Applying Exponential Weighting Moving Average Control Parameter Adaptation Technique with Generalized Differential Evolution

Saku Kukkonen
Centre of Computational Engineering
and Integrated Design (CEID)
Lappeenranta University of Technology
P.O. Box 20, FIN-53851 Lappeenranta, Finland
Email: dr.saku.kukkonen@gmail.com

Carlos A. Coello Coello
CINVESTAV-IPN
Computer Science Department
Av. IPN No. 2508, San Pedro Zacatenco
México D.F. 07300, México
Email: ccoello@cs.cinvestav.mx

***Abstract*—In this paper, an Exponential Weighting Moving Average (EWMA) control parameter adaptation technique is tested with Generalized Differential Evolution 3 (GDE3) using a set of multi-objective test problems and performance metrics. The results with and without EWMA control parameter adaptation are compared. EWMA has been earlier proposed with the original unconstrained single-objective Differential Evolution (DE), and EWMA adapts crossover and mutation control parameter values.**

From the results, it is observed that if good initial control parameter values are used, then there is not clear performance difference between GDE3 with and without EWMA. However, if the initial control parameter values are not good, then EWMA gives clear improvement in performance.

Based on the results, the same control parameter adaptation technique can be used both in the case of single- and multi-objective optimization with GDE3. However, different initial control parameter values should be used in different cases, and recommendations for initial values are given at the end of the paper.

I. Introduction

Generalized Differential Evolution 3 (GDE3) is a general purpose Evolutionary Algorithm (EA) for problems having different number of constraints and objectives. GDE3 is an extension of Differential Evolution (DE) that is meant for unconstrained single-objective optimization. GDE3 has been applied for several real optimization problems and in scientific comparative studies [1]. As many other EAs, GDE3 has control parameters, whose values have to be set before optimization. Setting these parameter values is not an easy task even for an experienced user, let alone for a non-experienced practitioner. Therefore, there is a need to automate the selection of control parameter values.

In this paper, one recent proposal for control parameter adaptation with DE has been implemented and tested with GDE3. The proposal is namely Exponential Weighting Moving Average (EWMA) control parameter adaptation technique that has provided good results with single-objective DE [2]. Here, GDE3 with and without EWMA control parameter adaptation is tested. Comparison is done experimentally using multi-objective problems and performance metrics defined for the CEC 2007 Special Session on Performance Assessment of Multi-Objective Optimization Algorithms [3]. The problems have two, three, or five objectives, and the number of decision variables varies from 3 to 30. Difficulty of the problems varies in terms of separability, modality, and geometry of the Pareto-front.

The remainder of this paper is organized as follows: Multi-objective optimization with constraints is briefly defined in Section II. Section III describes the multi-objective optimization method and control parameter adaptation technique used. Section IV describes experiments and results. Finally, our conclusions, discussion, and some possible paths for future research are provided in Section V.

II. Multi-Objective Optimization with Constraints

A multi-objective optimization problem (MOOP) with constraints can be presented in the form [4, p. 37]:

$$\begin{array}{ll} \text{minimize} & \{f_1(\vec{x}), f_2(\vec{x}), \ldots, f_M(\vec{x})\} \\ \text{subject to} & g_1(\vec{x}) \leq 0 \\ & g_2(\vec{x}) \leq 0 \\ & \quad\vdots \\ & g_K(\vec{x}) \leq 0 \end{array}$$

Thus, there are M functions to be optimized and K inequality constraints. Maximization problems can be converted to minimization problems, and all the constraints can be converted into the form $g_k(\vec{x}) \leq 0$. Thereby the formulation above is without loss of generality.

The objective of Pareto-optimization is to find an approximation of the Pareto-front, *i.e.*, to find a set of solutions that are not dominated by any other solution. Weak dominance relation $\preceq$ between two vectors is defined in such a way that $\vec{x}$ weakly dominates $\vec{y}$, *i.e.*, $\vec{x} \preceq \vec{y}$ iff $\forall i : f_i(\vec{x}) \leq f_i(\vec{y})$. Dominance relation $\prec$ between two vectors is defined such a way that $\vec{x}$ dominates $\vec{y}$, *i.e.*, $\vec{x} \prec \vec{y}$ iff $\vec{x} \preceq \vec{y} \ \wedge \ \exists i : f_i(\vec{x}) < f_i(\vec{y})$. The dominance relationship can be extended to take into consideration constraint values and objective values at the same time. Constraint-domination $\prec_c$ is defined in this paper so that $\vec{x}$ constraint-dominates $\vec{y}$, *i.e.*, $\vec{x} \prec_c \vec{y}$ iff any of the following conditions is true [1]:

- $\vec{x}$ and $\vec{y}$ are feasible and $\vec{x}$ dominates $\vec{y}$ in objective function space.
- $\vec{x}$ is feasible and $\vec{y}$ is not.
- $\vec{x}$ and $\vec{y}$ are infeasible and $\vec{x}$ dominates $\vec{y}$ in constraint function violation space.

The definition for weak constraint-domination $\preceq_c$ is analogous by the dominance relation changed to weak dominance in the above definition. Constraint-domination is a special case of a more general concept including goals and priorities that is presented in [5].

III. Methods

A. Differential Evolution

The Differential Evolution (DE) algorithm [6], [7] was introduced by Storn and Price in 1995. The design principles of DE are simplicity, efficiency, and the use of floating-point encoding instead of binary numbers. As a typical EA, DE has a random initial population that is then improved using selection, mutation, and crossover operations. Several ways exist to determine a stopping criterion for EAs but usually a predefined upper limit (G_{max}) for the number of generations to be computed provides an appropriate stopping condition. Other control parameters for DE are the crossover control parameter (CR), the mutation factor (F), and the population size (NP).

