Least sensitive tolerance allocation

Mohamed H. Gadallah

Department of Mechanical Design and Production,
Faculty of Engineering,
Cairo University,
Giza, 12613, Cairo, Egypt
Fax: 202-35673025
E-mail: mohamed@aucegypt.edu

Abstract: A new formulation to the tolerance allocation based on minimum sensitivity is given. Several synthesis methodologies are reviewed to highlight their importance in details. A brief but complete literature review is given and conclusions are drawn. Objectives such as minimum cost functions, minimum sensitivity functions and minimum variance functions are formulated and a heuristic approach is used for optimisation. An example problem is given to illustrate the concept and results indicate that there are alternatives for tolerance allocation problems. Besides, there are margins of tolerance for which the measure of objective functions are least sensitive.

Keywords: tolerance allocation; least sensitivity; heuristics.


Biographical notes: Mohamed H. Gadallah is a Professor of Industrial Engineering and Operations Research, Department of Mechanical Design and Production, Faculty of Engineering, Cairo University. He received his PhD from McMaster University in Canada in 1995. He joined Indiana University – Purdue University and University of Illinois at Chicago in 2002 and 2004 as a Visiting Scholar. A well reputed Quality and Productivity Consultant in Automotive Industry in the USA, Canada, Qatar and Egypt, he spent part of his career at Cairo University, the American University in Cairo and Qatar University. His research interests include approximation modelling and optimisation using ANN and RSM. He is an Associate Editor to TCSME and on the editorial board of IQET.

1 Background

Literature on tolerance allocation is rich and abundant. Since the early ‘90s, several studies using traditional techniques are carried. In this study, a brief but complete review is given to the problem under several headings; these are statistical tolerance techniques, robust optimisation problem, GA-based tolerance allocation, neural network-based tolerance allocation, cost tolerance models and variance reduction techniques. A review is given next.
1.1 Statistical tolerance analysis

Nigam and Turner (1995) described the technique of applying statistical methods to tolerance analysis of assemblies. Various approaches are reviewed especially under the assumption of non-ideal probability distributions.

1.2 Robust optimisation problem

Parkinson et al. (1993) described a general rigorous approach for robust optimal design based on non-linear programming. The method encompasses controllable and non-controllable tolerances and includes both worst-case and statistical analysis. They also presented a general approach for worst case or statistical analysis. The outcome of the methodology can apply the method to find an optimum that will remain feasible when subject to variation, and/or the designer can minimise or constrain the effects of tolerances as one of the objectives or constraints of the design problem. Two main issues are addressed. The first is the issue of design feasibility. In a constrained design space, the effect of variations is to reduce the size of feasible region. We develop procedures to account for tolerances during design optimisation such that the final design remains feasible despite variations in parameters (variables). There are two viewpoints:

1. control variation by minimising sensitivities
2. control variation by trading off controllable tolerances with uncontrollable ones.

Post optimality analysis (POA) refers to investigating the effects of small changes such as tolerances, upon an optimal solution by evaluating derivatives of optimal functions or variables. The analysis assumes the same constraint set will be binding at the nominal and the robust optimum, whereas the two-step method does not. Constraint correlation is considered. If they are positively correlated, they will tend to be feasible at the same time. If they are negatively correlated, one will be feasible while the other will be infeasible.

Ramakrishnam and Rao (1994) investigated the application of robust design based on Taguchi’s (1987) design philosophy to large-scale non-linear programming problems. The formulation utilises the expected value of loss function as objectives. Approximate expressions for the gradients and fast re-analysis techniques are given.

Emch and Parkinson (1994) stressed that there is a need for better analytical tools to help develop robust designs. Only worst case tolerances are induced to introduce the variability. The study points out that despite the importance, variability is minimally addressed in the design phase (usually through safety factors). The reason is due to the lack of good design tools to incorporate both variability and nominal values. The study incorporates NLP with the model of the design to understand and control variability. The procedure was applied for linear tolerance analysis and extended next to non-linear cases. Feasibility robustness can be developed for worst case tolerances by increasing the value of the constraint functions by the amount of the total constraint variation during optimisation. This has the effect of reducing the size of the feasible region with a degradation in the value of the objective. One assumption behind is that the variation is constant from the nominal to the robust optimum presumes that the two are not very far apart and the derivatives do not change significantly from one to the other. Three methods are used; these are the continuously updated transmitted variation (CUTV),
constant gradient sign (CGS) and second order function approximation (SOFA). The first and second methods will be exact if monotonicity holds. The third method will be exact if monotonicity holds and the functions are quadratic.

