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, jsachin}@iitk.ac.in

KanGAL Report No. 2002001

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 problem having different complexities, the efficacy of NSGA-II in handling different types of decision variables, constraints, and multiple objectives are demonstrated. An investigation of multiple obtained solutions provides a number of interesting insights to the gearbox design problem, which are otherwise difficult to obtain using existing optimization techniques.

1 Introduction

A multi-speed gearbox design optimization problem introduces a number of challenges (Carroll and Johnson, 1984), particularly if the design involves optimization of both its kinematic structure and gear details. 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 satisfaction of a number of equality and inequality constraints. The resulting optimization problem becomes more 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 (Jain and Agogino, 1990). Although there exist a number of mixed-integer nonlinear programming algorithms (Rao, 1984), they are often computationally demanding and are avoided as far as possible. To tackle such problems, a multi-step optimization procedure is usually adopted. In each step, a part of the problem (for example, involving only integer variables) is first solved by fixing the rest of the problem (involving the real-valued decision variables, for example) to reasonable values. Thereafter, another part of the problem (real-valued decision variables) is solved by keeping the parameters of the first part (integer variables) to the obtained optimal values. This procedure is continued till all components of the overall problem is solved. For obvious reasons, although such a procedure enables classical optimization algorithms to be applied, the final obtained solutions may not necessarily correspond to the optimal solutions to the overall problem. Moreover, such decomposition techniques 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. Besides, there are other generic difficulties which we describe in the next paragraph.

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 (Chankong and Haimes, 1983; Ehrgott, 2000; Miettinen, 1999). 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). If an optimization problem involves an objective which is of different type than that of the rest of the objectives, that objective is converted into the other type, so that all objectives are of the identical type. For example, a simple way to convert a maximization problem into a minimization problem is to use the inverse function. Although 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 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 more flexible than the classical methods and allows to alleviate all difficulties mentioned above. Particularly, the proposed technique – 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 type easily.

NSGA-II is applied to increasingly more complex versions of the gearbox design problem which has been solved in the literature. Starting from the two-objective, 10-variable, real-parameter gearbox optimization problem (which has also been solved in Jain and Agogino, 1990), 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. The success of NSGA-II's application in the gearbox design problem suggests its immediate application to similar or more complicated and challenging engineering design problems.

In the reminder of the paper, we briefly outline the multi-speed gearbox design problem, stating the mathematical formulation of the problem. Difficulties faced by classical optimization methods are discussed next. Thereafter, a brief introduction to multi-objective evolutionary optimization techniques is highlighted. A specific optimization technique (NSGA-II) is described next. Thereafter, systematic applications of NSGA-II to five different gearbox design problems are demonstrated. Finally, conclusions of the present study are presented.

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. Usually, the output

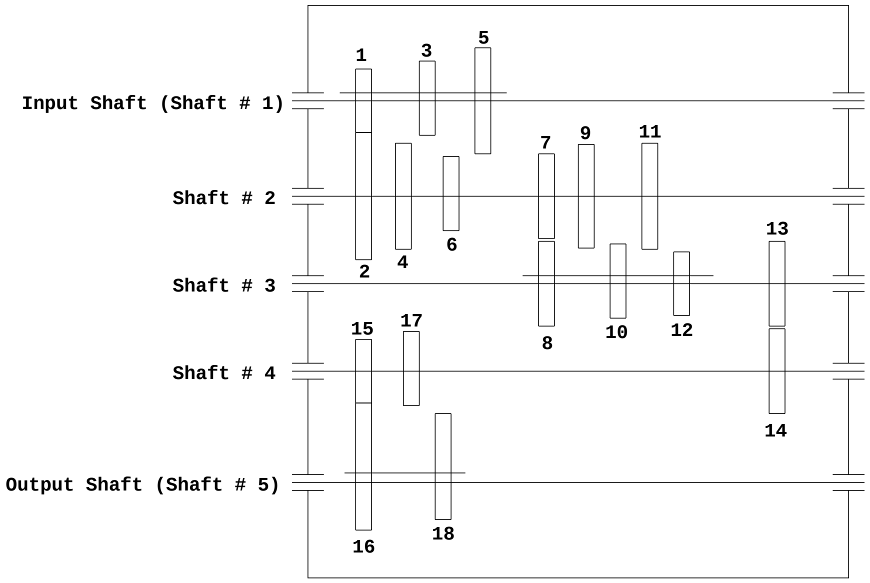

Figure 1: Schematic of a gear train.

speeds are desired to vary in a geometric progression. For example, with an input shaft 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, module etc.) 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 to 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 (Miettinen, 1999). The presence of such multiple Pareto-optimal solutions poses a difficulty in decision-making to a designer. Ironically, each Pareto-optimal solution is different and produces a different trade-off between conflicting objectives. In such situations, the difficulty arises in choosing a particular solution. One approach to solve this problem is to first find as many such Pareto-optimal solutions as possible before making any particular choice on importance of one objective over the other. Once a set of Pareto-optimal solutions is found, 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 pragmatic in the presence of multiple trade-off solutions, rather than the use of such information in arriving at a suitable weight vector before the optimization procedure and in finding the corresponding optimal solution.

In this paper, we concentrate on achieving the first task of finding multiple Pareto-optimal solutions to the gearbox design problem in the presence of more than one objective. 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 G 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 in this direction (Jain and Agogino, 1990; Rao and Eslampour, 1986) have considered two objectives – maximizing the power delivered f_1 and minimizing the total volume of gear material f_2. It is intuitive that a gearbox which maximizes the power delivered by a gearbox is likely to have a large volume, thereby contradicting the second objective. Thus, the consideration of these two conflicting objectives of design is likely to produce a set of Pareto-optimal solutions in the objective space.

There also exist a number of constraints associated to 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. The bending and wear stresses developed in the i-th gear-pair can be calculated as follows (Jain and Agogino, 1990):

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

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

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 in wheel to that in pinion (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 $m(n_{w_i} + n_{p_i})/2$) 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:

$$y_i = 0.52 \left(1 + \frac{20}{n_{w_i}}\right), \tag{3}$$

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.

