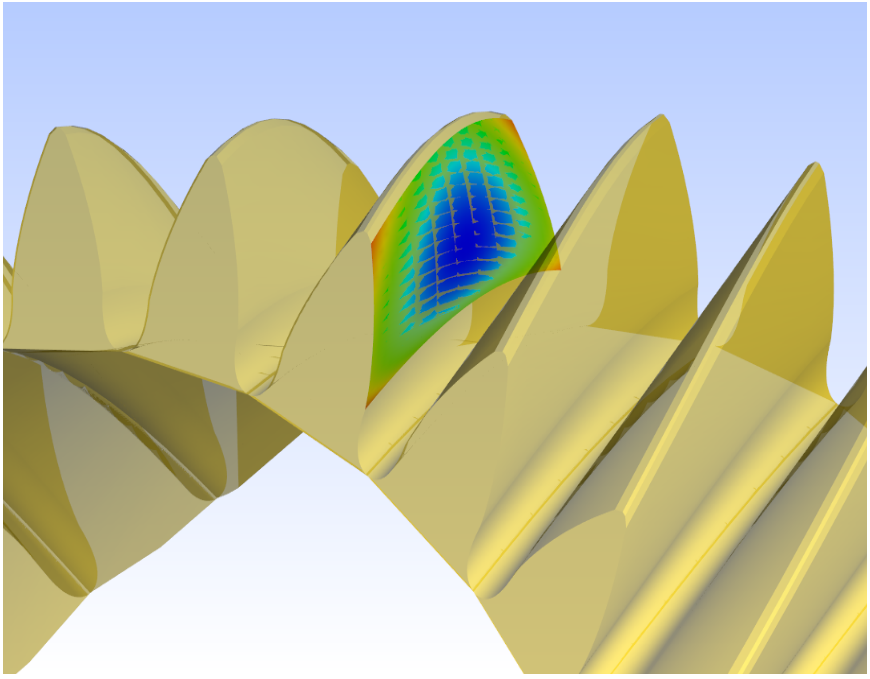
CHALMERS
UNIVERSITY OF TECHNOLOGY



Gear Design Multi-Objective Optimization for Automotive Transmissions

Master's Thesis in Automotive Engineering

AKSHAY D. S. BHAT

Department of Mechanics and Maritime Sciences
CHALMERS UNIVERSITY OF TECHNOLOGY
Gothenburg, Sweden 2019

Master's thesis 2019:96

Gear Design Multi-Objective Optimization for Automotive Transmissions

AKSHAY D. S. BHAT

Department of Mechanics and Maritime Sciences
Division of Combustion and Propulsion Systems
Chalmers University of Technology
Gothenburg, Sweden 2019

Gear Design Multi-Objective Optimization for Automotive Transmissions

AKSHAY D. S. BHAT

Industrial Supervisors: Omar D. Mohammed and Peter Falk, CEVT AB
University Examiner: Sven B. Andersson, Department of Mechanics and Maritime Sciences

Master's Thesis 2019:96
Department of Mechanics and Maritime Sciences
Division of Combustion and Propulsion Systems
Chalmers University of Technology
SE-412 96 Gothenburg
Telephone +46 31 772 1000

Cover: Gear tooth micro-geometry modification constructed in MASTA(source: CEVT AB).

Printed by Chalmers University of Technology
Gothenburg, Sweden 2019

Abstract

Transmission design and development have to take a lot of design variables and their resulting effects into account as early in the design stage as possible. Most of these resulting effects(also known as objective functions in this thesis) are conflicting in nature. There will be an optimal set of design variables that will result in minimum or maximum of these objective functions. It is very beneficial to arrive at this optimal set of design variables using a structured methodology rather than trial and error methods to obtain unique solutions and save time. Gear whine and durability are the two main conflicting factors considered in this thesis. The optimization methodology which is the aim of this thesis, is built to minimize the peak-peak transmission error, contact and root stresses by modifying the micro-geometry variables. The thesis was performed in collaboration with CEVT AB.

A design space is selected for the micro-geometry variables. WindowsLDP, a software used for gear tooth contact analysis, is used for Design of Experiments(DOE) to calculate the objective functions for multiple design points within the design space. Probability distribution and worst-case scenario distribution is applied for the objective functions to make them robust against torque. The data is then used to develop metamodels for each objective function in MATLAB using squared exponential Gaussian regression. They are then used in a multi-objective optimization algorithm in MATLAB to explore the design space and obtain a pareto/non-dominated set of solutions. These solutions are checked for design safety and ranked highest to lowest based on weights distributed between peak-peak transmission error and safety factors. The highest ranked micro-geometry values are substituted in the original gear model and the objective functions are calculated against torque and compared to the benchmark.

The highest ranked pareto optimal result shows that peak-peak transmission error, contact stress and gear 2 root stress minimized compared to the benchmark although there was a slight increase in gear 1 root stress. Other pareto optimal solutions show different levels of minimization for different objective functions. But this is normal as the objective functions are conflicting to each other. Also, once the design safety conditions are satisfied, reducing the peak-peak transmission error and hence the gear whine is very important.

Based on the results obtained, it can be concluded that the developed methodology managed to optimize the micro-geometry variables to minimize the objective functions in a very structured format and minimal time. It also arrived at a unique set of pareto solutions every time the optimization was completed, given that the conditions remained the same.

Keywords: micro-geometry, peak-peak transmission error, contact stress, root stress, metamodel, robustness measure, optimization

Acknowledgements

I would like to thank my supervisors Omar D. Mohammed and Peter Falk for their help, guidance and constructive feedback during this thesis. I would like to thank my examiner Sven B. Andersson for his input, support and mentoring. I would also like to thank Martin Hedström for giving me the opportunity to conduct my thesis at CEVT AB and many thanks to Magnus Flood and Viktor Pettersson for their encouragement throughout the thesis. A big thank you to all the staff of CEVT AB who made my work there possible.

A bow and a heartfelt thank you to my parents who encourage and help me at every step of my life and for their unending love and blessings.

Akshay D. S. Bhat, Gothenburg, August 2019

Contents

1
Introduction

Automotive transmissions are nowadays required to transmit higher loads via their gear train from powerful engines. Transmission gears are expected to be durable and have refined noise performance. These are conflicting cases observed at the system level of design of transmissions and are connected to specific parameters of a gear or gear train design in much earlier stages. Hence, modifying the parameters of a gear or a gear train during its design to get optimal values is of paramount importance. Macro-geometry of gears such as module, number of teeth, pressure angle, etc. and micro-geometry such as tooth profile relief and crown, lead crown and slope, etc. are some of the parameters that can be modified. A structured optimization approach which can weigh different attributes against each other is a great aid in the design process.

1.1 Literature Review

The motivation for this thesis is the necessity of a structured optimization routine as transmission gear design have many conflicting parameters such as transmission error and contact and root stresses. Minimizing each parameter by trial and error manually consumes valuable time and resources. An optimization methodology will result in faster design process and better quality product for the stakeholder. Also, it will yield in unique solutions regardless of it being used by different engineers with different thought process and levels of experience.

Several variables affect the performance of a transmission, or to be specific, performance of a gear system. Hence it is multidimensional. The effect of these several variables are simultaneous and must be studied in the same way while exploring the design space for optimum design [1]. The variables of gear design can be broadly classified as macro-geometry (module, number of teeth, pressure angle, etc.) and micro-geometry variables (tooth surface modifications). Macro-geometry variables are only changed in the initial steps of design as these changes affect the transmission as a system and may also affect other systems in the vehicle and their packaging. Micro-geometry variables being only tooth surface modifications in the order of microns do not affect the whole transmission and hence can be changed at later stages of design as well [2]. The contact properties of a meshing gear pair is very sensitive to operating conditions such as load and misalignment and also manufacturing errors. This is compensated by introducing tooth surface modification and hence is very important in gear design [3]. Gear tooth contact analysis softwares and FEM simulation softwares are widely used to calculate stresses, transmission error, tooth failures, etc., under given conditions. But they usually work to simulate a single design of the gear system at a time. With the development of softwares that deal with machine learning and optimization algorithms, a vast number of different design models can be simulated iteratively and with the results compared to one another, an optimum design solution can be selected limiting the need for costly and time consuming prototypes and testing [4]. The two available multi-objective optimization algorithms in MATLAB are genetic algorithm and direct search algorithm. Genetic algorithms are the most commonly used algorithms to solve multi-objective optimization problem as they do not converge to a local minima pre-maturely and thus have higher chance of finding the global or true minima. [4]

Optimization of gear tooth micro-goemetry modifications are studied in [1] - [8]. The micro-geometry is optimized by developing metamodels for the FEM simulation models in [4]. The computationally efficient metamodels are used in the optimization algorithm for the speedy calculation of the pareto optimal solutions. But the disadvantage of this methodology is that a range of torque/load conditions is not considered and different optimum design is generated for different torque values. Torque being an operating condition, having different designs for different torque values is not possible as one design is selected and manufactured. In [3], this problem does not exist as the optimization is done for a range of torque values and the pareto optimum design is calculated by introducing a robust counterpart of the objective function over the range of torques. But the disadvantage of this method is that it uses the original model in the contact analysis software to calculate the objective functions for all iterations. Hence a communication between the optimization algorithm and the contact analysis software is necessary. This is not possible in this thesis. Hence, there is a need for a new methodology that combines the attributes of both above stated methods. Need for metamodels due to non communication between optimizing algorithm and contact analysis software and optimum design for a range of torques led to the methodology developed in this thesis.

1.2 Aim

The aim of this thesis is to develop an optimization methodology to minimize gear whine, contact stress and root stress by modifying gear tooth surface micro-geometry within feasible design space using genetic algorithm optimization to obtain a set of pareto solutions. A post-optimization analysis of the pareto front ranks the optimal solutions according to a given set of conditions.

1.3 Methodology

The main steps in the thesis are:

- **Design of Experiments**
  A feasible design space or boundary is fixed for the variables. Multiple design points are generated within this design space using the full factorial sampling method. The objective functions are calculated for all the generated design points.

- **Robust Objective Function**
  The objective functions are converted to robust objective functions using two important robustness measures:
  - Probability-based measures
  - Worst-case scenario

- **Metamodelling**
  Metamodels or response surfaces are created using a metamodelling technique called Gaussian process. This simpler mathematical description of the more complex model of a gear train is used for the optimization process as it is computationally highly efficient.

- **Optimization**
  Multi-objective optimization is performed using gentic algorithm. The objective function, number of variables, upper and lower bounds must all be defined before running the optimization.

- **Post-optimization analysis**
  The pareto set of solutions obtained is analyzed by first calculating safety factor values for each solution. The solutions are then ranked based on the values of safety factors and peak-peak transmission error using TOPSIS algorithm.

1.4 Limitations

The thesis is limited to the optimization of one gear pair of the given transmission. Optimization of only gear tooth micro-geometry parameters is considered, as changes made to these parameters will not adversely affect the system design of the transmission which is important when the design phase is at a late stage. But the methodology developed is expected to aid further development and be able to incorporate more variables of the design process. WindowsLDP, a gear tooth contact analysis software contains the original model of the gear pair and is used for the DOE. The generation of DOE in this software is limited only to the five micro-geometry variables used in this thesis. MATLAB is used for the subsequent steps in the thesis. The accuracy of the gear pair model in WindowsLDP to calculate the objective functions is not a concern for this thesis, but the accuracy of the metamodel developed to mimic the WindowsLDP model is important.

2
Theory

Basic theory about gears, metamodeling and optimization is presented in this chapter.

2.1 Gear theory

Gears in transmissions are used to transmit power along the powertrain from the engine and/or electric motors to wheels. Gears on different shafts have different diameters and number of teeth to vary the range of torque and speed available from the engine and/or electric motors to the range required at the wheels. Ideally, for a gear pair there should not be any power transmission losses. But mainly due to frictional losses, the power transmission efficiency is around 98% [9]. The gear theory in this report will only look at some aspects of gear design required for the thesis given in subsections 2.1.1 - 2.1.9. For more detailed description of gear nomenclature Dudley's handbook [10] and AGMA manual [11] can be referred.