Michelena and Agogino (1994) provided a monotonicity analysis-based methodology to facilitate the solution of n-type parameter design problems. Identification of active sets of constraints allows a problem reduction to be used where solution to the original problem is obtained by solving a set of problems with fewer degrees of freedom.

Bates and Wynn (1996) developed a comprehensive strategy for robust design. Jeang (1999) used RSM to analyse data for an experimental model to determine optimal component tolerances in an assembly. RSM is a combination of mathematical and statistical techniques that provide both the optimal tolerances and the critical variables in an assembly.

Das (2000) formulated a derivative-based approximation for evaluating the expected value of the objective function on the non-linear manifold defined by the state equations for the system. The procedure presented bypasses the need for multiple solutions of the state equations and provides a cheaper and more optimisable approximation to the expectation.

1.3 GA-based tolerance allocation

Lee and Johnson (1993) developed a truncated Monte Carlo simulation and GA used for integrated analysis and synthesis. The simulation results indicated better performance and reduction in optimal cost when compared with results published previously. The study also stressed the time savings over regular mathematical programming algorithms. The nature of uncertainty in tolerance problems is formulated as stochastic optimisation and expressed as expected values of the objective and constraint functions. A pdf is defined and a multi-variate integration is carried. Stochastic optimisation methods can be classified into three categories: common non-linear programming with simulation, approximation methods and stochastic quasi-gradient methods. Specific values were given for the precision of coding and fitness of inversion, population size, cross-over probability and mutation probability. How other values affect the solutions is not a clear question. Several sample sizes ranging from 10 to 30,000 were taken.

Several tests are performed for various sampling numbers. The results indicate that the standard deviation is inversely proportional to the square root of the sample numbers. The relation between the computation time and number of sampling points was studied after each experimental run. The study concluded that there is no guarantee that better performance will always be obtained in absence of theoretical evaluation.

1.4 Neural network-based tolerance allocation

Kopardekar and Anand (1995) presented a neural network based approach for the tolerance allocation problem considering machine capabilities and mean shifts. The network is trained using BP method and is used to predict part tolerances. Neural networks are physical cellular systems that can acquire, store and utilise experimental knowledge. The neural network model consists of a number of interconnected processing elements called neurons. The interconnection of neurons is called axons. Similarly, neural networks are systems of simple but interconnected processing elements. Originally, neural networks were developed for both discrete and continuous mapping. A
three-multi-layer fully connected network with BP learning method is used. The network is used for predicting the component tolerances. A (4, 3, 3) was used meaning four input layers, three hidden layers and three output layers. The BP algorithm was trained using ten input patterns and an exhaustive search method was used to find the least cost combination. Several limitations of the approach are listed, these are:

- The approach requires some data with known outputs.
- The approach was not faster than exhaustive search. The viability of the method may become apparent for higher number assemblies.

1.5 Cost tolerance models

Chase et al. (1990) presented a discrete optimisation problem using several methods in comparison with exhaustive search methods. Several tolerance allocation methods are investigated, these include Lagrange multipliers. Other combinatorial process selection methods are also studied; among them exhaustive search, zero-one discrete search method and uni-variate search method. Several cost tolerance are reviewed, these include the reciprocal, reciprocal squared, exponential, reciprocal power, the piece wise linear and the empirical data. Lagrange multipliers has several limitations, these are:

- it cannot treat discontinuous cost tolerance functions, as it needs continuous first derivatives
- it cannot treat cost tolerance functions with preferred tolerance limits
- it has difficulty when applied to assemblies with interdependent tolerance loops or chains (assemblies described with several functional requirements)

The study concluded by a comparative evaluation for the exhaustive search, Lagrange multipliers and combinatorial methods. Zero-one method is claimed to be non-practical for large size problems. The SQP based on non-linear programming can treat multiple functional requirements but cannot guarantee the global minimum. This power is not available for exhaustive and uni-variate search methods.

Ngoi and Min (1999) presented a new approach to optimum tolerancing of components for both uni-lateral and bi-lateral requirements. A model showing the relationship between components is constructed from design. A set of linear equations is formulated based on design and assembly requirements. These equations are then solved to determine the optimal values.

Dong et al. (1994) presented new production cost tolerance models and a hybrid model for tolerance synthesis. The introduced models fit empirical cost tolerance data with minimum errors. The paper focuses on modelling error reduction.

Rajasekera (1995) stressed the fact that computer algorithms be integrated into the design tools to compute the costs at various stages of the design process.