In addition, there are variable bounds which are usually specified by the user for each decision variable. The complete optimization problem is given below:

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

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 foul 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:

$$n_1 + n_2 = n_3 + n_4 = n_5 + n_6, \quad \text{(between shaft 1 and 2)} \quad (5)$$
$$n_7 + n_8 = n_9 + n_{10} = n_{11} + n_{12}, \quad \text{(between shaft 2 and 3)} \quad (6)$$
$$n_{15} + n_{16} = n_{17} + n_{18}, \quad \text{(between shaft 4 and 5)} \quad (7)$$
$$n_2 \le n_7 + n_8, \quad (8)$$
$$n_4 \le n_7 + n_8, \quad (9)$$
$$n_6 \le n_7 + n_8, \quad (10)$$
$$n_7 \le n_1 + n_2, \quad (11)$$
$$n_9 \le n_1 + n_2, \quad (12)$$
$$n_{11} \le n_1 + n_2, \quad (13)$$
$$n_8 \le n_{13} + n_{14}, \quad (14)$$
$$n_{10} \le n_{13} + n_{14}, \quad (15)$$
$$n_{12} \le n_{13} + n_{14}, \quad (16)$$
$$n_{13} \le n_7 + n_8, \quad (17)$$
$$n_{14} \le n_{15} + n_{16}, \quad (18)$$

$$
\begin{aligned}
n_{15} &\leq n_{13} + n_{14}, & (19)\\
n_{17} &\leq n_{13} + n_{14}, & (20)\\
|\Omega_k - \Omega_k^I| &\leq \epsilon\Omega_k^I, \quad k = 1, 2, \ldots, K, & (21)\\
\frac{n_i w}{n_{p_i}}, &\leq r^{\max} \quad \text{for } i = 1, 2, \ldots, G, & (22)\\
n_{w_i} &\geq n^{(L)}, \quad \text{for } i = 1, 2, \ldots, G, & (23)\\
n_{p_i} &\geq n^{(L)}, \quad \text{for } i = 1, 2, \ldots, G, & (24)\\
n_i &\in \mathcal{I}, \quad \text{for } i = 1, 2, \ldots, 2G. & (25)
\end{aligned}
$$

In a gear-pair, n_{w_i} is assigned the larger number of teeth between the mating gears. Equations 5 to 7 take care of the center distance constraint, as mentioned above. Inequalities 8 to 20 take care of the gear-shaft clearances. Inequality 21 must be satisfied for all output shaft speeds (Jain and Agogino, 1990). 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 some 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 5 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 5 linear equality constraints and 96 nonlinear inequality constraints.

2.3 Past Studies and Their Shortcomings

There have been some efforts to solve this complex optimal gearbox design problem, involving multiple objectives, many nonlinear constraints, and mixed type of variables (discrete, integer, and real-valued). However, such studies (Jain and Agogino, 1990; Rao and Eslampour, 1986) have decomposed the overall problem into tractable subproblems to make them suitable to be solved by classical methods. Essentially, the following decomposition is used:

1. The design problem having gear teeth as the only design variables is 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. The outcome of this optimal design procedure is kept fixed for the next subproblem.

2. Thereafter, a gearbox with gear thickness 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 gear teeth and module are kept fixed.

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. It is clear that without the consideration of the gear thickness and/or module, more important objectives such as the power delivered by the gearbox and overall gearbox weight or size are meaningless.

Having obtained the optimal number of teeth for gears (which incidentally also minimized the error in output speed deviation from ideal in Jain and Agogino (1990)), the second subproblem focused on the mechanical strength considerations and optimized two important objectives (power delivered and the overall volume of the gearbox). Interestingly, the gear teeth values obtained by optimizing a completely different set of objectives were adopted in optimizing the second subproblem. This way, the possibility of using mixed types of variables (integer and real-valued) in one subproblem was avoided. This also enabled the use of a classical optimization method (monotonicity based technique) to arrive at the Pareto-optimal solutions to the second subproblem. It is important to note that there are at least two difficulties with the above decomposition 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 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.

With more than one objectives, the classical approaches pose another difficulty. In terms of handling multiple objectives, classical methods require all objectives to be of the same type. Since the second subproblem of the above gearbox design problem 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 have 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, such as a repetitive application of the weighted-sum approach as used in Jain and Agogino (1990), 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 (Deb, 2001) discusses a number of MOEAs and presents a number of application case studies. Figure 2 shows a schematic diagram of the two-step procedure for a generic multi-criterion optimal design procedure.

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 in Step 1. Since an MOEA works with a population of solutions in each iteration, a number of such solutions may be possible to obtain in one simulation run. Thereafter, in Step 2 of the ideal procedure, the obtained solutions are

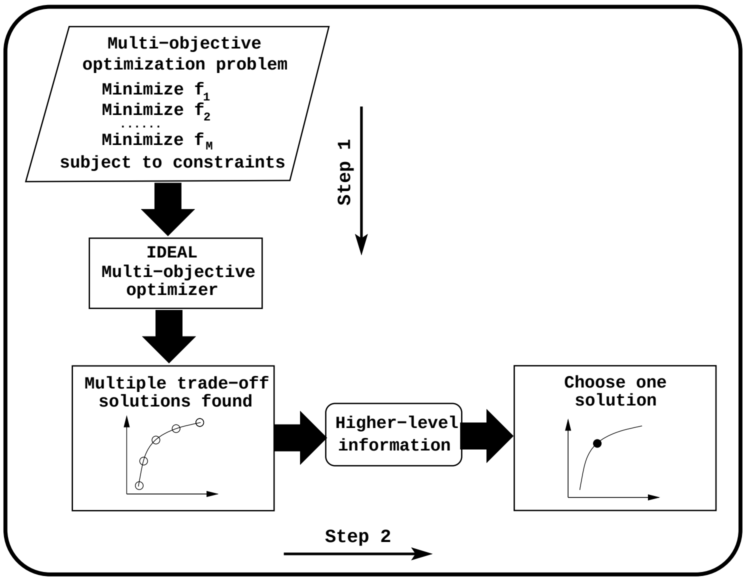

Figure 2: Schematic of an ideal multi-objective optimization procedure.

evaluated and a particular solution is chosen. Although the problem information (such as weight-vector information) may have to be used in Step 2, 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 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 Deb et al. (2000). 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. In the tournament between a pair of solutions, the solution with a better non-domination level wins. 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. When $N/2$ tournaments are played, the population is shuffled in a different random order and the whole procedure is repeated one more time to create a *mating pool* (with the winners of each tournament) of N solutions.

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, thereby allowing $100(1-p_c)\%$ recombinations to make offspring variables exactly the same as that in parents. For handling real-valued decision variables, recombination is performed variable-wise using the SBX operator (Deb and Agrawal, 1995). 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:

$$x_i^{(1,t+1)} = 0.5\left[(1+\beta_{q_i})x_i^{(1,t)} + (1-\beta_{q_i})x_i^{(2,t)}\right], \tag{26}$$

$$x_i^{(2,t+1)} = 0.5\left[(1-\beta_{q_i})x_i^{(1,t)} + (1+\beta_{q_i})x_i^{(2,t)}\right]. \tag{27}$$

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 $u_i \in [0,1]$:

$$\beta_{q_i} = \begin{cases} (2u_i)^{\frac{1}{\eta_c+1}}, & \text{if } u_i \le 0.5; \\ \left(\frac{1}{2(1-u_i)}\right)^{\frac{1}{\eta_c+1}}, & \text{otherwise.} \end{cases} \tag{28}$$

The parameter η_c is the distribution index of the SBX operator. This parameter enables the user to set the extent of search around the parents. If a large η_c value is used, the reduced extent of search results. For handling bounded variables, Equation 28 is slightly modified (Deb and Goyal, 1998), 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 (Goldberg, 1989), 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:

$$y_i^{(1,t+1)} = x_i^{(1,t+1)} + (x_i^{(U)} - x_i^{(L)})\bar{\delta}_i, \tag{29}$$

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:

$$\bar{\delta}_i = \begin{cases} (2r_i)^{1/(\eta_m+1)} - 1, & \text{if } r_i < 0.5, \\ 1 - [2(1-r_i)]^{1/(\eta_m+1)}, & \text{if } r_i \ge 0.5. \end{cases} \tag{30}$$

The parameter η_m has a similar effect to that of η_c. A similar procedure is also followed on the second child created by the crossover operator. For variables coded in binary strings, a bit-wise mutation, which flips a bit with the probability p_m is used.

