DETECTION AND CLASSIFICATION OF DEFECTS IN METAL PARTS USING ATOS PROFESSIONAL 2019

by

Ryan Gorman

A thesis submitted to the faculty of
The University of North Carolina at Charlotte
in partial fulfillment of the requirements
for the degree of Master of Science in
Mechanical Engineering

Charlotte

2022

Approved by:

_____
Dr. Edward Morse

_____
Dr. Jimmie Miller

_____
Dr. Brigid Mullany

ABSTRACT

RYAN GORMAN. Detection and Classification of Defects in Metal Parts using ATOS Professional 2019 (Under the direction of DR. ED MORSE)

When inspecting a manufactured part to determine whether it falls within tolerance, it is often important to distinguish the difference between a defect and geometric variation. A part that exbibits excessive geometric variation will fail to meet the specifications of the design, such as a nominally flat surface that is bowed in the middle or a sphere that is not perfectly round. By comparison, a part that contains defects may meet the design tolerances, but contains imperfections that can range in severity from a cosmetic deformity to one that completely compromises the utility of the part. The role of defect detection in the manufacturing process is to make that distinction between the presence of a defect and geometric variation, and to classify the severity of detected defects so an educated decision can be made about whether the part is suitable for use.

While custom applications can be utilized in some environments for defect classification and detection, they can be expensive and time consuming to create and maintain. Due to this issue, there are a variety of commercial applications that have been created in an attempt to streamline and simplify the defect workflow. The ATOS Professional application is one such software, and has the advantage of being directly integrated with the GOM family of measuring instruments.

The goal of this project is to gain a thorough understanding of the defect detection and classification tools available in ATOS Professional 2019, and provide documentation

for use of these tools in future projects. To assist in this study, an artifact with manufactured defects was created in the UNC Charlotte machine shop. A full investigation of this artifact is accomplished by changing the parameters of the surface defect map and surface defect classification tools in ATOS Professional 2019. The data is then analyzed to make recommendations based on the ideal parameters for this artifact, and these parameters are validated on other artifacts to see how well they translate.

Through this research, no correlation was found between the four main variables used to create the defect map when considering the percent error, but a correlation was present between the variables and potential false positives detected. Orientation and inversion were also found to have a small, yet noticeable effect on the results of both the defect map and surface defect classification tools. The parameters determined for the defect artifact translated well to other manufactured artifacts, but were considerably less successful when applied to real world “organic” artifacts.

It is recommended based on this research to utilize the ATOS Professional defect toolset on artifacts with relatively flat surfaces with no dominant surface structure. Careful attention should be paid to the quality of the mesh and orientation of the part in the software prior to analysis. The surface defect map tool should have the maximum defect size parameter set to approximately the width of the widest anticipated defect of interest on the part, the number of directions parameter set to 1, and the type parameter set to depressions only or bulges only, not both. The XYZ parameter should be set based on the orientation of the part.

DEDICATION

I dedicate this work to Katie Cannon and Roxy for their unwavering support.

ACKNOWLEDGEMENTS

Thanks to Dr Morse, Dr Miller and Dr Mullany for being part of my thesis committee and advising me on completion of the thesis. I also thank Rilyn Fox for his contributions to the processing of the data for this project, Jesse Redford for his discussions about AI defect processing, and the team at the Kansas City National Security Complex for their guidance on the work completed.

This work was supported in part by the Department of Energy's Kansas City National Security Campus managed by Honeywell FM&T through the Plant Directed Research and Development program, contract number 705029.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1 INTRODUCTION

In manufacturing processes, metrology techniques are typically applied near the end of the process to ensure a manufactured part is within tolerance and provide quality control. Integrating these techniques directly into manufacturing processes allows for parts to be checked in real time as they are created, which allows for fewer product recalls, fixes to be applied to failing parts before they are shipped out, and an overall reduction in parts that have to be scrapped. One major goal of these quality control inspections is the detection of defects, which can present a unique set of challenges not present in regular part inspection.

An important part of a defect detection inspection process is the ability to distinguish the difference between a defect and geometric variation. A part that exbibits excessive geometric variation will fail to meet the specifications of the design, such as a nominally flat surface that is bowed in the middle or a sphere that is not perfectly round. By comparison, a part that contains defects may meet the design tolerances, but contains imperfections that can range in severity from a cosmetic deformity to one that completely compromises the utility of the part. The role of defect detection in the manufacturing process is to make that distinction between the presence of a defect and geometric variation, and to classify the severity of detected defects so an educated decision can be made about whether the part is suitable for use.

While some organizations still utilize a human to manually check their parts for defects, the proliferation of computers has allowed for defect detection to be performed

automatically. Open-source programming languages such as python allow for anyone with the time and knowledge to write their own application to detect and classify defects. However, the process of writing a custom application can be time consuming and costly, and by the time all of the edge cases, bugs and compatibility issues have been fixed, there may be a new set of software capabilities required to keep up with evolving technologies. As a solution to this problem, some manufacturers of measuring instruments offer a software application integrated with their instruments. This thesis intends to explore the defect related capabilities of one such software, ATOS Professional 2019.

The main factors explored in this work are the parameter settings used by the defect related tools in ATOS Professional. These four parameters are XYZ direction, maximum defect size, number of directions, and type. The parameters are explored by applying the tools to a measurement taken by the ATOS Scanbox 4105 of an artifact machined to contain defects, and several artifacts containing defects taken from the UNC Charlotte machine shop scrap bin.

Chapter 2 addresses the literature reviewed as background for this project, while Chapter 3 addresses the motivation for conducting this research. Chapter 4 serves as an overview of the experimental setup and procedure, and Chapter 5 reviews the data processing conducted on the raw data. This processed data is displayed as the results section of Chapter 6, and Chapter 7 draws conclusions from these results. Finally, Chapter 8 makes recommendations for future research based on the conclusions drawn in Chapter 7, and Chapter 9 presents ideas for future experiments to build off of these results.

CHAPTER 2 LITERATURE REVIEW

To detect defects in a manufactured part, one must first take a measurement of the surface to be analyzed. The technique chosen to measure the surface of the parts for these experiments was structured light scanning, which is a non-contact measurement technique. Systems that employ this technique traditionally consist of two cameras and a projector, although some systems function with only one camera [1]. The projector projects a specific pattern of structured light onto the artifact to be measured, and the cameras take multiple images of it to capture the way the light pattern deforms on the artifact surface. These images are then stitched together into a high-density point cloud using a process known as triangulation [2]. While other measuring instruments output a point cloud at the end of this process, the ATOS Software takes the extra step of automatically combining the captured points into a tessellated surface, saved in an .stl file.

When attempting to inspect an artifact for defects, it is important to define the terms used to describe the imperfections of interest. ISO Standard 3290-1:2014 defines a surface defect as “element, irregularity, or group of elements and irregularities of the real surface, unintentionally or accidentally caused during manufacture, storage, handling or use of the surface” [3]. For individual defects, ISO 9802:2022 defines a scratch as a “score mark made by a hard body” [4], while ISO 25537:2008 defines a dig as “deep, short scratch” [5]. This definition of a scratch was used to describe all of the defects on the measured artifacts that fit this criteria, while the dig type defects were nominally described as dents.

An interest in detecting defects in manufactured parts is still relatively new, with manufacturers in the United States following the craftsmanship model invented in Europe up until the early 19th century [6]. This model operated on the principle that the craftsman should provide his own quality control for the work he does, as selling defective work would risk losing customers. In the late 19th century, Frederick Taylor introduced the idea of Taylorism, which greatly increased productivity by assigning specialized roles to workers and eliminating unnecessary parts of the manufacturing process [7]. While productivity was increased by the implementation of this model, the quality of manufactured goods decreased. To remedy this, the first inspection departments were created by factory managers to find defective parts and stop them from reaching customers [6].

During World War II, it was important that the equipment manufactured for the war effort was safe for the troops to use. To attempt to ensure that equipment was not defective, the United States armed forces utilized huge inspection forces. This quickly proved to be unpopular as personnel retention was low, so the armed forces switched to sampling inspection methods instead of going unit by unit [6]. Sampling has now become an important aspect of quality control in modern manufacturing.

With the advent of sampling inspection, manufacturers could afford to inspect parts more thoroughly for defects. One of the earliest defect detection instruments was the supersonic flaw detector, a device built to bounce supersonic waves off the bonding agent used in Hanford slugs to determine the quality of the bond in a non-destructive manner [7]. Throughout the 1960s and 1970s, ultrasonic instruments proved to be highly

useful in the field of defect detection [8]. While originally the process was confined to military and research applications, technological advances allowed for different methods and techniques for defect detection to be explored. New methods of utilizing technologies such as X-rays [9] and laser scanners [10] allowed researchers to employ defect detection in a wider range of situations.

One of the biggest technological advances happening around this time was the rapid improvement and increasing availability of computers, which touched every aspect of every industry. Researchers began to explore the integration of computers into quality processes, from using them to aid in reducing noise and defects in imaging applications [11] to some early defect detection techniques utilizing early image processing tools to analyze light intensity [12]. As processing power and capability increased, early versions of techniques used today began to be experimented with, such as the use of computer vision algorithms in inspection techniques [13] and full-blown automated inspection processes [14]. These techniques have only become more prevalent and well researched over the last couple of decades.

Modern defect detection techniques are often focused on the use of AI and neural networks to detect defects with speed and accuracy on a wide variety of specimens. Deep learning neural networks have been proven to outperform traditional computer vision algorithms in head-to-head competitions in both processing time and accuracy, and these networks can be applied to existing defect detection processes to greatly increase productivity [15]. The advent of these neural networks has also allowed for great strides to be made in the automation of these vision-based techniques. Implementation of AI

elements has allowed for greater accuracy, lower costs, and overall, less manpower spent on the development of these systems [16]. These tools can also be expanded in scope to apply classification to the detected defects as well, with some possessing classification capabilities right of the box with no alterations, as seen for example in the paper by Aydin et al in which researchers detected and classified defects on railway fasteners using a neural network and image processing [17].

Although in some cases humans may still be able to outperform machines in the realm of defect detection, in defect classification tasks automated systems outperform humans in terms of consistency, reliability and speed [18]. While these systems may perform better by measurable metrics, conventional defect classification systems still require human expertise to designate the pass/fail condition and the numerical thresholds to identify defects. The introduction of advanced algorithms has reduced the need for this human expertise to set these threshold values, and in some cases completely eliminated it entirely. Despite the fact that the success of these algorithms is entirely dependent on the training they receive beforehand, utilization of a robust dataset can allow for extremely accurate detection and classification completely independent of human intervention [18].

While the main focus of this research is on the defect detection and classification tools built in to the ATOS Professional software, it is still important to note the capabilities of these external software tools, as they may serve useful in a direct comparison between these two options. The weaknesses of these systems may also carry over between tools. For example, in an experiment conducted on the neural network known as ResNet-101, the classification layer known as fc1000 erroneously classified ~

2% of defects without the added feature layer known as pool5 [16]. A pooling layer is a two-dimensional filter applied to a defect map that summarizes the features of the region contained within the filter, and is used to reduce the overall dimensions that the neural network needs to consider [19]. This could prove analogous to filtering options available within ATOS Professional.

Furthermore, in the review of modern defect detection models that utilize neural networks, researchers note that while these models can provide fantastic accuracy, larger datasets require models built from the ground up to support the specific inspection the operator wants, as well as high quality and reliable data for all parts that need to be inspected [15]. These issues and others could potentially arise when evaluating the capabilities of the ATOS suite of tools.

CHAPTER 3 EXPERIMENT MOTIVATION

The primary motivation for this experiment was to gain a thorough understanding of the defect tools available in ATOS Professional 2019 for use in experiments in the metrology research labs at UNC Charlotte. After discovering these tools existed in the software for a research project, it was realized that there was relatively little documentation on how to utilize them properly. Reaching out to the software support team responsible for this application returned an answer that they too were not particularly well versed in the use of this tool. Through their parent company, a training program that contained some basic information and guidance was purchased, but this program had a relatively narrow scope and was lacking a lot of the information needed to employ the toolset for the actual artifacts of interest to the Charlotte group.

Using the information gleaned from the training program, these tools were utilized using the generic settings and classification equation for some basic experiments, producing mixed results. While defects were able to be identified and classified using these basic parameters, the results yielded were less than ideal. A sample of these results can be seen below in FIGURE 3-1. As can be seen from the example, the software has marked all the defects as one type, and has identified significantly more defects (over 6000) than were expected to be seen on the surface of the part.

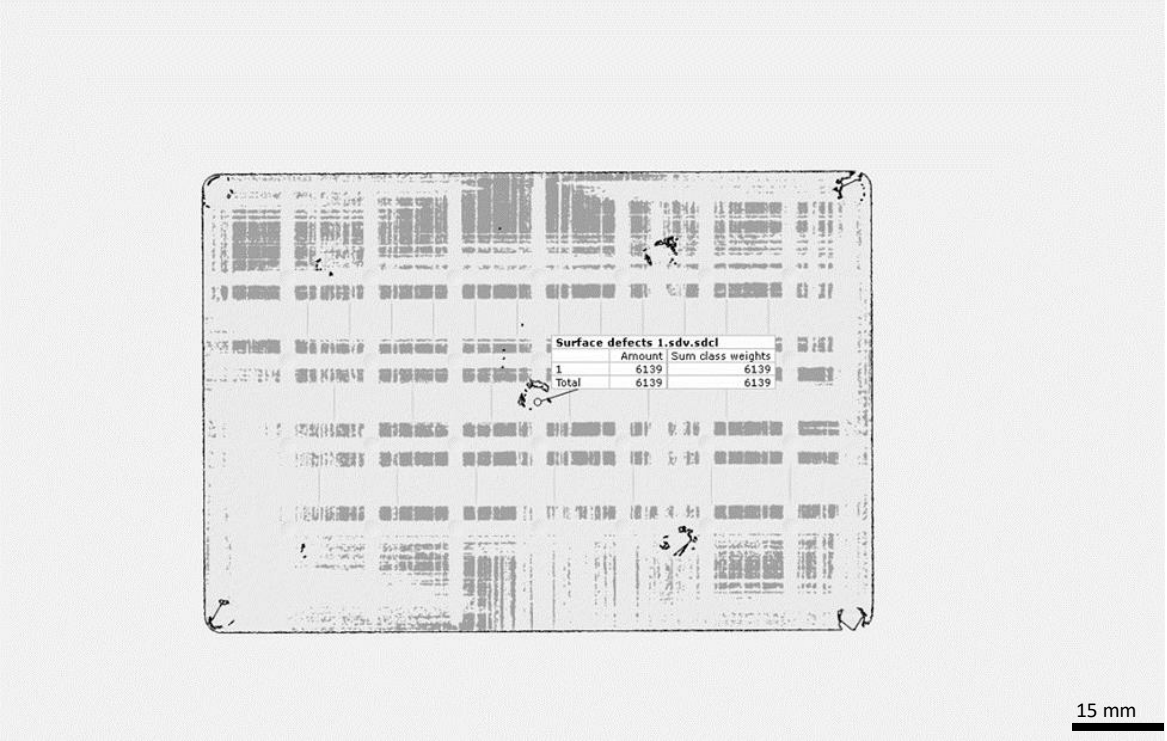

FIGURE 3-1: Basic Classification Example

It was seen that while these basic parameters provided a framework for use of these tools, there were issues in the implementation details. The software would often overestimate the size of the defects and misrepresent their depth. Furthermore, the default classification equation required modification to work on artifacts outside of its original intended purpose, and would fail to properly classify defects by severity as intended even with these changes made. Therefore, this project was proposed and adopted to further the understanding of these tools.

CHAPTER 4 EXPERIMENT PROCEDURE

4.1 Equipment and Software Used

The hardware and software used to perform these experiments are listed here; the majority of components studied in detail are on the software side. To conduct the measurements for the experiments an ATOS Scanbox 4105 was used, with the sensor being an ATOS Capsule 8M. The instrument was paired with its companion software, ATOS Professional, specifically the 2019 version. This application allows for interfacing directly with the instrument to take measurements and perform inspections on measured parts. While the measurements used for this experiment did come from this instrument, it would be most accurate to say that the experiment is being performed mainly on the software.

Seven artifacts were chosen for use with this experiment, with two specifically manufactured to contain regularly-shaped defects and the other five “organic” artifacts, which were leftover scrap from machining processes that naturally contained defects. The first defect artifact was manufactured exclusively with scratches of increasing depth across the face, and was created to test the capabilities of the defect tool suite at varying depths. The second defect artifact was manufactured with 25 dents and 28 scratches evenly spaced on the top face of the same nominal depth. This artifact was created to serve as a baseline for detecting and classifying the two main types of defects this project aimed to capture, dents (also called digs or dimples) and scratches. The five additional artifacts were pieces of scrap metal retrieved from the machine shop scrap bin, and were chosen for their varying defect placements, surface finishes and geometries. The artifact

names and a brief description of each artifact is included below in Table 1, and images of each artifact are included in APPENDIX A. Note that the five "organic" artifacts were down-selected from over a dozen candidates, which is the reason for the non-sequential sample numbering.

Table 1: Artifact Descriptions

Artifact Name	Description
Defect Artifact 1	Manufactured artifact with scratches varying in depth from approximately 25 μm – 500 μm numbered from 1-15 starting with the shallowest scratch. Dimensions of the sample were approximately 25 x 230 x 25 mm
Defect Artifact 2	Manufactured artifact with 25 dents and 28 scratches. Scratches were approximately 125 μm deep and 250 μm wide, while dents were approximately 125 μm deep and 2.5 mm in diameter. Sample dimensions approximately 70 x 101 x 22 mm
Sample 2	Organic artifact with a squared step on the top face. Sample dimensions approximately 38 x 49 x 38 mm
Sample 5	Organic, square artifact with defects on the top face. Additional defects added by operator. Sample dimensions approximately 38 x 32 x 6 mm
Sample 10	Organic, fin shaped artifact with many deformations and bends in the surface. Sample dimensions approximately 279 x 248 x 83 mm
Sample 11	Organic, triangular artifact with a ridged surface finish. Ridges were approximately 70 μm deep. Sample dimensions approximately 51 x 197 x 64 mm
Sample 13	Organic, flat artifact with a bend in the metal. Additional defects added by operator. Sample dimensions approximately 79 x 117 x 2 mm

In addition to the GOM equipment and software, Microsoft Excel and Google

Sheets were used for basic data processing and to generate many of the figures in this paper.

4.2 Setup

As the physical experimentation side of the project was purely for data collection to be evaluated in the software, the physical setup was generally the same for each experiment conducted. Each part was prepared by being sprayed by hand with a rattle can of SKD Developer spray, and GOM 0.4 mm reference point stickers were applied to the surface. After the samples were prepared, they were individually placed in the approximate center of the rotation table in the Scanbox. To scan each part, the sensor was placed at approximately 45 degrees of elevation, and the part was rotated 12 times in 30-degree increments to simulate a standard quick inspection process. Additional shots were then added as needed to fully capture the top face of the part. Each point cloud then underwent polygonization, a proprietary process in the ATOS software that stitches together individual measurements by utilizing an algorithm that produces a tessellated surface with relatively uniform triangle size. The ATOS polygonization function was set to use the "more details" postprocessing option, and afterwards the models were trimmed to exclude all extraneous data such as the rotation table and other pieces of the measuring apparatus. An example of this experimental setup can be seen below in FIGURE 4-1.

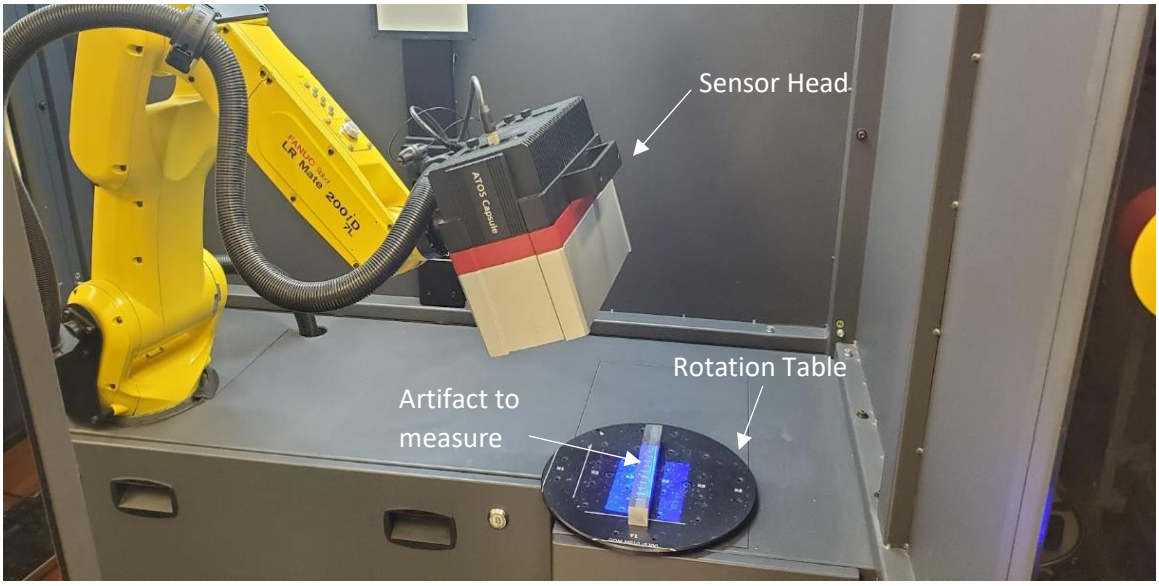

FIGURE 4-1: Example Experimental Setup

In order to focus on the software side of the experiment, many aspects of the inspection process were standardized across the parts to minimize variation. All meshes were imported into a clean ATOS file as .stl files. The parts were then aligned to the XY plane to simulate the position in which they are scanned by default. Each set of different tool parameters utilized were saved as an individual file. A software screenshot showing this setup is shown in FIGURE 4-2.

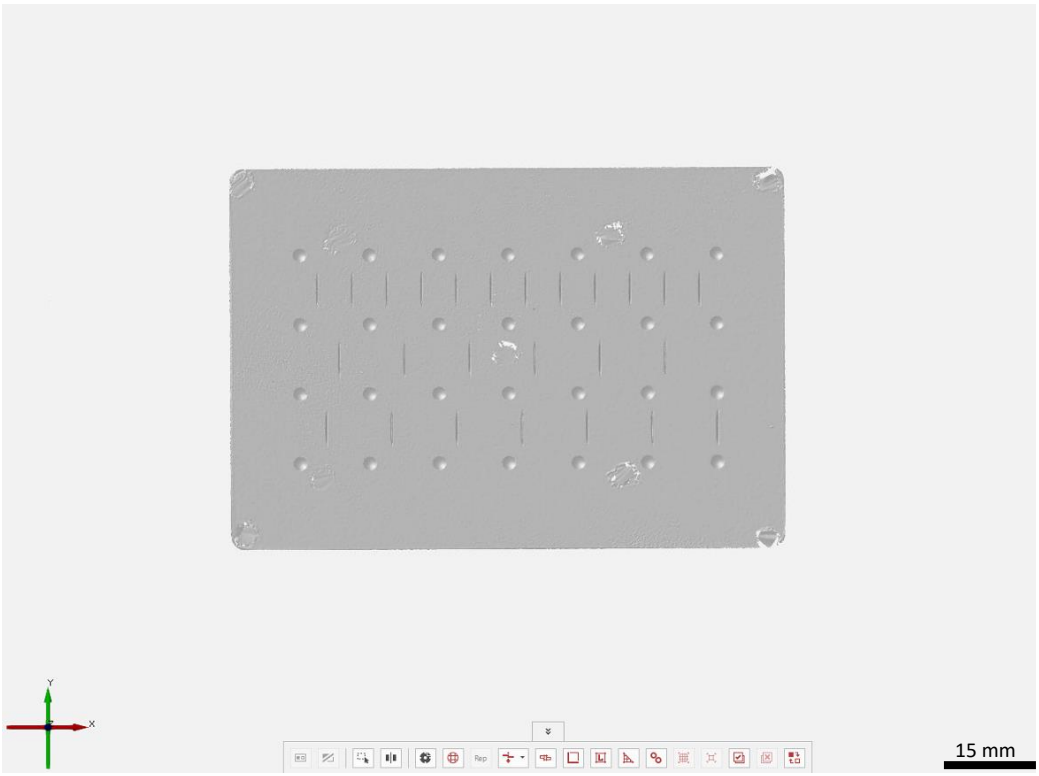

FIGURE 4-2: Example Software Import Setup

4.3 General Experimental Procedure

The main steps to the experiments are described here. The first step is the measurement of the artifact as described above, producing a mesh to be analyzed. The second step is the production of a defect map, which is a map of values superimposed on this mesh. It is in this step that several experiments were run in which the software parameters are varied. One attribute of the defect map is that regions are identified that may be considered as defects. A metric used in this thesis is the number of regions identified, which are referred to as "candidate false positives." The final step is to apply a classification algorithm to the defect map, identifying areas that are considered defects. The physical and software experimental setups were identical for all experiments as seen in FIGURE 4-1 and FIGURE 4-2. Each of the one hundred and forty permutations of the software parameters was applied and then saved as its own file. After these permutations were created, the surface defect classification tool was applied and the results were recorded.

4.4 Procedure Variations

Using the general procedure described above, the various software parameters were adjusted in an organized set of experiments and the results saved for subsequent analysis.

The first experiment focused on the “maximum defect size” parameter, which defines the length of the line that the surface defect map tool runs across the surface of a part to detect defects. This parameter seemed like the single most important guideline to present, as modifying the value changed not only the time needed to compute the map,

but the coverage of defects and presence of "candidate false positives" as well. For defect artifact 2, this parameter was set in steps of 1 mm between 1 and 4 mm. 2.5 and 6.5 mm were also investigated, as they were the maximum length of the dent and scratch defects on the face of the artifact respectively.

The second experiment focused on the “direction” parameter. The software offers the ability to run the previously aforementioned maximum defect line along the surface of the part in a three-dimensional cartesian coordinate system. Considering defects are computed at least in part on a directional basis by the surface defect map tool, it seemed important to experiment with all available orientations for any given position. While all directions may have not been feasible in all part orientations, the directions that were available were applied for all conducted trials.

The third experiment was the simplest conducted, focusing on the “number of directions” parameter. While it initially seemed unimportant, the number of directions parameter allows the operator to choose whether the max defect line is run a second time perpendicular to the original orientation, or just once across the face of the part. While the line is set to run twice by default, this can increase the number of candidate false positives detected by the map. For each trial, the parameter was set to run the line both once and twice.

The fourth experiment addressed the “type” parameter. Depending on whether the parameter is set to detect depressions or bulges, the maximum defect line will be run along either the top face, or the underside of the face in question to detect defects. Selecting both will cause the line to search for both kinds of defects. While the tool is set

to both by default, in many cases selecting both will cause not only a large number of candidate false positives, but also significant noise to be detected even in the defects that are correctly identified. For each conducted trial, the parameter was set to the either depressions only, bulges only, or both.

FIGURE 4-3 below displays a sample configuration window for the surface defect map tool that displays the four main parameters as they are configurable in the software.

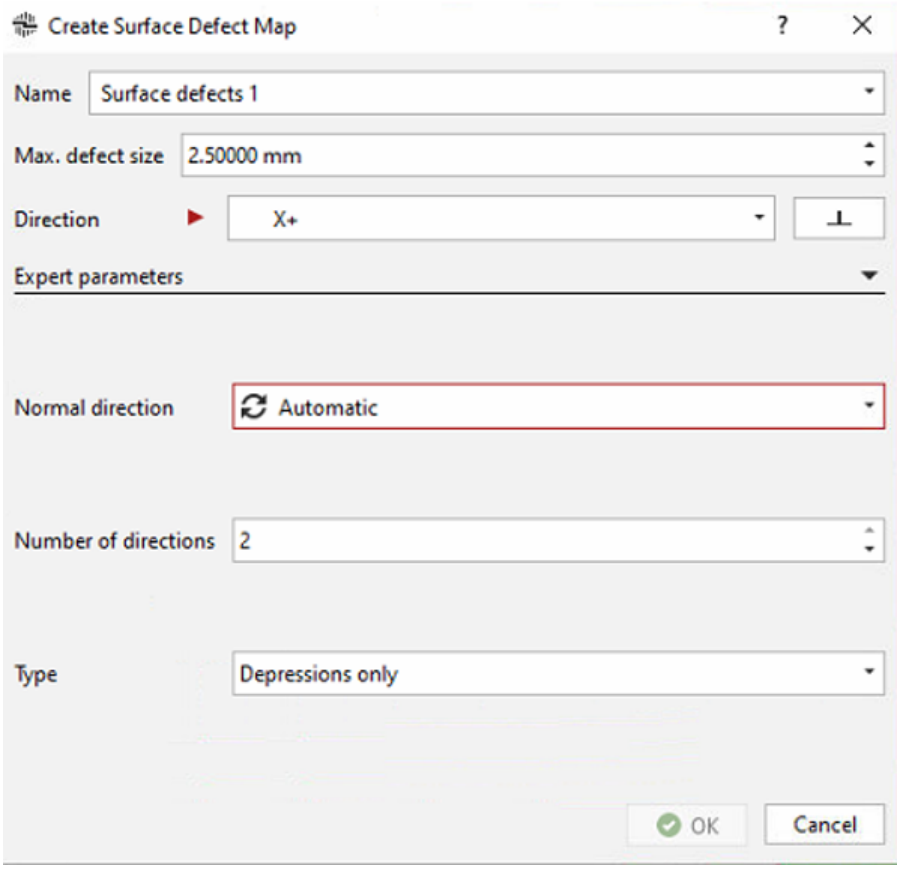

FIGURE 4-3: Example Defect Map Configuration Window

The fifth experiment was unique out of the conducted experiments, as it contained the only changes made to the setup throughout the experimental suite. As the surface defect map is based on the direction and length of the line, it would stand to reason that changing the orientation of the part would change the way the defects were calculated. For all previously run parameters, another set of trials was done where the orientation was changed from 0-90 degrees in increments of 15 degrees by rotating the artifact about the Z axis in the software, effectively changing the X and Y directions of evaluation.

