Development of an Integrated Inspection System for Additive and Subtractive Hybrid Manufacturing

by

Cody Berry

A Thesis Submitted in Partial Fulfillment of the Requirements for the Degree of Master of Applied Science in The Faculty of Engineering and Applied Science Mechanical Engineering

University of Ontario Institute of Technology

January, 2018

© Cody Berry, 2018
ABSTRACT

Using closed-loop inspection systems to modify coordinate metrology tasks for specific engineering applications has been a demanding research subject over the last decade. The objective of this thesis is to develop an Integrated Inspection System (IIS) that uses these tasks for applications in systems combining additive and subtractive manufacturing technologies, commonly referred to as hybrid manufacturing. Hybrid manufacturing has opened a new and innovative avenue in product development, and also in product repair and maintenance. One of the areas where it can excel is in the repair of dies and moulds. This is due to the ability to add material to worn out areas, then subtract the excess material to return to the ideal geometry without affecting a large area of the piece. Previously developed coordinate metrology techniques can be modified to aid in the minimization of the cost of repair. A method using skin modelling techniques and weighted total least squares was utilized to determine the geometry of the repaired zone. These methods were combined to create the developed system which minimizes the repair cost. The best result of this method was then used as an initial condition for an optimization algorithm resulting in the optimal solution. The developed system produced cost reductions in all tested circumstances, with the best results found in surfaces with large, non-uniform errors. The developed system can be implemented and customized for various hybrid manufacturing applications in the pursuit of lower repair costs.
DEDICATION

To my parents, for all their support.
ACKNOWLEDGMENTS

I would like to thank my supervisor, Dr. Ahmad Barari, for his guidance and support. I would also like to thank my committee members, Dr. Ibrahim Dincer and Dr. Haoxiang Lang, for their time.
TABLE OF CONTENTS

Abstract ................................................................................................................................ii

Dedication ........................................................................................................................... iii

Acknowledgments ............................................................................................................... iv

List of Tables ...................................................................................................................... vii

List of Figures .................................................................................................................... viii

List of Symbols and Abbreviations ....................................................................................... x

Symbols ............................................................................................................................ x

Abbreviations ................................................................................................................. xii

1. Introduction ................................................................................................................. 1

2. Literature Review ........................................................................................................ 5

   2.1. Sources of Computational Uncertainty in Coordinate Metrology ....................... 5

   2.2. Point Measurement Planning .............................................................................. 7

   2.3. Substitute Geometry Evaluation ....................................................................... 13

   2.4. Deviation Zone Estimation ............................................................................... 17

   2.5. Cost of Repair – Reasoning for integrative ...................................................... 19

   2.6. Directed Energy Deposition .............................................................................. 20

       2.6.1. Laser Metal Deposition (LMD) ................................................................. 20

       2.6.2. Electrospark Deposition .......................................................................... 26

       2.6.3. Micro Plasma Transferred Arc ................................................................. 26

       2.6.4. Powder Bed Fusion .................................................................................. 27

   2.7. Possible Paths Going Forward ............................................................................ 28
3. Methodology .................................................................................................................. 29
  3.1. Adaptive Sampling towards PMP ................................................................. 29
  3.2. Weighted Total Least Squares Fitting toward SGE ........................................ 29
    3.2.1. Volume Calculation ..................................................................................... 36
  3.3. Calibration of Skin Modelling towards DZE ................................................... 41
  3.4. Integrated Process ......................................................................................... 44
    3.4.1. Weighting Process ................................................................................... 46
    3.4.2. Bump Code for Oscillations ...................................................................... 55
  3.5. Heuristic Optimization ................................................................................. 56
4. Implementation and Results .................................................................................... 58
  4.1. Validation ....................................................................................................... 58
  4.2. Results ........................................................................................................... 68
    4.2.1. 3D Printed Surface ................................................................................... 68
    4.2.2. Gouged Wax Piece .................................................................................. 72
    4.2.3. Metal Piece with Two Holes ..................................................................... 75
    4.2.4. NURBS Surface ........................................................................................ 77
    4.2.5. Staircase Model ....................................................................................... 80
    4.2.6. Results Comparison ................................................................................ 91
5. Conclusions and Recommendations ....................................................................... 92
  5.1. Conclusions ................................................................................................. 92
  5.2. Recommendations ....................................................................................... 93
6. References ........................................................................................................... 94
LIST OF TABLES

Table 3-1 Comparison of Small Discrepancy Tests ........................................................... 51
Table 3-2 Comparison of Large Discrepancy Tests ........................................................... 54
Table 3-3 Results of optimization for two offset planes ................................................... 57
Table 4-1 Final Results for each case ................................................................................ 91
LIST OF FIGURES

Figure 1-1 Traditional method of coordinate metrology .................................................... 2
Figure 1-2 Proposed revision of Coordinate Metrology tasks ............................................. 4
Figure 2-3 Hammersley distribution for 20 points ............................................................... 8
Figure 2-4 Spiral Method for Point Selection [19] ............................................................. 11
Figure 2-5 HamSpi Method for Point Selection [19] ......................................................... 12
Figure 2-6 Example of a Skin Model [34] .......................................................................... 18
Figure 2-7: Micrographs of cross-sections through selected samples [40] ...................... 22
Figure 2-8 Hardness profiles in laser welds [44] ............................................................... 25
Figure 3-1 Example case for Volume Calculation ............................................................. 36
Figure 3-2 Base of Pyramid for Volume Calculation ......................................................... 38
Figure 3-3 Error Graph for Skin Modelling with unfiltered results ................................. 42
Figure 3-4 Error Graph for Skin Modelling with filtered results ...................................... 43
Figure 3-5 Integrated process flowchart ........................................................................ 45
Figure 3-6 Net Volume Method – small discrepancy test results ..................................... 49
Figure 3-7 Absolute Volume Method – small discrepancy test results ............................ 50
Figure 3-8 Volume Dependent Weighting – small discrepancy test ............................... 50
Figure 3-9 Centroid Method – small discrepancy results ............................................... 51
Figure 3-10 Net Volume Method – large discrepancy test results .................................... 52
Figure 3-11 Absolute Volume Method – large discrepancy test results ......................... 53
Figure 3-12 Volume Dependent Weighting – large discrepancy test ............................. 53
Figure 3-13 Centroid Method – large discrepancy test .................................................... 54
Figure 3-14 Optimization results for two offset plates .................................................... 55
Figure 3-15 Two Offset Planes with Fminsearch results ................................................ 57
Figure 4-1 Results of Integrated Process on Ideal Plane .................................................. 59
Figure 4-2 Two offset planes ............................................................................................. 60
Figure 4-3 Trigonometric Model for Volume Calculation Validation ............................... 61
Figure 4-4 Results for Volume Calculation Verification .................................................... 64
Figure 4-5 Plane orientation when weights are equal ...................................................... 65
Figure 4-6 Plane orientation for additive cost of 4, subtractive cost of 2 ....................... 66
Figure 4-7 Plane orientation for additive cost of 8, subtractive cost of 2 ....................... 66
Figure 4-8 Plane orientation for additive cost of 16, subtractive cost of 2 ....................... 67
Figure 4-9 Plane orientation for additive cost of 32, subtractive cost of 2 ....................... 67
Figure 4-10 3D Printed Flat Surface .................................................................................. 69
Figure 4-11 Skin Model for 3D Printed Surface ............................................................... 70
Figure 4-12 Repair Cost for 3D Printed Surface ............................................................... 70
Figure 4-13 3D Printed Surface – Fminsearch Results .................................................... 71
Figure 4-14 Gouged Wax Surface ..................................................................................... 72
Figure 4-15 Gouged Wax – Skin Model ............................................................................ 73
LIST OF SYMBOLS AND ABBREVIATIONS

SYMBOLS

\( \bar{P} \) – Centroid
\( \bar{x}, \bar{y}, \bar{z} \) – Mean of Points
\( \nabla \) – Gradient
a,b,c,d – Distance from Data to Fit Plane
\( A_b \) – Base Area
\( A_t \) – Projected Area
B – Project Point
d\( \perp \) – Perpendicular Distance of Point to Plane
e – Distance from Edge of Test Plane to Middle of Fit Plane Along Fit Plane
F – Objective Function
g – Distance between Candidates
G – Optimization Bound
h – Height
i – Index of summation
I – Intersection Point
j – Length of Plane
K – nxn Square Matrix
\( K_{uu}, K_{vv} \) – Error Propagation Factors
l – Euclidean Distance Between Edges of Planes
M – Mean Vector
\( M^TM \) – Covariance Matrix
n – Number of Points
P – Point
p – Point
q – Current Result
Q – Eigenvectors
r – Distance from Origin
S – Singular Values
t – Scalar Parameter
U – Left Singular Vectors
u – Vertical Distance Between Planes
V – Right Singular Vectors
\( V_n \) – Negative Volume
\( V_p \) – Positive Volume
\( V_{pr} \) – Prism Volume
\( V_{py} \) – Pyramid Volume
\( V_T \) – Total Volume
w – Weight
X – Objective Function
x, y, z – Point Coordinates
y – Width of Plane
\( z_c \) – Centroid of Heights
\( \alpha, \beta, \gamma \) – Principal Angles
\( \alpha\epsilon \) – Randomization Factor
\( \Gamma \) – Light Absorption Coefficient
\( \eta \) – Attractiveness Coefficient
\( \Lambda \) – Eigenvalues
\( \lambda \) – Static Point
\( \Pi \) – Plane
\( \pi \) – Vector Containing Principal Angles
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>µPTA</td>
<td>Micro Plasma Transferred Arc</td>
</tr>
<tr>
<td>AVM</td>
<td>Absolute Volume Method</td>
</tr>
<tr>
<td>CAD</td>
<td>Computer Aided Design</td>
</tr>
<tr>
<td>CAM</td>
<td>Computer Aided Manufacturing</td>
</tr>
<tr>
<td>CM</td>
<td>Centroid Method</td>
</tr>
<tr>
<td>CMM</td>
<td>Coordinate Measuring Machine</td>
</tr>
<tr>
<td>CNC</td>
<td>Computer Numerical Control</td>
</tr>
<tr>
<td>DED</td>
<td>Directed Energy Deposition</td>
</tr>
<tr>
<td>DZE</td>
<td>Deviation Zone Estimation</td>
</tr>
<tr>
<td>ED</td>
<td>Electrospark Deposition</td>
</tr>
<tr>
<td>FEM</td>
<td>Finite Element Method</td>
</tr>
<tr>
<td>GMAW</td>
<td>Gas Metal Arc Welding</td>
</tr>
<tr>
<td>IIS</td>
<td>Integrated Inspection System</td>
</tr>
<tr>
<td>IR</td>
<td>Infrared</td>
</tr>
<tr>
<td>ISO</td>
<td>International Organization for Standardization</td>
</tr>
<tr>
<td>LMD</td>
<td>Laser Metal Deposition</td>
</tr>
<tr>
<td>NSR</td>
<td>Neighbourhood Search for Representative</td>
</tr>
<tr>
<td>NURBS</td>
<td>Non-Uniform Rational B-Spline</td>
</tr>
<tr>
<td>NVM</td>
<td>Net Volume Method</td>
</tr>
<tr>
<td>PDF</td>
<td>Probability Density Function</td>
</tr>
<tr>
<td>PMP</td>
<td>Point Measurement Planning</td>
</tr>
<tr>
<td>SGE</td>
<td>Substitute Geometry Evaluation</td>
</tr>
<tr>
<td>SLS</td>
<td>Selective Laser Sintering</td>
</tr>
<tr>
<td>TIG</td>
<td>Tungsten Inert Gas</td>
</tr>
<tr>
<td>TLS</td>
<td>Total Least Squares</td>
</tr>
<tr>
<td>VDW</td>
<td>Volume Dependent Weighting</td>
</tr>
<tr>
<td>WTLS</td>
<td>Weighted Total Least Squares</td>
</tr>
</tbody>
</table>
1. INTRODUCTION

Coordinate metrology is the process of taking a manufactured piece, and verifying that it was made to conform to standards. In coordinate metrology, manufactured parts are measured and analyzed to ensure they have been manufactured to tolerance. This process involves three major tasks: Point Measurement Planning (PMP), Substitute Geometry Evaluation (SGE) and Deviation Zone Estimation (DZE). PMP involves determining the location of data points on a surface. SGE is the fitting of the ideal CAD model to the data points found in PMP. DZE is using the differences between the data points and the ideal surface to determine the overall error in the part. Using these tasks, the choice to either dispose of the part or perform error compensation techniques, such as re-machining or additive repair can be made. In general, these processes are performed sequentially, and data in a future process is not used to refine past processes. This can be seen in Figure 1-1.