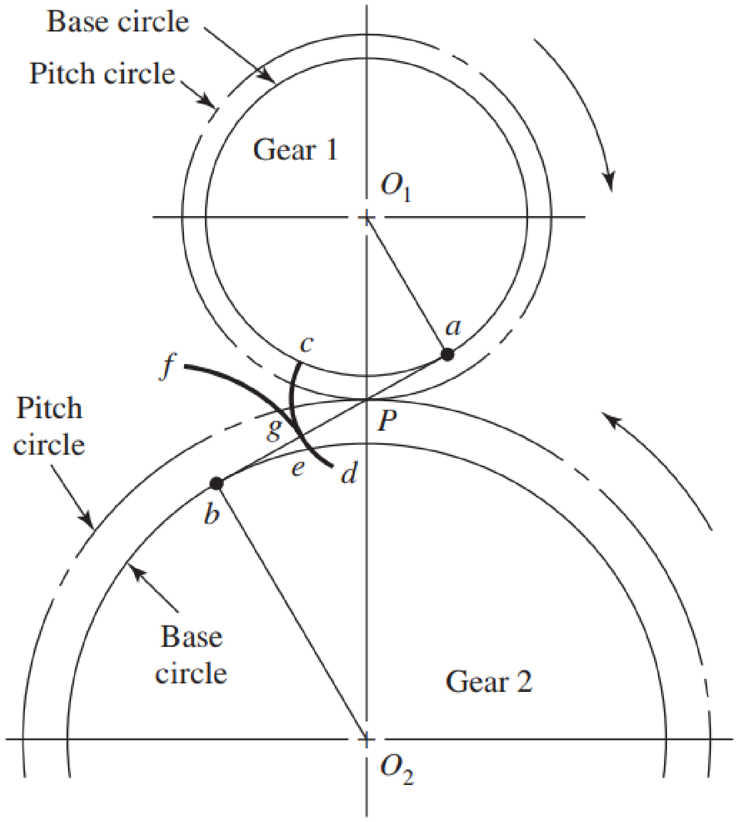
2.1.1 Speed Ratio

The speed ratio of a gear pair is defined as the ratio of speed (ω) of driving gear to speed of driven gear or ratio of number of teeth (Z) of driven gear to number of teeth of driving gear.

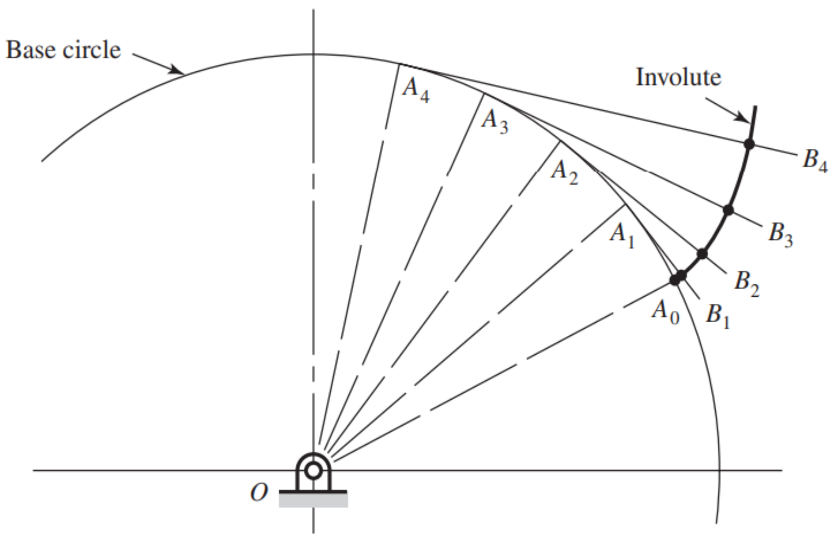
$$i = \frac{\omega_{driving}}{\omega_{driven}} = \frac{Z_{driven}}{Z_{driving}} \tag{2.1}$$

2.1.2 Conjugate Action and Involute Profile

One of the basic considerations while designing gears is that the angular velocity of the driving gear should be transferred smoothly to the driven gear with a constant speed ratio. This is called conjugate action and is one of the basic law of gearing. According to [10], it is defined as *"normals to the profiles of mating teeth must, at all points of contact, pass through a fixed point located on the line of centers"*. The commonly used way to achieve this is by making the shape of the active profiles an involute curve. A curve marked out by a point on a straight line when the straight line is rolling on the base circle without any slip is called an involute curve. Active profiles are the surfaces of the gear teeth that makes contact to transmit the motion.



(a) Conjugate action [12]



(b) Construction of involute [12]

Figure 2.1: Illustration of involute action and its construction

2.1.3 Face width

The length of the gear teeth along the axial plane is called face width. [11]

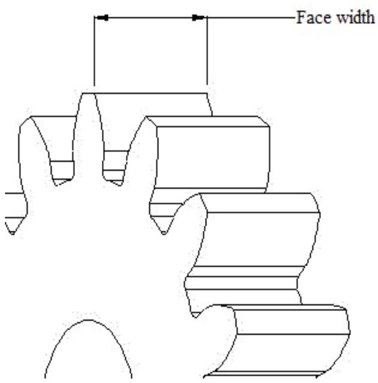


Figure 2.2: Face width

2.1.4 Involute roll angle

According to the American Gear Manufacturer's Association [11], the involute roll angle is defined as *"the angle whose arc on the base circle of radius unity equals the tangent of the pressure angle at a selected point on the involute."*

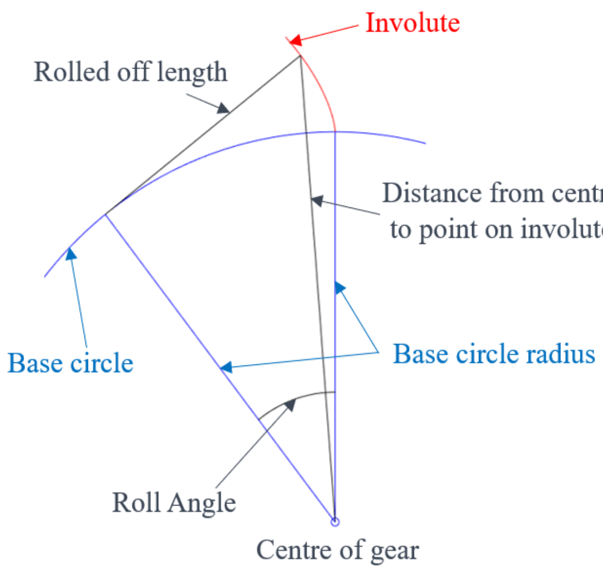


Figure 2.3: Roll Angle

2.1.5 Start and End of Active profile

For a given gear, the start of active profile (SAP) is the lowest possible point on the flank of the tooth at which contact with the mating tooth tip starts and end of active profile (EAP) is the last point of contact as shown in Fig. 2.4. [13]

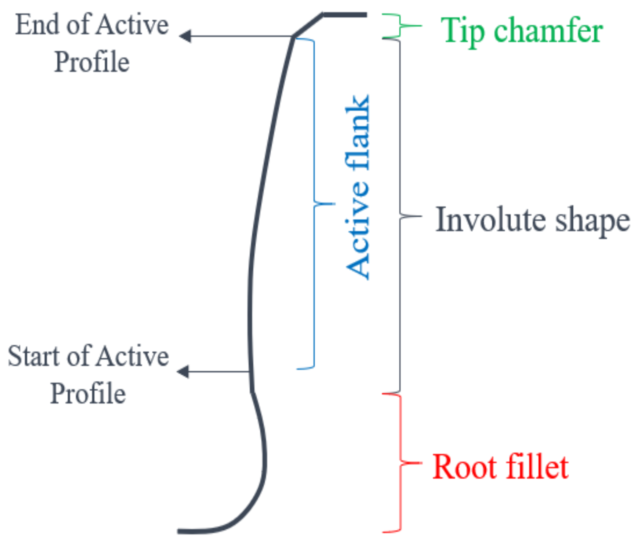


Figure 2.4: Start and End of Active Profile

2.1.6 Micro-geometry Modifications

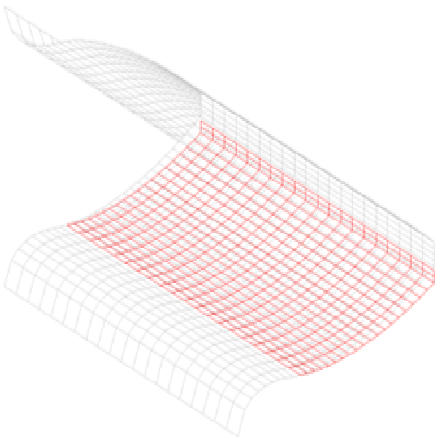
As discussed in section 2.1.2, conjugate action between mating gears is one of the basic considerations required in gear design. This can be achieved by having perfect involute profiles for the gear teeth. The disadvantage of having perfect involute is that it provides conjugate action only when load applied is zero. Once a load is applied to the gears, the conjugate action is no longer present due to the deflections caused by the load. Hence, modification of the active tooth profile is required to negate the effects of the deflection. This modification to the tooth profile done by removing very small amount of material (in the order of microns) from the active tooth profile is called micro-geometry modification. [14]

The micro-geometry modifications discussed in this thesis are:

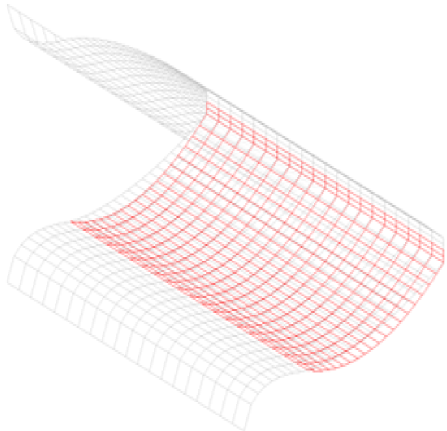
- Profile slope $f_{h\alpha}$
- Profile crown c_{α}
- Lead slope $f_{h\beta}$
- Lead crown c_{β}
- Bias

Slope modifications are linear while crown modifications are parabolic in nature. Considering micro-geometry modifications for a driving gear, the SAP is close to root of the gear tooth and EAP is close to tip of the gear tooth. Profile slope ($f_{h\alpha}$) modification is made by adding or removing material at EAP of the gear tooth. Profile crown (c_α) modification also known as profile barrelling is made by removing material both at SAP and EAP of the gear tooth. Lead slope ($f_{h\beta}$) modification is made by adding or removing material from one end of the face width of the tooth. Lead crown (c_β) modification is made by removing material from both ends of the face width of the gear tooth [7]. Bias modification is basically twisting the active profile with one end appearing to be twisted in clockwise direction and the other end appearing to be twisted in anti-clockwise direction [15]. Fig. 2.5 shows the different micro-geometry modifications.

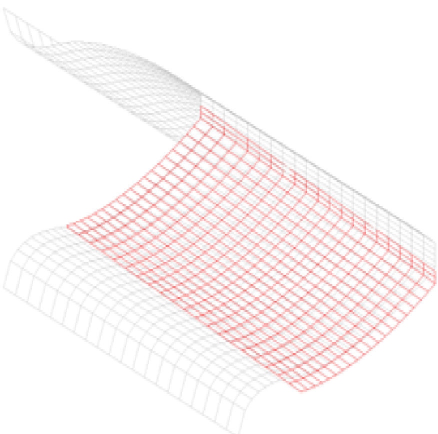
Compensation for tooth bending deflections, manufacturing errors and load sharing across the profile of mating teeth is affected by profile modifications. Misalignment are compensated by lead modifications. Manufacturing errors affect both profile and lead modifications. Bias modifications also affect load sharing. [14] [16]



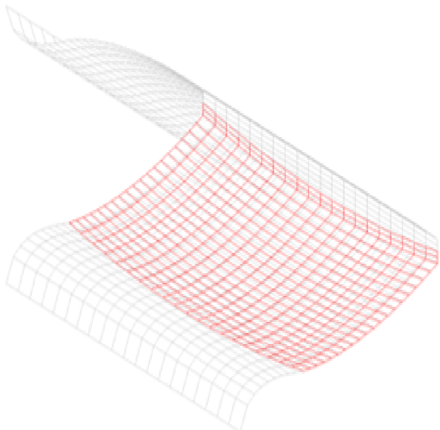
(a) Profile Slope



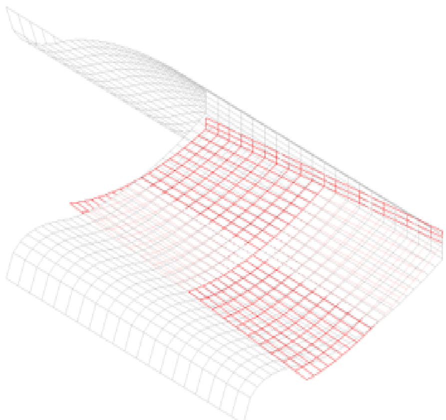
(b) Profile Crown



(c) Lead Slope



(d) Lead Crown



(e) Bias

Figure 2.5: Micro-geometry modifications included in thesis

2.1.7 Gear whine and transmission error

Gear whine is caused by two mating gears in motion. The vibrations caused due to change in force are transmitted from the gears to the transmission casing via shafts and bearings which then vibrates and generates noise in the air. It is also transferred to rest of the structure. According to studies [17] - [19], transmission error in the mating gears is one of the primary causes of gear whine. Hence, reducing the transmission error will reduce the gear whine.

