Multi-Speed Gearbox Design Using Multi-Objective Evolutionary Algorithms

Kalyanmoy Deb and Sachin Jain
Kanpur Genetic Algorithms Laboratory (KanGAL)
Department of Mechanical Engineering
Indian Institute of Technology Kanpur
Kanpur, PIN 208 016, India
deb@iitk.ac.in

Abstract

Optimal design of a multi-speed gearbox involves different types of decision variables and objectives. Due to lack of efficient classical optimization techniques, such problems are usually decomposed into tractable subproblems and solved. Moreover, in most cases the explicit mathematical expressions of the problem formulation is exploited to arrive at the optimal solutions. In this paper, we demonstrate the use of a multi-objective evolutionary algorithm, which is capable of solving the original problem involving mixed discrete and real-valued parameters and more than one objectives, and is capable of finding multiple non-dominated solutions in a single simulation run. On a number of instantiations of the gearbox design problem having different complexities, the efficacy of NSGA-II in handling different types of decision variables, constraints, and multiple objectives are demonstrated. A highlight of the suggested procedure is that a post-optimal investigation of the obtained solutions allows a designer to discover important design principles which are otherwise difficult to obtain using other means.

1 Introduction

A multi-speed gearbox design optimization problem introduces a number of challenges [1], particularly if the design involves optimization of both its kinematic structure and gear geometries. The resulting optimization problem involves design variables which can be both integer-valued (for gear teeth), discrete-valued (for gear module), and real-valued (for gear thickness). The feasibility of a design solution depends on the satisfaction of a number of equality and inequality constraints involving strength and gear meshing restrictions. The resulting optimization problem becomes further complicated due to the presence of multiple conflicting objectives, often desired to be optimized in such real-world problems.

While tackling such complex problems using classical optimization methods, researchers usually decompose the overall problem into tractable subproblems [9]. To tackle such problems, a multi-step optimization procedure is usually adopted. In each step, a part of the design problem (for example, involving only integer variables) is first solved by fixing the rest of the variables to reasonable values. Thereafter, another part of the problem (say, involving real-valued decision

variables) is solved by fixing previous variables to the values obtained in the first subproblem. This procedure is continued till all components of the overall problem is solved. For obvious reasons, although such a procedure enables classical optimization techniques to be applied, the final obtained solutions may not necessarily correspond to the true optimal solutions of the overall problem. Moreover, such decomposition technique may not be applicable to problems which cannot be decomposed into meaningful subproblems. Some classical optimization methods also face difficulties in searching a constrained search space having multiple equality and nonlinear inequality constraints. Generic methods such as the penalty function based approaches largely depends on the chosen penalty parameter values.

Classical optimization methods for handling multiple objectives (weighted-sum approaches or goal-programming methods) require the user to supply a weight vector or a preference vector, before solving the problem [2, 7, 11, 12]. Since such a weight or preference vector scalarizes multiple objectives into a single objective, the outcome of the optimization process is usually a single optimal solution. In order to obtain a set of so-called Pareto-optimal solutions for multi-objective optimization, these methods must have to be applied many times with different weight or preference vectors. Moreover, most classical approaches demand that all objectives are of the same type (either all are of minimization type or of maximization type). One common way to convert a maximization problem into a minimization problem is to use the inverse function. This conversion facilitates the use of a classical optimization method, but causes a difficulty in multi-objective optimization. Such conversions do not emphasize the complete range of the transformed objective uniformly. Thus, a number of well-distributed (or trade-off) solutions in the original objective space may be difficult to obtain with a uniformly set of weight vectors used in the converted objective space. In this paper, we propose a multi-objective evolutionary optimization technique, which is a more flexible technique and alleviates most difficulties mentioned above. Particularly, the proposed technique – the modified non-dominated sorting GA or NSGA-II – is capable of handling the following complexities often arises in an optimization problem:

1. NSGA-II is capable of handling discrete (including integer) and real-valued decision variables together without any special consideration,

2. NSGA-II is capable of handling constraints in a simple way and without having to set any new parameter, such as penalty parameters.

3. NSGA-II is capable of handling multiple objectives of mixed types easily.

Here, the NSGA-II is applied increasingly to more complex versions of the gearbox design problem than what has been attempted in the literature [9]. Starting from the two-objective, 10-variable,

real-parameter gearbox optimization problem, NSGA-II shows its ease of application and efficacy to more difficult problems, leading to a four-objective, 29-variable, 101-constraint, mixed-integer gearbox optimization problem. In addition to the optimal solutions, the paper also suggests a procedure for the possibility of revealing important design principles for designing a gearbox. The success of NSGA-II's application in this problem suggests its immediate application to similar or other more complicated and challenging engineering design problems.

2 Multi-Speed Gearbox Design

In the multi-speed gearbox design problem, the purpose is to achieve different speeds in the output shaft, whereas the input shaft rotates at a fixed angular velocity. To facilitate multiple speed reductions, the gearbox contains a number of intermediate shafts. Figure 1 shows a typical 18-speed gearbox having 18 gears, placed in a two-dimensional array. Gears in shafts 1, 3 and 5 can be translated to get meshed with corresponding gears in shafts 2 and 4. With an input shaft

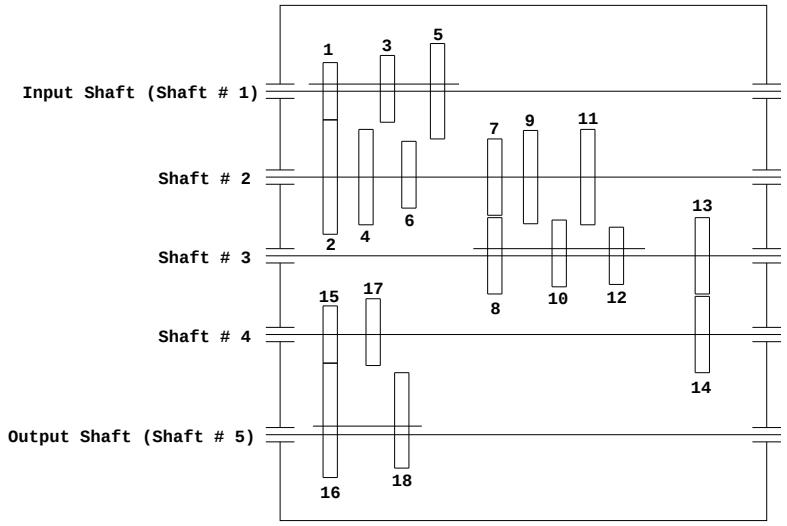


Figure 1: Schematic of a gear train.

speed of 1400 rpm and with a ratio between consecutive output speeds of 1.14, the i-th *desired* output shaft speed would be $1400/(1.14)^{i-1}$. For the above gearbox with 18 options in the output shaft speed, the desired lowest output speed becomes $1400/(1.14)^{17}$ or 150 rpm.

The optimal design of such a gearbox then involves searching for a set of gears (with adequate number of teeth, thickness, and module) which would achieve output shaft speeds closer to the desired values. Moreover, the design should be such that the gear sizes make a feasible plan (such as to maintain identical center distance between all combination of mating gears between two shafts) and gears are strong from bending and wear considerations. Finding a feasible design

satisfying all the above constraints is by itself a difficult task. However, it is often desired to find a solution which is not only feasible but also optimizes an objective of design. In this problem, a number of such objectives are possible: (i) minimization of overall gear material used (which is directly related to the weight and cost of the gearbox), (ii) maximization of power delivered by the gearbox, (iii) minimization of the center distance between input and output shafts, and (iv) minimization of error between achieved and desired output shaft speeds. When multiple *conflicting* objectives are considered (such as (i) and (ii)), there are usually more than one optimal solution to the resulting problem. Such solutions are called Pareto-optimal solutions or efficient solutions or non-inferior solutions [11]. Although only one of the solutions have to be recommended for fabrication, one approach to multi-objective problem solving is to first find a widely distributed set of Pareto-optimal solutions before making any particular choice on the importance of one objective over the other. Thereafter, one solution can be picked using some higher-level decision-making criteria. Here, we would not like to belabor the discussion on what higher-level decision-making criteria will be the best for a gearbox design problem, as such a discussion would be subjective to the user and would depend on many non-technical aspects which are beyond the scope of the present discussion. Instead, we argue that such a decision-making becomes easier and more pragmatic than the usual practice of optimizing a single-objective function derived by prefixing a relative preference factor of each objective. As a by-product of our approach, we shall demonstrate how an analysis of multiple optimal solutions can reveal important design principles. In the following subsections, we outline the formulation of the optimal gearbox design problem in a systematic manner.

2.1 Number of Teeth Fixed

When the number of teeth in each gear is kept fixed, the decision variables are all gear-pair thickness values t_i (for $i = 1, 2, \ldots, G$) and the power delivered p. For the gearbox layout shown in Figure 1, there are nine gear-pairs, or $G = 9$. The module is kept the same for all gear-pairs. Earlier studies [9, 13] have considered two objectives – maximizing the power delivered f_1 and minimizing the total volume of gear material f_2. A little thought will reveal that these two objectives are conflicting to each other. There also exist a number of constraints associated with this problem. Each gear must satisfy the bending stress σ_b and wear stress σ_w constraints: (i) $\sigma_b \leq S_b$ and (ii) $\sigma_w \leq S_w$, where S_b and S_w are the permissible bending and wear stresses, respectively. The bending and wear stresses developed in the i-th gear-pair can be calculated as follows [9]:

$$\sigma_{b_i} = \frac{97500 p k_c k_d (r_i + 1)}{\omega_i a_i t_i m r_i y_i \cos\beta}, \tag{1}$$

$$\sigma_{w_i} = \frac{0.59(r_i+1)}{r_i a_i}\sqrt{\frac{97500 p k_c k_d E(r_i+1)}{\omega_i t_i \sin 2\beta}}, \quad (2)$$

where k_c (= 1.5) is the stress concentration factor and k_d (= 1.1) is the dynamic load factor. The parameter r_i is the transmission ratio defined as the ratio of the number of teeth (n_{w_i}) in wheel to the number of teeth (n_{p_i}) in pinion (or, n_{w_i}/n_{p_i}) for the i-th gear-pair. The parameter ω_i is the speed of the wheel in rpm, a_i is the center distance for the corresponding gear-pair (calculated as $m(n_{w_i}+n_{p_i})/2$, where m is the module) in cm and t_i is the thickness of the gear-pair in cm. The parameter y_i is the form factor defined as follows: $y_i = 0.52\,(1+20/n_{w_i})$, where n_{w_i} is the number of teeth for the wheel in the corresponding gear-pair. The pressure angle is β (= 20 degrees) and the parameter E is the Young's modulus of the gear material in kgf/cm^2. The complete optimization problem is given below:

$$\left.\begin{array}{rll} \text{Maximize} & f_1 = p, & \\ \text{Minimize} & f_2 = \frac{\pi m^2}{4}\sum_{i=1}^{G}\left(n_{p_i}^2 + n_{w_i}^2\right) t_i, & \\ \text{Subject to} & \sigma_{b_i} \le S_b, & \text{for } i = 1,2,\ldots,G, \\ & \sigma_{w_i} \le S_w, & \text{for } i = 1,2,\ldots,G, \\ & t^{(L)} \le t_i \le t^{(U)}, & \text{for } i = 1,2,\ldots,G, \\ & p^{(L)} \le p \le p^{(U)}. & \end{array}\right\} \quad (3)$$

Thus, for the gear layout shown in Figure 1, there are a total of 10 real-valued decision variables and 38 inequality constraints. The problem can be made more complex by treating the module (m) also as a decision variable. Since any arbitrary module is not desired for difficulties in fabrication, the module can be treated as a discrete-valued variable taking only a few selected values.