Ann and Seng (1996) presented a simple dimensional chain identification method employing non-linear programming approach. The model construction begins with the first cut and proceeds chronologically to the last cut. In the definition of associated working dimensions/tolerances between any surfaces, there should be a certain
convention to follow. Optivar software makes use of Hooks and Jeeve sub-routines to solve the resulting set of equations. The objective function takes the form of a summation form made of weighing factors. This reflects the possibility of assigning different weights to different tolerances.

Ji et al. (2000) presented a new approach based on fuzzy comprehensive evaluation and GA. The new optimal model exploits DFA/DFM by combining the functional sensitivity factors and machinability factors of parts. Fuzzy comprehensive evaluation is used to evaluate the machinability of parts and a new mathematical model is developed and solved using genetic algorithm (GA). The study pointed out that it is difficult to determine the cost tolerance relationship for every process. The machinability of parts depends on the dimensions, the geometrical structure, the material machinability and the machining accuracy. These factors are treated as fuzzy and a two-order fuzzy comprehensive evaluation is used to process these factors. Membership degrees are determined directly by experts or by membership functions. The assembly tolerance is a summation function of a constant value and a factor reflecting the degree of importance of each design tolerance on an assembly.

GA is different from traditional search methods used in optimisation problems. GA works with a coding of design variables as opposed to the variables themselves. Continuity of parameter space is not a priori requirements. Generally, GA searches from a population of points (using a sample of points). This eliminates nearly the possibility of being trapped in a local minimum. GA possesses several attractive features such as flexibility, parallelism, simplicity and versatility.

GA uses several kinds of representations such as binary digit, floating-point representation and permutations of a list. The use of each representation is very much dependent on the search space. The constrained optimisation function is transformed into an unconstrained problem using an exterior penalty function method. The method is applied to a part with biased variations.

Diplaris and Sfantsikopoulos (2000) developed new analytical cost-tolerance models. This model is more inclusive and has wider field of applications. It takes into account the size of the tolerance dimension, the tolerance size and the size of initial tolerance prior to machining.

1.6 Variance reduction techniques

Webb and Parkinson (1995) discussed the properties and optimisation of the linear variance function. This is a measure of how variations in variables (and parameters) are transmitted to design functions. Mathematical properties of this function are discussed. There are at least two levels at which variation can be addressed during the design phase. The first level involves maintaining design feasibility. Variations in variables can cause the violation of design constraints. Such designs are said to have ‘feasibility robustness’. The second level involves reducing the sensitivity of the design to the transmitted variation. Besides keeping the design feasible, we may wish to minimise the sensitivity of the design to variation. This is known as ‘sensitivity robustness’. A key relationship in both of these levels is the linear variance function. Some of the properties of the variance functions are detailed:
the variance function has a minimum at every stationary point in the performance function

b another fact is use, for a quadratic function, second derivatives are constant and higher derivatives do not exist.

Parkinson (1998) described a method to aid robust design in the presence of design parameter uncertainty. This is based on probabilistic simulation. Optimum values of nominal design parameters are obtained via non-linear optimisation algorithms in the presence of constraints. The method is used to obtain robust design point when a performance measure is defined by an analytical expression. The method also employs a non-linear minimisation of a Lagrangian expression by use of a simplex search optimisation algorithm.

Approaches employing neural networks and response surface models need further validation as the neural networks employs BP, RBN or probabilistic approaches. It is in fact an approximate model between inputs and outputs. Response surface models are approximate technique by nature. How the value of the resulting tolerance values compares with accuracy of RSM is an important question. If the RSM employs a linear model, probably the accuracy of solution is inferior to the resulting solution.

1.7 Problem statement

In this study, the problem of least cost tolerance allocation is given (Gadallah, 2000; Zhang and Wang, 1993). Data pertinent to the example problem is given in appendix together with the functional constraints expressed in terms of variables by equation (1) and equation (2), respectively. The cost tolerance model used in this study is similar to that used by Zhang and Wang (1993). Instead of the usual approach, three formulations are introduced, these are:

1 the minimum sensitivity formulation
2 the minimum variance formulation
3 minimisation of the second derivative formulation.

Solutions based on the three formulations are given and compared.

Another advance in this study is:

1 The constraints represent limits on the available feasible domain. A way of maximising the utilisation of the domain would be very much needed.
2 The solution of this procedure is done in n-stages corresponding to the availability of several constraints.
3 The difference in solutions of the least cost formulation and the least sensitive formulation represents the margin how much variables are allowed to change (increase or decrease without much changing the value of objective function).

Figure 1 gives a measure of objective function versus iteration number for Stage 1 and Stage 2, respectively. In Stage 1, X1, X3 and X5 are chosen such as the sub-production cost is minimum. In Stage 2, X2 and X4 are chosen to minimise the objective function such that the production cost is minimum and the constraints is close to the allowable
limit. The figure is annotated in two forms to differentiate between Stage 1 and Stage 2, respectively. When the objective function refers to another measure other than the minimum cost function, we call that a measure of objective function.