As computing power has increased, it has become possible to integrate the coordinate metrology tasks. This integration allows the different steps of the process to inform each other, with results of one section feeding into another. The processes need to be modified so that they can take in new data and modify the original results. Some work has already been done to achieve the integration of these tasks. The cost and feasibility of repair is also something that needs to be considered with these processes.
Repair cost is a large factor when determining whether or not a part needs to be repaired or replaced. By minimizing the cost of repair, it becomes a more attractive option. This minimization can also be used in the manufacturing process if something goes wrong. With traditional techniques, material could only be removed when an error was introduced into a part. This limited the ability to minimize the cost of repair. For example, in the undercut situation, where material had been left over after a pass that should have been removed, it would be fairly trivial to remove the excess material. However, if the
overcut situation occurred, and too much material had been removed, there was no easy and cost effective way to repair the error. With the advent of hybrid manufacturing, that is manufacturing systems that incorporate both additive and subtractive processes, a paradigm shift has begun to occur. Now, in the overcut condition, you can simply add more material and then remove it to return to tolerance. With the ability to quickly switch between adding and subtracting material, the way parts are manufactured is changing. This also means that how parts are repaired is changed. Whenever there is shift as drastic as this in a well used and long lasting system, other systems designed to support it have to change also. Therefore, a new system to minimize the cost of repair using hybrid manufacturing technologies is proposed.

In order to accomplish this minimization, several processes will need to be integrated with one another. PMP methods will be used to get a sample set that accurately represents the part to be repaired. A DZE method will be used to take the sample set and convert it into a representative geometry that can be analyzed in greater detail. SGE will be used to determine where a plane needs to be fit in order to minimize the cost of repair. By combining all of these processes, an integrated system will be created. The end goal of this thesis is to have a developed framework that can be used to minimize the cost of repair for a planar surface. An example of this modified system can be seen in Figure 1-2.
This thesis will proceed as follows. First, a review of the work done in the fields of coordinate metrology and additive repair will be conducted. The methodologies that will be used to develop the integrated process will then be created and explained. The integrated system will then be developed. The created processes will then be validated using some basic data sets. Then, various case studies will be conducted using the
developed system, and the results will be examined and discussed. Finally, proposals for the improvement of the system will be examined.

2. LITERATURE REVIEW

In this section, a review of existing literature related to this thesis will be conducted. First, different aspects of coordinate metrology will be discussed. Then, different methods of additive repair will be discussed.

2.1. SOURCES OF COMPUTATIONAL UNCERTAINTY IN COORDINATE METROLOGY

Reducing uncertainty is the goal of coordinate metrology. By creating knowledge corresponding to the geometric and dimensional features of a work piece, the uncertainty of the manufacturing processes used to create it can be reduced. Because all measurement devices, such as tactile probes and high density laser scanners, are imperfect, there are also uncertainties introduced through measuring a work piece. The sources of inspection uncertainties are well described in a paper by Barari et al. [1]. This includes errors due to the kinematics of the coordinate measuring machine (CMM), the effects of the probing sensor, operator error, and incorrect datum selection, environmental effects including temperature, vibration, and light, as well as computational error. The effects of computation errors have been underestimated by researchers.
In 2011, Barari [2] showed that the uncertainty due to computational tasks can be as large as the total uncertainty caused by all other physical sources. This shows a need for greater research into minimizing the effects of computational uncertainty.

Operator skill also plays an important role in inspection, and how they decide to inspect a part can lead to large changes in the final results. There are several major parameters that need to be decided by the operator: the number of sample points, the location of sample points, fitting criteria, and the method for representing the deviation zone. In an integrated inspection system (IIS), the link between the traditional computational tasks allows selection of these parameters using data that would otherwise go unused in common practice. Upstream activities should be used to determine the optimum settings of these tasks. For example, the information from the manufacturing process should be a factor in determining the number and density of sampled points. This would require CAM data alongside the model or CAD information of the part, as shown in [3]–[6]. It was also shown that a feedback system can be an effective way to share useful data between the computational tasks [7], [8].

Previous researchers have reported several very important closed loops between DZE-PMP, SGE-PMP, and DZE-SGE. Examples of closed loops between SGE and PMP can be seen in [9]–[11]. In [9], capturing a new sample point or rejecting an already captured sample point was dynamically decided based on progress in the fitting process which eventually became a guided search for sampling. The information required for sampling was generated and revised dynamically by estimating the probability density function of geometric deviations using the Parzen-window method. A similar approach was adopted
in [11] but instead of trying to recognize the real probability density function of the
geometric deviations through the process, it was assumed that the probability density
function has a Gaussian distribution. [10] and [12] used a virtual sampling technique to
analyze high density laser scan data and combine PMP and SGE. Instead of using a PDF,
sites were chosen using pre-existing sampling strategies, then the average of all the
samples in the neighbourhood were chosen as a representative point. These new points
were then used to fit a plane, and sampling was redone until the method converged.

2.2. POINT MEASUREMENT PLANNING

In 1935, Van der Corput [13] developed a sampling method in one dimension that
minimized discrepancy. This method was expanded on by Roth [14] in 1954 to incorporate
two dimensional data sets, and was then expanded upon again by Hammersley [15] in
1960 for n dimensions. Wang et. al [16] showed that the Hammersley distribution could
be used to reduce the number of sampled points required to achieve a level of acceptable
accuracy quadratically. These expansions managed to keep discrepancy low, but were
very strict and systematic in their implementations. Halton and Zaremba [17] developed
an easier to implement method, but was restricted in that the number of points sampled
had to be a power of two. These methods were tested to show results were consistent
and followed a normal distribution [18]. Another developed sampling strategy involved
both systematic and random components. As shown by Cochran [19], this stratified
random method was discussed. It breaks down the entire sampling area into different
“windows”, and then samples an equal number of points randomly within these windows.
This causes an increase in surface coverage of the sampled points, which might be missed in a systematic approach, while minimizing the inconsistencies of results due to random sampling.

In 1993, Woo and Liang [20] were one of the first groups to determine the efficacy of using a low discrepancy sample set in PMP. They applied the Hammersley distribution[15], shown in Figure 2-1 standard test surfaces with known errors. Then they compared the results against standard uniform sampling methods. They found that the Hammersley distribution could achieve the same accuracy as uniform sampling methods. The reduction in sample size was up to the square root of the original sample size.

In 1997, Lee et al. [21] did further testing focused on the Hammersley distribution. Instead of utilizing flat surfaces with known errors, they introduced measured features as well to
test the effects of added dimensionality. Their tests used random and uniform sampling methods on a hemisphere and cone. They combined Hammersley and other methods to determine if there was a benefit of doing so. The results showed that the Hammersley distribution was still able to reduce the number of sample points. There were also improvements when Hammersley was combined with a stratified technique that separated.

In 1995, Woo et al. [22] did further testing on different low discrepancy sequences, comparing them to random and structured sample sets. They used both Hammersley and the newer Halton-Zaremba sequence. They found that there was significant improvement when using the low discrepancy methods in the majority of cases. The Halton-Zaremba sequence was also found to be limited to sample sizes that were powers of two.

In 1997, Chan, King and Stout [23] showed the effects of using different sampling strategies on the same piece, comparing results between each one. They focused on the sampling of circular features, and utilized both systematic techniques and random techniques. Their data showed that the sampling technique chosen would affect the results of sampling, but this effect was also dependent on the measurement machine used, and the software used as well.

In 1998, Edgeworth and Wilhelm [24] began looking at iterative processes for determining the location of sample points. Instead of a static selection of points, they proposed a method for fitting to an original set of data points, then selecting a new set of data points based on this new fit. This model allowed for easy identification of manufacturing errors.
Instead of the usual static models, their proposed method looked at the surface normals of the part to find any continuous errors.

In 2003, Badar et al. [25] proposed the use of search algorithms to locate points on surface geometry. This involved a completely different approach than the standard static models in use. Their proposed method involved using various search algorithms in conjunction with a CNC controlled coordinate measuring machine. By choosing a set of starting sample points, a search would look for either local minima or maxima. This provided a more accurate view of the form error, and could better show errors in the piece. Where a static model could miss irregularities not contained in the predefined fit area, these new methods could search around their starting points to determine where errors were.

In 2006, Collins et al. [26] introduced some new sampling methods that were more focused on 3D models. These new methods attempted to consider the type of machining used when choosing points. As they were focusing on flat and revolved surfaces, they were looking for methods that could more accurately represent the origin of a part. They used the Spiral method, based on the Archimedes spiral, and the HamSpi method, which is the Hammersley and the Spiral method combined. The time-efficiency of the methods was comparable to the Hammersley distribution, but they focused on the origin of the part. This was important due to increased deviations in parts made with the techniques analyzed. The need for these methods highlights the weaknesses of the Hammersley and other static methods.
In 2016, Tao et. al [27] looked at minimizing the amount of sample points required to accurately represent splines. Taking a high density scan, they utilized bi-Akima spline interpolation to determine the connections between points. This allows for a smaller sample set to accurately represent the curve, showing that data set reduction is possible without a reduction in accuracy.

**FIGURE 2-2**  SPIRAL METHOD FOR POINT SELECTION [26]
In 2017, Lin et al. [28] looked at a least square fitting algorithm for circular cams. They compared the conventional method of breaking the cam into individual segments and comparing each segment to its corresponding parametric function, and their proposed method which looked at the cam as a whole using a least squares analysis. This allowed for the profile to be closed and continuous, and provided higher accuracy evaluations of the profile.
In 2017, Wang et. al [29] examined adaptive sampling techniques, and utilized simulated sampling to determine the efficiency of their methods. They managed to increase the efficiency of sampling of structured surfaces while reducing the time it took to sample the surface.