The sixth and final experiment was procedurally identical to Experiment 1. The mesh for defect artifact 2 was inverted using the "invert selected normals" function in the ATOS software. The same trials from Experiment 1 were then conducted to investigate how the software handled inverted defects of equal magnitude.

4.5 Issues to Note

There are some potential issues to note during the experimental phase that may cause some discrepancies in results interpretation. While the defects on the manufactured defect artifacts were precisely manufactured and therefore could have their dimensions estimated with a good degree of certainty, the estimations of defect size were less precise for the organic artifacts. As such, the maximum defect size could not be set as accurately during the validation experiments. This could cause errors in the calculation of the defect map. For Experiment 2, parts that contained distortions in the face being measured for defects could not be aligned perfectly to the XY plane. Due to this discrepancy, there could be error introduced by the direction of the touching line differing from other

specimens. This error could also be present in Experiment 5, as the orientation of the part when rotated is directly dependent on the initial alignment in ATOS Professional 2019.

4.6 Defect Classification and AI

The project was initially conceived with an additional experiment in which the parts classified using the built in ATOS Professional defect classification tool would also be processed using a custom AI defect detection program. The results of this program would have then been compared to the results from the built-in tools, and the possibility of determining more precise filtering parameters from this dataset would have been explored. However, while the AI program worked well on images, adapting it to run on 3D models in a different computing environment proved to be a massive undertaking, as there were numerous bugs and compatibility issues to work through. Due to the time constraints of the master's thesis format, the decision was made to drop this experiment and pass the work completed on it as a recommendation for future work.

CHAPTER 5 DATA PROCESSING

5.1 Mesh Processing

Due to the insulated nature of the GOM software package, it would have been necessary to perform almost all of the data processing in the software even if the experiments were not designed around their use. The software does not allow for the export of the raw data taken from an attached instrument. To create a mesh in ATOS Professional, the individual measurements that make up the raw data must be stitched together using polygonization. When using the polygonization function, there are five options for the detail level of the resulting mesh. These levels are merely descriptive, and have no quantitative values associated with them to assist the operator. The five options available are No postprocessing, More details, Standard, Less details, and Smallest data volume. It was decided that since the experiments being conducted were primarily concerned with the detection of defects of various shapes and sizes on the surface of a part, the best option would be to use the "More details" postprocessing option, as it is realistically used in some industry applications and would provide the highest resolution view of the surface of the parts being inspected. An example of one of the polygonized meshes can be seen below in FIGURE 5-1.

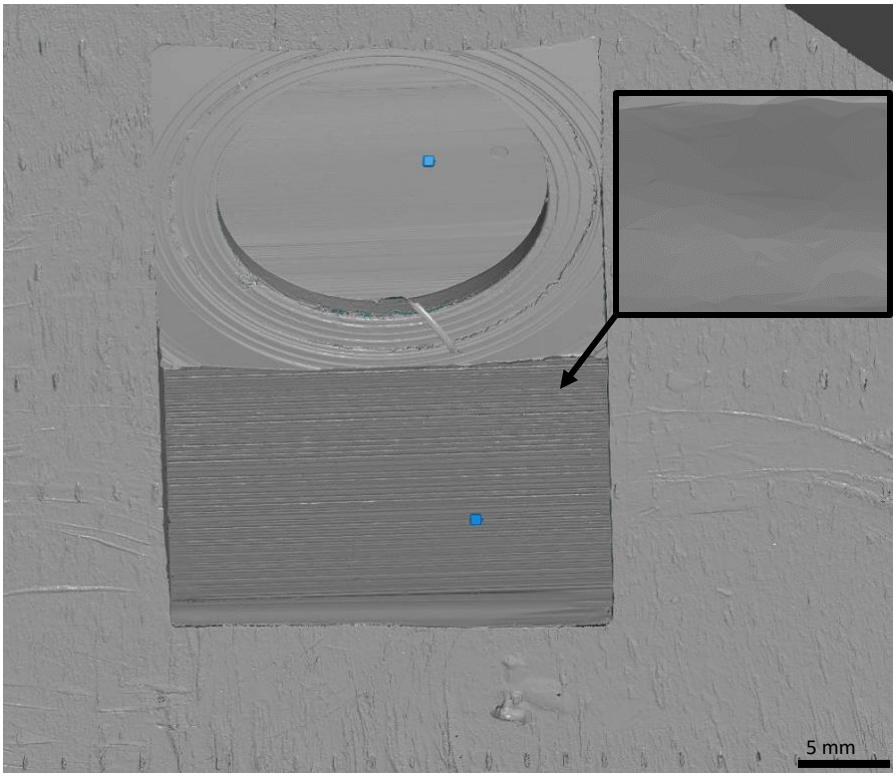

FIGURE 5-1: Mesh polygonization example. The inset shows a closer view of the triangulated point cloud

After polygonization, each mesh was exported from the software and then reimported as an .stl into a new file to remove any extraneous elements the software may have held over from the acquisition portion. A plane was then fit to each mesh using a gaussian plane fit while removing outliers beyond 3 sigma of the data. This plane serves as the basis to create a local coordinate system on the right corner of the mesh, with a line and edge point serving as the two other components of the coordinate system. For parts that were not flat, these elements were constructed on the flattest available section of the mesh. Once this coordinate system was created, an alignment by coordinate systems was created to align the part to the XY plane. This is the default orientation a

part would be scanned in directly from an automated instrument like the ATOS Scanbox.

Once the mesh was aligned, a "deviation to geometry" check was built comparing the plane to the mesh. This check provides the Z deviation of the measured mesh when compared to a geometric feature, which in this case was the fitted plane. The surface defect map tool was then applied using the various parameters described in the experimental procedure of CHAPTER 4. Each experiment was saved as its own file to streamline the data processing procedure. The surface defect classification tool was then built on top of the surface defect map, and the classification formula described further in CHAPTER 6 was applied. It was at this point that the dataset was split again, with further inspection elements being built on a partial section of the mesh. An example of a full mesh inspection can be seen below in FIGURE 5-2, which includes the outline of the partial mesh used for additional analysis.

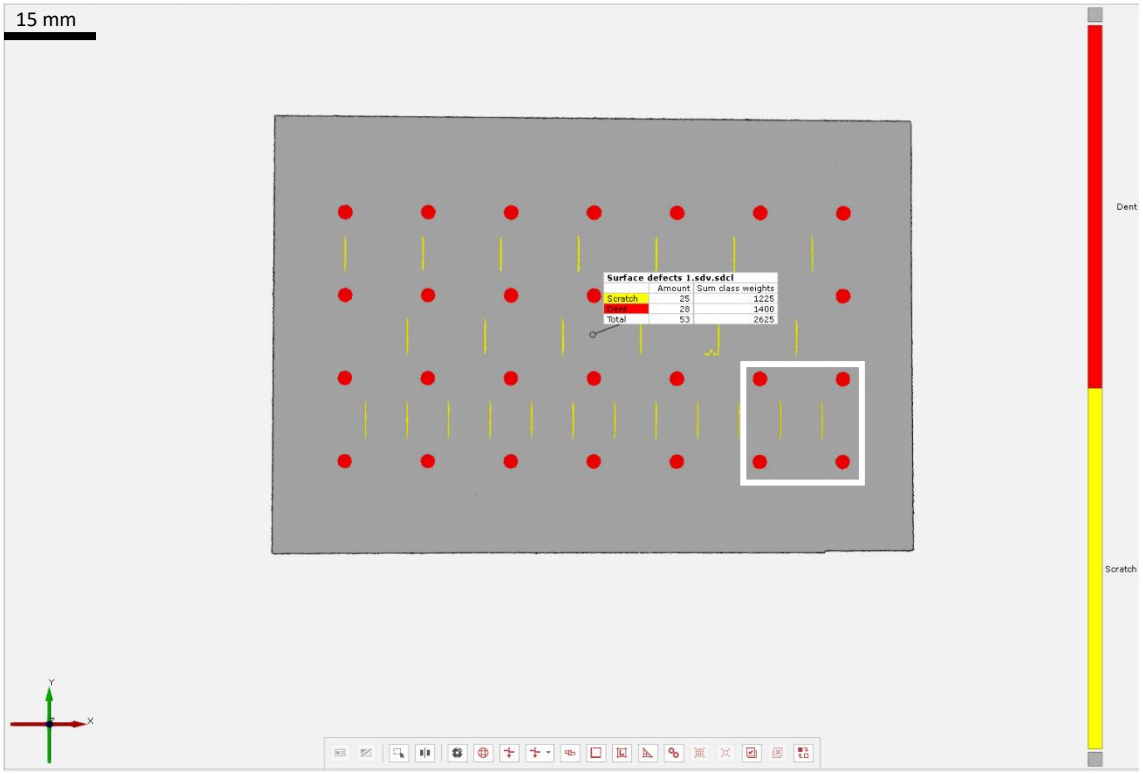

FIGURE 5-2: Example full mesh surface classification. The section used for the partial mesh inspection is indicated in white

Because the full mesh contained many similar defect features, the partial mesh was used for many analyses to speed the processing. For the partial mesh inspection, the bottom right corner of the defect artifact was chosen as a representative section of the whole mesh, as it contained a good density of defects.

The following analysis method was used to determine the change in deviations related to the construction of the defect map from the actual mesh surface. The defect regions were selected, and a number of points were placed covering the surface area of the defect using the equidistant points function at a point-to-point distance of 0.5 mm. The points were placed in the same location across both the deviation to geometry

feature and the surface defect map. Each point was labelled, and the data was exported to Excel. A diagram of the defect locations and names for defect artifact 2 is included in APPENDIX B, and an example of this partial mesh inspection is included in FIGURE 5-3.

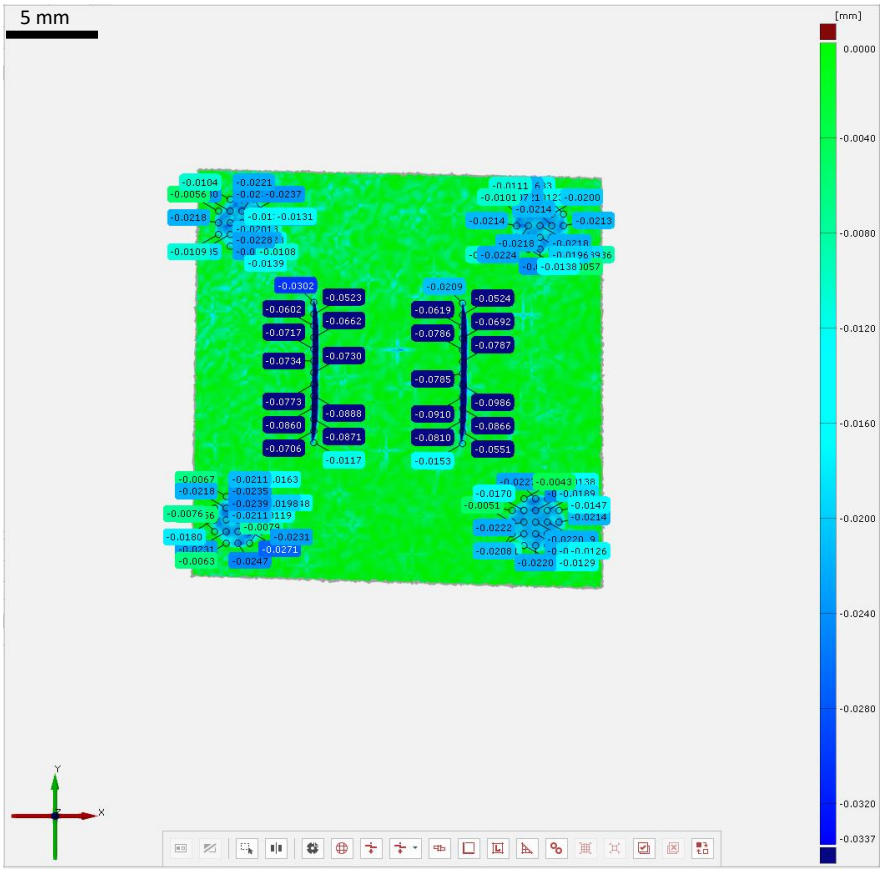

FIGURE 5-3: Example partial mesh inspection

5.2 Excel Processing

Once the CSV files of deviation values were exported from the partial mesh inspection plans, the deviation of each point was compared to the deviation from geometry check. The deviation value reported by the mesh was treated as the ground

truth value to further simulate an inspection conducted solely in ATOS Professional. By dividing the two values, the relationship between the two numbers was expressed as a percent error. Combining all the values that make up a defect and averaging this percentage gave the "total defect coverage" for that set of parameters. This term is used to represent the magnitude change between the mesh and the defect map produced using each different parameter set.

For the full mesh inspection, the surface defect classification tool was applied directly to the unfiltered mech. The classification tool utilizes an equation which calculates the local slope to get a representative value for each of the defects detected by the surface defect map. This equation is given in section 6.8 and is explained in greater detail there. The table of values was then exported, and an analysis performed to determine appropriate parameter settings for rules to filter out the false positives. After these filtering rules were applied, the number of false positive defects and true positives were recorded in Excel as well. The data was averaged for each type of parameter tested, and conclusions were drawn on the optimal set of parameters using the results.

Once a defect was detected, it was designated as a candidate true or false positive. To filter past this point and determine whether a defect is a true or false positive, several guidelines were utilized. Visible defects on the surface of the artifact were identified prior to inspection as the ground truth set of true positives. If the available area of defect is more than 80 percent identified, the defect was counted as a true positive. If less than 80 percent of the defect was identified, an approximation was calculated by estimating the total percentage of defect identified. A false positive was defined as any defect

identified by the software that was not visible in a human inspection. Any positive that overlapped with an identifiable defect was a false positive as well.

In the results chapter that follows, the two primary metrics that are used to evaluate the influence of software settings are the percentage error from filtering and the candidate true and false positive counts from the creation of the defect map.

CHAPTER 6 RESULTS

6.1 Experiment 1

The measured percent error in the deviation reported in the Z direction between the mesh and the surface defect map for the partial mesh inspections can be seen below in FIGURE 6-1. For all scatter plots in the results section, five percent error bars were used as standard deviation error bars made many of the figures significantly harder to read.

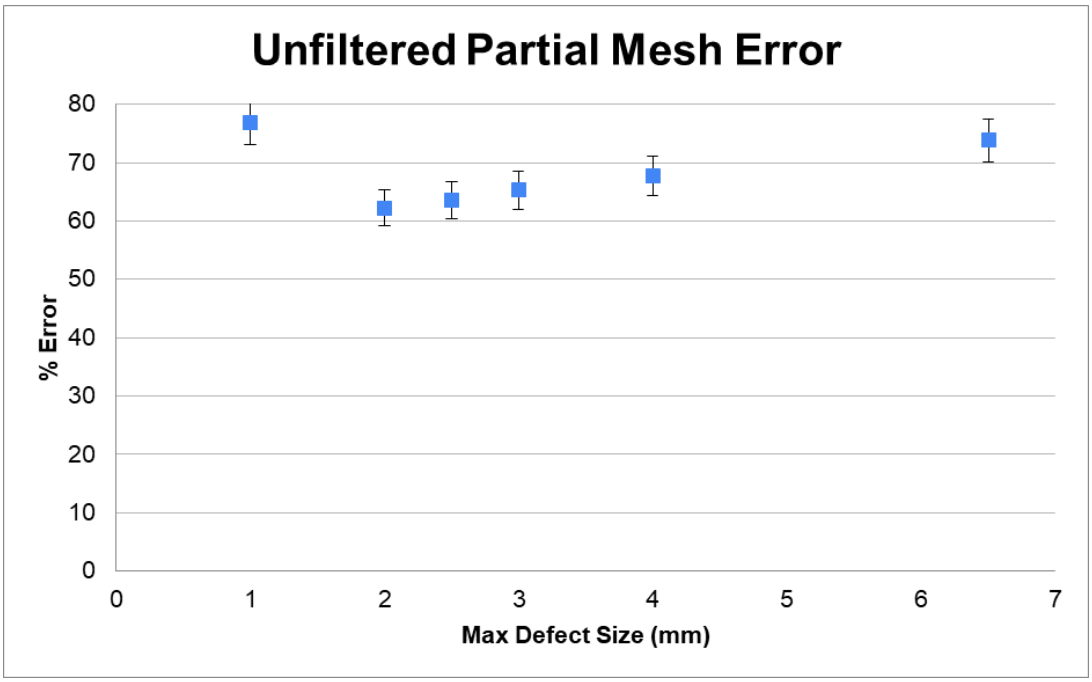

FIGURE 6-1: Unfiltered error for Experiment 1 partial mesh inspection

The raw data for these experiment as is contained significant outliers in the dataset. Because these outliers persisted across all trials involving the partial mesh inspections, the decision was made to generate a second dataset for these experiments utilizing a 2-sigma filter. An example of the plot after the filter is applied can be seen

below in FIGURE 6-2, and FIGURE 6-3 breaks down the data even further into the reported error for each defect.

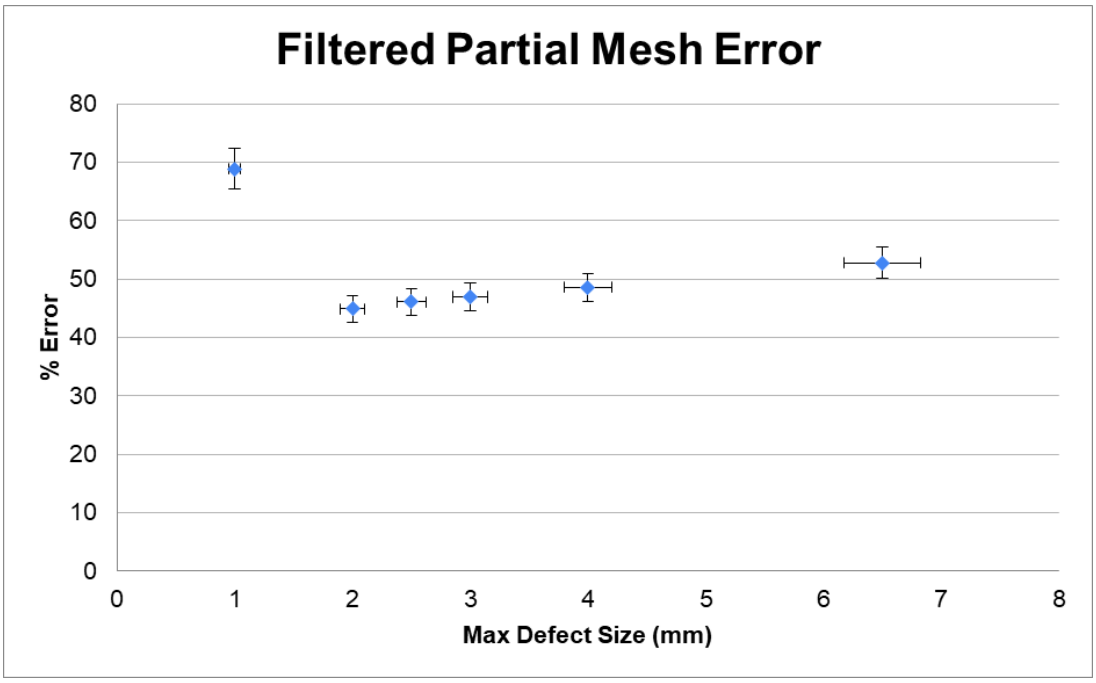

FIGURE 6-2: Filtered error for Experiment 1 partial mesh inspection

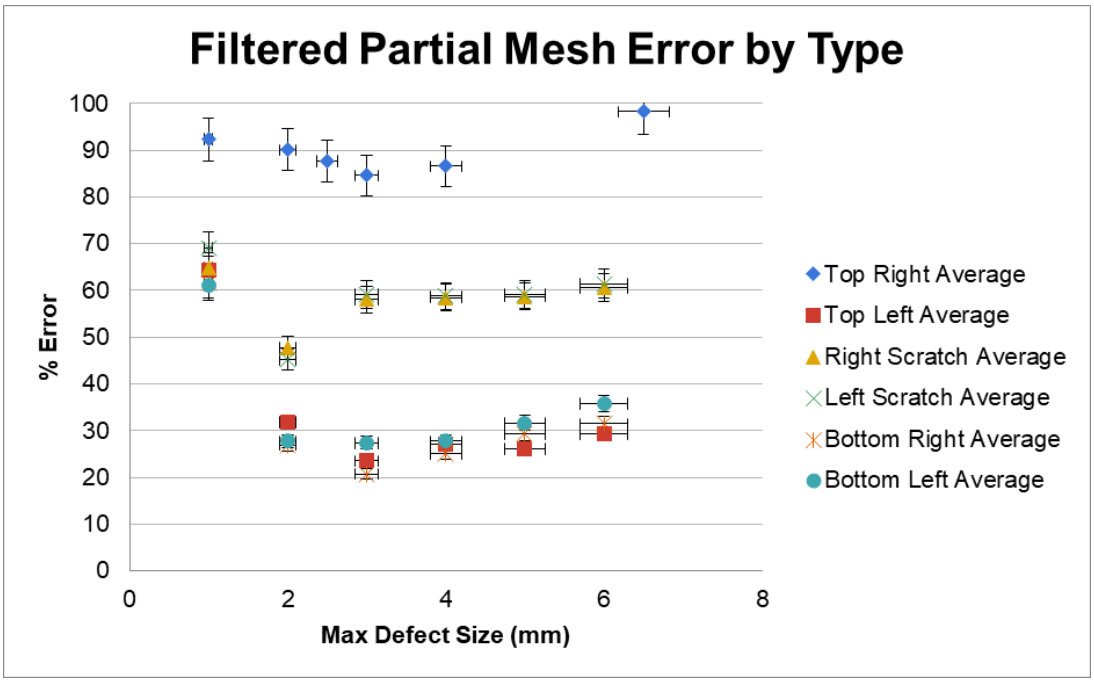

FIGURE 6-3: Filtered error sorted by defect location and type for Experiment 1

Observing these plots, it is instantly noticeable that the lowest amount of error occurred around the 2-3 mm range, with an upward trend in error observed on both sides of this range. The maximum width of any defect on the surface of the artifact in the X direction was 2.5 mm, and the maximum length in the Y direction was 6.5 mm. As such, these data points were included as well. While 2.5 mm represented the lowest calculated error on the raw data plot, after applying the filter the error was slightly lower at 2 mm. Both plots also had 6.5 mm as the length with the 2nd most error, only trailing 1 mm. There was no notable deviation from these trends for any individual defect, with each one following the same basic trend. However, the scratch defects exhibited significantly higher error overall than the dent defects, with the exception of the dent in the top right corner of the sample that registered more error than any other defect examined.

A similar trial was also conducted for the full mesh, with the candidate true positives recorded in FIGURE 6-4 and the candidate false positives recorded in FIGURE 6-5.

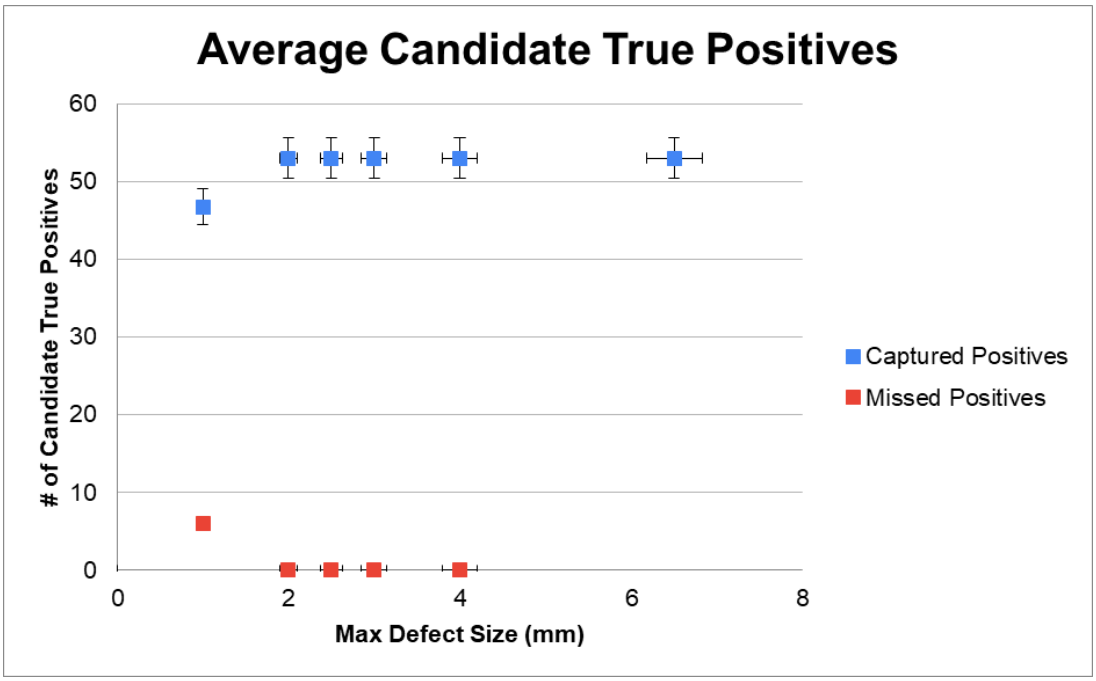

FIGURE 6-4: Average candidate true positives for Experiment 1 full mesh inspection

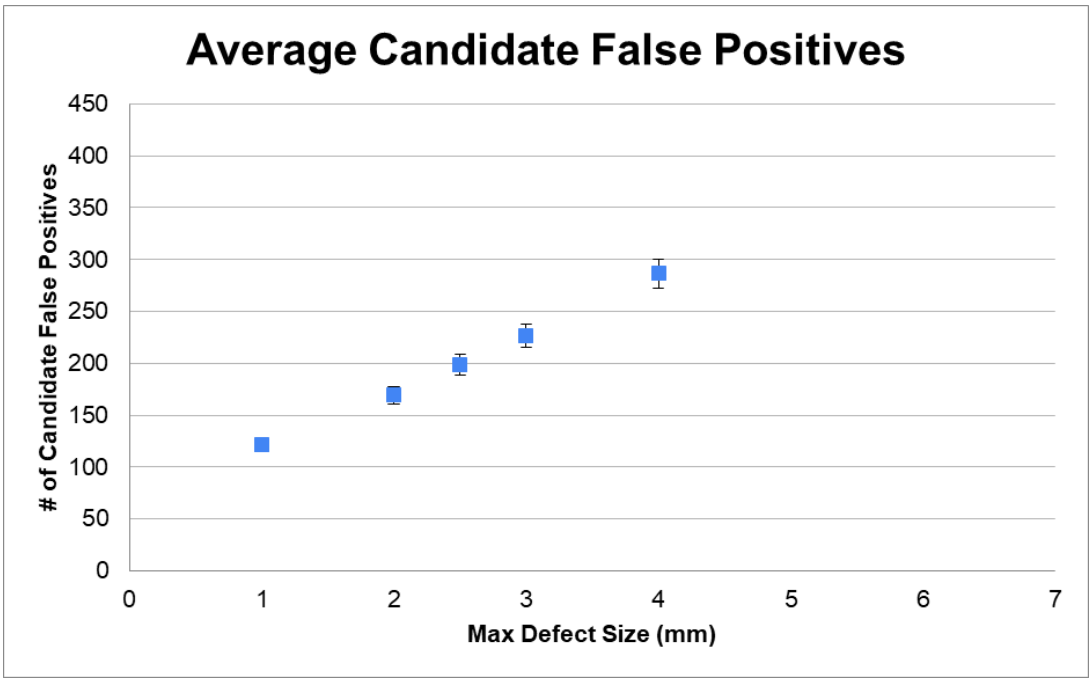

FIGURE 6-5: Average candidate false positives for Experiment 1 full mesh inspection

Both figures displayed a clear upward trend that increased with the length of the line. The candidate true positives leveled off at 2 mm, while the candidate false positives continually increased as the line increased. There were no instances in which a defect was detected in two or more pieces, so this case was not considered when totaling the positives for this artifact.

6.2 Experiment 2

The partial mesh data for the XYZ direction parameter, both unfiltered and filtered through a 2-sigma filter, can be seen below in FIGURE 6-6 and FIGURE 6-7.

