A Study on the Convergence of Multiobjective Evolutionary Algorithms

Tushar Goel[1], Nielen Stander[2]
Livermore Software Technology Corporation, Livermore CA

High computational cost has been a major impediment to the widespread use of evolutionary algorithms in industry. While the clock time for optimization using the GA can be reduced by parallelization, the computational cost can only be improved by reducing the number of function evaluations. For single objective optimization problems, the convergence curve can be utilized to obtain a suitable compromise between the computational cost and the quality of the solution. A non-domination criterion based metric that tracks the growth of an archive of non-dominated solutions over a few generations is proposed to generate a convergence curve for multi-objective evolutionary algorithms. Two analytical and two crashworthiness optimization problems were used to demonstrate the practical utility of this measure. It was observed that, similar to single-objective optimization problems, there were significant advances towards the POF in the early phase of evolution and relatively smaller improvements were obtained as the population matured. This information was used to terminate the search to obtain a good trade-off between the computational cost and the quality of the solutions. The paper also demonstrates the successful use of compute clusters for parallel processing to significantly reduce the clock time for optimization.

I. Introduction

Most practical engineering problems involve multiple design objectives and constraints. The optimization of such systems with more than one objective function is called multi-objective optimization. These objectives are often in conflict. Contrary to the single-objective optimization problem (SOP), the multi-objective optimization problem (MOP) does not result in a single optimum solution. Instead, it results in a set of optimum solutions that represent different trade-offs among the objectives. These solutions are known as Pareto optimal solutions or constitute the Pareto optimal solution set [1]. The function space representation of the Pareto optimal solution set is known as the Pareto optimal front. The most common strategy to find Pareto optimal solutions is to convert the multi-objective optimization problem to a single objective optimization problem and then find a single trade-off solution. There are multiple ways of converting a MOP to a SOP, namely, the weighted sum strategy, inverted utility functions, goal programming, ε-constraint strategy etc [1]-[3]. This biggest drawback of these conversion strategies is that each optimization simulation results in a single trade-off. Besides, multiple runs may not yield sufficiently diverse trade-off solutions [2].

Genetic algorithms (GAs) have been demonstrated to efficiently solve multi-objective optimization problems because they result in a diverse set of trade-off solutions in a single simulation run [2]. However, the GA requires tens of thousands of simulations to converge to the global Pareto optimal front. Though a recent study highlighted the application of a multi-objective GA with a small number of simulations [4], the high computational cost remains the biggest potential drawback for solving engineering problems that may involve impact or other expensive simulations to analyze the problem.

The continual reduction in the cost of the hardware, the clusters of processors have become increasingly common. Such systems are particularly useful for running genetic algorithms that are inherently suited to parallelization. One can reduce the clock time significantly by concurrently running many individuals (simulations) in a GA population. The time to analyze each design can be reduced further if the analysis code can utilize multiple processors as for instance is the case with the MPP version of LS-DYNA® [5] used in this study. While this two level parallelization has a multiplicative effect on the clock time-savings, the ever-increasing complexity of computational models may easily offset the potential gains of parallelization [6].

---
[1] Member AIAA
[2] Member AIAA

Most recent efforts in reducing the computational cost have been focused on developing more efficient algorithms (e.g., micro GA [7], ParEGO [8]), improving efficiency of existing algorithms [9], using a combination of local search methods with GA [10]-[12], or using approximation models [13]-[20] for fitness evaluation but the influence of a stopping criterion on the computational cost of the GA has largely been overlooked. The maximum number of generations is by far the most popular termination criterion for both single- and multi-objective GA. There have been some efforts in developing the bounds on the number of generations for single objective genetic algorithms [21]-[23]. Some early research work [24]-[27] also provided conditions for convergence for the multi-objective evolutionary algorithms (MOEA).

This paper attempts to reduce the total computation time for solving industrial multi-objective optimization problems by exploiting the convergence properties of the multi-objective evolutionary algorithm [28]. It is well known that for single-objective optimization, there is a considerable improvement in the quality of solutions in the initial phases of the GA simulation and incremental improvements are observed later on. One can terminate the GA search once a reasonable improvement is obtained thus saving a significant computational expense required to converge to the global optimal solution. In the last few years, researchers have started exploring efficient stopping criteria for multi-objective evolutionary algorithms [28]-[34].

Rudenko and Schoenauer [29] proposed the standard deviation of maximum crowding distance criterion to detect stabilization of the NSGA-II algorithm for two objective optimization problems. Marti et al. [30], [34] estimated the rate of improvement (MGBM criterion) using the measured rate of improvement based on the ratio of dominated solutions in a population as well as Kalman filters. The simulations were stopped when the estimated rate fell below the threshold value. While this measure performed reasonably on low dimension problems, the authors noted difficulty with a 10-objectives problem. This measure required a user-defined parameter to control the estimated rate of improvement.

Trautmann et al. [31] used the Kolmogorov-Smirnov test to statistically ascertain the convergence by comparing the generational distance [35], hyper-volume [36], and spread [37] metrics for a few generations. In a follow-up paper, Wagner et al. [32] advanced this idea to develop an online convergence detection (OCD) criterion by carrying out χ^2-variance test and t-test for linear behavior on different performance metrics. They stopped simulation when either of the tests indicated convergence. Later Naujoks and Trautmann applied the OCD criterion for the design of airfoil surfaces that minimize drag at different flow conditions. This method is complex and requires selecting many parameters.

While researchers have assumed that the MOEAs improve the solutions in the early phase of evolution, this was never demonstrated. In this paper, i) the convergence properties of multi-objective evolutionary algorithms were studied, and ii) a simple termination criterion/criteria were developed to stop the GA simulations when a reasonably good local Pareto optimal front was obtained instead of attempting to locate the global Pareto optimal front. Specifically, the elitist non-dominated sorting genetic algorithm (NSGA-II) [38] augmented with an archiving [16] strategy (NSGA-IIa) was used as the MOEA. Two analytical problems and two engineering problems were used to illustrate the main concepts. The engineering problems were represented by two crashworthiness examples analyzed using the nonlinear dynamic finite element analysis program LS-DYNA. It was demonstrated that there were significant improvements in the quality of local Pareto optimal front in the early stage of simulations. Using a stopping criterion based on novel convergence and diversity metric, the evolutionary search was terminated such that a reasonable trade-off in the quality of solutions and the computational cost was found.

The paper is organized as follows. The analytical and crashworthiness optimization test problems are described in the next section. The details of the simulation setup and performance metrics are furnished in the Section **Test Setup and Performance Metrics**. The main results of this study are documented in the **Results** section, and major findings are summarized in the **Summary** section.

II. Test Problems

This section lists different test problems used to illustrate the convergence behavior of the NSGA-IIa. The salient features and challenges of the two analytical problems and two crashworthiness problems are described in detail as follows.

A. Analytical Problems

1. *TNK [39] Problem*

Minimize:

$$f_1(\mathbf{x}) = x_1,$$
$$f_2(\mathbf{x}) = x_2,$$

Subject to:

$$C_1(\mathbf{x}) \equiv x_1^2 + x_2^2 - 1 - 0.1\cos\left(16\arctan(x_1/x_2)\right) \geq 0,$$
$$C_2(\mathbf{x}) \equiv 0.5 - (x_1 - 0.5)^2 - (x_2 - 0.5)^2 \geq 0,$$
$$0 \leq x_1, x_2 \leq \pi.$$

This problem results in a non-convex, discontinuous global Pareto optimal front that coincides with the boundary of the first constraint. It is important to note that not all the points on the constraint boundary are part of the global Pareto optimal front. This problem poses a difficulty in finding diverse solutions on the Pareto front.

2. OSY [40] Problem

Minimize:

$$f_1(\mathbf{x}) = -\left[25(x_1 - 2)^2 + (x_2 - 2)^2 + (x_3 - 1)^2 + (x_4 - 4)^2 + (x_5 - 1)^2\right]$$
$$f_2(\mathbf{x}) = x_1^2 + x_2^2 + x_3^2 + x_4^2 + x_5^2 + x_6^2,$$

Subject to:

$$C_1(\mathbf{x}) \equiv x_1 + x_2 - 2 \geq 0,$$
$$C_2(\mathbf{x}) \equiv 6 - x_1 - x_2 \geq 0,$$
$$C_3(\mathbf{x}) \equiv 2 + x_1 - x_2 \geq 0,$$
$$C_4(\mathbf{x}) \equiv 2 - x_1 + 3x_2 \geq 0,$$
$$C_5(\mathbf{x}) \equiv 4 - (x_3 - 3)^2 - x_4 \geq 0,$$
$$C_6(\mathbf{x}) \equiv (x_5 - 3)^2 + x_6 - 4 \geq 0,$$
$$0 \leq x_1, x_2, x_6 \leq 10,\ 1 \leq x_3, x_5 \leq 5,\ 0 \leq x_4 \leq 6.$$

This problem also has a known Pareto optimal front that lies on the constraint boundary. There are five connected sections of the Pareto front where different constraints are active. This problem poses great difficulty in finding the entire spread of the Pareto optimal front.

B. Crashworthiness Problems

1. Knee Bolster Design

The first crashworthiness problem employs a finite element simulation of a typical automotive instrument panel (shown in Figure 1) impacting the knees [41]. The spherical objects that represent the knees move in the direction determined from prior physical tests. The instrument panel (IP) comprises of a knee bolster that also serves as a steering column cover with a styled surface, and two energy absorption brackets attached to the cross vehicle IP structure.