Transmission error is defined as “the difference between the actual position of the output gear and the position it would occupy if the gear drive were perfectly conjugate”[20] and is as shown in Fig. 2.6.

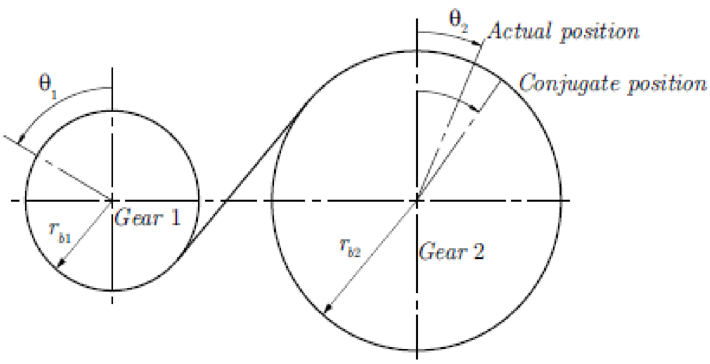


Figure 2.6: Transmission Error [21]

2.1.8 Contact Stress

When two mated bodies are in contact, they are under a type of stress called Hertzian contact stress. Two gears in mesh can be represented by two cylinders in contact. In an ideal situation, there is a line contact between the two bodies. Stress being force divided by area, line contact results in infinite stress. But in real conditions, the bodies in contact deform slightly due to elastic deformation and a small area of contact is created. This keeps the contact stress limited. Pitting and scuffing failure modes are caused due to high contact stress on gear tooth flank. [22]

2.1.9 Root Stress

Root stress occur at the two root fillets of a loaded gear tooth. The gear tooth can be simplified to a beam with a point load to calculate the root stress. On the working side of the tooth flank (where the contact of the mating teeth occurs), the root stress is tensile and on the non-working side, it is compressive. The maximum root stress should not exceed the permissible bending stress of the material used, else crack initiation and propogation will occur leading to failure [23]. Contact stress and root stress can be visualized in Fig. 2.7.

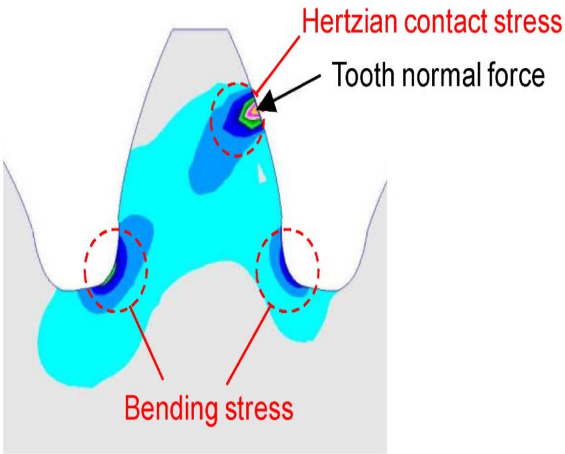


Figure 2.7: Contact Stress and Root Stress [24]

2.2 Metamodeling

Simulation of a process, most of the times, consumes a lot of time to calculate all the data required for a given design point. The simulation software also might be designed to calculate many parameters that might not be required for the optimization process and is a waste of computation power and time. Also, the optimization algorithm if setup in a different software will have to communicate with the simulation software back and forth for every design point calculation. Moreover, the setting up of the communication itself can be difficult or sometimes impossible to achieve.

All the problems stated above can be solved single-handedly by introducing the concept of metamodeling. Also popularly known as response surface modeling, it is the process of replacing complex simulation model with a simpler representation that predicts the solutions for a given set of data (design space). The developed metamodel can be validated by comparing the values calculated by the simulation model and the values predicted by the metamodel for the same set of design points. Metamodeling has been applied by studies [25] and [26] to full vehicle structure optimization problem. Both are very good examples to show that it is not feasible to run full simulation models for every point to be explored in the given design space. The slight loss of accuracy that comes with metamodeling is accepted because of the large amount of computation time saved that directly translates to reduced costs.

If the relation between the variables and objective functions can be fitted with a predetermined parameterized function, polynomial models can be used for metamodeling. In this thesis the objective functions express a high degree of non-linearity with the five micro-geometry variables, so much that the form of their relationship is unknown. In such cases, Gaussian process algorithm may be used to develop metamodels. [4]

The squared exponential Gaussian process is defined by it's kernel or covariance function given in Eq. 2.2 [27]. Here, σ_f is signal standard deviation and σ_l is a scale that defines the uncorrelation of objective functions depending on distances between input points x_i.

$$k(x_i, x_j|\theta) = \sigma_f^2 \cdot exp[-\frac{1}{2}\frac{(x_i - x_j)^T(x_i - x_j)}{\sigma_l^2}] \tag{2.2}$$

2.3 Optimization

A multi-objective optimization is the process of optimizing two or more objective functions simultaneously. The objective functions have to be either minimized or maximized. When the objective functions are conflicting to each other, determining the optimum solution becomes quite challenging. An optimization problem may or may not be constrained. Constraints could be equality and/or non-equality constraints and/or rectangular constraints with upper and lower bounds. Typically, a multi-objective optimization problem is defined as

$$\begin{array}{rll} \text{Minimize/Maximize,} & f_m(x), & \text{m = 1,2,....,M;} \\ \text{subject to,} & g_j(x) \geq 0 & \text{j = 1,2,....,J;} \\ & h_k(x) = 0 & \text{k = 1,2,....,K;} \\ & x_i^{(L)} \leq x_i \leq x_i^{(U)} & \text{i = 1,2,....,I;} \end{array}$$

where $x \in R^n$ for n number of decision variables is the solution set of all real numbers. A subset of this which satisfies all the constraints given is called the feasible decision variable space and is n-dimensional. All the x values in this feasible space is a potential solution. Each feasible solution can be used to calculate a feasible objective function. This way an objective function space is created which is M-dimensional (number of objective functions). So for every point in the feasible decision variable space, there is a point corresponding to it in the objective function space. This objective function space consists of the minima or maxima required. The optimization algorithm must sweep through this space to determine the required minima or maxima and the point corresponding to this minima or maxima in the feasible decision variable space will constitute as the optimal solution. [28]

A multi-objective optimization will not yield a single optimal solution but rather a set of optimal solutions. This is called the pareto front. It consists of all found non-dominated solutions. A visualization of the pareto front for two objective functions, both to be minimized, is given in Fig. 2.8.

According to [28], the domination is defined as follows:

"A solution x_1 is said to dominate the other solution x_2, if both the following conditions are true:

- *The solution x_1 is no worse than x_2 in all objectives. Thus, the solutions are compared based on their objective function values.*
- *The solution x_1 is strictly better than x_2 in at least one objective function."*

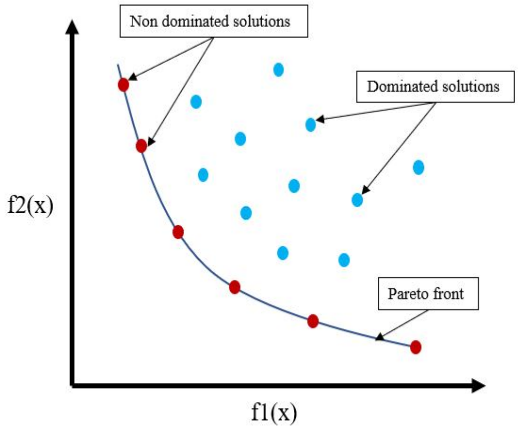


Figure 2.8: Visualization of pareto front

The optimization algorithm used in this thesis is a multi-objective optimization genetic algorithm from MATLAB. Genetic algorithms copy the biological evolution from nature. A random set of solutions called population from the given design space evolves towards better, more optimal solutions iteratively. Each iteration is called a generation. A solution generating better objective function is considered to be fitter than the one generating worse objective function. This solution is then selected for the next generation and the worse solution is eliminated. A new population is formed by carrying over the better solutions for the next iteration. Modification of the better solutions and reproduction of solutions from two solutions is done by mutation and crossover. These newly formed solutions become part of the population for the next iteration and objective functions are calculated and compared to find new better solutions. This process continues until the termination of the iteration is reached and a set of pareto optimal solutions is obtained. [29]

3

Metamodel based Multi-objective Optimization

Basic gear design data (example) used and the methodology developed for the optimization are presented in this chapter.

The optimization process deals with the following:

- Objective functions to minimize : peak-peak transmission error, contact stress, root stress of gear 1 and gear 2
- Decision variables are micro-geometry parameters : profile slope, profile crown, lead slope, lead crown, bias
- Uncertain variable: torque
- Constraints for decision variables : rectangular constraints with upper and lower bounds
- Metamodeling method : squared exponential Gaussian process
- Optimization algorithm : multi-objective genetic algorithm using MATLAB function gamultiobj
- Post-optimization analysis : Technique for Order of Preference by Similarity to Ideal Solution (TOPSIS) method based on calculated safety factors and transmission errors

The basic design data of a sample helical gear set are given in table 3.1. Gear 1 is driving and gear 2 is the driven.

Table 3.1: Basic design data of sample helical gear pair

Design Parameter	**Gear 1**	**Gear 2**
Number of teeth	17	79
Normal module (mm)	2.63	
Pressure angle (degree)	21.25	
Helix angle (degree)	-30.25	30.25
Operating center distance (mm)	142.0	
Tip diameter (mm)	59.2	236.7
Root diameter (mm)	44.2	222.06
Face width (mm)	36.0	33.0
Pitch diameter (mm)	51.75	240.52

An assumption is made in the gear tooth contact analysis in WindowsLDP that the load distribution on the 1% area of the outer edge of the active profile is neglected to exclude numerical errors that occur at corner contact.

3.1 Design of Experiments

A design space can be defined by defining values or functions that mark the outer boundary. Any values outside this design space will be considered unfeasible for the process. To build a metamodel or response surface, information within the design space is required. If large amount of information is provided, the metamodel will predict solutions more accurately but takes longer computation time and if small amount of information is provided, vice versa. A set of data points within the design space must be generated to calculate the necessary information. This can be achieved randomly, but random generation may not cover the entire design space and may render the metamodel inaccurate in these unexplored regions. There are many algorithms that generate design points such as latin hypercube sampling, full factorial, monte carlo sampling, etc. Full factorial sampling is used in this thesis as it is the available option in WindowsLDP.

Full factorial sampling method consists of factors for which the data points must be generated and levels that decide how many data points are allotted for each factor. All possible combinations between the factors are generated. The advantage of this is that the design space is explored thoroughly and systematically. If the level of the factor is increased to slightly increase the resolution of the information, a very large increase in the number of data points happen which makes the process computationally inefficient.

The response functions are calculated with the generated design samples. This whole set of information within the design space constitute the Design of Experiment. DOE can be defined as *"the method to determine relationship between factors affecting a process and the output of the process."* [30]

The five micro-geometry variables given in section 2.1.6 are the factors for the full factorial sampling. WindowsLDP can generate a maximum of 32000 sample points in it's full factorial algorithm. It is done in the robustness tab in WindowsLDP. Upper boundary value and lower boundary value is chosen for each of the micro-geometry variable. Then levels are selected for each variable. If a level of 10 is selected for the variable profile slope ($f_{h\alpha}$), then a set of 10 equidistant values is created with the lowest and highest value in the set being the defined lower and upper bounds for that variable.

3.2 Robustness Measures

Robustness to the objective function can be achieved by introducing robustness measures against some uncertain variables for the design process. For example, torque, misalignment, manufacturing tolerances, etc are uncertain conditions to the design process. Making the objective functions robust against these conditions will provide a solution set of micro-geometry values that can be used under all uncertain conditions with less unpredictable changes in system behaviour. Introducing robustness measures also increases the computation efficiency but the resulting loss of information will give a less accurate solution.