2.2 Number of Teeth Varying

When the number of teeth in gears are also kept as decision variables, the resulting problem must involve additional constraints considering the following three aspects:

1. The center distance between two shafts must be maintained by all corresponding gear-pairs,
2. No gear should interfere with any shaft. These constraints arise because we have considered a two-dimensional layout of the gears having extended shafts. Some of these constraints can be eliminated by using a three-dimensional gearbox. However, here we consider the two-dimensional gearbox shown in Figure 1.
3. All obtained output shaft speeds Ω_k ($k = 1,2,\ldots,K$) must be within a permissible error limit ϵ from the ideal speeds Ω_k^I, and
4. The maximum gear-ratio in any gear-pair must not exceed a limit $r^{\max}$.

Moreover, the very nature of these additional decision variables (number of teeth) requires them to be treated as non-zero positive *integers*. For the gear-train layout shown in Figure 1, we formulate the above constraints:

$$n_1 + n_2 = n_3 + n_4 = n_5 + n_6, \quad \text{(between shaft 1 and 2)} \tag{4}$$

$$n_7 + n_8 = n_9 + n_{10} = n_{11} + n_{12}, \quad \text{(between shaft 2 and 3)} \tag{5}$$

$$n_{15} + n_{16} = n_{17} + n_{18}, \quad \text{(between shaft 4 and 5)} \tag{6}$$

$$n_2 \leq n_7 + n_8, \qquad n_4 \leq n_7 + n_8, \qquad n_6 \leq n_7 + n_8, \qquad n_7 \leq n_1 + n_2, \tag{7}$$

$$n_9 \leq n_1 + n_2, \qquad n_{11} \leq n_1 + n_2, \qquad n_8 \leq n_{13} + n_{14}, \qquad n_{10} \leq n_{13} + n_{14}, \tag{8}$$

$$n_{12} \leq n_{13} + n_{14}, \qquad n_{13} \leq n_7 + n_8, \qquad n_{14} \leq n_{15} + n_{16}, \qquad n_{15} \leq n_{13} + n_{14}, \tag{9}$$

$$n_{17} \leq n_{13} + n_{14}, \tag{10}$$

$$|\Omega_k - \Omega_k^I| \leq \epsilon \Omega_k^I, \quad k = 1, 2, \ldots, K, \tag{11}$$

$$\frac{n_{w_i}}{n_{p_i}}, \leq r^{\max} \quad \text{for } i = 1, 2, \ldots, G, \tag{12}$$

$$n_{w_i} \geq n^{(L)}, \quad \text{for } i = 1, 2, \ldots, G, \tag{13}$$

$$n_{p_i} \geq n^{(L)}, \quad \text{for } i = 1, 2, \ldots, G, \tag{14}$$

$$n_i \in \mathcal{Z}, \quad \text{for } i = 1, 2, \ldots, 2G. \tag{15}$$

In a mating gear-pair, n_{w_i} is assigned to be the one with the larger number of teeth. Equations 4 to 6 take care of the center distance constraint, as mentioned above. Inequalities 7 to 10 take care of the gear-shaft clearances. Inequality 11 must be satisfied for all output shaft speeds [9]. There are a total of K (18 in this case) such constraints. The next set of constraints take care of the maximum permissible gear ratio requirement in all gear-pairs. The next two set of constraints restrict the number of teeth in each gear to have a reasonable lower bound and the final constraint enforces that the number of teeth be integers. Thus, for the gearbox layout shown in Figure 1, there are five equality constraints and 58 inequality constraints involving the gear teeth variables. If the gear thickness, teeth, power, and module are all kept as variables, there are 10 real-valued, one discrete-valued, and 18 integer-valued variables in the gearbox design problem. Moreover, there are five linear equality constraints and 96 nonlinear inequality constraints.

2.3 Past Studies and Their Shortcomings

Past efforts to solve this complex optimal gearbox design problem [9, 13] have decomposed the overall problem into tractable subproblems to enable them to be solved by classical methods. Essentially, the following decomposition was used:

1. The design problem having gear teeth as the only design variables was solved first in order

to minimize two objectives: (i) the error in output shaft speeds from ideal speeds and (ii) the center distance between input and output shafts.

2. A gearbox with gear thicknesses as only decision variables is optimized for two different objectives: (i) minimization of the volume of gear material and (ii) maximization of the power delivered. Here, the number of teeth in each gear is kept fixed to that obtained in the first subproblem.

The above decomposition procedure of the original problem is a clever one, particularly because the first subproblem does not involve any mechanical strength considerations of the gears and is simply a kinematic layout optimization problem involving integer variables only. Two appropriate objectives involving number of gear teeth are chosen to suit the optimization problem. Having obtained the optimal number of teeth for gears (which incidentally also minimized the error in output speed deviation from ideal in [9]), the second subproblem focused on the mechanical strength considerations and optimized two important objectives (power delivered and the overall volume of the gearbox). It is important to note that there are at least two difficulties with the above decomposition-based approach:

1. First of all, such a clever decomposition is difficult to generalize for any arbitrary problem. A great deal of specific problem knowledge as well as expertise with the optimization literature is needed to come up with a dichotomy of a problem for the use of suitable methods.

2. Secondly, even after cleverly decomposing a problem and solving each subproblem, the obtained final solution need not to be the true optimal solution of the original problem. This is because, in complex problems having interacting variables, the individual optimal solutions to parts of the complete problem need not necessarily be combined to obtain the true overall optimal solution.

Having objectives of different types the classical approaches pose another difficulty. Since the second subproblem involves maximization of delivered power and minimization of volume of the gearbox, both past studies converted the first objective into the minimization of the inverse of the delivered power. Although, in principle, optimization of such a converted objective should result in the same optimal solution as that would be obtained by optimizing the original objective in the case of a single objective problem, the same may not be true for multiple objective optimization. In order to find a well-distributed set of Pareto-optimal solutions using a classical generating approach [11], such as a repetitive application of the weighted-sum approach as used in [9], a uniform set of weight vectors in the converted case may not produce a well-distributed set of

Pareto-optimal solutions of the original problem. In the above case, the weighted objective

$$F = w_1 \,(\text{volume}) + w_2 \left(\frac{1}{p}\right)$$

will produce more dense solutions having smaller values of p with a uniform set of weight vectors $(w_1, w_2)^T$ than that having a large value of p.

Some of the above difficulties associated with the classical approaches can be alleviated using an evolutionary optimization approach, and a more systematic, flexible, and a combined optimization task can be achieved. In the following section, we describe a specific multi-objective evolutionary optimization algorithm for this purpose and present simulation results in the next section.

3 Multi-Objective Evolutionary Algorithms

Over the past decade or so, there have been an increasing interest in using evolutionary algorithms (EAs) in multi-objective optimization. The main reason of their use is the ability of EAs to find multiple optimal solutions simultaneously in a single simulation run. With the advantage of inherent parallel processing of EAs, multi-objective EAs (MOEAs) can efficiently find a set of well-distributed non-dominated solutions close to the true Pareto-optimal front. The first author's recent book on the topic [3] discusses a number of MOEAs and presents a number of application case studies. Instead of using any weight information right in the beginning of the design process, MOEAs first find a set of well-distributed set of trade-off solutions. Since an MOEA works with a population of solutions in each iteration, a number of such solutions are possible to obtain in one simulation run. Thereafter, the obtained solutions are analyzed and a particular solution is chosen. Although some problem information may have to be used here, it is argued that the presence of a number of trade-off solutions facilitates the use of the problem information better than the absence of any such solution, as is the case for the classical preference-based approaches to multi-criterion optimization. In the following subsection, we briefly describe the specific MOEA used in this study.

3.1 Non-dominated Sorting Genetic Algorithm (NSGA-II)

NSGA-II begins it search with a population P of N randomly created (feasible or infeasible) solutions. The population members are then divided into a number of subpopulations $\mathcal{F}_i$ according to an increasing level of non-domination (so that $P = \cup_i \mathcal{F}_i$). This procedure requires repetitive identification of non-dominated solutions in the population and exclusion of identified non-dominated solutions from the population. This procedure requires $O(MN^2)$ computational

complexity, where M is the number of objectives. Thereafter, each solution in a subpopulation $\mathcal{F}_i$, is assigned a *crowding distance* value, equal to the sum of the difference in the normalized objective function values between the two neighboring solutions in each objective. Thus, this distance value for a solution captures the extent of crowding by other non-dominated solutions. If the neighboring solutions are close to a solution in the objective space, the corresponding crowding distance is small. Based on the non-domination level and crowding distance of each population member, three operators – tournament selection, recombination and mutation – are applied to the population one after another. The details of NSGA-II can be found in [5]. We describe these operations in the following paragraphs.

Tournament selection: The population is shuffled in a random order and two solutions at a time are picked from the top of the list for a tournament selection. The solution with a better non-domination level wins the tournament. In the case of a tie in the non-domination level, the solution with the larger crowding distance value wins. This way, better and well-dispersed non-dominated solutions are emphasized during the tournament selection procedure, thereby ensuring that a good spread of non-dominated solutions will be maintained at the end of a NSGA-II run.

Recombination: Two solutions (called *parents*) are picked randomly from the mating pool at a time and two new solutions (called *offspring*) are created using the following recombination operator. Each recombination is performed with a probability p_c. For handling real-valued decision variables, recombination is performed variable-wise using the SBX operator [4]. The i-th variable of two parents $x_i^{(1,t)}$ and $x_i^{(2,t)}$ is used to calculate two new values for the offspring solution as follows:

$$x_i^{(1,t+1)} = 0.5\left[(1+\beta_{q_i})x_i^{(1,t)} + (1-\beta_{q_i})x_i^{(2,t)}\right], \quad (16)$$

$$x_i^{(2,t+1)} = 0.5\left[(1-\beta_{q_i})x_i^{(1,t)} + (1+\beta_{q_i})x_i^{(2,t)}\right]. \quad (17)$$

where β_{q_i} is computed so as to assign more probability for creating new solutions closer to the parent solutions using a uniform random number $u_i \in [0,1]$:

$$\beta_{q_i} = \begin{cases} (2u_i)^{\frac{1}{\eta_c+1}}, & \text{if } u_i \leq 0.5; \\ \left(\frac{1}{2(1-u_i)}\right)^{\frac{1}{\eta_c+1}}, & \text{otherwise.} \end{cases} \quad (18)$$

The parameter η_c is the distribution index of the SBX operator. For handling bounded variables, Equation 18 is slightly modified [6], so that there is zero probability of creating a solution outside the variable bounds. For variables coded in binary strings, a single-point recombination, which swaps the *bits* on the right side of a randomly chosen cross-site along the string [8], is used here.

