{(n+w-1)!(m-w)!} - E(K|\{n_1,\ldots,n_w\})$$

Furthermore it follows from (8) that

$$(41) \qquad \Sigma_{m=w}^{\infty} \frac{(m+1)!m!}{(n+m-1)!(m-w)!} = \Sigma_{m=w+1}^{\infty} \frac{m!(m-1)!}{((n-1)+m-1)!(m-(w+1))!}.$$

$$= \frac{(w+1)!w!(n-w-4)!}{(n-2)!(n-3)!}.$$

Hence, substituting (41) in (40), (11) in (40), and (40) and (11) in (39) we find (13), which concludes the proof of Corollary 1.2.

To prove (14) we observe that it follows from the general posterior density of $K,\theta_1,\ldots,\theta_K$ (4) that

$$(42) \qquad E(\theta_{n_j} | \{n_1,\ldots,n_w\}) = \frac{(n-1)!(n-2)!}{w!(w-1)!(n-w-2)! \prod_{i=1}^{w} n_i!} \times$$

$$\times \Sigma_{m=w}^{\infty} (m-1)! \int \ldots \int_{I_m} \sum_{(g_1,\ldots,g_w)\in S_m[w]} \theta_{g_j} \prod_{i=1}^{w} \theta_{g_i}^{n_i} \prod_{i=1}^{m} d\theta_i$$

$$= \Sigma_{m=w}^{\infty} \frac{(n-1)!(n-2)!(m-1)!}{w!(w-1)!(n-w-2)! \prod_{i=1}^{w} n_i!} \sum_{(g_1,\ldots,g_w)\in S_m[w]} \frac{(n_j+1) \prod_{i=1}^{w} n_i!}{(n+m)!}$$

$$= \frac{(n-1)!(n-2)!(n_j+1)}{w!(w-1)!(n-w-2)!} \Sigma_{m=w}^{\infty} \frac{(m-1)!m!}{(m-w)!(n+m)!}$$

$$= \frac{(n_j+1)(n-w-1)}{n(n-1)}.$$

With respect to Corollary 1.4 we use the fact that the expected value of a sum of random variables is equal to the sum of the expected values. Application of this result to the expected values (14) immediately yields (15).
Finally

$$(43) \qquad \sigma^2(\Omega | \{n_1,\ldots,n_w\}) = E(\Omega^2 | \{n_1,\ldots,n_w\}) - (E(\Omega | \{n_1,\ldots,n_w\}))^2$$

$$= E(\Sigma_{i=1}^{w} \Sigma_{j=1}^{w} \theta_{n_i} \theta_{n_j} | \{n_1,\ldots,n_w\}) - (E(\Omega | \{n_1,\ldots,n_w\}))^2$$

$$= \Sigma_{j=1}^{w} E(\theta_{n_j}^2 | \{n_1,\ldots,n_w\}) + \Sigma_{i=1}^{w} \Sigma_{\substack{j=1 \\ j\neq i}}^{w} E(\theta_{n_i}\theta_{n_j} | \{n_1,\ldots,n_w\}) -$$

$$- (E(\Omega | \{n_1,\ldots,n_w\}))^2.$$

Analogously to (42) we have that

$$(44) \qquad E(\theta_{n_j}^2 | \{n_1,\ldots,n_w\}) = \frac{(n-1)!(n-2)!(n_j+1)(n_j+2)}{w!(w-1)!(n-w-2)!} \Sigma_{m=w}^{\infty} \frac{(m-1)!m!}{(m-w)!(n+m-1)!}$$

$$= \frac{(n_j+1)(n_j+2)(n-w-1)(n-w)}{(n-1)n^2(n+1)},$$

and

$$(45) \qquad E(\theta_{n_i}\theta_{n_j} | \{n_1,\ldots,n_w\}) =$$

$$= \frac{(n-1)!(n-2)!(n_i+1)(n_j+1)}{w!(w-1)!(n-w-2)!} \Sigma_{m=w}^{\infty} \frac{(m-1)!m!}{(m-w)!(n+m+1)!}$$

$$= \frac{(n_i+1)(n_j+1)(n-w-1)(n-w)}{(n-1)n^2(n+1)}.$$

Substitution of (15), (44) and (45) in (43) yields the desired result (16).

APPENDIX B.

Proof of Theorem 2.

The proof of (29) is given in [Boender & Zielinski 1982; Boender & Rinnooy Kan 1893a].

The proof of (30) is obtained by substituting (23) in (28).