In this thesis, only torque is considered to be the uncertain variable. Two types of robustness measures are used:

- Probability based measures
- Worst-case scenario measures

3.2.1 Probability based measures

Probability based measures distribute the importance of the objective function over the uncertain variables. A distribution function is used for this and it is done by determining for which value of the uncertain variable the objective function is important and for which value it is not. For an automotive transmission the noise levels must be kept low at cruising speeds that is usually at low torque values for passenger comfort. At high torque values, durability is more important than minimizing noise. Hence peak-peak transmission error (PPTE) must be prioritized to be minimum at low torque values. This is achieved by introducing a triangular distribution function and is given by Eq. 3.1 [31]. This will result in the probability function over torque looking like Fig. 3.1. The robust equivalent of the peak-peak transmission error is calculated as given in Eq. 3.2 and is solved in MATLAB using the trapz function. Here, T_u and T_l are the upper and lower torque values.

$$p_t(T) = \frac{2(T_u - T)}{(T_u - T_l)^2} \tag{3.1}$$

$$f_{robust}(PPTE) = \int_T f(PPTE) p_t(T) dT \tag{3.2}$$

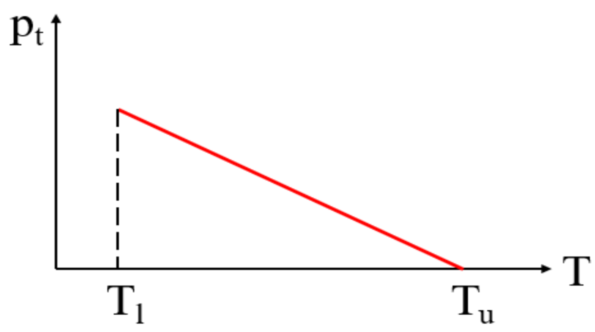


Figure 3.1: Triangular distribution function [3]

3.2.2 Worst-case scenario measures

Damage is induced to the gears at high torque values. Most often, contact stress (CS) and root stress (RS) are highest at highest torque level. Hence they carry very high importance for minimization. Worst-case scenario gives maximum importance to the highest torque level and zero importance to the rest. It is given by Eq. 3.3 and the probability function will look like Fig. 3.2. Also the robust equivalent of the contact stress and root stress is calculated as given in Eq. 3.4 and Eq. 3.5.

$$p_w(T) = \begin{cases} 1, & \text{if } f(stress) = max f(stress) \\ 0, & \text{otherwise} \end{cases} \tag{3.3}$$

$$f_{robust}(CS) = max(f(CS)) \tag{3.4}$$

$$f_{robust}(RS) = max(f(RS)) \tag{3.5}$$

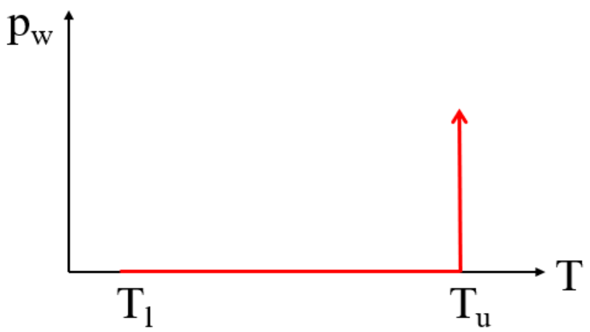


Figure 3.2: Worst case scenario distribution function [8]

3.3 Metamodeling

As discussed in section 2.2, the relation between the objective functions and the micro-geometry variables was found to be highly non-linear. Hence, Gaussian process regression method was used to develop the response surface or metamodel. Metamodels were created for each objective function, that is four metamodels for four objective functions.

The regression learner in MATLAB was used for this process. The input to regression learner should be in tabular form. The first five columns in the table are the micro-geometry values and the last column is one of the objective functions. After the input is selected, one of the two available validations is selected. The two available validations are cross-fold validation and hold out validation. It is recommended by MATLAB to select hold out validation for very large data sets. Hold out validation for 20% data was selected for the metamodel training in this thesis. Validation is important in developing the metamodel to prevent over-fitting of the Gaussian function. In the next step, the regression method is selected and in this case, squared exponential Gaussian process. The regression is carried out based on the data given to the regression trainer.

The model generated is in the form of an object in MATLAB is then exported into the work space and saved. It is then later called by the optimization algorithm as a function. This process is repeated for all the objective functions. The process of generating the metamodel is as shown in Fig. 3.3 to 3.5.

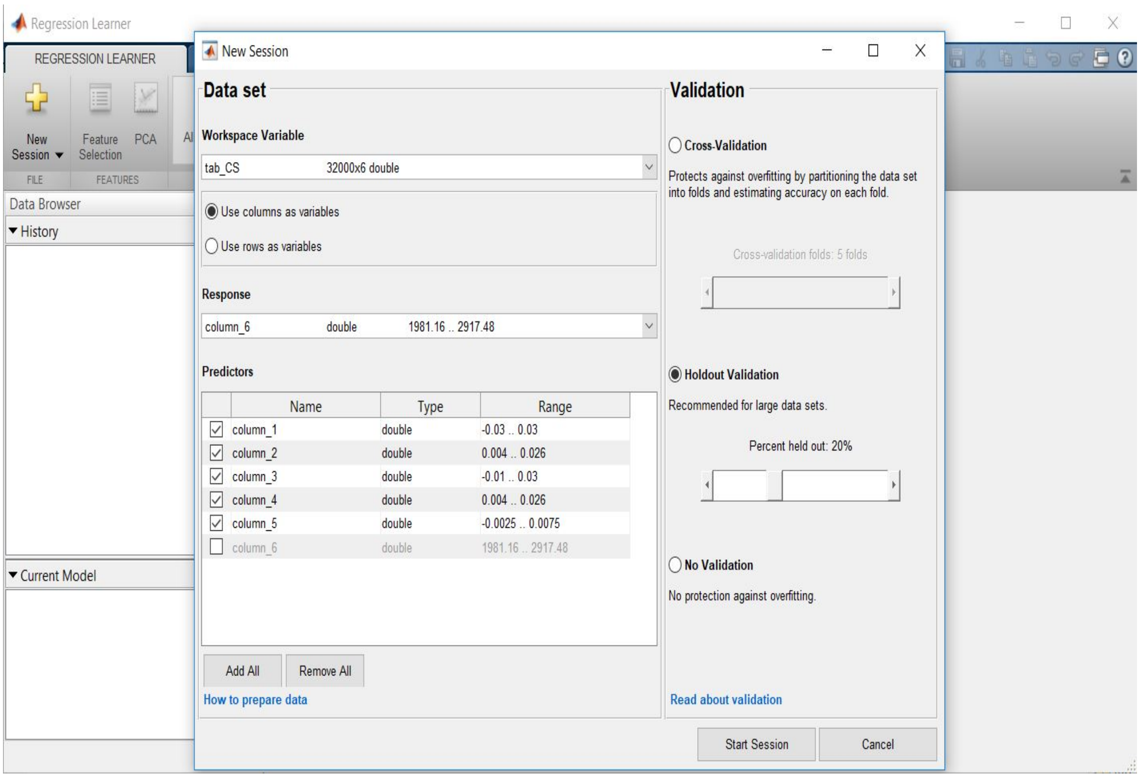


Figure 3.3: Regression Learner Predictors and Responses

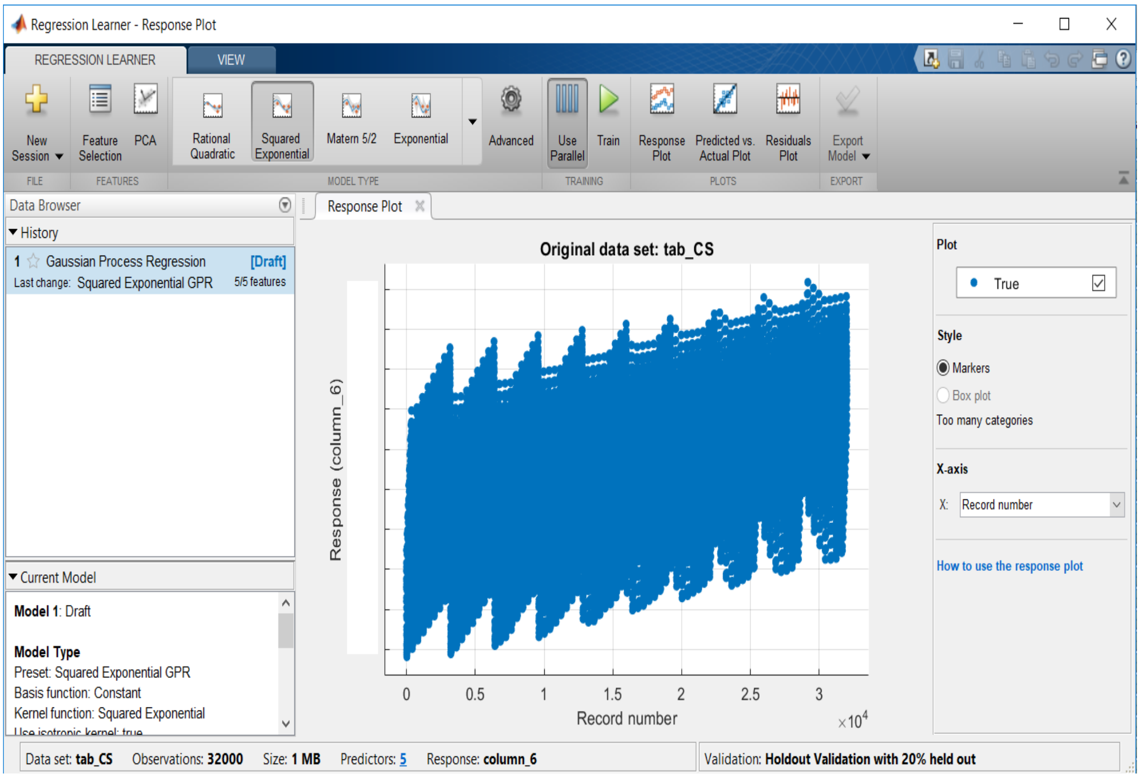


Figure 3.4: Response plot and model training

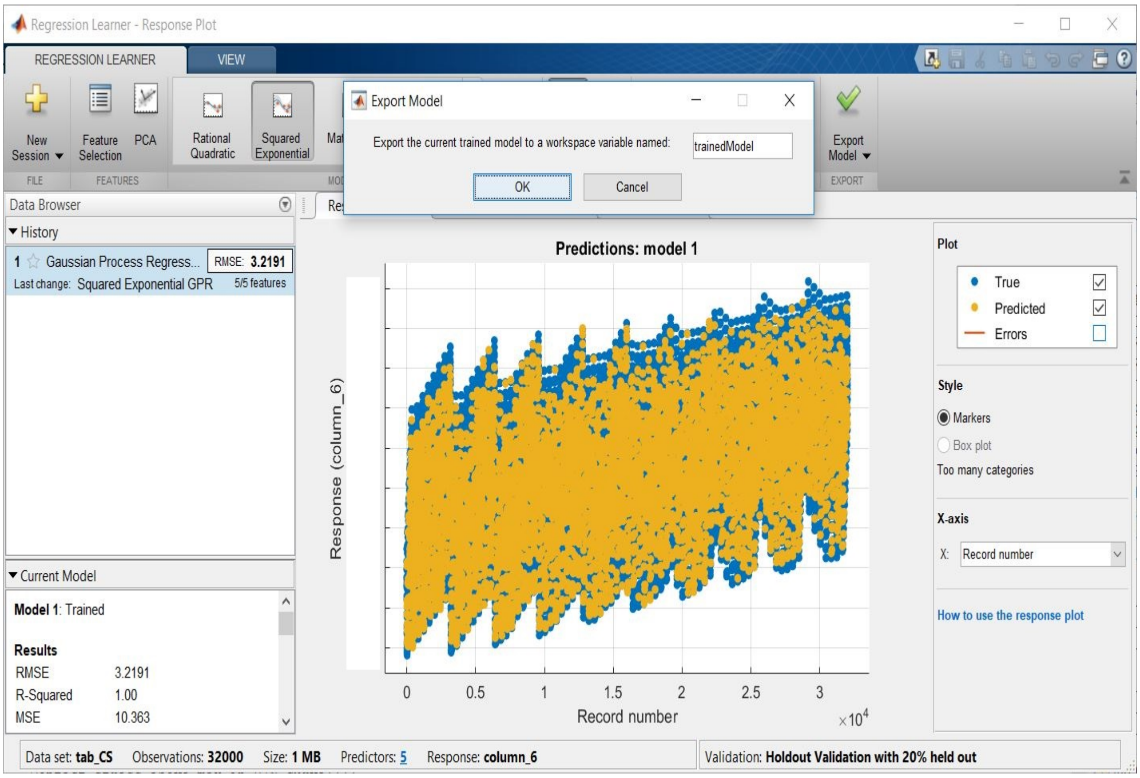


Figure 3.5: Export model

3.4 Metamodel Verification

Prior to optimization, the generated metamodel must be verified to check if it is mimicking the real model accurately. 52 design points within the design space were generated randomly. The objective functions were calculated in WindowsLDP for these design points. The gear model in WindowsLDP represent the real model in this case. The values generated by it is assumed to be the correct values. The objective functions were then calculated for the same design points using the generated metamodels. These values were then compared to the WindowsLDP generated values to determine if the metamodels are accurate enough to continue on to the optimization process. The results of the verification can be found in section 4.1.