Mutation: Each variable in an offspring created by the recombination operator is perturbed in its vicinity with a mutation probability p_m. For real-valued variables, the mutated value is calculated as follows:

$$y_i^{(1,t+1)} = x_i^{(1,t+1)} + (x_i^{(U)} - x_i^{(L)})\bar{\delta}_i, \tag{19}$$

where the parameter $\bar{\delta}_i$ is calculated from the zero-mean polynomial probability distribution using a uniform random number $r_i \in [0,1]$, as follows:

$$\bar{\delta}_i = \begin{cases} (2r_i)^{1/(\eta_m+1)} - 1, & \text{if } r_i < 0.5, \\ 1 - [2(1-r_i)]^{1/(\eta_m+1)}, & \text{if } r_i \geq 0.5. \end{cases} \tag{20}$$

The parameter η_m has a similar effect to that of η_c. For variables coded in binary strings, a bit-wise mutation, which flips a bit with the probability p_m, is used.

Elite-preservation: After the offspring population Q is calculated by a serial application of the above three operators, both parent P and offspring Q are combined to form a new population R of size $2N$. The final operation of NSGA-II selects the best N solutions from this combined population R in a systematic manner. The population R is sorted according to different non-domination levels and each subpopulation from the best non-domination level is accepted one at a time, till no further subpopulation can be considered. If the final subpopulation is too big to be accepted in its totality, only those with the largest crowding distance values are chosen to fill up the population. This *elite-preserving* procedure allows preservation of good solutions from both parent and offspring populations. This completes one *generation* of the NSGA-II procedure. The overall computational complexity of one generation of NSGA-II to $O(MN^2)$.

The final accepted population of size N becomes the new parent population and the above procedure is continued for a number of generations till a termination criterion is satisfied. The continuous emphasis of non-dominated solutions and preservation of less-crowded solutions in the population cause NSGA-II to converge closer to the Pareto-optimal solutions with a well-distributed set of solutions.

3.2 Constraint Handling

For handling constraints in NSGA-II, we simply use a modified definition of the *domination* principle:

Definition 1 *A solution $\mathbf{x}^{(i)}$ is said to 'constraint-dominate' a solution $\mathbf{x}^{(j)}$ (or $\mathbf{x}^{(i)} \preceq_c \mathbf{x}^{(j)}$), if any of the following conditions are true:*

1. *Solution $\mathbf{x}^{(i)}$ is feasible and solution $\mathbf{x}^{(j)}$ is not.*

2. *Solutions* $\mathbf{x}^{(i)}$ *and* $\mathbf{x}^{(j)}$ *are both infeasible, but solution* $\mathbf{x}^{(i)}$ *has a smaller constraint violation.*

3. *Solutions* $\mathbf{x}^{(i)}$ *and* $\mathbf{x}^{(j)}$ *are feasible and solution* $\mathbf{x}^{(i)}$ *dominates solution* $\mathbf{x}^{(j)}$ *in the usual sense.*

The advantages of using the above definition in sorting a population into different levels of non-domination are as follows: (i) any infeasible solution is dominated by any feasible solution, (ii) between two infeasible solutions, the one with a smaller overall constraint violation (or in some sense, closer to a constraint boundary) dominates the other. Here, constraints are first normalized and constraint violations, if any, are simply added together to compute the overall constraint violation, and (iii) between two feasible solutions, the usual domination principle is used.

4 Computational Results

In this section, we investigate the efficacy of NSGA-II in optimizing the gear-train design problem depicted in Figure 1. We present a total of five cases, starting from a two-objective, 10-variable, 38-constraint real-parameter gearbox optimization problem and ending with a four-objective, 29-variable, 101-constraint, mixed-integer gearbox optimization problem. In all cases, the same NSGA-II code (A C code implementation which can be downloaded from the web site `http://www.iitk.ac.in/kangal/soft.htm`) is used with minor changes in the parameter settings.

The lower and upper limits of all gear thickness values are kept as 0.5 and 2.5 cm, respectively. The power delivered p is varied between 1 to 50 kW. The maximum speed-ratio $r^{\max} = 3.5$ is used here. The lower limit on gear teeth $n^{(L)} = 18$ is enforced. The bending strength of $S_b = 2,500$ kgf/cm^2 and wear strength $S_w = 17,500$ kgf/cm^2 are used. For the gear material, the Young's modulus is $E = 2.1(10^6)$ kgf/cm^2. The input speed of $1,400$ rpm is used and a speed reduction factor of 1.14 in the output speed is chosen so that the maximum desired speed is $1,400$ rpm and the minimum desired speed is about 150 rpm (this requires $K = 18$ different output speeds).

The probabilities of recombination and mutation operators used are $p_c = 0.9$ and $p_m = 0.3$, respectively, unless otherwise stated. The remaining parameters are modified from one case to another and are mentioned in the respective places.

4.1 Case I: Varying Gear Thickness Only

First, we fix the gearbox layout (by fixing the number of gear teeth and the module) as suggested in [9]. The following n_i values are used:

	Gear-Pair Number i								
	1	2	3	4	5	6	7	8	9
n_{p_i}	20	33	28	32	34	33	33	20	26
n_{w_i}	56	43	48	37	35	36	34	54	48

The above setting causes a maximum error of 6.28% in the output speeds. The module is kept fixed at 0.28 cm. Only thickness of each gear pair (nine of them) and power delivered are used as decision variables. All variable bounds are handled by initializing and creating solutions within specified bounds. Thus, no additional constraint is necessary to ensure satisfiability of the variable bounds. Moreover, since NSGA-II does not require any objective function and constraints to be differentiable, we can use the following strategy to reduce the number of stress constraints. Instead of having nine bending stress constraints, we can use only one constraint: $\max_{i=1}^{9} \sigma_{b_i} \leq S_b$. Similarly, nine wear stress constraints can be reduced to just one constraint. This way, NSGA-II considers the original gearbox design problem as a problem having 10 real-valued decision variables, two inequality constraints, and two objectives. Normalized constraint violations of both constraints are added to get the overall constraint violation, which is used in the constraint-domination calculation.

Figure 2 shows 10,000 random *feasible* solutions (obtained from approximately 160,000 randomly created solutions) in the objective space. NSGA-II is run with a population size of 100 for 500 generations, requiring a total of 50,000 function evaluations. Three independent runs from different initial populations have found very similar results. The average time taken to complete a run on a 1 GHz Pentium III processor is only 4.9 sec. The figure also shows 100 obtained non-dominated solutions using NSGA-II in small shaded circles. Thus, it is clear that the initial random population with 100 members did not contain any Pareto-optimal solution. It took NSGA-II about 500 generations to move from the initial random population to converge near to the Pareto-optimal set. The figure also shows four Pareto-optimal solutions reported in [9] in triangles. These solutions are as follows[1]:

	Solution Number			
	1	2	3	4
Power (kW)	1.71	3.37	4.47	8.55
Volume (cc)	777.76	919.54	1082.14	1858.76

In addition to a good convergence near the Pareto-optimal front, the NSGA-II is also able to maintain a good distribution of 100 solutions over the entire Pareto-optimal front in a single simulation run. Although the variable p is varied in $[1, 50]$ kW, no solution beyond $p = 8.56$ kW is found to lie on the obtained non-dominated front. We tabulate seven well-dispersed solutions (marked in Figure 2 with boxes) in Table 1. Two extreme solutions are exactly the same as those

[1] It is important to mention that for solution 3 the volume reported in [9] was 1076.77 cc. However, our calculation finds the volume to be somewhat different (1082.14 cc).

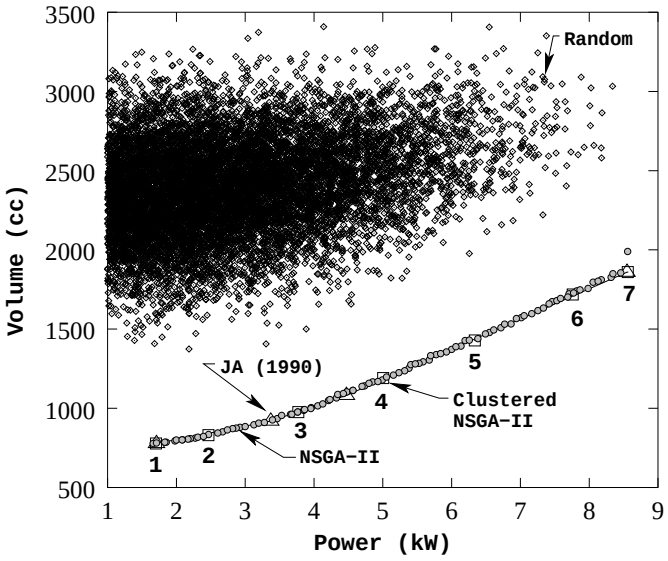


Figure 2: A random set of 10,000 feasible solutions, four Pareto-optimal solutions obtained in [9], and 100 non-dominated solutions obtained by NSGA-II are shown. Only thickness of 18 gears are varied here.

Table 1: Seven non-dominated solutions obtained by NSGA-II in Case I.

Sol. Id.	t_1	t_2	t_3	t_4	t_5	t_6	t_7	t_8	t_9	p	f_1	f_2
	(cm)									(kW)	(kW)	(cc)
1	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	1.703	1.703	777.757
2	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.727	0.536	2.466	2.466	830.759
3	0.500	0.500	0.500	0.528	0.500	0.505	0.563	1.105	0.832	3.772	3.772	975.983
4	0.501	0.500	0.500	0.694	0.659	0.627	0.761	1.472	1.090	5.006	5.006	1191.129
5	0.561	0.500	0.500	0.866	0.805	0.763	0.953	1.853	1.378	6.338	6.338	1428.466
6	0.692	0.500	0.500	1.083	0.982	0.936	1.215	2.267	1.690	7.761	7.761	1718.401
7	0.758	0.500	0.500	1.167	1.079	1.028	1.272	2.500	1.864	8.560	8.560	1860.073

found in [9]. Interestingly, all of these solutions have identical values for t_2 and t_3. Monotonic increase in other variables with increasing delivered power is also evident from the table.

Having found a set of non-dominated solutions on or close to the Pareto-optimal front, we now make the problem more flexible by allowing some additional parameters as decision variables in order to get a better insight into the gearbox design problem. It is important here to highlight that there exist no results to these problems using any classical method.