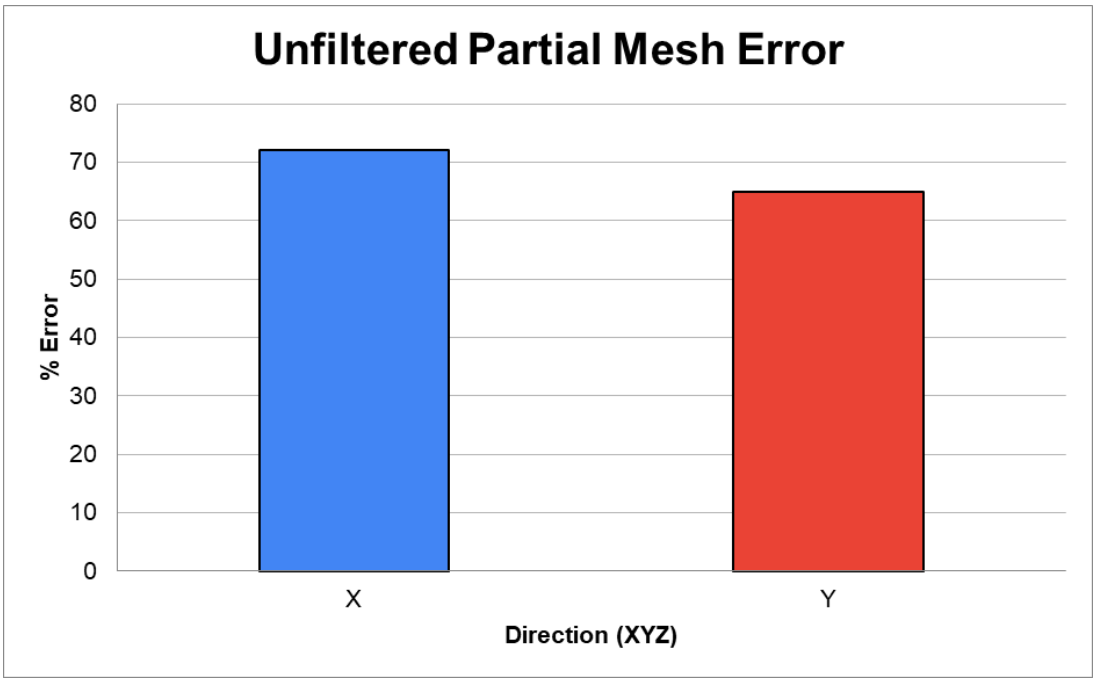

FIGURE 6-6: Unfiltered error by XYZ direction for Experiment 2 partial mesh inspection

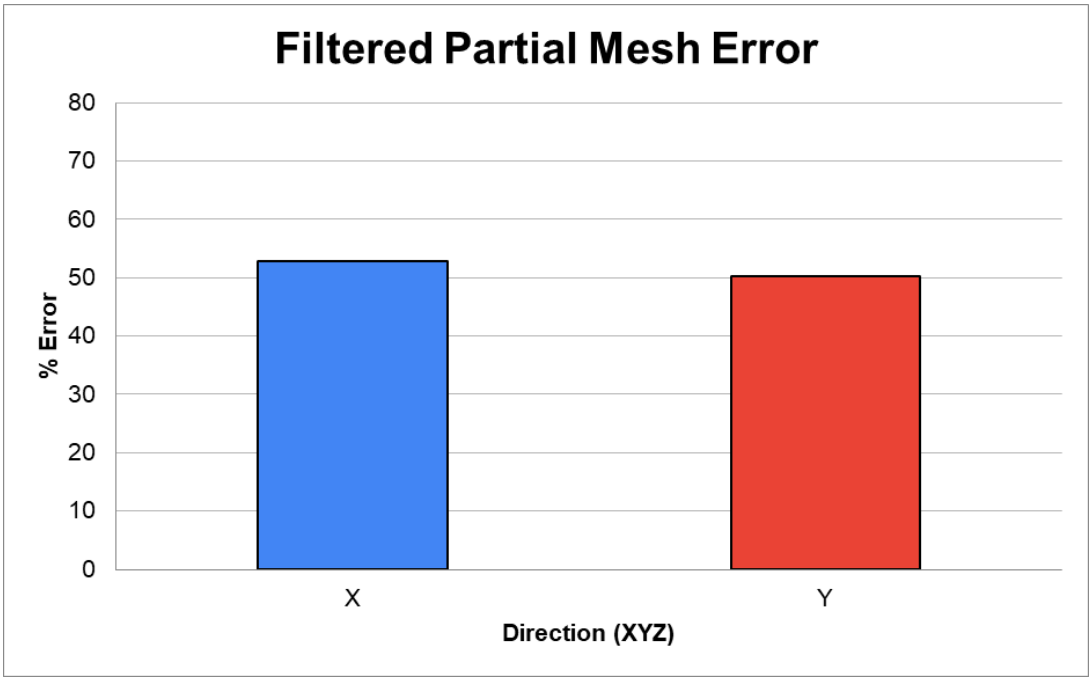

FIGURE 6-7: Filtered error by XYZ direction for Experiment 2 partial mesh inspection

Looking at this data, it appears that the Y direction contains less error than the X across both the filtered and unfiltered datasets. While the overall average does provide Y with slightly less error, breaking it down by defect in FIGURE 6-8 provides a much less clear picture.

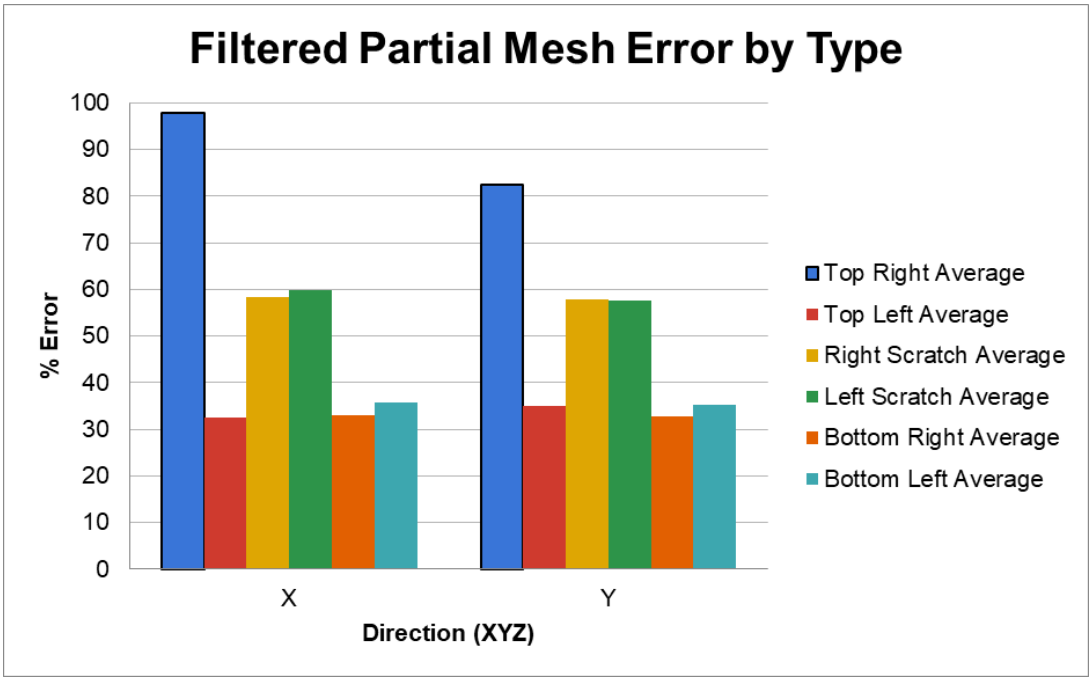

FIGURE 6-8: Filtered error sorted by defect type for Experiment 2 partial mesh inspection

It is immediately noticeable from this chart that the majority of the difference between X and Y comes from the top right defect, with every other defect registering within 3 percentage points of each other. Removing this defect from the dataset reduces the difference between the two directions to approximately 0.1 percent.

FIGURE 6-9 and FIGURE 6-10 show the candidate true/false positive data for the full block trial.

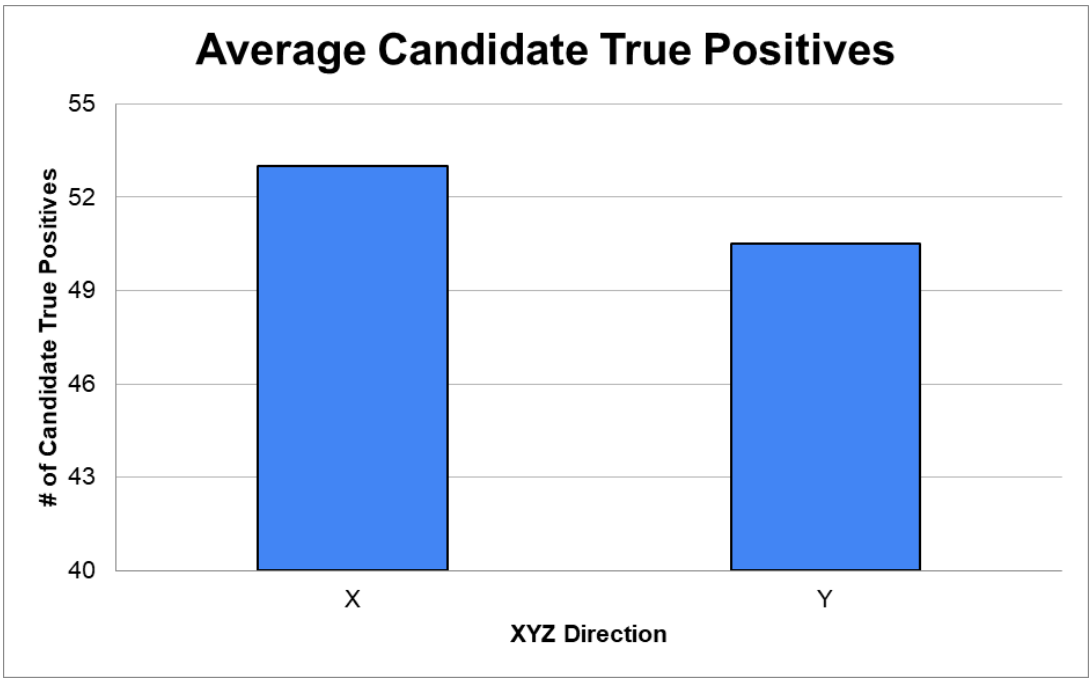

FIGURE 6-9: Candidate true positive data for Experiment 2 full mesh inspection

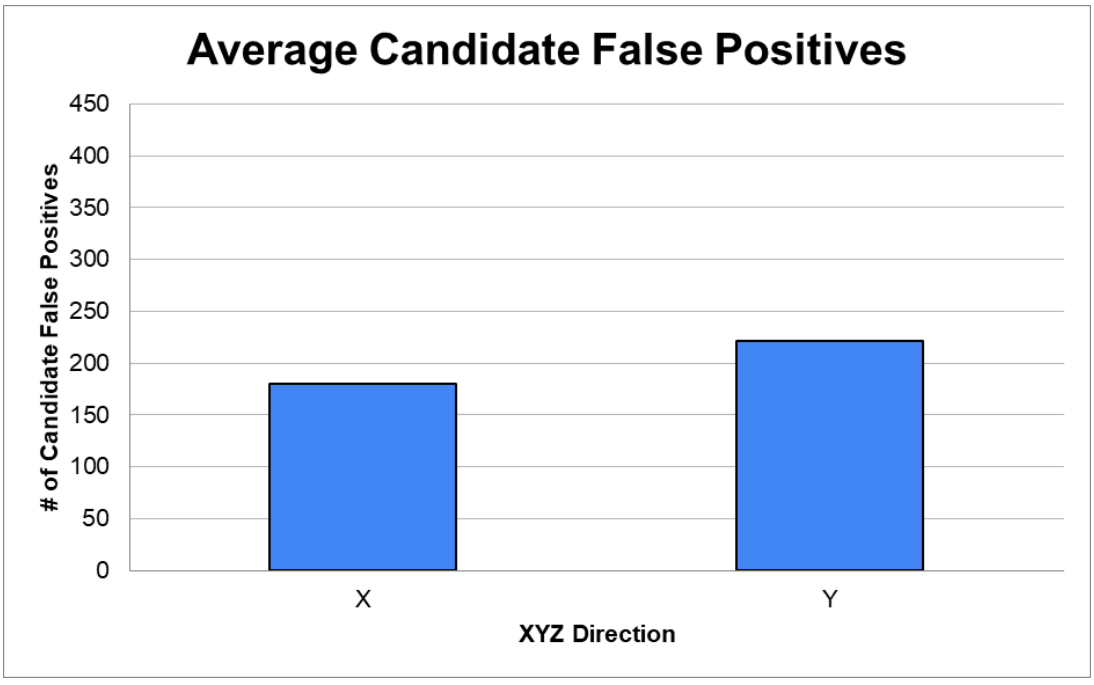

FIGURE 6-10: Candidate false positive data for Experiment 2 full mesh inspection

This dataset appears to provide a clearer trend than the partial block data, with X containing both more candidate true positives and significantly less candidate false positives than Y. However, deconstructing this average once again makes drawing a clear distinction between X and Y impossible. The Y direction contained less candidate true positives due to one datapoint, and while the difference was more noticeable in the candidate false positives dataset, the differential was very inconsistent across the trials, as seen in FIGURE 6-11.

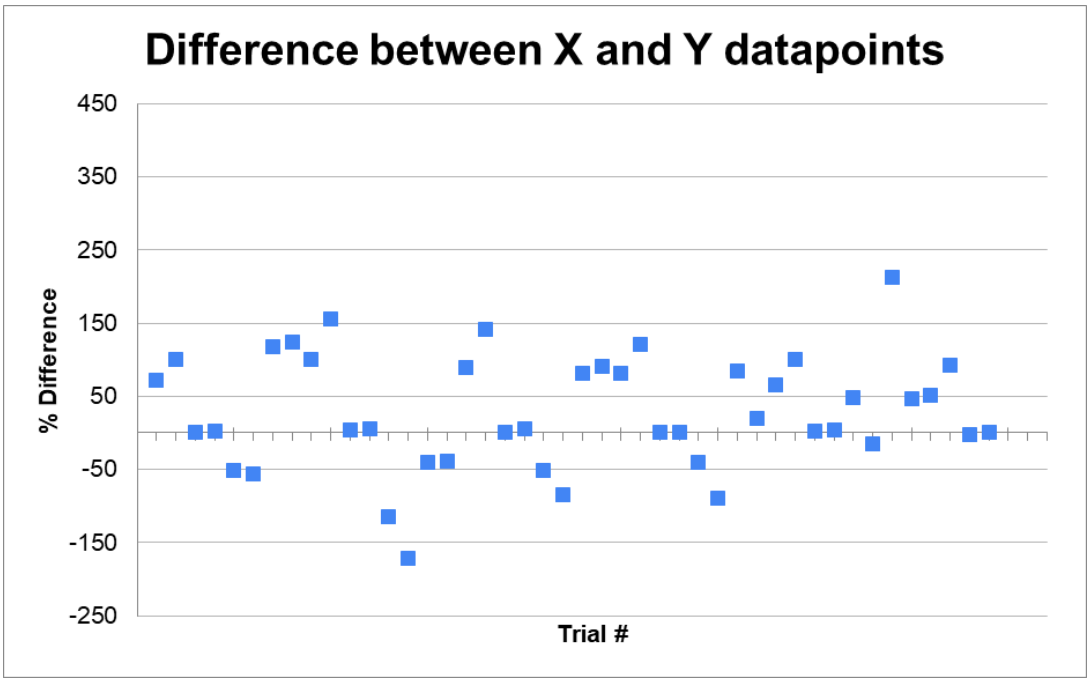

FIGURE 6-11: Calculated difference between Y and X values for all datapoints in the Experiment 2 full mesh inspection

Considering the lack of definitive trend across all the figures presented, it appears there is no definitive preferred direction to use with the defect map tool with a default alignment.

6.3 Experiment 3

The partial mesh data for the number of directions parameter, both unfiltered and filtered through a 2-sigma filter, can be seen below in FIGURE 6-12 and FIGURE 6-13.

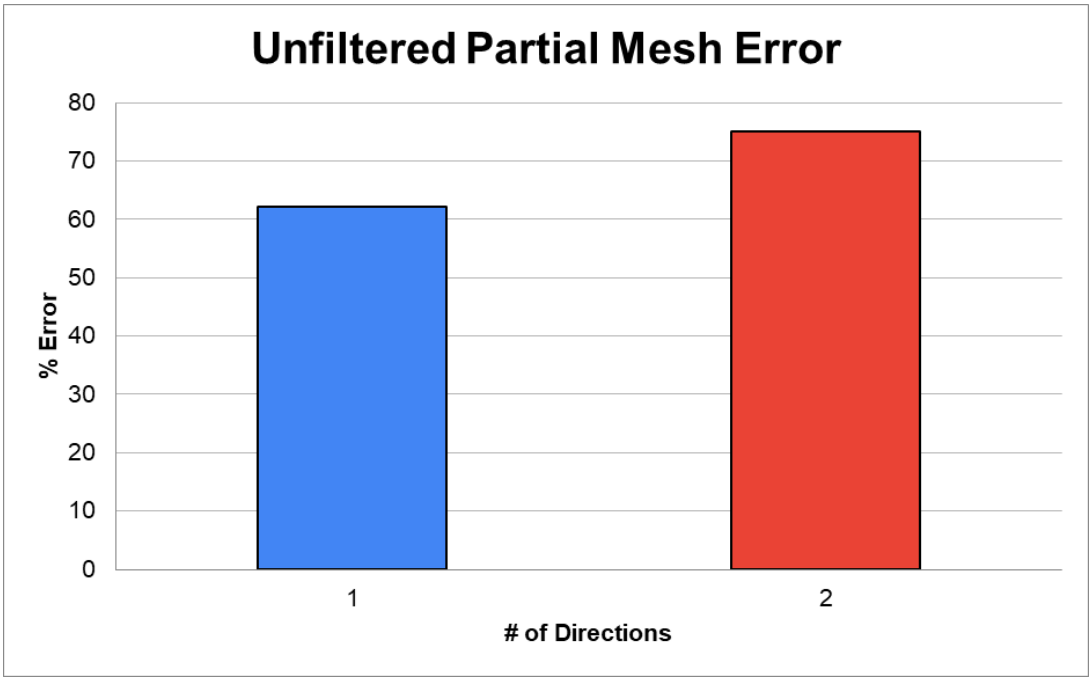

FIGURE 6-12: Unfiltered error for Experiment 3 partial mesh inspection

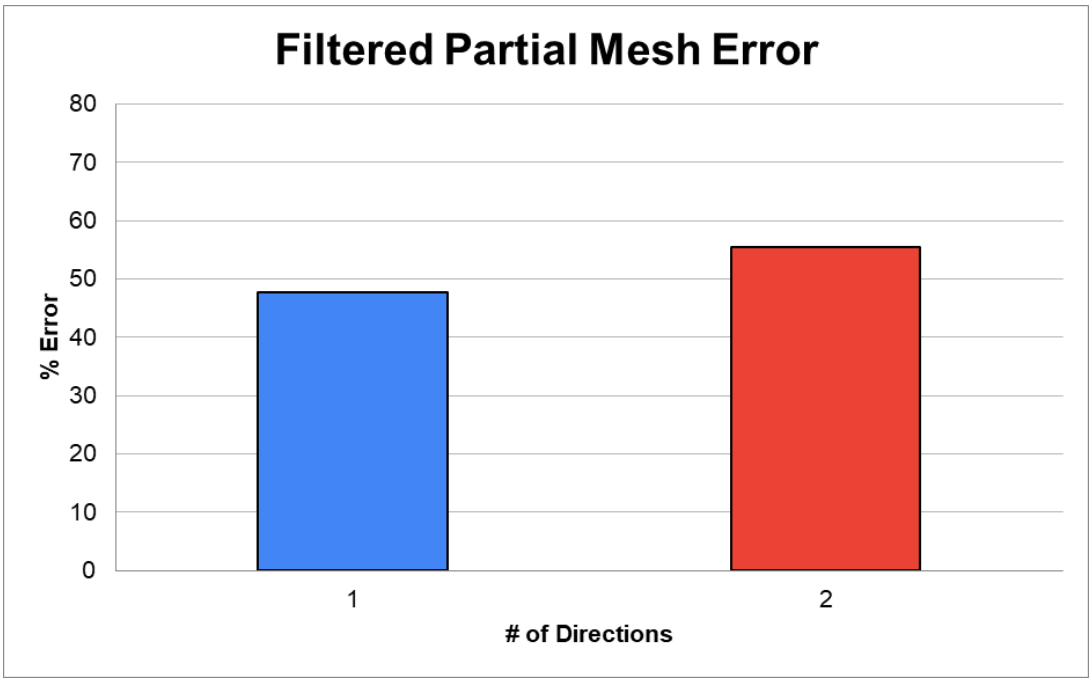

FIGURE 6-13: Filtered error for Experiment 3 partial mesh inspection

While the pattern exhibited by this data set is very similar to that seen in the data from Experiment 2, there are a couple of key differences. Compared to Experiment 2, the difference between the two bars is more pronounced, with 7.63% more error compared to only 2.63% more in Experiment 2. This difference becomes even more visible when broken down by defect in FIGURE 6-14.

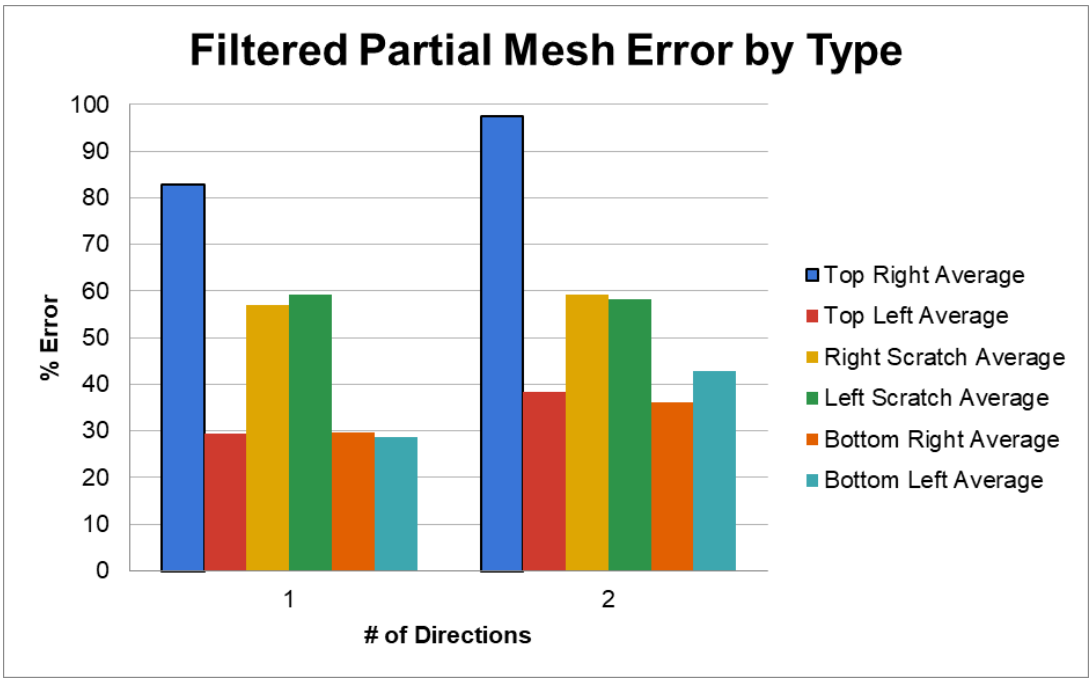

FIGURE 6-14: Filtered error for Experiment 3 partial mesh inspection by defect type

The largest difference existed at the top right defect again, but there were several other defects with greater than a 5% difference, such as the top left and bottom right dents. While the scratches reported very similarly across both trials, there was a notable increase in error for all the dents. A similar pattern can be seen in the candidate true/false positive data in FIGURE 6-15 and FIGURE 6-16.

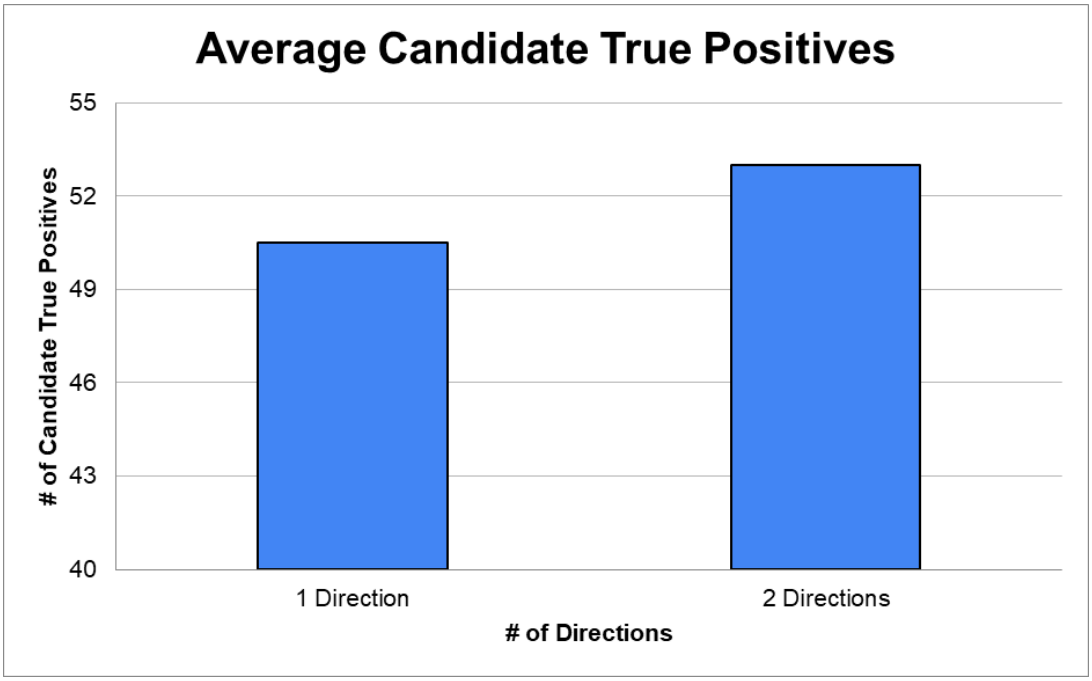

FIGURE 6-15: Average candidate true positives for Experiment 3 full mesh inspection

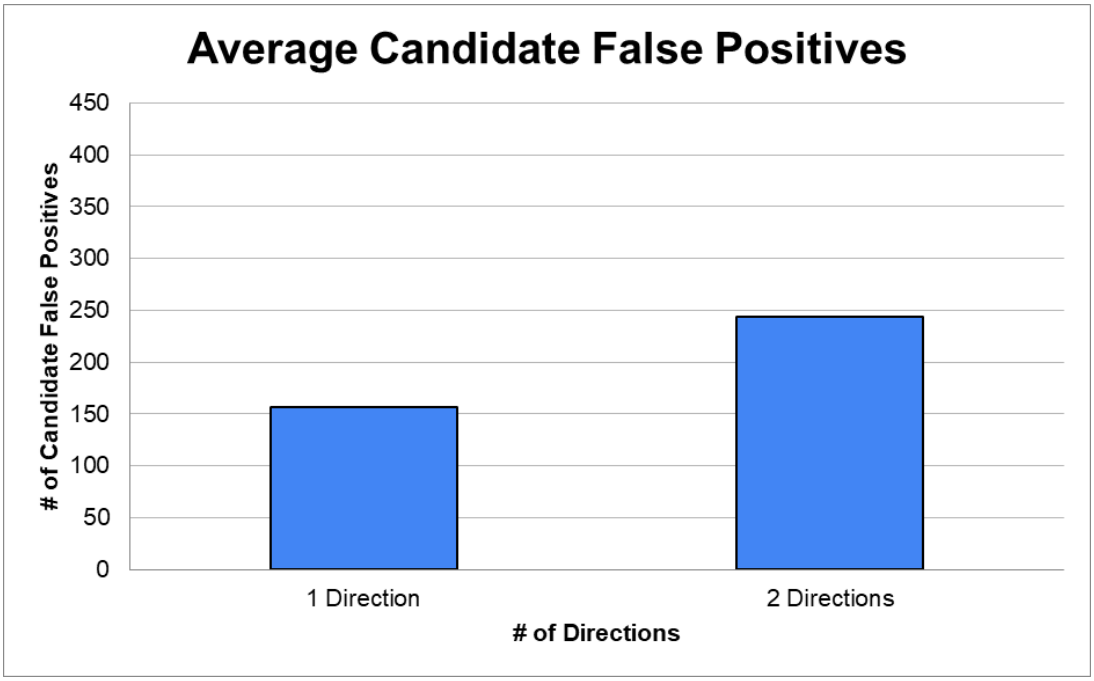

FIGURE 6-16: Average candidate false positives for Experiment 3 full mesh inspection

Once again, the overall dataset is very similar to that of Experiment 2. However, the trends observed again seem more definitive in the Experiment 3 data. While only one trial recorded less than 53 candidate true positives, it was noted by the operator that reducing the direction parameter from 2 to 1 significantly reduced the number of defects detected for both candidate true and false positives. This is especially evident in FIGURE 6-16, as not only did 1 direction average 86.9 less candidate false positives per trial, but there was not a single trial in which the detected candidate defects were lower for 1 direction when compared to 2 directions. If the surface defect map tool is struggling to pick up all of the expected defects, it may make sense to increase this parameter to 2 directions. Otherwise, 1 direction is the clear choice to both reduce both error and recorded candidate false positives.

The partial mesh data for the type parameter, both unfiltered and filtered through a 2-sigma filter, can be seen below in FIGURE 6-17 and FIGURE 6-18. The DB column in this dataset refers to the option that checks for both depressions and bulges.