The uncertainty of coordinate metrology results drastically changes by any change in sampling data sets. It is shown in [20], [21] how any change in sampling strategy affects inspection results in flatness evaluation. From the sensitivity of this simple problem to changes in sampling strategy, it can be assumed that for more complex geometries the effect is even more significant. This sensitivity analysis has been under investigation by many experiments in [10] when [10] showed the effects of different sampling strategies on the inspection results.

There has been a lot of work done in the field of sampling strategy and point measurement planning. A lot of this work has occurred in trying to find the optimal static strategy to encompass most cases. However, there have been attempts to change from a static model, either to a quasi-static model or to a completely adaptive model.

2.3. SUBSTITUTE GEOMETRY EVALUATION

In 1968, Williamson[32] least squares fitting was used extensively to fit lines to data sets. For a line, this involves minimizing the vertical distance of each point to the line. This was the standard method for many years as it allowed a y value to be estimated given a set of x values. Williamson showed that the method was fast and accurate. Another important characteristic was its constant convergence, whereas other methods at the time, such as
least squares cubic, would sometimes fail to converge. So, least squares fitting was shown to provide consistent results.

van Huffel [33] conducted a review in errors in variables modeling techniques. While the book focused on the mathematical fitting strategies, there were also papers on the fitting of idealized geometry to point clouds for coordinate metrology. Zwick [34] did a comparison between total least squares, linear least squares, and minimax fitting. He showed that linear least squares and minimax fitting had efficient solutions in two-dimensional space, but had inefficient solutions in three-dimensional space. Total least squares, however, had efficient solutions in both spaces. This meant that it was more applicable for a wider variety of problems, especially when computing power was limited. Zwick did note that linear least squares and minimax fitting were gaining interest again due to the properties of the techniques, linear least squares being less aggressive and minimax fitting being very aggressive. This led to different results after fitting that could affect how a part is treated.

Nielsen [35] introduced a refinement to the least squares estimation using Lagrange multipliers. During this time, the least squares problem was still an optimization problem, and Nielsen wanted to determine a way to solve both linear and nonlinear constrained problems. The Lagrange optimization algorithm he proposed allowed for these solutions without affecting the results of the estimation.

Nassef and ElMaraghy [36] conducted a comparison between total least squares and minimax fitting with a focus on point selection for different feature geometries. Their
main concern was the effect of the sampling strategy on the error found by the fitting algorithm. They initially discussed the differences of TLS and minimax fitting, and go over some past criticisms. The major criticism of TLS fitting was its apparent overestimation of the geometric deviations and so, minimax fitting gained popularity because of its more aggressive approach. However, after further research they determined TLS fitting was reporting deviation values closer to the true values than minimax fitting. Nassef and ElMaraghy found that the geometric features of the part should determine the sampling method used.

Krystek and Anton [37] proposed a TLS method involving weighted points in two dimensions. They tested their method against known datasets and determined that it functioned similarly to other TLS methods. The main advantage over other implementations was the ability to perform regression on all data sets. Before this, TLS suffered when the result was a vertical line as there were discontinuities in the equation. This made TLS fitting more robust while maintaining the same quality of results found in other studies.

Malengo and Pennecchi [38] extended the weighted total least squares algorithm to curves as well as straight lines. They also wanted to incorporate straight-line fitting so that one package would be able to solve for both curves and straight lines. The algorithm they developed was good at dealing with both uncertain and correlated variables while minimizing uncertainties in the measurements. They verified their results against a Monte Carlo method and determined their algorithm maintained much lower uncertainty values
while the Monte Carlo method had increased uncertainty as the models’ nonlinearity increased.

Shakarji and Srinivasan [39] looked at tolerance standards involving TLS fitting of lines, planes and parallel planes. Their goal was to provide algorithms for each of the geometries that were simple and easy to implement, and did not need optimization algorithms to solve. They presented and proved these algorithms, and verified that the algorithms gave the same or better results than the traditional method of solving TLS fits. Shakarji and Srinivasan showed that the iterative method was no longer necessary and instead, by solving a few equations, you could find the TLS fit.

Shakarji et al. [40] produced a review of TLS fitting and its role in coordinate metrology. They discuss why TLS fitting has remained an attractive option over the years, the ease of use of the algorithm, and its applications for nonlinear problems. One of their main points is that ISO standards are beginning to use TLS fits, which means they will need to be used to check for standards. Near the end of the paper, they look at the uses of weighted total least squares in fitting. They state that by weighting points in low-density areas higher than those in high-density areas, the discrepancy of data sets will be removed as a source of error.

Considering fitting to a tolerance zone instead of fitting to an ideal geometry was introduced in [1], [41]. This consideration increases the nonlinearity of the optimization problem during the SGE process, however the results of this fitting can be much more practical and efficient to avoid rejecting an acceptable part. Complexity resulting from
nonlinearity of the objective function in the fitting process can cause trapping in local minima which produces false results for inspection. [42] Presents an iterative approach for SGE that avoids this issue by re-energizing the current solution to check for other minima in the area.

The use and development of TLS fitting over the past 50 years is well documented and it has been shown to be useful in many cases in two and three dimensions for several different objectives. With the advent of weighted total least squares, it is possible to modify the TLS results without modifying the original data set. This can be applied to coordinate metrology, as well to change the substitute geometry to optimize for different cases.

2.4. DEVIATION ZONE ESTIMATION

Skin models were introduced in the field of tolerance analysis [43]. In order to conform to ISO standards, parts must be within certain tolerances. To verify if a part conforms, the differences between the substitute geometry and the measured data are analyzed. Therefore, a detailed model of the geometric deviations is developed. This model is a non-ideal representation of the geometric deviations and is called a skin model.

Jamiolahmadi and Barari [44] utilized a finite difference method approach to develop the deviation zone. They tested their method on several different data sets where known errors occurred and evaluated the efficiency and accuracy of their method. The number of points needed to accurately define the geometric deviations of a surface was minimized using their developed method without introducing extra uncertainty.
Barari et al. [4] modeled the geometric deviations of different manufactured surfaces using NURBs surfaces. They achieved this by partitioning a machine’s workspace into sections with quasistatic errors, so that a set of linear transformations that represent the error are obtained. When these linear transformations are applied to the ideal geometry, an estimation of the geometric deviations can be obtained.

Inspection results typically are presented by evaluating a zone for geometric deviations on the measured part. Detailed deviation zone also referred to as a skin model is the most accurate way to represent this result. The skin model presents a continuous function for geometric deviations of the entire surface including the geometric deviation of a limited number of points that are sampled and the geometric deviation of infinite number of points that were never sampled. Having such a comprehensive model allows precise planning for any downstream activities including finishing operations, functionality
analysis, assembly planning, manufacturing process control, and manufacturing error compensation which is the objective of this thesis. Estimation of the detailed deviation zone for inspected surfaces is performed by using a Delaunay triangulation approach in [45] which is based on a bilinear interpolation of geometric deviations corresponding to sampled points. Some of the other works are using a finite difference approach to model the detailed deviation zone including [44], [46]–[48]. A more accurate method to develop the detailed deviation zone was recently presented in [49]. In this research, a finite element method is utilized to develop the skin model of the part. In this thesis, in order to develop our desired IIS this latest approach is adopted.

2.5. COST OF REPAIR – REASONING FOR INTEGRATIVE

The two most prevalent technologies that are currently used in die and mould repair are tungsten inert gas (TIG) welding and gas metal arc welding (GMAW). Both processes have been used for decades, but have a couple of flaws. First, when the part defects are relatively small, the bead thickness of these techniques will be too large to repair just the defect. This can require extra time machining out the defect so that the resulting hole is large enough for the technique to work. The other major issue is that these processes are very heat-intensive. To make sure the weld material fuses to the base material, constant intense heat is needed over a large area. This intense heat can cause the base material, which is usually a high alloyed tool steel, to be subjected to heat cracking or to form alloys with the weld material that have undesirable properties [50]. Other options have been investigated because of these undesirable effects.
The main set of processes are directed energy deposition (DED) methods. These processes involved taking some weld material, introducing it into the defect, and applying a large amount of energy to a very small area to cause the new material to bond to the base material. This solves both issues presented by the traditional methods because the small melt area allows the weld material to be inserted into smaller cracks, and the small heated zone allows heat to dissipate quickly into the base material without it becoming overly hot.

Powder bed fusion has also been considered a useful method. However, it is known more for part creation than part repair and has only been considered useful a handful of times. Another option would be standard material extrusion, though currently there are no direct extrusion devices capable of extruding the material required to repair dies and molds. Most extrusion devices are focused on the extrusion of metals with melting temperatures lower than 300C.

2.6. DIRECTED ENERGY DEPOSITION

This section will look at techniques involving Directed Energy Deposition methods. These are methods that deploy material directly to the area that needs it, and then cures or sets the material in-situ.

2.6.1. LASER METAL DEPOSITION (LMD)

LMD is one of the most researched technologies when it comes to DED processes. This is due primarily to lasers being readily available in the late 90’s when deficiencies in TIG and
GMAW were realized. These deficiencies lead to further research in LMD technologies and is currently one of the standard methods of repair.

Vedani et al. [51] looked into the difficulties of LMD when applied to both nitride and chrome plated 1.2738 steel samples. They found that if a surface is treated in any way that changes the surface chemistry dramatically it could cause many errors after the repair. In the case of chrome plating, the weld metal was over-alloyed by the dissolution of the chrome-plated layers causing hot cracking in the interface zone. The chrome layer also experienced cracking outside of the repair area, which was likely caused by welding stress. It was determined that repair of chrome-plated surfaces through LMD was unfeasible. For nitrided surfaces, the main issue was the increased porosity caused by the release of nitrogen gas during the welding process. To counteract the porosity they used a laser to re-melt the area that was going to be repaired to release the gas before the welding occurred. This destroyed the nitrided layer, but this was determined to be a small cost to pay when up against replacing the entire mold.

Pinkerton et al., [50] looked at some of the downsides of LMD and also looked at the effects of slot geometry as it related to reparability. They stated that porosity was one of the major issues affecting LMD as it is a powder-based method. The higher the porosity, the lower the strength of a repair, and thus reduced life for the mould. They also looked at the interface zone in detail. As can be seen in Figure 2-5, the interface zone, which is where the weld material and base material mix, has very little porosity and is fairly uniform. This shows off a good feature of LMD, its “self-quenching” property. This property occurs because the melt zone, despite having extremely high temperatures, has
a very small area and so the heat can be conducted away easily leading the part to cool rapidly. They also found that the repair was hardest in the interface zone. Finally, they looked at the effects of slot geometry on LMD. They found that for optimal results a slot would need to have sloped walls so the flow of powder is not blocked and to allow direct laser irradiation. However, the lowest vertex in a V-slot also provides issues as porosity is increased in the area, likely due to bead size of the powder.