3.5 Optimization

Optimization process was carried out to determine at what micro-geometry values the objective functions are at their minimum. The transmission error, contact stress and root stresses all have to be minimized and the metamodels of all these objective functions are contained within a vector. This vector is one of the input for the optimization function of MATLAB. The optimization function is *"gamultiobj"*. It's form is given by Eq. 3.6. [32]

$$x = gamultiobj(func, nvars, A, b, Aeq, beq, lb, ub, nonlcon, options) \tag{3.6}$$

The micro-geometry values that will be obtained at the end of the optimization iteration is x. The objective function vector is input in the place of *func*. The number of variables of the optimization problem in this case is five and *nvars* represents this. The design space of the micro-geometry variables can be defined by various constraints such as linear inequality constraints given by $A * x \leq b$, linear equality constraints given by $Aeq * x \leq beq$, non-linear constraints given by *nonlcon*. In this thesis, these constraints are not used to define the design space and hence are omitted. The other type of constraints used to set the design space are bound constraints defined by a minimum value in *lb* and a maximum value by *ub*. For five micro-geometry variables, the constraints *lb* and *ub* are vectors containing five values. The *options* in this function is used to change different available options for the genetic algorithm function such as changing the values of different algorithm stopping criteria, display settings while the algorithm is running, an initial solution as input to help start the algorithm (the default being zero), etc., and can be found in the MATLAB documentation for *gamultiobj*. [32]

A stopping criterion is very important to be setup while running a genetic algorithm optimization as it does not converge to a stop. The two stopping criteria considered in this thesis are function tolerance limit and maximum number of generations allowed. Function tolerance limit is the average relative change of the objective function values between iterations. Maximum number of generation is the maximum number of iterations that the algorithm will run. The function tolerance limit was set to 0 and 5000 maximum generations. The function tolerance condition was not satisfied and the algorithm stopped after running for 5000 iterations. A pareto set (non-dominated set) of solutions was obtained at the end of the optimization process.

3.6 Post-optimization analysis

The pareto solutions are all optimal solutions. When compared to each other, an improvement in one of the objective function will result in the deterioration of one or more other objective functions. But like any design, the designs obtained in the pareto set must satisfy the safety factor conditions setup for the gear set. So, as part of the post-optimization analysis, the safety factors for all the pareto solutions are calculated. They are calculated based on the S-N curve of the material and duty cycle applied for the gear set. The solutions that satisfy the critical cut-off safety factor values are separated from the remaining solutions for further processing.

TOPSIS - Technique for Order of Preference by Similarity to Ideal Solution, a well known method for multi-criteria decision making is used to rank the solutions from best to worst. The peak-peak transmission values and the safety factor values, which are the two criteria chosen for the process, are contained in a matrix called the decision matrix. The values in this decision matrix are then normalized to make comparison easier as they are of different units. This is now called the normalized matrix. Now weight values have to be multiplied to the normalized matrix. The total weight is considered to be 1 and equal weight distribution is when the two criteria are 0.5 each. In this thesis, the weight values were decided based on a set of conditions. The conditions are:

- If all the safety factor values are close to the cut-off values, the criterion safety factor gets a higher weight compared to the criterion peak-peak transmission error. This is because at this point choosing a solution with lesser stress values becomes more important than choosing the one with lesser peak-peak transmission value.

- If the safety factor values are much higher than the cut-off values, the criterion peak-peak transmission error gets a higher weight compared to the criterion safety factor. This is because, in an opposite way, choosing lesser peak-peak transmission value is more desirable than lesser stress values.

After the decision of weight values for each criterion, they were multiplied with the normalized matrix and the resultant matrix is called weighted matrix. Each element of this weighted matrix is represented by p_{ij}, where $i = 1, 2, ...m$ and $j = 1, 2, ...n$. In the next step, the ideal best and the ideal worst solution was chosen. The minimum of peak-peak transmission value and the maximum of safety value are the ideal best solutions given by p_j^+. In the opposite way, the maximum and minimum of the former and the latter are the ideal worst solutions given by p_j^-. The shortest distance or the euclidean distance of each element of the weighted matrix to ideal best and ideal worst was calculated. It is given by Eq. 3.7 - 3.8. The relative closeness of the pareto solution to the ideal best solution was calculated using the Eq. 3.9. This was considered as the score for the ranking of solutions with 1 being the best score and 0 the worst. The pareto solutions were then sorted according to the score in the descending order and presented as the result of the optimization routine. [33]

$$d_i^+ = \sqrt{\sum_{j=1}^{n}(p_j^+ - p_{ij})} \tag{3.7}$$

$$d_i^- = \sqrt{\sum_{j=1}^{n}(p_j^- - p_{ij})} \tag{3.8}$$

$$score = \frac{d_i^-}{d_i^+ + d_i^-} \tag{3.9}$$

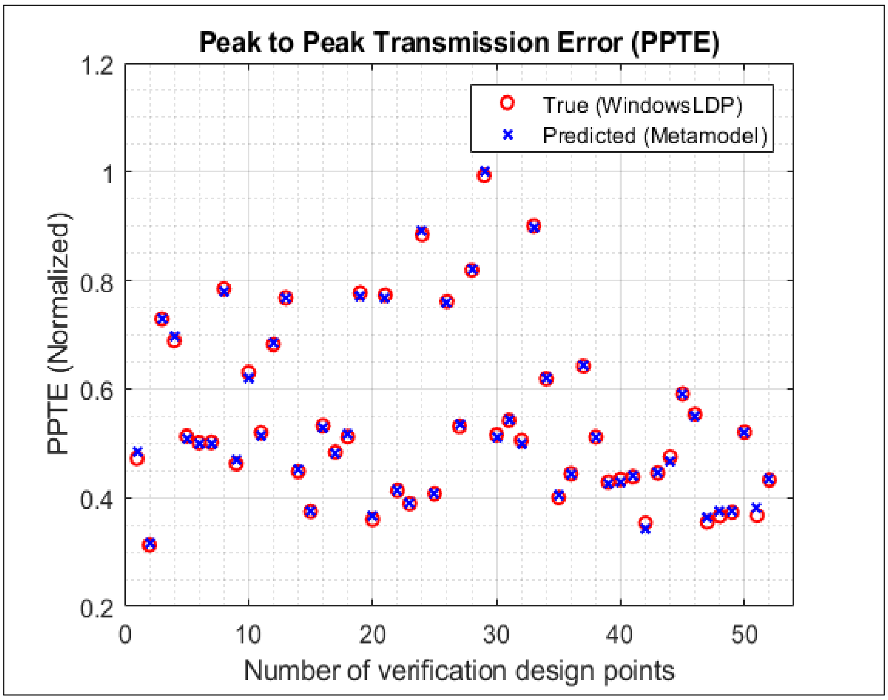
4
Results and Discussion

The results of the thesis are presented in this chapter along with discussion about the same.

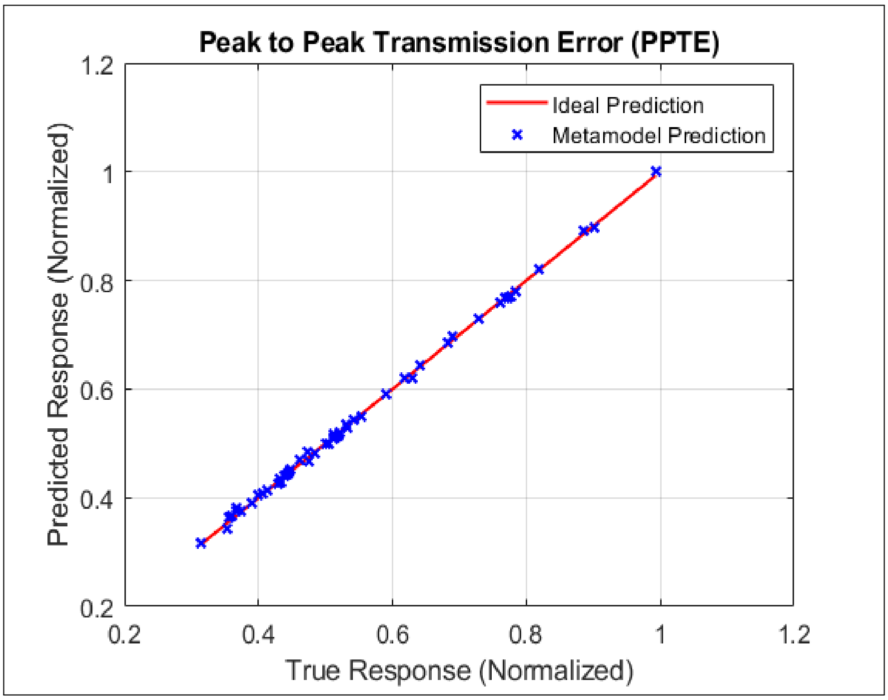
4.1 Metamodel Verification

The metamodels have to mimic the true model within the design space and calculate approximately equal objective function values for a given set of variables. In Fig. 4.1a, the x-axis represents the different design points within the design space. Each design point represents different micro-geometry values. The y-axis represents peak-peak transmission error (PPTE) with normalized values. For each design point, there are two values of PPTE, one calculated by the model in WindowsLDP (red circle) and another calculated by the metamodel (blue cross). This can also be visualized in another way as shown in Fig. 4.1b where the metamodel values are plotted against the true values(WindowsLDP Values). The red line indicates the ideal prediction. The blue crosses indicate the metamodel calculated values. Similarly, graphs for contact stress and root stress metamodels can be found in Fig. 4.2 to 4.4.

In this thesis, the objective function values calculated from WindowsLDP model is considered to be true. Lower deviations from the true values indicate higher accuracy of the metamodel. In Fig. 4.1a, the predicted values of the PPTE metamodel are equal or approximately equal to the true values for all the design points. In Fig. 4.1b, all the predicted values lie close to the ideal or true prediction line for smaller deviation and hence better accuracy of the metamodel. After a metamodel is generated, it is important to make this check otherwise the subsequent steps will carry errors and the optimal solution will not be correct. Through these graphical checks, the metamodels for PPTE, contact and root stresses are all deemed to be accurate in this case.

