Optimization of Stochastic Problems with Probability Functions via Differential Evolution

Bayrammyrat Myradov [a†]

[a] *Consultant*

Chance constrained programming, quantile/Value-at-Risk (VaR) optimization and integral quantile / Conditional Value-at-Risk (CVaR) optimization problems as Stochastic Programming Problems with Probability Functions (SPP-PF) are one of the most widely studied optimization problems in recent years. As a rule real-life SPP-PF is nonsmooth nonconvex optimization problem with complex geometry of objective function. Moreover, often it cannot be formulated in normative mathematical programming form.

Differential evolution based approach together with Monte Carlo sampling technique of the chance constraint(s) evaluation for solution of SPP-PF is proposed in this study.

Convergence of differential evolution algorithm and some practical implementation aspects of the proposed approach are discussed.

The several case studies such as the blending problem, the quantile linear programming problem, optimization of runway area demonstrate the effectiveness of the proposed approach.

Keywords: stochastic programming; probability function; differential evolution; global optimization; Monte Carlo sampling.

1. Introduction

Many practically significant optimization problems in business, industry, engineering, economics, finance and other fields are reasonable to formulate with probability functions (see e.g. [1-9]).

Different types of the Stochastic Programming Problem with Probability Function (SPP-PF) were introduced by Charnes and Cooper [1], Kataoka [2], Miller and Wagner [3]. Since last more than fifty years the studies of similar problems have grown with theoretical developments. Research on algorithms and applications on these problems has also been very active, especially in recent years. As basic references to theory, algorithms and applications of theirs we refer to [4-10].

At present four subclasses of SPP-PF such as Chance Constrained Programming Problem (CCPP), Stochastic Optimization Problem with Probabilistic Objective Function, VaR problem or optimization problem with quantile objective function and CVaR problem or optimization problem with integral quantile function are well known [1-9].

There are indirect via reformulating of the initial stochastic problem to its deterministic equivalence and direct methods as well as numerical optimization algorithms for obtaining optimal solution of different SPP-PF (see e.g. [4-16]). The peculiar properties of these methods and algorithms depend on structure and characteristics of each SPP-PF.

If SPP-PF is a convex optimization problem then its only one optimum theoretically can be found using gradient-based algorithms (GBA). The theory of a convex SPP-PF is widely studied (see [4-9]). But even in this case it is not easy to obtain optimum practically for some convex SPP-PF, for example, when objective function of the problem has wide flatness. Kall and Mayer have interpreted it as hidden nonconvexity of the convex optimization problem [8]. Clearly, this phenomenon is not good for classical GBA. But to avoid this one is impossible. Therefore, it is necessary to develop special procedures [8]. Kall and Mayer also noted that the steepness of the function is an additional difficulty of stochastic linear programs with probability function [8]. Flatness and steepness show that even if

[†] *bbmrdv@gmail.com*

SPP-PF has the good mathematical properties such as convexity and differentiability, computational implementation of GBA and a reach of true value of optimum by GBA may not be easy.

Despite of many studies of differentiability properties of probability functions, investigations of structure and properties of these problems, their optimality conditions, developing of several methods and approaches, SPP-PF can be efficiently solved just in some notable cases (see e.g. [1-9]).

It should be noted also that even for stochastic linear programming problem with chance constraints, it is difficult theoretically to prove an existence of its optimal solution when so-called technological matrix (TM) or right-hand side (RHS) is continuous or infinite aside from a few select distributions [8, 9].

Due to SPP-PF being a general one nonconvex problem, recent attention of researchers has focused on the approximation solution methods. The sample-average approximation (SAA) of joint CCPP was studied by Luedtke and Ahmed [17]. Pagnoncelli et al also studied sample approximations of chance constrained problems [18]. Branda studied reformulation of nonlinear stochastic programs with several joint chance constraints by stochastic programs with suitably chosen penalty-type objectives and discussed solving the investigated problems using Monte Carlo simulation techniques for the cases when the set of feasible solution is finite or bounded infinite [19]. He also discussed applications of sample-approximation techniques to the problems with generalized integrated chance constraints and proposed rates of convergence for the set of feasible solutions [20].

There are several studies on using metaheuristic methods of optimization for solving CCPP. Aringhieri discussed the possibility of using Tabu Search and simulation for finding solution of joint CCPP in case when TM is stochastic [21]. Tanner and Beier proposed metaheuristic based on Tabu Search for finding good, feasible solutions to joint CCPPs with discretely distributed parameters [22], Poojari and Varghese have developed and investigated genetic algorithm based technique for solving CCPP [23].

Grzybowski considered joint CCPP with random both TM and RHS. He proposed Monte Carlo method combined with Global Optimization algorithms such as genetic algorithms, evolutionary programming methods, Simulated Annealing Algorithm to solve the initial problem. He noted although the solution of the problem via proposed approach "may be called satisfactory rather than optimal" but "taking into account the very complex nature of the optimization problem, the approach seems to be very attractive" [24].

So, next main difficulties of solving real-life complex SPP-PF can be marked as summarizing of this short overview [4-24]:

- Often it is impossible to reformulate real-life SPP-PF to its deterministic equivalence, especially when TM is stochastic;
- Flatness, steepness, the ravines and other "bad" modular surface of objective function makes this problem more difficult to solve;
- This problem is nonconvex nonsmooth one in general;
- Often it is impossible theoretically to prove of global optima existence for this problem;
- Often this problem cannot be formulated in classical normative mathematical form.

2. A general formulation of SPP-PF and preliminaries.

Let's formulate one-stage stochastic programming problem with probability function in general form as Problem G:

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

where: $x = (x_1, x_2, \dots, x_N)^T, X = \{x \mid x_n \in [x_n^L, x_n^U], n = \overline{1, N}\}$.

Subject to:

$$\mathcal{P}_j(x) \overset{\text{def}}{=} \mathbb{P}\left\{\omega \in \Omega\colon\ b^L_{i,j} \leq \tilde{g}_{i,j}(x,\omega) \leq b^U_{i,j}, i = \overline{1,I_j}\right\},\ \mathcal{P}_j(x) \geq \alpha_j,\ j = \overline{1,J} \qquad (2)$$

Here: $x \in \mathbb{R}^N$ is a decision variable vector; $\omega\colon \Omega \to \mathbb{R}^l$ is a random vector on probability space $(\Omega, \mathcal{F}, P)$;$\alpha_j \in (0,1)$ are given level of probability; $\tilde{g}_{i,j}(x,\omega)$ are stochastic functions; $b^L_{i,j}$, x^L_n and $b^U_{i,j}$, x^U_n are lower and upper levels of the $\tilde{g}_{i,j}(x,\omega)$ and x respectively; N - number of decision variables; I_j - number of the chance constraints jointed in j$^{\text{th}}$ individual chance constraint; J is number of individual chance constraints; $\mathbb{P}\{\bullet\}$ is a probability function.