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. At this stage, only the required number of vacant population slots are filled from the final acceptable subpopulation using the crowding distance values. This *elite-preserving* mechanism allows preservation of good solutions from both parent and offspring populations. This completes one *generation* of the NSGA-II procedure. The crowding distance computation and elite-preserving operation are computationally less expensive than the non-dominated sorting operation, thereby amounting 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. In all simulations presented in this paper, we continue till a pre-specified number of generations are elapsed.

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, a simple yet efficient procedure is adopted here. Instead of using any penalty function to handle the constraints (which introduces an additional difficulty of choosing an appropriate set of penalty parameters) or having to change the above procedure drastically, we simply use a modified definition of the *domination* principle:

Definition 1 *A solution $\mathbf{x}^{(i)}$ is said to 'constrain-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:

1. Any infeasible solution is dominated by any feasible solution.
2. 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.
3. 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. These parameters are kept the same for all cases. The remaining parameters are modified from one case to another and are mentioned in the respective cases.

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 Jain and Agogino (1990). 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.

An interesting feature of NSGA-II is that it requires the user to specify lower and upper bounds of each decision variables, within which the search would be restricted. In this gearbox design problem, all variable bounds are, thus, handled in a natural way and 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 9 bending stress constraints, we can use the following 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.

Figure 3 shows 10,000 random *feasible* solutions (obtained from approximately 1,60,000 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 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 and on an average may have contained only about 6 feasible solutions. 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 Jain and Agogino (1990) 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

[1]It is important to mention that for solution 3 the volume reported in Jain and Agogino (1990) was 1076.77 cc. However, our calculation finds the volume to be somewhat different (1082.14 cc).

It is once again clear from the figure that NSGA-II is able to converge on or very close to the true Pareto-optimal front. Moreover, NSGA-II is also able to maintain a good distribution of 100 solutions over the entire Pareto-optimal front. Although the variable p is varied in $[1, 50]$ kW, no

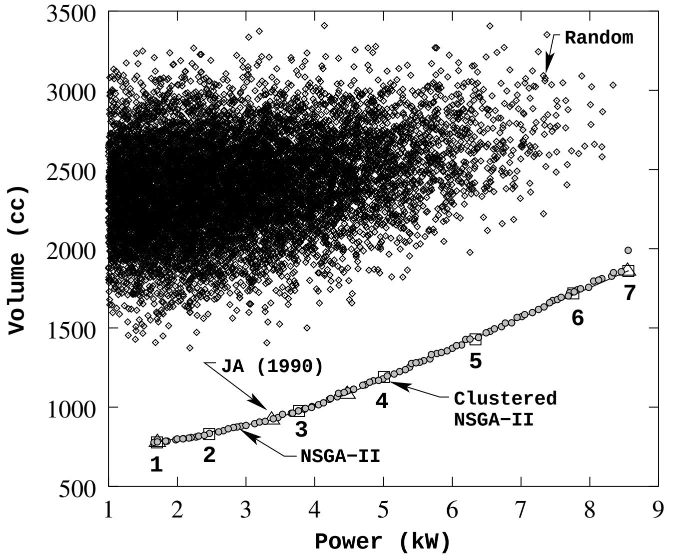

Figure 3: A random set of 10,000 feasible solutions, four Pareto-optimal solutions obtained in Jain and Agogino (1990), and 100 non-dominated solutions obtained by NSGA-II are shown. Only thickness of 18 gears are varied here.

solution beyond $p = 8.56$ kW is found to be feasible in the obtained non-dominated front.

We continue the NSGA-II run for 3,000 generations and obtain 100 non-dominated solutions which are closer to the non-dominated front reported in Jain and Agogino (1990). Of these 100 solutions, seven well-dispersed solutions are shown in Table 1 and are marked in Figure 3 with boxes. Some of these solutions are exactly the same as those found in Jain and Agogino (1990).

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

Sol.	t_1	t_2	t_3	t_4	t_5	t_6	t_7	t_8	t_9	p	f_1	f_2
Id.	(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

4.2 Case II: Varying Gear Module and Thickness

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 for reasons mentioned earlier there exist no results using any classical method to any of the following versions Cases II to V) of the gearbox design problem.

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. A single-point crossover with probability of 0.9 and a bit-wise mutation with probability of 0.2 are used for the binary substring. For real-parameter decision variables, identical operators and parameters as those used in the previous case are adopted here.

Figure 4 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. Four solutions of Case I reported in Jain and Agogino (1990) are also marked for comparison. 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. Although the upper limit of p is 50 kW, no solution having delivered power greater than 18.28 kW is found to be feasible and non-dominated.

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

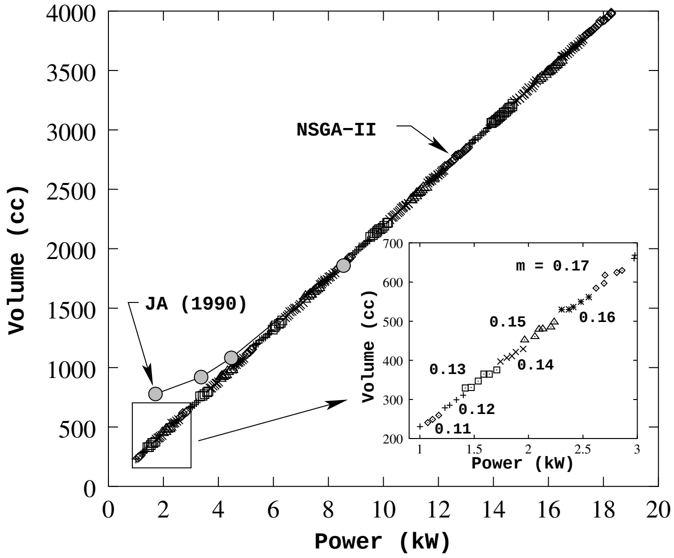

Figure 4: NSGA-II solutions for the case of module being an additional variable are shown against four Pareto-optimal solutions of Case I (as reported in Jain and Agogino (1990)). The inset figure shows that obtained solutions with increasing power requires monotonically increasing modules.

front corresponds to a module of 0.41 cm (the 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 5 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 6 that a monotonic increase in the module is needed for larger power transmission. Another interesting aspect is that the rate of increase in the value of module requirement reduces for large power transmissions.