For the proof of (31) we substitute (25) in (28). Then

$$(46)\qquad E(E(L_3|(n+1,W))|(n,w)) \geq E(L_3|(n,w)) - c_3 \frac{w(n-w-1)}{n(n-1)} + 1$$

$$\geq E(L_3|(n,w)) + c_3 \frac{((n-1)/2)^2 - (n-1)^2/2}{n^2(n-1)} + 1$$

$$\geq E(L_3|(n,w)) - \frac{c_3}{4n} + 1,$$

from which (31) follows.

Substitution of (27) in (28) yields

$$(47)\qquad E(E(L_4|(n+1,W))|(n,w)) \geq E(L_4|(n,w)) - 2c_4 \frac{w(w+1)(n-w-1)}{n^2(n-1)(n+1)} + 1$$

$$\geq E(L_4|(n,w)) + 2c_4 \frac{(w+1)^2(w+1-n)}{n^2(n-1)(n+1)} + 1$$

$$\geq E(L_4|(n,w) - 2c_4 \frac{(2n/3)^2(2n/3 - n)}{n^2(n-1)(n+1)} + 1$$

$$\geq E(L_4|(n,w)) - \frac{c_4}{3(n-1)} + 1,$$

which proves (32).

References

Becker, R.W. & G.V. Lago (1970), A global optimization algorithm, in: Proceedings of the 8th Allerton conference on circuits and system theory.

Berbee, H.C.P., C.G.E. Boender, A.H.G. Rinnooy Kan, C.L. Scheffer, R.L. Smith & J. Telgen (1985), Hit-and-run algorithms for the identification of nonredundant linear inequalities, report 8504/0, Econometric Institute, Erasmus University Rotterdam.

Boender, C.G.E., A.H.G. Rinnooy Kan, L. Stougie & G.T. Timmer (1982), A stochastic method for global optimization, Mathematical Programming 22, 125-140.

Boender, C.G.E. & R. Zielinski (1982), A sequential Bayesian approach to estimating the dimension of multinomial distribution, preprint 256, Institute of Mathematics, Polish Academy of Sciences, to appear in the proceedings of the 18th Banach Center Semester of Mathematical Statistics, Warsaw.

Boender, C.G.E. & A.H.G. Rinnooy Kan (1983a), A Bayesian analysis of the number of cells of a multinomial distribution, The Statistician 32, 240-248.

Boender, C.G.E. & A.H.G. Rinnooy Kan (1983b), Bayesian multinomial estimation of animal population size, report 8322/0, Econometric Institute, Erasmus University Rotterdam.

Boender, C.G.E. (1984), The generalized multinomial distribution: A Bayesian analysis and applications, Ph.D. Thesis, Erasmus Universiteit Rotterdam.

Dixon, L.C.W. & G.P. Szegö (eds.) (1975), Towards global optimisation, North-Holland, Amsterdam.

Dixon, L.C.W. & G.P. Szegö (eds.) (1978), Towards global optimisation 2, North-Holland, Amsterdam.

De Groot, M.H. (1970), Optimal statistical decisions, McGraw-Hill, New York.

Rinnooy Kan, A.H.G. & G.T. Timmer (1984), A stochastic approach to global optimization, report 8419/0, Econometric Institute, Erasmus Universiteit Rotterdam.

Rinnooy Kan, A.H.G. & G.T. Timmer (1985), "Stochastic Global Optimization Methods, Parts I & II", Econometric Institute, Erasmus University Rotterdam.

Timmer, G.T. (1984), Global optimization: a stochastic approach, Ph.D. Thesis, Erasmus University Rotterdam.

Wilks, S.S. (1962), Mathematical Statistics, Wiley, New York.

Zielinski, R. (1981), A statistical estimate of the structure of multi-extremal problems, Mathematical Programming 21, 348-356.

LIST OF REPORTS 1985

8500 "Publications of the Econometric Institute Second Half 1984: List of Reprints 378-400, Abstracts of Reports".

8501/O **R.M. Karp, J.K. Lenstra, C.J.H. McDiarmid and A.H.G. Rinnooy Kan,** "Probabilistic analysis of combinatorial algorithms: an annotated bibliography", 26 pages.

8502/E **M.E. Homan,** "Verschillen in consumptie tussen één en tweekostwinnerhuishoudens: een eerste analyse", 13 pages.

8503/O **A.W.J. Kolen,** "The round-trip p-center and covering problem on a tree", 17 pages.

8504/O **H.C.P. Berbee, C.G.E. Boender, A.H.G. Rinnooy Kan, C.L. Scheffer, R.L. Smith and J. Telgen,** "Hit-and-run algorithms for the identification of nonredundant linear inequalities", 32 pages.

8505/S **H. Brozius and L. de Haan,** "On limiting laws for the convex hull of a sample", 10 pages.