A significant portion of the lower torso energy of the occupant is absorbed by appropriate deformation of these brackets. The wrap-around of the knee around the steering column is delayed by adding a device, known as the yoke, to the knee bolster system. The shape of the brackets and yoke are optimized without interfering with the styled elements. The eleven design variables are shown in Figure 2 and the ranges are given in Table 1. To keep the computational expense low, only the driver side instrument panel was modeled using approximately 25000 elements; and the crash was simulated for 40ms, by which time the knees had been brought to rest.

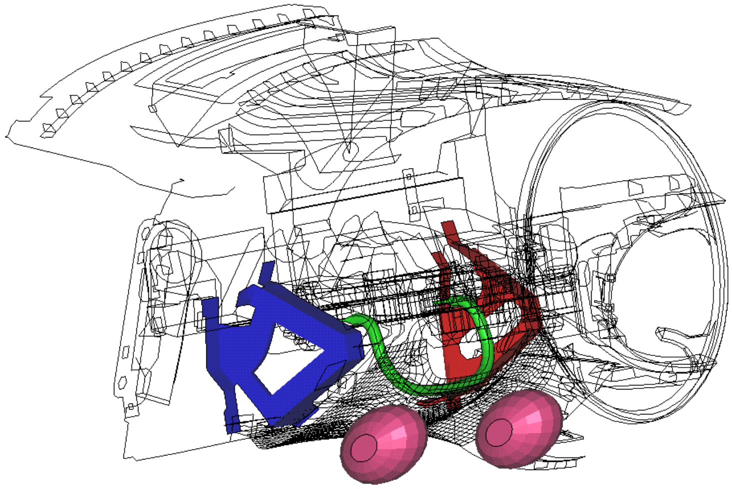


Figure 1: Automotive instrument panel with knee bolster system used for knee-impact analysis. (Courtesy: Ford Motor Company)

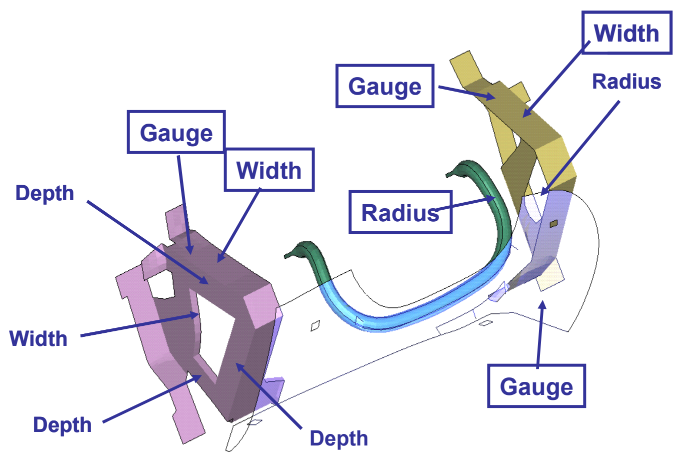


Figure 2: Design variables of the knee bolster system.

Table 1: Design variables used to simulate knee-impact of automotive instrument panel structure.

Name	Lower	Baseline	Upper
L-Bracket gauge	0.7	1.1	3.0
T-Flange depth	20.0	28.3	50.0
F-Flange depth	20.0	27.5	50.0
B-Flange depth	15.0	22.3	50.0
I-Flange width	5.0	7.0	25.0
L-Flange width	20.0	32.0	50.0
R-Bracket gauge	0.7	1.1	3.0
R-Flange width	20.0	32.0	50.0
R-Bracket radius	10.0	15.0	25.0
Bolster gauge	1.0	3.5	6.0
Yolk radius	2.0	4.0	8.0

The design optimization problem accounting for the optimal occupant kinematics was formulated as follows:

Minimize:

Maximum knee force,
Average knee displacement or intrusion,
Mass

Subject to:

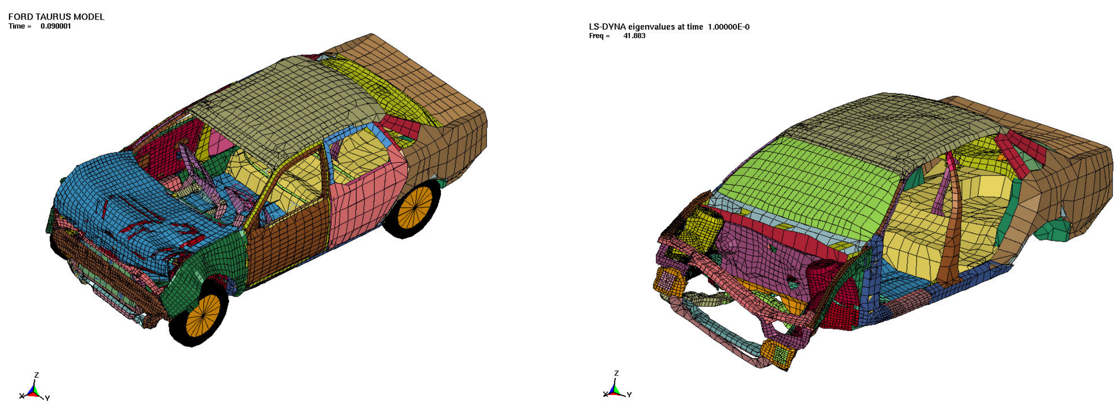
Table 2: Design constraints for the Knee-impact analysis problem.

	Upper bound
Kinetic energy	154000
Yoke displacement	85

All responses were scaled. Knee forces were the peak SAE filtered (60 Hz) forces whereas all the displacements were represented by the maximum intrusion. LS-DYNA [5] was used to simulate different designs. Each simulation requires approximately 20 minutes on a dual-core Intel Xeon (2.66 GHz) processor with 4 GB memory.

2. Multi-disciplinary optimization (MDO)

The performance of a National Highway Transportation and Safety Association (NHTSA) vehicle was optimized for crashworthiness. The full frontal impact crash was simulated using a finite element model containing approximately 30,000 elements, shown in Figure 3(A). A modal analysis was performed on a so-called 'body-in-white' model that had approximately 18,000 elements. The vibration model was depicted in Figure 3(B) in the first torsional vibration mode.



A) Crash Model B) Vibration model

Figure 3: Finite element models for the multi-disciplinary optimization problem.

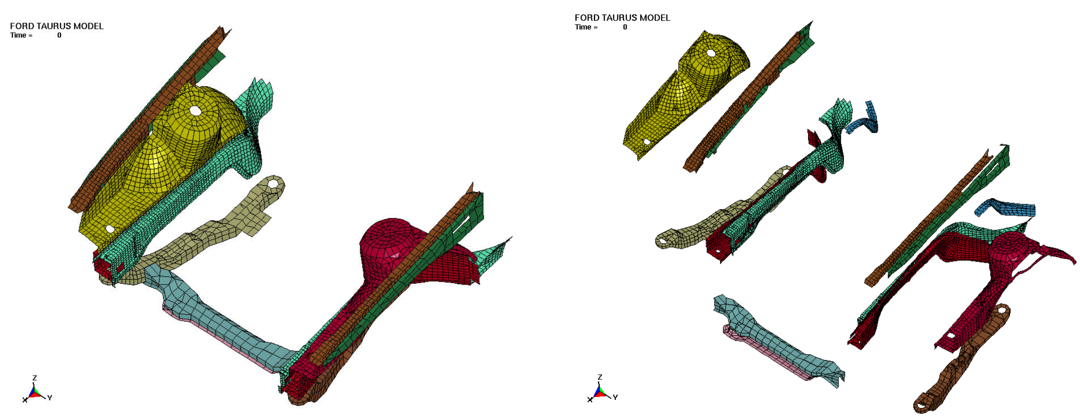


Figure 4: Thickness design variables (with exploded view).

The tracking methodology applied to the torsional mode was described in the paper by Craig et al. [42]. The design variables represented gauges of the structural components in the engine compartment of the vehicle (Figure 4). Seven gauge variables namely apron, rail-inner, rail-outer, shotgun-inner, shotgun-outer, cradle rail and cradle cross member were selected to optimize the performance. Twelve parts comprising aprons, rails, shotguns, cradle rails and the cradle cross member were affected by selected design variables. LS-DYNA [5] was used for both the crash and modal analysis simulations, in explicit and implicit analysis modes respectively.

The vehicle performance is characterized by the intrusion, stage pulses, mass, and torsional frequency. A multi-disciplinary, multi-objective optimization problem is formulated as follows:

Minimize:
Mass,
Intrusion

Subject to:

Table 3: Design constraints for the MDO problem.

	Lower bound	Upper bound
Maximum intrusion ($\boldsymbol{x}_{\text{crash}}$)	-	551.27mm
Stage 1 pulse($\boldsymbol{x}_{\text{crash}}$)	14.512 g	-
Stage 2 pulse($\boldsymbol{x}_{\text{crash}}$)	17.586 g	-
Stage 3 pulse($\boldsymbol{x}_{\text{crash}}$)	20.745 g	-
Torsional mode frequency($\boldsymbol{x}_{\text{NVH}}$)	38.27Hz	39.27Hz

Table 4: Starting values and bounds on the design variables of the MDO problem.