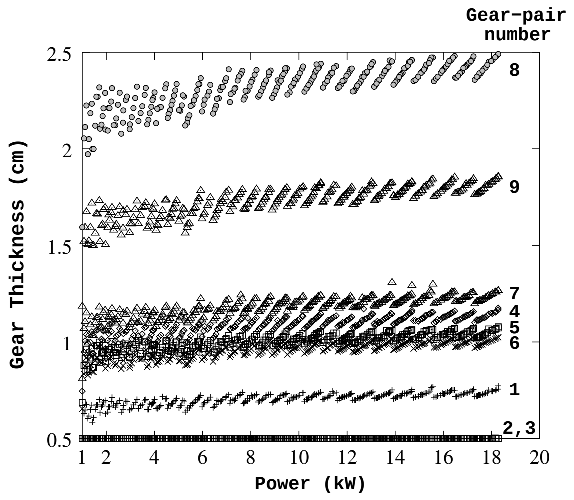

Figure 5: Variation of gear thickness in obtained solutions.

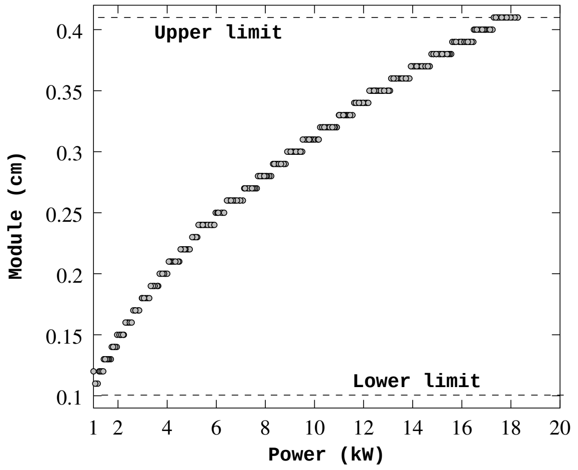

Figure 6: Variation of gear module in obtained solutions.

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:

$$p = A\sigma_{b,w}m^2, \tag{31}$$

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 Jain and Agogino (1990) that the bending stress constraints are always critical than the wear stress constraints. Figure 7 plots the bending stress σ_b for all obtained non-dominated solutions. It

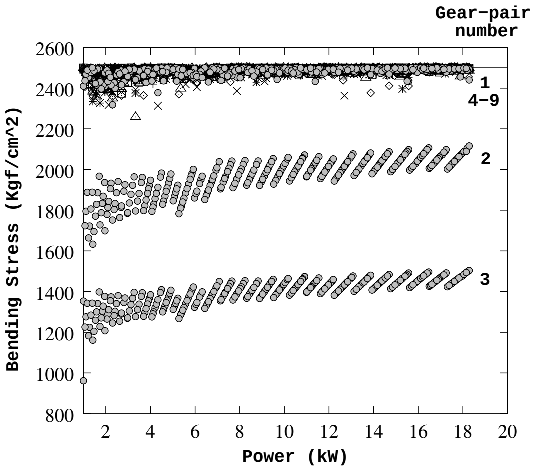

Figure 7: The bending stress of NSGA-II solutions are shown against delivered power.

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 5, 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 7 that the critical constraints are the seven inequality bending stress constraints 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 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 31, leading to $m^2 \propto p$. The exponent of the polynomial variation of module with power as observed in Figure 6 is close to 0.5, thereby supporting the above argument[2] Therefore, the above equation yields $f_1 \propto p$. It is also interesting to observe from Figure 4 that the linear Pareto-optimal front will pass through the origin, if extrapolated, thereby supporting our argument above. Since, we have kept a lower bound on p, NSGA-II could not find any solution with p smaller than that lower bound ($p^{(L)} = 1$ kW).

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 a given solution and are flexible in representing mixed types of variable, such problems are easy to handle.

4.3 Case III: Varying Gear Teeth and Thickness

To make the gearbox design problem more flexible, we now 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 5 to 7, we use following two strategies with NSGA-II:

Variable elimination: We use the first three linear equality constraints to eliminate five decision variables. For example, the first equation can be used to eliminate two variables as follows:

$$
\begin{aligned}
n_4 &= n_1 + n_2 - n_3, && (32) \\
n_6 &= n_1 + n_2 - n_5. && (33)
\end{aligned}
$$

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 $n_5 > n_3 > n_1$. This way several permutations of these gear sizes will not be considered 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:

$$
\begin{aligned}
n_3 &= n_1 + \Delta n_{1,3}, && (34) \\
n_5 &= n_1 + \Delta n_{1,3} + \Delta n_{3,5}. && (35)
\end{aligned}
$$

[2]It is also interesting to note that for a single bevel or hypoid gear-pair, Krenzer and Hotchkiss (1986) have also suggested the use of gear diameters increasing polynomially with transmitted torque having an exponent of 0.34.

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 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 1 and n_i variables to vary in $[18, 81]$ in steps of 1. 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 $\eta_m = 50$. Figure 8 shows the obtained non-dominated solutions for $\epsilon = 0.05$ and 0.10. It is clear from the figure that if a large error in the output shaft speeds from the ideal is permit-

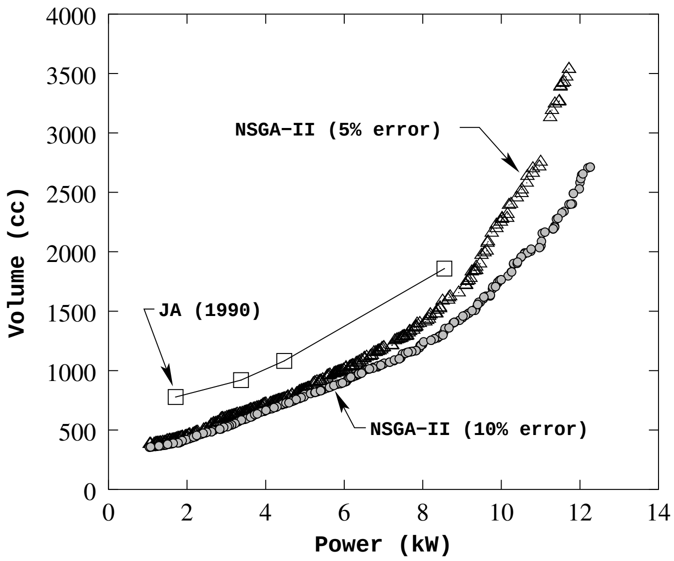

Figure 8: NSGA-II solutions are shown with two limiting errors in output shaft speeds. Gear teeth and thickness are varied here.

ted, better non-dominated solutions are expected. The figure also marks the four Pareto-optimal solutions reported elsewhere (Jain and Agogino, 1990) 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. For example, in the case of $\epsilon = 0.10$, the obtained non-dominated solutions have a span on the delivered power in $[1.04, 12.26]$ kW, compared to $[1.66, 8.56]$ kW obtained without the gear teeth being considered as decision variables (Case I).