At each generation G, DE goes through each D dimensional decision vector $\vec{x}_{i,G}$ of the population and creates the corresponding trial vector $\vec{u}_{i,G}$ as follows [8]:

$$
\begin{array}{l}
r_1, r_2, r_3 \in \{1, 2, \ldots, NP\}\,, \text{(randomly selected,} \\
\quad \text{except mutually different and different from } i) \\
j_{rand} = \text{floor}\,(rand_i[0,1) \cdot D) + 1 \\
\text{for}(j = 1; j \leq D; j = j + 1) \\
\{ \\
\qquad \text{if}(rand_j[0,1) < CR \vee j = j_{rand}) \\
\qquad\qquad u_{j,i,G} = x_{j,r_3,G} + F \cdot (x_{j,r_1,G} - x_{j,r_2,G}) \\
\qquad \text{else} \\
\qquad\qquad u_{j,i,G} = x_{j,i,G} \\
\}
\end{array}
$$

This is the most common DE version, DE/rand/1/bin, also known as the classic DE. Functions $rand_i[0,1)$ and $rand_j[0,1)$ return a random number drawn from the uniform distribution between 0 and 1 for each i and j. Both CR and F remain fixed during the entire execution of the algorithm. Parameter $CR \in [0,1]$, which controls the crossover operation, represents the probability that an element for the trial vector is chosen from a linear combination of three randomly chosen vectors and not from the old vector $\vec{x}_{i,G}$. The condition "$j = j_{rand}$" ensures that at least one element of the trial vector is different compared to the elements of the old vector. Parameter F is a scaling factor for mutation and its value range is $(0, 1+]$ (*i.e.* larger than 0 and upper limit is around 1 although there is no hard upper limit). In practice, CR controls rotational invariance of the search, and a small value for it (*e.g.*, 0.1) is useful with separable problems while larger values (*e.g.*, 0.9) are useful for non-separable problems. Parameter F controls the speed and robustness of the search, *i.e.*, a lower value for F increases the convergence rate but it also increases the risk of getting stuck into a local optimum. Parameters CR and NP have the similar effect on the convergence rate as F has. [9]

After the mutation and crossover operations, the trial vector $\vec{u}_{i,G}$ is compared to the old vector $\vec{x}_{i,G}$. If the trial vector has an equal or better objective value, then it replaces the old vector in the next generation. This can be presented as follows in the case of minimization of an objective [8]:

$$
\vec{x}_{i,G+1} = \begin{cases} \vec{u}_{i,G} & \text{if} \quad f(\vec{u}_{i,G}) \leq f(\vec{x}_{i,G}) \\ \vec{x}_{i,G} & \text{otherwise} \end{cases}
$$

DE is an elitist method since the best population member is always preserved and the average objective value of the population will never deteriorate.

B. Generalized Differential Evolution

The first version of Generalized Differential Evolution (GDE) extended DE for constrained multi-objective optimization, and it modified only the selection rule of the basic DE [10]. The basic idea in the selection rule of GDE is that the trial vector is selected to replace the old vector in the next generation if it weakly constraint-dominates the old vector. There was no explicit sorting of non-dominated solutions [11, pp. 33 – 44] during the optimization process or any mechanism for maintaining the distribution and extent of solutions. Also, there was no extra repository for non-dominated solutions.

The second version, GDE2, made the selection based on crowdedness when the trial and old vector were feasible and non-dominating with respect to each other in the objective function space [12]. This improved the extent and distribution of the obtained set of solutions but slowed down the convergence of the overall population because it favored isolated solutions far from the Pareto-front until all the solutions had converged near the Pareto-front.

The third version is GDE3 [13], [14]. Besides the selection, another part of the basic DE has also been modified. Now, in the case of feasible and non-dominating solutions, both solutions are saved for the population of next generation. Before continuing to the next generation, the size of the population is reduced using non-dominated sorting and pruning based on diversity. The pruning technique used in the original GDE3 is based on crowding distance, which provides a good crowding estimation in the case of two objectives. However, crowding distance fails to approximate crowdedness of solutions when the number of objectives is more than two [14]. Since, the problem set adopted in [3] consists of problems with more than two objectives, a more general diversity maintenance technique proposed in [15] is used. The technique is based on a crowding estimation using the nearest neighbors of solutions in an Euclidean sense, and an efficient nearest neighbors search technique.

All the GDE versions can handle different number of M objectives and number of K constraints, including the cases where $M = 0$ (constraint satisfaction problem) and $K = 0$ (unconstrained problem). When $M = 1$ and $K = 0$, the versions are identical to the original DE, and this is why they are referred to as *Generalized* DEs.

C. Exponential Weighting Moving Average Control Parameter Adaptation Technique

Exponential Weighting Moving Average (EWMA) control parameter adaptation technique has been recently proposed for single-objective DE. The technique was first proposed for adaptation of CR and F separately, and then the technique was proposed to be used for adapting both CR and F simultaneously [2]. According to the results presented in [2], EWMA applied to adapt both CR and F simultaneously outperforms the classic DE and separate control parameter adaptation approaches. While EWMA adapts CR and F, population size NP and number of generations G_{max} are kept fixed.

EWMA performs control parameter adaptation by updating weighted average history of successful control parameter values. Two new variables, $EWMA_{CR}$ and $EWMA_F$, are used to store successful control parameter history of CR and F, respectively. EWMA values start from the same values as CR and F. EWMA values are then later on updated according to successful CR and F values meaning that their values are updated whenever new trial solution replaces the corresponding target solution. CR and F values are generated anew at the beginning of each generation using current EWMA values added with a random value. Formally, at beginning of a generation, new CR and F values are created as follows:

$$\begin{aligned} CR &= EWMA_{CR} + rand[-x, x] \\ F &= EWMA_F + rand[-x, x], \end{aligned}$$

where $rand[-x, x]$ is a uniformly distributed random variable in the range $[-x, x]$. After creation, new values are checked and changed if they exceed given lower and upper bounds. Also, a c value describing the effect of crossover and mutation operations to the standard deviation of the population is calculated and a check is enforced so that defined lower and upper bounds are not violated. The c value is calculated as [16]:

$$c = \sqrt{2F^2CR - 2CR/NP + CR^2/NP + 1}$$

Checking and updating control parameter values is done as follows:

if $F < F_{min} \vee F > F_{max}$,
then $F = EWMA_F$
if $CR < CR_{min} \vee CR > CR_{max} \vee c < c_{min} \vee c > c_{max}$,
then $CR = EWMA_{CR}$

Checking and updating is done in this order so the c value is calculated after checking and possibly correcting the F value.