Schmidt et al. [52] also looked at slot geometry, but instead focused on stainless steel and titanium alloys. They found that while slot geometry was important, the only thing that had to be ensured was that the powder jet could access the slot. Although better results were achieved in situations where the laser path and the wall of the slot were not parallel, acceptable results were achieved in those situations where the powder jet had ample access. They also determined that titanium alloys did not require a trailing argon nozzle if low heat was used.
Pleterski et al. [53] looked at how repairs needed to be prepped to ensure long part life after repair. They concluded that better repairs were achieved with larger material removal to prep the part. They also found that the cleaner the surface, the more likely a repair was going to be successful. This was likely due to removing contaminants from the surface that could modify the alloy chemistry, and removing oxidation layers that could prevent proper bonding.

Borrego et al., [54] showed that LMD allows for repair of dies and molds with relatively small changes of material composition in the repaired zone when the same material was used as filler. In their case, they used H13 and P20 steels. They found that after repair, there were still defects that affected the nominal stress in the piece. This meant that the fatigue resistance was diminished in the repaired piece when compared to a new piece. However, these stresses were close to parallel of the loading axis, and the effect was minimized because of this. Figure 2-6 shows the hardness levels in a repaired piece when subjected to multiple laser powers. (a) shows the hardness generated when the laser power was 104.4 W and with a diameter 0.5 mm, (b) at power of 106.6 W and a diameter of 0.5mm, (c) at a power of 111.6 W and diameter of 0.6 mm and (d) a power of 113.4 W at a diameter of 0.6 mm. The increase in hardness at the interface zone is likely due to microstructure changes in the zone.

Guijun et al., [55] looked at the software and process control side of LMD. They implemented a closed loop controller to modulate laser power, which was controlled by an IR sensor. The IR sensor constantly monitored the weld pool to maintain a steady temperature. They were attempting to repair the knife-edge on a turbine blade. They
found that without proper control of the weld pool, excess material was deposited at the edges of the blade and severely impacted their dimensional accuracy. In any location where deceleration was required, the laser was held in position for a longer time, causing the temperature of the weld pool to increase. This, combined with the excess material deposition, caused cracking in the final piece. When closed loop control was used, the laser power was decreased during the deceleration of the table. This allowed for consistent material deposition along the length of the knife-edge, and allowed a constant temperature to be used throughout each layer. This resulted in an increase in dimensional accuracy, and eliminated hot cracking. They concluded that the weld pool temperature was a very important aspect to control, if hot cracking was to be eliminated in LMD.

Leunda et al. [56] looked at LMD as a surface coating method. They found that preheating the substrate prevented hot cracking when performing LMD, as there was reduced thermal shock. However, it was important to keep the temperature low as high temperatures could induce cracking as residual tensile stresses in the substrate were released. They also found that the coating area could be subjected to a post-deposition heat-treatment, if the temperature was controlled to not allow the substrate hardness to be affected.
Rombouts et al. [57] looked at the surface finish of repaired LMD parts. They found that the surface quality of LMD repaired parts, like that of fused deposition modelling parts, suffered from the “staircase effect”. While previously these surfaces went through machining to have these effects reduced, they proposed a new re-melting process. When testing, the re-melting process provided substantial improvements in surface roughness and surface uniformity, if a high enough laser power was utilized to penetrate the surface of the weld area.

Nie et al. [58] looked at the effects of wire-fed LMD systems against powder-jet systems. They found that wire-fed systems had far reduced porosity because the material was solid as it was being melted, as opposed to the bead nature of powder fed systems. However, wire-fed systems presented their own challenges because proper wire feed speeds and
temperatures had to be maintained. They found that if the weld material was too cold it would not fuse to the base metal, if the weld material was too hot it would liquefy before coming into contact with the base metal and be affected by gravity and atmospheric conditions, which would lead to asymmetric weld paths. Another downside is wire-fed systems can only use ductile metals to deposit, only powder-fed can handle more brittle materials thus far.

2.6.2. Electrospark Deposition

Electrospark deposition is fairly new and untouched by research. In ED, a very high current is pulsed very quickly through an electrode and into the workpiece. This generates extremely high temperatures and causes part of the electrode to be deposited onto the workpiece. The extremely high temperature, but small work time, causes the part to self-quench on every pass. So, very little of the heat is passed to the base metal. This is beneficial because it means the base metal has little to no changes.[59]

Tusek et al. [60] examined some of these properties. They found you must use an inert gas for this process. For both the electrode and the melted surface to ensure proper bonding, stellite 6 provided the highest quality deposits with regard to porosity, with the surface roughness of the deposits ranging between 0.8 and 5.5 µm. They concluded by saying many portions of the process had not yet been examined scientifically.

2.6.3. Micro Plasma Transferred Arc

Plasma transferred arc welding is a relatively new technique for welding. It involves a non-consumable electrode that is used to create a plasma arc between itself and the work
piece. The plasma arc creates a melt pool that a filler material is introduced to. A couple of benefits of this technique is it requires a very low amount of power to work, and it has a lower cost to buy initially. It also has the ability to work with wire-fed systems and powder-fed systems.[61]

Jhavar et al. [61] first proposed the micro plasma transferred arc method for die and mould remanufacturing. They first determined the deposition parameters needed for a successful single bead and then they progressed to multi bead tests. Mechanical property tests showed that the results were equivalent to results derived from LMD and ED. It also had low heat transference to the base metal. They determined that this method would require further study to ensure its effectiveness as a repair process.

2.6.4. Powder Bed Fusion

Powder bed fusion is a technique mainly used to produce new parts through selective laser sintering (SLS). However, it has been used successfully a handful of times to repair existing parts but it is not usually considered for repairs. This is because of the complex set up required which restricts SLS from completing repairs on complex geometries. However, one team has managed to use it to build onto existing cut down parts.

Andersson et al.[62] looked at using SLS to repair burner nozzles. They examined the process from beginning to end, including personal user safety and end result compliance. To achieve their results, they cut the burner nozzle off of the piece and then placed the remaining part into the SLS machine. Through an optical system, they lined up the CAD model with the remaining piece, and then printed a new nozzle through with SLS. When
subjected to fatigue tests, it was found that the part failed in the base metal, far below the interface zone, and thus was unlikely to have been an effect of the repair. The parts were also found to be serviceable, being near equivalent to a new piece in dimensional accuracy, heat transfer characteristics, and lifespan.

2.7. POSSIBLE PATHS GOING FORWARD

With these technologies, there are many paths forward. With LMD, the major topic to be examined is closed loop control, and this applies to all the techniques listed above. Barring a fully closed loop manufacturing system that integrates additive, subtractive, and verification technology, adding closed loop control to pre-existing additive systems looks to be a promising way of increasing the quality.

There are also paths forward in the newer repair technologies such as ED and µPTA. Both technologies need to be put through the rigorous testing as LMD was put through before it was accepted by the industry. µPTA seems to be the lowest cost option.

An interesting topic would be adapting SLS to part repair. Due to its current very rigid nature, requiring a vat of metal powder and a wiper, it would be difficult to achieve repairs on complex geometry. However, the work by Andersson et al. has shown that when used correctly, the results can often be better than other repair methods.
3. METHODOLOGY

In this section, the proposed integrated inspection process will be developed. The adapted method used for sampling will be explained. Then, the theory behind the fitting algorithm will be developed. The modifications made to the skin modelling code that was adapted for use in this project will then be discussed. Finally, the method of integrating each part into one cohesive system will be explained.

3.1. ADAPTIVE SAMPLING TOWARDS PMP

The adaptive sampling method being used to trim down the number of points in each data set, while still maintaining accuracy, was adapted from a paper by Lalehpour et al [12]. In this method, points were selected using an algorithm called Neighbourhood Search for Representative (NSR). In this algorithm, a stratified random sampling strategy was masked over the full data set. This entailed separating out the full range of the data set into separate “windows” and then choosing a point at random within each window. The points selected from this mask became sites. These sites then had their neighbourhood, an area around the site with a radius related to the full length of the data set, examined for points. The mean of these points became the “representative” point of that area.

3.2. WEIGHTED TOTAL LEAST SQUARES FITTING TOWARD SGE

Total least squares fitting takes a plane and fits it to a set of points to minimize the square of the orthogonal distance from each point to the plane. To do this, the plane and point
set will be defined. The plane will be represented by \( \Pi(\alpha, \beta, \gamma, r) \) where \( \alpha, \beta \) and \( \gamma \) are the principal angles defining the plane, and \( r \) is the orthogonal distance from the plane to the origin. The point set will be \( P(x, y, z) \), where \( x, y \) and \( z \) are the coordinates of each point.

Once these are defined, the objective function to minimize is as follows:

\[
X^2(\Pi,P) = \sum_{i=1}^{n} d_{\perp}^2
\]  

(3-1)

where \( d_{\perp} \) is the orthogonal distance of a point to a plane, as shown in equation 3-2:

\[
d_{\perp} = x_i \cos \alpha + y_i \cos \beta + z_i \cos \gamma - r
\]  

(3-2)

where \( x_i, y_i, \) and \( z_i \) are the coordinates of a single point. Substituting this equation into Equation 3-1, the minimization problem becomes Equation 3-3:

\[
X^2(\Pi,P) = \sum_{i=1}^{n} (x_i \cos \alpha + y_i \cos \beta + z_i \cos \gamma - r)^2
\]  

(3-3)

As this is a minimization problem, the first derivative of the equation is taken and set to zero. This is used to determine the point where the function is smallest. The second derivative of the equation is also found in order to determine if the function has global minima or maxima.

\[
\frac{\partial X^2}{\partial r} = \frac{\partial \sum_{i=1}^{n} (x_i \cos \alpha + y_i \cos \beta + z_i \cos \gamma - r)^2}{\partial r} = 0
\]  

(3-4)

\[
\frac{\partial X^2}{\partial r} = -2 \sum_{i=1}^{n} x_i \cos \alpha + y_i \cos \beta + z_i \cos \gamma - r = 0
\]  

(3-5)
\[
\frac{\partial^2 X^2}{\partial r^2} = 2 \tag{3-6}
\]