4.2 Case II: Varying Gear Module and Thickness

We first allow the module of the gear train to vary in the range $[0.10, 0.41]$ cm in steps of 0.01 cm. This way, a discrete set of 32 different modules are considered. We handle this decision variable using a five-bit binary substring. Thus, a complete solution to the NSGA-II is a binary substring

coding the gear module and 10 real-parameter values representing nine gear-pair thicknesses and the power delivered. Here, a bit-wise mutation with probability of 0.2 is used for the binary substring. Figure 3 shows the obtained solutions with a population size of 300. NSGA-II is run for 10,000 generations, taking an average of 660 seconds on the same machine as used in Case I. Four solutions of Case I (in circles) reported in [9] are also marked for comparison in the figure. It is clear that NSGA-II with the flexibility of choosing a suitable module produces a better set of non-dominated solutions, particularly in the range of gearboxes with smaller delivered power. Here, we have obtained solutions having the power varied in the range $[1.00, 18.28]$ kW, as compared to $[1.66, 8.56]$ kW obtained in Case I. It is also interesting to note that the entire non-dominated front is now arranged in an ascending order of magnitude of the module. The inset figure demonstrates this fact. Starting from a module very close to the chosen lower bound, solutions with a gradual increase in module are found. The module values are marked in the inset figure. The top-right portion of the obtained front corresponds to a module of 0.41 cm (the

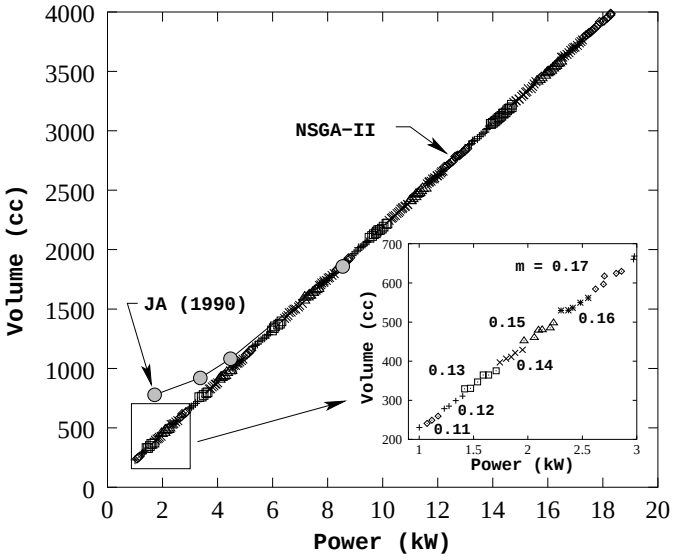


Figure 3: NSGA-II solutions for the case of module being an additional variable are shown against four Pareto-optimal solutions of Case I (as reported in [9]).

chosen upper bound on module) spanning the power delivered in the range $[17.30, 18.28]$ kW.

These results provide an useful insight to the gearbox design problem. If a large power transmission is desired, it is optimal to use gears with a large module rather than to use gears with large thickness and a small module. Figure 4 shows the variation of thickness of each gear-pair in solutions obtained by NSGA-II. It is clear from the plot that the gear thickness values do not change much from one solution to another, but it is clear from Figure 5 that a monotonic increase in the module is needed for larger power transmission. Another interesting aspect is that

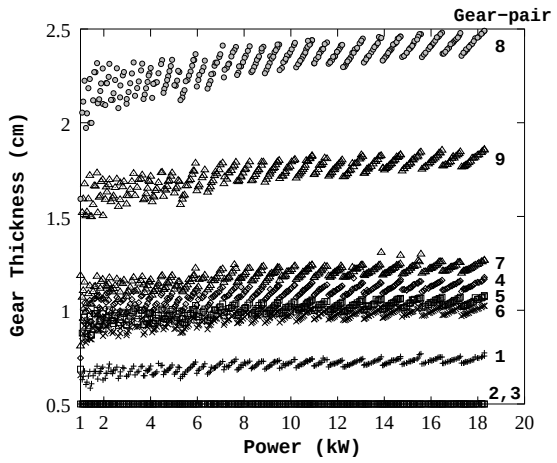


Figure 4: Variation of gear thickness in obtained solutions.

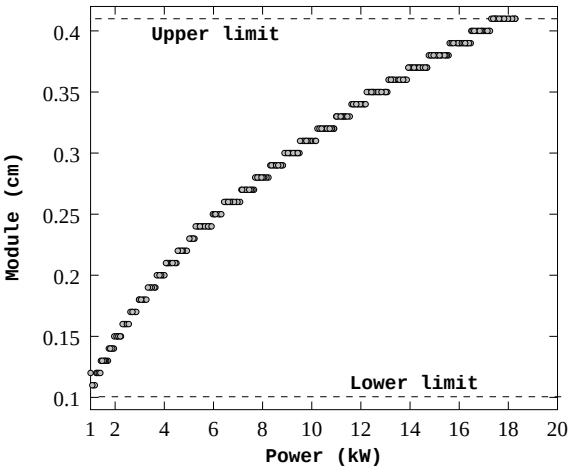


Figure 5: Variation of gear module in obtained solutions.

the rate of increase in the value of module requirement reduces for large power transmissions. Such information are useful to designers and practitioners and are often difficult to achieve by simply analyzing the mathematical expressions describing the objective functions and constraints.

However, by assuming that thickness of all gears are more or less the same for all Pareto-optimal solutions, equation 1 or 2 reveals the following relationship between the power and module:

$$p = A\sigma_{b,w}m^2, \tag{21}$$

where A is a constant which depends on the critical gear-pair configuration. The parameter $\sigma_{b,w}$ represents either σ_b or σ_w, depending on the corresponding constraints which are active at the Pareto-optimal solutions. For the gear teeth configuration used here, it has been shown in [9] that the bending stress constraints are always more critical than the wear stress constraints. Figure 6 plots the bending stress σ_b for all obtained non-dominated solutions. It is clear from the plot that except for gear-pairs 2 and 3, all other gear-pairs almost satisfy the bending stress constraint with the equality sign (bending strength being almost equal to $2,500$ kgf/cm^2). Thus, in those cases, equation 1 can be used to calculate the gear thickness values. As also evident from Figure 4, the allocated thicknesses in gear-pairs 2 and 3 are equal to the chosen lower limit (or 0.5 cm). It is also obvious from the Figure 6 that the critical constraints are the seven inequality bending stress constraints (1 and 4 to 9) satisfied with an equality sign. Thus, from equation 1, we can write $p = \kappa m^2 t_i$ for these seven gear-pairs, where κ is a constant. Now, in the expression for f_1 (volume), we can replace t_i (for $i = 1, 4, 5, 6, 7, 8, 9$) by using $p/(\kappa m^2)$. For $i = 2$ and 3, t_i is

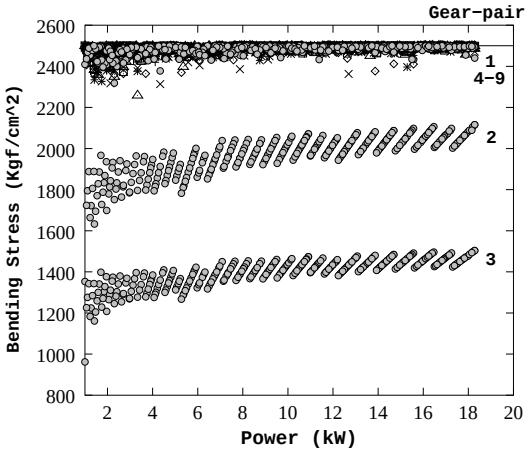


Figure 6: The bending stress of NSGA-II solutions are shown against the delivered power.

constant and is equal to $t^{(L)} = 0.5$ cm. Thus, the relationship between f_1 and p is as follows:

$$\begin{aligned} f_1 &= \frac{cm^2}{4} \sum_{\substack{i=1 \\ i \neq 2,3}}^{9} (n_{w_i}^2 + n_{p_i}^2) \frac{p}{\kappa m^2} + \frac{cm^2}{4} \sum_{i=2}^{3} (n_{w_i}^2 + n_{p_i}^2) 0.5, \\ &= D_1 p + D_2 m^2. \end{aligned}$$

Since, there is only one module for the entire gear system, the module m at a particular power p must satisfy equation 21, leading to $p \propto m^2$ or, $m^2 \propto p$. The exponent of the polynomial variation of module with power as observed in Figure 5 is close to 0.5, thereby supporting the above argument[2]. Therefore, the above equation yields $f_1 \propto p$. Figure 3 verifies that the relationship between objectives for the Pareto-optimal solutions (passing through the origin, if extrapolated) is linear, thereby supporting our argument above.

It is important to mention that such mixed-integer programming problems (having both discrete and continuous decision variables) are difficult to handle using classical optimization methods. Since evolutionary algorithms require only function information for any given solution and are flexible in handling mixed types of variable, such problems are comparatively easier to handle.

4.3 Case III: Varying Gear Teeth and Thickness

Next, we vary the number of teeth in all 18 gears (nine gear-pairs), in addition to the nine gear-pair thicknesses and power delivered. We fix the module at 0.28 cm. In order to handle the constraints given in equations 4 to 6, we use following two strategies with NSGA-II:

Variable elimination: We use the first three linear equality constraints to eliminate five decision

[2]It is also interesting to note that for a single bevel or hypoid gear-pair, Krenzer and Hotchkiss [10]) have also suggested the use of gear diameters increasing polynomially with transmitted torque having an exponent of 0.34.

variables. For example, the first equation can be used to eliminate two variables as follows:

$$n_4 = n_1 + n_2 - n_3, \qquad n_6 = n_1 + n_2 - n_5. \tag{22}$$

Redundancy reduction: The redundancy in solutions is reduced by prefixing an order of gear sizes in a gear-pair cluster. For example, the redundancy in solutions can be reduced by assuming that $n_5 > n_3 > n_1$. By this way several permutations of these gear sizes will not be reconsidered during optimization.