**Figure 1** Measure of objective function versus iteration number using n-stage optimisation (see online version for colours)

![Stage-1 Stage-2](image)

**Table 1** Solutions of the optimisation problem

<table>
<thead>
<tr>
<th>O-T</th>
<th>L-S-O-F</th>
<th>C-V</th>
<th>O-P-C</th>
<th>Min. solution</th>
<th>Max. solution</th>
<th>N-O-I</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Based on minimum variance formulation (variance of $X_i = 0.0005$)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.0052, 0.0059, 0.0042, 0.0056, 0.0036</td>
<td>$22.809$</td>
<td>0.0246</td>
<td>${1, 1, 1, 1, 1}$</td>
<td>222.17</td>
<td>2,281</td>
</tr>
<tr>
<td>2</td>
<td>Based on minimum sensitivity formulation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.0052, 0.0059, 0.0042, 0.0056, 0.0036</td>
<td>$22.806$</td>
<td>0.0246</td>
<td>${1, 1, 1, 1, 1}$</td>
<td>1,423</td>
<td>-</td>
</tr>
<tr>
<td>3</td>
<td>Based on minimum cost formulation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.0066, 0.0055, 0.0042, 0.0056, 0.0030</td>
<td>$19.686$</td>
<td>0.0246</td>
<td>${3, 3, 3, 2, 1}$</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>4</td>
<td>Based on minimum second derivative</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.0052, 0.0059, 0.0049, 0.0042, 0.0043</td>
<td>$22.999$</td>
<td>0.0246</td>
<td>${1, 1, 1, 1, 1}$</td>
<td>Minimum second derivative = 620,000</td>
<td>-</td>
</tr>
</tbody>
</table>

Notes: 1 – minimum variance formulation; 2 – minimum sensitivity formulation; 3 – minimum cost formulation; 4 – minimum second derivative formulation; N-O-I = number of iterations; C-V = constraint value; L-S-O-F = least sensitive objective function; O-T = optimum tolerance; O-P-C = optimum process combination

Table 1 gives solutions to the optimisation problem using the four formulations. From the results, the minimum variance and sensitivity yield almost the same results in terms of decision variables and process combinations. The minimum and maximum values are 222.17 and 2,281 for the variance formulation (assuming the variance of $X_i = 0.0005$).
The resulting solution will vary as per the variance of decision variables. Based on the minimum cost formulation, the minimum production cost is $19.686 and process combinations are \{3, 3, 3, 2, 1\}. The highest value of objective function corresponds to the minimum second derivative formulation. It is realised that the solution utilises most of the available feasible domain in contrary to other formulations. The number of iterations in all formulations is reasonably small.

<table>
<thead>
<tr>
<th>O-T</th>
<th>L-S-O-F</th>
<th>C-V</th>
<th>O-P-C</th>
<th>Min. solution</th>
<th>Max. solution</th>
<th>N-O-I</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Based on minimum sensitivity formulation</td>
<td>0.0052, 0.0059, 0.0042, 0.0056, 0.0036</td>
<td>$22.806</td>
<td>0.0246</td>
<td>{1, 1, 1, 1, 1}</td>
<td>1423</td>
<td>-</td>
</tr>
<tr>
<td>2 Based on feasibility robustness formulation</td>
<td>0.0064, 0.0050, 0.0037, 0.0047, 0.0047</td>
<td>$22.750</td>
<td>0.0245</td>
<td>{1, 1, 1, 1, 1}</td>
<td>Partial sequential optimisation</td>
<td>2, 1</td>
</tr>
</tbody>
</table>

Notes: X1: 0.0052 \rightarrow 0.0064 (+0.0012); X2: 0.0059 \rightarrow 0.0050 (–0.0009); X3: 0.0042 \rightarrow 0.0037 (–0.0005); X4: 0.0056 \rightarrow 0.0047 (–0.0009); X5: 0.0036 \rightarrow 0.0047 (+0.0011)

The objective function in case of feasibility robustness is smaller than minimum sensitivity formulation. The feasible space utilised is larger than in the case of minimum sensitivity formulation. There is a gain (loss) in case of X1, X2, X3, X4 and X5 equal to +23%, –18%, 13%, –17% and +30%, respectively. In partial sequential optimisation, the number of stages = 2; but the idea can be extended to multi-stages (equivalent to the number of functional constraints).