8506/O **J.B.G. Frenk and A.H.G. Rinnooy Kan,** "On the rate of convergence, to optimality of the LPT rule - postscript", 5 pages.

8507/E **P. Kooiman, H.K. van Dijk and A.R. Thurik,** "Likelihood diagnostics and Bayesian analysis of a micro-economic disequilibrium model for retail services", 35 pages.

8508/O **C.G.E. Boender, A.H.G. Rinnooy Kan and J.R. de Wit,** "A Bayesian procedure for the (s,Q) inventory problem", 26 pages.

8509/E **B.M.S. van Praag, S. Dubnoff and N.L. van der Sar,** "From judgments to norms: measuring the social meaning of income, age and educaton", 35 pages.

8510/S **B. Bode, J. Koerts and A.R. Thurik,** "On shopkeepers' pricing behaviour", 19 pages.

8511/M **H. Bart, I. Gohberg and M.A. Kaashoek,** "Exponentially dichotomous operators and inverse Fourier transforms", 70 pages.

8512/M **J. Brinkhuis,** "Normal integral bases and embedding of fields", 13 pages.

8513/E **P.M.C. de Boer,** "On the relationship between Revankar's and Lu-Fletcher's production function", 5 pages.

8514/S **L. de Haan,** "Extremes in higher dimensions: the model and some statistics", 15 pages.

8515 Publications of the Econometric Intitute First Half 1985: List of Reprints 401-414, Abstracts of Reports.

8516/A **A.R. Thurik and A. Kleijweg,** "Cyclical effects in retail labour productivity", 20 pages.

8517/C **B.M.S. van Praag and J. van Weeren,** "The impact of past experiences and anticipated future on individual income judgements", 24 pages.

8518/A **A.R. Thurik and J. Koerts,** "Behaviour of retail entrepreneurs". 16 pages.

8519/B **M. Hazewinkel, J.F. Kaashoek and B. Leynse,** "Pattern formation for a one dimensional evolution equation based on Thom's River Basin Model". 24 pages.

8520/A **H.K. van Dijk, T. Kloek and C.G.E. Boender,** "Posterior moments computed by mixed integration", 25 pages.

8521/B **J. Brinkhuis,** "Testing concavity and quasi-concavity is easy", 12 pages.

8522/A **R. Harkema,** "Minimum sample size requirements for maximum likelihood estimation of some demand models", 25 pages.

8523/B **A.C.F. Vorst,** "The general linear group of discrete Hodge algebras", 10 pages.

8524/A **A.M.H. Gerards and A.W.J. Kolen,** "Polyhedral combinatorics in combinatorial optimization", 25 pages.

8525/A **B.J. Lageweg, J.K. Lenstra, A.H.G. Rinnooy Kan and L. Stougie,** "Stochastic integer programming by dynamic programming", 19 pages.

8526/A **J.B. Orlin,** "A dual version of Tardo's algorithm for linear programming", 9 pages.

8527/A **B.M.S. van Praag,** "Household cost functions and equivalence scales", 24 pages (report 8424/E revised)

8528/A **S. Schim van der Loeff,** "Limited information maximum likelihood estimation of a subsystem of nonlinear equations", .. pages.

8529/A **B.S. van der Laan and A.S. Louter,** "On the number and the amount of damage of a passenger car traffic accidents in the Netherlands", ..pages.

8530/A **B.S. van der Laan and A.S Louter,** "A statistical model for the costs of passenger car traffic accidents", .. pages.

8531/A **B.M.S. van Praag, J. de Leeuw and T. Kloek,** "The population-sample decomposition approach to multivariate estimation methods", 34 pages.

8532/A **M.E. Homan, B.M.S. van Praag and A.J.M. Hagenaars,** "Household cost functions and the value of home production in one-and two earner families", .. pages.

8533/B **R.J. Stroeker,** "An inequality for YFF's analogue of the Brocard angel of a plane triangle", 18 pages.

8534/A **J. Bouman,** "Testing nonnested linear hypothesis: A Bayesian approach based on imcomletely specified prior distributions". pages.

8535/A **P.M.C. de Boer, R. Harkema and B.J. van Heeswijk,** "Estimating foreign trade functions, a comment and a correction". pages.

8536/A **C.G.E. Boender and A.H.G. Rinnooy Kan,** "Bayesian stopping rules for multistart global optimization methods", 30 pages.

Until report 8515 this series was devided in E(conometrics), S(tatistics), M(athematics) and O(perations Research). From report 8516 on it will be

A. Economics, Econometrics and Operations Research

B. Mathematics

C. Miscellaneous