The latter condition is further made simpler by using a different set of decision variables. For example, instead of using n_1, n_3, and n_5 as decision variables, we may use n_1, $\Delta n_{1,3}$, and $\Delta n_{3,5}$ as decision variables and derive n_3 and n_5 as follows:

$$n_3 = n_1 + \Delta n_{1,3}, \qquad n_5 = n_1 + \Delta n_{1,3} + \Delta n_{3,5}. \tag{23}$$

By restricting these incremental variables (Δn_i) to be greater than zero, we ensure that $n_5 > n_3 > n_1$. With these considerations, the total number of independent variables kept for representing number of gear teeth are 13, thereby making the total number of variables to 23, causing a reduction of five integer variables. In the NSGA-II implementation, we allow these incremental variables (Δn_i) to vary in $[1, 31]$ in steps of one and n_i variables to vary in $[18, 81]$ in steps of one. The recombination and mutation operators treat these variables as real-valued at first. Before the objective functions and constraints are evaluated, they are rounded to their nearest integers. For these integer variables, we use $\eta_c = 2$ and $\eta_m = 50$ in equations 18 and 20, respectively. Figure 7 shows the obtained non-dominated solutions for $\epsilon = 0.05$ and 0.10. It is clear from the figure that if a large error (ϵ) in the output shaft speeds from the ideal is permitted, better non-dominated solutions are expected. The figure also marks the four Pareto-optimal solutions reported elsewhere [9] obtained by using a fixed number of teeth and module in gears (Case I). Since the number of teeth in gears are also kept as variables here, a more flexible search is permitted, thereby obtaining a better set of non-dominated solutions. Table 2 shows the gear teeth and thickness of four selected solutions (uniformly spaced across the obtained front). It is clear from the table that for gearboxes with a large delivered power, thickness and teeth in gears are also large. The table also shows the corresponding gear details for $\epsilon = 0.05$. For the gear sizes, a trend similar to that found in the $\epsilon = 0.10$ case is also observed here. Although the latter solutions cause a maximum of $\epsilon = 0.05$ error in the output speeds from the desired, each of four solutions tabulated above found a somewhat similar yet different gear teeth combination to achieve the same objective. In order to investigate the efficacy of NSGA-II in handling constraints, we calculate the percentage maximum error in the output shaft speed for each obtained non-dominated solution.

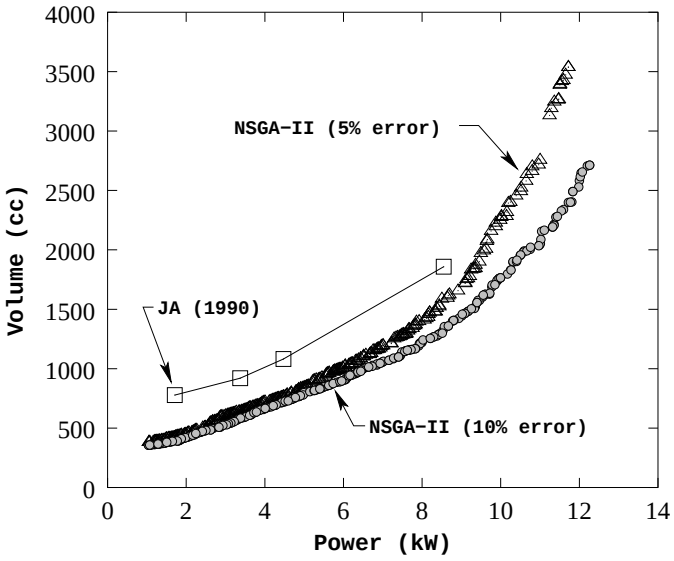


Figure 7: NSGA-II solutions are shown with two limiting errors in output speeds. Gear teeth and thickness are varied here.

Table 2: Gear details of four selected solutions in Case III with a maximum of 10% error. The results with at most 5% error are shown in brackets.

Sol. No.	Power (kW)	Volume (cc)		Gear-Pair Number								
				1	2	3	4	5	6	7	8	9
1	1.040 (1.050)	355.03 (381.95)	n_{p_i}	18 (18)	27 (28)	18 (18)	18 (19)	19 (21)	20 (22)	18 (18)	18 (18)	23 (23)
			n_{w_i}	37 (38)	28 (28)	37 (38)	23 (27)	22 (25)	21 (24)	32 (28)	33 (35)	28 (30)
			t_i	0.50 (0.50)	0.50 (0.50)	0.50 (0.50)	0.50 (0.50)	0.50 (0.50)	0.50 (0.50)	0.50 (0.50)	0.51 (0.50)	0.50 (0.50)
2	5.503 (5.547)	846.52 (941.52)	n_{p_i}	21 (19)	28 (30)	18 (19)	28 (28)	26 (30)	25 (29)	19 (22)	18 (22)	24 (18)
			n_{w_i}	39 (41)	32 (30)	42 (41)	29 (33)	31 (31)	32 (32)	38 (29)	43 (30)	37 (47)
			t_i	0.50 (0.50)	0.50 (0.50)	0.50 (0.50)	0.55 (0.64)	0.51 (0.59)	0.50 (0.55)	0.92 (1.10)	1.88 (1.80)	1.29 (1.29)
3	9.008 (9.094)	1443.68 (1718.20)	n_{p_i}	22 (25)	28 (30)	18 (18)	29 (32)	27 (30)	26 (28)	20 (22)	18 (20)	24 (27)
			n_{w_i}	39 (43)	33 (38)	43 (50)	36 (32)	38 (34)	39 (36)	39 (35)	55 (65)	49 (58)
			t_i	0.75 (0.50)	0.50 (0.50)	0.50 (0.50)	0.68 (0.73)	0.66 (0.70)	0.66 (0.67)	1.06 (1.25)	2.46 (2.30)	1.78 (1.65)
4	12.262 (11.720)	2712.36 (3537.46)	n_{p_i}	29 (27)	31 (39)	19 (24)	37 (37)	35 (40)	33 (39)	30 (36)	23 (27)	31 (36)
			n_{w_i}	44 (54)	42 (42)	54 (57)	38 (44)	40 (41)	42 (42)	47 (45)	80 (81)	72 (72)
			t_i	1.01 (0.93)	0.50 (1.25)	0.50 (1.57)	0.79 (1.07)	0.75 (0.95)	0.72 (0.91)	1.08 (1.30)	2.50 (2.49)	1.82 (1.82)

Figure 8 shows the proportion of obtained non-dominated solutions having different percentage maximum error. For clarity, all solutions with a percentage maximum error in a block of 0.25% are grouped together. The figure shows that a large proportion of non-dominated solutions maintain a percentage maximum error close to the maximum permissible value. This indicates that the constraint handling strategy used in NSGA-II is effective in exploiting the chosen maximum error limit in the output speeds.

With a specific number of gear teeth used in Case I, it was observed that the bending stress constraints are always more critical than the wear stress constraints. However, in this case (when the number of gear teeth are also kept as decision variables), the bending stress need not always

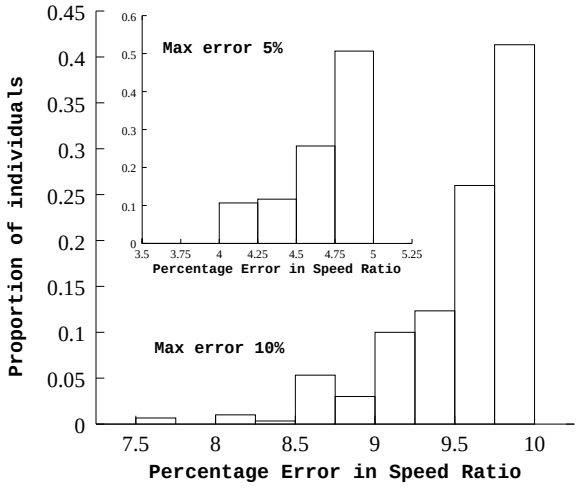


Figure 8: Proportion of obtained solutions having different maximum percentage error in speed ratios is shown for two cases of optimization (with 5% and 10% limiting error).

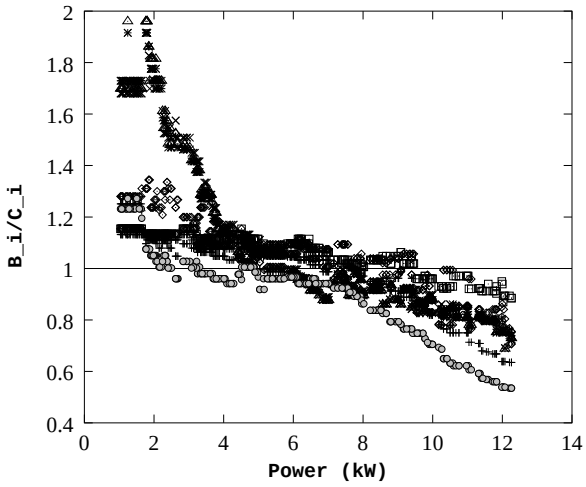


Figure 9: The ratio B_i/C_i for all nine gear-pairs observed in different obtained solutions.

be critical. To investigate which constraint is critical in which scenario, we have plotted B_i/C_i for all nine gear-pairs versus p in Figure 9 for all obtained non-dominated solutions (where the bending constraints with the equality sign are expressed as $p = B_i t_i$ and the wear constraints with the equality sign are expressed as $p = C_i t_i$ [9]) It can be seen from the figure that the ratio B_i/C_i is more than one for small values of p indicating that wear stress constraints are more critical than bending stress constraints. Using the wear stress constraints, the parametric relationship of the Pareto-optimal front can be written as follows:

$$f_2 = \gamma \sum_{i=1}^{9} (n_{p_i}^2 + n_{w_i}^2) t_i, \quad p = \alpha_i t_i \frac{n_{w_i} n_{p_i}^2}{n_{p_i} + n_{w_i}}.$$

A dimensional study on the design variables reveal a linear relationship between the two objectives f_2 and p, as observed in Figure 7. On the other hand, for solutions with a large value of p, bending stress constraints are more critical. With a large power requirement, the gear sizes also need to be large and are critical for failure due to bending than wear. Using the bending stress constraints, the parametric relationship becomes

$$f_2 = \gamma \sum_{i=1}^{9} (n_{p_i}^2 + n_{w_i}^2) t_i, \quad p = \beta_i t_i \frac{n_{p_i}^2}{n_{w_i}} \left(1 + \frac{20}{n_{w_i}}\right).$$

As can be seen from these equations, for an increase in the number of gear teeth and thickness the rate of change of f_2 is more here than that in the former case. Figure 7 depicts this non-linear relationship among the obtained solutions for large-powered gearboxes.

4.4 Case IV: Varying Gear Module, Teeth and Thickness

Next, we make the gearbox design problem even more flexible by allowing the module also as a decision variable. The range and step used for this variable are kept the same as those used in Case II ($m \in [0.10, 0.41)$ cm in steps of 0.01 cm). This increases the total number of variables to 29. An $\epsilon = 0.10$ is used in this study. Rest all parameters are kept the same as those used in the previous case. Figure 10 shows the 300 non-dominated solutions obtained using NSGA-II after 10,000 generations. On the same computer mentioned earlier, an NSGA-II run took an average of 660 seconds. As in the Case II, a monotonically increasing magnitude of module is observed

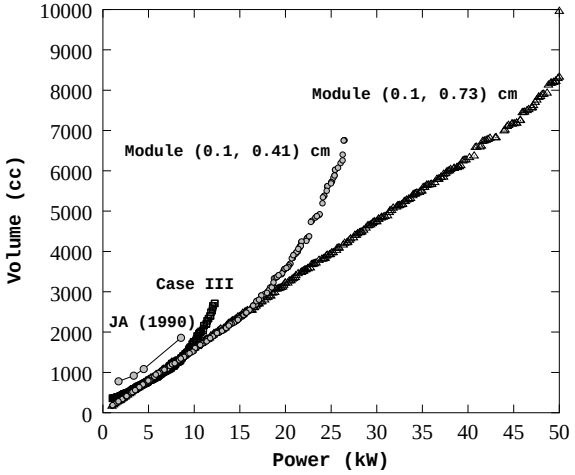


Figure 10: NSGA-II solutions for varying gear module, teeth, and thickness are shown. Non-dominated fronts for two different ranges of modules are compared with that obtained in Case III (with a fixed module of 0.28 cm). Open circles indicate results when the module is allowed to vary within [0.10,0.41] cm and the triangles indicate results for modules varying in [0.10,0.73] cm.