Name	Lower	Baseline	Upper
Rail inner	1.0	2.0	3.0
Rail outer	1.0	1.5	3.0
Cradle rails	1.0	1.93	3.0
Aprons	1.0	1.3	2.5
Shotgun inner	1.0	1.3	2.5
Shotgun outer	1.0	1.3	2.5
Cradle cross member	1.0	1.93	3.0

The bounds and the baseline design thickness values on the design variables were specified in Table 4. The three stage pulses were calculated from the SAE filtered (60Hz) acceleration $\ddot{x}$ and displacement x of a left rear sill node in the following fashion:

$$\text{Stage } j \text{ pulse} = -\frac{k}{(d_2 - d_1)}\int_{d_1}^{d_2} \ddot{x}\mathrm{d}x; k = 0.5 \text{ for } j = 1, k = 1.0 \text{ otherwise;}$$

with the limits $(d_1;d_2)$ = (0;184); (184;334); (334;Max(x)) for i = 1,2,3 respectively, all displacement units were given in *mm* and the minus sign converted acceleration to deceleration.

In summary, the optimization problem aims to minimize the mass and intrusion without compromising the crash and vibration criteria. The objectives and constraints were scaled using the target values to balance the violations of the different constraints. Each crash simulation took about 3,400 to 3,900 seconds on one core of the IBM x3455 machine and the modal analysis took about 40 seconds.

II. Test Setup and Performance Metrics

A. Setup

All problems were simulated using the LS-OPT® [43] optimization program. For all test problems, the NSGA-IIa algorithm with the binary tournament selection operator, SBX crossover and polynomial mutation operators in real-coded space was used. The parameters used for GA simulations are given in Table 5. Since the function evaluation expense of the analytical problems was negligible, a single processor was used for simulations. On the other hand, the computational cost of analyzing crashworthiness problems was quite high so two different clusters of processors were used.

Table 5: NSGA-II parameters for different problems.

	Pop. Size	Max. Gen	P_{cross}	P_{mut}	DI_{cross}	DI_{mut}
TNK	100	200	1.00	0.50	10	100
OSY	100	300	1.00	0.17	10	100
KNEE	80	120	1.00	0.10	20	20
MDO	80	100	0.99	0.15	5	5

The analyses for the MDO problem were run on the x3455 cluster of the IBM Poughkeepsie Benchmark Center. Each node of this cluster had two Dual-Core AMD Opteron(TM) 2220 SE processors with the clock rate of 2.8 GHz. This cluster had a total of 40 nodes (160 cores). The job distribution was handled by the queuing system Tivoli Workload Scheduler LoadLeveler®. Since the chosen population size was 80 in this experiment, LS-OPT submitted 80 crash jobs and 80 NVH jobs to LoadLeveler at each generation, and LoadLeveler arranged all jobs to optimize the use of the computing resource. The smaller of the two crashworthiness problems, the Knee-bolster design problem was simulated on the in-house cluster of 16 nodes. Each node had a single Dual-Core Intel Xeon processor with a clock rate of 3.6 GHz (32 cores). A shared memory of 16GB was available for all processors. The Sun Grid Engine® v 6.1 queuing system was used for job scheduling.

B. Performance Metrics

The multi-objective optimization results in a set of candidate optimal solutions that span the function space. Unlike single objective optimization, where a simple convergence criterion is sufficient to assess the performance, the multi-objective optimization results are measured using two criteria, convergence to the global Pareto optimal front (POF) and the diversity of the solutions [2]. Special metrics used to quantify convergence and diversity of the solution sets are described as follows:

a) **Convergence to Pareto optimal front**: To compare the sets of non-dominated solutions[3] [2] from various generations, the number of solutions that are dominated in each set is computed using a weak non-domination criterion. The smaller the number of dominated solutions, the better is the convergence.

b) **Diversity**: The diversity of the Pareto front estimates how wide-spread the trade-off solutions are, (spread) and how uniformly the entire trade-off front is sampled (uniformity). These two measures of diversity are estimated by the following metrics:

   i. **Spread**: The spread of the front is calculated as the diagonal of the largest hypercube in the function space that encompasses all points. A large spread is desired to find diverse trade-off solutions. It is important to note that the spread measure is derived only based on the extreme solutions such that this criterion is susceptible to the presence of a few isolated points that can artificially improve the spread metric. An equivalent criterion might be the volume of such a hypercube.

   ii. **Uniformity measure**: The uniformity measure is a complimentary criterion (to the spread metric) that detects the presence of poorly distributed solutions by estimating how uniformly the points are distributed in the Pareto optimal set. The uniformity measure [2] is defined as,

$$\Delta = \sum_{i=1}^{N} \frac{\left|d_i - \bar{d}\right|}{N}; \bar{d} = \frac{1}{N}\sum_{i=1}^{N} d_i . \tag{1}$$

   where d_i is the crowding distance[4] [2] of the solution in the function or variable space. The boundary points are assigned a crowding distance of twice the distance to the nearest neighbor.

[3] A non-dominated solution is not dominated by any other solution in the set. A solution is considered to be dominated if it is not better than comparing solution in all objectives and is worse in at least one objective.

This measure is similar to the standard deviation of the crowding distance and hence a small value of the uniformity measure is desired to achieve a good distribution of points.

C. Stopping Criterion

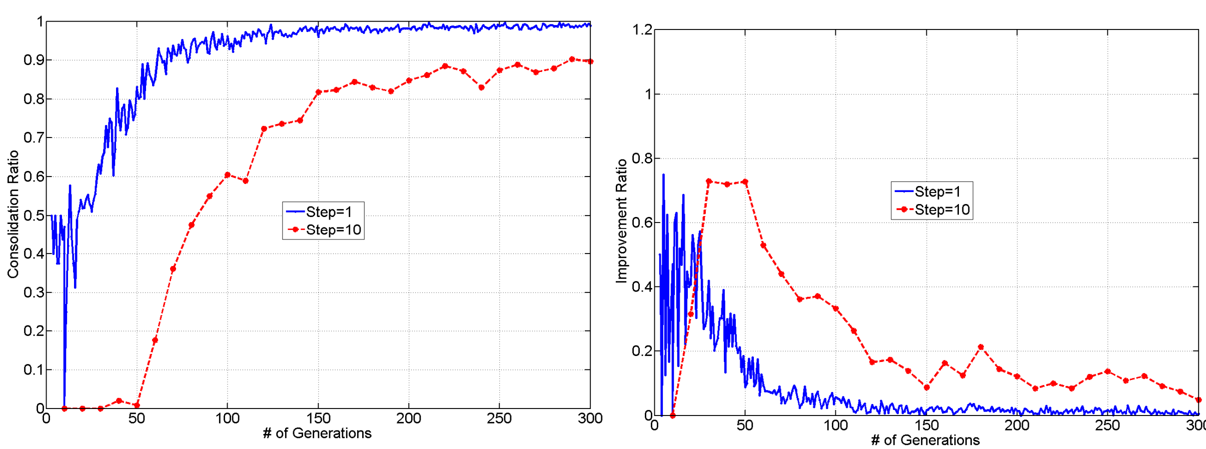
A novel dominance based criterion is proposed in this paper. In the proposed approach, an external archive of non-dominated solutions is maintained. This archive is updated after each generation: i) to remove the solutions that are dominated by the new evolved solutions; and ii) to add the new solutions that are also non-dominated with respect to the archive. Duplicate copies of the non-dominated solutions are also eliminated.

Now to estimate the stopping criterion, the archive at a generation is compared with an older archive (obtained at some earlier generation). The number of solutions in the old archive that were dominated by the newer archive (dominated solutions) and the number of older archive members that are also present in the new archive (non-dominated solutions), were computed. Note that the new archive included all older archive members so none of the new archive members would be dominated. The number of dominated solutions and non-dominated solutions are scaled by the size of the archive to facilitate a fair comparison among archives at different generations while using these metrics to determine the stopping criterion. The scaled number of dominated solutions is termed *improvement-ratio* that indicates the improvement in the quality of solutions; and the scaled number of non-dominated solutions is called the *consolidation-ratio* that represents the proportion of potentially converged solutions.

It is noted that the improvement-ratio is similar to the *error ratio* metric, proposed by van Veldhuizen [35], which calculated the percentage of the non-dominated solution set that was NOT part of the global POF. The main difference is that the error ratio required *a priori* knowledge of the global POF where as the improvement ratio does not require such information. Thus the information ratio metric can be universally used for both analytical and engineering problems.

A sample calculation of the two metrics with the generation number for the OSY problem was shown in Figure 5. It was obvious from Figure 5, that the archive population stabilized with evolution as convergence was approached; and the number of dominated solution (improvement ratio) approached zero whereas the sizes of new and old archives (consolidation ratio) became comparable i.e., the consolidation ratio approached unity. The improvement ratio and the consolidation ratio did not reach exact zero or unit values respectively due to numerical precision limitations. Nevertheless, a decreasing improvement-ratio and/or increasing consolidation-ratio indicated the convergence. Of the two metrics, one can clearly see that the improvement ratio was noisier than the consolidation ratio; and hence the latter was chosen as a more robust convergence metric.

It was also observed from Figure 5, that both metrics are very strongly influenced by the locally random behavior of evolution when the new and old archive are separated by one generation (step = 1). The affect of noise is filtered by increasing the generation gap between the new and old archives to ten (step = 10) i.e., the convergence metrics were computed every ten generations.



A) Consolidation ratio

B) Improvement ratio

Figure 5: Consolidation ratio and improvement ratio metrics for the OSY problem.

[4] Crowding distance is defined as half the perimeter of the largest hypercube around a point that does not encompass any other solution.