Let's consider objective function$F(x)$ is defined, bounded and can be formulated as (see [1-10]):

Type A: A1) linear deterministic: $(c^T x)$, A2) nonlinear deterministic:$f(x)$, and A3) stochastic $(c^T x) + f(x) + \mathbb{E}\{\tilde{f}(x,\omega)\}$ ones;

and Type B: B1) probabilistic: $P_\varphi(x) \overset{\text{def}}{=} \mathbb{P}\{\tilde{f}(x,\omega) \leq \varphi\}$, B2) quantile: $\varphi_\alpha(x) \overset{\text{def}}{=} \min\,(\varphi\colon P_\varphi(x) \geq \alpha)$, and B3) integral quantile: $\psi_\alpha(x) \overset{\text{def}}{=} \{\tilde{f}(x,w)\}_\alpha \overset{\text{def}}{=} \frac{1}{1-\alpha}\int_\alpha^1 \varphi_\beta(x)d\beta$ functions. Here $\varphi \in \mathbb{R}^1$ is α - quantile of $\tilde{f}(x,\omega)$.

The constraints (2) are called as combined chance (probabilistic) constraints. If $I_j = 1$ and $J = 1$ then Problem G has single chance constraint. If $I_j = 1$ and $J > 1$ then this problem has J individual chance constraints (ICCs) only. If $I_j > 1$ and $J = 1$ it has single joint chance constraint (JCC) with I chance constraints. If $I_j > 1$ and $J > 1$ it has J ICCs with I_j JCCs in each one.

The feasible set of Problem G is defined as $X^f := \{X \in \mathbb{R}^N : \mathcal{P}_j(x) \geq \alpha_j\}$. Assume that the feasible set of Problem G is not empty, i.e. $X^f \neq \emptyset$.

CCPP is the Problem G with $F(x)$of Type A and the chance constraint(s) (2). This is the most studied subclass of Problem G.

Stochastic optimization problem with probabilistic optimization function is formulated as [5-6]:

$$F(x) \overset{\text{def}}{=} P_\varphi(x) \to max_{x\in X} \qquad (3)$$

VaR problem or optimization problem with quantile objective function $\varphi_\propto(x)$ is formulated as [5-6]:

$$F(x) \overset{\text{def}}{=} \varphi_\alpha(x) \overset{\text{def}}{=} \min\,(\varphi\colon P_\varphi(x) \geq \alpha) \to min_{x\in X} \qquad (4)$$

CVaR problem or optimization problem with integral quantile function is formulated as [5-6]:

$$F(x) \overset{\text{def}}{=} \psi_\alpha(x) \to min_{x\in X} \qquad (5)$$

As mentioned above, in fact quite often real-life Problem G is a hard nonconvex nonsmooth optimization problem and objective function $F(x)$ has many local extrema. In this case locating its global extremum is a very challenging task. In some cases $F(x)$ may have several global extrema with same or near the same values of$F(x)$**,** though the values of decision variable x can be appreciably differ from each other. Therefore, often it is useful to study suboptimal (near optimal) solutions of Problem G.

Definition 1: $F(x^\varepsilon)$ is called ε – suboptimal or near optima value of objective function $F(x)$ if $\|\{F(x^s)\} - \{F(x^o)\}\| \leq \varepsilon$**,** where $x^o = \arg\min_{x\in X}\{F(x)\}$ and $\varepsilon > 0$.

The set of ε - suboptimal values is defined as $X^\varepsilon := \{x \in \mathbb{R}^N : F(x^s) - F(x^o) \leq \varepsilon\}$. Obviously that is $X^\varepsilon \subset X^f$. Suppose we can choose an appropriate ε, which can meet the required accuracy and make $X^\varepsilon \neq \emptyset$.

The set of X^{ε} appears greatly, for example, when $F(x)$ has wide flat optimality zone.

Absence of a priori information on properties and character of global behavior of objective function $F(x)$ and structure of its local extrema is the usual for complex real-life SPP-PF. Therefore such "good" mathematical properties of $F(x)$as convexity and differentiability are not required in this study. Obviously this essentially complicates an optimization problem solution.

As it is well known if $F(x)$ is multimodal, noisy and/or nondifferentiable then GBAs as a rule get trapped in one of its local extrema. And such extremum can quite be false when $F(x)$ is noisy. Usually local optimization methods of convex programming are reasonably tried to adapt for nonconvex problems. For example, generalizations of the stochastic quasigradient method to several classes of nonconvex nonsmooth stochastic optimization problems are considered by Ermoliev and Norkin [10]. Also the use of special procedures like multistarts is very popular in local optimization methods for finding of the conditional global optimum of the problem. But approach for adaptation of local optimization methods to nonconvex nonsmooth problems is not often effective in practice.

Therefore developing global optimization methods is more useful for complex real-life optimization problems in practice. Taking into account high complexity of real-life Problem G in general such as its nonlinearity, nondifferentiability, nonsmoothness, nonscalability, discontinuity, analytically non defined form, multimodality in the presence of ravines, plateaus, valleys, basins and other "bad" modular surface of objective function may be derivative free methods of optimization is one of the more suitable methods for solving or at least for finding of near-optimal or good-enough solutions of this problem. Especially derivative free methods of optimization may be most preferred ones for optimization of real-life complex multi structural system with random changing of its structure. These methods do not require calculation of any form of gradients (sub-, quasi-, pseudo-, generalized etc.). They are developing very intensively in recent years. Theoretically they can substitute for gradient – based methods, but not vice-versa.

In recent years, metaheuristics as derivative free methods for global optimization have become more and more attractive for solving hard optimization problems.

Usually these methods are classified into two main groups: population-based metaheuristics (e.g. Genetic Algorithms, Particle Swarm Optimization and Differential Evolution (DE)) and neighborhood metaheuristics (e.g. Simulated Annealing and Tabu Search).

There are some disadvantages of these methods such as lack of mathematically strong convergence proof with theoretical estimation of convergence rate and they are usually require high enough computing time.

The purpose of this study is to illustrate the efficiency of using differential evolution method as one of the modern derivative free and metaheuristic method for finding of optimal/suboptimal (near optimal)/good-enough solution of Problem G.

3. Differential evolution method

DE is a stochastic multi-agent evolutionary method of global optimization. It was developed by Storn and Price [25]. The most popular and basic (classical, canonical, standard) DE algorithm (DEA) is named DE/rand/1/bin.