The second derivative is a positive constant, so the function is always convex. This guarantees that there is a minimum value. Now, \( r \) is isolated so that it can be replaced in the minimization function:

\[
0 = \cos \alpha \sum_{i=1}^{n} x_i + \cos \beta \sum_{i=1}^{n} y_i + \cos \gamma \sum_{i=1}^{n} z_i - \sum_{i=1}^{n} r \tag{3-7}
\]

\[
0 = \cos \alpha \sum_{i=1}^{n} x_i + \cos \beta \sum_{i=1}^{n} y_i + \cos \gamma \sum_{i=1}^{n} z_i - nr \tag{3-8}
\]

\[
r = \cos \alpha \frac{\sum_{i=1}^{n} x_i}{n} + \cos \beta \frac{\sum_{i=1}^{n} y_i}{n} + \cos \gamma \frac{\sum_{i=1}^{n} z_i}{n} \tag{3-9}
\]

\[
r = \cos \alpha \bar{x} + \cos \beta \bar{y} + \cos \gamma \bar{z} \tag{3-10}
\]

where \( \bar{x}, \bar{y}, \) and \( \bar{z} \) are values the represent the average \( x, y, \) and \( z \) values for the whole data set. At this point, a simple partial differential equation cannot be done to determine the minimum values of the three principal angles of the plane. Instead, the minimization problem is redefined it is as follows:

\[
F(\pi) = \sum_{i=1}^{n} \left( \pi \cdot (p_i - \bar{P}) \right)^2 = \|Mp\|^2 \tag{3-11}
\]
where \( \pi \) is the plane, \( p_i \) is a vector containing the coordinates of the current point, and \( \bar{P} \) represents the centroid. To determine the minimum of this function, the Lagrange multiplier method is used. For this, \( F(\pi) \) subject to \( G(\pi) = 0 \), where \( G(\pi) = |\pi|^2 - 1 \) is set, which becomes Equation 3-12:

\[
\nabla F(\pi) = \lambda \nabla G(\pi) \quad \text{(3-12)}
\]

where \( \nabla F(\pi) \) is the gradient of \( F(\pi) \) and \( \lambda \) is a stationary point that minimizes the function. This breaks down into Equations 3-13 to 3-16:

\[
\nabla F = \begin{bmatrix}
\frac{\partial F}{\partial a} \\
\frac{\partial F}{\partial b} \\
\frac{\partial F}{\partial c}
\end{bmatrix}, \nabla G = 2\pi
\quad \text{(3-13)}
\]

\[
\nabla F = 2 \sum_{i=1}^{n} \pi \cdot (p_i - \bar{P})(x_i - \bar{x}) = 2M^TM\pi
\quad \text{(3-14)}
\]

\[
M^TM = \begin{bmatrix}
\sum_{i=1}^{n} (x_i - \bar{x})^2 & \sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y}) & \sum_{i=1}^{n} (x_i - \bar{x})(z_i - \bar{z}) \\
\sum_{i=1}^{n} (y_i - \bar{y})(x_i - \bar{x}) & \sum_{i=1}^{n} (y_i - \bar{y})^2 & \sum_{i=1}^{n} (y_i - \bar{y})(z_i - \bar{z}) \\
\sum_{i=1}^{n} (z_i - \bar{z})(x_i - \bar{x}) & \sum_{i=1}^{n} (z_i - \bar{z})(y_i - \bar{y}) & \sum_{i=1}^{n} (z_i - \bar{z})^2
\end{bmatrix}
\quad \text{(3-15)}
\]
Equation 3-16 shows that the vector that minimizes the function is the smallest eigenvector of $M^T M$. Weighted total least squares follows this pattern almost exactly, with a few modifications. The minimization problem is defined again, however, a weight factor is added.

$$X^2(\Pi, P) = \sum_{i=1}^{n} w_i (x_i \cos \alpha + y_i \cos \beta + z_i \cos \gamma - r)^2$$  \hspace{1cm} (3-17)

Here, $w_i$ is a set of weights, each associated with a point, where each weight is a positive value. The partial derivative is then taken with respect to $r$ in order to determine whether the centroid is still a required point on the TLS plane.

$$\frac{\partial X^2}{\partial r} = \frac{\partial}{\partial r} \sum_{i=1}^{n} w_i (x_i \cos \alpha + y_i \cos \beta + z_i \cos \gamma - r)^2 = 0$$ \hspace{1cm} (3-18)

$$\frac{\partial X^2}{\partial r} = -2 \sum_{i=1}^{n} w_i (x_i \cos \alpha + y_i \cos \beta + z_i \cos \gamma - r) = 0$$ \hspace{1cm} (3-19)

$$0 = \cos \alpha \sum_{i=1}^{n} w_i x_i + \cos \beta \sum_{i=1}^{n} w_i y_i + \cos \gamma \sum_{i=1}^{n} w_i z_i - \sum_{i=1}^{n} w_i r$$ \hspace{1cm} (3-20)
\[ r = \cos \alpha \cdot \frac{\sum_{i=1}^{n} w_i x_i}{\sum_{i=1}^{n} w_i} + \cos \beta \cdot \frac{\sum_{i=1}^{n} w_i y_i}{\sum_{i=1}^{n} w_i} + \cos \gamma \cdot \frac{\sum_{i=1}^{n} w_i z_i}{\sum_{i=1}^{n} w_i} \] (3-21)

Equation 3-21 shows that the weighted centroid is a point on the fit plane. The rest of the solution follows the same path as the non-weighted version, however, the value of \( M^T M \) needs to be determined. Equation 3-22 also represents the minimization problem as redefined in Equation 3-11, but with added weights:

\[ F(\pi) = \sum_{i=1}^{n} w_i (\pi \cdot (p_i - \bar{P}))^2 = \|Mp\|^2 \] (3-22)

where \( p \) is the vector that defines the principal angles of the plane. \( M \) then becomes the Equation 3-23:

\[
M = \begin{bmatrix}
\sqrt{w_1} (x_1 - \bar{x}) & \sqrt{w_1} (y_1 - \bar{y}) & \sqrt{w_1} (z_1 - \bar{z}) \\
\sqrt{w_2} (x_2 - \bar{x}) & \sqrt{w_2} (y_2 - \bar{y}) & \sqrt{w_2} (z_2 - \bar{z}) \\
\vdots & \vdots & \vdots \\
\sqrt{w_n} (x_n - \bar{x}) & \sqrt{w_n} (y_n - \bar{y}) & \sqrt{w_n} (z_n - \bar{z})
\end{bmatrix}
\] (3-23)

If we use this to find the new \( M^T M \) we get, it becomes Equation 3-24:

\[
M^T M = \begin{bmatrix}
\sum_{i=1}^{n} w_i (x_i - \bar{x})^2 & \sum_{i=1}^{n} w_i (x_i - \bar{x})(y_i - \bar{y}) & \sum_{i=1}^{n} w_i (x_i - \bar{x})(z_i - \bar{z}) \\
\sum_{i=1}^{n} w_i (y_i - \bar{y})(x_i - \bar{x}) & \sum_{i=1}^{n} w_i (y_i - \bar{y})^2 & \sum_{i=1}^{n} w_i (y_i - \bar{y})(z_i - \bar{z}) \\
\sum_{i=1}^{n} w_i (z_i - \bar{z})(x_i - \bar{x}) & \sum_{i=1}^{n} w_i (z_i - \bar{z})(y_i - \bar{y}) & \sum_{i=1}^{n} w_i (z_i - \bar{z})^2
\end{bmatrix}
\] (3-24)
Solving for the smallest eigenvector of $M^TM$ provides the orientation of the fit plane. However, this problem can be solved without having to calculate $M^TM$. The singular value decomposition of any mxn matrix $A$ is Equation 3-25:

$$A = USV$$  \hspace{1cm} (3-25)

therefore, the Equation 3-26 can be found:

$$A^TA = V(S^TS)V^T$$ \hspace{1cm} (3-26)

where $V$ is the right singular vectors of $A$, $U$ is the left singular vectors of $A$ and $S$ is the singular values of $A$. The eigendecomposition of a matrix $K$ that is an nxn square matrix with $n$ linearly independent variables is Equation 3-27:

$$K = Q\Lambda Q^{-1}$$ \hspace{1cm} (3-27)

where $Q$ contains the eigenvectors of $K$ and $\Lambda$ contains the corresponding eigenvalues of $K$. If $K = A^TA$ is set, when $K$ is a symmetric positive definite matrix, equations 3-26 and 3-27 are equal, and $V$ is equal to $Q$. Thus, because $V$ contains the eigenvectors of $A$, we can simply choose the vector associated with the smallest singular value, and this will be the vector that determines the principal angles for the fit plane. By removing the need for a transpose and multiplication, the numerical accuracy of the method is increased.
3.2.1. **Volume Calculation**

The volume contained between a triangle making up the skin model and the current fit plane is used to calculate the weight of each point in the skin model. As the skin model can contain a large number of triangles, it was necessary to determine a quick way of calculating the volume. To do this, the full volume was first separated into two distinct volumes with the separation plane being a plane parallel to the fit plane, translated along the z axis to be at the level of the lowest vertex in the triangle. This results in Figure 3-1.

**Figure 3-1 Example case for Volume Calculation**
The top volume becomes an irregular pyramid and the bottom volume becomes a triangular prism. Therefore the total volume, $V_T$, can be calculated using Equation 3-28:

$$V_T = V_{pr} + V_{py}$$  \hspace{1cm} (3-28)

where $V_{pr}$ is the volume of the prism and $V_{py}$ is the volume of the pyramid. The volume of the prism is calculated using the projected area of the triangle to the fit plane, multiplied by the height of the prism, in this case the distance from the lowest vertex, $z_1$, to the fit plane, which in this case has a height of zero.

$$V_{pr} = A_T \cdot z_1$$  \hspace{1cm} (3-29)

where $A_T$ is the area of the projected triangle and is equivalent to Equation 3-30:

$$A_T = \frac{|B_2B_3| \cdot h}{2}$$  \hspace{1cm} (3-30)

where $|B_2B_3|$ is the length of the vector connecting $B_2$ and $B_3$, and $h$ is the height of the projected triangle. The volume of any pyramid is given by the area of its base, multiplied by its height and divided by three:

$$V_{py} = \frac{A_b \cdot h}{3}$$  \hspace{1cm} (3-31)

where $A_b$ is the base of the pyramid and $h$ is the height. The base of the pyramid in this case is made up of two projected lines, a plane parallel to the projected plane, and an edge of the original triangle. The two projected lines are parallel to one another and
perpendicular to the projected plane. In Figure 3-2, the base of the pyramid of the example in Figure 3-1 is shown.

![Figure 3-2 Base of Pyramid for Volume Calculation](image)

This arrangement causes the base of the pyramid to be a right trapezoid, and thus the area of the base can be calculated via Equation 3-32:

$$A_b = \frac{1}{2} a(h_1 + h_2) \quad (3-32)$$

$$A_b = \frac{1}{2} |\overrightarrow{B_2B_3}|((z_2 - z_1) + (z_3 - z_1)) \quad (3-33)$$

where each z value is the height of its corresponding point. As the triangular prism and the pyramid share the face containing the height of both the projected triangle and the pyramid, Equation 3-30 can be rearranged to get the height of the pyramid:

$$h = \frac{2A_T}{|\overrightarrow{B_2B_3}|} \quad (3-34)$$

making Equation 3-31 become Equation 3-35:
Subbing in Equation 3-33, the formula becomes Equation 3-36:

\[
V_{py} = \frac{A_b \cdot \frac{2A_T}{B_2B_3}}{3} \tag{3-35}
\]

Which when simplified, becomes Equation 3-37

\[
V_{py} = \frac{1}{2} \left| B_2B_3 \right| \cdot \left( \left( z_2 - z_1 \right) + \left( z_3 - z_1 \right) \right) \cdot \frac{2A_T}{B_2B_3} \tag{3-36}
\]

\[
V_{py} = \frac{\left( \left( z_2 - z_1 \right) + \left( z_3 - z_1 \right) \right) \cdot A_T}{3} \tag{3-37}
\]

Taking this equation and Equation 3-30 and putting them both into Equation 3-28, the total volume of the example can be found using Equations 3-38 to 3-40:

\[
V_T = A_T \cdot z_1 + \frac{\left( \left( z_2 - z_1 \right) + \left( z_3 - z_1 \right) \right) \cdot A_T}{3} \tag{3-38}
\]

\[
V_T = \frac{3A_T \cdot z_1}{3} + \frac{\left( \left( z_2 - z_1 \right) + \left( z_3 - z_1 \right) \right) \cdot A_T}{3} \tag{3-39}
\]

\[
V_T = A_T \left( \frac{z_1 + z_2 + z_3}{3} \right) \tag{3-40}
\]

This shows that the volume between any triangle and plane can be determined via the projected area of the triangle to the plane, and the height of the centroid of the triangle.