EWMA values are updated based on control parameter values that were used to create the successful trial solution. The update is done in the following way:

$$\begin{aligned} EWMA_{CR} &= \alpha \cdot CR + (1 - \alpha) \cdot EWMA_{CR} \\ EWMA_F &= \alpha \cdot F + (1 - \alpha) \cdot EWMA_F \end{aligned}$$

Thus, the idea is to update EWMA values to the direction of successful CR and F values remembering also the history of previous EWMA values.

The EWMA adaptation technique involves several new parameters, but these are going to have chosen fixed values that are not to be changed by the user.

Something nice about the presented adaptation technique is that it is simple and only modifies the control parameter part of the algorithm. In GDE3, EWMA values are updated whenever the trial vector weakly constraint-dominates the target vector. This way the function is identical to [2] in the case of single-objective optimization and therefore also the results in [2] are directly applicable.

IV. Experiments

A. Configuration

GDE3 with and without the EWMA control parameter adaptation was implemented in the ANSI-C programming language. With both approaches, the same initial control parameter values were used. Tests with each test problem were repeated 100 times and the same random number generator seeds were used for both approaches in order to have the same initial random populations. Two different approaches for initial control parameter values were used: in one approach the fixed initial control parameter values were used through repetitions and in another approach different random values (but the same for both approaches) were used through repetitions. For fixed control parameter values three different setups were used: $CR = 0.1$ & $F = 0.5$; $CR = 0.2$ & $F = 0.2$; and $CR = 0.9$ & $F = 0.9$. The first setup ($CR = 0.1$ & $F = 0.5$) is the same as used in [17] and this setup was found with a preliminary test to be quite optimal for the given set of test problems. The second setup ($CR = 0.2$ & $F = 0.2$) is found to be suitable in general in the case of multi-objective problems [1]. The third setup ($CR = 0.9$ & $F = 0.9$) is the one often recommended for single-objective optimization and also used in [2].

For the population size and number of generations, fixed values $NP = 100$, and $G_{max} = 4999$ were used. These values were the same as used in [17]. The size of the population was set according to the smallest desired approximation set size given in [3]. With chosen NP and G_{max} values, the number of function evaluations (FES) is exactly 500 000, which was an upper limit given in [3]. The performance was measured after 5 000, 50 000, and 500 000 FES.

In [3], approximation sets of different sizes were demanded for different problems. In GDE3, the size of the approximation set is usually the same as NP. Now, NP was kept fixed for all the problems, and solutions for the approximation set were collected during generations. Populations of the 200 last generations (49 last generations in the case of 5 000 FES) were merged together and non-dominated solutions were selected from this merged set of solutions. If the size of the non-dominated set was larger than the desired approximation set size, then the set was reduced to the desired size using the pruning technique described in [15].

The performance was measured using the Hypervolume and the R indicator as in [3], [17]. These indicators measure overall quality of the obtained sets of non-dominated solutions. Both indicators were used in unary form with the help of provided reference sets, and for both indicators, smaller values mean better results (if the value is less than 0, then the obtained solution is better than the corresponding reference set).

The variable values used in the EWMA technique were the same as used in [2], and these values were $x = 0.1$, $\alpha = 0.1$, $F_{min} = 0.2$, $F_{max} = 1.0$, $CR_{min} = 0.0$, $CR_{max} = 1.0$, $c_{min} = 1.0$, and $c_{max} = 1.5$. The values were deliberately kept same as in [2] to have identical control value adaptation in both single- and multi-objective cases without problem specific tuning. It could be possible to have some better values for a given set of problems but then there would be danger of loosing generality.

Each test problem given in [3] has a predefined range of variables, thus there are boundary constraints. In the case of boundary constraint violations, violating variable values were reflected back from the violated boundary by the amount of violation as in [1], [9], [17]

B. Results of Experiments

The problems given in [3] were solved 100 times and achieved results are presented in Tables I–XII. The tables show the mean values of indicators. The results in Tables I–VI correspond to the case when $CR = 0.1$ and $F = 0.5$. The results in Tables VII–XII correspond to the case when random values $CR \in [0.0, 1.0]$ and $F \in [0.2, 1.0]$ were used as initial values for the approaches. The tables contain also statistical significance comparison between approaches. The value of Wilcoxon rank-sum test [1] is shown under mean indicator values. If one approach has obtained significantly better results according to Wilcoxon rank-sum test (with 1% significance level) than the other one, then the better value is printed with boldface.

From Tables I–VI it can be seen that when good control parameter values are used for CR and F, there is no clear evidence if EWMA provides clear improvement to the results. It appears that with a small number of FES, fixed parameter values provide slightly better results, but with a large number of FES, EWMA provides slightly better results. Anyway, the results are contradicting and a clear judgment cannot be made.

From Tables VII–XII it can be seen that when control parameter values for CR and F are randomly selected, EWMA provides a clear improvement to the results. Improvement is especially clear for larger numbers of FES.

[1]Wilcoxon rank-sum test was chosen since the test does not expect the random variables to follow a normal distribution. According to the histograms of the indicator values, the performance metric values are not normally distributed.

Numerical results for the cases $CR = 0.2$ & $F = 0.2$ and $CR = 0.9$ & $F = 0.9$ are not shown here but in the first case, the results were similar compared to the shown case of good control parameter values. In the second case, EWMA was even more clearly better than in the case of randomly selected control parameter values.

As anticipated, better initial control parameter values lead to better performance according to the performance metrics. However, with EWMA the performance difference between good and bad control parameter values is very small and in some few cases EWMA with bad initial control parameter values resulted better performance metric values than with good initial control parameter values. One should note that EWMA adjusts control parameter values only based on dominance relation between the trial and old solution, but the performance metrics here measure overall performance.

Figure 1 shows history of $EWMA_{CR}$ and $EWMA_F$ values as box plots during generations when random CR and F values were used with OKA2. It can be seen that $EWMA_{CR}$ and $EWMA_F$ did not converge to some certain values but rather to some value ranges. This same was observed with the other problems. Value ranges for $EWMA_{CR}$ were smaller and having smaller values than for $EWMA_F$.