DEA starts from generating of a random initial population. As a rule:

$$x_{n,i_{NP},0} = x_n^L + rand_{n,i_{NP}}(0,1)(x_n^U - x_n^L),\ i_{NP} = \overline{1,NP}\,, n = \overline{1,N} \quad (6)$$

here, $rand_{n,i_{NP}}(0,1)$ is uniformly distributed random number between 0 and 1; NP is the population size.

After initialization, mutation vectors $v_{n,i_{NP},g}$for classical DEA are generated according to differential mutation strategies with crossover method as:

$$v_{n,i_{NP},g} = x_{n,r_1,g} + FM \cdot (x_{n,r_2,g} - x_{n,r_3,g}) \tag{7}$$

where: FM is the mutation factor; $r_1, r_2, r_3 \in [1, NP]$ are randomly chosen integers, and $r_1 \neq r_2 \neq r_3 \neq i_{NP}$; $g = \overline{0, GMax}$ is number of generation, $GMax$ is a maximal number of generations.

In the crossover operation, a recombination of the mutation (donor) vector $v_{n,i_{NP},g}$ and the target (parent) vector $x_{n,i_{NP},g}$ produce a trial vector$u_{n,i_{NP},g}$. Basic DEA employs the binomial crossover defined as:

$$u_{n,i_{NP},g} = \begin{cases} v_{n,i_{NP},g} & if\ (rand_n(0,1) \leq CR)\ or\ (n = n_{rand}) \\ x_{n,i_{NP},g} & otherwise \end{cases} \tag{8}$$

where: $n_{rand} \in [1, N]$ is the randomly chosen index.

The selection operation is used to choose the population in next generation between $u_{n,i_{NP},g}$ and $x_{n,i_{NP},g}$:

$$x_{n,i_{NP},g+1} = \begin{cases} u_{n,i_{NP},g} & if\ F(u_{n,i_{NP},g}) \leq F(x_{n,i_{NP},g}) \\ x_{n,i_{NP},g} & otherwise \end{cases} \tag{9}$$

The control parameters for DEA are: i) the population size NP, ii) mutation factor $FM \in]0, 1]$ as a rule, iii) the crossover rate $CR \in [0,1]$.

There are many different strategies and modifications of DE based methods and algorithms developments as well as studies of the control parameters influence for performance of differential evolution algorithms (see e.g. [26-29]).

Premature convergence and stagnation are the main drawbacks of DEAs. There are several well-known different strategies for preventing premature convergence such as insect prevention, uniform crossover, preselection or crowding, fitness sharing and increasing of population size. As a rule these strategies can help avoid of stagnation problem also. But some of these strategies can be harmful in practical implementation for optimization problem with certain properties of objective functions. For instance, uniform crossover is harmful with nonseparable functions but the parameter CR controls the degree of rotational invariance of DE, ipso facto making suitable of using DE for harder nonseparable optimization problems. In such cases (if $CR = 1$) stagnation can be prevented by randomized mutation scale factor FM (see [29]).

Zaharie showed that premature convergence can be prevented if mutation and crossover induce an increase of the population variance and found appropriate values for the control parameters of DEA [26]. Price et al. discussed a practical approach to global optimization by DE [27]. Feoktistov emphasized the next three keys to DEA success: i) spontaneous self-adaptability, ii) diversity control and iii) continuous improvement [28]. Ronkkonen investigated continuous multimodal global optimization with DE based methods [29].

Jeyakumar and Shanmugavelayutham provided an empirical study and analyzed convergence nature of 14 Differential Evolution (DE) variants by measuring their so-called convergence speed and quality measure to solve 14 global optimization problems [30].

Some significant theoretical results concerning convergence and stability of DE were also obtained. Dasgupta et al. studied the stability and convergence of the population-dynamics in DE and drew attention on some similarity of DE with classical gradient descent method [31]. Ghosh et al. established asymptotic convergence behavior of a classical DEA by applying the concepts of Lyapunov stability theorems [32]. Zaharie reviewed the theoretical results concerning DE convergence, stability and probability distributions of the populations generated during evolution [33]. Researchers continue their studies for increasing of convergence rate of DEA without compromising of quality of solution. For example, recently Ali et al. in [34] and Shamekhi in [35] proposed the modified DEAs that are faster than original one.

It is possible to say the DE extends the gradient search strategy to nondifferential objective function with complex geometry. In DE the direction that leads to one feasible point to another one is defined as the differential variation.

Numerous empirical studies show DE has capability to escape from the local optima. Many experiments showed that as a rule an evolution of population fits to dynamics a random “cloud” of points moving as the whole along of geometry of objective function iterating its distinctive features. In case of hit in a ravine "cloud" takes the form of this ravine and distribution of points becomes such that the population mean of a difference of two random vectors appears directed along of the long side of a ravine. It provides fast movement along the narrow extended ravines whereas for gradient-based methods in similar conditions it is typical oscillatory dynamics «from a wall to a wall». So, DE has capacity for dynamically to model features of geometry of objective function thanks to "internal" random numbers generator consummated as a difference between randomly chosen vectors of a current population. It is remarkable ability of DE for rapidly passing complex ravines providing efficiency even in case of a difficult geometry of objective function.

Perhaps thanks to this feature and the similarity with gradient-based methods, DE is sure and relatively fast convergence metaheuristic and provides high-quality solution for a hard optimization problem with reasonable computational time compared to other modern heuristics methods.

So, practical implementation of different DEAs shows it is enough effective for global optimization.

But strong proof of DEA’s global convergence is a difficult task. Theoretical analyses of the conditions for the convergence of evolutionary algorithms were studied, for example, by Rudolph [36] and He and Yu [37]. He and Yu noted any evolutionary algorithm with the elitist selection strategy satisfies the condition which means that once population enters the optimal set then its offspring has little possibility to escape from the optimal set, i.e. this set is attractive [37]. Their results can be applied to DEAs.

The sufficient conditions for global convergence of DEA were studied by Hu et al. just recently [38]. As noted in [38] the conclusions presented here are more relaxed and easier to check than in [36,37]. From our point of view, the transfer of the theoretical results of that study in application of DEA for solving SPP-PF is acceptable since $F(x)$ is defined and bounded and it is assumed that $X^{\varepsilon} \neq \emptyset$.

Definition 2. (Based on [38]). Let $\{X_g, g = 0,1,2,3,\dots\}$ be a population sequence generated by using DEA to solve the optimization Problem G. Then DE converges to the conditional global extremum of this problem, if and only if $\lim_{g\to\infty} P\{X_g \cap X^{\varepsilon} \neq \emptyset\} \to 1$.

A sufficient condition for the convergence of DE is given by next theorem.