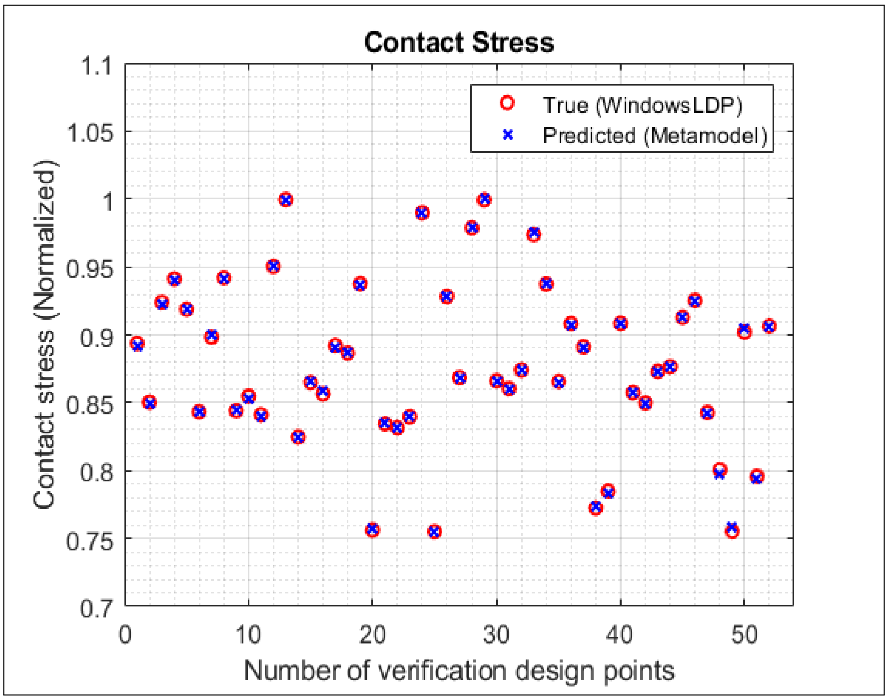


(a) Values calculated from WindowsLDP model and Metamodel for each design point

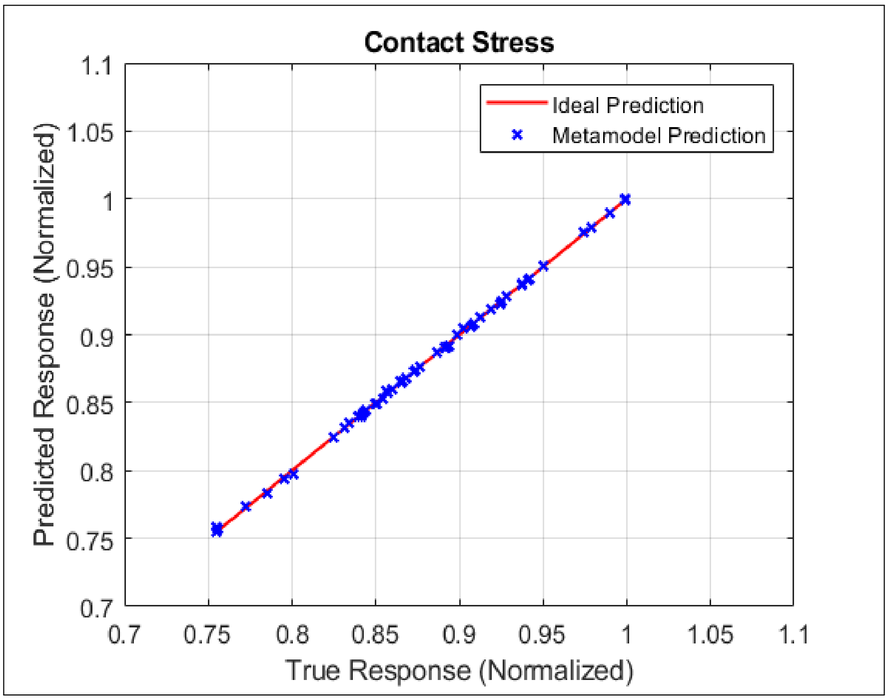


(b) Predicted Response Vs. True Response

Figure 4.1: Verification of PPTE Metamodel

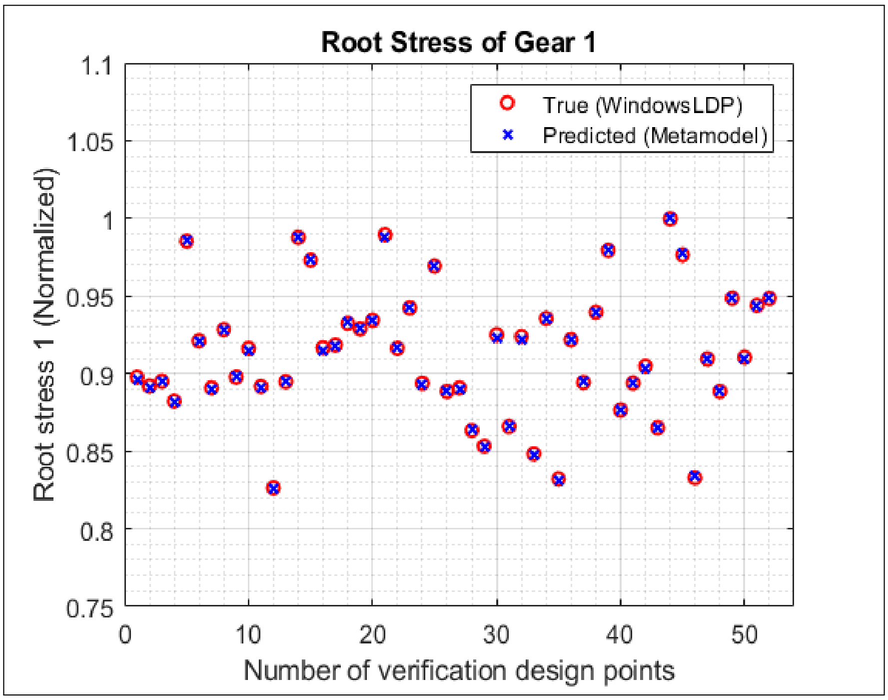


(a) Values calculated from WindowsLDP model and Metamodel for each design point

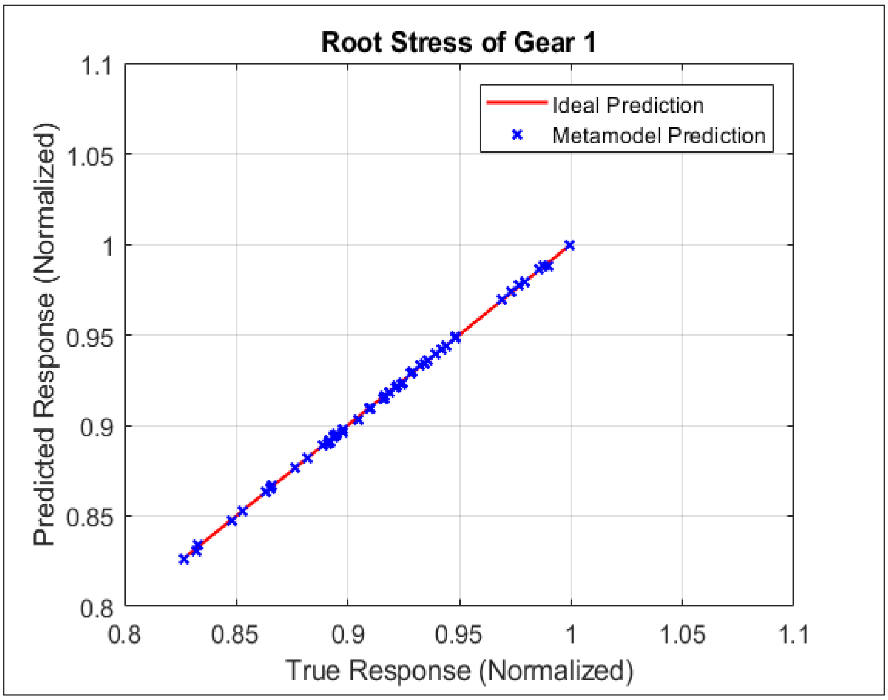


(b) Predicted Response Vs. True Response

Figure 4.2: Verification of Contact Stress Metamodel

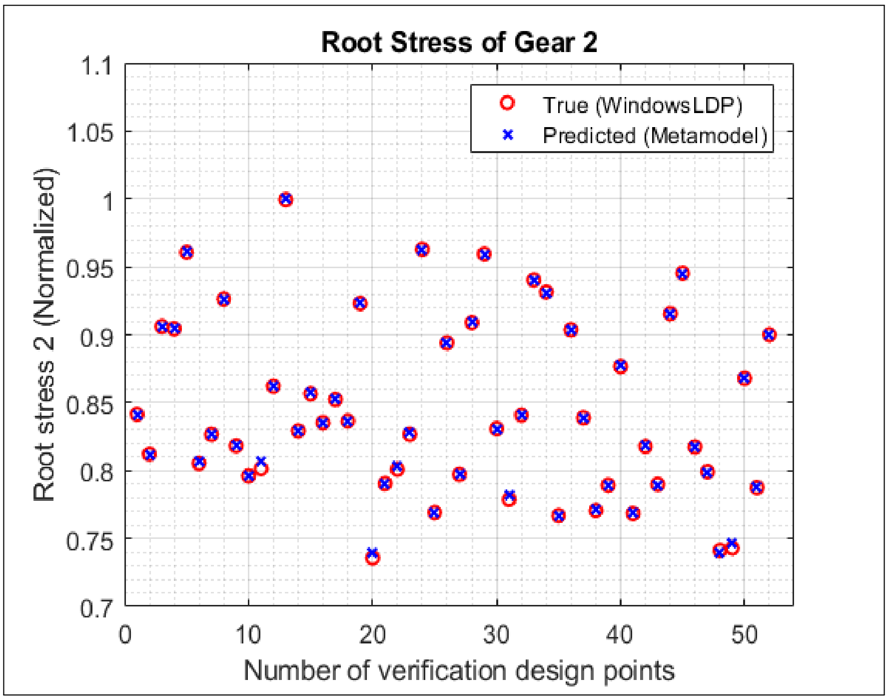


(a) Values calculated from WindowsLDP model and Metamodel for each design point

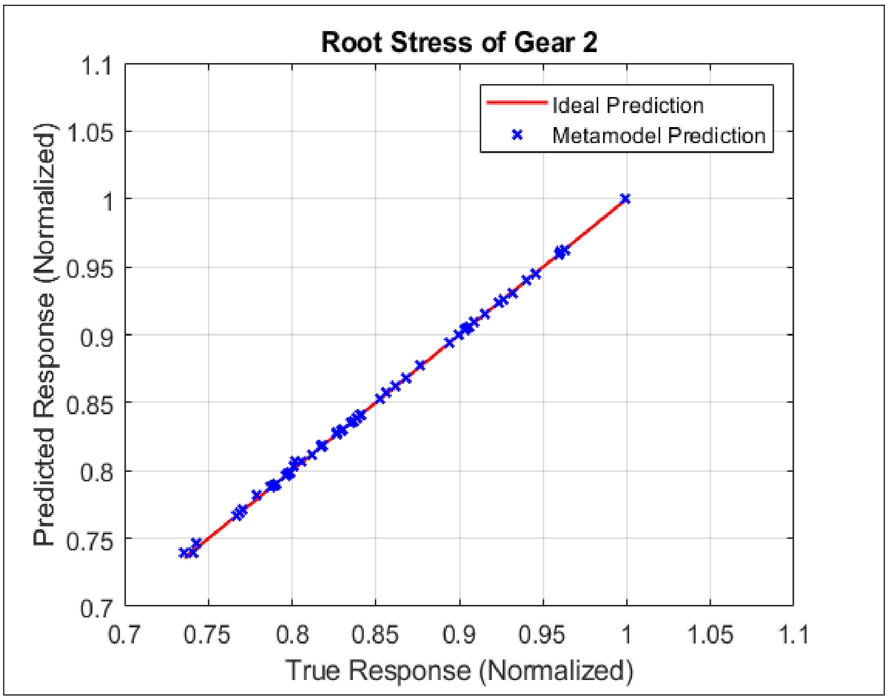


(b) Predicted Response Vs. True Response

Figure 4.3: Verification of Root Stress 1 Metamodel



(a) Values calculated from WindowsLDP model and Metamodel for each design point



(b) Predicted Response Vs. True Response

Figure 4.4: Verification of Root Stress 2 Metamodel

4.2 Benchmark vs Optimized Design

The objective functions of the optimized micro-geometry design is compared to the benchmark and a case of zero micro-geometry (only five that were considered for this thesis were made to be zero). In Fig. 4.5 to 4.8, the x-axis is represented by torque and the y-axis is represented by PPTE, contact stress and root stresses with normalized values. The optimized design is represented by Pareto Rank 1 as it is the highest ranked solution out of the set of pareto optimal solutions.

The micro-geometry values from highest ranked pareto optimal solution are substituted in the WindowsLDP model and the objective functions are calculated and plotted against the torque range. This is then compared to the benchmark and zero micro-geometry cases. Zero micro-geometry case is used just to emphasize the importance of micro-geometry in gear tooth design under varying loads and misalignments. In Fig. 4.5, 4.6 and 4.8, an improvement is seen in PPTE, contact stress and gear 2 root stress compared to the benchmark. A small increase is seen in gear 1 root stress in Fig. 4.7. This is considered to be normal as all the objective functions are conflicting to each other and improving one of them may worsen the others. Also, safety factor calculations and checks to determine the safety of design are integrated into the post optimization code. Hence, any increase in stresses compared to the benchmark is okay if the design satisfies the safety factor conditions. Transmission error can be minimized up to the limits of design safety thus preventing over designing in this aspect. In this case, the disadvantage of having zero micro-geometry can be seen in PPTE and maximum contact stress as they are very high. Even though the root stresses are lower, the gears will fail much faster than the other designs in terms of high contact stress due to pitting problems.