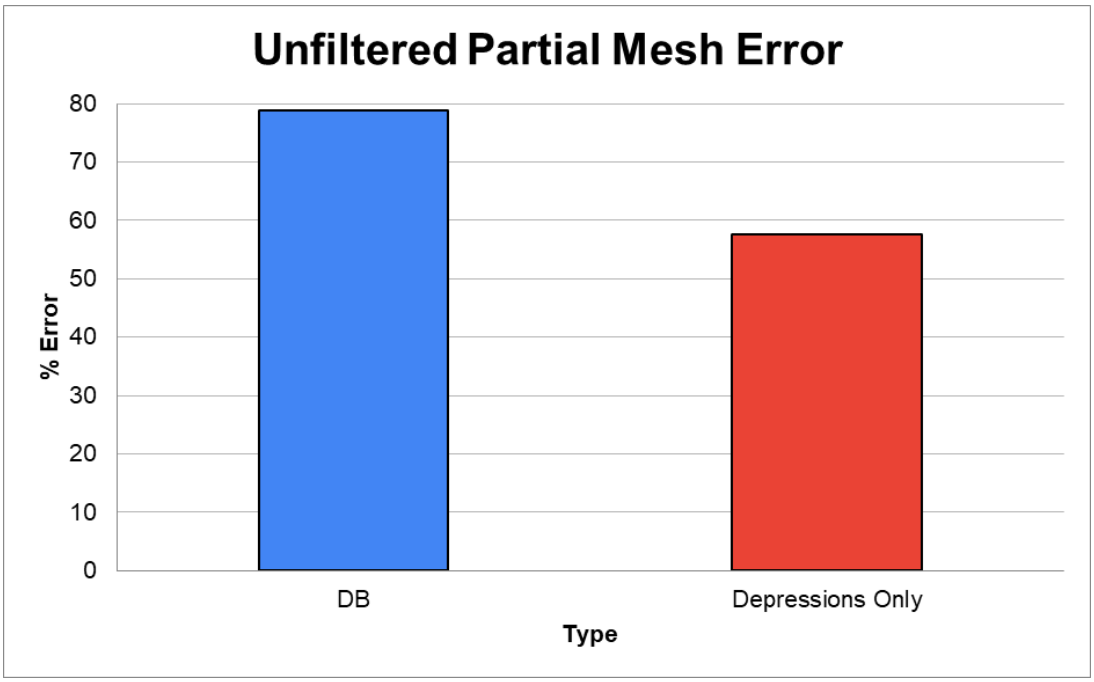

FIGURE 6-17: Unfiltered error for Experiment 4 partial mesh inspection

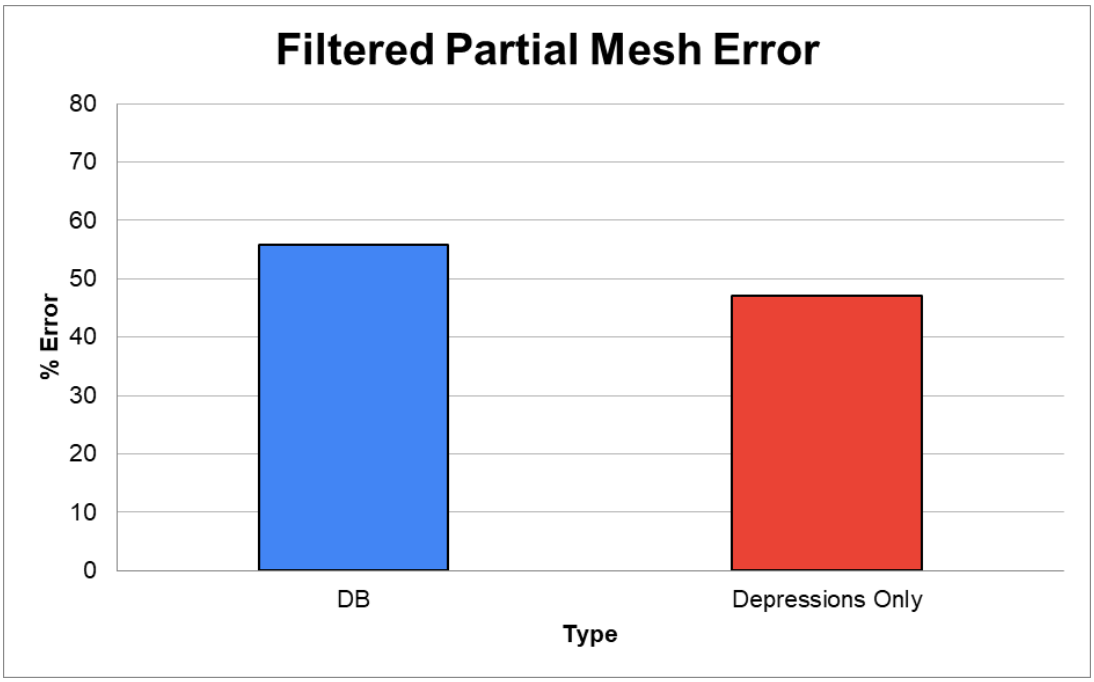

FIGURE 6-18: Filtered error for Experiment 4 partial mesh inspection

Looking at this chart, one of the biggest discrepancies between the unfiltered and filtered datasets emerges, with the 2-sigma filter removing 23% error from the DB average. Breaking the dataset down by defect in FIGURE 6-19 further exposes this trend.

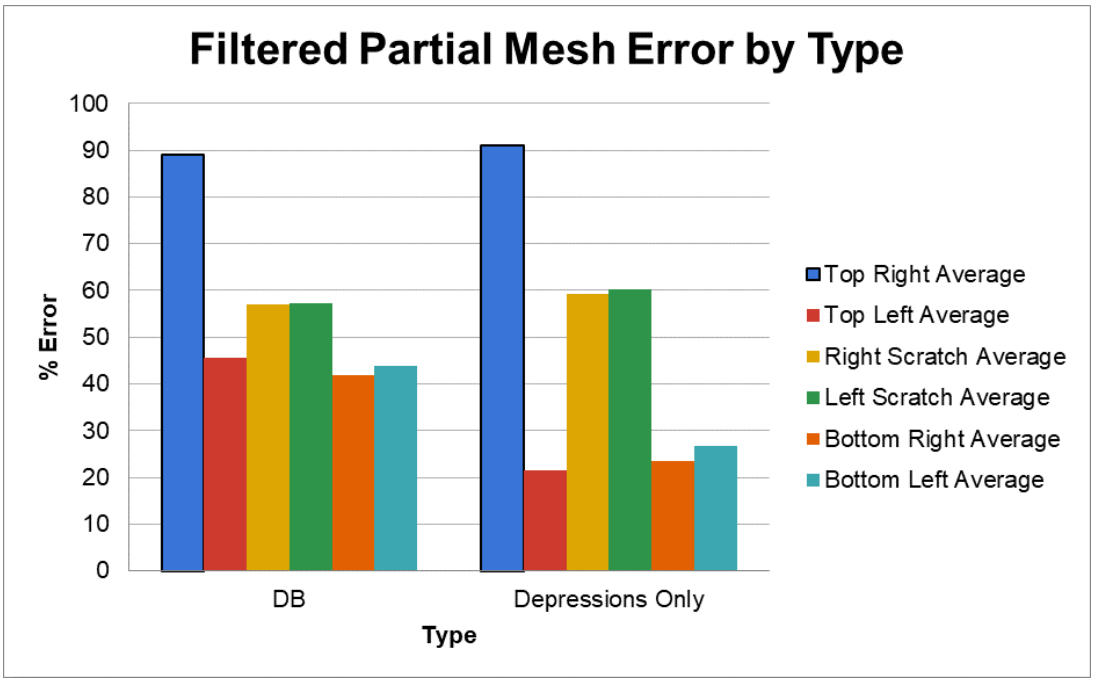

FIGURE 6-19: Filtered error for Experiment 4 partial mesh inspection sorted by defect type

The scratch trials are close for both options, but the parameters that are not close are separated by almost 20% error. Across the conducted experiments, it is noticeable that the defects that register the most error are the scratches and the top right dent. This make sense for the scratches, as their geometry provides many different edges for the defect map to interpret in different ways. The only observed difference between the top right defect and the other defects is that it is the closest defect to the origin. The previously observed discrepancy between the trials that changed and the ones that didn't is even more visible in the full mesh data contained in FIGURE 6-20 and FIGURE 6-21.

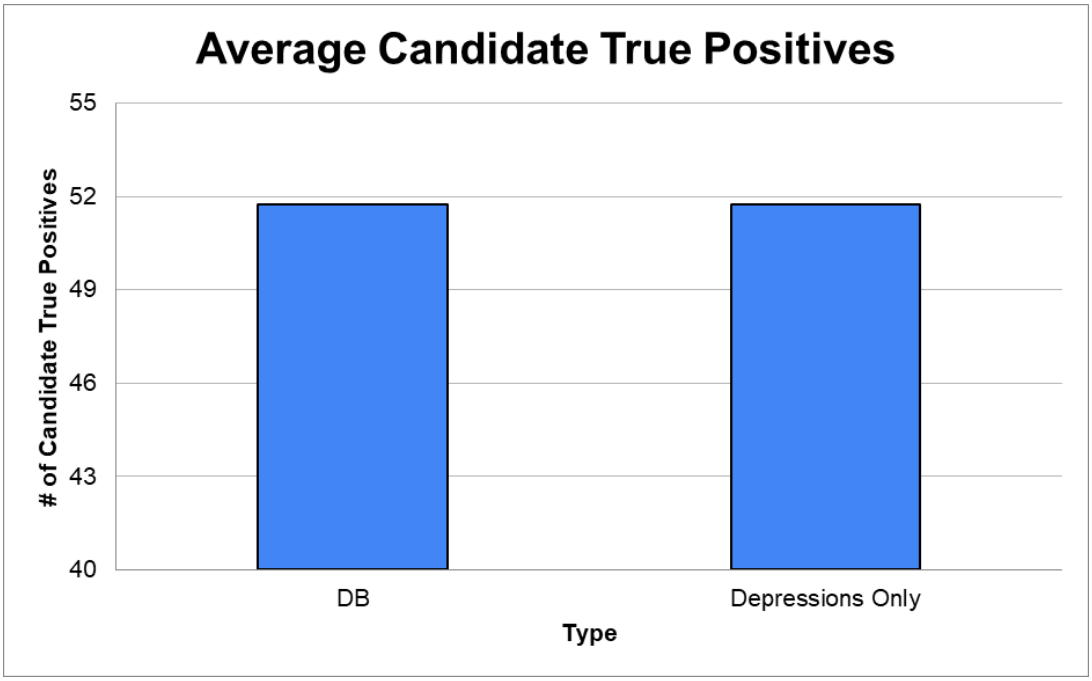

FIGURE 6-20: Average candidate true positives for Experiment 4 full mesh inspection

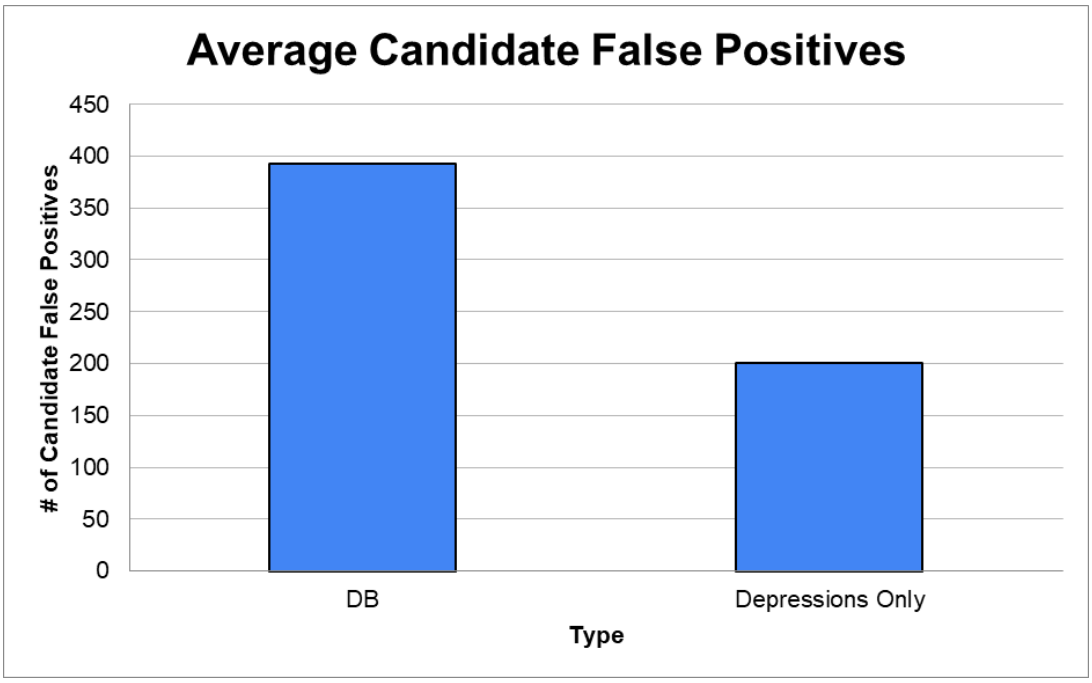

FIGURE 6-21: Average candidate false positives for Experiment 4 full mesh inspection

The full mesh inspection was not subject to any filters, and in this dataset both options registered the same number of true positives. However, when comparing the candidate false positives the DB option registered almost 100% more, on average containing 392.95 compared to only 200.3 for the depressions only option.

6.5 Experiment 5

The partial mesh data for the orientation experiment, both unfiltered and filtered through a 2-sigma filter, can be seen below in FIGURE 6-22 and FIGURE 6-23. Due to the number of permutations needed to conduct this experiment, the max defect size line was standardized at 2.5 mm for all trials.

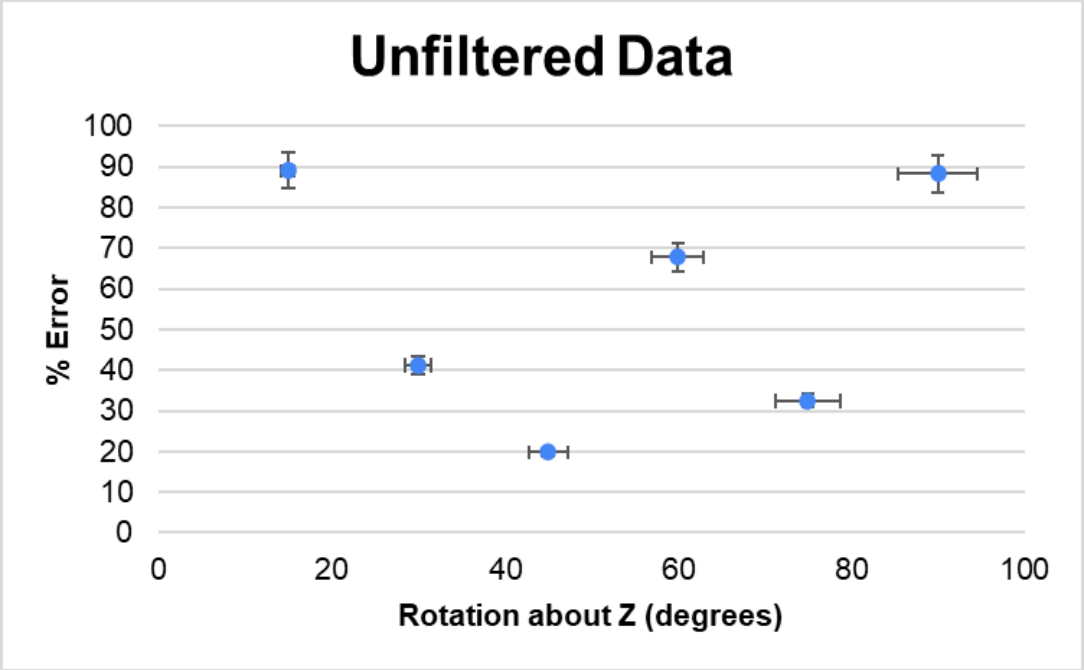

FIGURE 6-22: Unfiltered error by orientation for Experiment 5 partial mesh inspection

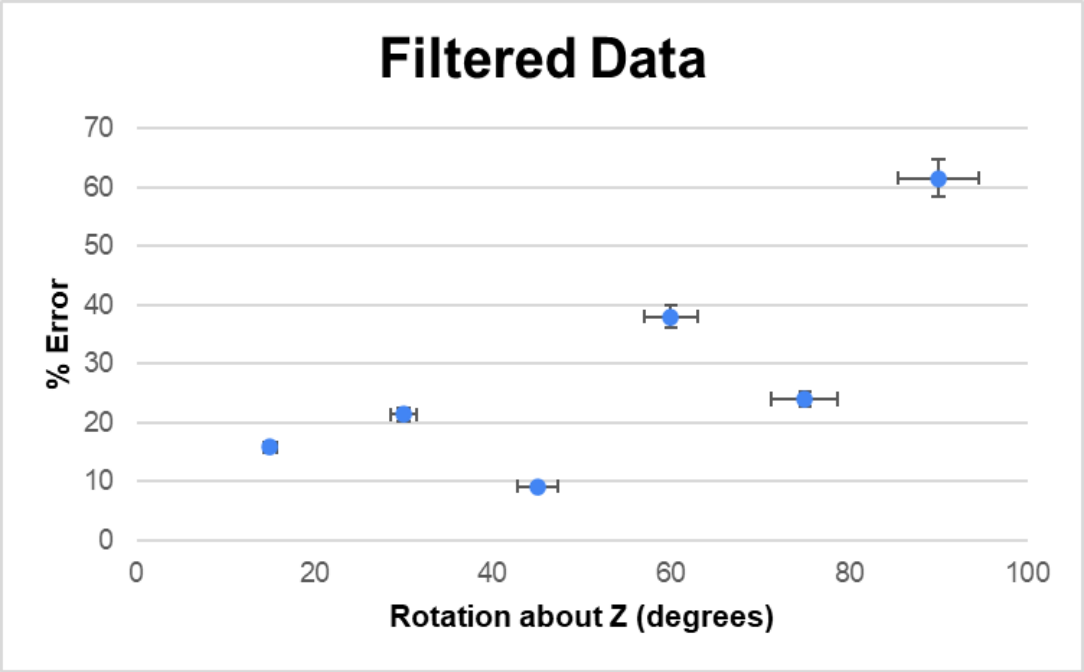

FIGURE 6-23: Filtered error by orientation for Experiment 5 partial mesh inspection

The datapoint at 15 degrees contained the most outliers, with 74% error removed between the filtered and unfiltered dataset. This dataset averaged the results from both the XYZ directions and number of directions parameter, so choice of these parameters did not affect the results.

The filtered data broken down by defect is visible in FIGURE 6-24.

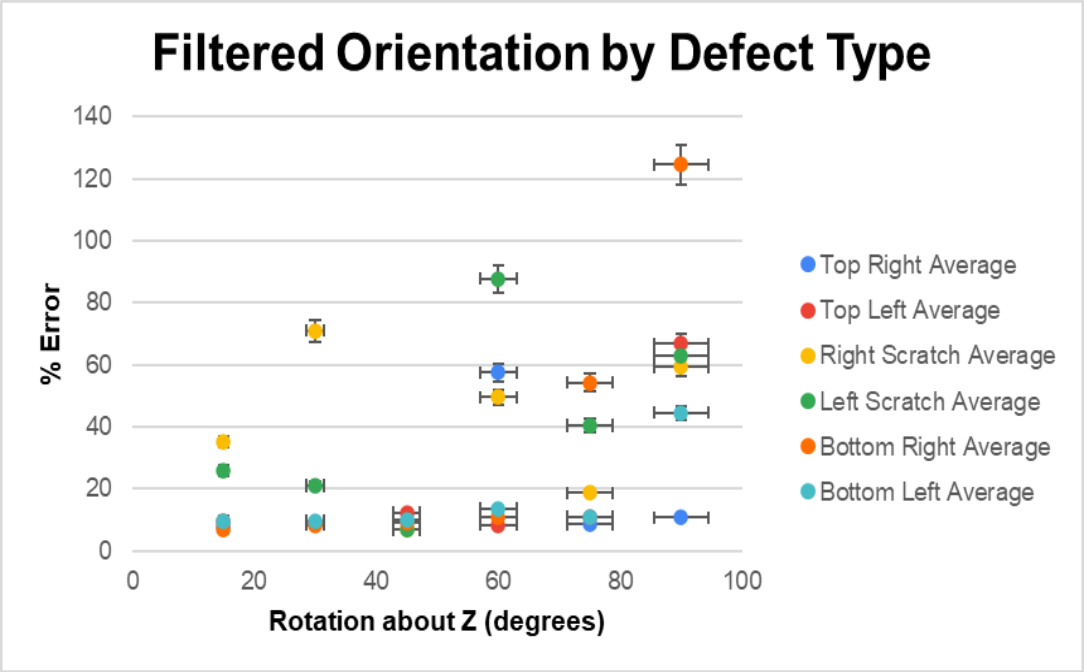

FIGURE 6-24: Filtered error sorted by defect type for Experiment 5 partial mesh inspection

The data clearly indicates that the rotation to 90 degrees produces the largest jump in error, with 4 of the 6 defects examined possessing at or over 60% error when compared to the mesh. There did seem to be some correlation between rotation and error, as the right scratch exhibited the most error at 30 degrees, and the left scratch exhibited the most error at 60 degrees. This pattern was exhibited to a lesser extent in the dents as well, as the bottom left dent exhibited the most error at 30 degrees and the top right dent displayed the most error at 60 degrees. The scratches also displayed the most error at the 30, 60, and 90 degree locations, which is most likely due to the angled geometry of the scratches compared to the round dents. To check for noticeable differences by direction, the orientation data was isolated to just X and Y in FIGURE 6-25.

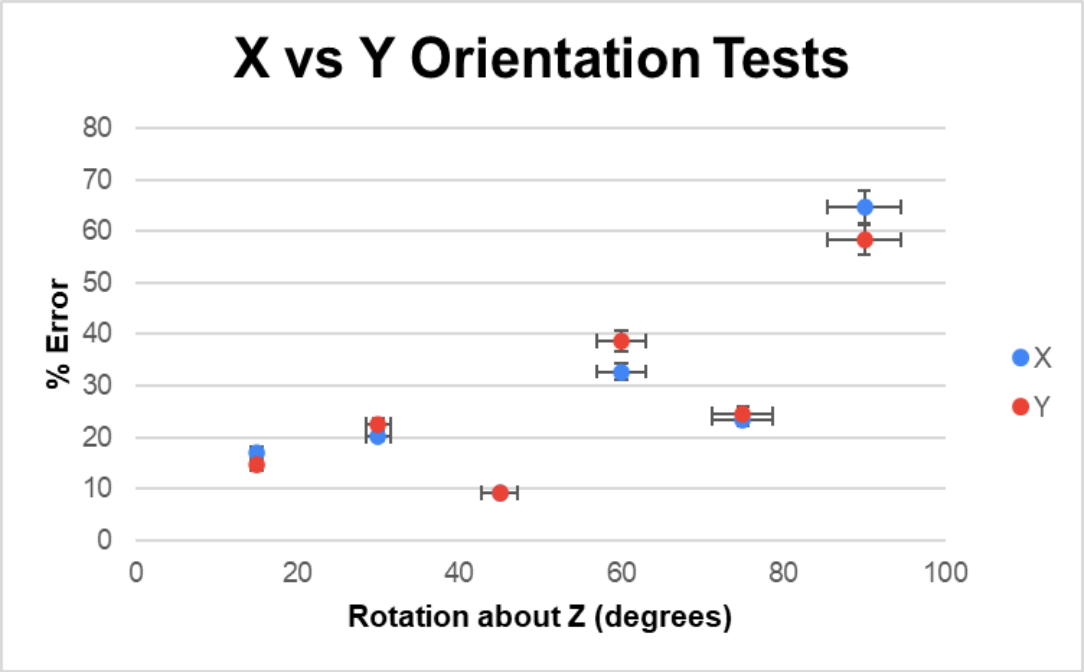

FIGURE 6-25: Filtered error by orientation isolated to just X and Y for Experiment 5 partial mesh inspection

No trend is visible looking at this figure, with the only deviations of more than 5% occurring at 60 and 90 degrees. This suggests any difference between the two cannot be attributed to the direction of the line even with the change in orientation. The observed trend for the number of directions parameter was also consistent, as visible in FIGURE 6-26.

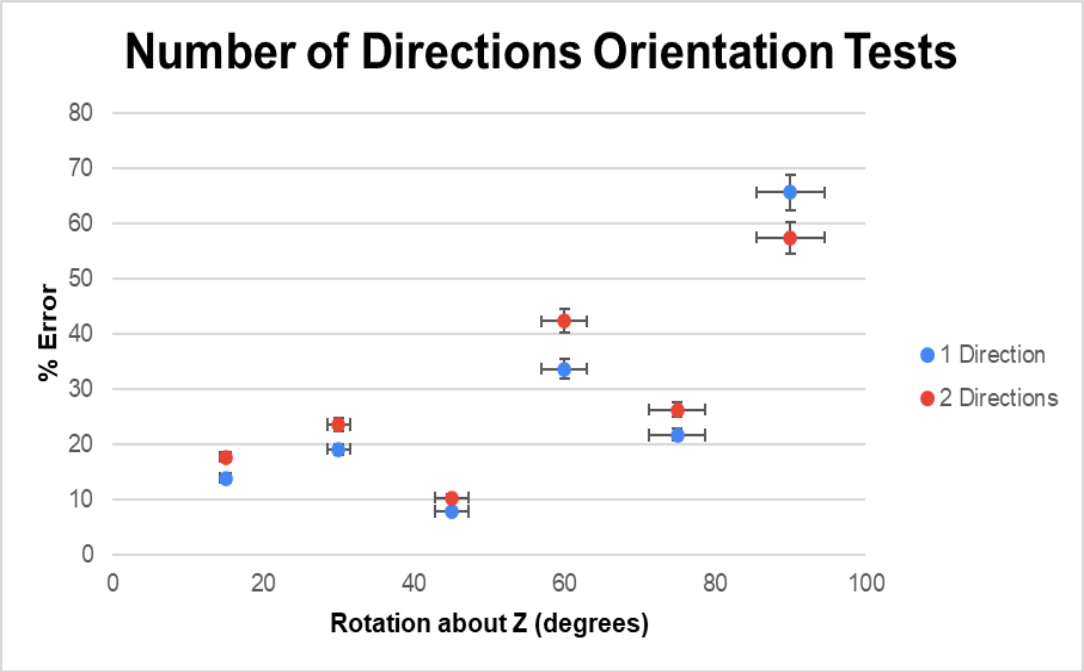

FIGURE 6-26: Filtered error by orientation isolated to just number of directions for Experiment 5 partial mesh inspection

The one direction option performed better than the two directions option for all orientations besides 90 degrees. Figures for the full mesh inspections are visible in FIGURE 6-27 and FIGURE 6-28.

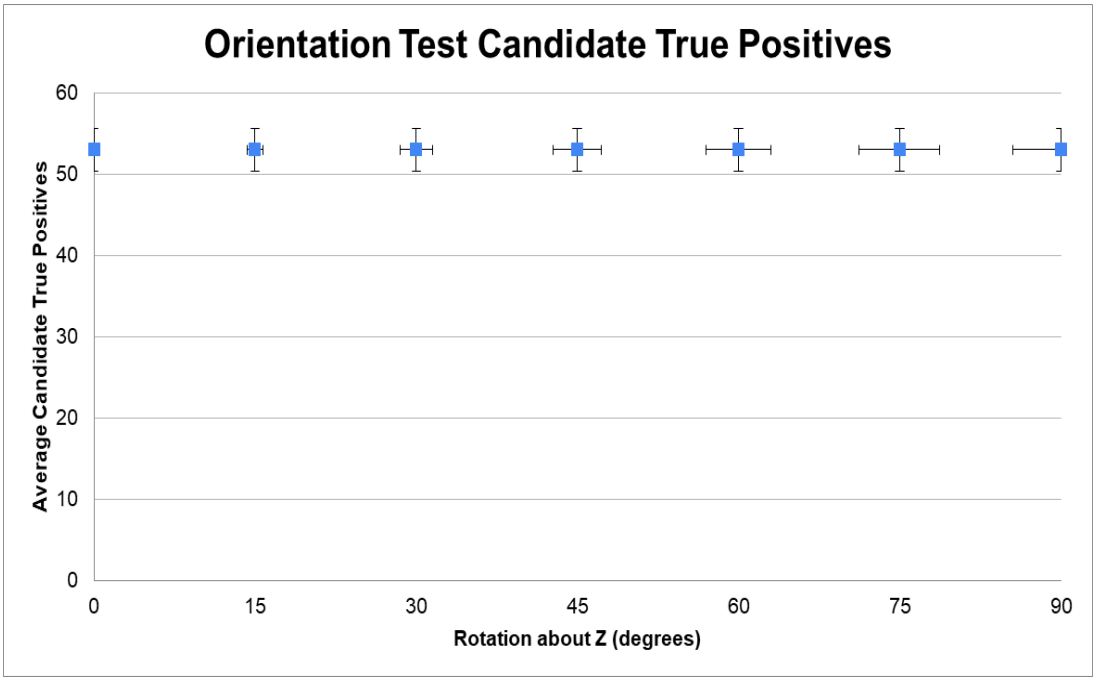

FIGURE 6-27: Average candidate true positives for Experiment 5 full mesh inspection

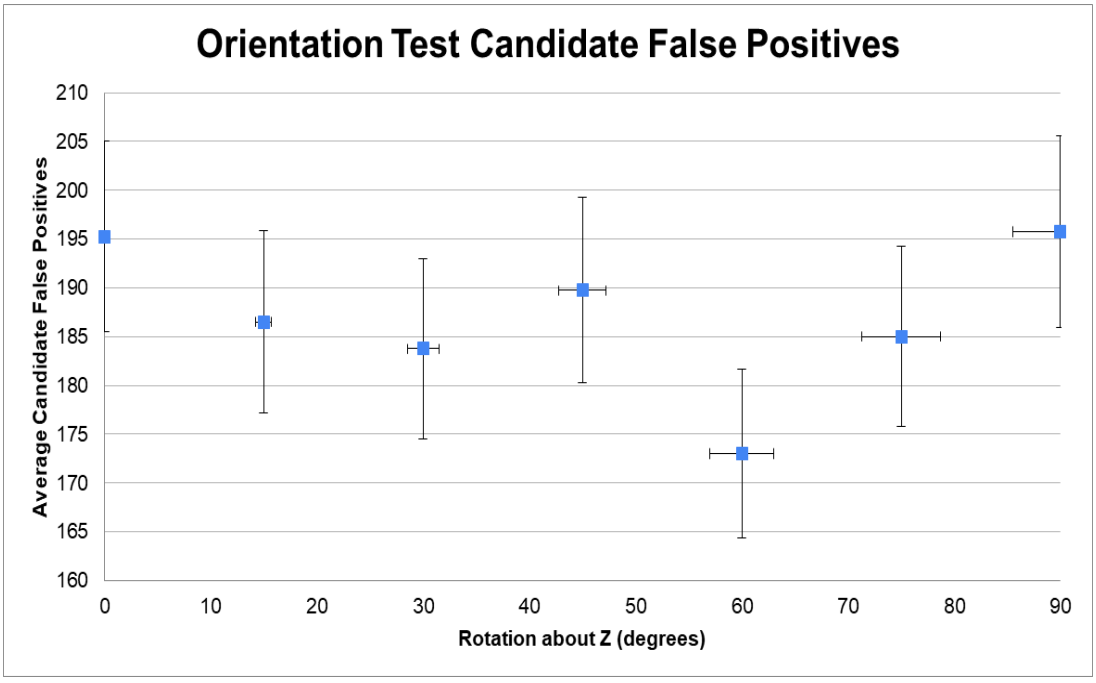

FIGURE 6-28: Average candidate false positives for Experiment 5 full mesh inspection

There was no difference in the candidate true positives recorded at any given orientation. The candidate false positives were also fairly consistent, with the biggest difference between any orientation existing at 60 degrees. Besides the outlier at 60 degrees, the positions at 0 and 90 degrees represented the most candidate false positives, which is consistent with the findings of the partial inspection. Dividing the data up by direction in FIGURE 6-29 reveals an interesting trend.