In the case where a triangle is intersected with the plane, this process still finds the total volume contained between the triangle and the plane. However, this volume is the
summation of the positive volume above the plane and the negative volume below the plane. This value is sufficient for determining the weights to be applied for weighted total least squares, however it does not account for the cost of repair. To do this, the individual volumes need to be determined. First, the intersection points of the line and the plane need to be determined. The side of the plane each point lies on is determined by checking the sign of the z coordinate as the fit plane is assumed to have a height of 0. Vectors are created from points with opposite signs, and the intersection points are found using parametric interpolation. The parameter \( t \) is found using the z coordinates of two points on either side of the plane, and the height of the plate itself:

\[
t = \frac{z_1 - z_0}{z_1 - z_2}
\]  

(3-41)

where \( z_0 \) is the height of the fit plane, and \( z_1 \) and \( z_2 \) are the heights of the two points that make up the vector. In this case, \( z_0 \) will always be 0, as the dataset is rigid body transformed so that the fit plane is the XY plane. With this parameter, the intersection point can be determined via the Equation 3-42:

\[
P_I = (1 - t) \ast P_1 + t \ast P_2
\]  

(3-42)

where \( P_1 \) and \( P_2 \) are points of the triangle. Using these intersection points and the vertices of the original triangle, three new triangles are formed. With these triangles, the appropriate positive and negative volumes can be quickly determined and stored. The method as written assumes the fit plane has a z height of zero. This works due to the rigid body transformation that takes place on the data set to transform it to the fit plane during
the skin modelling step. As this process is also repeated during each refitting of the weighted data set, a z-height of zero can be assumed for each iteration.

### 3.3. CALIBRATION OF SKIN MODELLING TOWARDS DZE

The skin model, using finite element methods, was developed by Barari et al. [44]. In this method, a subset of the given data set is chosen to be processed.

A plane is then fit to this data set using standard TLS methods. The rotation matrix associated with a transformation of the fit plane to base plane is found, and the data set is rigid body transformed to the new base plane. We define the axes of this coordinate system as u-v-d, where u and v represent the two major axes of the data set, and d represents the geometric deviations of each point from the fit plane. The u-v plane is then separated into a number of windows equal to the square root of the number of points in the reduced sample size. Once these points are located within their windows, a search is run to find the values with the largest deviations from the u-v plane. These points become the boundary conditions for a Delaunay triangulation. A grid with twice the number of windows is then overlaid, and extra points are added around the rectangular boundary of the fine grid. The Delaunay triangulation is then carried out, and any triangle large then a threshold defined by the function is further subdivided by making the centroid of the large triangle a new site for the Delaunay triangulation. This is repeated until all triangles are smaller than the threshold value. When this is complete, the FEM analysis is begun. The in depth proof of this method can be found in the paper by Barari et al. [44]. The main equation that governs how the skin model is as follows:
\[
\frac{\partial}{\partial u} \left( K_{uu} \frac{\partial S}{\partial u} \right) + \frac{\partial}{\partial v} \left( K_{vv} \frac{\partial S}{\partial v} \right) + E = 0 \quad (3-43)
\]

Where \( K_{uu} \) and \( K_{vv} \) are pre-set error propagation values. In the original paper, these values were both set to one for the sake of problem simplification. However, this can lead to an increased amount of error as the data sets are changed. To solve this, error modeling was conducted to determine how the function reacted to changing K values, and an optimization algorithm was used to find the best K values for each data set. To conduct the error modeling, tests were run on various data sets with the K values being set over a range encompassing both negative and positive values. There was a uniform step between each K value. The results for one data set are shown in Figure 3-3.

![Figure 3-3 Error Graph for Skin Modelling with Unfiltered Results](image)

The graph shows there is some degree of symmetry to the error, but when the signs of \( K_{uu} \) and \( K_{vv} \) are not the same, large error is introduced into the system. By removing the K values that cause excess error, the useful area of the graph is brought out.
In Figure 3-4, there are a few things that aid in choosing and directing an optimization algorithm. First, it can be seen that the error values are mirrored about the $K_{uu}$ and $K_{vv}$ axes. This means that the optimization algorithm can be told to ignore any non-positive values. It can also be seen that the areas where the $K$ values differ greatly from one another cause large amounts of error. This allows the algorithm to start away from those areas, causing the computation time to decrease. With this information, a particle swarm optimization algorithm was chosen. With a particle swarm optimization, the function to be optimized does not need to be continuous. This allows the optimization to take place without the need for derivation, which would be difficult to implement with the way the skin model is developed. The chosen algorithm is the firefly optimization algorithm [63]. This algorithm mimics the behaviour of fireflies in the wild, with each individual’s
brightness being determined by the value it outputs. In this case, smaller values induce higher brightness in an individual. This attracts other individuals more heavily to the better result. Each individual goes through an attraction phase, where they are moved around the solution area using Equation 3-44:

\[ q_i = q_i + \eta e^{-\Gamma g_{ij}} (q_j - q_i) + \alpha \epsilon_i \]  \hspace{1cm} (3-44)

where \( q_i \) represents the solution of the current candidate, \( \eta \) represents the attractiveness of another candidate, \( q_j \), \( g_{ij} \) represents the distance between the two candidates, \( \Gamma \) is a light absorption coefficient that dictates the amount of light “absorbed” by the medium, and \( \alpha \epsilon_i \) is a randomization parameter drawn from a Gaussian distribution. After a set number of iterations, the current values of each individual is checked, and the K values associated with the lowest error are used. This can be allowed to run for as many iterations as desired, but after experimentation, 10 iterations was determined to provide adequate results.

3.4. INTEGRATED PROCESS

With all the processes defined, combining them into one process can be done. The flowchart in Figure 3-5 details the process in order.
In the revised process, each of the traditional coordinate metrology steps is replaced with the corresponding modified version. First, the discrete 3D data points are retrieved by a
high density laser scanner. This provides several hundreds of thousands to millions of points, depending on the part in question. These data points are then passed to the adaptive sampling method discussed in 3.1. This reduces the number of sample points while maintaining the accuracy of the sample. With this smaller data set, the computation time for the remaining steps is drastically reduced. The reduced data set is then passed to a TLS fitting algorithm. This algorithm fits a baseline optimum fit plane. With this plane, the skin model can be calculated, creating a discretized triangulated surface. With this skin model and the TLS fit plane, the initial cost of repair can be calculated. This provides a baseline to compare the results of the added weighting process described in section 3.4.1. This process calculates a weight for each vertex of the skin model, which is itself a point from the reduced data set. These weights are then used to create a new fit plane using Weighted Total Least Squares. Using this new fit plane, a new set of costs for each segment of the skin model is calculated, and the modified total cost of repair is found. This process then repeats until a termination condition is met. This termination condition is discussed in section 3.4.2.

3.4.1. Weighting Process

The only piece of this process that has not yet been described is the weights that are used in weighted total least squares. In order to determine the weights, each of the parts needs to be used. Using the triangulated surface created by the skin model code and the fit plane from the weighted total least squares code, the volume encapsulated by each section of the surface can be calculated. In order to modify the fit plane, the weights of
the points need to be changed every iteration. By using the volume of the triangle as a factor in the weight calculation, this criteria is met. Ideally, by continually fitting and reweighting the data set, an optimal solution will arise. Four weighting methods were used in this study.

The first, and simplest of the methods, which will be called the Net Volume Method (NVM), used the net volume of each triangle and multiplied this value by a cost multiplier. This cost multiplier was associated with either the additive or subtractive repair method and was dependent on the position of the centroid of the triangle. This calculation for cost is shown in Equation 3-45.

\[
Cost = A_t \cdot z_c \cdot \text{Cost Factor} \quad (3-45)
\]

Where \(A_t\) is the projected area of the triangle to the plane, \(z_c\) is the height of the centroid of the triangle, and the cost factor is a positive scalar value representing the per volume cost of a method of repair. This cost value was then split equally between each node of the triangle.

The second method was similar to the first, however, instead of using the net volume of the triangle and applying a single cost factor, both cost factors were applied. This will be referred to as the Absolute Volume Method (AVM). Depending on the location of all three nodes, this method takes into account volumes on either side of the plane, and does not assume that any triangle must be repaired either by additive or subtractive methods, but never both. Using Equation 3-46 accounts for all scenarios.
\[ \text{Cost} = V_p \cdot \text{Sub. Cost Factor} + |V_n \cdot \text{Add. Cost Factor}| \]  

(3-46)

Where \( V_p \) is the positive volume of the triangle, and \( V_n \) is the negative volume of the triangle. This equation accounts for both styles of repair in each triangle, and in the case that one method is unnecessary as all points exist above or below the plane, the opposing volume drops to zero, removing it from the weight. When the cost is calculated, it is divided evenly among the vertices of the triangle.

The third method involved each vertex receiving a fraction of its associated volume. It will be referred to as Volume Dependent Weighting (VDW). This method takes into account the type of volume that is being worked on, and also the position of the associated vertexes. For example, if there were two points above the fit plane and one point below, then the points above the fit plane would share the weight associated with the positive volume, and the point below the plane would solely get the weight associated with the negative volume.

The final method utilized the centroid of each triangle as opposed to the vertices and will be referred to as the Centroid Method (CM). For this method, the triangulation was still carried out, however, instead of refitting the WTLS plane to the data set, it was refit to the centroids of the triangles forming the skin model. These were weighted using Equation 3-44.

For each weighting method, the data set used was the two offset planes used previously to validate the volume calculation code, shown in Figure 4-2. Two cases for this data set will be examined. In one case, the difference between the two costs of repair will be small,
with additive repair having a cost of 5 per volume of material, and subtractive having a cost of 2. The second case will have a large difference between the two values, with additive being increased to 50.

First, the small discrepancy test will be carried out. In the Figures Figure 3-6 through Figure 3-9, the total cost will be plotted against the iteration of the method. The methods will be presented in the order they were introduced.

![Net Volume Method - Small Discrepancy](image)

**FIGURE 3-6 Net Volume Method – Small Discrepancy Test Results**
**FIGURE 3-7** ABSOLUTE VOLUME METHOD – SMALL DISCREPANCY TEST RESULTS

**FIGURE 3-8** VOLUME DEPENDENT WEIGHTING – SMALL DISCREPANCY TEST
In all cases, oscillatory patterns are found. While the NVM and AVM oscillate around a fixed point, VDW and CM have changing patterns. VDW and CM both also have very high peaks, with CM being larger at its height by more than 400. In Table 3-1, values from these figures are compared.