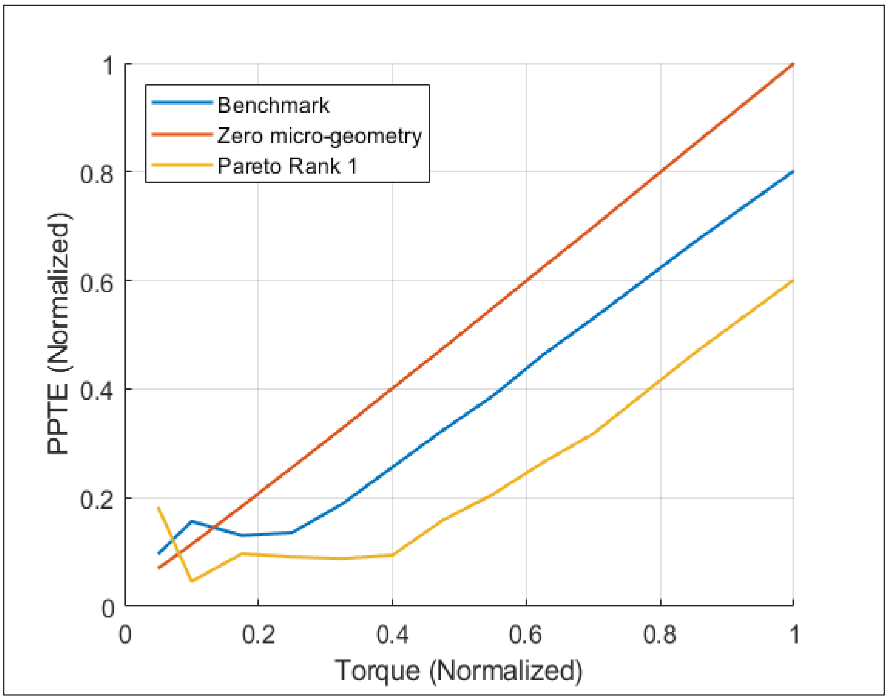


Figure 4.5: Peak-Peak Transmission Error

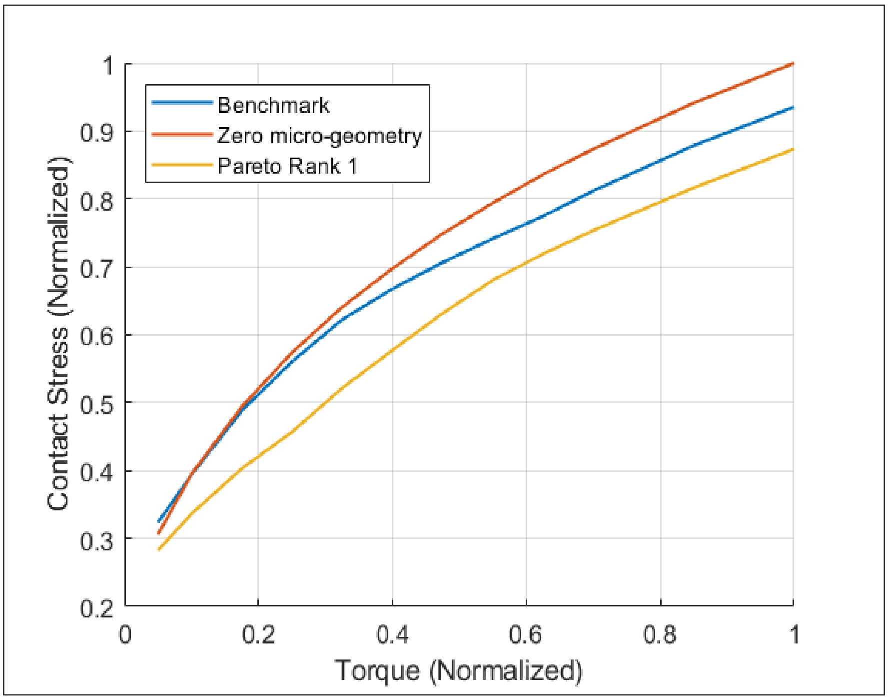


Figure 4.6: Contact Stress

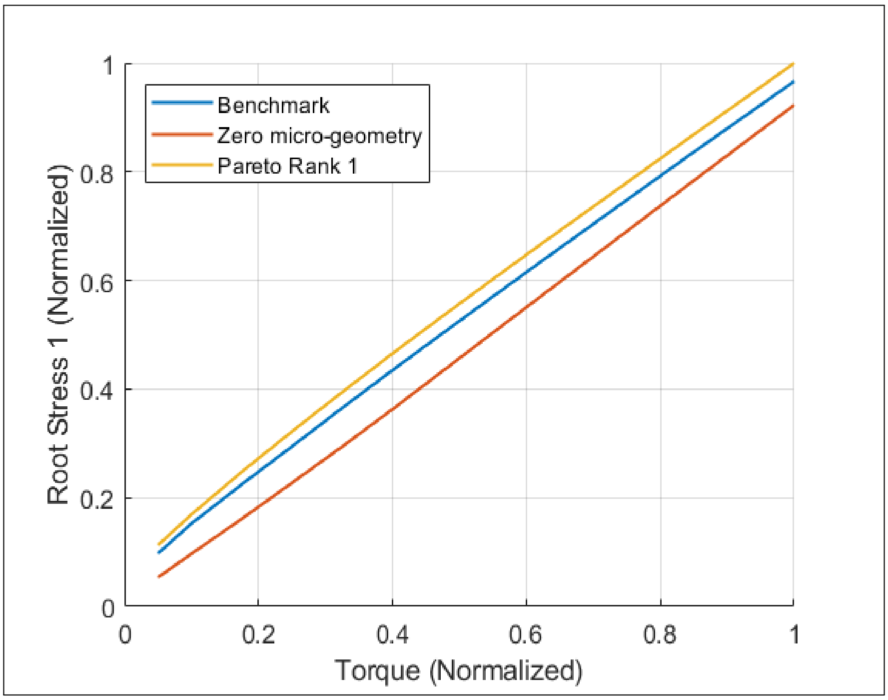


Figure 4.7: Root Stress of Gear 1

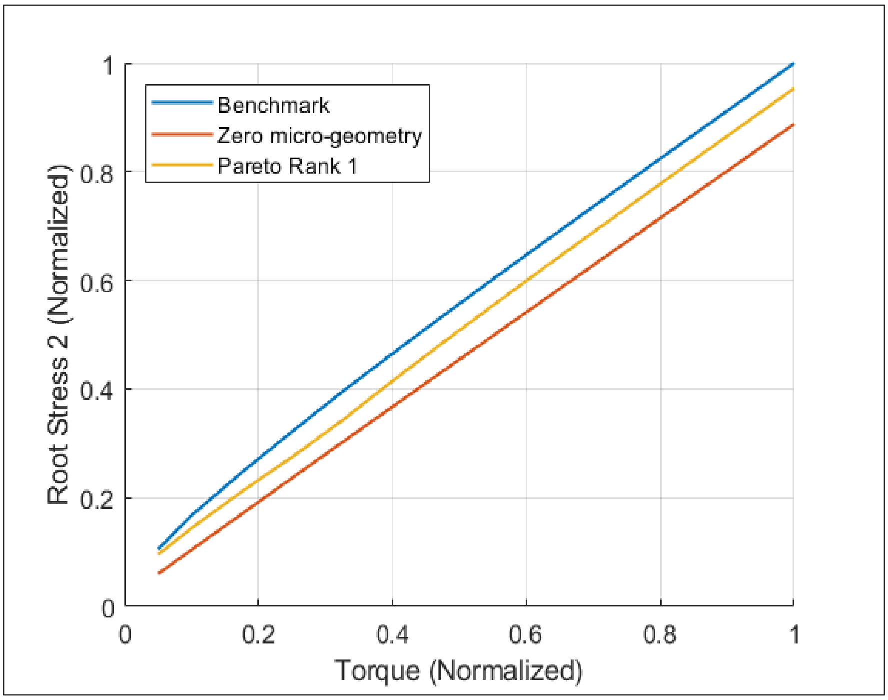


Figure 4.8: Root Stress of Gear 2

4.3 Pareto Solutions Comparison

The objective functions of the top five ranked pareto optimal designs are compared against each other in this section to visualize the conflicting nature of the objective functions and the set of non-dominated solutions.

Comparing pareto rank 1 and pareto rank 4 objective functions in Fig. 4.9 to 4.12, rank 1 has lower PPTE in low range torque while rank 4 has lower PPTE in high range torque. Also, the maximum contact stress and gear 1 maximum root stress is lower for rank 4 compared to rank 1 and vice versa for gear 2 maximum root stress. This clearly shows the conflicting nature of the objective functions and the difference caused because of this in the two optimal solutions. Ranking one solution above another depends on the requirements and engineer preference. If the contact stress has to be minimized at the cost of higher PPTE in low range torque, then rank 4 will ranked above rank 1. But in this thesis, preference is given to minimize PPTE at lower torque range and hence is given a higher weight compared to the stresses. But nevertheless the smaller weight given to the stresses comes into picture in the third solution as it has the highest contact and gear 2 root stresses even though it has relatively similar PPTE and is the reason for it to be ranked below the top 2 pareto solutions.