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 large delivered power, thickness and teeth in gears are large. Table 3 shows the corresponding gear details for $\epsilon = 0.05$. A similar trend in the gear sizes, as observed in the $\epsilon = 0.10$ case, is 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 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 9 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. For example, solutions between 9.75% to 10% are grouped together, and

Table 2: Gear details of four selected solutions in Case III (with 10% error)

Sol. No.	Power (kW)	Volume (cc)		Gear-Pair Number, i								
				1	2	3	4	5	6	7	8	9
1	1.040	355.03	n_{p_i}	18	27	18	18	19	20	18	18	23
			n_{w_i}	37	28	37	23	22	21	32	33	28
			t_i	0.500	0.500	0.500	0.500	0.500	0.500	0.501	0.506	0.500
2	5.503	846.52	n_{p_i}	21	28	18	28	26	25	19	18	24
			n_{w_i}	39	32	42	29	31	32	38	43	37
			t_i	0.501	0.500	0.500	0.552	0.510	0.500	0.920	1.884	1.288
3	9.008	1443.68	n_{p_i}	22	28	18	29	27	26	20	18	24
			n_{w_i}	39	33	43	36	38	39	39	55	49
			t_i	0.751	0.500	0.500	0.684	0.663	0.656	1.063	2.456	1.782
4	12.262	2712.36	n_{p_i}	29	31	19	37	35	33	30	23	31
			n_{w_i}	44	42	54	38	40	42	47	80	72
			t_i	1.013	0.500	0.500	0.794	0.747	0.721	1.079	2.500	1.816

Table 3: Gear details of four selected solutions in Case III (with 5% error)

Sol. No.	Power (kW)	Volume (cc)		Gear-Pair Number, i								
				1	2	3	4	5	6	7	8	9
1	1.050	381.95	n_{p_i}	18	28	18	19	21	22	18	18	23
			n_{w_i}	38	28	38	27	25	24	28	35	30
			t_i	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500
2	5.547	941.52	n_{p_i}	19	30	19	28	30	29	22	22	18
			n_{w_i}	41	30	41	33	31	32	29	30	47
			t_i	0.500	0.500	0.501	0.642	0.592	0.553	1.100	1.802	1.295
3	9.094	1718.20	n_{p_i}	25	30	18	32	30	28	22	20	27
			n_{w_i}	43	38	50	32	34	36	35	65	58
			t_i	0.597	0.502	0.500	0.726	0.698	0.674	1.245	2.300	1.649
4	11.720	3537.46	n_{p_i}	27	39	24	37	40	39	36	27	36
			n_{w_i}	54	42	57	44	41	42	45	81	72
			t_i	0.934	1.251	1.571	1.069	0.947	0.910	1.304	2.491	1.821

so on. 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.

In the previous two cases with a specific gear tooth values used in Case I, it was observed that the bending stress constraints are more critical than the wear stress constraints. However, in this case (when the gear teeth are kept as decision variables), the bending stress need not always be critical. To investigate which constraint is critical, we have plotted B_i/C_i versus p in Figure 10 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 $p = C_i t_i$ (Jain and Agogino, 1990)). It can be seen from the figure that the ratio B_i/C_i is more than one for small values of p and is less than one for large values of p. This means that for small values of p, the wear stress constraints are more critical than bending stress constraints. Since we have kept a lower bound on the gear thickness, it is also intuitive that the gearboxes with small delivered power are over-designed for bending considerations. Gear thickness values smaller than this lower bound would be adequate to deliver the same power. However, the optimality

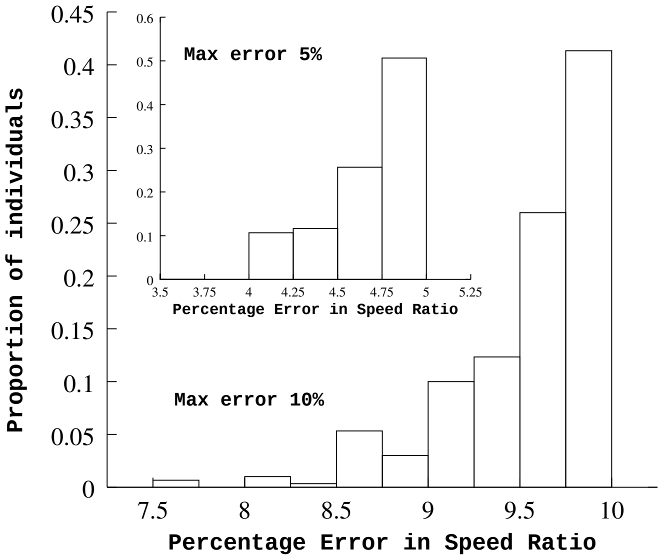

Figure 9: Proportion of obtained solutions having different maximum percentage error in speed ratios is shown for two cases of optimization (with 5% and 10% limiting error). Most solutions exhibit error in output shaft speed close to their permissible limits.

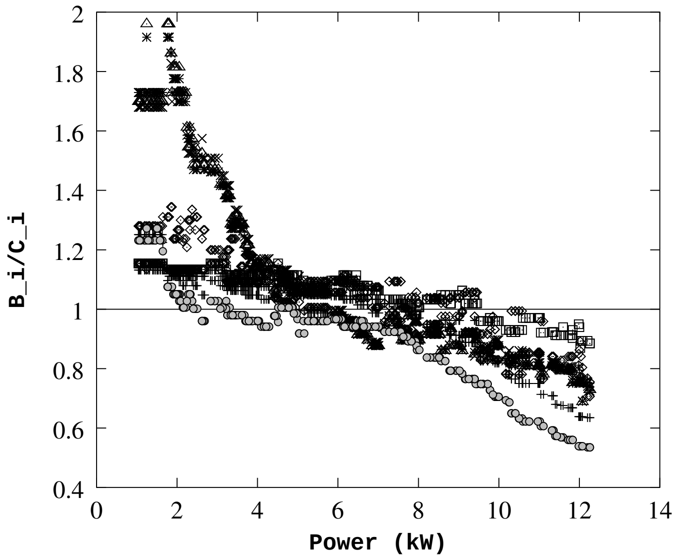

Figure 10: The ratio B_i/C_i for all nine gear-pairs observed in different obtained solutions.

considerations require that such gears satisfy the wear stress constraints with the equality sign.

On the other hand, for solutions with a large value of p, bending stress constraints are more critical. With large power requirement, the gear sizes are also large and are critical for failure due to bending than wear. For such gears, the optimality condition requires that the bending stress constraints be satisfied with the equality sign.

In the former case (for $B_i/C_i > 1$), we observe the following relationship between power and decision variables from equation 2:

$$p = \alpha_i t_i \frac{n_{w_i} n_{p_i}^2}{n_{p_i} + n_{w_i}}, \tag{36}$$

where α_i is a value independent of the decision variables. For the latter case (for $B_i/C_i < 1$), we use equation 1:

$$p = \beta_i t_i \frac{n_{p_i}^2}{n_{w_i}} \left(1 + \frac{20}{n_{w_i}}\right), \tag{37}$$

where β_i is a value independent of the decision variables. For a fixed module, the volume (f_2) varies as follows:

$$f_2 = \gamma \sum_{i=1}^{9} (n_{p_i}^2 + n_{w_i}^2) t_i, \tag{38}$$

where γ is a constant. Now, it can be observed from the above three equations that the rate of increase in f_2 with respect to p when all gear teeth and thicknesses are increased is more in the latter case (for larger p values). This is exactly what we observe in the obtained non-dominated NSGA-II solutions in Figure 8.

