

**UNIVERSITÄT DORTMUND**

---

REIHE COMPUTATIONAL INTELLIGENCE

---

SONDERFORSCHUNGSBEREICH 531

---

Design und Management komplexer technischer Prozesse  
und Systeme mit Methoden der Computational Intelligence

---

Controlled Model Assisted Evolution Strategy with  
Adaptive Preselection

Frank Hoffmann and Sebastian Hölemann

Nr. CI-209/06

Interner Bericht

ISSN 1433-3325

June 2006

Sekretariat des SFB 531 · Universität Dortmund · Fachbereich Informatik/XI  
44221 Dortmund · Germany

---

Diese Arbeit ist im Sonderforschungsbereich 531, „Computational Intelligence“, der Universität Dortmund entstanden und wurde auf seine Veranlassung unter Verwendung der ihm von der Deutschen Forschungsgemeinschaft zur Verfügung gestellten Mittel gedruckt.

# Controlled Model Assisted Evolution Strategy with Adaptive Preselection

Frank Hoffmann, *Member, IEEE*, and Sebastian Hölemann

**Abstract**—The utility of evolutionary algorithms for direct optimization of real processes or complex simulations is often limited by the large number of required fitness evaluations. Model assisted evolutionary algorithms economize on actual fitness evaluations by partially selecting individuals on the basis of a computationally less complex fitness model. We propose a novel model management scheme to regulate the number of preselected individuals to achieve optimal evolutionary progress with a minimal number of fitness evaluations. The number of preselected individuals is adapted to the model quality expressed by its ability to correctly predict the best individuals. The method achieves a substantial reduction of fitness evaluations on a set of benchmarks not only in comparison to a standard evolution strategy but also with respect to other model assisted optimization schemes.

## I. INTRODUCTION

Evolutionary algorithms solve global, complex, high-dimensional, multi-modal optimization problems without an explicit analytical description of the underlying fitness function. This property makes them particularly suitable for direct optimization of real processes, for example in the context of hard-ware-in-the-loop control system design. However, due to the large number of evaluations evolutionary algorithms meet limitations if the computation of the fitness functions is time-consuming or expensive. This observation motivates the substitution of costly true fitness evaluations by an approximate fitness model which is generated from observations of the true fitness function [1], [2], [3]. It is assumed that the computational cost to generate and query the model is negligible compared to the cost of true evaluations.

Fitness modeling has been investigated in several publications as a means to accelerate evolutionary optimization [1], [2]. Model management is concerned with the decision of which individuals are evaluated on the true fitness function and which ones are solely selected based on their predicted fitness values. Increasing the number of evaluated individuals provides additional training data to generate a more accurate model, thereby reducing the risk of premature convergence of optimization due to misleading minima of the fitness model.

The paper is organized as follows. Section II introduces the concept of model assisted evolution strategies. Section III compares the utility and performance of several instance based learning schemes in the context of fitness modeling. Section IV describes the novel scheme to regulate the number of preselected individuals within the model assisted evolution strategy with the objective of optimal exploitation of fitness

evaluations. The results in section V demonstrate that the approach is efficient and robust as it reduces the number of evaluations substantially. The paper concludes with a summary in section VI.

## II. MODEL ASSISTED EVOLUTION STRATEGY

The idea of a model assisted evolution strategy (MAES) is to generate a larger number  $\lambda_p > \lambda$  of offspring but only to evaluate the most promising candidates on the true fitness function [4].

Preselection picks the  $\lambda$  best individuals according to the fitness predicted by the model, which are then evaluated on the true fitness function to determine  $\mu$  parents. That way all parents are guaranteed to be tested on the true fitness function which avoids convergence of the optimization to misleading minima introduced by the fitness model. The  $\lambda$  true fitness evaluations at each generation provide additional training data to refine the fitness model. Notice, that the total number of true fitness evaluations, namely  $\lambda$  tests per generation, remains the same. The advance in convergence of the MAES compared to a standard ES is attributed to the larger number of  $\lambda_p$  candidates, which increases the probability to find solutions that are superior to the previous generation. Therefore, an MAES is expected to find better solutions with the same amount of computational effort for fitness evaluations. However, this improvement is only achievable if the actually best  $\mu$  individuals survive the model based preselection. The current quality of the fitness model does not depend primarily on the residual model error, but rather on its ability to correctly predict the ranking of individuals in the context of selection. In [4] the authors come to the conclusion, that a MAES already advantageous if the model based preselection performs better than a purely random selection, as on average superior individuals undergo the true fitness test. This trend is reversed in case of a misleading model which performs worse than random selection. This observation leads to the idea of controlled model assisted evolution strategies which dynamically adapt the number of offspring  $\lambda_p$  to the accuracy of the model during evolution. If the model performs better than random selection it is safe to increase  $\lambda_p$  and thereby the impact of the model on the evolution strategy. Vice versa  $\lambda_p$  is reduced if the model performs worse than random selection. The novelty of our approach is to regulate  $\lambda$  instead of  $\lambda_p$  based on the model quality. From a practical perspective evolutionary progress should be measured in terms of true fitness evaluations rather than number of generations as the former effectively determine the cost and temporal demands of the optimization. Our results demonstrate that controlling