Theorem 1. (Based on [38]) Consider using DEA to solve the optimization Problem G. In the $g_k th$ target population X_{g_k}, there exists at least one individual x, which corresponds to the trial individual u by a reproduction operator, such that $P\{u \in X^{\varepsilon}\} \geq \zeta(g_k) > 0$ and the series $\sum_{k=1}^{\infty} \zeta(g_k)$ diverges; then DE converges to the optimal solution set X^{ε}. Where $\{g_k, k = 1,2,\dots\}$ denotes any subsequence of natural number set,$P\{u \in X^{\varepsilon}\}$ denotes the probability that u belongs to the suboptimal solution set X^{ε}, and $\zeta(g_k)$ is a small positive value which may change as g_k.

Proof of this theorem is the same as in [38].

Corollary 1 [38]. In Theorem 1, if $\zeta(g_k)$ equals ever to a positive constant $\zeta > 0$, then DE convergence to the optimal solutionX^{ε}.

Theorem 1 requires $P\{u \in X^{\varepsilon}\}$ should be large enough. Hu et al. have recently studied the problem of development different convergent DE variants in theory which holding convergence in probability [38–40]. They also noted there are other DEA variants which satisfies the convergence condition (e.g. [41-42]). As we can see, often simple modifications of classical DEA is required for making DEA convergent one.

4. Practical implementation aspects of DEA for solving of SPP-PF.

4.1 Description of modified DEA.

In spite of many studies of DEAs' application for different optimization problems, especially for unimodal and multimodal deterministic ones (see e.g. [25-29], [34-35]), the use of DE for solving of Problem G is not known.

As mentioned above there are DEAs with global convergence in probability. Therefore we have decided to make use of one of such convergent algorithms in this first study of application DEAs for solving of Problem G.

Modified classical DEA with global convergence in probability according to Theorem 1 via modified mutation vectors which have been proposed by Hu et al. in [40] is used for solving of Problem G in this study. Main reason for use of this algorithm is that the subspace clustering mutation operator can be easily incorporated into any state-of-the-art DEAs, thereby developing their convergent variants in theory as Hu et al. emphasized that in [40].

Modified mutation vectors $v^{M}_{n,i_{NP},g}$are generated as:

$$v^{M}_{n,i_{NP},g} = \begin{cases} v_{n,i_{NP},g} & if\ r_1 \in rand(1, \lfloor NP(1+R_b)\rfloor) \leq NP \\ x_{n,R_{btop},g} + rand(0,1)\left(x_{n,b_1,g} - x_{n,b_2,g}\right) & otherwise \end{cases} \tag{10}$$

Where: R_b is the increasing factor of the random integer r_1 region;$x_{n,R_{btop},g}$ is an individual selected by randomly sampling from the top R_b of the gth population; $x_{n,b_1,g}$ and $x_{n,b_2,g}$ are two boundary individuals, each element of which is equal to the upper or lower boundary value with an equal probability [40].

According to [40] the probability that at least one donor individual v locates in the $\varepsilon - suboptimal$ set can be estimated as:

$$P\{v \in X^{\varepsilon}\} \geq NP \cdot R_b \cdot P^{N}_{min} \cdot \frac{\mu(X^{\varepsilon})}{\mu(X^{f})} \tag{11}$$

$$P^{N}_{min} = \min\{P^{N}_{n} > 0, n = 1, N\} \tag{12}$$

Here: P^{N}_{n} is the probability of each donor individual locating in any subspace; $\mu(\cdot)$ is denotes the measure of a measurable set.

Since $0 < CR < 1$ that [40]:

$$P\{u \in X^{\varepsilon}\} \geq NP \cdot R_b \cdot P^{N}_{min} \cdot \frac{\mu(X^{\varepsilon})}{\mu(X^{f})} \cdot (1 - CR) > 0 \tag{13}$$

So if $\zeta(g_k)$ takes [40]:

$$\zeta(g_k) \equiv\ NP \cdot R_b \cdot P^{N}_{min} \cdot \frac{\mu(X^{\varepsilon})}{\mu(X^{f})} \cdot (1 - CR) > 0 \tag{14}$$

Then $\sum_{k=1}^{\infty} \zeta(g_k)$ diverges.

Therefore DEA with modified mutation vectors (10) holds global convergence.

4.2 Evaluation of chance constraint's satisfaction.

The use of effective optimization algorithm is necessary but not sufficient for solving of Problem G with chance constraint(s).

The evaluation of chance constraint's satisfaction is the quite difficult separate task (see [4-9]). Therefore the use of Monte Carlo sampling technique for solving this task is reasonable when there are no other effective methods.

Consider the chance constraint:

$$\mathbb{P}\{\tilde{g}(x,\omega) \geq 0\} \geq \alpha \qquad (15)$$

Let the set A be the set of the events when the constraint $\tilde{g}(x,\omega) \geq 0$ holds true. The characteristic function is:

$$\mathbb{1}_A(x) = \begin{cases} 1 & if \quad x \in A \\ 0 & if \quad x \notin A \end{cases} \qquad (16)$$

The process of the chance constraint evaluation is considered as a Bernoulli trial process.

Denote by $N_s$the number of satisfying of the chance constraint (15) in a fixed number of trials N_{tr} and by $\hat{\alpha}$ an empirical estimation of α. Then:

$$N_s = \sum_{i_{tr}=1}^{N_{tr}} \mathbb{1}_A(x) \qquad (17)$$

And

$$\hat{\alpha} = N_s/N_{tr} = \mathbb{E}\{\sum_{i_{tr}=1}^{N_{tr}} \mathbb{1}_A(x)\} \qquad (18)$$

This means $\mathbb{P}\{\tilde{g}(x,\omega) \geq 0\} \geq \alpha$ can be replaced by $\hat{\alpha} \geq \alpha$ because according to the Law of Large Numbers $\hat{\alpha}$ tends to α with probability 1 as $N_{tr} \to \infty$. Moreover, $\hat{\alpha}$ converges with probability 1 to α uniformly on any compact subset of X.

Denote by N_f the number of fails to satisfy of the constraint (15) in a fixed number of trials N_{tr}.In order to $\hat{\alpha} \geq \alpha$ there is a need to:

$$N_f \leq N_{tr} - N_s \leq \lceil N_{tr}(1-\alpha) \rceil \qquad (19)$$

This inequality is used for evaluation of the chance constraint (15) satisfaction by computer simulation via Monte Carlo sampling technique.

This approach of the chance constraint(s) evaluation has several advantages such as simplicity, independency from distribution function of random variables, capacity for work with random variables in TM and RHS as well as with correlated random variables. It should be noted we do not need to simulate all N_{tr} continuously. If during the simulating process the number of fails will achieve N_f for givenN_{tr} and α then the process of evaluation of (15) will be finished.