It is proposed that the consolidation ratio metrics can be used as a stopping criterion to terminate the MOEA search once significant improvements in the quality of solutions are achieved to reduce the computational cost. The proposed criteria can be implemented with any multi-objective evolutionary algorithm in the following two ways:

i) Firstly, one can stop the simulation when the consolidation ratio reaches a pre-determined cut-off value e.g., when the consolidation-ratio is greater than 0.66. Of course, the choice of the exact cut-off value may depend on the problem and may require some trial and error.

ii) The second method is particularly attractive in industrial setups where each analysis takes much longer (hours) than the data-processing using genetic algorithms (seconds). Besides, the use of computer clusters is regulated to restrict a single user taking all the processing capability for long periods. In this strategy, the user could initialize the GA simulation with a reasonable number of generations (e.g., 40 generations) and track the consolidation ratio to identify the formation of a "knee" in the convergence curve. The onset of the "knee" indicates smaller improvements in the quality of solutions with evolution. Of course this does not guarantee the Pareto optimal front. If the convergence is not achieved after the initial run, the user can iteratively restart the MOEA simulation by adding a few generations (e.g., 10 generations) at a time. The number of restarts would be controlled by the desired convergence. This iterative strategy would result in the least computational cost to achieve a good local Pareto optimal front. This strategy was adopted here.

It is important to note differences in the proposed convergence metric and those available in literature. The main differences are:

i. While all other criteria [29]-[34] have used population data from each generation, the proposed criterion maintains an external archive[5] of non-dominated solutions at each generation that avoids the issue of *Pareto-drift* [16]. The archive also significantly reduces the noise in the non-dominated solution data and provides robust results.

ii. The proposed criterion uses a non-dominance based criterion where as other researchers have used either diversity based criteria ([29], [31], [32]), or dominance based criteria ([30], [34]). The non-domination based criterion is less noisy compared to the dominance based criterion.

iii. Other researchers have computed the so-called "online stopping criteria" at each generation but the proposed criterion is estimated every few generations to reduce the possibility of convergence to a local POF and reduce the effect of noise.

III. Results

In this section, the optimization results for the four test problems are described in detail.

A. Time Savings from Parallelization

The two crashworthiness optimization problems were analyzed using clusters of processors to reduce the clock time. For the MDO problem, a total of 7956 distinct designs were analyzed using LS-DYNA and 3823 designs were found feasible. The total number of designs was not 8000 (population 80, 100 generations) because simulations for duplicate designs were reused. The savings by distributing function evaluations on a cluster using the IBM cluster was about 7,800 hours. That is, by running 80 simulations on 80 cores of the cluster concurrently, the elapsed simulation time was reduced from about 7950 hours on a single core to 100 hours. Similarly, for the knee-bolster design problem, 9599 designs were simulated and nearly 8817 designs were found feasible. In this case, only one design was found to be an exact duplicate and hence was not simulated. The total elapsed time, by running 16 jobs in parallel and allocating two cores to each job, was 130 hours compared to 3200 hours on a single processor. Clearly, there were significant savings in the clock time by running the jobs in parallel.

B. Convergence to Pareto Optimal Front

As discussed earlier, the convergence of the GA simulation was tracked by the growth of the archive in Figure 6. The size of the archive usually increased with generations. Only for the MDO problem, the archive size decreased intermittently as significantly improved solutions evolved that dominated many solutions from the archives at previous generations. This phenomenon was more likely to appear when the size of the archive was smaller than the population size. The number of solutions of the archive that remained non-dominated also increased monotonically for all problems indicating the progress towards the global Pareto optimal front. The number of solutions newly added to the archive was expectedly high in the initial phase, and it stabilized with the increase in generations.

[5] One can also use the non-dominated solutions at any generation instead of the archive.

Similarly it was observed that the number of old archive member solutions that were dominated by the new individuals reduced with evolution. These results indicate the stabilization of the population in terms of convergence.

To further quantify the convergence to the global POF, the stabilization ratio and the improvement ratio metrics proposed here were plotted in Figure 7. The consolidation-ratio, indicative of surviving old archive members, increased sharply in the initial phase indicating significant improvements in the convergence characteristics of the archive. As the evolution was carried out much longer, the consolidation-ratio stabilized to a very high proportion (approx 90%). This “knee-shaped” convergence curve for multi-objective optimization is very similar to that observed for single-objective optimization using genetic algorithms. The GA simulation could have been stopped soon after the formation of the knee in the consolidation-ratio curve to achieve significant cost reduction.

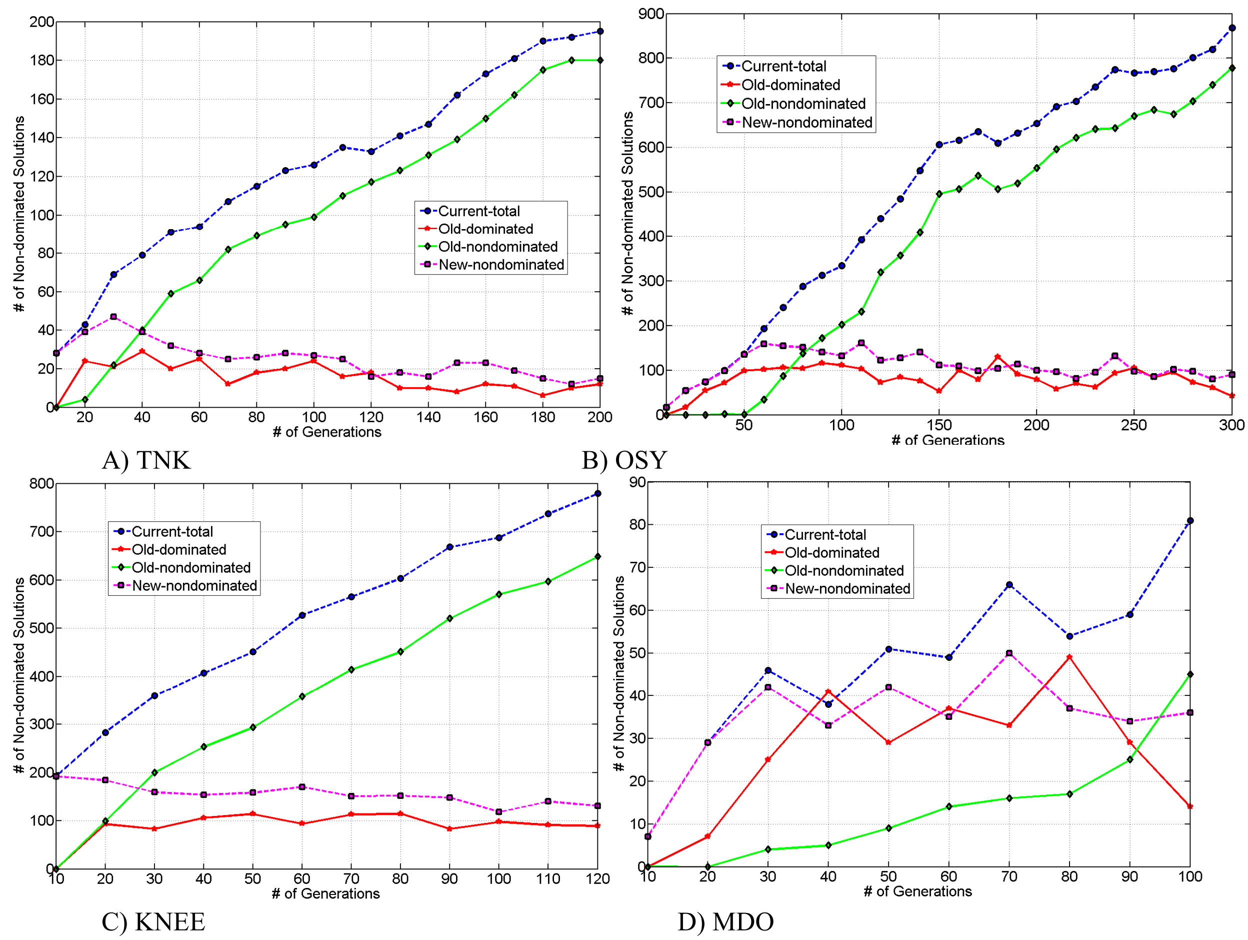
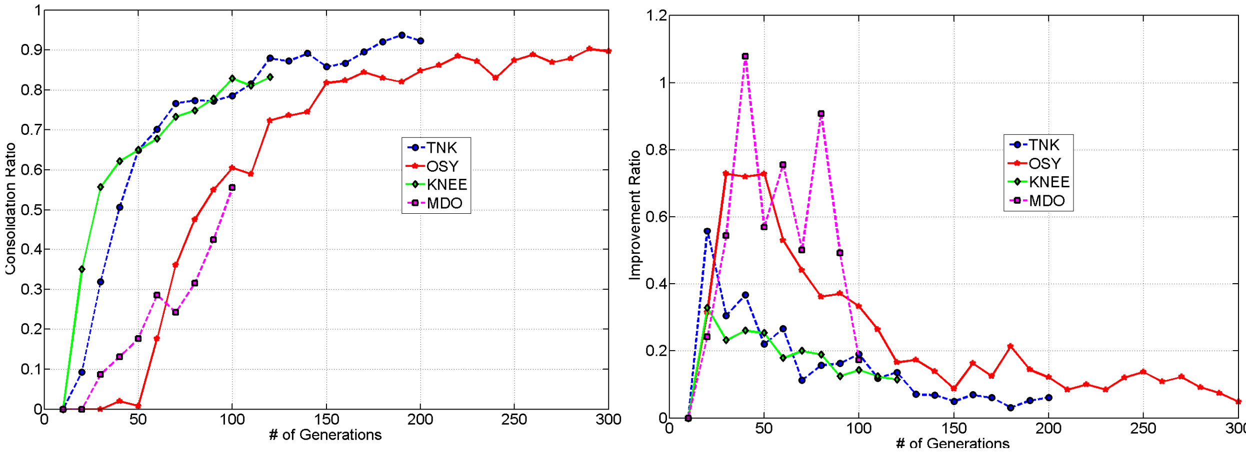


Figure 6: Growth of the archive of non-dominated solutions. Current-total indicates the total number of solutions in this generation, Old-dominated are the number of solutions that were part of the archive previously but were removed from the archive because they were dominated by the new individuals, Old-non-dominated solutions were members of the archive that remained non-dominated with respect to the new solutions, and New-non-dominated solutions are newly found non-dominated solutions.