Figure 2 gives a schematic showing the allowable variations in X2-X4 space. According to the solutions obtained, X2 should reduce from 0.0059 mm to 0.0050 mm and X4 should reduce from 0.0056 mm to 0.0047 mm. These reductions allow the system to be least sensitive at the expense of increasing the cost. Similarly, in the X1-X3-X5 domain, X1 and X5 should increase from 0.0052 to 0.0064 and 0.0036 to 0.0047. The X3 should reduce from 0.0042 to 0.0037.
Figure 3 gives a comparison among the four formulations. The minimum variance formulation exhibits the highest objective function value at the early start of optimisation. The least cost formulation shows a decreasing trend towards the minimum. The least sensitivity formulations show two regions: the first, a fluctuating trend and the second, a decaying trend towards the minimum. Figure 4 gives a comparison for the variance formulation with variances on decision variables ranging from 0.0001 to 0.0005. It is true that experience on problem variations dictate values for variance. The approach taken in this paper can assume the variances on variables as unknown (Ramakrishnam and Rao, 1994). This can be done at the expense of the resulting number of variables.

Figure 3  Comparison among various formulations

![Comparison among various formulations](image1)

Figure 4  Comparison among the minimum variance formulation (see online version for colours)

![Comparison among the minimum variance formulation](image2)

1.8  Heuristic approach

An optimisation based on heuristics is used to solve the three problems. The approach is briefly given in Appendix 2.
2 Conclusions

The procedure given in the study does not only offer an optimum for the problem of
tolerance allocation. It offers a method that gives a margin or allowance for variation of
each variable. Through n-stage optimisation, each constraint is utilised to the maximum.
The optimum reached cannot be termed as global, since the scheme presented has a
sequential nature. In the minimum sensitive, variance and second derivative
formulations, the solutions obtained are higher than the least cost formulation.

This procedure is very much suitable for multi-constraint problems. That is why we
call it n-stage optimisation corresponding to n-constraints. When the number of
constraints is one, the solution is equivalent to the least cost formulation. This might
increase the computational time for multi-constraint problems.

The margin of solutions for the variables is an allowance where the objective function
is least sensitive to variations. These variations are the solutions of the problem and not
known a priori. The least sensitive formulation was introduced before with respect to
general design problems and not for the tolerance allocation problems. Now, the
variations take a range and not just a point.

In this paper, the least cost tolerance allocation is herein modified using the least
sensitivity and least variance formulations. The least cost problem results in optimal
tolerance values that minimise the production cost. The least sensitivity and variance
formulations result in the tolerance values that minimise the sensitivity and variance of
production cost. An n-stage optimisation was finally introduced that result in maximum
utilisation of design constraints. The difference between the values of the minimum cost
and n-stage optimisation represents a margin, through which variations on variables will
not result in violation of constraints. This margin of variations might be towards the
positive or negative sides. A better-needed method would be to find the bi-lateral margin
of variations. Several comments can be made:

1 The allocation problem is specific to problems with multiple requirements. For
   single requirement problems, the solution should conform to those of usual
   problems.

2 The usual allocation problems obtain a solution that minimises the measure of
   objective functions. Here, the solution maximises the utilisation of constraints and
   still maintains a degree of robustness.

References

Chase, K., Greenwood, W., Loosli, B. and Hauglund, L. (1990) ‘Least cost tolerance allocation of
mechanical assemblies with automated process selection’, *Manufacturing Review*, Vol. 3,
No. 1, pp.49–59.


Appendix 1

Cost-tolerance function coefficients

<table>
<thead>
<tr>
<th>Component</th>
<th>Process 1</th>
<th>Process 2</th>
<th>Process 3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$a$</td>
<td>$b$</td>
<td>$a$</td>
</tr>
<tr>
<td>1</td>
<td>3.0</td>
<td>0.012</td>
<td>2.0</td>
</tr>
<tr>
<td>2</td>
<td>-0.33</td>
<td>0.0093</td>
<td>-8.0</td>
</tr>
<tr>
<td>3</td>
<td>3.0</td>
<td>0.003</td>
<td>2.0</td>
</tr>
<tr>
<td>4</td>
<td>4.0</td>
<td>0.008</td>
<td>3.0</td>
</tr>
<tr>
<td>5</td>
<td>6.0</td>
<td>0.004</td>
<td>5.0</td>
</tr>
</tbody>
</table>


Functional constraints:

\[
x_1 + x_2 + x_3 + x_4 + x_5 \leq 0.025
\]  
\[
x_1 + x_3 + x_2 \leq 0.015
\]
Appendix 2

Algorithm employed (see online version for colours)

3-Levels

T1

1

0

-1

5-Levels

T1

-1

-0.5

0

+0.5

1

6-Levels