$\lambda$  substantially reduces the number of fitness evaluations not only with respect to a standard ES, but also with respect to other variants of CMAES.

In order to evaluate the quality of the preselection it seems at first necessary to determine the true fitness values of all  $\lambda_p$  individuals in order to compare the predicted and the true ranking. Instead the quality of the model is determined with respect to the  $\mu$  out of  $\lambda$ -selection for which the true fitness values of the preselected individuals are known. The  $\lambda$  individuals are ranked according to their true fitness and a weight  $(\lambda - i)$  is associated with the  $i$ -th individual which reflects the notion that the model should identify the highest ranked individuals. The quality of the model does not only consider the number of correctly selected individuals but also their rank, and is calculated by summing their weights

$$Q = \sum_{i=1}^{\lambda} g_i(\lambda - i) \quad (1)$$

in which  $g_i = 1$  if the  $i$ -th individual is correctly selected and  $g_i = 0$  otherwise. The quality distinguishes according to the true rank, e.g. missing out on the de facto best individual is more costly than missing out on the second best individual. In case none of the de facto parents is selected the quality becomes zero, in case all  $\mu$  parents are selected correctly the quality becomes

$$Q^{max} = \sum_{i=1}^{\mu} (\lambda - i) = \mu\lambda - \frac{\mu(\mu + 1)}{2} \quad (2)$$

For a random selection of individuals the expected value of model quality becomes

$$Q^{rand} = \frac{\mu^2}{\lambda} \frac{2\lambda - \mu - 1}{2} \quad (3)$$

These quality measures form the basis for adapting the number of individuals that are subject to the true fitness evaluation to quality of the model as described in section IV.

### III. COMPARATIVE ANALYSIS OF FITNESS MODELS

Fitness modeling is concerned of approximating an unknown target function based on training data generated from true fitness evaluations. This section compares the utility and performance of different supervised learning methods in the context of fitness approximation. In the case of MAES new fitness values only become available during evolution, which requires online learning with the model subject to stepwise refinement. This requirement excludes off-line learning methods, such as neural networks which necessitate retraining with the entire batch of data. In the context of model based preselection one does not seek a global approximation but rather emphasizes local models that approximate the fitness function in the region of search space populated by the current generation [5]. Local models do not build an explicit representation of the underlying function but merely store the training data. For each query they generate a local model that is valid in the vicinity of the query point. The contribution of a stored training pair to the local model is weighted by its distance to the query point. The local methods differ in

the type of local model used for approximation and in the distance based kernel function that determines the weights. In the following we describe local linear weighted regression in detail, as this method was eventually employed within the controlled model assisted evolution strategy described in the next two sections.

Local linear regression (LWR) is derived from standard regression techniques. The data points are weighted by their distance to the query point in the regression step. The idea is to attribute more relevance to data points that are close to the query point and thereby increasing their influence on the approximating function. That way the approximating function locally adapts to the vicinity of the query point  $\mathbf{q}$  whereas the quality of approximation decreases with increasing distance to  $\mathbf{q}$ . This makes sense as a local fitness model suffices to predict the ranking of the current generation that usually occupies a limited region of search space. Local regression minimizes an error function of the form

$$\epsilon = \sum_{i=1}^k \left( f(\mathbf{x}^{(i)}) - \hat{f}(\mathbf{x}^{(i)}) \right)^2 \cdot w_i \quad (4)$$

in which the weight  $w_i = w_i(\Delta(\mathbf{x}^{(i)}, \mathbf{q}))$  is a function of the distance between the  $i$ -th data point and the query point. In our case the weight is determined by a Gaussian kernel function

$$w_i(\Delta(\mathbf{x}^{(i)}, \mathbf{q})) = e^{-\Delta(\mathbf{x}^{(i)}, \mathbf{q})^2} \quad (5)$$

and the distance between two points is defined by a scaled Euclidean distance

$$\Delta(\mathbf{x}, \mathbf{q}) = \frac{\sqrt{\sum_{l=1}^n (x_l - q_l)^2}}{h} \quad (6)$$