**Table 3-1 Comparison of Small Discrepancy Tests**

<table>
<thead>
<tr>
<th>Weight Method</th>
<th>Lowest Value</th>
<th>Highest Value</th>
<th>Average Value</th>
<th>Peak to Peak Percent Variation</th>
</tr>
</thead>
<tbody>
<tr>
<td>NVM - Figure 3-6</td>
<td>45.94</td>
<td>67.50</td>
<td>56.30</td>
<td>47%</td>
</tr>
<tr>
<td>AVM - Figure 3-7</td>
<td>46.02</td>
<td>71.21</td>
<td>58.80</td>
<td>55%</td>
</tr>
<tr>
<td>VDW - Figure 3-8</td>
<td>46.26</td>
<td>120.52</td>
<td>80.06</td>
<td>161%</td>
</tr>
<tr>
<td>CM - Figure 3-9</td>
<td>59.05</td>
<td>25539</td>
<td>8175</td>
<td>43147%</td>
</tr>
</tbody>
</table>
The results show that NVM finds the lowest result and also has the lowest variation iteration to iteration. All methods except CM find a suitably small value for total cost. The next test will compare this methods when the cost factors have a large amount of discrepancy between them.

**Figure 3-10 Net Volume Method – Large Discrepancy**

*Net Volume Method - Large Discrepancy*

*Figure 3-10 Net Volume Method – Large Discrepancy Test Results*
Figure 3-11 Absolute Volume Method – Large Discrepancy Test Results

Figure 3-12 Volume Dependent Weighting – Large Discrepancy Test Results
The results after the large discrepancy test are much different from the previous test. While CM, VDW and NVM exhibit similar behavior to the previous test, AVM shows marked improvement. Table 3-2 Comparison of Large Discrepancy Tests compares the results of these tests.

<table>
<thead>
<tr>
<th>Weight Method</th>
<th>Lowest Value</th>
<th>Highest Value</th>
<th>Average Value</th>
<th>Peak to Peak Percent Variation</th>
</tr>
</thead>
<tbody>
<tr>
<td>NVM - Figure 3-10</td>
<td>146.55</td>
<td>438.69</td>
<td>241.10</td>
<td>199%</td>
</tr>
<tr>
<td>AVM - Figure 3-11</td>
<td>145.14</td>
<td>438.69</td>
<td>160.06</td>
<td>202%</td>
</tr>
<tr>
<td>VDW - Figure 3-12</td>
<td>68.10</td>
<td>931.10</td>
<td>348.77</td>
<td>1267%</td>
</tr>
<tr>
<td>CM - Figure 3-13</td>
<td>265.10</td>
<td>313690</td>
<td>61438</td>
<td>118230%</td>
</tr>
</tbody>
</table>

AVM now has a lower value than NVM, however, the difference is still minimal. They also both never have a higher result than the initial TLS fitting. VDW and CM still have issues...
with this, and way overshoot the result. From these results, AVM has been chosen to be used for the process. While NVM has marginally lower results when there is small discrepancy between the different costs of repair, it does not take into account the actual cost of repair when calculating weight.

3.4.2. Bump Code for Oscillations

To deal with the constant oscillations, if a repeating pattern is detected in the results, a randomized bump is introduced into the orientation vector. This causes large fluctuations in the cost, but they are quickly compensated for. This causes the optimization to be bumped out of any local minima it may be trapped and allow it to possibly find the global minimum. The result of this method can be seen in Figure 3-14.

![Image: Figure 3-14 Optimization results for two offset plates]
The bump code was triggered seven times in Figure 3-14. When triggered, the total cost was immediately raised to extraordinary levels. This value then drops rapidly on the next iteration. While the patterns after each bump were similar, after the sixth bump, a difference in the response is seen. This indicates that the method found a different result and was bumped out of a local minima.

3.5. HEURISTIC OPTIMIZATION

Due to the inability of the WTLS method to converge on a solution, a heuristic optimization method was used to determine if the WTLS method was finding a result close to the minimum value. In order to validate this, after the WTLS method had finished, the best result was taken and used as the initial condition for the MATLAB function fminsearch. This function uses the Nelder-Mead method for optimization. Using the functions already developed, this method is given the orientation vectors of the best fit plane from WTLS as its initial guess. In Figure 3-15, the results of the two offset planes with the optimization in use is shown.
FIGURE 3-15 TWO OFFSET PLANES WITH FMINSEARCH RESULTS

The asterisk value is the best point found from WTLS, and the line below the WTLS results is the results given by fminsearch. Table 3-3 Results of optimization for two offset planes compares the best values of both methods.

<table>
<thead>
<tr>
<th>Piece</th>
<th>TLS</th>
<th>WTLS</th>
<th>Fminsearch</th>
<th>WTLS over TLS improvement</th>
<th>Fminsearch over WTLS Improvement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Two Offset Planes</td>
<td>73.52</td>
<td>56.29</td>
<td>55.89</td>
<td>23.44%</td>
<td>0.71%</td>
</tr>
</tbody>
</table>
The values show that WTLS, in this instance, found a result close to a minimum value. However, there was still improvement over the WTLS results. Therefore, the optimization algorithm will be used alongside the WTLS method to find the best fit plane.

4. IMPLEMENTATION AND RESULTS

This section will deal with the validation and results of the developed method. The first section will look at a test case used to verify that each portion of the method is working correctly. Then, the results of 5 different case studies will be examined.

4.1. VALIDATION

In order to test the developed methodology, a test platform was created and different test cases were used to determine the validity of the developed method. The sample sets used were a perfect plane and two equally sized planes offset from one another in space. With the plane data set, the method was being checked for any added error. During each subsequent fitting and reweighting, minimal changes in the orientation of the fit plane and the total cost of repair for the part should be noticed. These small fluctuations would be caused by errors induced by successive calculations and floating point error, and
should be so small as to not have a large effect on the fitting process. In Figure 4-1, the cost of repair is plotted against the number of iterations.

From this figure, it can be seen that the total cost of repair does not change even after 200 iterations, which is the expected result. This shows that the method is not adding in extra error during its operation in the most trivial case.

In order to validate with the two offset planes, seen in Figure 4-2, two different methods were undertaken to have a comparison.
FIGURE 4-2 TWO OFFSET PLANES

One method used a CAD program, and the other relied on trigonometry. The main thing to be examined was the total volume determined via each method, with a focus on determining the sum of both the positive and negative volumes. The results for these methods can be compared to determine if the implemented volume calculation method functions correctly. For all methods, the fit plane was first determined via TLS fitting, and the data set was rigid body transformed so that the normal distance to the fit plane could be easily determined. Then, some defining features of the plane were found, namely the distance and location of the extreme points of the plane, and the areas where the plane
intersected with the data set. With this data, the trigonometric solution could be found.

The data was used to create Figure 4-3.

FIGURE 4-3 TRIGONOMETRIC MODEL FOR VOLUME CALCULATION VALIDATION

In this figure, the value of $a$ is found through the fitting algorithm and the value of $c+d$ is set parametrically. In order to determine the volumes, the value of $c$ and $d$ need to be determined trigonometrically. As $ac$ and $bd$ form similar triangles, the value of $c$ can be found using the Equations 4-1 and 4-2.

\[
\frac{a + b}{a} = \frac{c + d}{c} \quad (4-1)
\]

\[
c = \frac{c + d}{\left(\frac{a + b}{a}\right)} \quad (4-2)
\]

With this value, the volume of both $ac$ and $bd$ can be determined using the width of the plane, $c+d$, which is set in the code. This provides a partial representation of what will be calculated by the code as the skin model not account for gaps in the data set. Therefore, it will bridge the gap, and calculate the volume of the space encompassed by the gap. To
determine the volume across the gap, the length of b is reused, and the length of the hypotenuse is found. To do this, the length of the gap is calculated using Equation 4-3.

\[ l = \sqrt{u^2 + (y - 2j)^2} \]  

(4-3)

Where l is the Euclidean distance of one edge of the gap to the other. Half of the length of the gap is then the hypotenuse of the triangles formed by b and the plane. Then, Pythagorean’s theorem can be used to determine the height of the triangle, and the area calculated.

\[ e = \sqrt{\left(\frac{l}{2}\right)^2 - b^2} \]  

(4-4)

Where e is half the length of the gap, parallel to the fit plane. With the area, the volume of the gap portion can be calculated with the width of the planes.

As an example, a planar width, y, of 10, length, j, of 2.5, and a z offset, u, of 2. Therefore, Equation 3-46, using values substituted from the fit plane can be used to find the value of c.

\[ c = \frac{2.5}{\left(\frac{0.2762 + 0.3461}{0.2762}\right)} = 1.1096 \]  

(4-5)

Then, the value of e needs to be determined to find the volume encompassed by the gap.

\[ l = \sqrt{2^2 + (10 - 2(2.5))^2} = 5.385 \]  

(4-6)
With $c$ determined, the total volume of either the volume above or below the plane will be the summation of those sections. The equation to calculate the volume of either section is detailed in Equation 4-8, with the total volume in this case being twice the volume determined.

$$V = \left( \frac{a \sqrt{c^2 - a^2}}{2} + \frac{b \sqrt{(j - c)^2 - b^2}}{2} + \frac{b \cdot e}{2} \right) \cdot y$$  \hspace{1cm} (4-8)

Where $y$ is the width of the plane. Using the example values above, the value of either the positive or negative volumes is found.

$$V = \left( \frac{0.2762 \cdot 1.1096}{2} + \frac{1.3904 \cdot 0.3461}{2} + \frac{0.3461 \cdot 2.67}{2} \right) \cdot 10 = 8.4353$$  \hspace{1cm} (4-9)

The data used to develop the planes trigonometrically is also used to find the solution in the CAD software. In Figure 4-4, the original data and the fit plane are represented in the CAD program. This allowed the test case to be validated using pre-existing and proven methods. The volume value shown in the figure is representative of the volume between the offset planes data set and the TLS fit plane. The white segments are those volumes above the fit plane, and the black segments are the volumes below the fit plane.
FIGURE 4-4 RESULTS FOR VOLUME CALCULATION VERIFICATION

The results of the method described above were 8.4363 mm^3. This is in line with both of the validation methods, barring any rounding or calculation error inherent in the equations or programs.

Another piece of the program that requires validation is the weighted total least squares fitting. As the weights allow the fit plane to be reoriented and translated in space, it is important to verify that the process works correctly and in an expected way. To test this, the impact of volumes on the fitting process will be removed, and weights will be determined by the side of the plane a point happens to rest on. By varying the weight each side gets, a change in the plane should be noticed. To begin with, equal weights were given to points on either side of the plane. In essence, this would mean that none of the points would be given any extra importance, and therefore the plane should not move
from its initial location. This should essentially provide the same results as traditional total least squares. In Figure 4-5, the data set and fit planes after twenty iterations are shown.

![Figure 4-5: Plane Orientation when Weights are Equal](image)

**Figure 4-5 Plane orientation when weights are equal**

The blue line emerging from the plane represents the normal vector of the plane. In this case, it can be seen that there is no movement of the fit plane, and it is behaving as expected. The weight value of all points under the plane is then increased exponentially
to determine the effect. Figures 4-6 and 4-9 show the plane shift when the weight is 4, 8, 16, and 32.