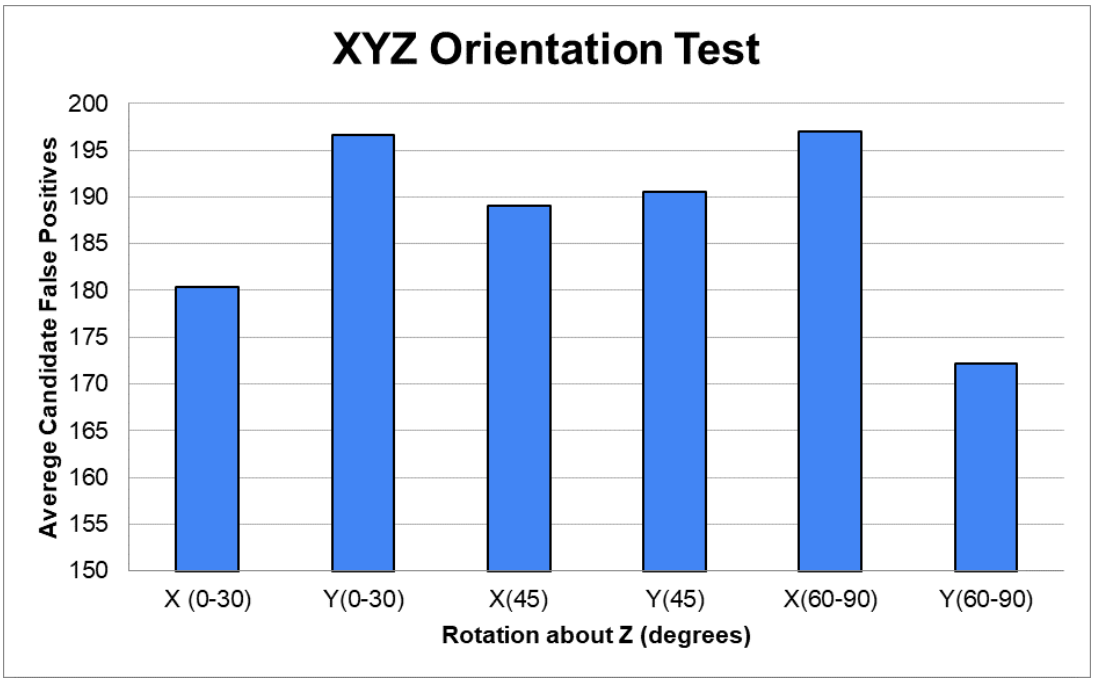

FIGURE 6-29: Average candidate false positives for Experiment 5 by XYZ direction

There is a clear difference in candidate false positives depending on what direction is used with what orientation. Performance at 45 degrees was approximately the same, with X registering significantly less defects between 0-30 degrees, and Y registering less between 60-90 degrees. While no clear difference appeared in any other

consideration of XYZ direction, it may be worth considering setting the parameter according to the orientation of the part in the software.

As with the partial mesh inspection, the trend of the 1 direction parameter performing better than 2 directions held true for the full mesh as well, as can be seen in FIGURE 6-30 and FIGURE 6-31.

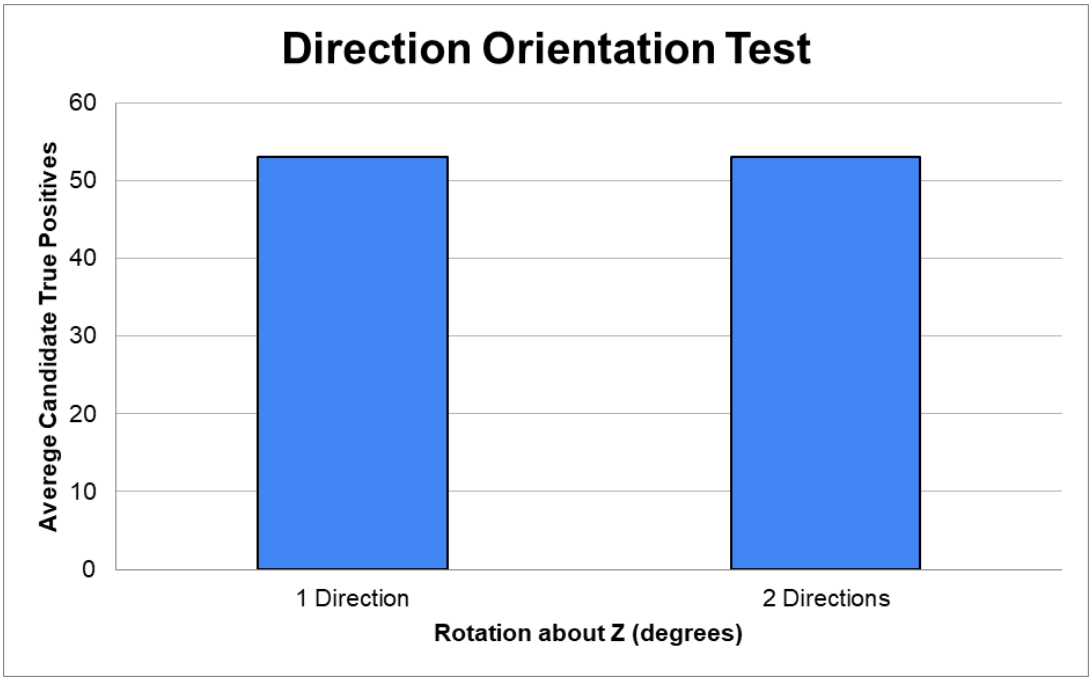

FIGURE 6-30: Average candidate true positives for Experiment 5 by number of directions

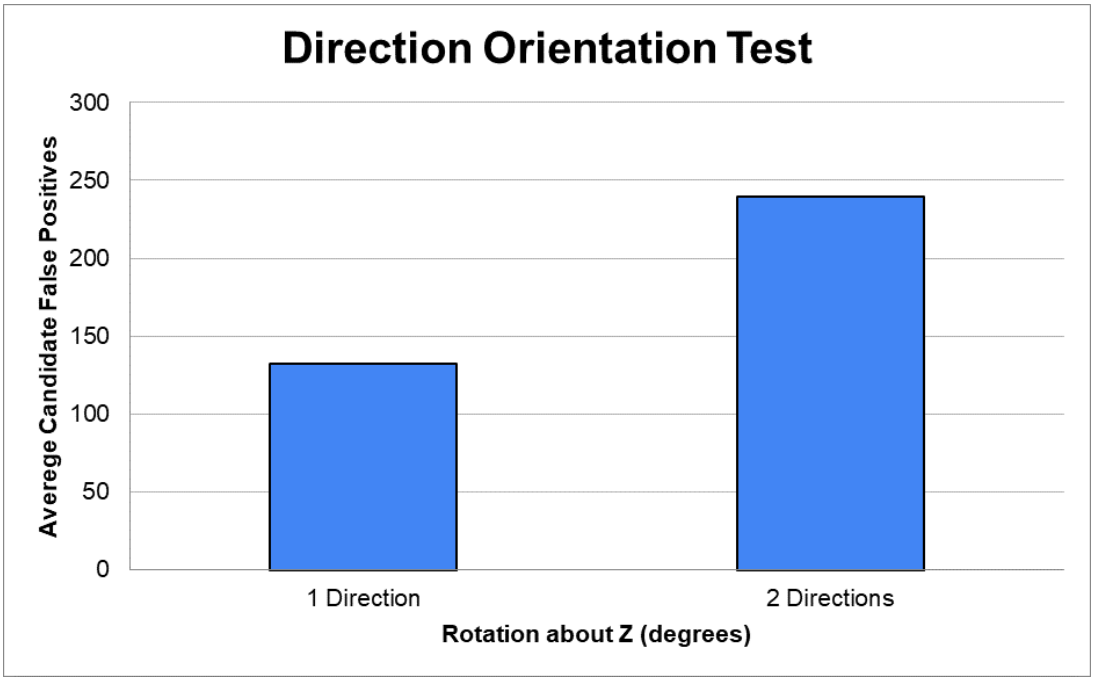

FIGURE 6-31: Average candidate false positives for Experiment 5 by number of directions

6.6 Experiment 6

The unfiltered and filtered data for the inverted mesh experiment compared to their non-inverted counterparts can be seen in FIGURE 6-32 and FIGURE 6-33. As it was in the orientation experiments, the max defect size line was set to 2.5 mm for these experiments. Note that the non-inverted data is referred to as the standard dataset in the figures for this section

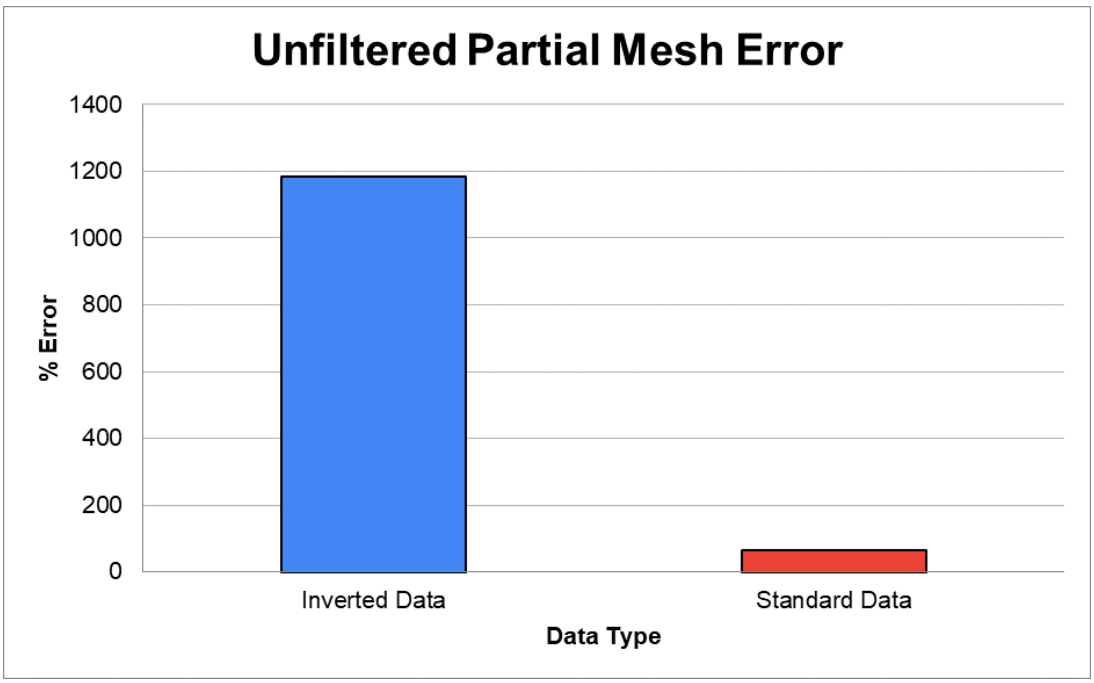

FIGURE 6-32: Unfiltered error for Experiment 6 partial mesh inspection

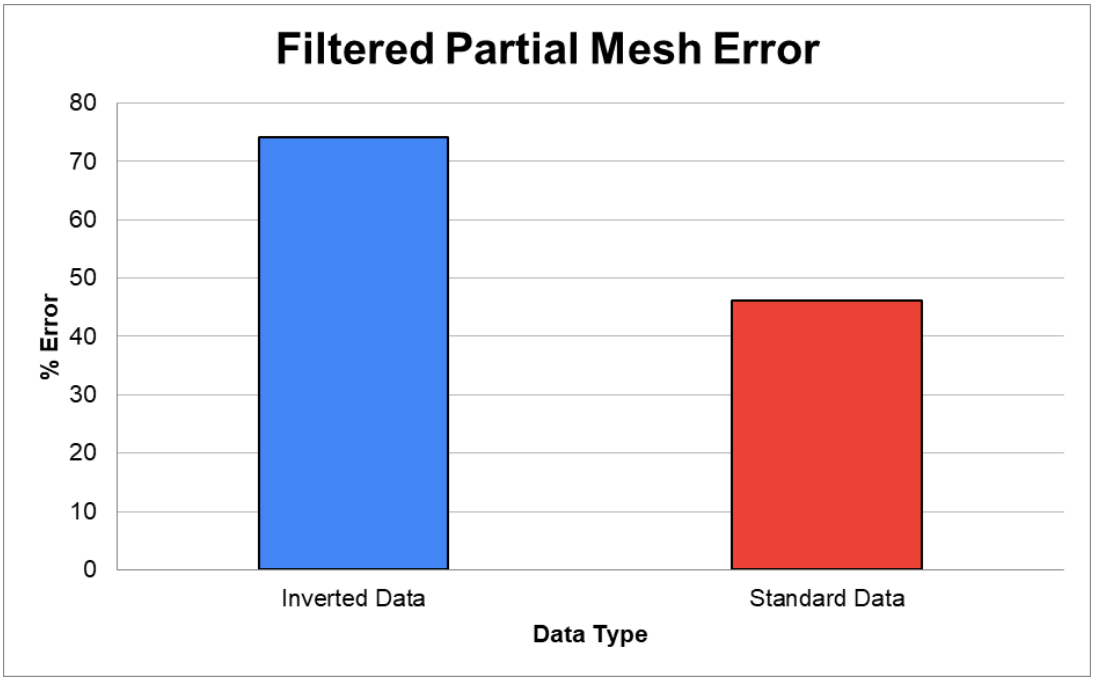

FIGURE 6-33: Filtered error for Experiment 6 partial mesh inspection

The difference between the two datasets is enormous, with the unfiltered inverted data registering almost 1200% error. Even with the significant outliers filtered out, there was almost a 28% difference in error between the inverted and standard data. Dividing the dataset up by defect in FIGURE 6-34 makes the source of the difference clearer.

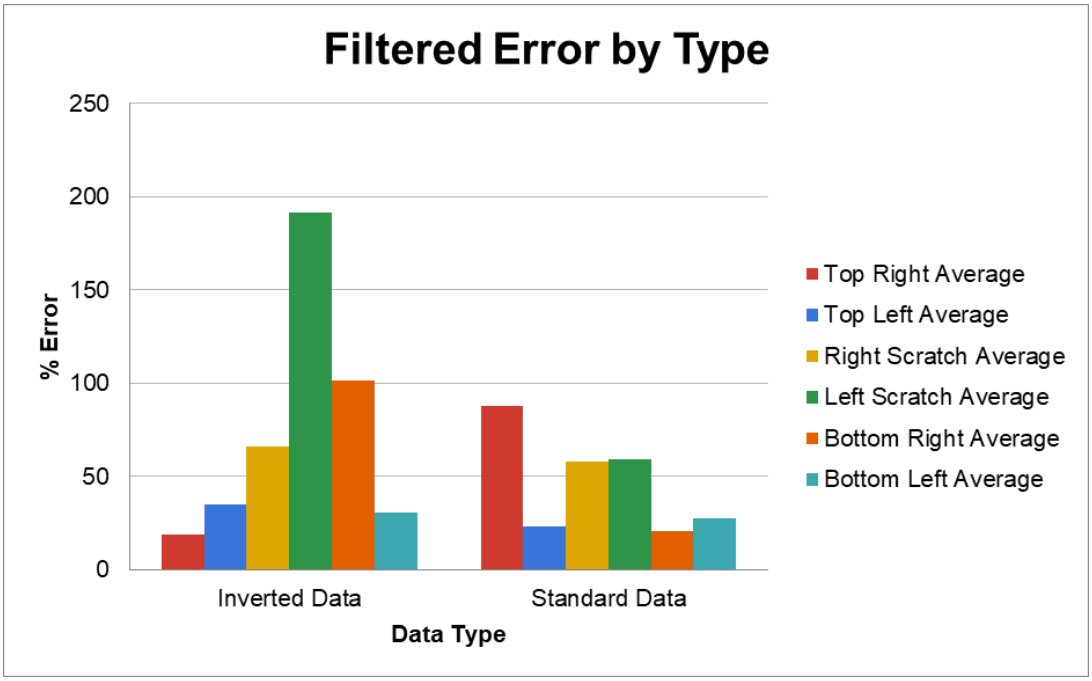

FIGURE 6-34: Filtered error for Experiment 6 partial mesh inspection by defect type

It was originally planned to include a version of this figure with the unfiltered data as well, but at 6571% the reported error for the left scratch was so big that the scale made the rest of the data unreadable. Looking at the filtered version in the figure above, the left scratch still represents a clear outlier. However, the right scratch and bottom left dent are the only defects even within 10% of the standard data, with the bottom right dent representing the next clearest outlier at almost 80% difference in error between the datasets. Interestingly, the one defect the inverted data performed better on than that standard data was the top right dent, which was the defect that recorded the most error in most of the other conducted experiments.

Looking at the XYZ directional parameter and number of directions parameter, the previously drawn conclusions are also present in this dataset despite significantly

more error than the standard data. These charts can be seen in FIGURE 6-35 and FIGURE 6-36.

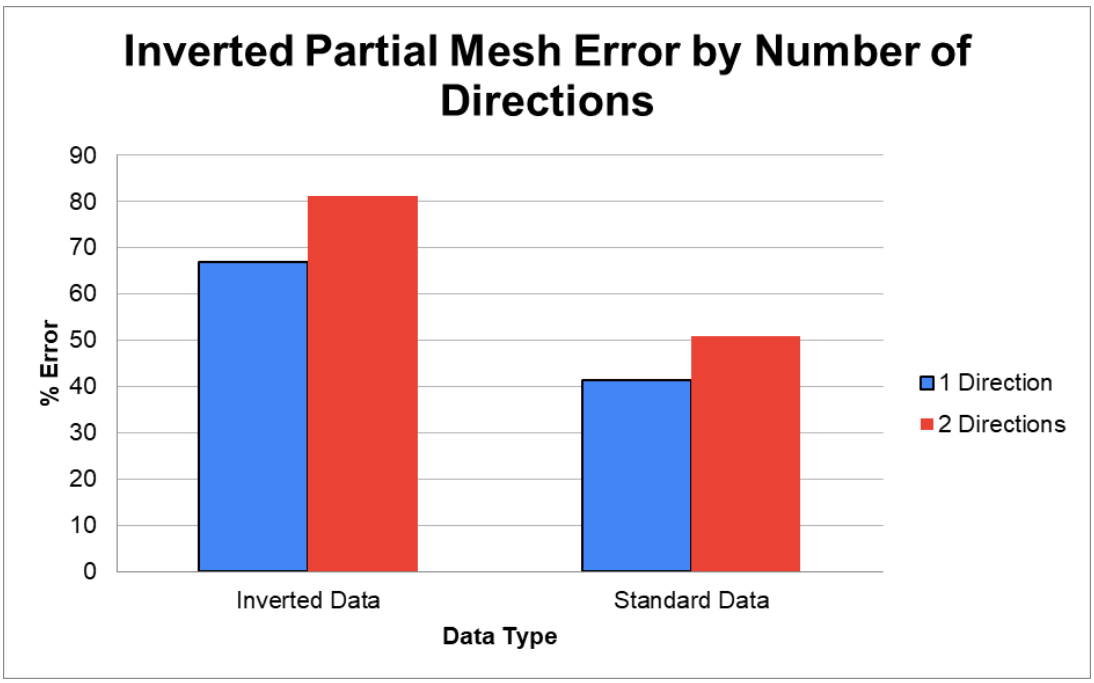

FIGURE 6-35: Filtered error for Experiment 6 partial mesh inspection by number of directions

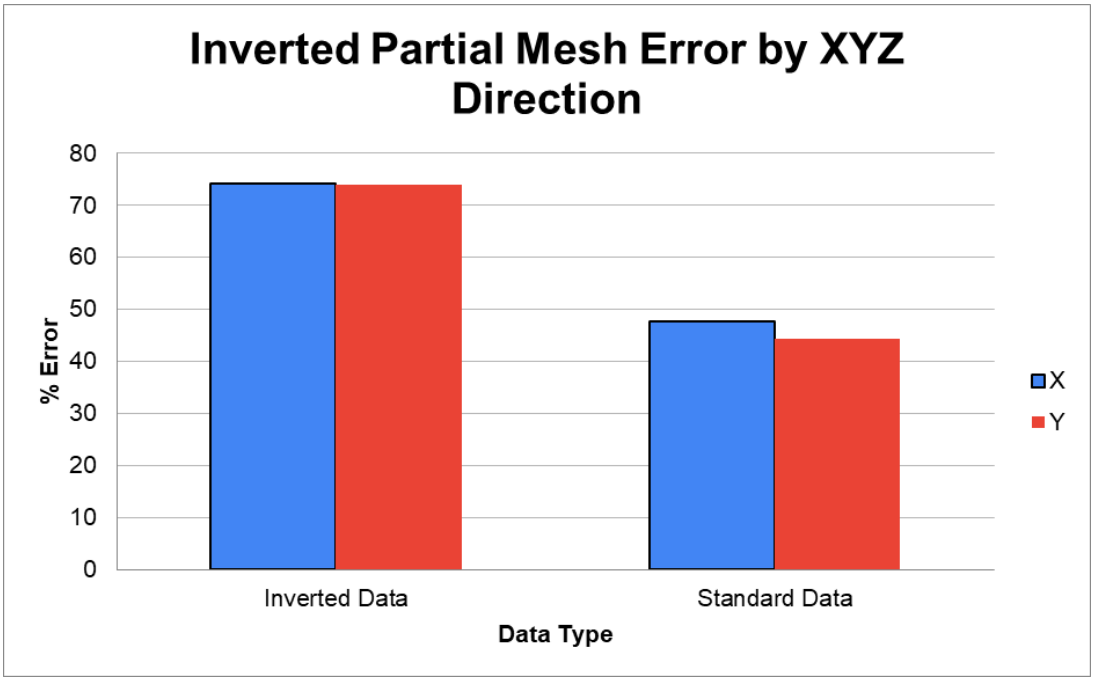

FIGURE 6-36: Filtered error for Experiment 6 partial mesh inspection by XYZ direction

FIGURE 6-37 shows the candidate true and false positives for the full inverted mesh compared to the non-inverted data.

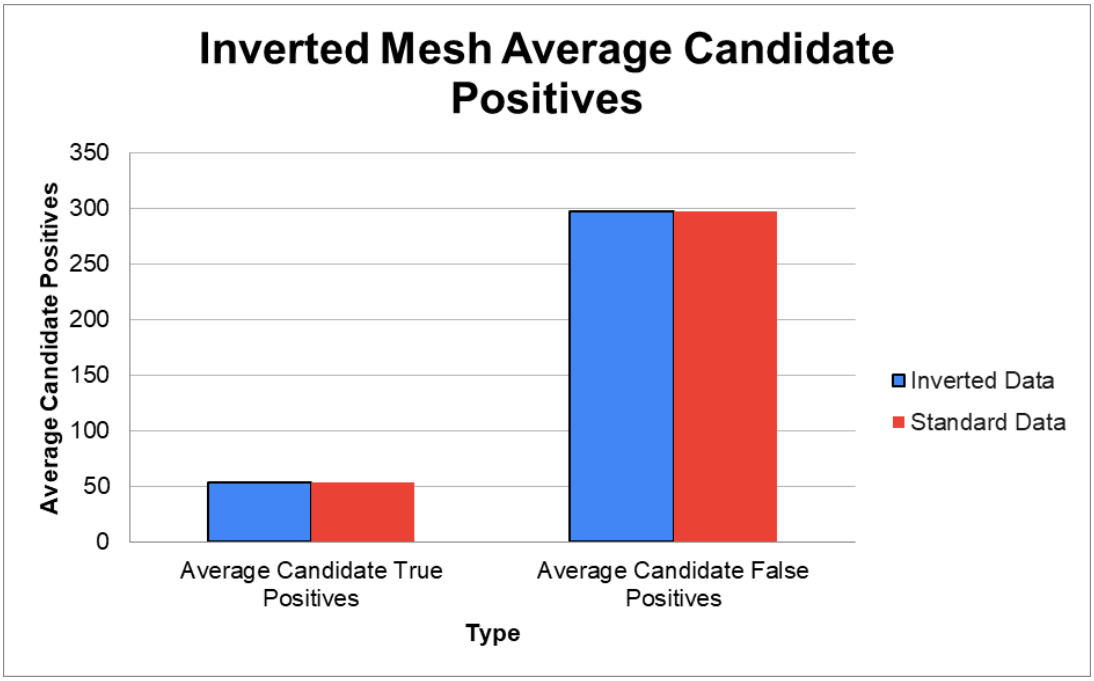

FIGURE 6-37: Candidate true/false positives for inverted mesh compared to non-inverted dataset

The large jump in error seen in the partial mesh inspection is not seen here, as both the candidate true and false positive numbers are almost identical to that of the standard dataset. This holds true for the number of directions and XYZ data in FIGURE 6-38 and FIGURE 6-39 as well. Note that the candidate true positives for these trials were the exact same as the values in FIGURE 6-37, so the values were excluded to promote visual clarity.