4.4 Case IV: Varying Gear Module, Teeth and Thickness

Next, we make the gearbox design problem even more flexible by keeping 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. However, five equality constraints are used to reduce the total number of variables to 24. An $\epsilon = 0.10$ is used in this study. Rest all variables and parameters are kept the same as those used in the previous case. Figure 11 shows the 300 non-dominated solutions obtained using NSGA-II after 10,000 generations. On the same computer mentioned earlier, each NSGA-II run took an average of 660 seconds. As in Case II, a monotonically increasing magnitude of module is observed for solutions having increasing requirement in delivered power. Once again, the solution with the largest power delivered requires the maximum permissible module of 0.41 cm. Figure 11 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.

Figure 12 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. Thereafter, the module remains fixed to the upper limit. For the range of non-dominated solutions in which the module is different, the corresponding non-dominated front is *linear*, as evident from Figure 11 and as explained in Case II. Thereafter, when the module remains fixed, 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 gets *dominated* by gearboxes obtained using a larger module and is, thus, can be avoided in practice.

To demonstrate that the nonlinear part of the non-dominated region arises due to the module getting fixed to its upper bound, 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 11. It is evident that now the obtained

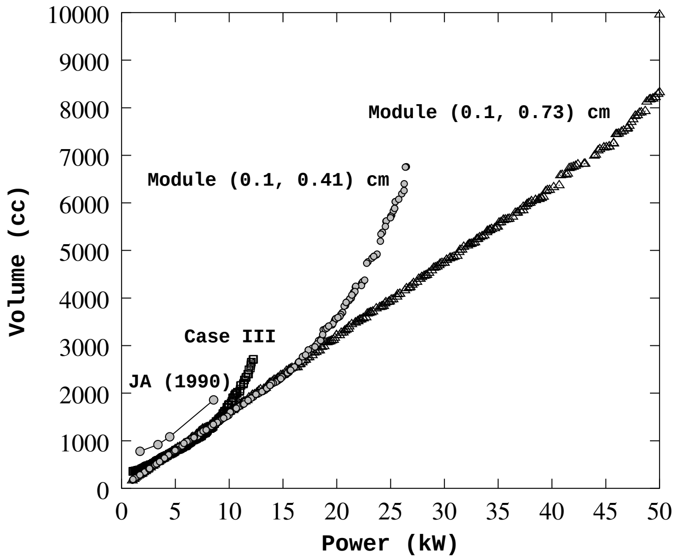

Figure 11: NSGA-II solutions for varying gear module, teeth, and thickness are shown. Non-dominated fronts for two different ranges of modules are compared with that obtained in Case III (with a fixed module of 0.28 cm).

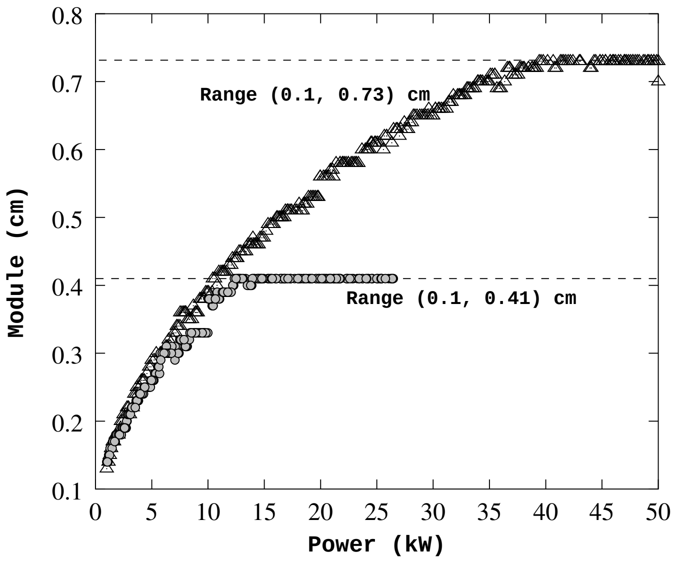

Figure 12: Module m obtained in different non-dominated solutions for two cases.

front remains linear till the module of the gearbox becomes equal to the chosen upper limit. Figure 12 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).

Figures 13 and 14 show that the variation of number of teeth and thickness in all gears with delivered power for the $m^{(U)} = 0.41$ cm case. Two distinct phases are again observed in these figures. For gearboxes with a power capacity of about 15 kW, the gear teeth and gear thickness in all pairs except 8 and 9 are more or less constant. During this part of the obtained non-dominated front, solutions mainly vary due to the variation in the module (see Figure 12) and as discussed in Case II the relationship between two objectives (volume and power) is linear. However, in the remaining part of the obtained non-dominated front, the module is fixed to the upper limit and the number of teeth increases. These conditions are similar to that discussed in Case III, and the relationship between volume and power is found to be nonlinear in a manner similar to that observed in Case III. Interestingly, in the nonlinear part of the non-dominated front, the main variation in gearboxes delivering large powers happen mainly due to variations in thickness and number of teeth in gear-pairs 7, 8 and 9, and not due to an increase in the module.

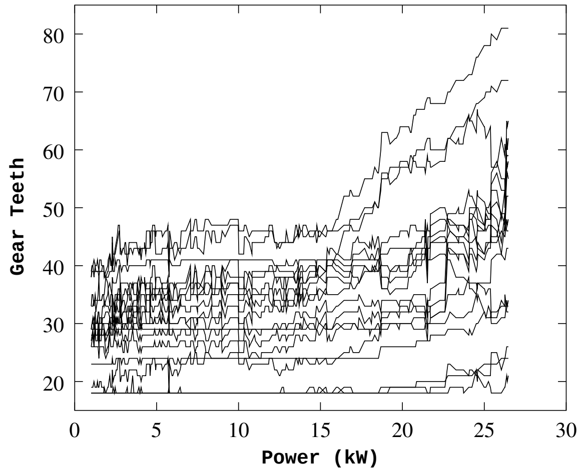

Figure 13: Variation of gear teeth in obtained solutions. Values are joined with a line for clarity.

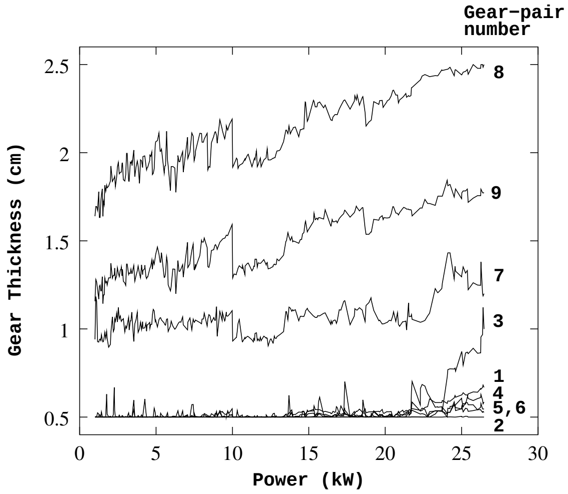

Figure 14: Variation of gear thickness in obtained solutions. Values are joined with a line for clarity.