The scaling factor  $h$  is adapted to the distribution of data points which strongly depends on the state of convergence of the population. In principle, the optimal value for  $h$  can be determined by means of leave-one-out cross-validation. For the sake of computational efficiency we suggest a simpler adaptation mechanism in which  $h$  is simply given by the distance between the query point  $\mathbf{q}$  and its nearest neighbor.

In local linear weighted regression the approximating function

$$\hat{f}(\mathbf{x}) = \beta_0 + \sum_{l=1}^n \beta_l x_l \quad (7)$$

is linear in the data  $\mathbf{x}$  as well as the parameters  $\beta$  of the model. The error functional from equation 4 is given by a quadratic form

$$\epsilon(\beta) = (\mathbf{y} - \mathbf{H} \cdot \beta)^T \mathbf{W} (\mathbf{y} - \mathbf{H} \cdot \beta) \quad (8)$$

in which

$$\mathbf{H} = \begin{pmatrix} 1 & x_1^{(1)} & \cdots & x_n^{(1)} \\ 1 & x_1^{(2)} & \cdots & x_n^{(2)} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_1^{(k)} & \cdots & x_n^{(k)} \end{pmatrix} \quad (9)$$

is the matrix of data points and  $\mathbf{W} = \text{diag}((w_1, \dots, w_k)^T)$  is a diagonal weight matrix with  $w_i = (\Delta(\mathbf{x}^{(i)}, \mathbf{q}))$ . The optimal

parameter vector  $\beta$  that minimizes 8 is obtained by means of least squares

$$\beta = (\mathbf{H}^T \mathbf{W} \mathbf{H})^{-1} \mathbf{H}^T \mathbf{W} \mathbf{y} \quad (10)$$

Distance weighted averaging operates with a constant regression function

$$\hat{f}(\mathbf{x}) = \beta_0 \quad (11)$$

Nearest neighbor constitute the computationally most simple type of local models as the predicted value at the query point is taken from its nearest neighbor (1-NN) or as the average of its  $N$  nearest neighbors. In other words, the weight with which a training example contributes is either zero or one and is not directly modulated by the distance.

Whereas local linear regression is able to extrapolate beyond the distribution of data points distance weighted average and nearest neighbor methods only interpolate between data points. This constitutes problems as the current population usually exhibits a different distribution than previous generations from which the training examples emerge. Our experimental analysis confirms this concern.

Gaussian processes have been widely used in fitness modeling [6],[3]. Gaussian processes offer the advantage that in addition to the fitness estimate itself they also provide an estimate of the uncertainty of the model at the particular query point. Based on the information about the variance of the estimate it is possible to compute a probability of improvement (POI) which states the likelihood that the true but unknown fitness value actually exceeds a threshold  $f_{min}$ . In [6] the preselection step identifies precisely those candidates that possess the largest probability of improvement with respect to the current best fitness. A detailed discussion of Gaussian processes is beyond the scope of this paper. The interested reader is referred to [7] for a detailed introduction.

In the following the performance of the methods nearest neighbors (NN), distance weighted averaging (DWA), linear weighted regression (LWR) and Gaussian processes (GP) is evaluated in the context of MAES on a set of four artificial benchmarks problems, namely the sphere function

$$f(\mathbf{x}) = \sum_{l=1}^n x_l^2 \quad (12)$$

the Rosenbrock function

$$f(\mathbf{x}) = \sum_{l=1}^{n-1} \left( 10(x_l^2 - x_{l+1})^2 + (x_l - 1)^2 \right) \quad (13)$$

the Griewank function

$$f(\mathbf{x}) = 1 + \frac{1}{100} \sum_{l=1}^n x_l^2 - \prod_{l=1}^n \cos\left(\frac{x_l}{\sqrt{l}}\right) \quad (14)$$

and the Ackley function

$$f(\mathbf{x}) = -20 \exp\left(-0.2 \frac{1}{n} \sqrt{\sum_{l=1}^n x_l^2}\right) - \exp\left(\frac{1}{n} \sum_{l=1}^n \cos(2\pi x_l)\right) + 20 + e \quad (15)$$