**Figure 4-6** Plane orientation for additive cost of 4, subtractive cost of 2

**Figure 4-7** Plane orientation for additive cost of 8, subtractive cost of 2
In these figures, as the weight of the points under that plane increases, there is a definite increasing translation of the plane. This shows that the weighted total least squares technique is working for basic cases, and shows expected behavior for this test.
4.2. RESULTS

Five different case studies will be examined. In each case, some form of damage has been applied to a planar surface. Each case will be allowed to run for 200 iterations, after which the total cost of repair determined from Total Least Square Fitting and Weighted Total Least square fitting will be examined. This number of iterations was used as it was greater than the maximum number of iterations that fminsearch took to reach its termination condition, and also generally allowed the code to “bump”, meaning stabilization of the results had occurred. The additive and subtractive cost of repair per unit volume will be 5 and 2 respectively. These costs are unitless, and are simply used to drive the algorithm. For real world application, a currency and unit of measurement would need to be defined to get accurate real world costs.

4.2.1. 3D PRINTED SURFACE

The first surface will be a flat 3D printed surface. In the configuration this part was printed in, the errors on the surface should be minimal. However, as 3D printing is a layer based process, there are divots in the surface where each layer connected. Also, there are extra pieces of plastic left over from the printing process. Currently, these defects would be covered using some form of post-processing, such as the acetone vapour bath technique used on ABS pieces. In a hybrid process, these errors would be removed after the additive process using a milling head. The part used is shown in Figure 4-10.
In the Figure 4-11 and Figure 4-12, the developed skin model and results of the method are shown. The data used to develop the skin model was found using a high density laser scanner. Figure 4-13 shows the results found using fminsearch.
FIGURE 4-11 SKIN MODEL FOR 3D PRINTED SURFACE

3D Printed Surface - Skin Model

FIGURE 4-12 REPAIR COST FOR 3D PRINTED SURFACE

3D Printed Surface - Repair Cost
In Figure 4-12, the oscillatory nature of the method can still be seen. However, because of the curving edges of the surface, it does not quickly find a stable point. The results are damped, and the cost begins to arrive at a value before destabilizing again. As stable oscillations were not achieved, the “bump” code was never triggered. Also, because the results were damped, the minimum value from WTLS was found very quickly.
4.2.2. **Gouged Wax Piece**

In Figure 4-14, a wax piece with several gouges is shown. This mimics extreme damage to a planar surface. This would be a typical use case for the developed process. The data from this model was collected using a high density laser scanner. Figure 4-15 shows the skin model created from this data.
**Figure 4-15** Gouged Wax – Skin Model

**Figure 4-16** Gouged Wax – Repair Cost
In Figure 4-15 the skin model and repair costs found using the developed method are seen. The skin model is mainly planar, unlike the 3D printed surface, and there are only cuts into the model. In Figure 4-16, there is a similar pattern as in Figure 4-12. Instead of immediately entering stable oscillation, the program instead acts like a damped system. In this case however, it is damped enough that the bump code triggers around iteration 165. After the bump, the repair cost immediately drops to below the level of the first iteration, and the minimum value from WTLS is found. It then begins oscillating again. Figure 4-17 shows the results from fminsearch.
4.2.3. **Metal Piece with Two Holes**

The third case is a traditionally manufactured aluminum piece. It is planar, with two through holes. On the surface, typical machining errors can be seen. The piece is shown in Figure 4-18.

![Figure 4-18 Metal Piece with Two Holes](image)

**Figure 4-18 Metal Piece with Two Holes**

Figures 4-19 and 4-20 show the skin model and optimization results for this piece. Of note, it can be seen that the skin model formed over the holes, however, the mesh has low density in these areas. While the data was collected using a high density
laser scanner, no data was collected in these areas. The extra data is due to how the skin model code generates vertices.
As is seen in Figure 4-20, this model behaves very similarly to the previous one. The oscillations are damped and stabilize slowly. When the bump code is triggered in this instance, the next iteration has a value very close to that gained from pure TLS fitting. However, despite similarities with the previous results, the minimum value for this piece is found very quickly and before the bump code is triggered.

4.2.4. NURBS SURFACE

The following piece is an exaggeration of a planar surface’s defects to determine the efficacy of the method at a larger scale. The surface is 3D printed, and the individual layers can be clearly seen in Figure 4-22.
The skin model developed for the NURBS surface is seen in Figure 4-23. The data used to develop this skin model was collected using a high density laser scanner.
The results for this model are far different than the results for the previous two. The bump code is triggered multiple times, which can be seen in Figure 4-24, and the values after bump are not always consistent. Another interesting thing to note is that the minimum value from WTLS is found just before a bump, not just after like with the previous two models.

4.2.5. Staircase Model

The final piece is a 3D printed piece that has five differently angled surfaces. Each surface increases in angle 15 degrees, for a range of 15 to 75 degrees. The purpose of this is to analyze the “staircase effect” inherent in 3D printed pieces. The lower the angle from the
horizontal, the rougher a 3D printed surface is. This would be a typical problem in a hybrid manufacturing system. Figure 4-26 shows the piece.

**Figure 4-26 Staircase Piece**

Figures Figure 4-27Figure 4-31 show the skin models developed for each of the angled surfaces.
Staircase, 15 Degrees - Skin Model

Figure 4-27 Staircase Piece, 15 Degrees – Skin Model

Staircase, 30 Degrees - Skin Model

Figure 4-28 Staircase Piece, 30 Degrees – Skin Model
The deviations in each of the skin models in Figures Figure 4-27Figure 4-31 is very small, as is consistent with the layer height of each piece. However, these deviations are very important to the results. In Figures Figure 4-32 through Figure 4-36, the results of optimization for each of the developed skin models are shown.
**Figure 4-32 Staircase Piece, 15 Degrees – Repair Cost**

**Figure 4-33 Staircase Piece, 30 Degrees – Repair Cost**
In Figures Figure 4-32Figure 4-36, the patterns vary wildly. Despite all the surfaces being very similar, the shift in the layers greatly affects the results. In the 15 degree case, multiple patterns are shown. After the first two bumps, similar patterns are settled in to, indicating similar local minima. However, after the third bump, the pattern shifts, indicating a new local minima. This is also reflected in the 30 and 45 degree cases, but despite having very similar geometries, the results are vastly different. Finally, in the 60 and 75 degree cases, both of which would be the “smoothest”, there is a return to repeated patterns, indicating similar local minima that are found repeatedly. This seems to indicate that this method finds better results for data sets that are less planar.
FIGURE 4-37 STAIRCASE, 15 DEGREES – FMINSEARCH RESULTS

Staircase, 15 Degrees - Fminsearch Results

FIGURE 4-38 STAIRCASE, 30 DEGREES – FMINSEARCH RESULTS

Staircase, 30 Degrees - Fminsearch Results
Figure 4-39 Staircase, 45 Degrees – Fminsearch Results

Figure 4-40 Staircase, 60 Degrees – Fminsearch Results
Figure 4-41 show the results for fminsearch for each of the staircase results. In each case, the improvement was minimal because of the relatively planar nature of each case.
4.2.6. **Results Comparison**

<table>
<thead>
<tr>
<th>Piece</th>
<th>TLS</th>
<th>WTLS</th>
<th>Fminsearch</th>
<th>WTLS over TLS improvement</th>
<th>Fminsearch over WTLS improvement</th>
</tr>
</thead>
<tbody>
<tr>
<td>3D Printed Surface</td>
<td>7251.25</td>
<td>7180.35</td>
<td>7060.64</td>
<td>0.98%</td>
<td>1.67%</td>
</tr>
<tr>
<td>Gouged Wax</td>
<td>1982.34</td>
<td>1926.57</td>
<td>1869.97</td>
<td>2.81%</td>
<td>2.94%</td>
</tr>
<tr>
<td>Metal, Two Hole</td>
<td>300.677</td>
<td>269.596</td>
<td>235.067</td>
<td>10.34%</td>
<td>12.81%</td>
</tr>
<tr>
<td>NURBS Surface</td>
<td>16214.4</td>
<td>13950.4</td>
<td>12952.9</td>
<td>13.96%</td>
<td>7.15%</td>
</tr>
<tr>
<td>Staircase, 15 Degrees</td>
<td>33.49</td>
<td>26.74</td>
<td>26.468</td>
<td>20.16%</td>
<td>1.02%</td>
</tr>
<tr>
<td>Staircase, 30 Degrees</td>
<td>67.32</td>
<td>58.63</td>
<td>57.077</td>
<td>12.91%</td>
<td>2.65%</td>
</tr>
<tr>
<td>Staircase, 45 Degrees</td>
<td>100.91</td>
<td>81.75</td>
<td>81.275</td>
<td>18.99%</td>
<td>0.58%</td>
</tr>
<tr>
<td>Staircase, 60 Degrees</td>
<td>101.539</td>
<td>88.54</td>
<td>86.26</td>
<td>12.80%</td>
<td>2.58%</td>
</tr>
<tr>
<td>Staircase, 75 Degrees</td>
<td>100.527</td>
<td>90.82</td>
<td>86.81</td>
<td>9.66%</td>
<td>4.42%</td>
</tr>
</tbody>
</table>

In Table 4-1, the results of each case study is shown. The results after TLS fitting, the WTLS process, and fminsearch are shown. The cost savings of each method are then compared, WTLS being compared to TLS, and fminsearch being compared with WTLS, to see the amount of improvement found in each case. In almost all cases, fminsearch improves less.
over WTLS than WTLS over TLS. The best results are found in the surfaces that have the least amount of smoothness initially.

5. CONCLUSIONS AND RECOMMENDATIONS

In this thesis, a method to minimize the repair cost of a planar manufactured part through a modification of the fit plane was developed. The results of the developed method are discussed, along with viable use cases and methods for improving and adding more functionality to the developed method.

5.1. CONCLUSIONS

This thesis presents a method that could be used to minimize the cost of repair for planar surfaces which had been rendered unusable either by manufacturing defect or by marring that had occurred through use. The developed method involved traditional methods, such as total least squares fitting and heuristic optimization algorithms, and non-traditional methods, such as skin modelling and weighted total least squares fitting. Various algorithms were developed and utilized to determine the volumes that would need to be added or subtracted in order to repair a piece. These values were then used to weight the points of the original data set toward changing the fit plane for the sake of minimizing cost. The best result from this method was then used in a heuristic optimization algorithm as the initial condition to determine the best result. Each method was validated analytically and experimentally. Cost savings in excess of 10% were found in most cases with the developed method. This method worked best in situations where
there were a large number of irregularities on a part. In pieces where the majority of the surface was unharmed and planar, the cost savings were minimal. These results show that the developed method could be used and refined to aid in the repair and fabrication of parts using hybrid manufacturing. It also has shown that weighted total least squares fitting can be used to intelligently determine an initial starting point for an optimization algorithm, and can, in some instances, find a value close to the optimal result.

5.2. RECOMMENDATIONS

An improvement to the developed method is the expansion of the method to include non-planar surfaces. By adapting the developed for use on sculptured surfaces and 3D features, it would possibly provide a simple way of minimizing cost on hybrid manufacturing techniques. Another improvement would be to take into account the topology of the errors being analyzed. While this is less of a factor for subtractive repair, this is a large problem in additive repair. The surface topology directly affects the cost of this method, and it cannot be assumed to be static. By incorporating topographical analysis, this process could account for more additive techniques and their intricacies.
6. REFERENCES