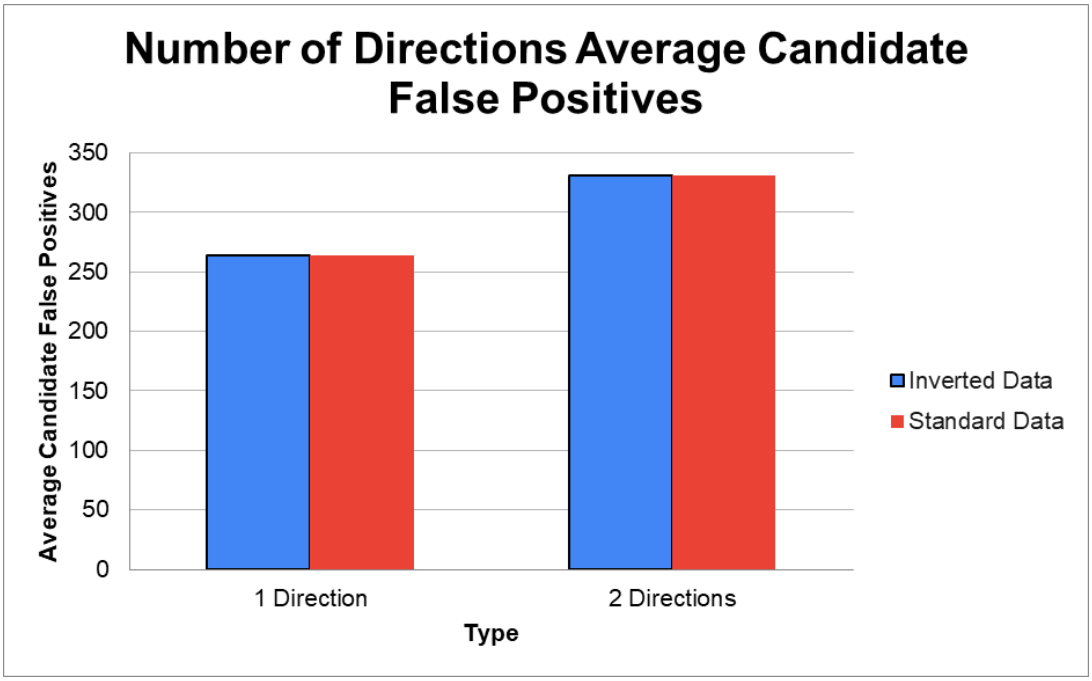

FIGURE 6-38: Candidate false positives for inverted mesh compared to standard dataset for directions parameter

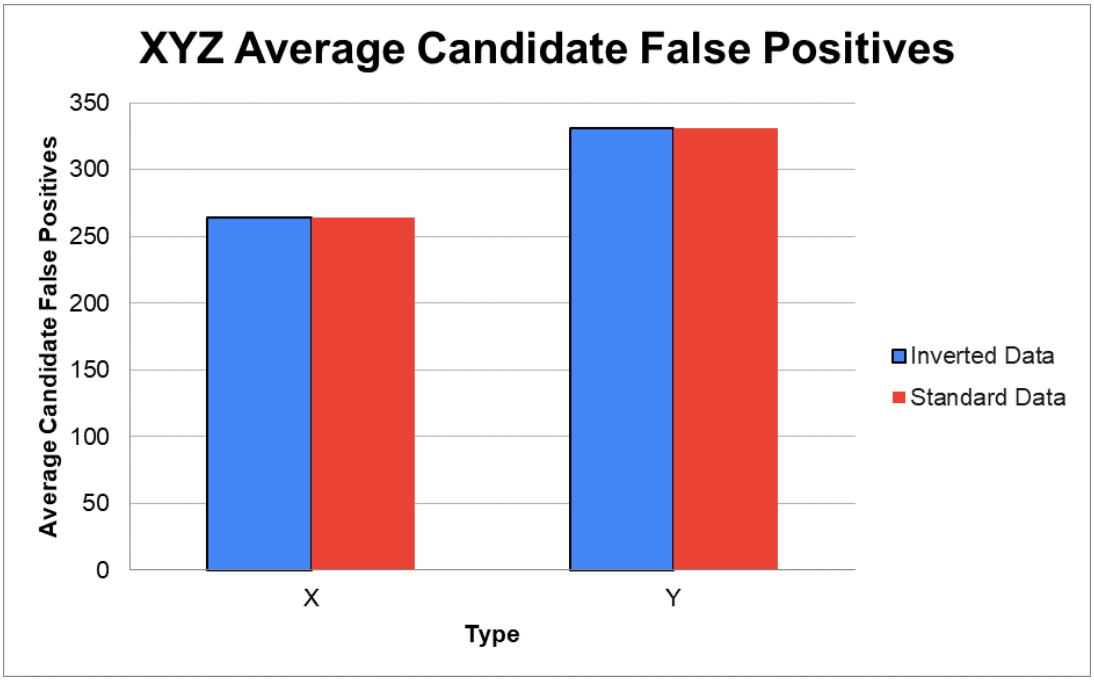

FIGURE 6-39: Candidate false positives for inverted mesh compared to standard dataset for XYZ direction

6.7 Correlation between Variables

After the results for percent error were collected, the Pearson correlation value was calculated for each of the four main variables. Linear regression was then used to generate a t-statistic and a p-value, and these values were then displayed below in Table 2.

Table 2: Percent Error Regression Statistics

Correlation type	T-Statistic	P-Value
Max Length/Type	20.3	≈ 0
Max Length/Number of Directions	98.4	≈ 0
Max Length/XYZ Direction	11.5	≈ 0
Type/Number of Directions	11.0	≈ 0
Type/XYZ Direction	23.5	≈ 0
Number of Directions/XYZ Direction	11.6	≈ 0

As observed in the table, there is no observable correlation between any of the variables tested for percent error. The dataset tested used only the calculated optimal parameters, but other tests that examined the other states of the variables similarly concluded that no correlation existed.

Correlation was also tested between the type, number of directions, and XYZ direction for the false positives in Table 3.

Table 3: False Positives Regression Statistics

Correlation type	T-Statistic	P-Value
Type/Number of Directions	3.09	0.091
Type/XYZ Direction	2.26	0.152
Number of Directions/XYZ Direction	9.21	0.011

The table shows that a correlation between all of these variables does exist when counting false positives. The strongest correlation is between the type selected and the XYZ direction, while the smallest correlation is between the number of directions and XYZ direction parameters.

6.8 Classification

The displayed equation below shows the formula used to classify defects using the surface defect tool. This formula uses the built-in functions in ATOS Professional to assign a value to a defect by measuring its slope.

$$x = grad(\text{atan}\ ((pow(max\_depth\_height\_3sigma * -100,2))/width))$$

x represents the formula result, while max_depth_height and width are measured parameters of the defects collected automatically by the defect classification tool. $grad$, $atan$ and pow are built-in math functions. The result of this formula was placed in a table, along with other information about each defect such as length, width, and other parameters. No units are assigned to these values by the ATOS software. FIGURE 6-40 and FIGURE 6-41 plot this length and area data as histograms.

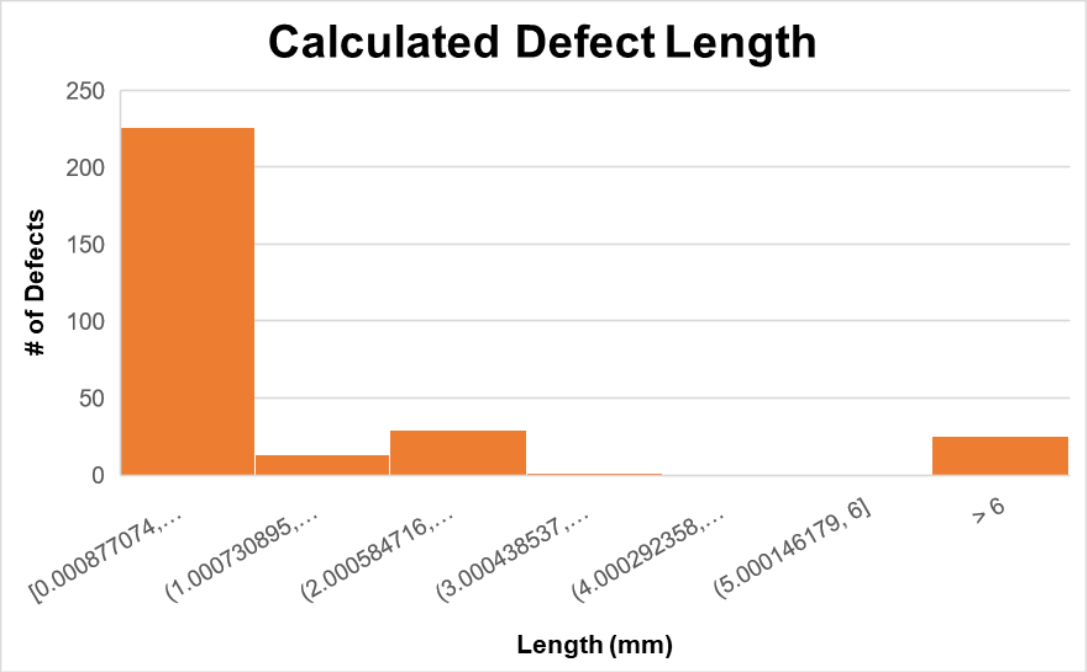

FIGURE 6-40: Histogram generated from calculated defect length

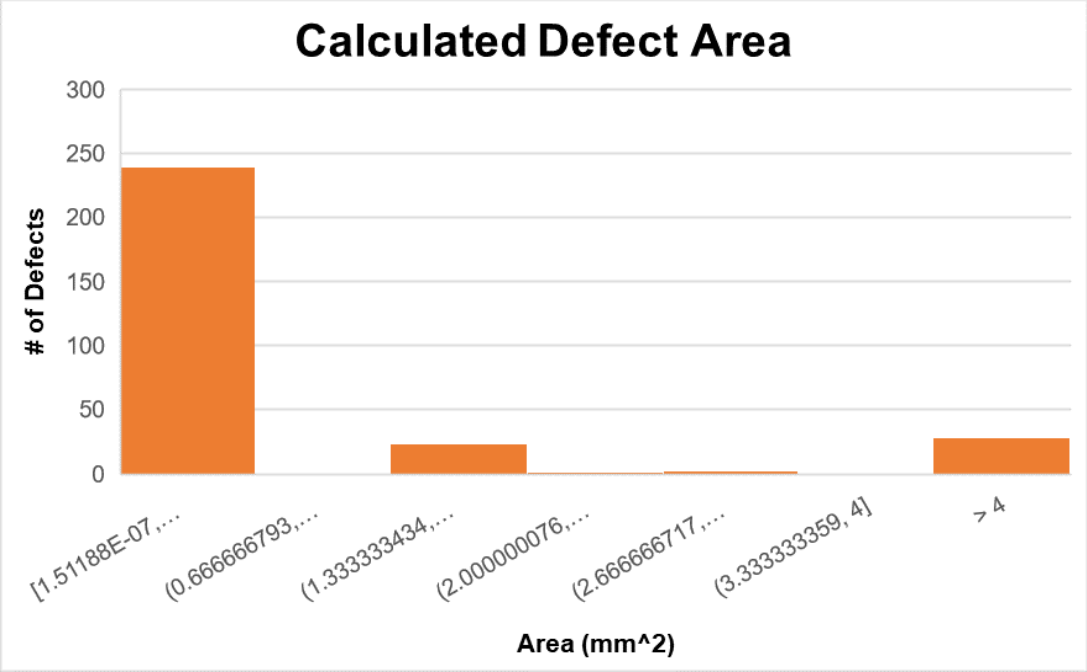

FIGURE 6-41: Histogram generated from calculated defect area

Knowing the approximate size of the defects on the artifact, filtering rules can be written using the data from these histograms. The length of the scratch defects was approximately 6.1 mm, and the area of the dent defects was approximately 4.9 mm^2. Looking at the figures, there is a clear separation of 25 defects in bin 7 for the length chart, and 28 defects in bin 7 for the area chart. This is the exact number of expected defects on the surface of the artifact, so rules were written to filter out all defects that did not meet either of these two criteria. Names were then assigned to the remaining bins to display the defects properly by type. The result of this classification is shown in FIGURE 6-42.

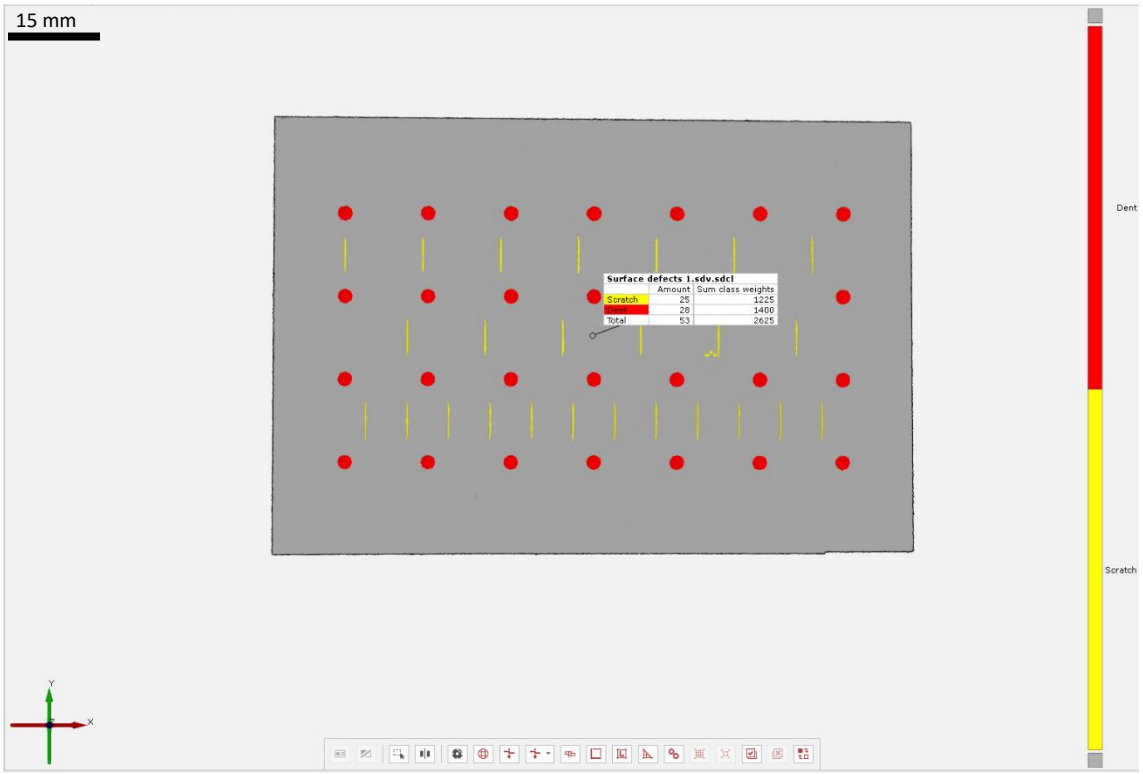

FIGURE 6-42: Surface defect classification of defect artifact 2 with filtering applied

With the filtering rules applied, all defects on the surface were correctly accounted for and classified, with the only imperfection found on a defect in the second row that the

defect map incorrectly identified. The method of setting up the defect classification by defect size proved to be very effective when applied to this ideal artifact.

6.9 Validation

The ideal parameters determined for defect artifact 2 were to set the max defect size to the width of the largest defect on the surface of the artifact, the number of directions to 1, and the type to either depressions or bulges, not both. These parameters were then applied to the other artifacts, starting with defect artifact 1.

For defect artifact 1, the 2-sigma filtered mean is plotted below in FIGURE 6-43. Note that the first 3 defects on the artifact were not picked up by the defect map, so the data is plotted from scratch 4 to 15. Labelled screenshots of all the referenced defects for the validation trials can be seen in CHAPTER 9APPENDIX BAPPENDIX B.

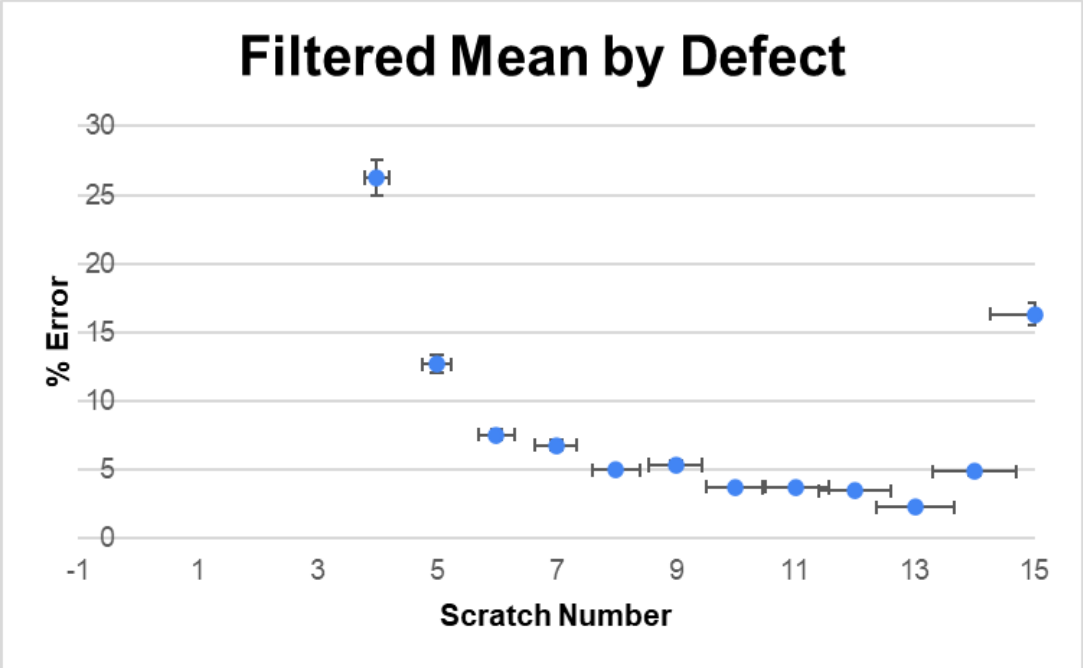

FIGURE 6-43: Filtered mean for defect artifact 1 plotted by defect

The data shows a high agreement between the mesh and deviation map between scratch 6 and 13, with a declining trend in error up until scratch 14 when it spikes back

up again. This spike can most likely be attributed to error in the method of data sampling rather than in the data though. When overlaying the points grid to measure defect deviation, the software does not allow the location, orientation, or number of points to be chosen, instead opting to lay the grid over a selected area defined by the operator and a chosen distance between each point. For defects 4-13, the software chose to lay a single row of points along the length of the scratch, while scratches 14 and 15 added a second row of points.

This inspection also encountered the first partial detection of a defect, which is something that showed up in all the validation trials. The decision was made to count a defect as fully detected if over 80% of its surface area was detected, with exceptions made if a key feature was missed or this detection came as part of a larger error. If less than 80% was detected, an estimate was made based on the sum of the parts of the defect that were detected, and it was noted that only a partial detection was achieved. The detection criteria for true positives was also updated to mean any true defect picked up by the instrument in the mesh, as the defect map cannot classify any defect that the instrument failed to pick up. Considering all of this, Table 4 contains the true and false positive data for defect artifact 1, and FIGURE 6-44 shows the final defect classification for the artifact.

Table 4: True/False Positives for Defect Artifact 1

Type	Number detected
True Positives	12
False Positives	333

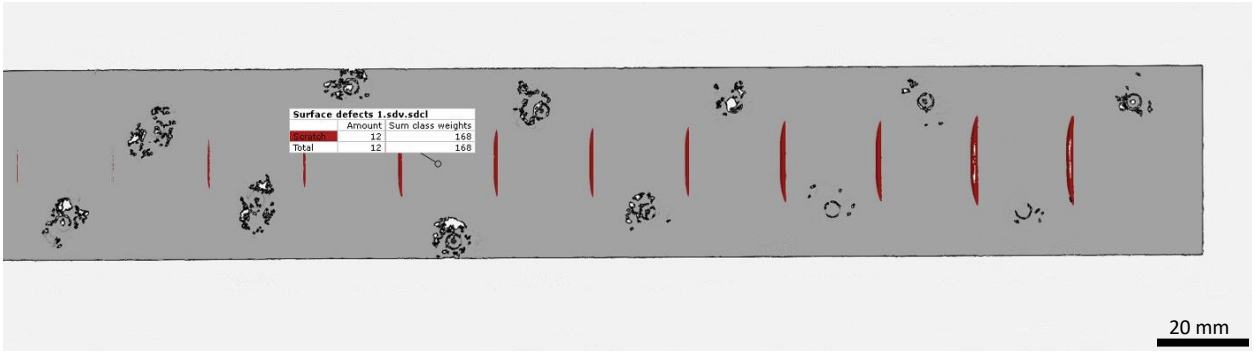

FIGURE 6-44: Surface defect classification for defect artifact 1

The software detected 12 out of 15 possible defects on the surface, and registered 333 false positives, which is in line with the results observed on defect artifact 2. The depth of the minimum detected scratch was approximately 28 microns, with none of the shallower scratches detected by the defect map. The classification process was also relatively simple due to the homogenous nature of the defects. Overall, the defect map and classification applied very well to this artifact.

For sample 2, there were issues immediately with the calculation of the defect map. FIGURE 6-45 displays the overall defect map before any analysis was applied to it, and Table 5 displays the true and false positive data for this mesh.

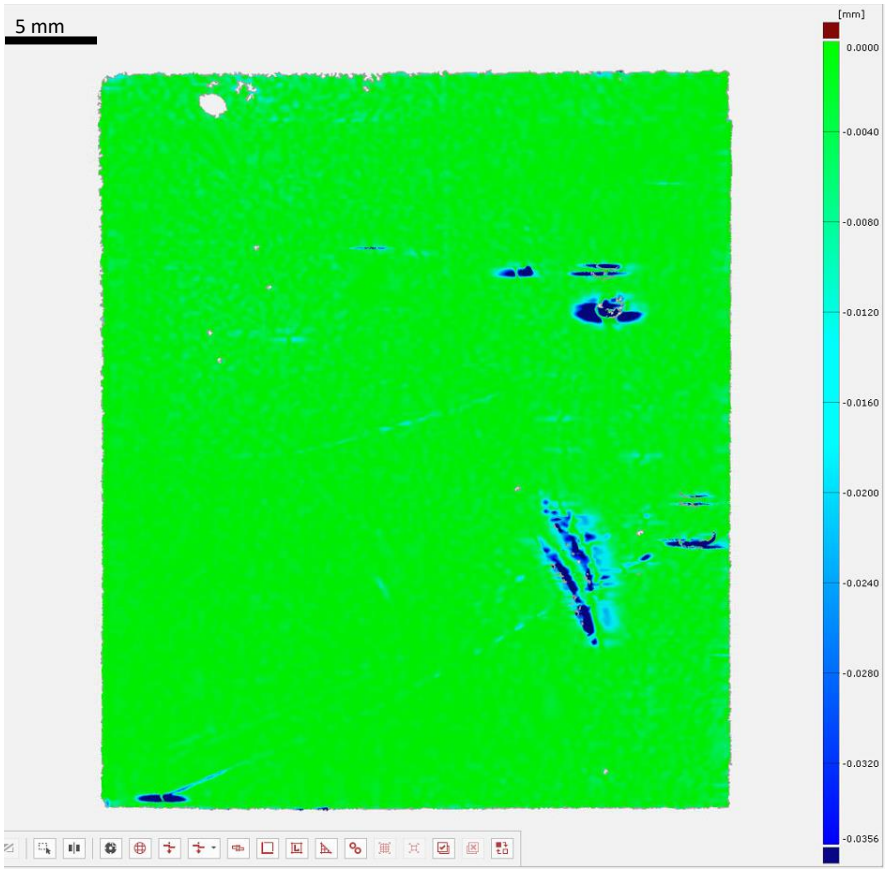

FIGURE 6-45: Surface defect map for Sample 2

Table 5: True/False Positives for Sample 2

Type	Number detected
True Positives	5.12
False Positives	78

While the map did detect 5.12 defects out of a possible 7, there were significant false positive defects surrounding the dents. The scratches accounted for the partial detections, with only 1 scratch being completely detected by the map. The mischaracterization of the defects on the surface of this artifact were mainly due to the

process that created the defects causing a raised burr of material on the edge of the dent or scratch. FIGURE 6-46 shows an example of this phenomenon, with the actual dent existing on the right side of the screenshot and the false positive on the left.

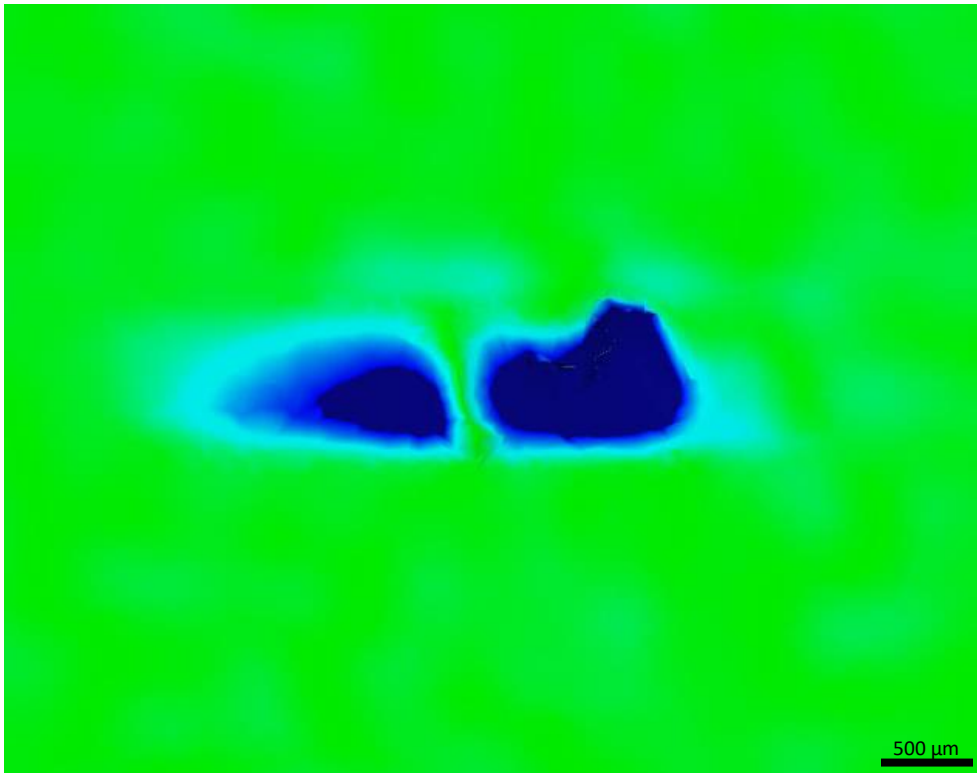

FIGURE 6-46: Example defect from the surface defect map for Sample 2

Calculating the defect map by examining the gradient of the defect causes the map to detect both sides of the burr as similarly sized dents, despite no such dent existing on the left side of it. Changing the type parameter to defects and bulges is supposed to alleviate this issue, and an example of this can be seen in FIGURE 6-47.

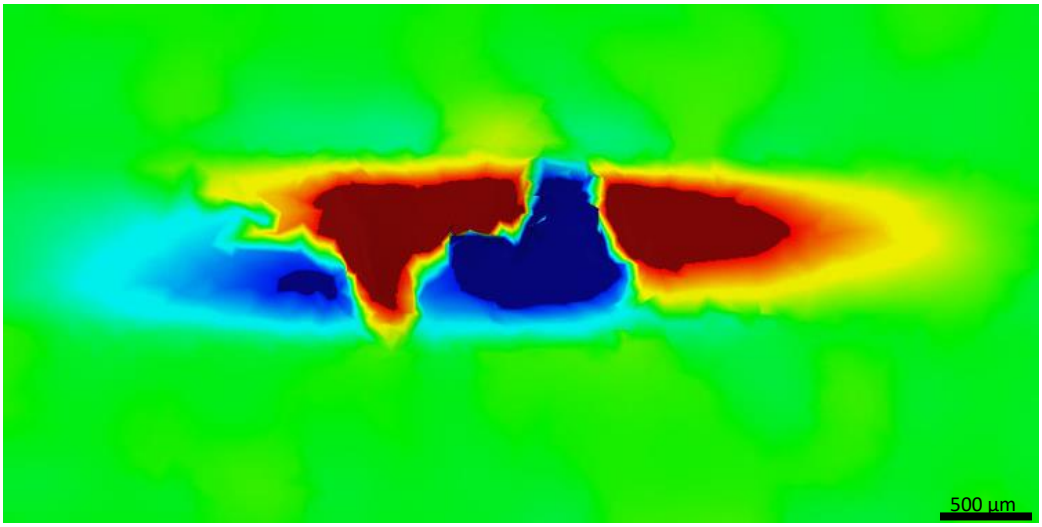

FIGURE 6-47: Example defect from the surface defect map for Sample 2 with type set to DB

As seen in this figure, while the burr was correctly identified as a bulge by the red color on the heatmap, it added significant noise around the edges of the dent. It also calculated the burr as much larger than it actually is on the part. As such, the decision was made to keep the type parameter set to depressions only.

Some of the false positives detected can be attributed to the mesh itself, as several of the defects had measured mesh elements that were not present on the artifact, and several others had components that were picked up incorrectly. Drop out, or holes in the mesh, was also an issue present in almost every measured defect. However, elements like some of the scratches were not affected by these issues and were still not picked up. This could have been due to the relative depth of the defect itself, because as noted in the previous trial the defect map struggled with shallow defects.

The average percent error for the partial mesh investigation of Sample 2 can be seen below in FIGURE 6-48.

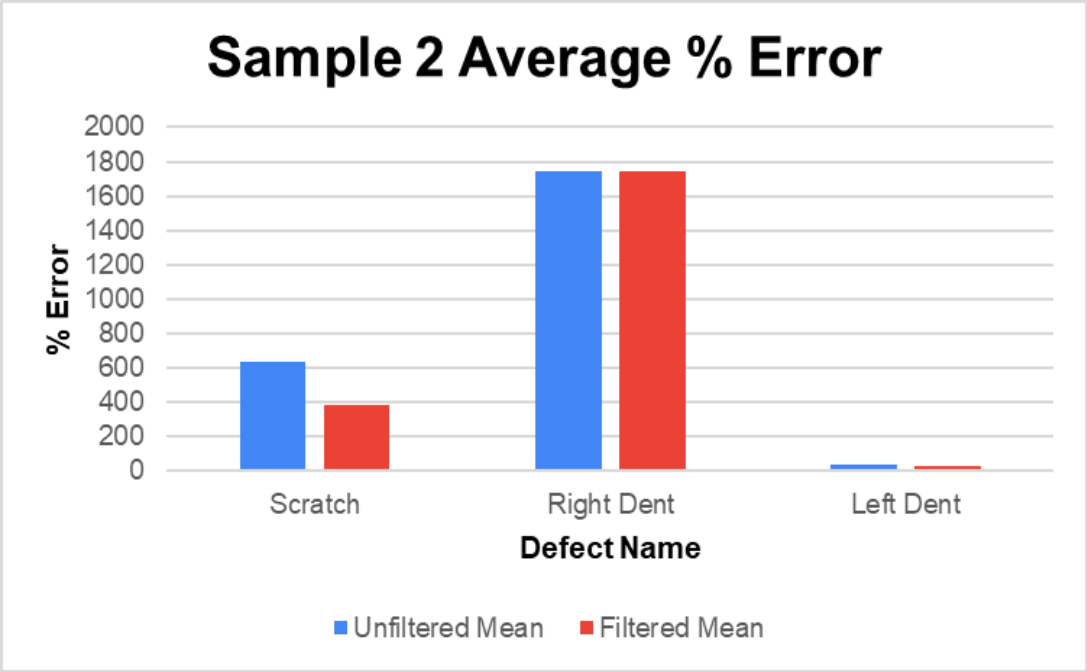

FIGURE 6-48: Average percent error from the surface defect map for Sample 2

The left dent was the defect with the least issues present in the mesh, so it predictably measured the least error at 37% unfiltered error and 33% filtered error. The scratch and the other dent had the most consistent mesh available of the other defects on the part, and even though this was the case, the error registered was significant enough to call into question the accuracy of the defect map in these areas. As expected, this caused significant issues with the classification tool, as can be seen in FIGURE 6-49.