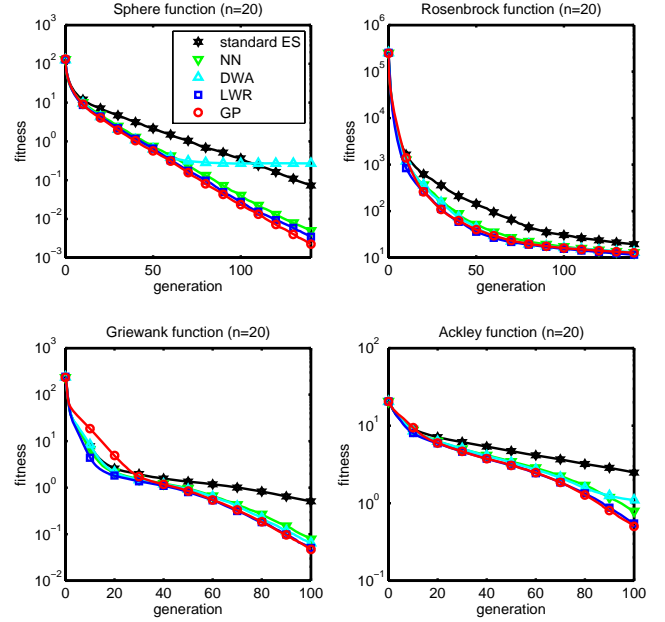


Fig. 1. Evolution of fitness under MAES for nearest neighbor (NN), distance weighted averaging (DWA), linear weighted regression (LWR) and Gaussian processes (GP) and a standard evolution strategy (ES) on a set of benchmark optimization problems.

The sphere and Rosenbrock function are unimodal, whereas the Griewank and Ackley function represent multi-modal optimization problems with potentially misleading local minima. All functions assume their global minimum at  $f_{min} = f(\mathbf{0}) = 0$  except for the Rosenbrock function with  $f_{min} = f(\mathbf{0}) = 1$ . In all cases the results are based on an average of 100 runs and the dimension of the search space is  $n = 20$ . Figure 1 shows the evolution of fitness for a MAES with constant  $\lambda = \lambda_p/2$ . It becomes obvious that for some optimization problems the choice of the approximation function has a significant impact on the convergence behavior of the MAES. All models perform equally well on the Rosenbrock function. On the other hand DWA is prone to cause premature convergence on the sphere function and in the long run is actually outperformed even by the standard ES. In general DWA and NN models are inferior to LWR and GP models. The later two perform similar across all problems, the GP shows slightly superior convergence in case of the sphere and the Ackley function. This minor advantage is annihilated by the significantly increased computational complexity of GP compared to LWR, which is the reason for analyzing the controlled model assisted evolution schemes on the basis of LWR models. In a practical application, that does not require the extra burden of multiple runs to obtain reliable statistics GP models might still be the first option.

A more detailed analysis of the interaction between model quality and correct selection of parents reveals the shortcomings of the DWA and NN models. The crucial factor of MAES is its ability to reliably preselect the de facto best individuals. For that purpose figure 2 depicts the ratio  $\mu_r/\mu$  of correctly selected parents, in which  $\mu_r$  is the fraction of

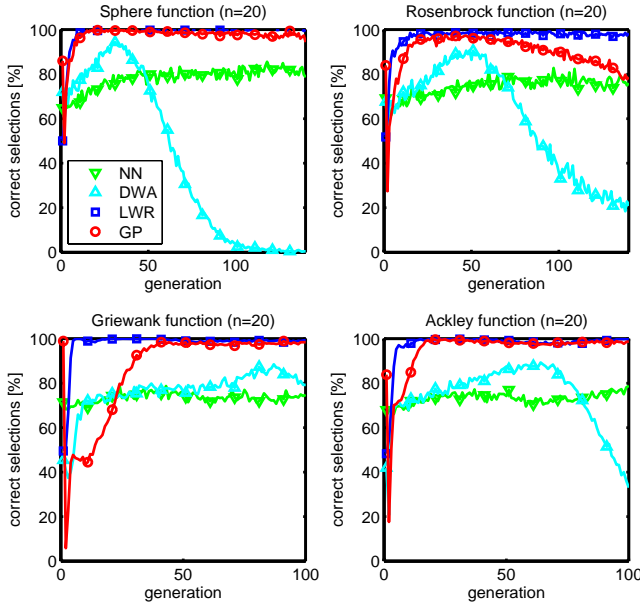


Fig. 2. Analysis of the accuracy of the preselection step in terms of the ratio of correctly selected parents under a MAES for nearest neighbor (NN), distance weighted averaging (DWA), linear weighted regression (LWR) and Gaussian processes (GP).