V. Conclusions and Discussion

Generalized Differential Evolution 3 (GDE3) with and without Exponential Weighting Moving Average (EWMA) control parameter adaptation has been experimentally evaluated using the problems and metrics defined for the CEC 2007 Special Session on Performance Assessment of Multi-Objective Optimization Algorithms. EWMA adapts values of crossover control parameter CR and mutation parameter F while population size and stopping condition are kept constant. EWMA has been earlier proposed for single-objective DE, but the technique can be adapted to GDE3 in such a way that methods are identical in the case of a single-objective problem. Since EWMA has already been evaluated in the case of single-objective optimization, EWMA was evaluated only for multi-objective optimization in this paper.

Tests were performed using fixed and randomly selected initial control parameter values. According to the numerical results, when initial control parameter values are good, it is not clear if GDE3 is better with or without EWMA control parameter adaptation. It can be said that the results are comparable with both ways. When initial control parameter values are not good, then GDE3 with EWMA provides clear improvement over GDE3 without EWMA.

Based on the results of this paper and results of single-objective DE with EWMA published in [2], it is advisable to use EWMA with GDE3 in the case of single- and multi-objective problems. However, different initial control parameter values should be used based on the number of objectives. In single-objective optimization, suitable control parameter

TABLE I

THE RESULTS FOR R INDICATOR (I_{R2}) ON TEST FUNCTIONS 1–7 WHEN $CR = 0.1$ & $F = 0.5$

FES		OKA2	SYMPART	S_ZDT1	S_ZDT2	S_ZDT4	R_ZDT4	S_ZDT6
5e3	GDE3	-7.4741e-05	**2.6338e-02**	4.5234e-02	**9.7147e-02**	**7.6181e-02**	**1.7341e-02**	**1.3424e-01**
	GDE3+EWMA	-3.6186e-05	3.1989e-02	4.8330e-02	1.0524e-01	8.3744e-02	2.0254e-02	1.3688e-01
	Wilcoxon	6.5742e-01	1.0859e-05	6.6689e-02	9.4194e-08	1.2387e-08	3.6312e-06	2.9686e-06
5e4	GDE3	-9.8804e-04	2.2664e-05	**1.1541e-04**	**1.3115e-03**	6.8281e-03	9.4912e-04	**2.0003e-02**
	GDE3+EWMA	**-1.0109e-03**	2.8472e-05	1.9509e-04	8.4648e-03	7.0592e-03	1.0807e-03	2.6633e-02
	Wilcoxon	4.4006e-04	3.3143e-01	1.1033e-04	3.7786e-14	9.1148e-01	3.1586e-01	2.0224e-06
5e5	GDE3	-1.0585e-03	1.5688e-06	6.2936e-06	**8.0821e-04**	2.2293e-05	7.2145e-04	-1.0788e-06
	GDE3+EWMA	-1.0595e-03	**1.4489e-06**	**2.4906e-06**	8.0125e-03	4.8615e-05	6.2342e-04	-9.1228e-07
	Wilcoxon	1.2048e-01	3.2081e-04	1.7887e-10	9.8530e-04	2.0341e-02	2.7102e-02	3.2217e-01

TABLE II

THE RESULTS FOR R INDICATOR (I_{R2}) ON TEST FUNCTIONS 8–13 WITH M = 3 AND $CR = 0.1$ & $F = 0.5$

FES		S_DTLZ2	R_DTLZ2	S_DTLZ3	WFG1	WFG8	WFG9
5e3	GDE3	2.2619e-04	4.2988e-04	4.7575e-04	8.0696e-02	-1.0764e-02	-4.0342e-03
	GDE3+EWMA	2.2195e-04	4.3675e-04	4.8801e-04	8.0386e-02	-1.0166e-02	-4.2093e-03
	Wilcoxon	2.2321e-01	3.4498e-01	4.3213e-01	1.4976e-01	5.5882e-02	9.6200e-01
5e4	GDE3	4.9709e-05	1.2210e-04	1.7305e-05	**5.5880e-02**	**-2.5996e-02**	-9.0437e-03
	GDE3+EWMA	**3.8010e-05**	**9.1176e-05**	1.6893e-05	6.1670e-02	-2.5526e-02	-9.0061e-03
	Wilcoxon	3.8066e-06	2.9575e-25	1.1032e-01	6.0236e-16	1.8116e-05	8.8634e-01
5e5	GDE3	6.6288e-06	2.2875e-05	**4.3575e-07**	**2.3419e-03**	-2.8498e-02	-9.3978e-03
	GDE3+EWMA	5.2820e-06	2.3095e-05	9.2820e-07	8.8549e-03	-2.8483e-02	-9.5445e-03
	Wilcoxon	1.7492e-02	8.9020e-01	3.4772e-10	3.9697e-30	6.8773e-01	8.4302e-02

TABLE III

THE RESULTS FOR R INDICATOR (I_{R2}) ON TEST FUNCTIONS 8–13 WITH M = 5 AND $CR = 0.1$ & $F = 0.5$

FES		S_DTLZ2	R_DTLZ2	S_DTLZ3	WFG1	WFG8	WFG9
5e3	GDE3	1.8990e-04	1.8701e-04	2.2249e-04	5.6353e-02	5.7327e-03	6.8174e-03
	GDE3+EWMA	2.0156e-04	1.8681e-04	2.2716e-04	5.6515e-02	6.0145e-03	6.4938e-03
	Wilcoxon	2.2978e-01	8.7285e-01	3.2779e-01	3.1061e-02	3.9584e-01	3.2418e-01
5e4	GDE3	3.6094e-05	5.6739e-05	1.2955e-05	**4.5486e-02**	-7.2376e-03	2.5055e-03
	GDE3+EWMA	**2.6311e-05**	**5.2087e-05**	1.3472e-05	4.7855e-02	-7.4687e-03	2.4960e-03
	Wilcoxon	3.1193e-15	3.0320e-04	6.3462e-01	5.7702e-17	1.3640e-02	9.3865e-01
5e5	GDE3	2.0631e-05	3.4022e-05	2.0241e-07	**4.5620e-03**	-1.1490e-02	1.9538e-03
	GDE3+EWMA	2.0061e-05	**2.2859e-05**	1.9460e-07	1.1460e-02	**-1.1627e-02**	**1.5916e-03**
	Wilcoxon	3.7708e-01	4.6801e-19	8.1549e-01	8.6930e-30	3.7760e-05	9.3931e-05