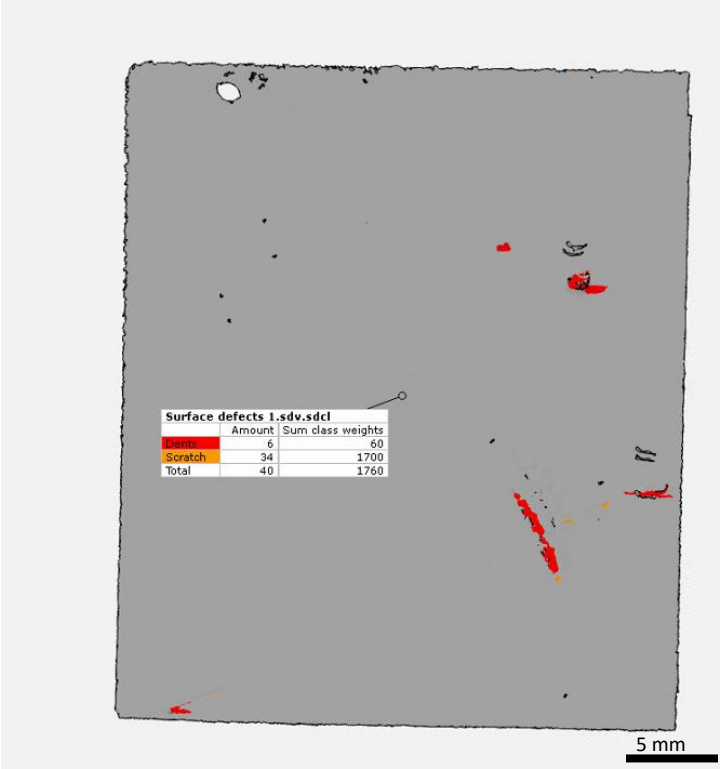

FIGURE 6-49: Surface defect classification for Sample 2

Even applying significant filtering to the detected defects, there are still many incorrectly classified defects and noise present in the classification. As such, the use of the defect map tool and classification was not successful on this mesh.

For Sample 5, the filtered and unfiltered mean are provided for the partial mesh inspection in FIGURE 6-50.

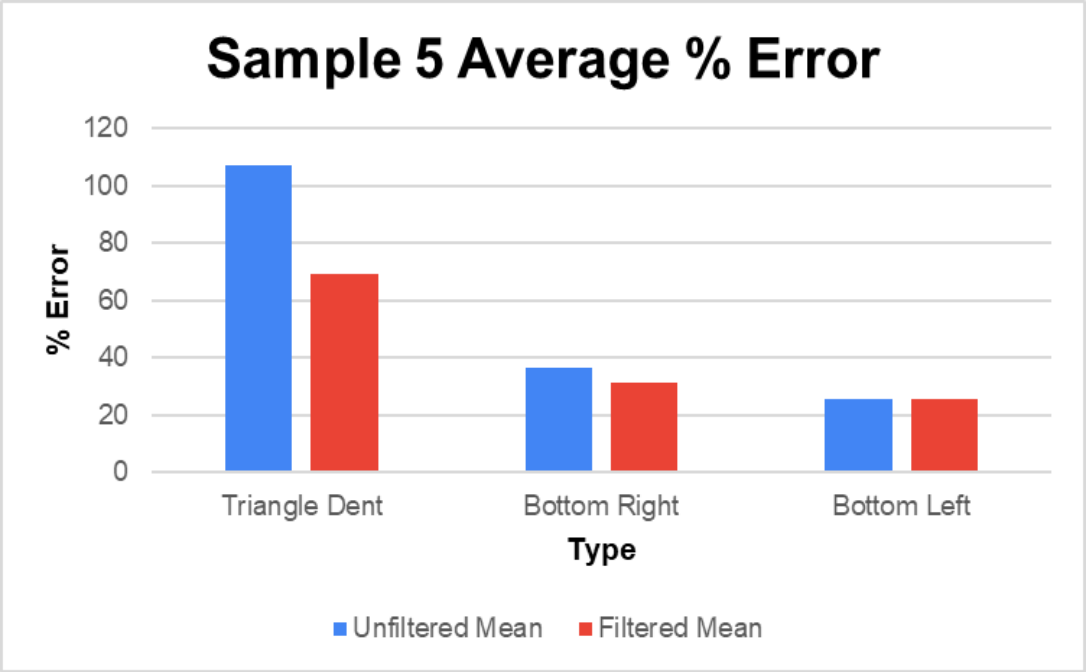

FIGURE 6-50: Average percent error from the surface defect map for Sample 5

This sample was very similar to Sample 2, but the instrument was able to measure it with less dropout and artifacting, which are conditions where the mesh drops out sections of data where there should be datapoints, or adds erroneous datapoints. As such, there is significantly less error in the measurement of each defect, with the highest filtered error registered by the triangle dent at 69% error. The triangle dent was also the only measured defect that suffered from dropout in the mesh. While the other two dents are not as accurate as the measured artifacts, the error measured is significantly lower than Sample 2 at 31 and 25% respectively.

The true/false positive data is displayed in Table 6, and the full surface classification is shown in FIGURE 6-51.

Table 6: True/False positives for Sample 5

Type	Number detected
True Positives	9.72
False Positives	151

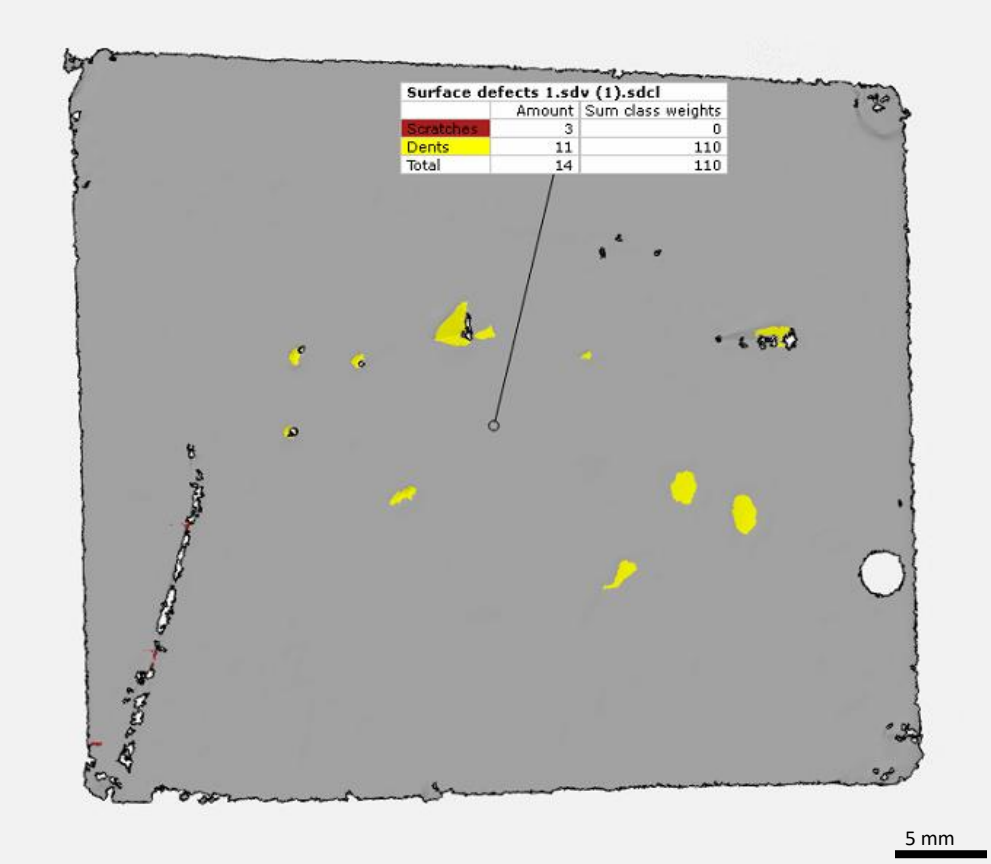

FIGURE 6-51: Surface defect classification for Sample 5

Note that there was no known way to combine partial defects or edit detected defects, so the number in the screenshot does not match the number of true detected

positives. The number of true positives detected was 9.72 out of a possible 15, with the map failing to pick up the scratch and the shallowest dents on the surface.

While it was significantly easier to sort the defects on this mesh, there were still a couple of issues with the classification step. The main issue that appeared in this step was the presence of false positives the system calculated to have a value of infinity. These defects did not appear in the manufactured trials, but were present for all of the organic trials in varying quantities. Because the formula result parameter was not a value that could be referenced when writing the filtering rules, these defects would inexplicably appear and disappear in the classification result, and exhibited strange behavior such as changing based on the spacing of the text in the formula code, and ignoring previous filtering rules altogether. The noise was eventually able to be filtered out of the classification, but it required significant time and effort to identify these defects and manually exclude them from the finished product. Although the accuracy level achieved was not as good as the results established for the manufactured artifacts, the classification tools were able to be utilized on this artifact with moderate success.

The other 3 artifacts all had issues that rendered the results from either the defect map or surface classification tools either inadequate for calculations to be made or in some cases completely unusable. Sample 10 suffered from major dropout that affected all 3 of the main scratches on the face of the artifact. FIGURE 6-52 shows the defect map as calculated for Sample 10.

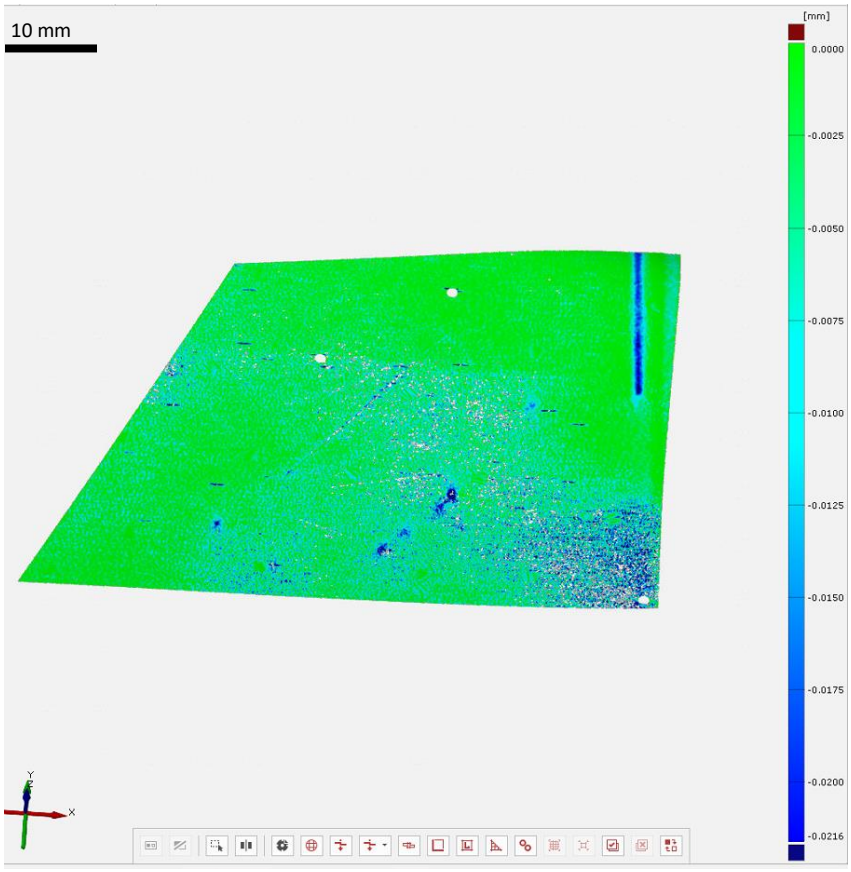

FIGURE 6-52: Surface defect map for Sample 10

While the surface defect map was able to broadly identify one of the scratches on the surface as can be seen in the figure, when the classification tool was applied it was revealed that it was actually detected as hundreds of miniscule defects as part of the 2540 defects the tool calculated on the surface of the artifact. It was not feasible to apply any classification to this part without isolating the inspected area to just the defects of interest, which defeats the purpose of attempting to classify the whole part. A partial inspection of the best detected scratch could not be performed either due to the lack of tools provided for inspection of a part without a CAD model in the GOM software. The deviation check used for previous samples relies on comparing the mesh to a projected

plane to calculated the deviation, so this calculation was completely incorrect due to the bends in the part.

A similar issue was encountered in the inspection of Sample 13, which was the artifact with a single bend in it. However, the end result was slightly different. The true/false positive data for Sample 13 can be seen below in Table 7, and the classification is visible in FIGURE 6-53.

Table 7: True/False positives for Sample 13

Type	Number detected
True Positives	11.22
False Positives	6472

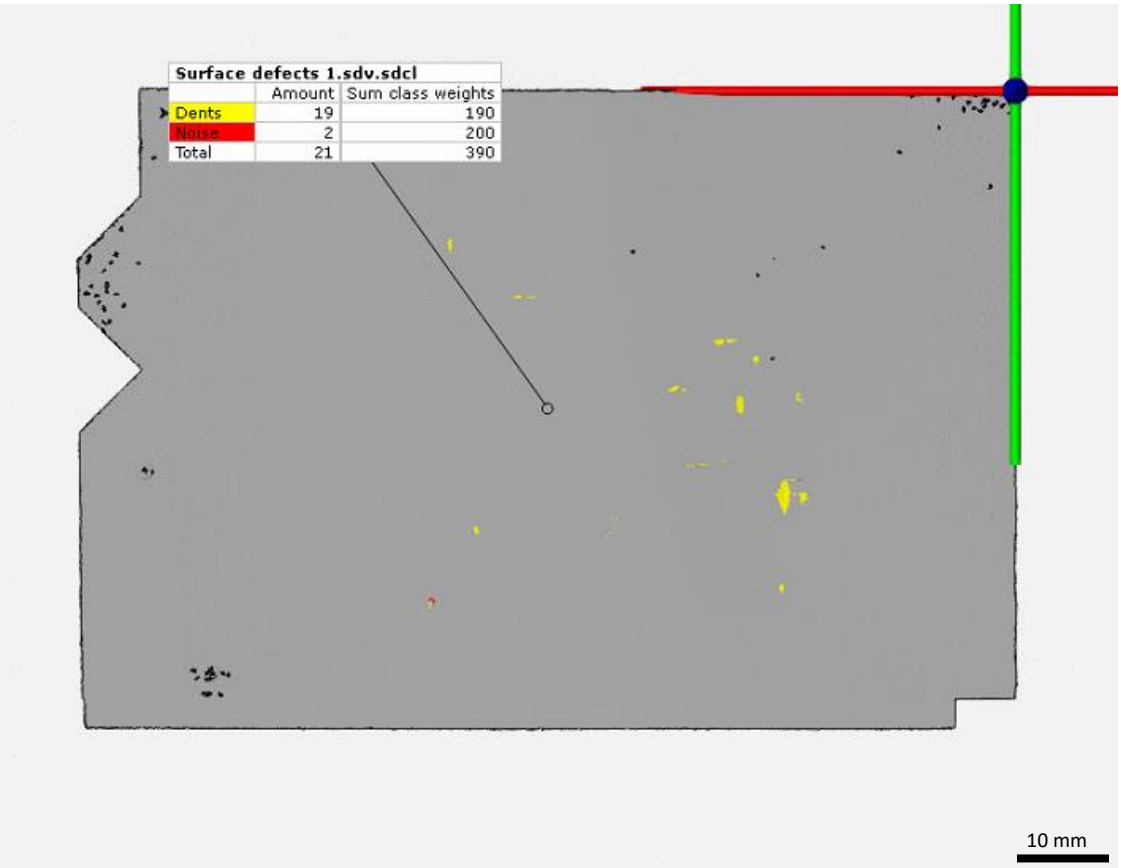

FIGURE 6-53: Surface defect classification for Sample 13

This sample was the second most successful classification of an organic artifact, with 11.22 defects correctly identified out of a possible 14. Compared to sample 10, the classification process was relatively simple as the software had correctly identified almost all of the defects as their proper size instead of as hundreds of parts of a whole. There was still a slight issue with the removal of the false positives, as can be seen by the 2 false positives still included in the figure that were identified as noise. However, the classification created from this data could be useful, unlike the results from Sample 10.

The partial inspection of Sample 13 ended in the same result as Sample 10, with the deviation to geometry check failing due to the bend in the part throwing off the

measured deviation. A fix for this issue was attempted where the check was only applied to the areas with defects situated away from the bend in the middle, but it still appeared to affect the results as no useful deviation check was possible.

FIGURE 6-54 shows the surface defect map calculated for Sample 11.

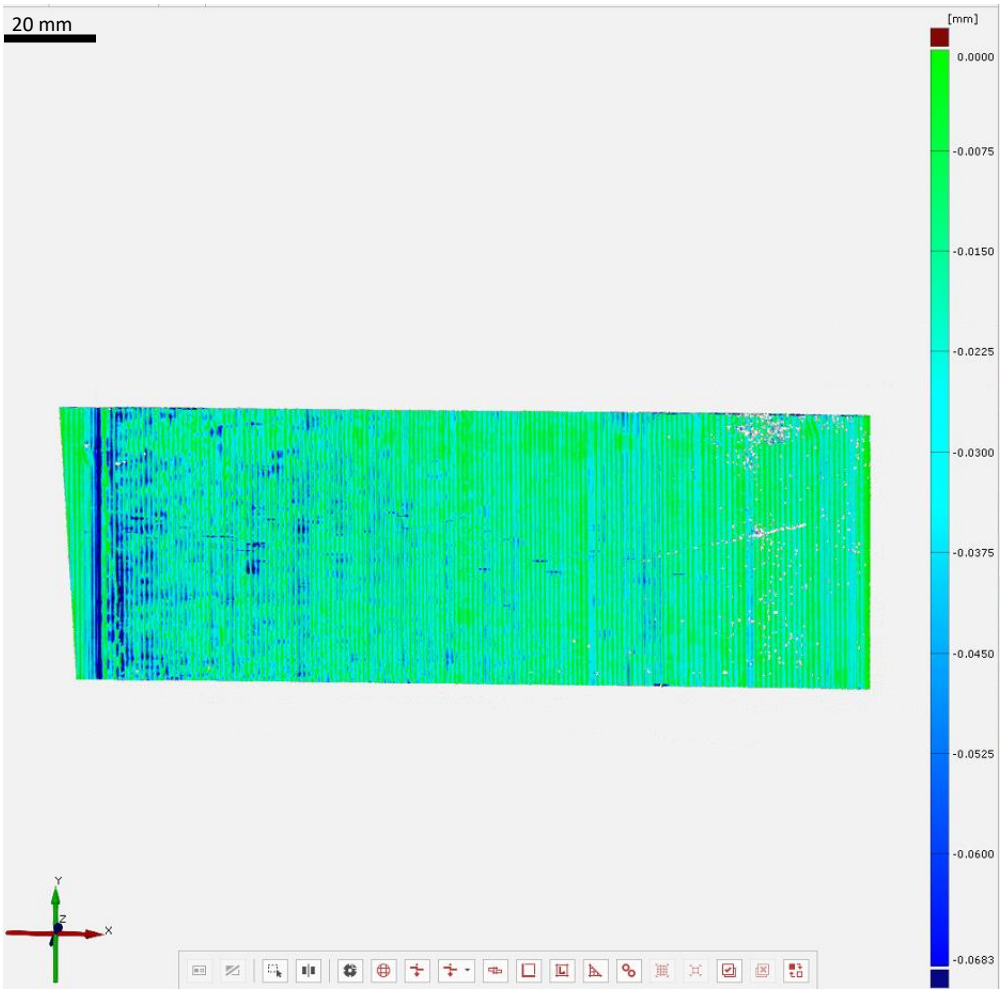

FIGURE 6-54: Surface defect map for Sample 11

As previously mentioned, Sample 11 has a ridged surface, with ridges that are approximately 65 microns deep. While some of the scratches on the surface suffered from dropout, large portions of these ridges were incorrectly detected as defects before

any filtering was applied. The results of the classification tool can be seen below in FIGURE 6-55.

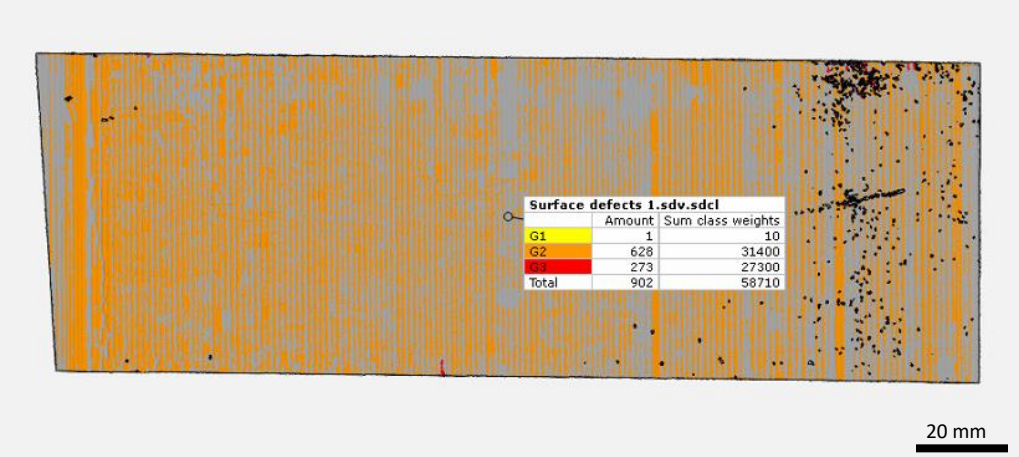

FIGURE 6-55: Surface defect classification for Sample 11

Looking at the classification without any filtering, it is clear that the software has detected the portions of the scratches as part of the ridges. While this may not be true for 100% of all the detected defect pieces on the surface, the majority of them have been detected this way and therefore would not be possible to classify. Additionally, the partial inspection was not possible due to the deviation check issue that was present for Sample 10 and Sample 13, as well as there being no identifiable defect detected to check against.

Considering these results, it is definitely not recommended to utilize the defect detection and classification tools on parts that have a consistently uneven surface such as the ridges of Sample 11. Additionally, while it may still be possible to use the tools on a bent surface, it is generally recommended that the surface be as flat as possible.

CHAPTER 7 CONCLUSIONS

Analysis of the correlation tables in the previous chapter shows that there is no correlation between any of the 4 main variables that go into calculating the surface defect map in terms of percent error when comparing the mesh and the defect map, but there is a correlation between those variables when counting false positives detected. While on its face this correlation may not seem that important, the validation trials showed that the presence of false positives can interfere with correctly identified defects, as well as just generally wreaking havoc with the classification filtering system. Therefore, it is important to minimize the number of false positives detected if possible.

However, it was also observed that reducing the parameters too much, specifically the max defect size line, can cause the defect map tool to not only underestimate the number of true defects on a mesh, but misrepresent the shape and size of these defects as well. Analyzing the artifact to be measured beforehand to be able to properly set these parameters is important to make sure that the tools function as accurately as possible.

Additionally, while it may not have been particularly important when analyzing the ideal artifacts, mesh quality was incredibly important to the success of the defect tools on real world artifacts. Meshes that contained dropout or artifacting significantly impacted the ability of both the defect map to calculate defects and the classification tool to classify them. Protrusions on the surface of the artifact also caused issues with both detection of false positives and misrepresentation of defects nearby. Surface structure

also seemed to be incredibly important to the ability to detect defects, with the ridged surface of Sample 11 rendering both tools unusable. Not enough different surface finishes were tested in these experiments to come to a definitive conclusion on which ones perform the best.

Parts that contain deformations in their geometry such as bends produced mixed results, with Sample 13 being classified successfully and Sample 10 failing to be classified. It seems likely that there is a relationship between a part containing a bend and an increase in the number of false positives detected on the part. There also appeared to be a relationship between the size of the artifact and the number of false positives detected. However, no formal experiments were done to confirm either of these hypotheses, so the observed trends could potentially be attributed to other factors instead.

Orientation had a small but noticeable effect on the performance of the defect tools. Rotating the part caused the placement of points overlaid by the point grid tool to rotate in turn with the rotation of the part about the Z axis. There was an observed decrease in false positives detected at all orientations from 15 to 75 degrees, while the percent error was at its lowest between 0 and 45 degrees. It appears that rotating the part about the 0 axis may decrease false positives, but potentially at the cost of defect map accuracy. There was also a slight observed difference in false positives based on which XYZ direction was chosen based on the orientation of the part.

For the inverted experiments, there was an increase in percent error for the scratches and the mesh overall of the inverted sample, while the dents remained approximately the same. Both the regular sample and the inverted one registered the

same amount of true and false positives.

Overall, the classification tool appears to be calibrated to overestimate the number of defects on a given part, as the software only missed clearly defined true positive defects when the max defect line was set to a smaller value than the defects themselves. This is observable in the fact that there were only 2 trials performed across all experiments conducted on defect artifact 2 that failed to capture the proper number of defects on the surface. Both defect tools performed the best on defects that were clearly defined, with dents consistently acquired and classified better overall than scratches and the clean, manufactured defects producing much more consistent results than the ones on the organic artifacts.

Finally, the formula component of the classification tool simultaneously has a lot of unexplored depth and arbitrary restrictions that make it difficult to work with. It functions using a custom combination of math and programming functions, which provides great flexibility in possible detection and filtering equations, but a lack of documentation makes the capabilities of the tool unclear. The requirement that the results recalculate after every change to the formula also makes it time consuming to make edits, and confusing as defects will change their reference number based on the parameters of the formula. The reference table is also unable to be exported from the tool itself, which makes creating any filtering rules based on the data found within it a time consuming and arduous process of manually copying the data to a spreadsheet by hand. Unexplained behavior was also observed in the case of the defects that were calculated as infinite, with their presence appearing and disappearing with changes to formula syntax

and ignoring previously defined classification rules. Further work in understanding the classification formula and filtering rule language will likely produce additional benefits in the use of the software's built-in tools.

CHAPTER 8 RECOMMENDATIONS

The most important recommendation to come of these experiments is to assess the artifact that the defect tools will be used on before attempting to apply them. If the part is to be scanned on a GOM structured light scanner for data acquisition, the defects must be acquirable according to the capabilities of the instrument, and the resulting mesh should be inspected for artifacting and dropout. Part geometry is important for all samples that the tools may be applied to, as the surface should be as flat as possible and free of any raised surface structure. Size may also be considered.

When applying the surface defect map tool to a part, aligning the part to a known orientation should be taken into consideration, as results do change based on part orientation. Changing the alignment after the fact may also break a previously constructed defect map based on how it was constructed, so the alignment should be constructed beforehand if possible. To set the defect map parameters themselves, the max defect size line should be set to approximately the width of the largest defect on the surface of the part. With a default XY alignment the XYZ direction set does not seem to be particularly important, although changing this orientation may cause the operator to choose X or Y. The number of directions parameter should be set to 1, and the type parameter should be set to either depressions or bulges only, not both.

To apply the surface classification tool, the formula used to classify defects should be entered into the tool first to provide the operator with a list of all the detected defects and their parameters, as this data is not available without a classification formula. While the resulting table can be used to write classification rules, it is advised to know

the parameters that will be used to classify the defects beforehand, as the table is hard to use and cannot be exported. It is also recommended to make liberal use of the exclude areas tool, both before applying classification rules to reduce the number of false positives that need to be filtered and after to remove extraneous defects not caught by the filters.

CHAPTER 9 FUTURE WORK

To build on the work done by these experiments, there are a number of additional experiments that could be performed to gain more knowledge on the process of defect acquisition and classification. The results from these experiments could further optimize the recommendations made for use of the defect tools. Several of the experiments proposed could serve to further validate the results obtained or compare them to other available options for defect classification.

The first experiment would be to take the artifacts measured for these experiments and apply an AI defect detection and classification tool to them to compare the results between the two. An experiment like this was part of the originally proposed experimental plan for this paper, but difficulties implementing the available tool for use with the measurements taken caused it to be scrapped. The existing meshes used for these experiments could be provided for use with the external tool with no additional measurements needed. The same data analysis could be run, with the percent error calculated between the defect map and the results of the AI tool at predefined points. This data could then be plotted to determine the accuracy of the ATOS defect tools.

The second experiment would be to investigate the performance of the defect tools on parts measured by external sources compared to the performance of the integrated structured light scanner. The previously measured artifacts could be measured by an external source and compared, or a completely new set of artifacts could be acquired. the environmental relative humidity influence on the structured light system. Besides the artifacts the experimental setup would be the same, with the data validated

through both a full and partial mesh inspection that measures percent error and true/false positives detected.

The third experiment would involve an investigation into the performance of the defect tools when provided with a CAD model to compare to. Much of the defect inspection workflow in ATOS Professional is designed with the operator having a CAD model in mind, and the surface defect comparison step of the process has the option of comparing the map to a CAD model before the classification tool is applied. An experiment to measure this would involve measuring artifacts that have CAD models both with and without the model. The same inspection process could then be applied, with the CAD model involved in one inspection and absent from the other.