preselected individuals from the set of the  $\mu$  actual parents assuming full knowledge. Notice, that a de facto parent is correctly selected as long as it appears among the  $\lambda$  preselected individuals as then the subsequent fitness evaluation reveals its true quality. The poor convergence behavior of DWA in case of the sphere function becomes apparent from the decline in correctly selected individuals after the 50th generation. In this case the model becomes counterproductive as the rate drops below the 50% that a random selection of a model free standard ES achieves on average. For the GP and the LWR model almost 100% of the true parents are preselected. Only in case of the Rosenbrock function the rate decreases for the GP model during the final generations. The NN model is reliable in so far as it permanently outperforms a random selection but is only marginal efficient with a correct selection rate of about 70%. The results confirm the hypothesis that the MAES improves the convergence as long as the preselection performs better than a random selection. They also show that the gain in convergence increases with the accuracy and robustness of the preselection process. Figure 3 depicts how the rate of incorrect selections depends on the number of preselected individuals. It reveals that the MAES might operate with a substantially smaller number of preselected individuals  $\lambda < \lambda_p/2$  without a substantial sacrifice on the rate of correctly selected parents. This observation advocates a strategy to regulate  $\lambda$  based on the quality of the fitness model in comparison to a random selection process to achieve an optimal trade-off between the cost of additional true fitness evaluations and the risk of rejecting potential parents due to an erroneous fitness model.

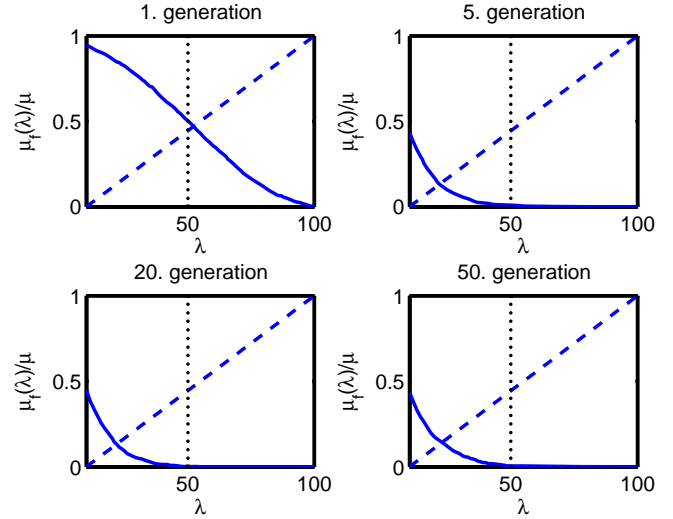


Fig. 3. Rate of incorrect selections as a function of  $\lambda$  for at 1st, 5th, 20th and 50th generation for  $\mu = 10$ ,  $\lambda_p = 100$ . The dashed line corresponds to  $\lambda = \lambda_p/2 = 50$ .

#### IV. $\lambda$ -CONTROLLED MODEL ASSISTED EVOLUTION STRATEGY

Our experiments indicated that the strategy of controlling the number of offspring  $\lambda_p$  is not successful for the task of regulating the impact of the fitness model on the evolutionary process. The data in figure 2 shows that at least for GP and LWR and  $\lambda = \lambda_p/2$  almost 100

In order to guarantee that the model based preselection performs better than random selection we propose a strategy, named  $\lambda$ -CMAES that regulates  $\lambda$  in order to maintain model quality. The number of offspring  $\lambda_p$  is kept constant. The key idea is to prefer model based selection over true fitness evaluations as long as the model is able to discriminate among the best individuals. At each generation  $\lambda$ -CMAES compares the actual model quality  $Q^g$  observed over the set of preselected individuals with the expected quality of a random selection according to Eq. 3. If  $Q^g > Q^{rand}$  the current model performs better than random selection and the number of evaluated individuals in the next generation  $g + 1$

$$\lambda^{g+1} = \max \left( \lambda^g - \frac{Q^{max} - Q^g}{\langle Q^{rand} \rangle} \delta_\lambda, \mu \right) \quad (16)$$

is reduced. The gain  $\delta_\lambda$  determines the rate of adaptation and is chosen small enough to suppress random fluctuations. In contrast, if the current model quality falls below the random selection quality  $Q^g < Q^{rand}$ , the evolution puts less trust into the model by increasing the number of true evaluations according to

$$\lambda^{g+1} = \min \left( \lambda^g + \frac{Q^{max} - Q^g}{Q^{max} - \langle Q^{rand} \rangle} \delta_\lambda, \lambda_p \right) \quad (17)$$