The improvement-ratio, that is the ratio of the number of previous archive members dominated by the new members (and hence discarded) and the total number of current archive members, was shown in Figure 7(B). As expected, this ratio was significantly high in the earlier generations when evolution resulted in better solutions. The ratio decreased with generations and showed signs of stabilization. The result for the MDO problem was noisier than the other problems but the trends supported the conclusion that the proportion of duplicate solutions decreased with evolution.

The results shown in this section clearly demonstrated that, a) similar to the single-objective evolutionary algorithms, significant improvements towards convergence were obtained in the early phase of multi-objective evolutionary algorithms, and b) either of the two metrics, the consolidation-ratio or the improvement-ratio could be used to decide on the convergence of the multi-objective evolutionary algorithm. It was noted that early termination of the simulation might not result in the global Pareto optimal front but the solution set obtained would likely be a good trade-off between the computational cost and the convergence to the global POF.



A) Consolidation-ratio B) Improvement ratio

Figure 7: The two convergence metrics proposed in this paper. The consolidation-ratio is the proportion of old solutions that have remained non-dominated in the current generation and the improvement-ratio is the proportion of old solutions that were dominated by new solutions.

C. Spread

The diversity of solutions was quantified using the spread of the archive at each generation. The objectives for the OSY problem were scaled by 250 and 50 units respectively, and the MDO problem were scaled by 0.1 units each to provide equivalence among all problems. The scaling only affected the scale of the curve. It was observed from Figure 8 that the spread of the candidate Pareto optimal front usually increased in the early stage of evolution and then stabilized. This metric also indicated the potential advantages of stopping the search earlier.

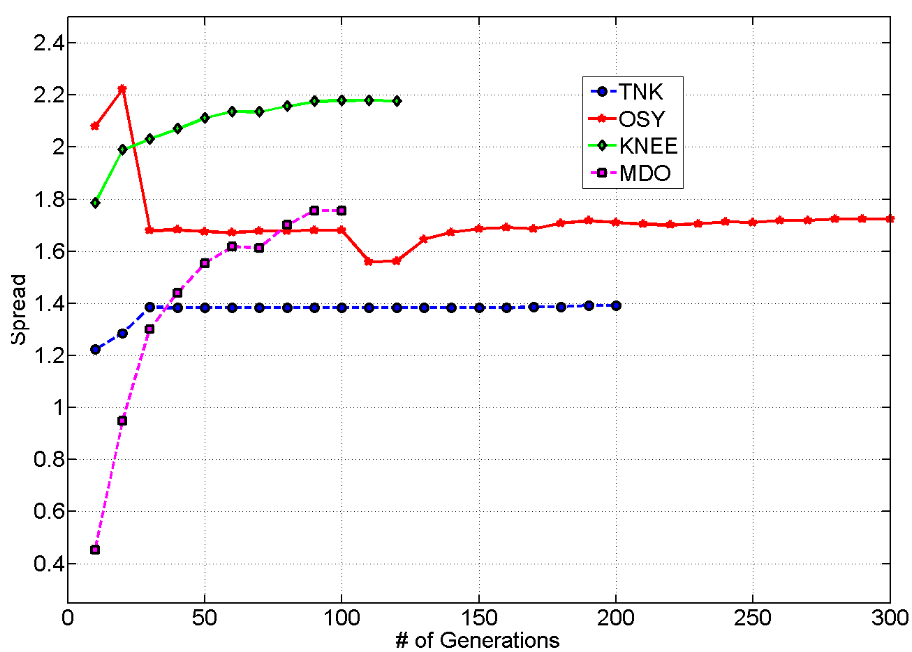
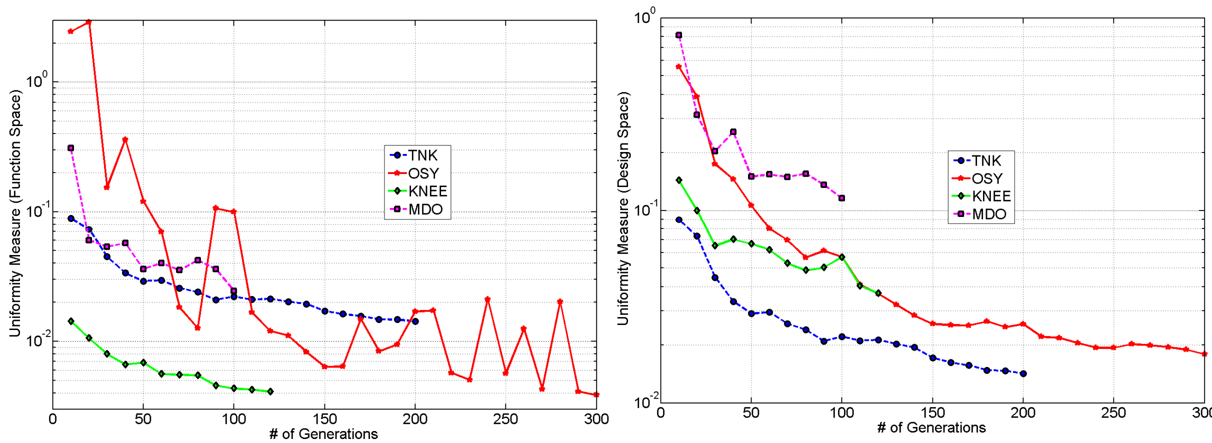


Figure 8: Spread of the candidate Pareto optimal front for different problems.

Apparently opposite trends for the OSY problem were largely due to the presence of a widely-spread local Pareto front (discussed later). The solutions in this non-dominated set were dominated by better solutions in the subsequent generations such that the spread decreased. It was also noted that the optimization search could not find complete Pareto front for the OSY problem in 300 generations. A gradual increase in the spread for this problem was observed because more solutions were being continuously evolved to traverse the Pareto optimal front. This behavior might not be atypical for many problems and indicated the potential lack of convergence that could be attributed to the early termination. However, the result obviously supported the hypothesis of significant savings in the computational cost by accepting a small compromise in the quality of solutions for this problem also.

D. Uniformity Measures

As shown in Figure 9, the uniformity measures in both function and variable spaces decreased with generations that indicated improvement in the distribution of solutions. The low value of the uniformity measure also confirmed the good distribution. The noisy behavior of the uniformity measure in the function space for the OSY problem was mainly due to the fact that the extreme solutions in the different regions of the Pareto optimal front were identified before the entire region was populated with the optimal solutions. Overall, the trends for the OSY function also agreed with the other problems.



A) Function space B) Variable space

Figure 9 Uniformity measures in function and variable space for all problems.

E. Validating the Stopping Criterion

The results shown above indicated that the GA simulations could be stopped earlier using the proposed criteria to achieve a substantial computational cost reduction. Using the convergence criteria (Figure 7) and accounting for the spread (Figure 8) and uniformity measures (Figure 9), the simulations for the TNK could be stopped after 80 generations, after 140 generations for the OSY problem and after 100 generations for crashworthiness optimization problems[6]. A comparison of the quality of the solutions in the archive at the proposed stopping generation and complete simulation results for all problems in shown in Figure 10.

As indicated by different metrics, there were significant improvements in the quality of trade-off solutions in the initial phase (notice the differences in trade-off curves obtained at 10^{th} and 40^{th} generations) for all the problems. Many solutions in the archive after 40 generations constituted a part of the global Pareto optimal front for the analytical problems. The entire POF was traversed after 80 generations for the TNK problem (Fig 7(A)) and more points were added to the Pareto front in the subsequent generations (Fig 7(B)).