TABLE IV

THE RESULTS FOR HYPERVOLUME INDICATOR $I_{\bar{H}}$ ON TEST FUNCTIONS 1–7 WHEN $CR = 0.1$ & $F = 0.5$

FES		OKA2	SYMPART	S_ZDT1	S_ZDT2	S_ZDT4	R_ZDT4	S_ZDT6
5e3	GDE3	4.3175e-04	**7.4974e-02**	1.5673e-01	**2.3549e-01**	**2.3089e-01**	**5.3839e-02**	**3.3980e-01**
	GDE3+EWMA	3.1182e-04	9.0588e-02	1.6824e-01	2.5942e-01	2.5543e-01	6.2261e-02	3.4660e-01
	Wilcoxon	3.6403e-01	1.0617e-05	1.2391e-02	8.1400e-07	3.5911e-09	3.6744e-06	5.1608e-06
5e4	GDE3	-1.1451e-03	6.7904e-05	**5.7510e-04**	**1.9345e-03**	2.0401e-02	3.2832e-03	**4.5048e-02**
	GDE3+EWMA	**-1.1702e-03**	8.5014e-05	9.0484e-04	1.0777e-02	2.1051e-02	3.6692e-03	6.0730e-02
	Wilcoxon	1.4147e-04	3.4748e-01	4.6817e-10	1.3997e-16	8.9213e-01	2.9736e-01	1.9269e-06
5e5	GDE3	-1.2265e-03	4.6798e-06	1.8072e-04	1.1643e-03	**5.2785e-05**	2.4401e-03	**-2.3044e-04**
	GDE3+EWMA	-1.2266e-03	**4.3086e-06**	1.8237e-04	9.7306e-03	1.1616e-04	**1.8021e-03**	-2.2790e-04
	Wilcoxon	9.7369e-01	2.7585e-04	2.2845e-02	1.8003e-02	1.7636e-13	2.7398e-05	4.2663e-34

TABLE V

THE RESULTS FOR HYPERVOLUME INDICATOR $I_{\bar{H}}$ ON TEST FUNCTIONS 8–13 WITH M = 3 AND $CR = 0.1$ & $F = 0.5$

FES		S_DTLZ2	R_DTLZ2	S_DTLZ3	WFG1	WFG8	WFG9
	GDE3	1.8887e-03	1.0813e-02	**7.0896e-03**	4.0785e-01	-7.3777e-02	-2.4175e-02
5e3	GDE3+EWMA	2.2946e-03	1.1268e-02	9.5670e-03	4.0616e-01	-7.0913e-02	-2.5211e-02
	Wilcoxon	3.1586e-01	3.2900e-01	6.8950e-12	1.0550e-01	8.9709e-02	8.9986e-01
	GDE3	1.1015e-04	5.2459e-04	2.6520e-06	**2.9004e-01**	**-1.6083e-01**	-5.7386e-02
5e4	GDE3+EWMA	**8.7776e-05**	**2.6033e-04**	2.9440e-06	3.1623e-01	-1.5834e-01	-5.7249e-02
	Wilcoxon	4.7527e-03	7.3976e-28	6.0020e-01	6.5339e-15	9.4548e-07	6.0873e-01
	GDE3	-3.4539e-06	9.1015e-06	**2.6768e-09**	**8.9850e-03**	**-1.7450e-01**	-6.0593e-02
5e5	GDE3+EWMA	-7.4541e-06	1.0390e-05	8.5377e-09	4.6398e-02	-1.7407e-01	-6.1344e-02
	Wilcoxon	2.9935e-02	1.4702e-01	2.6243e-08	1.3600e-30	2.3787e-03	1.0558e-02

TABLE VI

THE RESULTS FOR HYPERVOLUME INDICATOR $I_{\bar{H}}$ ON TEST FUNCTIONS 8–13 WITH M = 5 AND $CR = 0.1$ & $F = 0.5$

FES		S_DTLZ2	R_DTLZ2	S_DTLZ3	WFG1	WFG8	WFG9
	GDE3	1.4577e-03	3.2829e-03	**2.4495e-03**	6.3377e-01	-4.3075e-02	5.3230e-02
5e3	GDE3+EWMA	1.6588e-03	3.0503e-03	3.1149e-03	6.3537e-01	-3.6512e-02	4.8610e-02
	Wilcoxon	3.1350e-01	5.4640e-02	1.7701e-06	4.5510e-02	2.4349e-02	4.2218e-01
	GDE3	4.5072e-05	3.9071e-05	7.6578e-06	**5.2194e-01**	-2.6798e-01	-9.8845e-02
5e4	GDE3+EWMA	**2.2997e-05**	**-2.5670e-05**	9.0885e-06	5.4658e-01	-2.6485e-01	-9.7592e-02
	Wilcoxon	2.8480e-14	7.4803e-15	5.6006e-01	5.2016e-17	3.7314e-01	3.7183e-01
	GDE3	8.6042e-06	-1.2978e-04	1.4090e-09	**5.2020e-02**	-3.4919e-01	-1.1087e-01
5e5	GDE3+EWMA	7.8945e-06	**-1.6136e-04**	1.1693e-09	1.4329e-01	-3.4915e-01	**-1.1721e-01**
	Wilcoxon	3.3143e-01	3.5811e-29	3.1116e-01	1.4392e-30	9.0954e-01	9.6811e-10