Number of trials (N_{tr}), i.e. the sample size necessary to establish the mean value α to within $\pm\Delta$ with a confidence of β can be determined by well-known formula:

$$N_{tr} \geq \lceil Z_\beta^2 \frac{\alpha(1-\alpha)}{\Delta^2} \rceil \qquad (20)$$

where: Z_β is the $(1 - \frac{1}{2}\beta)$ percentile of a standard normal distribution.

Often it is useful to know interval estimating of α in practice for reasonable decision making.

Well-known Wilson score interval is often used for interval estimating of unknown probability by frequency:

$$CI_W = \frac{1}{1+\frac{1}{N_{tr}}Z_\beta^2}\left[\hat{\alpha} + \frac{1}{2N_{tr}}Z_\beta^2 \pm Z_\beta\sqrt{\frac{1}{N_{tr}}\hat{\alpha}(1-\hat{\alpha}) + \frac{1}{4N_{tr}^2}Z_\beta^2}\right] \qquad (21)$$

where: $\hat{\alpha}$ is the proportion of successes in a Bernoulli trial process estimated from the statistical sample N_{tr},Z_β is the $(1-\frac{1}{2}\beta)$ percentile of a standard normal distribution, β is the error percentile and N_{tr} is the sample size.

For example, the results of calculation at $\hat{\alpha} = 0.95$, $\beta = 5\%$, $N_{tr} = 10^5$ are lower bound of $\hat{\alpha}$ is $\hat{\alpha}_{LB_W} = 0.9486$ and upper bound of $\hat{\alpha}$ is$\hat{\alpha}_{UB_W} = 0.9513$. Whether this confidence interval $CI_W = (\hat{\alpha}_{LB_W}; \hat{\alpha}_{UB_W})$ is acceptable or not for decision maker depends on many factors. But undoubtedly any decision maker needs such information. In addition, decision maker may estimate "corrected" $\hat{\alpha}_c$ for achieving$\hat{\alpha}_{LB_W} = 0.95$. In this case it is $\hat{\alpha}_c = 0.9514$. It is supposed that Wilson score interval has good properties. But it is very conservative for fixed N_{tr} as $\hat{\alpha}$ decreases.

In the case of rare or common event when α is closer to 0 or 1 and/or there is no tolerance in the coverage levels it is recommended to use Clopper-Pearson interval (CI_{CP}) which is also called as exact method in spite of its conservativeness.

A more common way to represent CI_{CP} is using the equations below:

$$CI_{CP} \triangleq (\hat{\alpha}_{LB_{CD}}; \hat{\alpha}_{UB_{CP}}) \triangleq \begin{cases} \hat{\alpha}_{LB_{CP}} = 1 - BetaInv(1-\frac{\beta}{2}, N_{tr} - N_s + 1, N_s) \\ \hat{\alpha}_{UB_{CP}} = 1 - BetaInv\left(\frac{\beta}{2}, N_{tr} - N_s, N_s + 1\right) \end{cases} \qquad (22)$$

According to Clopper-Pearson interval estimation the population proportion falls in the range $\hat{\alpha}_{LB_{CP}}$ to $\hat{\alpha}_{UB_{CP}}$ where: $\hat{\alpha}_{LB_{CP}}$ is the Clopper-Pearson's confidence interval lower bound, $\hat{\alpha}_{UB_{CP}}$ is the Clopper-Pearson's confidence interval upper bound, β is the percent chance of making a Type I error, $1-\beta$ is the confidence, $BetaInv(\gamma, \mu, \sigma)$ is the inverse cumulative distribution for a beta random variable with γth probability and shape parameters μ and σ.

5. Numerical experiments.

There are the practical guidelines concerning the tuning of DEA for solving an optimization problem [see e.g. [26-29]).

One more problem under discussion in an execution of DEAs is the stopping criteria. There are several studies hereof (see, e.g. [43-44]).

Certainly holding of the knowledge of some characteristics of an objective function is preferably for choosing more appropriate type of DE strategy and controlling parameters including stopping criteria.

As a whole, finding of well balance between type of DE strategy and controlling parameters including stopping criteria which helps to prevent premature convergence and stagnation ensuring at that admissible convergence speed of real-life complex optimization problem is a challenging separate task and it will be discussed in another study.

The following values for the parameters are used on default in this study: NP=100; FM=0.5; CR=0.9; $MultistartMax$=1;$GMax$=10000; EPS=10^{-6}, R_b=0.1 if they are not changed in the cases below. The algorithm execution is stopped when the vector population meets the requirement $|F - F_{best}| \leq EPS$.

The results of the numerical experiments for the several cases are discussed in this paper. The calculations are executed on PC Intel Core 2 Duo CPU 2.2 GHz, 2.96 Gb RAM. The results discussed in this paper are obtained as a consequence of performance of software which have been written on FORTRAN and compiled on FORTRAN 95.

5.1 Case 1. Blending Problem

The blending problem is formulated as [18-19]:

$$\min_{x\in X} F(x) := \min_{x_1\geq 0, x_2\geq 0}(x_1 + x_2) \qquad (23)$$

$$s.t.\ \mathbb{P}\{\omega_1 x_1 + x_2 \geq 7, \omega_2 x_1 + x_2 \geq 4\} \geq \alpha,\ \alpha = 0.95;\ \omega_1 \sim rand(1{,}4), \omega_2 \sim rand\left({}^{1}/_{3}, 1\right) \qquad (24)$$

Initial population generated as uniformly random distribution: $x_{n,i_{NP},0} \sim rand_{n,i_{NP}}(0; 100)$, $n = \overline{1,N}$, $i_{NP} = \overline{1,NP}, N = 2$. It is considered$x_n \in [0; 100]$.

The numerical experiments are realized for two value of number of trials: $N_{tr[1]} = 10^5$ and $N_{tr[2]} = 5 \cdot 10^5$as well as for two value of number of populations$NP_{(1)} = 100$, $NP_{(2)} = 20$.

The results of the numerical experiments are presented in Table 1.

Table 1. The results of optimization for blending problem (Case 1).