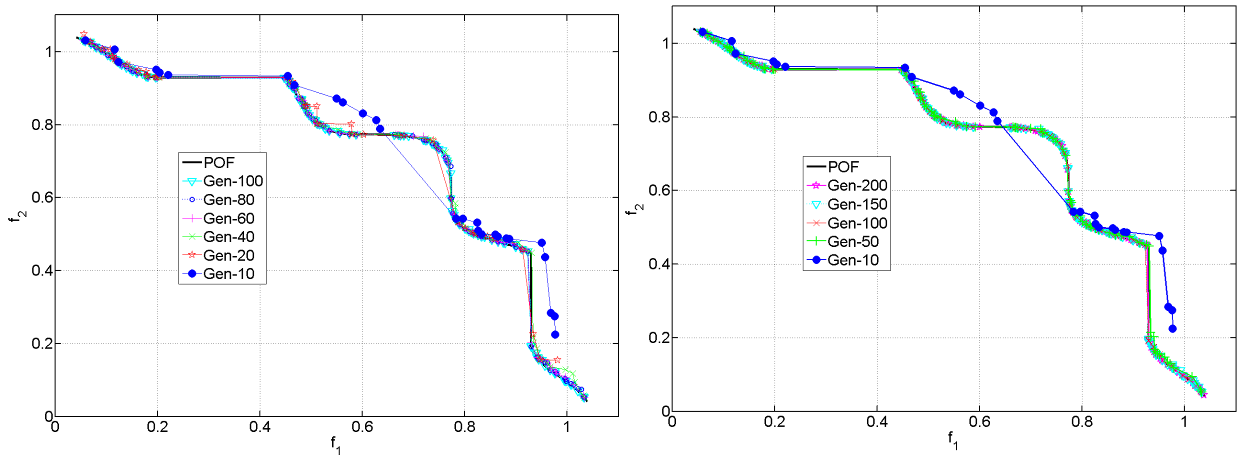
As described earlier, the OSY problem resulted in five distinct sections of the Pareto optimal front. The local Pareto optimal front obtained after the termination condition at 140 generations also comprised of five sections. However, only two sections (f_1 ranging from -260 to -130 units) were well identified after 140 generations for the OSY problem; the third section (f_1 ranging from -110 to -40 units) was sparsely approximated. The remaining two sections of the local Pareto optimal front (f_2 ranging from 4 to 28 units, and 50 to 80 units) were far from the global

[6] The result for the MDO problem was not very conclusive but more simulations could not be conducted due to the budget on the computational resource.

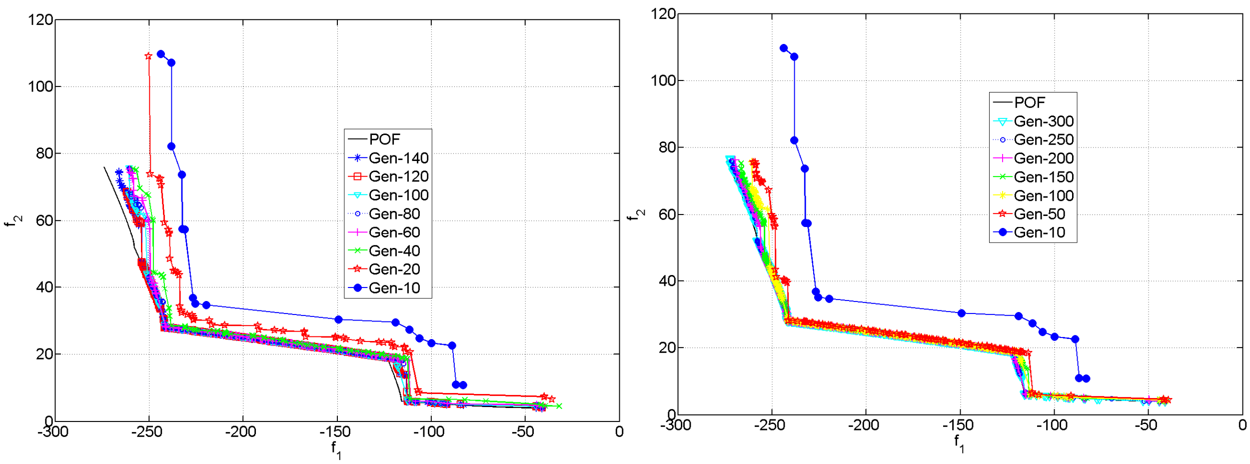
POF. The spread of the solutions in the third section was improved and convergence to the global POF was achieved by running the simulations up to 300 generations.

The pictorial representation of archive solutions at different generations for the knee-bolster design problem is depicted in Figure 7(C). While it is difficult to assess the convergence pictorially for this three objectives problem, initial improvements were obvious. There was a very noticeable overlap in the solutions from the 100^{th} and 120^{th} generations confirming the attainment of convergence to a local Pareto optimal front. The diversity of the solutions also appeared reasonably good and the same was noticed in the metrics discussed above.

As can be seen from Figure 7(D) for the MDO problem, there were noticeable improvements in the convergence of the solutions in the first 80 generations after which relatively smaller improvements were found. The improvements between trade-off solutions obtained between the 90^{th} and 100^{th} generations were relatively small[7]. Apparently, the trade-off solutions at the 100^{th} generation were better distributed than the solutions obtained at the 90^{th} generation.

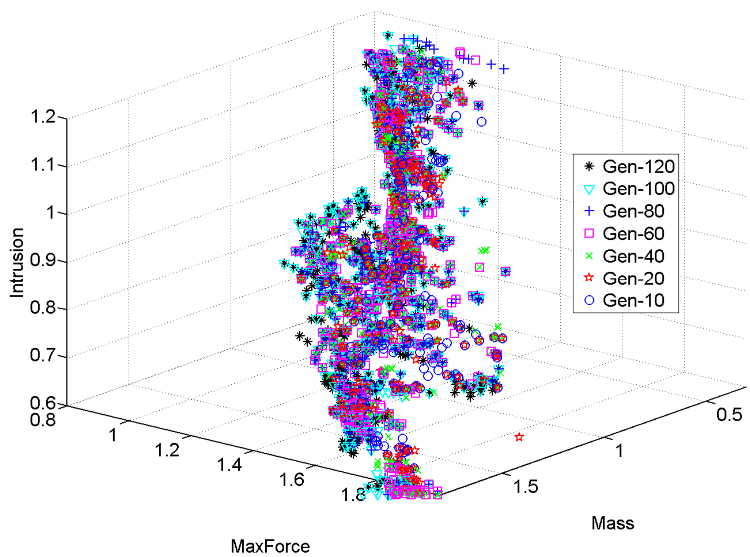


A) TNK

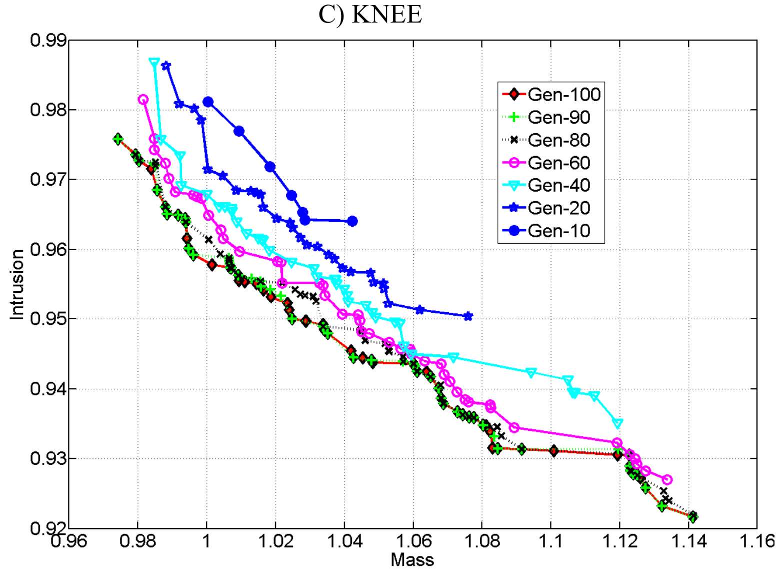


B) OSY

[7] The limitations on the computational budget restricted us from running more than 100 generations for the MDO problem.



C) KNEE



D) MDO

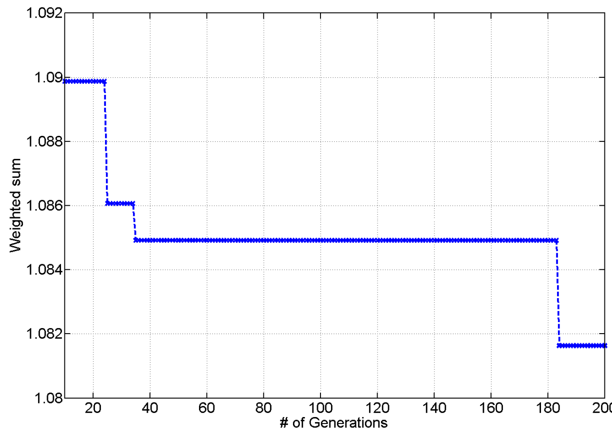
Figure 10: Non-dominated solutions at different generations. Gen-10 corresponds to 10th generation.

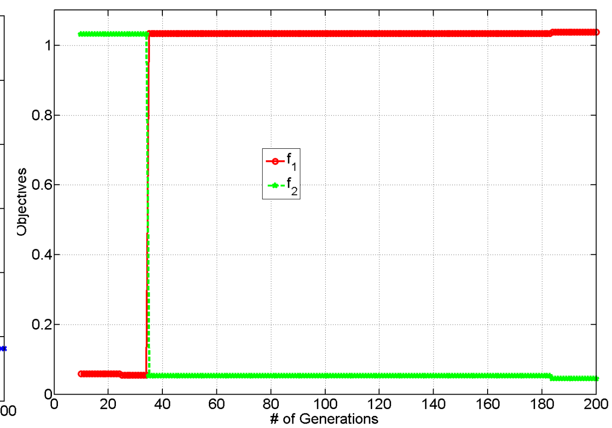
All the results presented here confirmed that i) significant improvements were obtained in the initial phase of evolution, and ii) the suggested stopping point based on the metrics proposed in this paper resulted in good trade-offs between the quality of solutions and the computational cost.