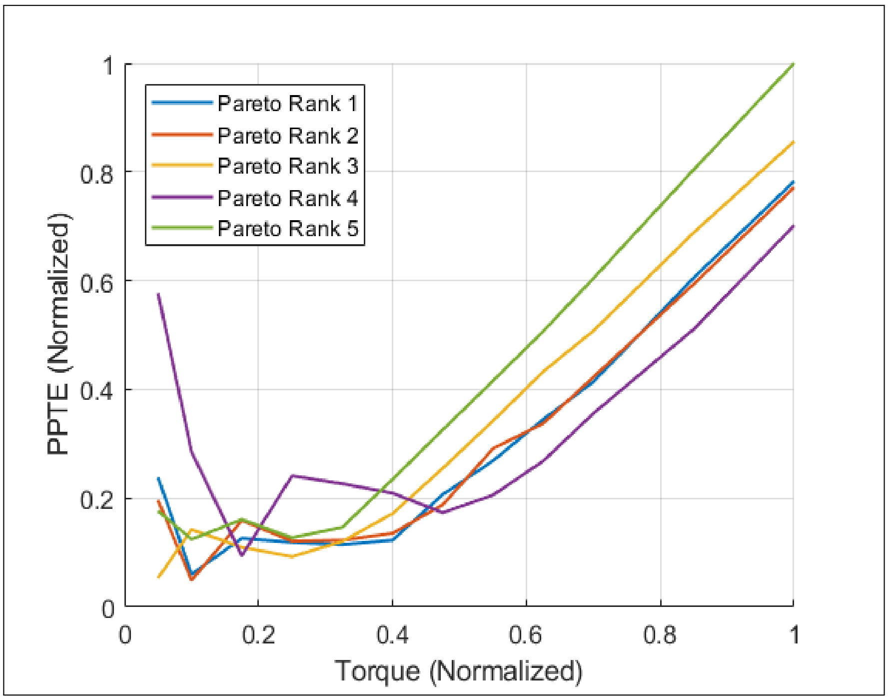


Figure 4.9: Peak-Peak Transmission Error

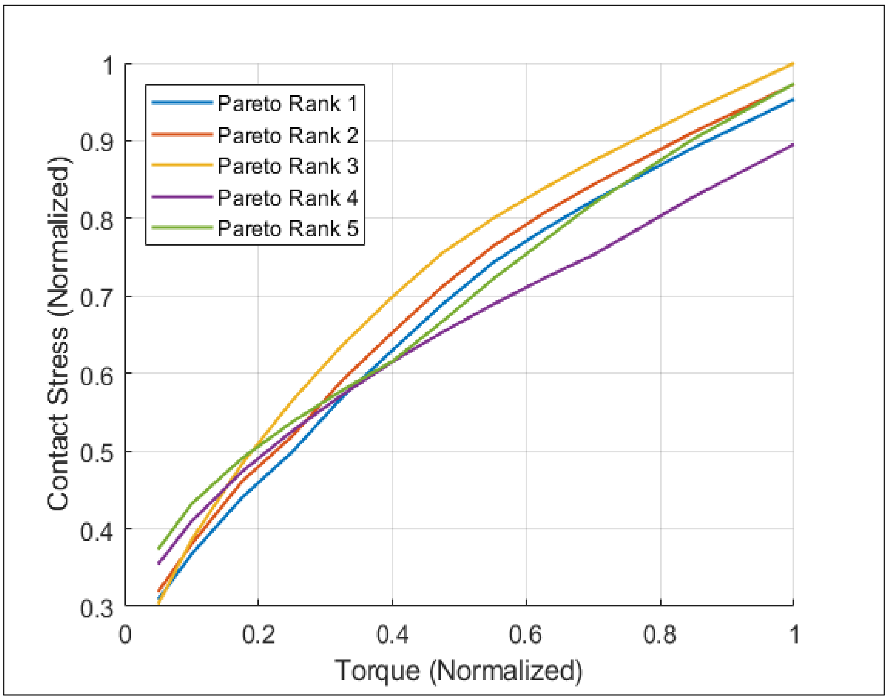


Figure 4.10: Contact Stress

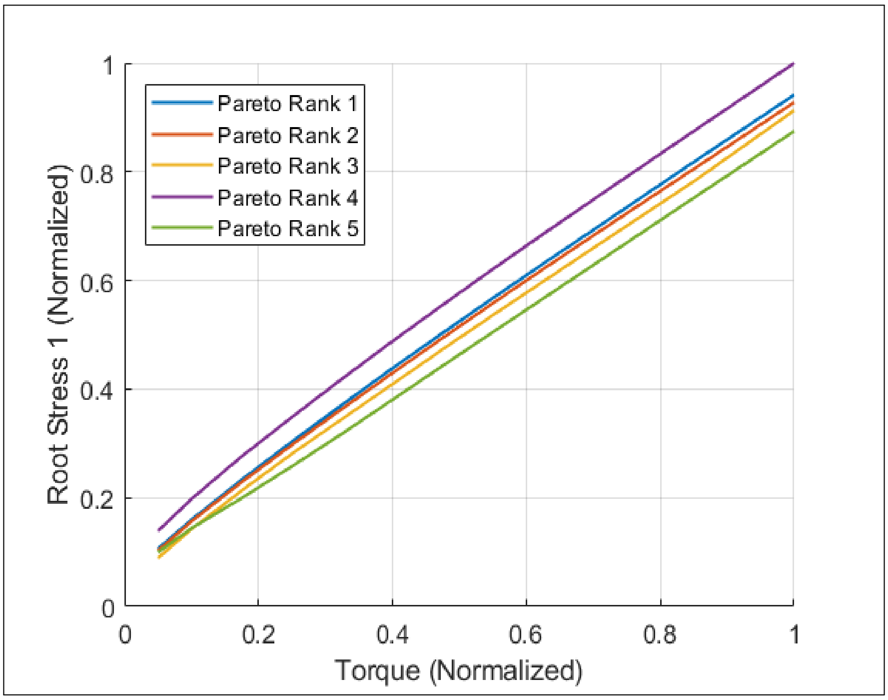


Figure 4.11: Root Stress of Gear 1

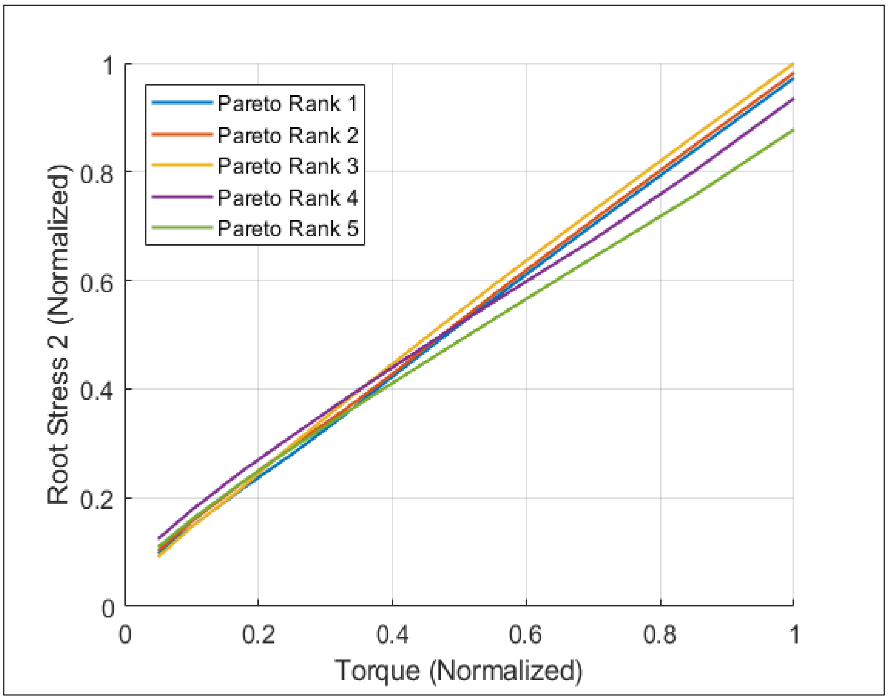


Figure 4.12: Root Stress of Gear 2

5

Conclusion and Future Work

Developing an optimization methodology for gear design was the main aim of this thesis. The methodology had to result in a pareto set of optimum micro-geometry design for the gear teeth such that gear whine, contact stress and root stresses were minimized. Safety factors for each of the pareto solution had to be calculated and checked for design safety. Also, the pareto solutions had to be ranked depending on weight distributed between the peak-peak transmission error and the safety factor. The methodology also had to be robust against the operating condition torque and also develop metamodels for the objective functions for reasons discussed in section 1.1.

The methodology for the optimization of gear design was successfully developed and is same as the methodology used in this thesis as outlined in section 1.3. A pareto set of optimum designs were obtained and after post-optimization process the solutions were ranked highest to lowest. The objective functions of the highest ranked solutions were compared against the benchmark. The peak-peak transmission error and thus gear whine was minimized. Contact stress and root stress of gear 2 was also minimized. But, the root stress of gear 1 was found to be slightly higher than the benchmark. The stresses also satisfied the design safety conditions. The methodology also helps to arrive at a unique solution and will not vary engineer to engineer. It also saves time and effort compared to trial and error methods followed previously.

As part of the future work, further development of the methodology in the thesis to incorporate more design variables is possible. Robustness against production tolerance can be included to determine a solution that is unaffected by the same. New metamodel generation methods can be explored to cut down the time involved in this process. Lesser number of design points can be used for generation of metamodels to save time at the cost of some accuracy.

Bibliography

[1] Ghribi Dhafer & Bruyere Jerome & Velex Philippe & Octrue Michel & Haddar Mohamed. Multi-objective optimization of gear tooth profile modifications. *Design and Modeling of Mechanical Systems*, pages 189–197, 2013.

[2] Chan IL Park. Multi-objective optimization of the tooth surface in helical gears using design of experiment and the response surface method. *Journal of Mechanical Science and Technology*, 24(3):823–829, 2010.

[3] Alessio Artoni & Massimo Guiggiani & Ahmet Kahraman & Jonny Harianto. Robust optimization of cylindrical gear tooth surface modifications within ranges of torque and misalignments. volume 135, 08 2013.

[4] Jakub A. Korta & Domenico Mundo. Multi-objective micro-geometry optimization of gear tooth supported by response surface methodology. *Mechanism and Machine Theory*, pages 278–295, 2017.

[5] Giorgio Bonori & Marco Barbieri & Francesco Pellicano. Optimum profile modifications of spur gears by means of genetic algorithms. *Journal of Sound and Vibration*, pages 603–616, 2008.

[6] Ghribi Dhafer & Bruyere Jerome & Velex Philippe & Octrue Michel & Haddar Mohamed. Robust optimization of gear tooth modifications using a genetic algorithm. *Condition Monitoring of Machinery in Non-Stationary Operations*, pages 589–597, 2012.

[7] Layue Zhao & Minggang Du & Yang Yang. Optimizing gear micro geometry for minimum transmission error when considering manufacturing deviation. *International Journal of Materials, Mechanics and Manufacturing*, 6(1), 2018.

[8] Alessio Artoni &Marco Gabiccini & Massimo Guiggiani & Ahmet Kahraman. Multi-objective ease-off optimization of hypoid gears for their efficiency, noise and durability performances. volume 133, 12 2011.

[9] Robert C. Juvinall & Kurt M. Marshek. *The Fundamentals of Machine Component Design.* John Wiley Sons, 2012.

[10] Stephen P. Radzevich. *Handbook of Practical Gear Design and Manufacture.* CRC Press, 2012.

[11] *Gear Nomenclature, Definitions of Terms with Symbols.* American Gear Manufacturers Association. ANSI/AGMA 1012-G05, 2005.

[12] Richard G. Budynas & J. Keith Nisbett. *Shigley's Mechanical Engineering Design.* McGraw-Hill, 2011.

[13] Dan Thurman. Calculating SAP and TIF. *Gear Technology*, 1999. Available at `https://www.geartechnology.com/issues/0999x/Thurman.pdf` (Visited on 2019/01/28).

[14] Jonny Harianto & Dr. Donald R. Houser. A methodology for obtaining optimum gear tooth micro-topographies for noise and stress minimization over a broad operating torque range. 01 2007.

[15] J. Lange. How are you dealing with the bias error in your helical gears? *Gear Technology*, 2009. Available at `https://www.geartechnology.com/issues/0509x/lange.pdf` (Visited on 2019/02/13).

[16] Donald R. Houser. The effect of manufacturing microgeometry variations on the load distribution factor and on gear contact and root stresses. 2009.

[17] Graduate R. W. Gregory & S. L. Harris & R. G. Munro. Dynamic behaviour of spur gears. *Proceedings of the Institution of Mechanical Engineers*, 178(1):207–218, 1963.

[18] A. Kahraman & Ravpreet Singh. Non-linear dynamics of a spur gear pair. *Journal of Sound and Vibration*, 142:49–75, 10 1990.

[19] Stephen L. Harris. Dynamic loads on the teeth of spur gears. *Proceedings of the Institution of Mechanical Engineers*, 172(1):87–112, 1958.

[20] D. B. Welbourn. Fundamental knowledge of gear noise - a survey. *Institution of Mechanical Engineers, Conference Publications*, pages 9–14, 01 1979.

[21] A. Ivar Nilsson. Gear whine noise excitation model. Master's thesis, Chalmers University of Technology, Gothenburg, Sweden, 2013.

[22] Xiaoyin Zhu. *Tutorial on Hertz Contact Stress*, 2012. Available at `https://wp.optics.arizona.edu/optomech/wp-content/uploads/sites/53/2016/10/OPTI-521-Tutorial-on-Hertz-contact-stress-Xiaoyin-Zhu.pdf` (Visited on 2019/03/09).

[23] *Calculation of load capacity of spur and helical gears. Part 3: Calculation of tooth bending strength.* ISO 6336-3, 2006. Available at `https://www.sis.se/api/document/preview/907849` (Visited on 2019/03/09).

[24] Thomas Tobie & Frank Hippenstiel & Hardy Mohrbacher. Optimizing gear performance by alloy modification of carburizing steels. *Metals*, 7:415, 10 2017.

[25] Miguel Costas & Jacobo Diaz & Luis Romera & Santiago Hernandez. A multi-objective surrogate-based optimization of the crashworthiness of a hybrid impact absorber. *International Journal of Mechanical Sciences*, 88, 11 2014.

[26] Jakub Korta & Rosario Raniolo & Marco Danti & Izabela Kowarska & Tadesusz Uhl. Multi-objective optimization of a car body structure. *SAE International Journal of Passenger Cars - Mechanical Systems*, 5:1143–1152, 05 2012.

[27] MathWorks. *Kernel (Covariance) Function Options.* Available at `https://se.mathworks.com/help/stats/kernel-covariance-function-options.html` (Visited on 2019/03/07).

[28] Kalyanmoy Deb. Multi-objective optimization using evolutionary algorithms: An introduction, 2011.

[29] Mohsen Ebrahimi & Milad M. Rabieh. Fast and accurate reservoir modeling: Genetic algorithm versus direct method. Jan 2012.

[30] K. Sundararajan. Isixsigma. *Design of Experiments - A primer.* Available at `https://www.isixsigma.com/tools-templates/design-of-experiments-doe/design-experiments-%E2%90%93-primer/` (Visited on 2019/04/03).

[31] Eric W. Weisstein. Triangular distribution. From MathWorld–A Wolfram Web Resource. `http://mathworld.wolfram.com/TriangularDistribution.html` (Visited on 2019/02/26).

[32] MathWorks. *gamultiobj*. Available at `https://se.mathworks.com/help/gads/gamultiobj.html` (Visited on 2019/03/18).

[33] Renato A.Krohlingab & André G.C.Pachecob. A-topsis – an approach based on topsis for ranking evolutionary algorithms. *Procedia Computer Science*, 55:308–317, 2015.