		$\{1\}N_{tr[1]}$		$\{2\}N_{tr[2]}$		$\{3\}N_{tr[1]}$		$\{4\}N_{tr[2]}$
α	**0.95**	**0.95**		**0.95**	α_c	**0.95137**		**0.95061**
ECI forα**:** $[\alpha^L; \alpha^U]$		**[0.9487;** **0.9513]**		**[0.9494;** **0.9506]**	$[\alpha_c^L; \alpha_c^U]$	**[0.95;** **0.9527]**		**[0.95;** **0.9512]**
$\hat{\alpha}$		**0.9507**		**0.9501**	$\hat{\alpha}_c$	**0.9521**		**0.9512**
ECI for $\hat{\alpha}$**:** $[\hat{\alpha}^L; \hat{\alpha}^U]$		**[0.9493;** **0.9520]**		**[0.9495;** **0.9507]**	$[\hat{\alpha}_c^L; \hat{\alpha}_c^U]$	**[0.9508;** **0.9534]**		**[0.9506;** **0.9518]**
	ES	**NS [1]**	**RE [1]**	**NS [2]**	**RE [2]**	**NS [3]**	**RE [3]**	**NS [4]**
$x_{1(1)}$ $x_{1(2)}$	**3.673**	**3.651** **3.812**	**-0.60%** **3.78%**	**3.654** **3.818**	**-0.52%** **3.95%**	**3.591** **3.818**	**-2.23%** **3.95%**	**3.591** **3.818**
$x_{2(1)}$ $x_{2(2)}$	**2.776**	**2.801** **2.715**	**0.90%** **-2.20%**	**2.801** **2.715**	**0.90%** **-2.20%**	**2.885** **2.715**	**3.93%** **-2.20%**	**2.885** **2.715**
$F_{(1)}$ $F_{(2)}$	**6.449**	**6.452** **6.533**	**0.05%** **1.30%**	**6.454** **6.533**	**0.08%** **1.30%**	**6.476** **6.533**	**0.42%** **1.30%**	**6.476** **6.533**
Time$_{(1)}$ NP=100 **Time$_{(2)}$ NP=20**		**47”** **6”**		**4’23”** **37”**		**25”** **6”**		**2’17”** **37”**

Here: ECI – exact confidence interval; **ES** – exact solution; **NS** – numerical solution; **RE** – relative error.

Exact confidence intervals are calculated by (22). The given level of probability α for the numerical experiments {3} and {4} are calculated so in order to equal its low bound to 0.95 with confidence level of 95%.

The results of experiments $\{3\}$ and $\{4\}$ are identical on x and F while α and α_c differ. Also the results at $NP_{(2)} = 20$ are almost identical for all experiments.

So the groundless increase of number of trials is useless from the point of view of quality of the obtained solution as well as time of its calculation. Application of the upper bound of the confidential interval for obtaining the guaranteed result with required probability leads to more conservative solution. The more population size leads to better quality of the solution but time for calculation is more.

As a whole the high quality solutions of blending optimization problem were obtained for reasonable computational time by proposed approach.

5.2 Case 2. Quantile linear programming problem

Let's consider the quantile linear programming problem (QLPP) [13]:

$$\varphi_\alpha \triangleq \min_{x \in X} F_\alpha(x), \quad x_\alpha \triangleq arg \min_{x \in X} F_\alpha(x) \tag{25}$$

$$F_\alpha(x) \triangleq \min\{\varphi : P_\varphi(x) \geq \alpha\}, \alpha \in (0, P^*) \tag{26}$$

$$P^* \triangleq sup_{x \in X} \mathbb{P}\{Q(x,\omega) \leq 0\} \tag{27}$$

$$P_\varphi(x) \triangleq \mathbb{P}\{F(x,\omega) \leq \varphi, \ Q(x,\omega) \leq 0\} \tag{28}$$

$$Q(x,\omega) \triangleq \max_{i=1,I}\{A_{1i}^T x + B_{1i}^T \omega + b_{1i}\} \tag{29}$$

$$F(x,\omega) \triangleq \max_{j=1,J}\{A_{2j}^T x + B_{2j}^T \omega + b_{2j}\} \tag{30}$$

Initial data: $A_1 = \begin{pmatrix} 6 & -6 \\ 1 & 1 \end{pmatrix}, B_1 = \begin{pmatrix} 6 & -6 \\ 1 & 1 \end{pmatrix}, \ b_1 = \begin{pmatrix} 0 \\ -5 \end{pmatrix}, A_2 = (10\ -10), \ B_2 = (-3 - 1), b_2 = (-2)$

$X \triangleq \{x : x \in \mathbb{R}^2, \ 0 \leq x_1 \leq 10, \ 0 \leq x_2 \leq 10\}, \omega \in \mathbb{R}^2 \sim \aleph(0, D), D \triangleq \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$

Table 2. The results of the solutions for QLPP (Case 2).

	x_1	x_2	φ	$\hat{\alpha}$	Time	MRE on x_2	MRE on φ
***Solutions by proposed approach with Multistart=3*(M1,M2,M3)**							
M1	**0**	**0.7250**	**13.97**	**0.9002**	**2'02"**	**13.28%**	**0.14%**
M2	**0**	**0.7942**	**13.96**	**0.9001**	**2'13"**		
M3	**0**	**0.7011**	**13.95**	**0.9001**	**3'22"**		
Solutions by four different algorithms presented in [13]							
1	**0**	**0.3218**	**18.656**	**Algorithm 1**			
2	**0**	**1.4400**	**14.674**	**Algorithm 2**			
3	**-0.0025**	**1.3370**	**13.906**	**Quasigradient algorithm**			
4	**0**	**2.0480**	**13.306**	**Solution on kernel**			

It is interesting to see an effect of multistarts of initial populations to the results of optimization for the Case 2. Maximum Relative Error (MRE) on x_2 with 3 multistarts (M1,M2 and M3) is 13.28% whereas MRE is 0.14% only for$F$. Multistarts of initial population are not important in this case. The comparison of the obtained results by proposed approach with other methods of solutions presented in [13] shows the high quality of the obtained results.

5.3. Case 3.Optimization of Runway Area

Optimization problem of runway area is formulated as [5-6]:

$$F(x_1, x_2, x_3) = 2x_3(x_0 + x_1 + x_2) \to min_{x_1, x_2, x_3} \quad (31)$$

s.t. $\mathbb{P}\{-x_1 \leq W \leq x_2,\ |Z| \leq x_3\} \geq \alpha$ (32)

$$x_1, x_2, x_3 \geq 0 \quad (33)$$

$$W = a_{11}W_x + a_{12}|W_z| \quad (34)$$

$$Z = a_{22}W_z \quad (35)$$

Initial data: $x_0 = 1500\, m$; $a_{11} = a_{12} = -20\, sec$; $a_{22} = 3\, sec$; $W_x, W_z \sim\aleph(0;\ 5\, m/sec)$.

The solution of this problem was obtained in [5-6] by reformulating of this initial problem to equivalence quantile optimization problem.

The results of the calculations by proposed approach are presented in Table 3 and Table 4. In general, the comparison of the results presented in Table 3 with other methods of solutions presented in [6] shows the high quality of the obtained results by proposed approach.