$\lambda$ -CMAES aims to adjust the quality of the preselection process to that of a random selection. This raises the question of how  $\lambda$ -CMAES then still performs better than a standard ES. Notice, that the selection quality is measured with respect

to the  $\lambda$  individuals, but that the model is expected to distinguish fitness between the overall  $\lambda_p$  offspring. A quality of  $Q^g > Q^{rand}$  means that the information provided by the model is not fully utilized for preselection, as some individuals with lower expected fitness are still evaluated on the true fitness function. It is more economic to spare these evaluations and consequently only evaluate individuals that the model fails to discriminate in terms of fitness. The results presented in the next section demonstrate that  $\lambda$ -CMAES achieves a faster yet robust convergence and that the expected benefit is actually realized. Despite the reduced number of fitness evaluations the evolution scheme generates a sufficient number of training data to improve the model.

In order to have a sufficient statistical basis for the evaluation of model quality according to eq. 1 the lower limit for  $\lambda$  was initially set to  $\lambda_\mu = 2\mu$ . However, in some of the benchmark problems  $\lambda$ -CMAES attains this lower limit within a few generations, which indicates a saturation of  $\lambda$  even though according to the control  $\lambda$  might be further reduced. Therefore, the lower limit for  $\lambda$  is reduced to  $\lambda = \mu$ . This limit might cause instability of the  $\lambda$ -regulation, in particular in the case  $\lambda = \mu$  all individuals are guaranteed to be correctly selected as parents. Therefore, the quality of the model in case of  $\lambda < 2\mu$  is evaluated on the basis of  $\lambda/2$  out  $\lambda$ -selection. That way there are at least  $\mu/2$  independent samples available to assess the model quality. The equations 2 and 3 that determine  $Q^{max}$  and  $Q^{rand}$  are modified as  $\mu$  is replaced by the nominal value

$$\tilde{\mu} = \begin{cases} \mu, & \text{if } \lambda \geq 2\mu \\ \frac{\lambda}{2}, & \text{if } \lambda < 2\mu \end{cases} \quad (18)$$

## V. RESULTS

The performance analysis of  $\lambda$ -CMAES uses the set of benchmark optimization problems introduced in section III and is based on the average of hundred test runs in each scenario. The plots in figure 4 compare the evolution of fitness between the  $\lambda$ -CMAES, a standard ES and a MAES with constant  $\lambda = \lambda_p/2$ . The plots also contain the evolution of  $\lambda(g)$  over the generations, the horizontal dashed lines correspond to  $\lambda = \mu$  and  $\lambda = 2\mu$ . The adaption rate is given by  $\delta_\lambda = \mu/2$  and the initial value is  $\lambda(0) = \lambda_p/2$ . As to be expected  $\lambda$ -CMAES performs better than a standard evolution strategy but is outperformed by MAES if progress is measured in terms of elapsed generations. The evolution of  $\lambda(g)$  indicates that  $\lambda$ -CMAES requires substantially fewer actual fitness evaluations compared to MAES. After a few generations the number of fitness evaluations decreases rapidly and then stabilizes at a value of approximately  $\lambda \approx 2\mu$  for the unimodal problems and  $\lambda \approx 1.5\mu$  for the multi-modal problems.

As the number of fitness evaluations per generation is no longer constant it only seems fair to compare  $\lambda$ -CMAES with MAES on the basis of actual fitness evaluations as these determine the eventual effort of an evolutionary optimization. Figure 5 depicts the evolution of fitness in terms of actual fitness evaluations and reveals the superiority of  $\lambda$ -CMAES. with the same number of evaluations  $\lambda$ -CMAES achieves