F. Selecting a Single Design

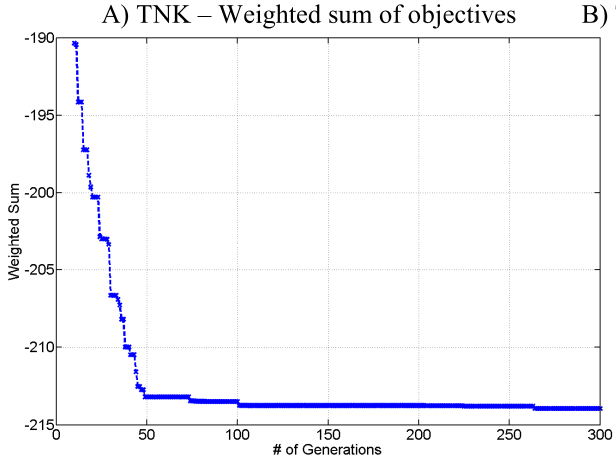
While it is useful to identify many choices to understand the nature of the problem and different possible trade-offs, a single solution is mostly selected as the final design. In one such scenario considered here, a design corresponding to equal importance of each objective was identified using a weighted sum of the objectives for each problem.

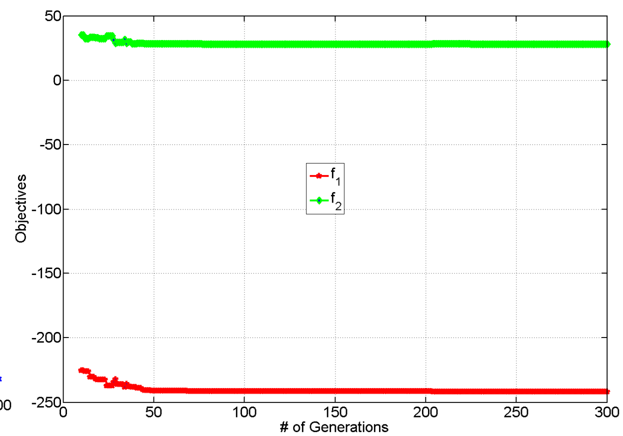
Figure 11 shows the weighted sum of the objectives and corresponding objective function values with generations. A gradual stepped reduction, typical of a GA based optimization, in the weighted sum of the objectives was observed for all test problems. Figure 11 indicated that the best design (according to the unit weight criterion) at any generation was not obtained by monotonous improvements in individual objectives. Instead, a trade-off between two objectives was reflected such that an occasional increase in one objective was coupled with a larger reduction in other objectives. The extreme behavior of the TNK problem was attributed to the non-convex nature of the problem where more than one single-objective optimum could be located for the same weight structure. The change in the objective function for the OSY problem was largely governed by the first objective that had a higher magnitude. For the knee problem, the intrusion increased slightly to accommodate a higher reduction in the maximum force. Similarly, there was an increase in the scaled intrusion to allow a higher mass reduction for the MDO problem. For all problems, significant improvements in the performance were obtained before the stopping generation indicated by the convergence metrics described in this paper.



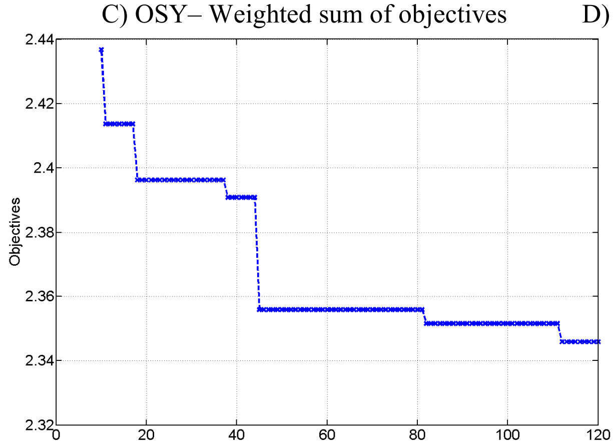


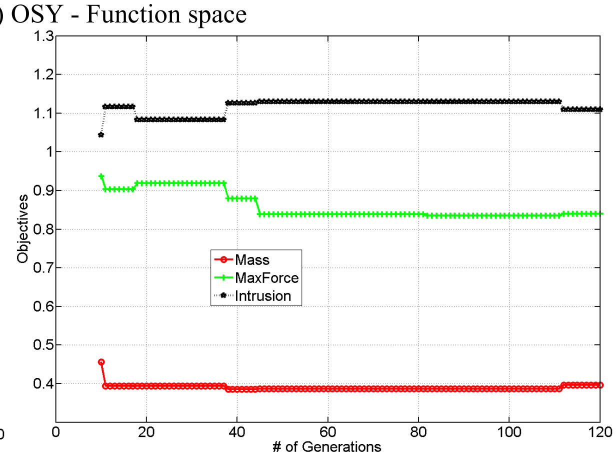
A) TNK – Weighted sum of objectives B) TNK – Function space



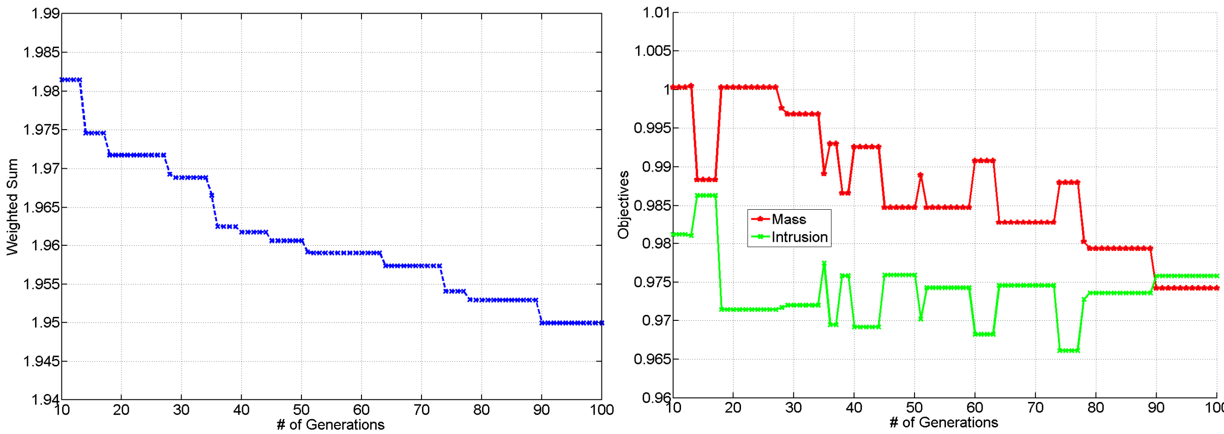


C) OSY– Weighted sum of objectives D) OSY - Function space





E) KNEE – Weighted sum of objectives F) KNEE - Function space



G) MDO – Weighted sum of objectives H) MDO - Function space

Figure 11: Generation number vs. minimum equally weighted sum of objectives.

IV. Summary

The issue of computational cost is very important for the application of genetic algorithms to industrial problems, particularly for the multi-objective optimization problems. The convergence characteristics of evolutionary algorithms for multi-objective optimization were studied in this paper and a simple and robust stopping criterion was developed. Specifically, the study was focused on a popular MOEA, the elitist non-dominated sorting genetic algorithm (NSGA-II).

The rate of convergence was assessed using a few metrics, proposed in this paper, which tracked the quality of solutions in an archive of non-dominated solutions set at each generation. The *consolidation-ratio* indicative of the proportion of the solutions in the archive that remained non-dominated with respect to new solutions was identified as a very reasonable metric to evaluate the relative improvements. Another metric *improvement-ratio* was the proportion of non-dominated solutions from previous generations that were dominated by the newly generated solutions. The diversity of the solutions was measured using a spread metric that estimated the length of the diagonal of the largest hyper-cube and uniformity measures that calculated the average deviation of the distances between the closest points. These metrics did not require *a priori* information about the global Pareto front and no additional parameter was required.

The performance of the different metrics was studied using two analytical problems and two crashworthiness problems. It was demonstrated that, similar to the single-objective GA, significant improvements in the convergence and the diversity of the solutions were obtained in the early phase of evolution. Relatively smaller improvements were observed in the later phases. The convergence curve obtained by plotting the consolidation-ratio against the number of generations resulted in the most prominent "knee" shape. This suggested that the search could be terminated at the on-set of "knee" formation to reduce the computational cost for industrial problems where the goal was often to find a significantly improved solution rather than the absolute best optimum.

A further analysis of the non-dominated solutions at different generations also indicated that the above metric resulted in a reasonable trade-off between the computational cost and the quality of the solutions. It was noted that early termination did not necessarily result in convergence to the global Pareto optimal front but the savings in the computational cost offered a very attractive trade-off for practical problems.

The analysis of the single solution chosen by minimizing the equally weighted sum of objectives also indicated a monotonic decrease in the weighted sum of objectives but individual objectives were selected to provide the best trade-offs. Lastly, a significant reduction in clock time was also realized by utilizing a parallel computing framework.

V. Acknowledgements

The authors would like to thank the IBM Poughkeepsie Benchmark Center for providing computing equipment and for their strong technical support. Dr Guangye Li from the IBM Deep Blue Computing Group at Houston, TX provided a lot of help with the simulations for the MDO problem.