Very often it is practically important to study of Problem G in “nine six” area, i.e. when α is closer to 1. As we can see in Table 3 the result of experiment {5} by proposed approach differs from the results obtained by other methods and it requires much more time for convergence of DEA.

Therefore, this experiment is studied separately. Given α is calculated in accordance with different permissible number of crash landing. The good-enough but not converged and converged solutions for each given α are shown in Table 4. Note that in this case ε - suboptimal or near optima and the good-enough solutions have exactly the same sense.

It should be noted $N_{tr} = 4.3 \cdot 10^6$ is estimated by formula (20) at $\beta = 0.05$ and $\Delta = 9.7 \cdot 10^{-7}$ for the experiment {5}. Opportunity of reasonable reducing required number of trials in this case via importance sampling or other techniques is not discussed here.

The good-enough solution when permissible number of crash landing equals to 4 obtained after 4 minutes of calculations whereas converged solution is achieved after 46 minutes.

Good enough solutions are found much faster than converged ones for other permissible numbers of crash landing also. At the same time there is no essential difference between good-enough and converged solutions.

As we can see the model is sensitive in small changing of α nearly 99.9999%. The proposed approach “catches” that the objective function increases on 48% from 0.402 to 0.595 under decreasing of acceptable crash landing from 10 to 4 (i.e. α changed very slightly from 0.9999977 to 0.9999991) at $N_{tr} = 4.3 \cdot 10^6$.

Table 3. Optimization of Runway Area at different given α. (Case 3).

	{1}	{2}	{3}	{4}	{5}
α	0.99	0.999	0.9999	0.99999	0.999999
N_{tr}	10^6	10^6	10^6	10^6	$4.3*10^6$
x_1, m	419.5	527.8	599.6	649.7	957.9
x_2, m	258.8	305.7	368.2	399.7	401.8
x_3, m	40.7	51.4	60.3	69.5	104.3
F, km^2	0.177	0.240	0.298	0.354	0.595
Time	8'25"	7'6"	5'20"	6'5"	46'5"
The results by other methods					
Analytical, F [5]	0.192	0.250	0.304	0.354	0.402
Analytical, F [45]	0,203	0,267	0,327	0,383	0,437
Confidence, F [5]	0.202	0.256	0.324	0.380	0.432
Stochastic Approximation,F [5]	0.170	0.236	0.304	0.366	-

Table 4. Optimization of Runway Area in depending on permissible number of crash landing at $N_{tr} = 4300000$ (Case 3).

Permissible number of crash landing	4		6		7		10	
$\alpha \approx$	0.9999991		0.9999986		0.9999984		0.9999977	
$\hat{\alpha}_{(\varepsilon)} \approx$ $\hat{\alpha}_{(c)} \approx$	0.9999993 0.9999991		0.9999986 0.9999986		0.9999984 0.9999984		0.9999977 0.9999977	
Solutions	ε	c	ε	c	ε	c	ε	c
x_1, m	958.7	957.9	825.8	825.8	956.5	956.5	919.2	928.7
x_2, m	402.1	401.8	459.0	459.0	517.6	471.1	473.9	470.8
x_3, m	104.1	104.3	105.0	105.0	69.74	69.25	74.79	69.36
F, km^2	0.595	0.595	0.585	0.585	0.415	0.405	0.433	0.402
Time	4'22"	46'5"	2'11"	7'43"	1'21"	16'54"	4'47"	19'2"

Here: ε - good-enough (ε-optimal) solution; c – converged solution;

6. Conclusions and future work.

The reasonability of using differential evolution method for optimization of stochastic problems with probability function(s), global convergence of differential evolution algorithms and their some practical implementation aspects is discussed in this paper.

The approach combines differential evaluation method with Monte Carlo sampling technique for chance constraint(s) evaluation is proposed for solving of stochastic programming problems with probability function(s).

Laboriousness, speed of convergence and quality of obtained solutions for the several case studies such as the blending, the quantile linear programming and optimization of runway area problems show efficiency of the proposed approach. These results have been compared with the results obtained by other methods. Also, there were found out some more interesting information at optimization of these problems. So, at the optimization of runway area problem, the proposed approach "catches" very sensitive range of the objective function values when this function increases on 48% from 0.402 to 0.595 under decreasing of acceptable crash landing from 10 to 4 at total landing of 4300000 (i.e. α changed very slightly from 0.9999977 to 0.9999991).

The cases also clearly demonstrate the importance of understanding the difference between given α and calculated $\hat{\alpha}$ as well as interval estimating of $\hat{\alpha}$ at using of numerical algorithms for solving of Problem G.

On the whole, from our point of view, the proposed approach can be suitable for finding of optimal / near-optimal / good-enough solutions of complex real-life stochastic problems with not well-behaved functions (nonlinear, nonconvex, nondifferentiable) having different "bad" properties such as multimodality, weak global structure, nonseparability, variable scaling, inhomogeneous search space and basin sizes, global to local optima contrast decay, plateaus.

The effectiveness of proposed approach for a complex real-life Problem G will be demonstrated in the nearest future in next study. This Problem G will be an optimization of real-life integrated water supply system with solar desalination unit formulated as stochastic simulation optimization problem with two stochastic objective functions and the chance constraints having correlated random variables in technological matrix. Also we will discuss challenges for improving of proposed approach.

References

1. Charnes A,Cooper W. Chance constrained programming. Management Science 1959; **6:**73–79.
2. Kataoka S.A stochastic programming model. Econometrica1963; **31:** 181–196.
3. Miller LB, Wagner H. Chance-constrained programming with joint constraints. Operations Research 1965; **13:**930–945.
4. Prékopa A. Stochastic Programming. Kluwer: Dordrecht, Boston, 1995; 600.
5. Kibzun AI, Kan YuS. Stochastic Programming Problems with Probability and Quantile Functions. Wiley: Chichester, 1996; 301.
6. Kibzun AI, Kan YuS. Stochastic Programming Problems with Probability Criteria. Fizmatlit:Moscow, 2009; 372(In Russian).
7. Prékopa A. Probabilistic Programming.In *Stochastic Programming. Handbooks in Operations Research and Management Science,*Ruszczyński A,. Shapiro A (eds), vol. 10. Elsevier: Amsterdam, 2003; 267-351.
8. Kall P, Mayer J. Stochastic linear programming. Models, Theory, and Computation. Springer:New York, 2005; 397.