Table 4 shows the gear details of five well-dispersed non-dominated solutions. The first three

Table 4: Gear details of five selected solutions in Case IV (with $\epsilon = 0.10$)

Sol. No.	Power (kW)	Volume (cc)	Module (cm)		Gear Pair Number, i: 1	2	3	4	5	6	7	8	9
1	1.006	169.55	0.14	n_{p_i}	19	28	18	26	27	26	18	18	23
				n_{w_i}	38	29	39	29	28	29	33	38	33
				t_i	0.500	0.500	0.501	0.500	0.500	0.500	0.941	1.640	1.157
2	8.528	1343.53	0.33	n_{p_i}	24	29	18	31	29	27	18	18	24
				n_{w_i}	41	36	47	31	33	35	35	46	40
				t_i	0.500	0.500	0.502	0.500	0.500	0.500	1.075	1.917	1.375
3	15.306	2390.97	0.41	n_{p_i}	24	29	18	34	32	30	19	18	24
				n_{w_i}	41	36	47	35	37	39	39	45	39
				t_i	0.502	0.500	0.544	0.535	0.522	0.501	1.073	2.297	1.610
4	23.271	4839.78	0.41	n_{p_i}	35	31	19	42	44	42	22	21	29
				n_{w_i}	45	49	61	47	45	47	38	72	64
				t_i	0.535	0.500	0.556	0.583	0.536	0.517	1.229	2.437	1.701
5	26.467	6758.97	0.41	n_{p_i}	32	34	21	58	55	52	26	24	33
				n_{w_i}	48	46	59	59	62	65	43	81	72
				t_i	0.669	0.500	1.000	0.585	0.547	0.527	1.201	2.500	1.773

solutions lie on the linear portion of the obtained non-dominated front. It is clear that the gear thicknesses for the first six gear-pairs are more or less the same in the first three solutions, as are also evident from Figure 14. Once again, the solutions exhibit different gear teeth combinations for achieving a maximum of 10% deviation in the output speeds from the desired values.

The variety in gear teeth combinations in the obtained solutions results mainly due to the flexibility in allowing 10% variation in output speeds from the desired speeds. When the same NSGA-II is run for a maximum of 1% deviation in the output speeds from the desired speeds, all 300 non-dominated solutions have identical gear teeth combination, shown in the following table:

	Gear-Pair Number i								
	1	2	3	4	5	6	7	8	9
n_{p_i}	37	48	29	60	57	53	23	43	54
n_{w_i}	68	57	76	61	64	68	68	72	61

The variability in solutions arise mainly due to a $m \propto \sqrt{p}$ change in module and due to changes in thickness of three gear-pairs (7, 8 and 9).

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:

$$f_3 = \frac{m}{2}\left[(n_1 + n_2) + (n_7 + n_8) + (n_{13} + n_{14}) + (n_{15} + n_{16})\right]. \tag{39}$$

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, for the first time, we show simulation results for the gearbox optimization problem with three of the four objectives. However, it can also be mentioned that the fourth objective (minimizing ϵ) is handled here by limiting its value within a specific value ($\epsilon \leq 0.1$ is used here).

NSGA-II with 300 population members takes an average of 700 seconds to run 10,000 generations on the same machine mentioned before. Although the solutions found are satisfactory, in order to achieve a better convergence, we have continued the NSGA-II run for 100,000 generations. Figure 15 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 (partially linear and partially nonlinear) in power delivered 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 dimensionality of the obtained non-dominated front in the three-dimensional objective 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.

To compare the differences in the obtained solutions between the two-objective and three-objectives cases, we plot module and thickness of three specific gear-pairs in Figures 16 and 17. It is clear that NSGA-II has found smaller module values in the case of an additional objective of

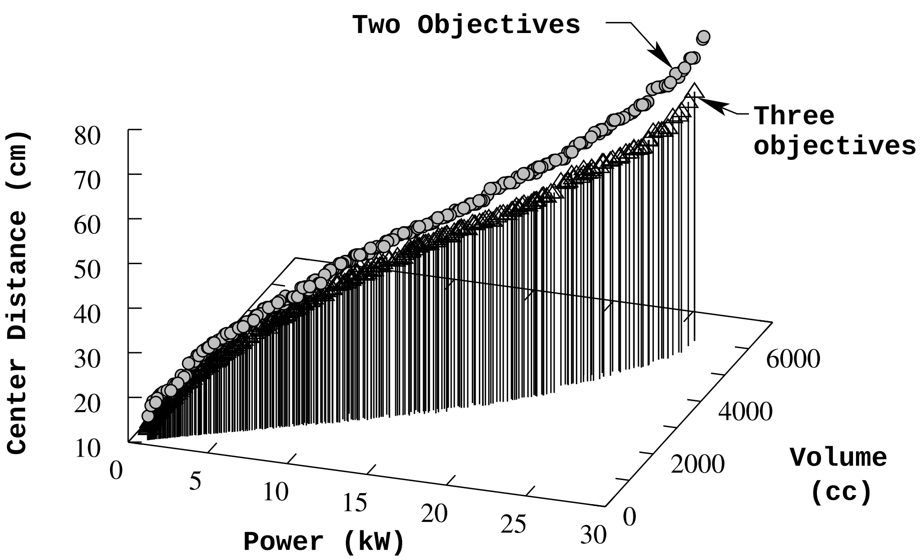

Figure 15: NSGA-II solutions for an additional objective of minimizing center distance are compared with NSGA-II solutions with two objectives (maximizing power and minimizing volume).

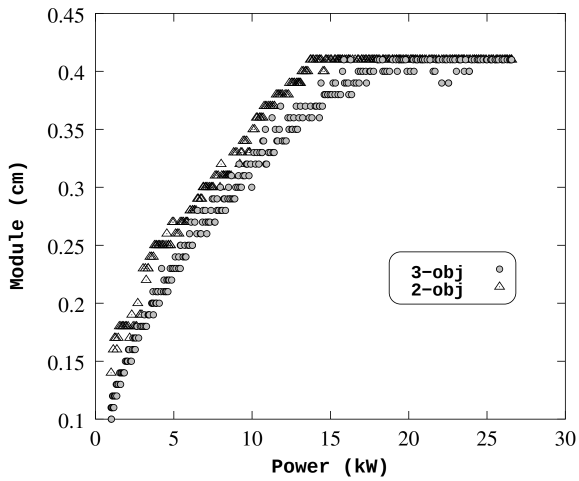

Figure 16: Comparison of gear modules in three and two-objective cases.