VI. References

[1] Miettenen K, Non-linear Multi-objective Optimization, Kluwer, 1999.
[2] Deb K, Multi-Objective Optimization Using Evolutionary Algorithms, Wiley, 2001, New York.
[3] Rao SS, Engineering Optimization: Theory and Practice, Wiley, 1996, New York.
[4] Goel T, Stander N, Lin Y-Y, "Efficient Resource Allocation for Genetic Algorithm Based Multi-Objective Optimization With 1000 Simulations", Structural and Multi-disciplinary Optimization, 2009, doi: 10.1007/s00158-009-0426-9.
[5] Hallquist, J.O. LS-DYNA® User's Manual, Version 971.
[6] Venkataraman S, Haftka RT, "Structural Optimization: What has Moore's Law Done for US?", Structural and Multi-disciplinary Optimization, 28 (6): 375-387, 2004.
[7] Coello CAC, Pulido GT, "Multiobjective Optimization Using a Micro Genetic Algorithm", In Proceedings of the Genetic and Evolutionary Computation Conference (GECCO-2001), San Francisco CA, 2001.
[8] Knowles J, "ParEGO: A Hybrid Algorithm with On-Line Landscape Approximation for Expensive Multi-Objective Problems", IEEE Transactions on Evolutionary Computation, 10(1): 50-66, 2006.
[9] Jensen MT, "Reducing the Run-time Complexity of Multiobjective EAs: The NSGA-II and other algorithms", IEEE Transactions on Evolutionary Computation, 7(5): 503, 515, 2003.
[10] Isibuchi M, Murata T, "A Multi-objective Genetic Local Search Algorithm and Its Application to Flowshop Scheduling", IEEE Transactions on Systems, Man and Cybernetics—Part C: Applications and Reviews, 28(3): 392-403, 1998.
[11] Deb K, Goel T, "A Hybrid Multi-objective Evolutionary Approach to Engineering Shape Design", In Proceedings of the First International Conference on Evolutionary Multi-Criterion Optimization (EMO'01), Zurich Switzerland, 385–399, March 2001.
[12] Goel T, Deb K, "Hybrid Methods for Multi-objective Evolutionary Algorithms", In proceedings of the 4th Asia-Pacific Conference on Simulated Evolution And Learning (SEAL-01), Singapore, 188–192, Nov 2001.
[13] El-Beltagy MA, Nair PB, Keane AJ, "Metamodeling Techniques for Evolutionary Optimization of Computationally Expensive Problems: Promises and Limitations", In Proceedings of the Genetic and Evolutionary Computation Conference (GECCO-1999), Orlando FL, 1999.
[14] Wilson B, Cappalleri D, Simpson TW, Frecker M, "Efficient Pareto Frontier Approximation using Surrogate Approximations", Optimization and Engineering, 2(1): 31-50, 2001.
[15] Giannakoglou KC, "Design of Optimal Aerodynamic Shapes Using Stochastic Optimization Methods and Computational Intelligence", Progress in Aerospace Sciences, 38(1): 43-76, 2001.
[16] Goel T, Vaidyanathan R, Haftka RT, Shyy W, "Response Surface Approximation of Pareto Optimal Front in Multiobjective Optimization", Computer Methods in Applied Mechanics and Engineering, 196(4-6): 879-893, 2007; also presented at the 10th AIAA/ISSMO Multidisciplinary Analysis and Optimization Conference (MAO, Sep-2004).
[17] Voutchkov I, Keane AJ, "Multi-objective Optimization Using Surrogates", In, Proceedings of the 7th International Conference on Adaptive Computing in Design and Manufacture. Bristol UK, April 2006.
[18] Ray T, Smith W, "A Surrogate Assisted Parallel Multi-objective Evolutionary Algorithm for Robust Engineering Design", Engineering Optimization, 38(8): 997-1011, 2006.
[19] Emmerich MTM, Giannakoglou KC, Naujoks B, "Single- and Multi-objective Evolutionary Optimization Assisted by Gaussian Random Field Metamodels", IEEE Transactions on Evolutionary Computation, 10(4): 421-439, 2006.
[20] Goel T, Dorney DJ, Haftka RT, Shyy W, "Improving the Hydrodynamic Performance of Diffuser Vanes via Shape Optimization", Computers and Fluids, 37(6): 705-723, 2008.
[21] Aytug H, Koehler GJ, "Stopping Criteria for Finite Length Genetic Algorithms", ORSA Journal of Computing, 8(2): 181-193, 1996.
[22] Aytug H, Koehler GJ, "New Stopping Criteria for Genetic Algorithms", European Journal of Operations Research, 126: 662-674, 1996.
[23] Pendharkar P, Koehler GJ, "A General Steady State Distribution based Stopping Criteria for Finite Length Genetic Algorithms", European Journal of Operations Research, 176: 1436-1451, 2007.
[24] van Veldhuizen DA, Lamont GB, "Evolutionary Computation and Convergence to a Pareto Front", In Proceedings of the 3rd Annual Conference on Genetic Programming, San Francisco CA, 221-228, August 1998.
[25] Hanne T, "On the Convergence of Multiobjective Evolutionary Algorithms", European Journal of Operational Research, 117(3): 553-564, 1999.

[26] Rudolph G Agapie A, “Convergence Properties of Some Multi-objective Evolutionary Algorithms”, In Proceedings of the IEEE Congress on Evolutionary Computation (CEC-2000), San Diego CA, 1010-1016, 2000.
[27] Rudolph G, “Self-Adaptive Mutations Lead to Premature Convergence”, IEEE Transactions on Evolutionary Computation, 5(4): 410-414, 2001.
[28] Li G, Goel T, Stander N, “Assessing the Convergence Properties of NSGA-II for Direct Crashworthiness Optimization”, 10th International LS-DYNA® User’s Conference, Detroit MI, Jun 8-10, 2008.
[29] Rudenko O, Schoenauer M, “A Steady Performance Stopping Criterion for Pareto-based Evolutionary Algorithm”, In Proceedings of the 6th International Conference on Multi-objective Programming and Goal Programming, Hammamet Tunisia, 2004.
[30] Marti L, Garcia J, Berlanga A, Molina JM, “A Cumulative Evidential Stopping Criterion for Multi-objective Optimization Evolutionary Algorithms”, In Proceedings of the 9th Genetic And Evolutionary Computation Conference (GECCO-07), London UK, 2835-2842, July 2007.
[31] Trautmann H, Ligges U, Mehnen J, Preuss M, “A Convergence Criterion for Multi-Objective Evolutionary Algorithm Based on Systematic Statistical Testing”, In Proceedings of the 10th Parallel Problem Solving From Nature (PPSN-X), Dortmund, Germany, 825-836, September 2008.
[32] Wagner T, Trautmann H, Naujoks B, “OCD: Online Convergence Detection for Evolutionary Multi-Objective Algorithms Based on Statistical Testing”, In Proceedings of the 5th International Conference on Multi-Criterion Optimization (EMO-09), Nantes France, 198-215, April 2009.
[33] Naujoks B, Trautmann H, “Online Convergence Detection for Multiobjective Aerodynamics Applications”, In Proceedings of the IEEE Congress on Evolutionary Computation (CEC-2009), Trondheim, Norway, May 2009.
[34] Martí, L., García, J., Berlanga, A., & Molina, J. M. (2009). An Approach to Stopping Criteria for Multi--objective Optimization Evolutionary Algorithms: The MGBM Criterion. In 2009 IEEE Conference on Evolutionary Computation (CEC 2009), Trondheim, Norway, May 2009.
[35] Van Veldhuizen DA, Multiobjective Evolutionary Algorithms: Classifications, Analyses, and New Innovations, PhD Thesis, Department of Electrical and Computer Engineering, Air Force Institute of Technology, Wright Patterson Air Force Base, OH, 1999.
[36] Zitzler E, Thiele L, “Multiobjective Evolutionary Algorithms: A Comparative Case Study and the Strength Pareto Approach”, IEEE Transactions on Evolutionary Computation, 3(4): 257-271, 1999.
[37] Deb K, Jain S, “Running Performance Metrics for Evolutionary Multi-Objective Optimization”, In Proceedings of the Fourth Asia-Pacific Conference on Simulated Evolution and Learning (SEAL-02), Singapore, 13-20, 2002.
[38] Deb K, Pratap A, Agarwal S, Meyarivan T, “A Fast and Elitist Multiobjective Genetic Algorithm: NSGA-II”, IEEE Transactions on Evolutionary Computation, 6(2): 182-197, 2002.
[39] Tanaka M, “GA-based Decision Support System for Multi-Criteria Optimization”, In Proceedings of the International Conference on Systems, Man and Cybernetics, 2: 1556-1561, 1995.
[40] Osyczka A, Kundu S, “A New Method to Solve Generalized Multicriteria Optimization Using the Simple Genetic Algorithm”, Structural Optimization, 10(2): 94-99, 1995.
[41] Akkerman A, Thyagarajan R, Stander N, Burger M, Kuhm R, Rajic H, “Shape Optimization for Crashworthiness Design using Response Surfaces”, In Proceedings of the International Workshop on Multidisciplinary Design Optimization, Pretoria South Africa, 270-279, 2000.
[42] Craig KJ, Stander N, Dooge DA, Varadappa S, “Automotive Crashworthiness Design Using Response Surface-Based Variable Screening and Optimization”, Engineering Computations, 22:38-61, 2005.
[43] Stander, N., Roux, W.J., Goel, T., Eggleston T. and Craig, K.J. LS-OPT® Version 4.0 User’s Manual, Livermore Software Technology Corporation, April 2009.