9. Shapiro A, Dentcheva D, Ruszczyński A. Lectures on Stochastic Programming: Modeling and Theory. MPS-SIAM series on optimization: Philadelphia,2009; 436.
10. Ermoliev YuM, Norkin VI. Solution of nonconvex nonsmooth stochastic optimization problems. Cybernetics and Systems Analysis 2003; **39** (5):701-715.
11. Kibzun AI, Matveev EL. Stochastic quasigradient algorithm to minimize the quantile function Automation and Remote Control 2010; **3:** 413-430.
12. Kibzun AI, Chernobrovov AI. Stochastic quasigradient algorithm to minimize the function of integral quantile. Automation and Remote Control 2012; **2:** 41-60.
13. Naumov AV, Ivanov SV. On stochastic linear programming problems with the quantile criterion. Automation and Remote Control 2011; **2:** 353-369.
14. Norkin V. On mixed integer reformulations of monotonic probabilistic programming problems with discrete functions. http://www.optimization-online.org/DB_HTML/2010/05/2619.html
15. Ivanov SV, Naumov AV. Algorithm to optimize the quantile criterion for the polyhedral loss function and discrete distribution of random parameters. Automation and Remote Control 2012; **1:** 105-117.
16. Kibzun AI, Naumov AV, Norkin VI. On reducing a quantile optimization problem with discrete distribution to a mixed integer programming problem. Automation and Remote Control 2013;**6:** 66-86.
17. Luedtke J, Ahmed S. A sample approximation approach for optimization with probabilistic constraints. SIAM J. Optimization 2008; **19:** 674–699.
18. Pagnoncelli B, Ahmed B, ShapiroAS. Sample average approximation method for chance constrained programming: theory and applications. Journal of Optimization Theory and Applications 2009; **142**(2): 399–416.
19. Branda M. Stochastic programming problems with generalized integrated chance constraints. Optimization: A Journal of Mathematical Programming and Operations Research 2012;**61**(8): 949-968.
20. Branda M. Chance constrained problems: penalty reformulation and performance of sample approximation technique. Kybernetika2012; **48** (1): 105-122.
21. Aringhieri R. A tabu search algorithm for solving chance-constrained programs. Journal of the ACM 2004; **5:**1–14.
22. Tanner MW, BeierEB. A General Heuristic Method for Joint Chance-Constrained Stochastic Programs with Discretely Distributed Parameters. http://www.optimization-online.org/DB_HTML/2007/08/1755.html
23. Poojari CA, Varghese B. Genetic algorithm based technique for solving chance constrained problems. European Journal of Operational Research 2008,**185 (3):**1128-1154.
24. Grzybowski AZ. Simulation analysis of global optimization algorithms as tools for solving chance constrained programming problems.ActaElectrotechnicaetInformatica2011; **11**(4):60–65.
25. Storn R, Price K. Differential Evolution – ASimple and Efficient Heuristic for Global Optimization over Continuous Spaces. Journal of Global Optimization 1997; **11:** 341-359.
26. Zaharie D. Critical values for the control parameters of differential evolution algorithms. In *Proceedings of 8th MENDEL International Conference on Soft Computing*. Brno,2002; 62–67.
27. Price KV, Storn RM,LampinenJA.: Differential Evolution: A Practical Approach to Global Optimization. Springer:Berlin,2005; 538.
28. Feoktistov V. Differential Evolution: In Search of Solutions. Springer: New York,2006; 195.
29. Ronkkonen J. Continuous Multimodal Global Optimization with Differential Evolution Based Methods. Diss. Lappeenranta University of Technology: Lappeenranta, 2009; 168.
30. Jeyakumar G, Shanmugavelayutham C. Convergence analysis of differential evolution variants on unconstrained global optimization functions. International Journal of Artificial Intelligence & Applications (IJAIA) 2011; **2**(2):116-127.

31. Dasgupta S, Das S, Biswas A, Abraham A. On stability and convergence of the population-dynamics in differential evolution. AI Commun. 2009; **22** (1):1-20.
32. Ghosh S, Das S, Vasilakos AV, Suresh K. On Convergence of Differential Evolution over a Class of Continuous Functions with Unique Global Optimum. IEEE Trans. on Systems, Man, and Cybernetics,Part B 2012; **42** (1): 107-124.
33. Zaharie D. Differential evolution from theoretical analysis to practical insights.In *Proceeding of 19th International Conference on Soft Computing.*Brno, 2013; 26-28.
34. Ali M, Pant M, Singh VP. Two modified differential evolution algorithms and their applications to engineering design problems. World Journal of Modelling and Simulation 2010; **6**(1):72-80.
35. Shamekhi A. An improved differential evolution optimization algorithm.International Journal of Research and Reviews in Applied Sciences (IJRRAS)2013; **15** (2):132-145
36. Rudolph G. Convergence of evolutionary algorithms in general search spaces.In *Proceedings of the Third IEEE Conference on Evolutionary Computation*. Piscataway, 1996; 50–54.
37. He J,Yu XH. Conditions for the convergence of evolutionary algorithms. Journal of Systems Architecture 2001;**47** (6): 601–612.
38. Hu Z, Xiong S, Su Q, Zhang X. Sufficient Conditions for Global Convergence of Differential Evolution Algorithm. Journal of Applied Mathematics 2013; Article ID 193196:1-14.
39. Hu Z,Xiong S, Fang Z,Su Q. A convergent Differential Evolution Algorithm with Hidden Adaptation Selection for Engineering Optimization. Mathematical Problems in Engineering 2014, Article ID 135652:1-11.
40. Hu Z, Xiong S, Wang X, Su Q, Liu M, Chen Z. Subspace Clustering Mutation Operator for Developing Convergent Differential Evolution Algorithm. Mathematical Problems in Engineering 2014, Article ID 154626:1-18.
41. Zhao Y, Wang J, Song Y. An improved differential evolution to continuous domains and its convergence.In *Proceedings of the 1st ACM/SIGEVO Summit on Genetic and Evolutionary Computation (GEC '09).*Shangai, 2009; 1061–1064.
42. Zhan Z, Zhang J. Enhance differential evolution with random walk, In *Proceedings of the 14th Annual Conference Companion on Genetic and Evolutionary Computation (GECCO'12).*Philadelphia,2012; 1513–1514.
43. Zielinski K,Laur R. Stopping criteria for differential evolution in constrained single-objective optimization. In *Advances in Differential Evolution,*vol. 143.Chakraborty UK (eds), Springer: Berlin, 2008; 111–138.
44. Roche D, Gil D, Giraldo J. Using statistical inference for designing termination conditions ensuring convergence of Evolutionary Algorithms. In *Proceedings of the Eleventh European Conference on the Synthesis and Simulation of Living Systems*, Boston, 2011; 680-687.
45. Dzotsoev A.A., Kan Yu.S., Shakhlevich P.K., Optimization of the area of a takeoff and landing runway // J. Computer and System Sci. Int., 2007, Volume 46, Issue 6, pp 881-886.