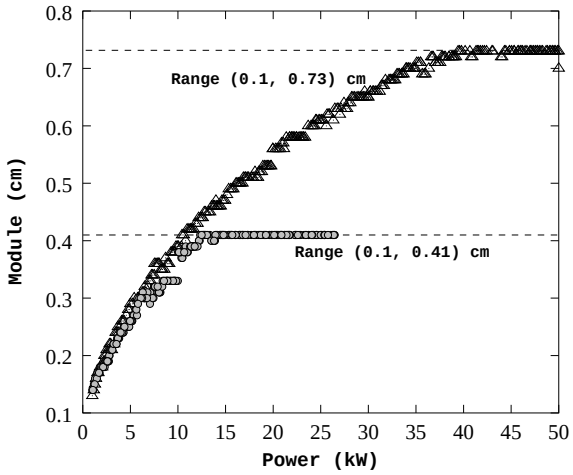


Figure 11: Module m obtained in different non-dominated solutions for two cases.

for solutions having increasing requirement in delivered power. Figure 10 also shows the non-dominated front obtained in Case III (optimized by keeping the module fixed). It is interesting to note that a better non-dominated set of solutions is possible to obtain by allowing the module to be a decision variable. Two distinct phases can be observed among the obtained non-dominated solutions. For gearboxes having a small delivered power the volume-power relationship is linear and for gearboxes having a large delivered power the relationship is nonlinear. These two phases have distinct properties.

Figure 11 shows that the variation of module in the obtained solutions. It is interesting to note

that module increases with power requirement in a manner similar to that observed in Case II till the chosen upper limit ($m^{(U)} = 0.41$ cm) on module is reached. An investigation on the gear thickness and the number of teeth for these solutions indicates that the same for gear-pairs 7, 8 and 9 increase monotonically with the delivered power. For these solutions, the corresponding non-dominated front is *linear*, as evident from Figure 10 and as explained in Case II. Thereafter, the module remains fixed to the upper limit. For these solutions, the optimized volume-power relationship is nonlinear in a manner similar to that observed in Case III. It is also interesting to note that the nonlinear portion of the non-dominated front will get *dominated*, if larger modules are allowed. To demonstrate this fact we have rerun NSGA-II with a relaxed upper bound (module varying in the range $(0.10, 0.73)$ cm). The corresponding non-dominated front obtained after 10,000 generations are shown using triangles in Figure 10. It is evident that now the obtained front remains linear till the module of the gearbox becomes equal to the chosen upper limit (0.73 cm). Figure 11 also shows that the corresponding module increases monotonically and in a very similar manner ($m \propto \sqrt{p}$) as that obtained in the earlier case (with $m^{(U)} = 0.41$ cm).

4.5 Case V: Three Objectives

To investigate further the performance of NSGA-II on more than two objectives, we formulate a third objective function of *minimizing* the center distance between input and output shafts. This objective function can be written as follows:

$$f_3 = \frac{m}{2}\left[(n_1 + n_2) + (n_7 + n_8) + (n_{13} + n_{14}) + (n_{15} + n_{16})\right]. \tag{24}$$

All constraints, decision variables, and other parameters as used in the previous case are also adopted here. The module is varied in $[0.10, 0.41]$ cm in steps of 0.01 cm. In section 2.3, we have discussed that previous studies on the gearbox design have considered a total of four different objectives. However, any particular optimization did not consider more than two objectives. Here and in the next section, for the first time, we show simulation results for the gearbox optimization problem with three and four objectives.

For the three-objective case, we have used $\epsilon = 0.1$. NSGA-II with 300 population members takes an average of 700 seconds to run 10,000 generations on the same machine mentioned before. Although solutions found are satisfactory, in order to achieve a better convergence, we have continued the NSGA-II run for 100,000 generations. Figure 12 shows 300 non-dominated solutions obtained in one such typical run of NSGA-II in triangles. The projection on the x-y plane shows a similar trade-off (partly linear and partly nonlinear) in delivered power and volume as that observed in the two-objective case (Case IV with $m \in [0.10, 0.41]$ cm). It is also interesting to note the dimension of the obtained non-dominated front in the three-dimensional objective

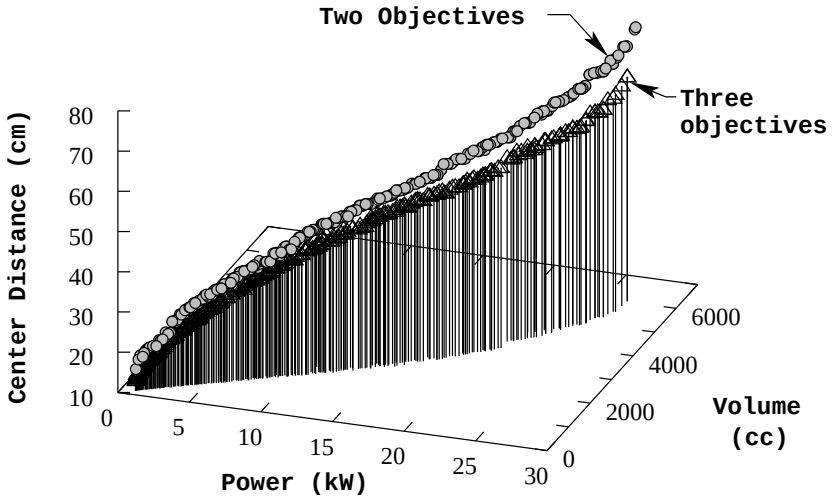


Figure 12: NSGA-II solutions for an additional objective of minimizing center distance are compared with NSGA-II solutions with two objectives (maximizing power and minimizing volume).

space. Although three objectives are considered, a two-dimensional curve is obtained as the non-dominated front. This may be because the center distance and overall volume of gear material are not likely to be conflicting to each other, thereby loosing one dimension in the non-dominated front. On this figure, we also superimpose the non-dominated solutions obtained in Case IV (two-objective case of maximizing power and minimizing volume, run up to 100,000 generations). The center distance f_3 is computed for each obtained solution and is plotted on the three-objective space. It is clear that when all three objectives are considered in NSGA-II, smaller values of the third objective (center distance) is found. This is because in Case IV no selection advantage was assigned to achieve a smaller center distance.

By investigating the variable values in the two cases (two and three objectives), it is observed that a smaller module and larger gear thicknesses are obtained for the three-objective case. However, the number of teeth in the gear-pairs are more or less the same in both cases. These properties of the obtained solutions can be explained by analyzing equation 24. Since the center distance between the input and output shafts have no consideration for the gear thickness and is directly proportional to the module, NSGA-II has optimized the center distance by reducing the module. To maintain the volume-power optimality, NSGA-II finds gearboxes with larger gear thickness values. Thus, the consideration of the third objective is found to be beneficial in terms of reducing the overall size of the gearbox.

4.6 Case VI: Four Objectives

Finally, we include the fourth objective of minimizing the maximum error in the output speeds from the ideal speeds. The number of variables, constraints, and NSGA-II parameters are kept the same as in Case V. Figure 13 shows the obtained non-dominated front on the three-objective space. It is observed that the projection of the non-dominated solutions on the volume-power

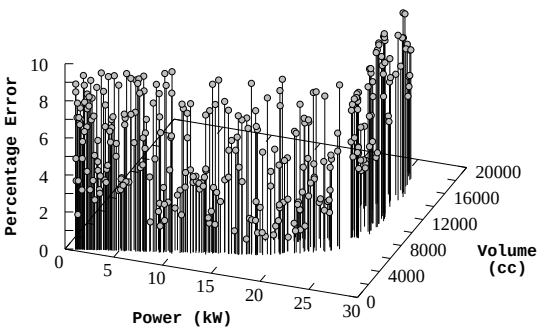


Figure 13: Non-dominated solutions for four objectives.

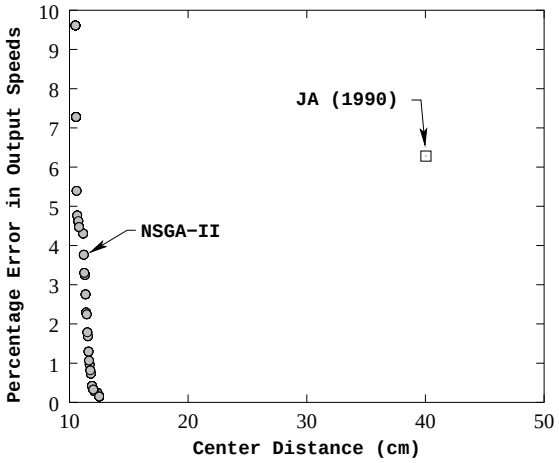


Figure 14: Trade-off between ϵ and center distance for two-objective NSGA-II and Jain and Agogino's [9] study.

space has an identical relationship as that observed in Case V and the optimal volume-power relationship is almost invariant to any level of ϵ. It is intuitive that gearboxes for a large ϵ consideration can be made from a number of different gear-teeth combinations, thereby allowing a wide variation in the center distance between input and output shafts. On the other hand, a gearbox with a very small maximum output speed error ϵ may not correspond to many gear teeth combinations, thereby reducing the range of allowable center distance between the input and output shafts.

4.7 Case VII: Output Speed Error Versus Center Distance

Figure 14 shows the obtained non-dominated front (shown in circles), when the gearbox design problem is solved only for two objectives: (i) minimize the maximum error in output speeds and (ii) minimize the center distance between the input and output shafts. All constraints and variables as used in the four-objective case are used here. The trade-off between error and center distance is clear from the figure. An analysis of the obtained solutions reveals that all of them have a fixed module of 0.1 cm (the lower bound) and a fixed delivered power of 1 kW (the lower bound). A gearbox with the smallest ϵ of 0.15% and a corresponding center distance of 12.5 cm is

found. Since both ϵ and center distance are minimized, the optimal gearboxes are also expected to be small, requiring a small volume and delivered power. The earlier study [9] with these two objectives and without any consideration of any mechanical strength constraints reported only one solution with a center distance of 40.04 cm and 6.28% of error in output speed (shown with a box). On the contrary, the study here suggests that there exist multiple trade-off solutions between these two objectives when all constraints are taken into account and the NSGA-II is capable of finding a number of such solutions.