TABLE VII

THE RESULTS FOR R INDICATOR (I_{R2}) ON TEST FUNCTIONS 1–7 WHEN $CR \in [0.0, 1.0]$ & $F \in [0.2, 1.0]$

FES		OKA2	SYMPART	S_ZDT1	S_ZDT2	S_ZDT4	R_ZDT4	S_ZDT6
	GDE3	9.0990e-05	3.5094e-02	6.2706e-02	1.2094e-01	9.6459e-02	3.2893e-02	1.4265e-01
5e3	GDE3+EWMA	-4.3284e-05	3.3678e-02	5.9992e-02	1.1947e-01	9.4186e-02	3.1326e-02	1.4187e-01
	Wilcoxon	5.5023e-01	5.5023e-01	2.0000e-01	4.5539e-01	5.6501e-01	2.6468e-01	2.8839e-01
	GDE3	-1.0158e-03	3.1885e-03	1.6729e-02	5.4585e-02	5.4063e-02	1.1513e-02	7.9917e-02
5e4	GDE3+EWMA	-1.0122e-03	**2.6746e-05**	**1.7832e-03**	**2.4310e-02**	**8.8899e-03**	**8.1009e-03**	**3.3218e-02**
	Wilcoxon	4.3644e-01	1.9589e-07	1.5245e-10	1.2564e-05	6.3391e-23	1.8976e-04	3.5383e-09
	GDE3	-1.0609e-03	7.2397e-05	2.7517e-03	3.0470e-02	2.4523e-02	3.1585e-03	2.4354e-02
5e5	GDE3+EWMA	-1.0599e-03	**1.5470e-06**	**2.4239e-06**	**2.2831e-02**	**5.2030e-05**	**1.4516e-03**	**-9.1228e-07**
	Wilcoxon	2.3139e-02	5.2712e-22	2.4079e-18	3.8528e-06	3.0608e-27	9.8076e-08	1.6828e-17

TABLE VIII

THE RESULTS FOR R INDICATOR (I_{R2}) ON TEST FUNCTIONS 8–13 WITH M = 3 AND $CR \in [0.0, 1.0]$ & $F \in [0.2, 1.0]$

FES		S_DTLZ2	R_DTLZ2	S_DTLZ3	WFG1	WFG8	WFG9
	GDE3	3.4579e-04	5.8574e-04	5.2718e-04	8.1095e-02	-6.2897e-03	1.8273e-03
5e3	GDE3+EWMA	2.7527e-04	5.8557e-04	4.8553e-04	8.1006e-02	-7.1069e-03	1.1363e-03
	Wilcoxon	4.2684e-02	9.2699e-01	1.7725e-02	5.2445e-01	1.8177e-01	2.5948e-01
	GDE3	6.6249e-05	3.2790e-04	2.2070e-04	6.8865e-02	-1.7439e-02	-7.2564e-03
5e4	GDE3+EWMA	**4.1566e-05**	2.7058e-04	**4.1792e-05**	**6.4482e-02**	**-2.3024e-02**	-8.4256e-03
	Wilcoxon	1.0924e-03	1.6367e-02	3.2137e-19	3.5667e-08	4.2630e-10	1.0050e-02
	GDE3	1.7416e-05	2.3852e-04	1.3866e-04	4.6443e-02	-2.1816e-02	-9.7485e-03
5e5	GDE3+EWMA	**6.0494e-06**	**1.5643e-04**	**7.6268e-07**	**9.7516e-03**	**-2.8486e-02**	**-1.1727e-02**
	Wilcoxon	1.4799e-11	4.0237e-05	6.2159e-30	4.3803e-19	3.5986e-22	1.2706e-05

TABLE IX

THE RESULTS FOR R INDICATOR (I_{R2}) ON TEST FUNCTIONS 8–13 WITH M = 5 AND $CR \in [0.0, 1.0]$ & $F \in [0.2, 1.0]$

FES		S_DTLZ2	R_DTLZ2	S_DTLZ3	WFG1	WFG8	WFG9
5e3	GDE3	2.6985e-04	2.1436e-04	2.5445e-04	5.7103e-02	9.1420e-03	1.0199e-02
	GDE3+EWMA	2.4499e-04	2.1287e-04	2.4646e-04	5.7063e-02	8.5124e-03	1.0271e-02
	Wilcoxon	1.8825e-01	8.6708e-01	3.5126e-01	8.1929e-01	1.1588e-01	8.3834e-01
5e4	GDE3	6.9061e-05	1.1395e-04	7.8860e-05	5.1477e-02	5.8853e-04	4.6295e-03
	GDE3+EWMA	**3.9147e-05**	**9.7612e-05**	**2.0456e-05**	**4.9750e-02**	**-3.6495e-03**	4.2931e-03
	Wilcoxon	1.6743e-09	8.4089e-03	1.9817e-18	2.6034e-06	7.0909e-07	2.0519e-01
5e5	GDE3	4.8841e-05	9.9389e-05	6.1632e-05	4.1544e-02	-3.9953e-03	3.1471e-03
	GDE3+EWMA	**2.1112e-05**	**4.1464e-05**	**1.9587e-07**	**1.5381e-02**	**-1.0855e-02**	**2.4592e-03**
	Wilcoxon	4.6451e-03	8.1327e-16	2.8330e-30	2.1069e-20	3.3280e-18	2.1106e-04

TABLE X

THE RESULTS FOR HYPERVOLUME INDICATOR $I_{\bar{H}}$ ON TEST FUNCTIONS 1–7 WHEN $CR \in [0.0, 1.0]$ & $F \in [0.2, 1.0]$

FES		OKA2	SYMPART	S_ZDT1	S_ZDT2	S_ZDT4	R_ZDT4	S_ZDT6
5e3	GDE3	2.5323e-04	9.8927e-02	2.2127e-01	3.1221e-01	2.9633e-01	1.0027e-01	3.6282e-01
	GDE3+EWMA	1.3165e-04	9.4993e-02	2.0911e-01	3.0623e-01	2.8917e-01	9.6022e-02	3.6028e-01
	Wilcoxon	5.8500e-01	5.5187e-01	1.5967e-01	4.8544e-01	5.9851e-01	3.3754e-01	2.2789e-01
5e4	GDE3	-1.1727e-03	9.1281e-03	6.0981e-02	1.1954e-01	1.6328e-01	3.4939e-02	1.9707e-01
	GDE3+EWMA	-1.1706e-03	**7.9881e-05**	**6.8085e-03**	**3.1003e-02**	**2.6436e-02**	**2.4691e-02**	**7.6408e-02**
	Wilcoxon	7.8904e-01	2.7522e-07	6.5885e-10	4.3318e-06	7.1585e-23	2.0111e-04	6.0097e-09
5e5	GDE3	-1.2244e-03	2.0987e-04	1.0373e-02	4.4906e-02	7.3195e-02	9.7340e-03	5.9312e-02
	GDE3+EWMA	**-1.2272e-03**	**4.5890e-06**	**1.8245e-04**	**2.7344e-02**	**1.2369e-04**	**4.3148e-03**	**-2.2767e-04**
	Wilcoxon	1.0553e-03	8.2778e-22	4.7871e-13	6.8117e-05	5.8259e-26	7.5873e-08	5.8632e-06