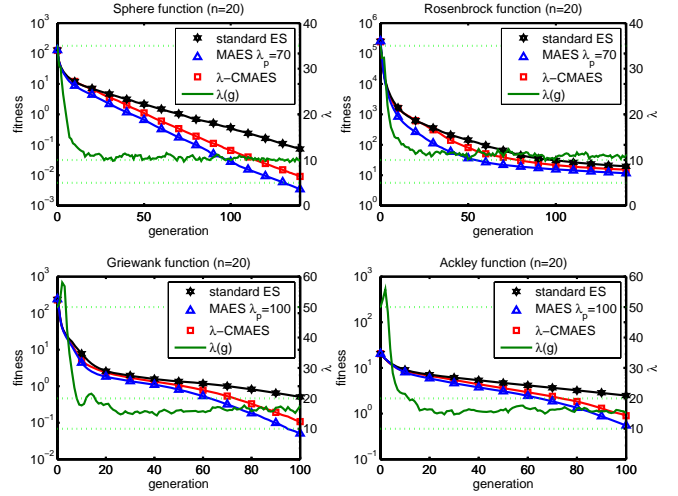


Fig. 4. Evolution of fitness for the benchmark functions in terms of number of generations for a standard ES, a MAES with constant  $\lambda$  and the  $\lambda$ -CMAES. The plot also shows the evolution of  $\lambda(g)$

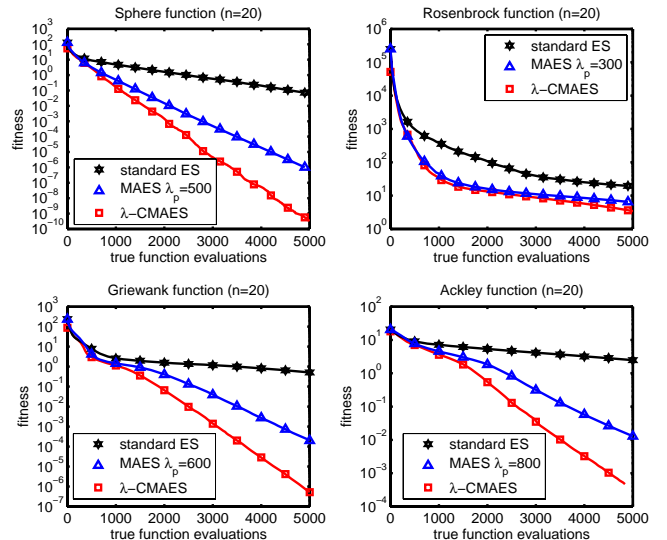


Fig. 5. Evolution of fitness in terms of number of fitness evaluations for a standard ES, a MAES with constant  $\lambda$  and the  $\lambda$ -CMAES.

fitness values that are several orders of magnitude better than a standard ES. Vice versa  $\lambda$ -CMAES achieves the same fitness value as an ES with less than one third of the computational cost. For the MAES the number of offspring  $\lambda_p$  is increased and set to optimal values  $\lambda_p \in [300 - 800]$  determined in an earlier independent analysis. The number of preselected individuals in MAES is kept constant at  $\lambda = 35$  for the unimodal and  $\lambda = 50$  for the multi-modal problems. Notice, that even though MAES requires more effort as the model estimates the fitness for more individuals its performance is clearly inferior to  $\lambda$ -CMAES.

Figure 6 shows that after a few generations  $\lambda$  fluctuates slightly around an average value. This raises the question if the control of  $\lambda$  is actually necessary or whether a MAES



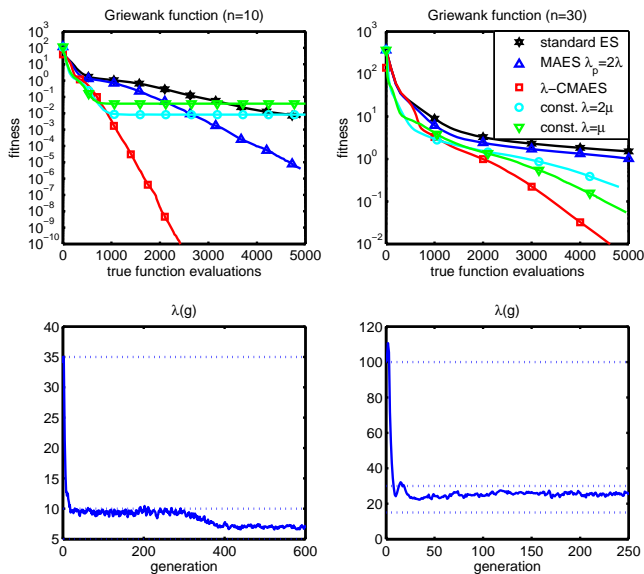


Fig. 6. Evolution of fitness in terms of number of fitness evaluations for a standard ES, three MAES with constant  $\lambda$  and the  $\lambda$ -CMAES for the Griewank function (upper graph). Number of preselected individuals (lower graph)