5 Conclusions

The multi-speed gearbox design optimization problem considered in this paper provides an adequate challenge to almost every aspect of optimization. The design problem involves (i) more than one (and as many as four) conflicting objectives of design, (ii) a total of 29 mixed type of decision variables (integer, discrete, and real-valued), and (iii) five equality and 96 nonlinear inequality constraints. A couple of previous studies using classical optimization methods had to decompose the overall problem into two tractable problems with at most 10 variables and 18 inequality constraints. This paper has shown that the application of a multi-objective evolutionary algorithm – NSGA-II – to the decomposed problems has able to find the same or better Pareto-optimal front as that obtained by classical methods. The other highlights of this paper are as follows:

1. NSGA-II has found as many as 300 different trade-off optimal solutions in one single simulation run, as opposed to the need of a repetitive use of any classical method to find multiple optimal solutions.

2. NSGA-II has been able to solve different complexities of the gearbox design problem without any major change in its parameter settings.

3. The obtained solutions have been analyzed to discover important design principles related to an optimal design of a gearbox:

    - For an identical output speed requirement, an optimal gearbox requires a larger module for larger delivered power. The optimal module varies proportional to the squared root of the desired delivered power.
    - However, in the event of impracticalities in using a large module for high-powered gears, it is better to fix the module to be as large as possible and use thicker gears rather than a gear with more teeth.
    - For low-powered gearboxes, the wear stress failure is more critical than bending stress failure and for high-powered gearboxes, the opposite is true.

Such important yet generic properties of a gearbox design are difficult to obtain by any other procedure, including an analysis of the mathematical programming problem describing the optimization problem. This is due to the fact that in order to reveal any such properties, it is first necessary to *find* a set of well-distributed Pareto-optimal solutions. Since the NSGA-II allows a tractable way of achieving a well-distributed set of non-dominated solutions close to the true Pareto-optimal front, such systematic applications and an analysis of the resulting solutions are possible.

References

[1] R. K. Carroll and G. E. Johnson. Optimal design of compact spur gear sets. *ASME Journal of Mechanisms, Transmissions, and Automation in Design*, 106:95–101, 1984.

[2] V. Chankong and Y. Y. Haimes. *Multiobjective Decision Making Theory and Methodology.* New York: North-Holland, 1983.

[3] K. Deb. *Multi-objective optimization using evolutionary algorithms.* Chichester, UK: Wiley, 2001.

[4] K. Deb and R. B. Agrawal. Simulated binary crossover for continuous search space. *Complex Systems*, 9(2):115–148, 1995.

[5] K. Deb, S. Agrawal, A. Pratap, and T. Meyarivan. A fast and elitist multi-objective genetic algorithm: NSGA-II. *IEEE Transactions on Evolutionary Computation*, 6(2):182–197, 2002.

[6] K. Deb and M. Goyal. A robust optimization procedure for mechanical component design based on genetic adaptive search. *Transactions of the ASME: Journal of Mechanical Design*, 120(2):162–164, 1998.

[7] M. Ehrgott. *Multicriteria Optimization.* Berlin: Springer, 2000.

[8] D. E. Goldberg. *Genetic Algorithms for Search, Optimization, and Machine Learning.* Reading, MA: Addison-Wesley, 1989.

[9] P. Jain and A. M. Agogino. Theory of design: An optimization perspective. *Mech. Mach. Theory*, 25(3):287–303, 1990.

[10] T. J. Krenzer and R. G. Hotchkiss. Bevel and hypoid gears. In J. E. Shigley and C. R. Mishcke, editors, *Standard handbook of machine design*, pages 34.1–34.61. New York: McGraw-Hill, 1986.

[11] K. Miettinen. *Nonlinear Multiobjective Optimization.* Kluwer, Boston, 1999.

[12] S. S. Rao. *Optimization: Theory and Applications.* Wiley, New York, 1984.

[13] S. S. Rao and H. R. Eslampour. Multistage multiobjective optimization of gearboxes. *Journal of Mechanisms, Transmissions, and Automation in Design*, 108:461–468, 1986.

Figure Captions

Figure 1: Schematic of a gear train.

Figure 2: A random set of 10,000 feasible solutions, four Pareto-optimal solutions obtained in [9], and 100 non-dominated solutions obtained by NSGA-II are shown. Only thickness of 18 gears are varied here.

Figure 3: NSGA-II solutions for the case of module being an additional variable are shown against four Pareto-optimal solutions of Case I (as reported in [9]).

Figure 4: Variation of gear thickness in obtained solutions.

Figure 5: Variation of gear module in obtained solutions.

Figure 6: The bending stress of NSGA-II solutions are shown against the delivered power.

Figure 7: NSGA-II solutions are shown with two limiting errors in output speeds. Gear teeth and thickness are varied here.

Figure 8: Proportion of obtained solutions having different maximum percentage error in speed ratios is shown for two cases of optimization (with 5% and 10% limiting error).

Figure 9: The ratio B_i/C_i for all nine gear-pairs observed in different obtained solutions.

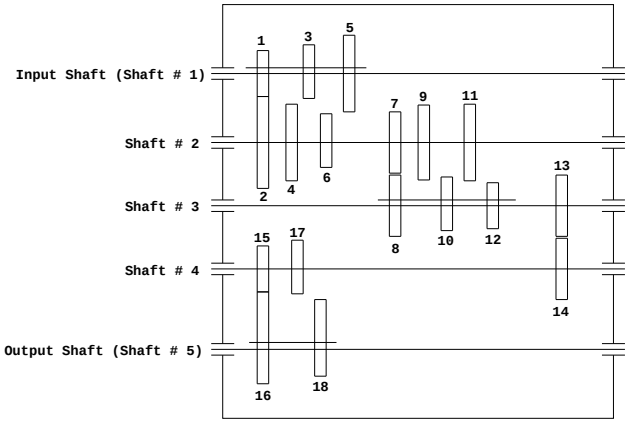
Figure 10: NSGA-II solutions for varying gear module, teeth, and thickness are shown. Non-dominated fronts for two different ranges of modules are compared with that obtained in Case III (with a fixed module of 0.28 cm). Open circles indicate results when the module is allowed to vary within [0.10,0.41] cm and the triangles indicate results for modules varying in [0.10,0.73] cm.

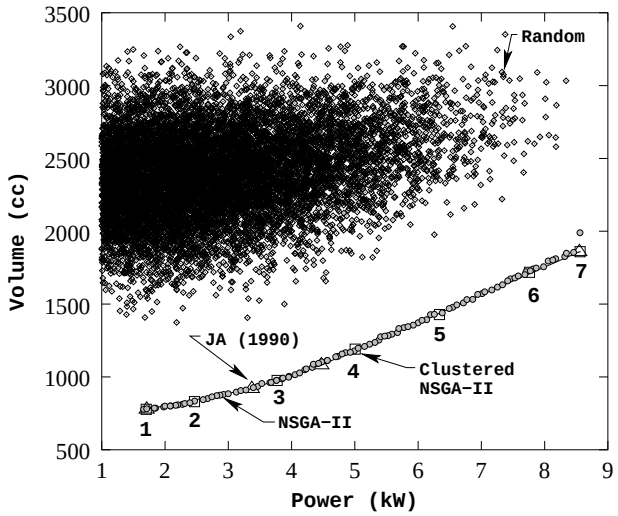
Figure 11: Module m obtained in different non-dominated solutions for two cases.

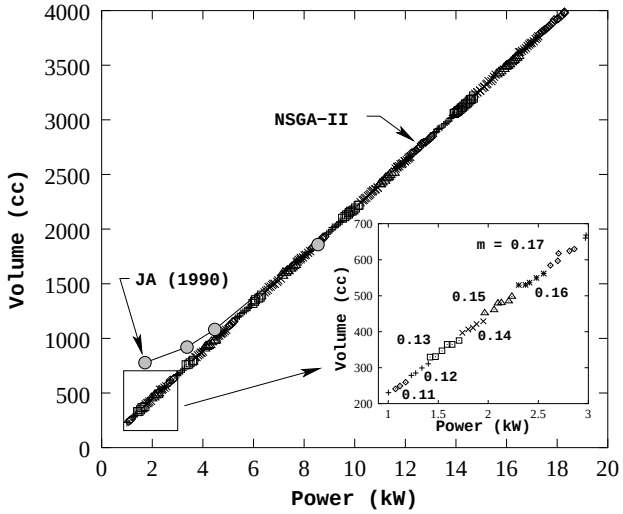
Figure 12: NSGA-II solutions for an additional objective of minimizing center distance are compared with NSGA-II solutions with two objectives (maximizing power and minimizing volume).

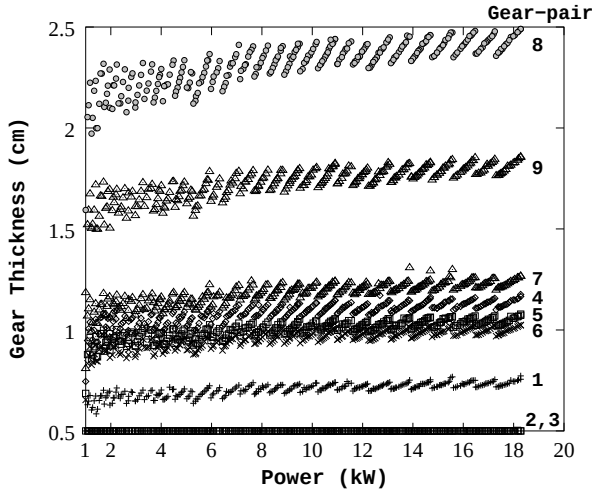
Figure 13: Non-dominated solutions for four objectives.

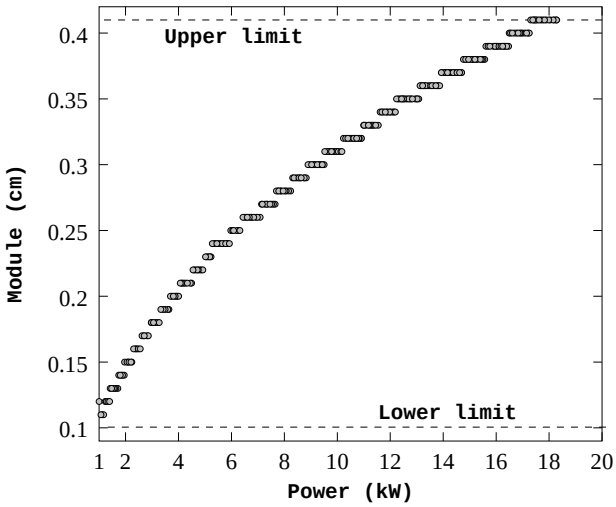
Figure 14: Trade-off between ϵ and center distance for two-objective NSGA-II and Jain and Agogino's [9] study.