TABLE XI

THE RESULTS FOR HYPERVOLUME INDICATOR $I_{\bar{H}}$ ON TEST FUNCTIONS 8–13 WITH M = 3 AND $CR \in [0.0, 1.0]$ & $F \in [0.2, 1.0]$

FES		S_DTLZ2	R_DTLZ2	S_DTLZ3	WFG1	WFG8	WFG9
5e3	GDE3	5.8959e-03	2.7836e-02	2.0730e-02	4.0926e-01	-4.9622e-02	1.2810e-02
	GDE3+EWMA	4.6389e-03	2.7487e-02	1.8234e-02	4.0882e-01	-5.3346e-02	8.4944e-03
	Wilcoxon	1.2582e-01	7.0036e-01	1.3450e-01	5.4211e-01	2.0695e-01	2.8508e-01
5e4	GDE3	9.3504e-04	9.4541e-03	7.4475e-03	3.4973e-01	-1.1333e-01	-4.4725e-02
	GDE3+EWMA	**2.9793e-04**	7.7124e-03	**8.3533e-04**	**3.2935e-01**	**-1.4400e-01**	**-5.1400e-02**
	Wilcoxon	6.6871e-11	1.5828e-02	1.0102e-20	3.1035e-08	3.0183e-10	5.5292e-03
5e5	GDE3	1.3654e-04	4.0382e-03	3.5082e-03	2.3789e-01	-1.3454e-01	-5.8255e-02
	GDE3+EWMA	**-2.5727e-06**	**2.9037e-03**	**3.3828e-09**	**5.1084e-02**	**-1.7417e-01**	**-7.1978e-02**
	Wilcoxon	2.7249e-17	2.6241e-05	2.9971e-30	3.8364e-19	2.5041e-23	2.4811e-08

TABLE XII

THE RESULTS FOR HYPERVOLUME INDICATOR $I_{\bar{H}}$ ON TEST FUNCTIONS 8–13 WITH M = 5 AND $CR \in [0.0, 1.0]$ & $F \in [0.2, 1.0]$

FES		S_DTLZ2	R_DTLZ2	S_DTLZ3	WFG1	WFG8	WFG9
5e3	GDE3	3.9840e-03	5.6178e-03	6.8908e-03	6.4070e-01	-2.7673e-03	1.1580e-01
	GDE3+EWMA	3.2311e-03	5.4464e-03	6.3435e-03	6.4046e-01	-8.1061e-03	1.1841e-01
	Wilcoxon	1.3324e-01	5.6832e-01	6.0702e-01	8.6131e-01	2.9286e-01	6.9133e-01
5e4	GDE3	4.9999e-04	1.8309e-03	2.0270e-03	5.8175e-01	-1.3082e-01	-3.8205e-02
	GDE3+EWMA	**2.0378e-04**	**1.4084e-03**	**1.7882e-04**	**5.6514e-01**	**-2.0052e-01**	-4.6213e-02
	Wilcoxon	2.4282e-16	2.7722e-03	8.5948e-21	2.8644e-06	4.0373e-06	1.5823e-01
5e5	GDE3	3.3757e-04	1.5353e-03	1.6889e-03	4.7174e-01	-1.8015e-01	-6.6889e-02
	GDE3+EWMA	**9.2885e-06**	**2.3457e-04**	**1.1889e-09**	**1.8951e-01**	**-3.3240e-01**	**-9.4063e-02**
	Wilcoxon	8.5941e-14	4.1942e-16	4.3099e-28	2.3621e-20	4.8987e-18	7.1863e-08

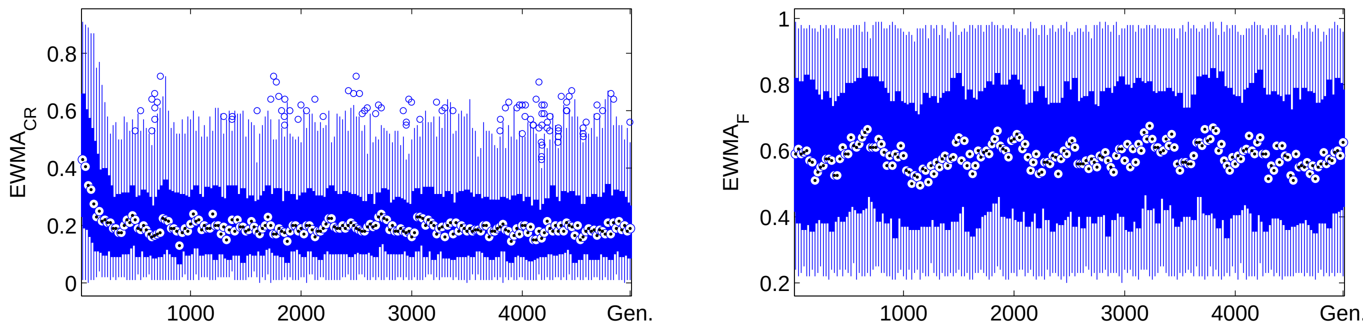

Fig. 1. Box plots of *EWMA* values on every 25 generations when random initial control parameter values were used with OKA2.

values could be $CR = 0.9$, $F = 0.9$, and $NP = 10 \cdot D$ [2]. In the case of multi-objective optimization, suitable control parameter values could be $CR = 0.2$, $F = 0.2$, and $NP = 100 \cdot (M - 1)$ [3].

As a future work remains to investigate and compare any other control parameter techniques that could be applicable with GDE3 in both single- and multi-objective optimization. Also, automatic stopping condition for GDE3 should be studied.