with a lower but constant value of  $\lambda$  does not achieve the same performance. It turns out that in general the optimal value of  $\lambda$  is problem specific and not necessarily constant throughout the evolution. The following results are based on the optimization of the Griewank-function for problem sizes  $n = 10$  and  $n = 30$ . For  $n = 10$  the parameters for the MAES and  $\lambda$ -CMAES are  $\mu = 5, \lambda = 35, \lambda_p = 70$  and  $\mu = 15, \lambda = 100, \lambda_p = 200$  for  $n = 30$ . The tests included two additional MAES with small constant values for  $\lambda = \mu$  and  $\lambda = 2\mu$  which correspond to the range to which  $\lambda$ -CMAES tends to converge. The results show that for the Griewank-function with dimension  $n = 10$  the value of  $\lambda$  further decreases after about 300 generations and is not constant throughout evolution. The MAES with small constant values of  $\lambda$  achieved similar performance as  $\lambda$ -CMAES in a large number of individual runs, but do not provide robust optimization. In particular for Griewank function with dimension  $n = 10$  the constant strategy often suffers from premature convergence of the MAES in local minima. Notice, that the density of local minima of the Griewank function increases for smaller dimensions.  $\lambda$ -CMAES proves robust with respect to local minima and the particular characteristics of the optimization problem.

## VI. CONCLUSION

This paper proposes a novel scheme for evolutionary optimization which attempts to exploit a limited number of actual fitness evaluations in the most effective manner. The model assisted evolution strategy reduces the number of actual fitness evaluations by preselecting offspring on the basis of a local fitness model. Our results show that weighted linear regression is most suitable for fitness function approximation due to its limited computational complexity, its ability to incorporate

new training data online, its capacity of extrapolating beyond the current distribution of data and finally its accuracy in the reliable preselection of the best individuals.

The novel model management scheme  $\lambda$ -CMAES permanently adapts the number of preselected individuals to the quality of the current model, thus using the fitness evaluations in the possible most efficient manner. The model quality is regulated such that it performs barely better than a random selection on the preselected individuals but nevertheless reliably discards inferior solutions in advance.  $\lambda$ -CMAES achieves a substantial reduction of fitness evaluations on a set of benchmarks not only in comparison to a standard evolution strategy but also with respect to other model assisted optimization schemes. Our analysis shows that the optimal rate of fitness evaluations is problem dependent and varies with the convergence of the population which confirms the necessity of adapting this parameter online.

## ACKNOWLEDGMENT

The authors would like to thank their colleague Daniel Schauten for his constructive comments and suggestions of this work.

## REFERENCES

- [1] Y. Jin, M. Olhofer, and B. Sendhoff, "A framework for evolutionary optimization with approximate fitness functions," *IEEE Transactions on Evolutionary Computation*, vol. 6, no. No. 5, pp. 481–494, October 2002.
- [2] Y. Jin, "A comprehensive survey of fitness approximation in evolutionary computation," *Soft Computing Journal*, vol. 9, no. 1, pp. 3–12, 2005.
- [3] Y. S. Ong, P. B. Nair, A. J. Keane, and K. W. Wong, "Surrogate-assisted evolutionary optimization frameworks for high-fidelity engineering design problems," in *Knowledge Incorporation in Evolutionary Computation*, ser. Studies in Fuzziness and Soft Computing, Y. Jin, Ed. Springer, 2004, pp. 307–332.
- [4] H. Ulmer, F. Streichert, and A. Zell, "Evolution strategies with controlled model assistance," in *Proceedings of the 2004 IEEE Congress on Evolutionary Computation*. Portland, Oregon: IEEE Press, 20–23 June 2004, pp. 1569–1576.
- [5] C. G. Atkeson, A. W. Moore, and S. Schaal, "Locally weighted learning," *Artificial Intelligence Review*, vol. 11, no. 1-5, pp. 11–73, 1997.
- [6] H. Ulmer, F. Streichert, and A. Zell, "Evolution strategies assisted by gaussian processes with improved pre-selection criterion," in *Proceedings of the 2003 Congress on Evolutionary Computation CEC2003, Canberra, Australia*. IEEE Press, 2003, pp. 692–699.
- [7] C. K. I. Williams, "Prediction with gaussian processes: From linear regression to linear prediction and beyond," in *Learning and Inference in Graphical Models*, M. I. Jordan, Ed. MIT Press, 1999, pp. 599–621.