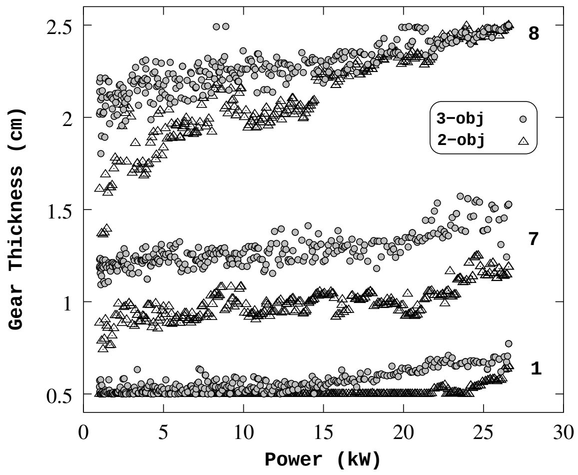

Figure 17: Comparison of gear thicknesses (gear-pairs 1, 7, and 8) in three and two-objective cases.

minimizing the center distance. Also, in this case the gears are thicker (although only three gear-pairs are shown in the figure for clarity, all gear-pairs exhibit the same trend). It is also observed that the gear teeth obtained in the three-objective case is almost similar to those observed in the two-objective case. These properties of the obtained solutions can be explained by analyzing equation 39. Since the center distance between the input and output shafts have no consideration for the gear thickness, NSGA-II has optimized the center distance by reducing the module. To maintain the volume-power optimality, NSGA-II finds gearboxes with larger thickness values (see Figure 17). Thus, the consideration of the third objective is found to be beneficial in terms of reducing the overall size of the gearbox.

It should be noted that the computational overhead and the complexities in handling additional objectives are not significant in NSGA-II.

4.6 Case VI: Four Objectives

Finally, we include an additional 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 18 shows the obtained non-dominated front on the three-objective space. The following interesting aspects can be observed from the figure:

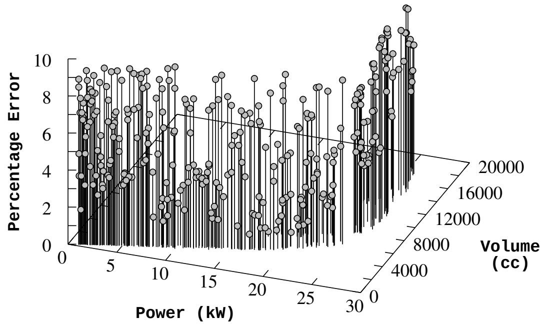

Figure 18: Non-dominated solutions for four objectives.

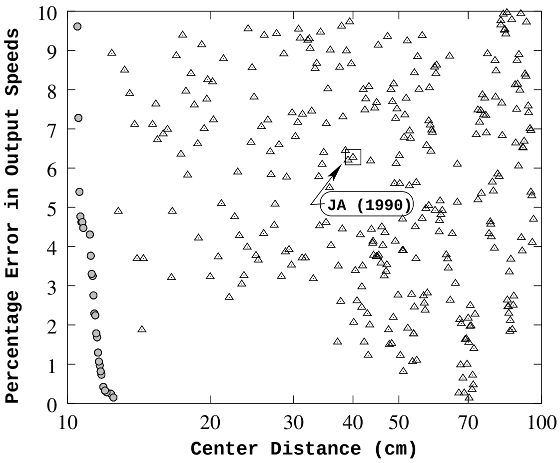

Figure 19: Trade-off between ϵ and center distance for two-objective NSGA-II, four-objective NSGA-II, and Jain and Agogino's (1990) study. The x-axis is plotted in logarithmic scale.

1. The resulting non-dominated front is a three-dimensional surface.
2. The projection of the non-dominated solutions on the volume-power space has an identical relationship as that observed in Case V.
3. 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, for a gearbox with a very small maximum output speed error ϵ, there may not be many gear teeth combinations possible, thereby reducing the range of allowable center distance between the input and output shafts.

Figure 19 shows the obtained non-dominated front (shown in circles), when the gearbox design problem is solved only for two objectives: (i) minimize 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 in the error and center distance is clear from the figure. An analysis of the obtained solutions reveal 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 non-dominated solutions (shown in triangles) obtained from the four-objective case are also superimposed in this figure. It is clear that the solutions from the two-objective case dominates that of the four-objective case.

The earlier study (Jain and Agogino, 1990) obtained only one solution (shown with a square) to the two-objective problem without any consideration of mechanical strength constraints. However, the above figure shows that there is trade-off between these two objectives when all other considerations are taken into account and NSGA-II is capable of finding it.

5 Conclusions

The multi-speed gearbox design optimization problem considered in this paper provides adequate challenge to almost every aspect of optimization. The design problem involves more than one conflicting objectives of design, a total of 29 decision variables (integer, discrete, and real-valued decision variables) which all must be varied in an ideal optimization scenario. The problem also contains 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 the more difficult one having 10 variables and 18 inequality constraints and solve them independently. This paper has shown that the application of a multi-objective evolutionary algorithm – NSGA-II – to the more complex decomposed problem has able to find the same Pareto-optimal front as that obtained by classical methods. The only difference is that NSGA-II has found 100 different solutions in one single simulation run, as opposed to the need of a repetitive use of any classical method to find multiple solutions. Having found the proof-of-principle solutions on the base-line version of the gearbox design problem, NSGA-II is then applied to more complex version of the problem by using more parameters as decision variables, finally leading to the complete optimization problem and with four objectives. In all cases, the flexibility and ease of application of NSGA-II are demonstrated and important insights of the optimal designs to the gearbox design problem are discovered:

1. No matter whether the number of teeth in gears are desired to be varied or not, an optimal gearbox requires a larger module for larger delivered power.

2. However, in the event of impracticalities in using a large module, it is better to fix the module to be as large as possible and increase the thickness of gears rather than to increase the number of gear teeth to design gearboxes with large delivered power. In these cases, to gain a unit increase in delivered power, a larger increase in volume of the gearbox than that required in the gearboxes with smaller delivered power will be required.

3. The optimal module requirement varies proportional to the squared root of the desired delivered power.

4. 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 means, including an analysis of the mathematical programming problem describing the optimization problem. This is due to the fact that in order to obtain any such properties, it is first necessary to *find* a set of well-dispersed Pareto-optimal solutions. Since NSGA-II allows a tractable way of achieving a well-distributed set of non-dominated solutions, such systematic applications and an analysis of the obtained results are possible.

The case studies performed in this paper on various complexities of the gearbox design problem are interesting and informative to any practitioner and should motivate other researchers to use NSGA-II in similar or more complex mechanical component design problems.

Acknowledgements

The work is executed under the Research and Development programme of the Ministry of Human Resources and Development (MHRD), Government of India.

References

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

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

Deb, K. (2001). *Multi-objective optimization using evolutionary algorithms.* Chichester, UK: Wiley.

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

Deb, K., Agrawal, S., Pratap, A., Meyarivan, T. (2000). A Fast Elitist Non-dominated sorting genetic algorithm for multi-objective optimization: NSGA-II. *Proceedings of the Parallel Problem Solving from Nature VI Conference*, 16-20 September. (Paris, France), (pp. 849–858).

Deb, K. and Goyal, M. (1998). 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.

Ehrgott, M. (2000). *Multicriteria Optimization.* Berlin: Springer.

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

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

Miettinen, K. (1999). *Nonlinear Multiobjective Optimization.* Boston: Kluwer.

Rao, S. S. (1984). *Optimization: Theory and applications*, New Delhi: Wiley Eastern.

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

Krenzer, T. J. and Hotchkiss, R. G. (1986). Bevel and hypoid gears. In J. E. Shigley and C. R. Mishcke (eds.) *Standard handbook of machine design.* New York: McGraw-Hill. (pp. 34.1–61).