Acknowledgments

The first author would like to acknowledge support of the South Savo Regional Fund of the Finnish Cultural Foundation (Suomen Kulttuurirahaston Etelä-Savon rahasto) and support of the Mexican Government thought the Foreign Ministry (Gobierno de México, a través de la Secretería de Relaciones Exteriores). The second author gratefully acknowledges support from CONACYT project no. 221551.

References

[1] S. Kukkonen, "Generalized Differential Evolution for global multi-objective optimization with constraints," Ph.D. dissertation, Lappeenranta University of Technology, Acta Universitatis Lappeenrantaensis 475, May 2012.

[2] J. Aalto and J. Lampinen, "A mutation and crossover adaptation mechanism for differential evolution algorithm," in *Proceedings of the 2014 IEEE Congress on Evolutionary Computation (CEC)*. Beijing, China: IEEE, July 2014, pp. 451–458.

[3] V. L. Huang, A. K. Qin, K. Deb, E. Zitzler, P. N. Suganthan, J. J. Liang, M. Preuss, and S. Huband, "Problem definitions for performance assessment on multi-objective optimization algorithms," School of EEE, Nanyang Technological University, Singapore, 639798, Technical Report, January 2007.

[4] K. Miettinen, *Nonlinear Multiobjective Optimization*. Boston: Kluwer Academic Publishers, 1998.

[5] C. M. Fonseca and P. J. Fleming, "Multiobjective optimization and multiple constraint handling with evolutionary algorithms–part I: A unified formulation," *IEEE Transactions on Systems, Man, and Cybernetics–Part A: Systems and Humans*, vol. 28, no. 1, pp. 26–37, Jan 1998.

[6] R. Storn and K. V. Price, "Differential Evolution - a simple and efficient heuristic for global optimization over continuous spaces," *Journal of Global Optimization*, vol. 11, no. 4, pp. 341–359, Dec 1997.

[2]Recommendation that the population size depends linearly from the number of decision variables, has been earlier given in different sources, *e.g.*, in [18], [19].

[3]If the final population is directly used as a result for multi-objective optimization, then it is good that the population size scales with the number of objectives. This kind of scaling has been performed earlier in [1].

[7] K. V. Price, R. Storn, and J. Lampinen, *Differential Evolution: A Practical Approach to Global Optimization*. Berlin: Springer-Verlag, 2005.

[8] K. V. Price, "An introduction to Differential Evolution," in *New Ideas in Optimization*, D. Corne, M. Dorigo, and F. Glover, Eds. London: McGraw-Hill, 1999, pp. 79–108.

[9] S. Kukkonen and J. Lampinen, "Constrained real-parameter optimization with Generalized Differential Evolution," in *Proceedings of the 2006 Congress on Evolutionary Computation (CEC 2006)*, G. G. Yen, S. M. Lucas, G. Fogel, G. Kendall, R. Salomon, B.-T. Zhang, C. A. Coello Coello, and T. P. Runarsson, Eds. Vancouver, BC, Canada: IEEE Press, July 2006, pp. 911–918.

[10] J. Lampinen, "DE's selection rule for multiobjective optimization," Lappeenranta University of Technology, Department of Information Technology, Technical Report, 2001.

[11] K. Deb, *Multi-Objective Optimization using Evolutionary Algorithms*. Chichester, England: John Wiley & Sons, 2001.

[12] S. Kukkonen and J. Lampinen, "An extension of Generalized Differential Evolution for multi-objective optimization with constraints," in *Proceedings of the 8th International Conference on Parallel Problem Solving from Nature (PPSN VIII)*, ser. Lecture Notes in Computer Science (LNCS), vol. 3242, Birmingham, England, Sept 2004, pp. 752–761.

[13] ——, "GDE3: The third evolution step of Generalized Differential Evolution," in *Proceedings of the 2005 Congress on Evolutionary Computation (CEC 2005)*. Edinburgh, Scotland: IEEE Press, Sept 2005, pp. 443–450.

[14] S. Kukkonen and K. Deb, "Improved pruning of non-dominated solutions based on crowding distance for bi-objective optimization problems," in *Proceedings of the 2006 Congress on Evolutionary Computation (CEC 2006)*, G. G. Yen, S. M. Lucas, G. Fogel, G. Kendall, R. Salomon, B.-T. Zhang, C. A. Coello Coello, and T. P. Runarsson, Eds. Vancouver, BC, Canada: IEEE Press, July 2006, pp. 3995–4002.

[15] ——, "A fast and effective method for pruning of non-dominated solutions in many-objective problems," in *Proceedings of the 9th International Conference on Parallel Problem Solving from Nature (PPSN IX)*, ser. Lecture Notes in Computer Science (LNCS), T. P. Runarsson, H.-G. Beyer, E. Burke, J. J. Merelo-Guervós, L. D. Whitley, and X. Yao, Eds., vol. 4193. Reykjavik, Iceland: Springer-Verlag, Berlin, Heidelberg, Sept 2006, pp. 553–562.

[16] D. Zaharie, "Critical values for the control parameters of Differential Evolution algorithms," in *Proceedings of Mendel 2002, 8th International Conference on Soft Computing*, R. Matoušek and P. Ošmera, Eds., Brno, Czech Republic, June 2002, pp. 62–67.

[17] S. Kukkonen and J. Lampinen, "Performance assessment of Generalized Differential Evolution 3 (GDE3) with a given set of problems," in *Proceedings of the 2007 IEEE Congress on Evolutionary Computation (CEC 2007)*, D. Srinivasan and L. Wang, Eds. Singapore: IEEE, Sept 2007, pp. 3593–3600.

[18] J. Lampinen and R. Storn, "Differential Evolution," in *New Optimization Techniques in Engineering. Studies in Fuzziness and Soft Computing*, G. C. Onwubolu and B. V. Babu, Eds. Springer-Verlag, 2004, vol. 141, pp. 123–166.

[19] J. Rönkkönen, S. Kukkonen, and K. V. Price, "Real-parameter optimization with Differential Evolution," in *Proceedings of the 2005 Congress on Evolutionary Computation (CEC 2005)*. Edinburgh, Scotland: IEEE Press, Sept 2005, pp. 506–513.