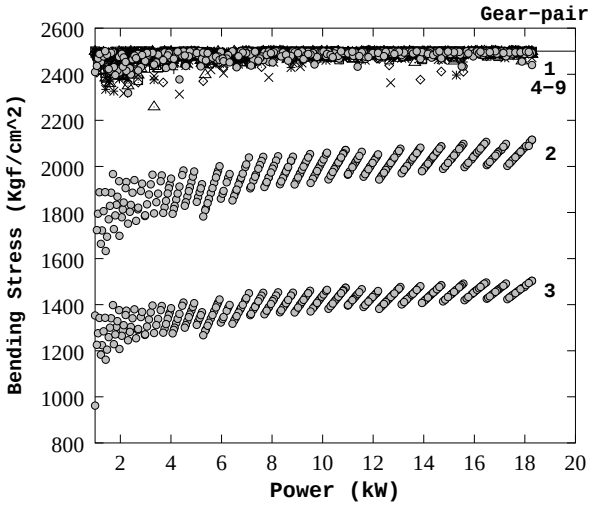

Input Shaft (Shaft # 1)
Shaft # 2
Shaft # 3
Shaft # 4
Output Shaft (Shaft # 5)
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18

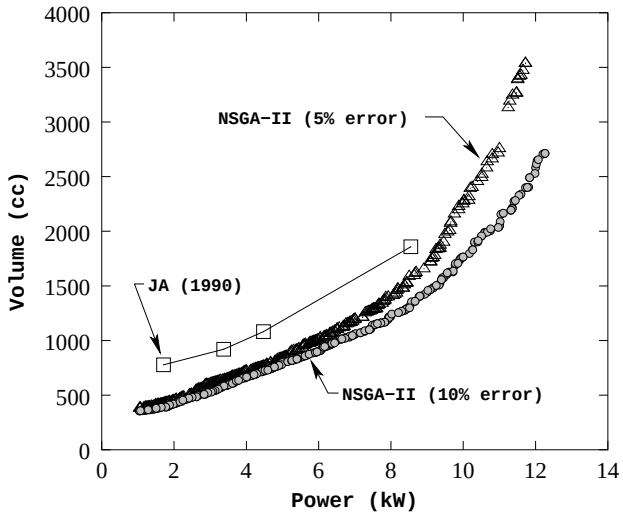

Random
JA (1990)
Clustered
NSGA-II
NSGA-II
1
2
3
4
5
6
7
Volume (cc)
Power (kW)
3500
3000
2500
2000
1500
1000
500
1
2
3
4
5
6
7
8
9

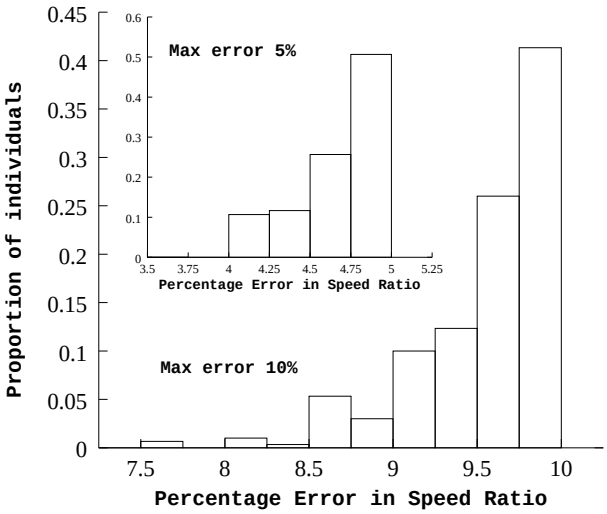

Volume (cc)
Power (kW)
NSGA-II
JA (1990)
m = 0.17
0.15
0.16
0.13
0.14
0.12
0.11

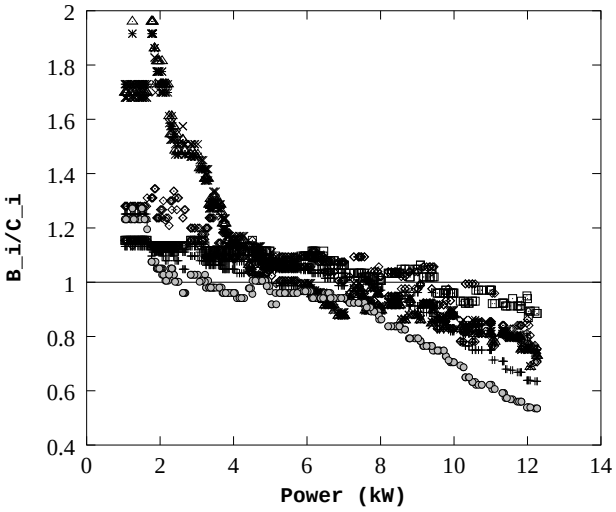

Gear−pair
8
9
7
4
5
6
1
2,3
Gear Thickness (cm)
Power (kW)
2.5
2
1.5
1
0.5
1
2
4
6
8
10
12
14
16
18
20

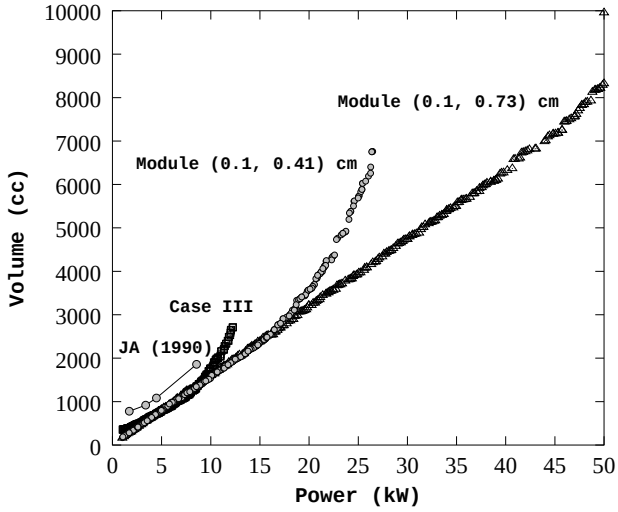

Upper limit
Lower limit
Module (cm)
Power (kW)
0.4
0.35
0.3
0.25
0.2
0.15
0.1
1
2
4
6
8
10
12
14
16
18
20

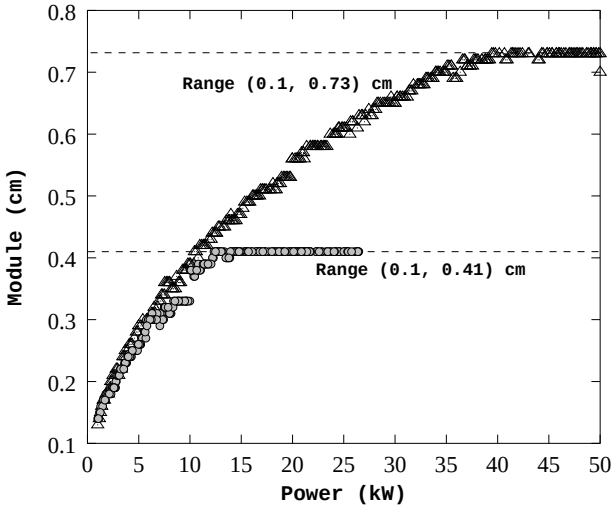

Gear−pair
1
4−9
2
3
Bending Stress (Kgf/cm^2)
Power (kW)
2600
2400
2200
2000
1800
1600
1400
1200
1000
800
2
4
6
8
10
12
14
16
18
20

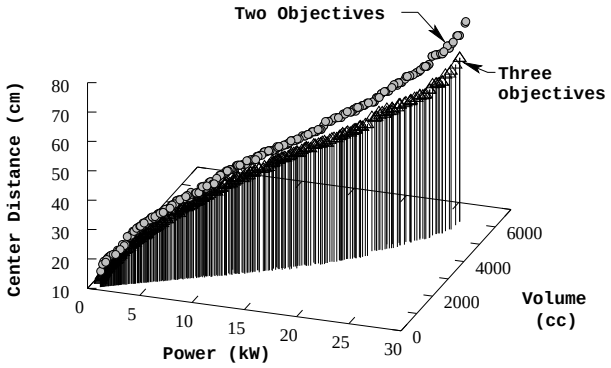

4000
3500
3000
2500
2000
1500
1000
500
0
Volume (cc)
0
2
4
6
8
10
12
14
Power (kW)
NSGA−II (5% error)
JA (1990)
NSGA−II (10% error)

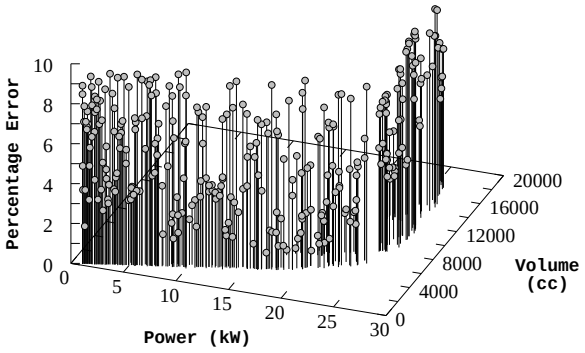

Max error 5%
Percentage Error in Speed Ratio
Max error 10%
Proportion of individuals
Percentage Error in Speed Ratio

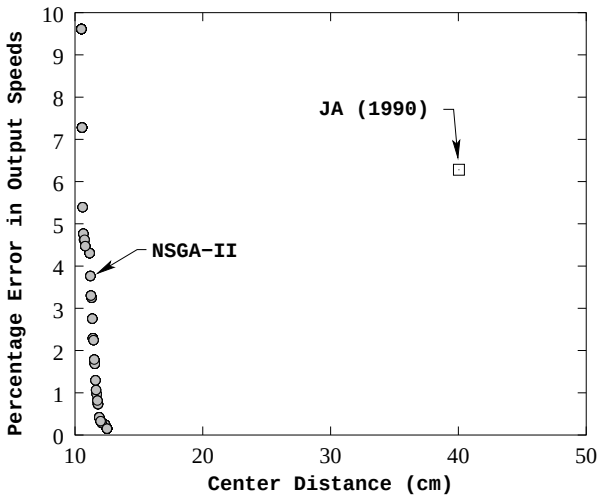

B_i/C_i
Power (kW)
2
1.8
1.6
1.4
1.2
1
0.8
0.6
0.4
0
2
4
6
8
10
12
14


Volume (cc)
Power (kW)
Module (0.1, 0.73) cm
Module (0.1, 0.41) cm
Case III
JA (1990)
0
1000
2000
3000
4000
5000
6000
7000
8000
9000
10000
0
5
10
15
20
25
30
35
40
45
50


Range (0.1, 0.73) cm
Range (0.1, 0.41) cm
Module (cm)
Power (kW)
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
5
10
15
20
25
30
35
40
45
50


Two Objectives
Three objectives
Center Distance (cm)
80
70
60
50
40
30
20
10
0
5
10
15
20
25
30
Power (kW)
6000
4000
2000
0
Volume (cc)


Percentage Error
10
8
6
4
2
0
0
5
10
15
20
25
30
Power (kW)
20000
16000
12000
8000
4000
0
Volume
(cc)


Percentage Error in Output Speeds
Center Distance (cm)
JA (1990)
NSGA−II
0
1
2
3
4
5
6
7
8
9
10
10
20
30
40
50