A fourth experiment could assess the effects of surface finish and spray on the capabilities of the defect tools. It was observed in CHAPTER 6 that a ridged surface finish greatly affected the ability of the defect map tool to identify defects independent of the classification by gradient. Requiring all the artifacts to be sprayed before scanning also affected the surface finish perceived by the instrument. Experiments could be run on artifacts with different surface finishes at different levels of spray to determine the effect on the accuracy of the defect map and the number of false positives detected by the classification tool.

One final proposed experiment would be to observe the results of changing the defect classification formula on the overall performance of the classification tool. The classification formula used in the experiments conducted for this investigation is based on an existing formula built in to the ATOS software designed to detect defects created

by the injection molding manufacturing process. Because the classification tool relies so heavily on the classification formula, changing this formula could result in an easier classification process for certain kinds of artifacts. An experiment could be designed where several different classification formulas are applied to different artifacts, and the results are compared.

REFERENCES

[1] Emmerich, J.-F. (n.d.). What is structured light scanning? Medit Blog. Retrieved November 10, 2022, from https://blog.medit.com/medit/what-is-structured-light-scanning

[2] GOM. (n.d.). Triangulation: Proven measuring principle with a future. GOM Metrology. Retrieved November 10, 2022, from https://www.gom.com/en/topics/triangulation#:~:text=The%20triangulation%2Dbased%20light%2Dsectioning,camera%20and%20the%20measuring%20object.

[3] International Organization for Standardization. (2014). Rolling bearings – Balls – Part 1: Steel balls (ISO 3290-1:2014). Retrieved December 5, 2022, from https://www.iso.org/standard/60132.html

[4] International Organization for Standardization. (2022). Raw optical glass – Vocabulary (ISO 9802:2022). Retrieved December 5, 2022, from https://www.iso.org/standard/77351.html

[5] International Organization for Standardization. (2008). Glass in building – Silvered, flat-glass mirror (ISO 25537:2008). Retrieved December 5, 2022, from https://www.iso.org/standard/42965.html

[6] Encyclopædia Britannica, inc. (n.d.). Taylorism. Encyclopædia Britannica. Retrieved November 10, 2022, from https://www.britannica.com/science/Taylorism

[7] Plott, R. F. (1945). SUPERSONIC FLAW DETECTOR, MODEL 1 T. Final Report on Part of PA No. 154-ML-54-3 (No. CT-2633). Chicago. Univ. Metallurgical Lab.

[8] Orner, J. W. (1960). Ultrasonic attenuation and physical properties of metals. WATERTOWN ARSENAL LABS MA.

[9] Fessler, P. A. (1974). Weld defect detection by ultrasonic and x-ray methods (No. GEPP-132). General Electric Co., St. Petersburg, Fla.(USA). Neutron Devices Dept..

[10] Ziekman, P., & Klumper, J. W. (1977). Laser Scanners for Defect Detection. In Applications of Holography and Optical Data Processing (pp. 661-670). Pergamon.

[11] Salahi, M., & Turner, C. W. (1984). Computer-aided imaging with optically-scanned acoustic transducers. In Acoustical imaging (pp. 513-523). Springer, Boston, MA.

[12] Ogura, M., Sakaue, K., Takagi, M., Adachi, Y., & Ikoma, T. (1981). Computer Aided Detection of Slow Motion of Defects in GaP Light Emitting Diodes. Japanese Journal of Applied Physics, 20(5), L363.

[13] Zhou, L., Chalana, V., & Kim, Y. (1998). PC-based machine vision system for real-time computer-aided potato inspection. International journal of imaging systems and technology, 9(6), 423-433.

[14] Zhang, J. B. (1996). Computer-aided visual inspection for integrated quality control. Computers in industry, 30(3), 185-192.

[15] Tulbure, A. A., Tulbure, A. A., & Dulf, E. H. (2022). A review on modern defect detection models using DCNNs–Deep convolutional neural networks. Journal of Advanced Research, 35, 33-48.

[16] Singh, S. A., & Desai, K. A. (2022). Automated surface defect detection framework using machine vision and convolutional neural networks. Journal of Intelligent Manufacturing, 1-17.

[17] Aydin, I., Sevi, M., Salur, M. U., & Akin, E. (2022). Defect classification of railway fasteners using image preprocessing and alightweight convolutional neural network. Turkish Journal of Electrical Engineering and Computer Sciences, 30(3), 891-907.

[18] Mullany, B., E. Savio, Han Haitjema, and R. Leach. "The implication and evaluation of geometrical imperfections on manufactured surfaces." CIRP Annals 71, no. 2 (2022): 717-739.

[19] CNN: Introduction to pooling layer. GeeksforGeeks. (2022, August 24). Retrieved November 10, 2022, from https://www.geeksforgeeks.org/cnn-introduction-to-pooling-layer/

APPENDIX A MEASURED ARTIFACTS

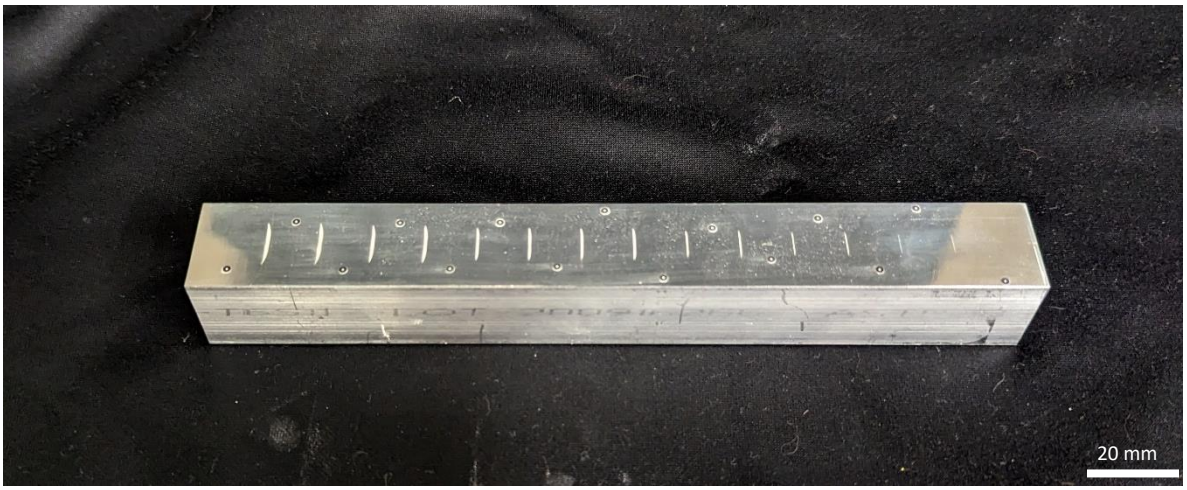

FIGURE A-1: Defect Artifact 1

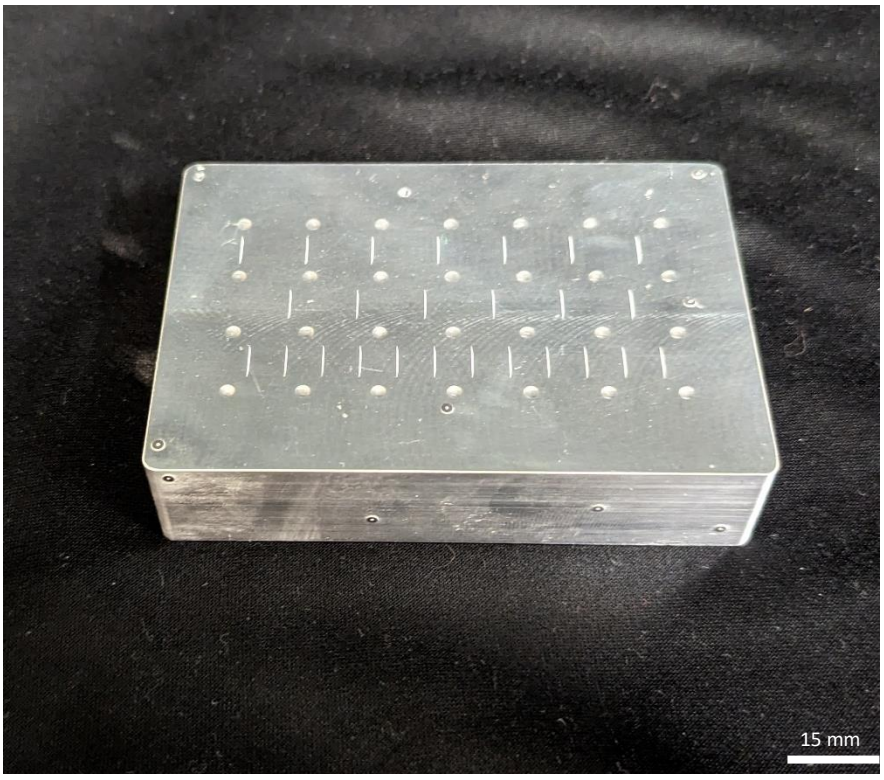

FIGURE A-2: Defect Artifact 2

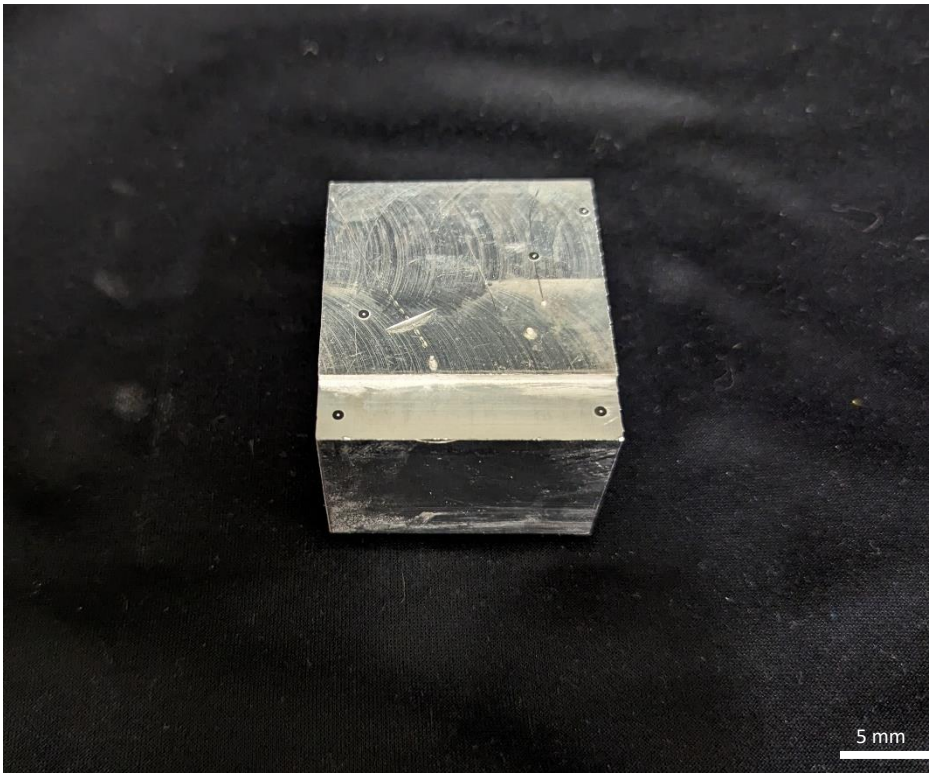

FIGURE A-3: Sample 2

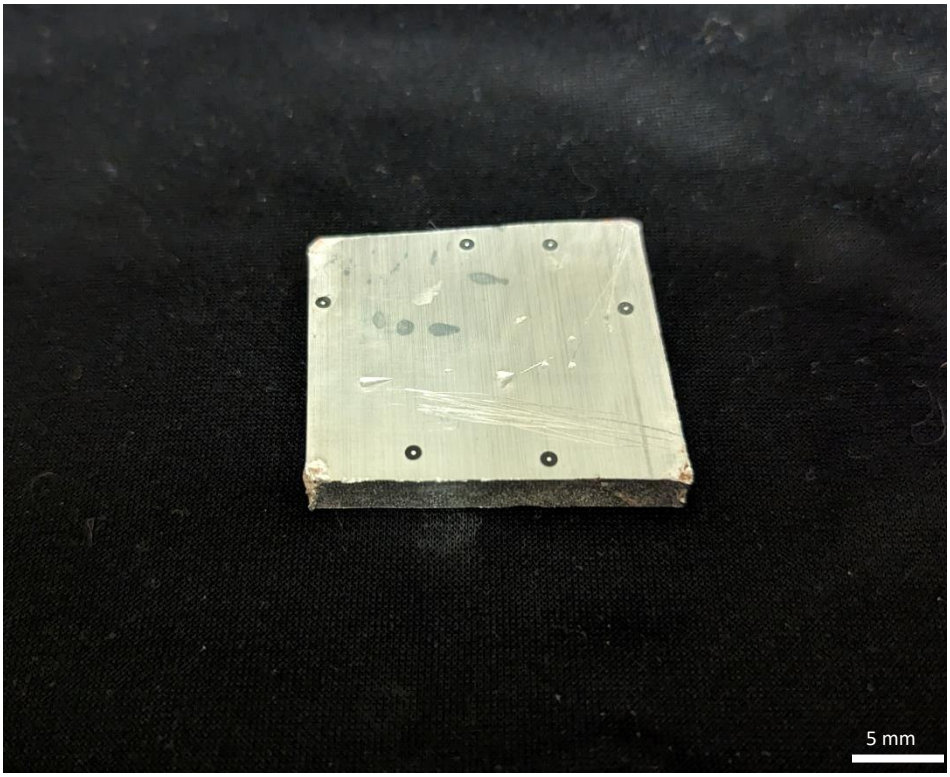

FIGURE A-4: Sample 5

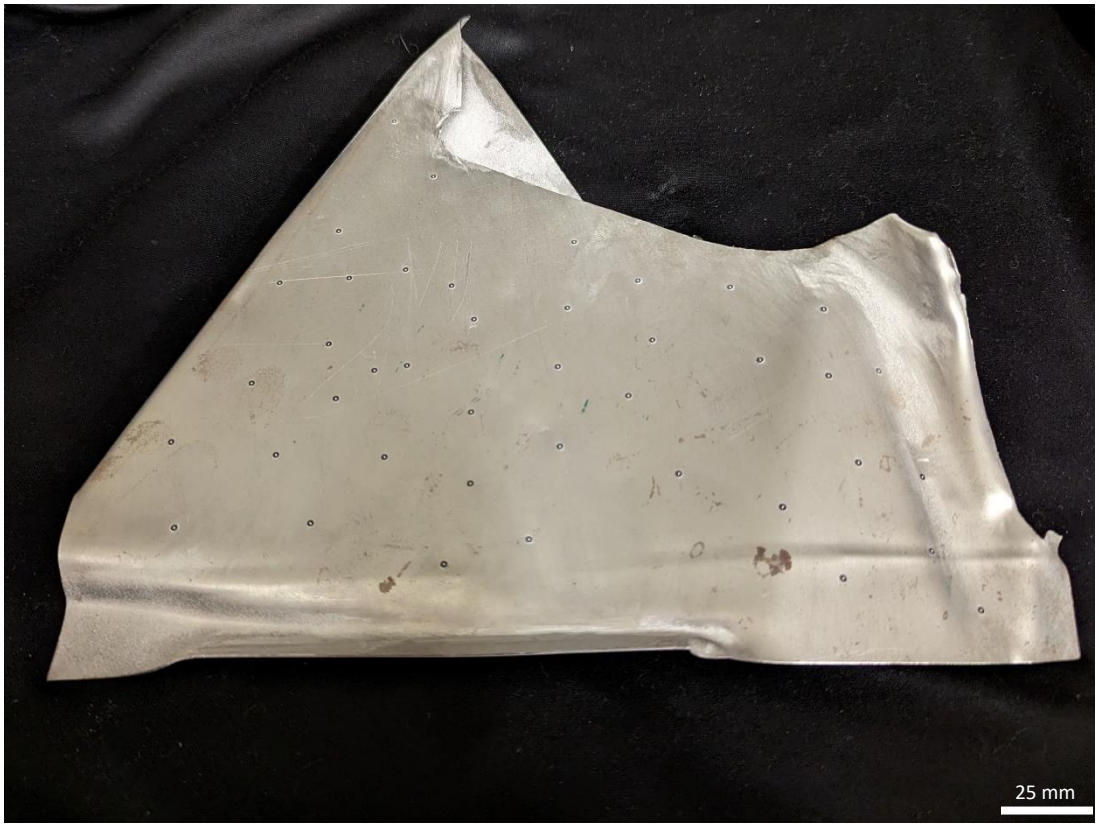

FIGURE A-5: Sample 10

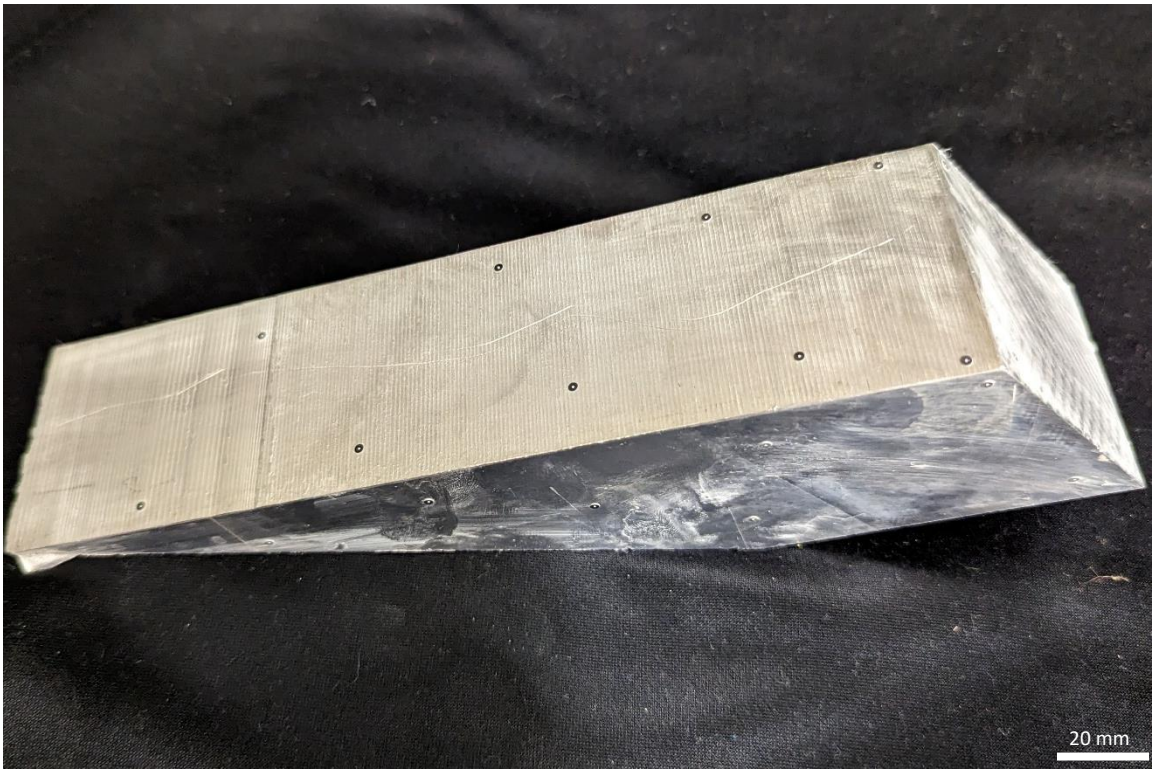

FIGURE A-6: Sample 11

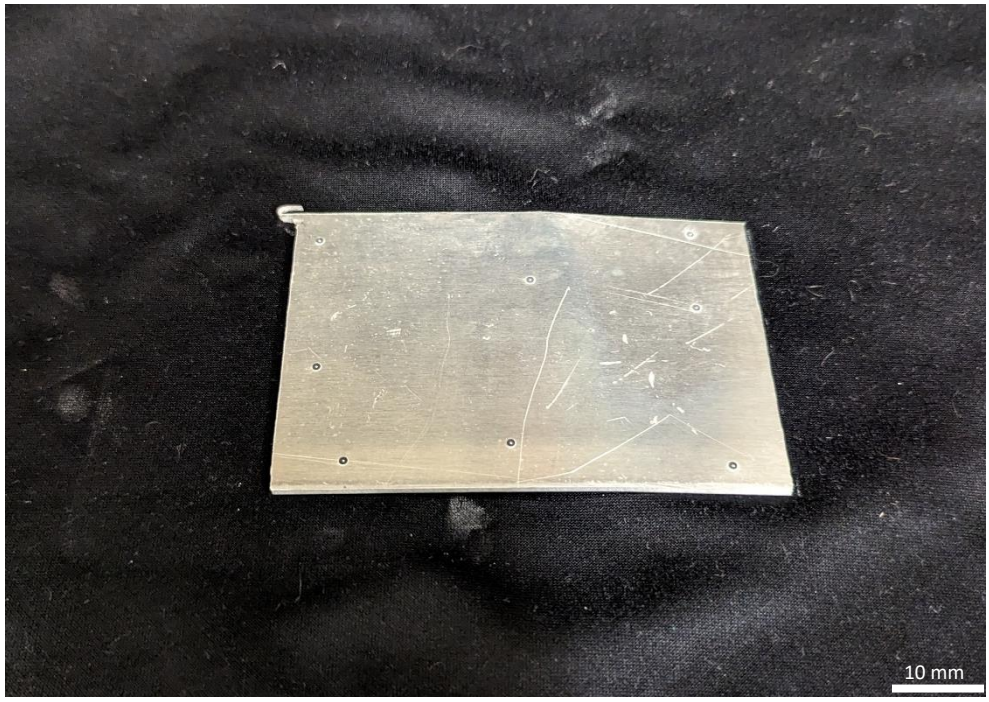

FIGURE A-7: Sample 13

APPENDIX B REFERENCED DEFECT NAMES AND LOCATIONS

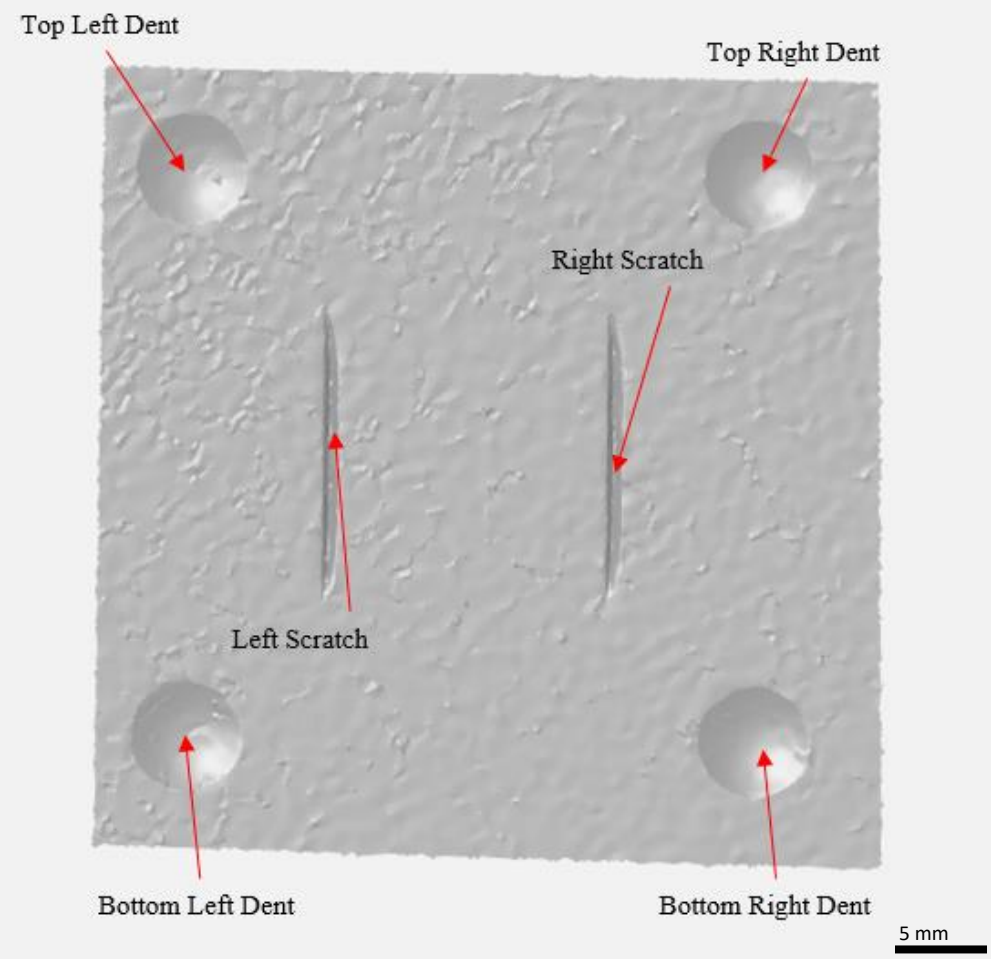

FIGURE B-1: Defect names and locations for partial mesh inspection

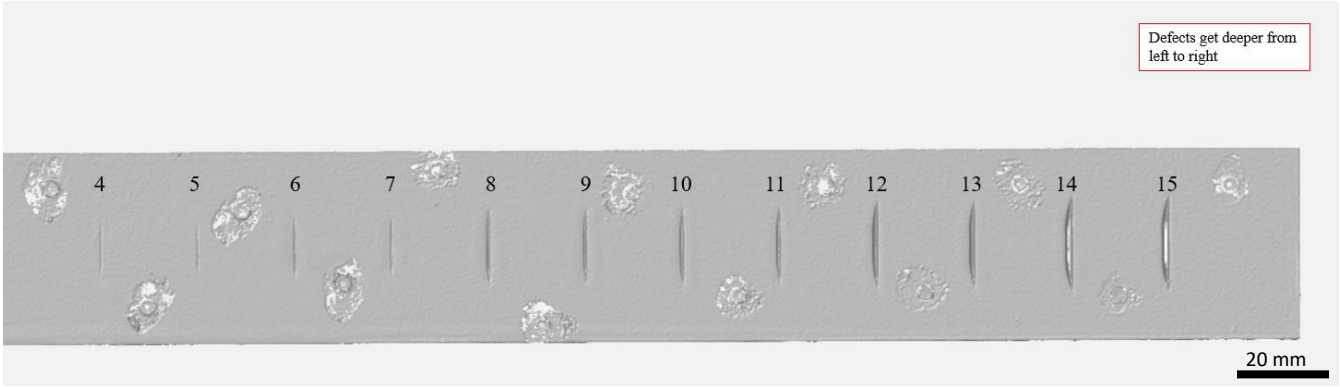

FIGURE B-2: Numbered defects for defect artifact 1

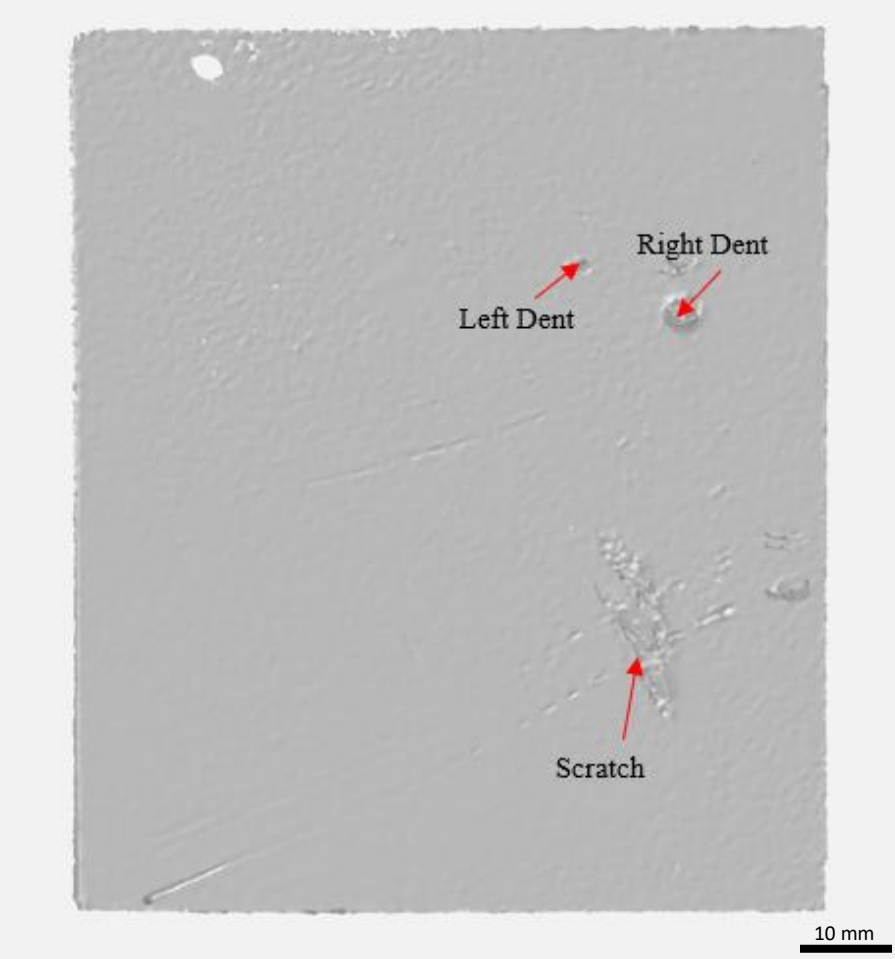

FIGURE B-3: Defect names and locations for Sample 2

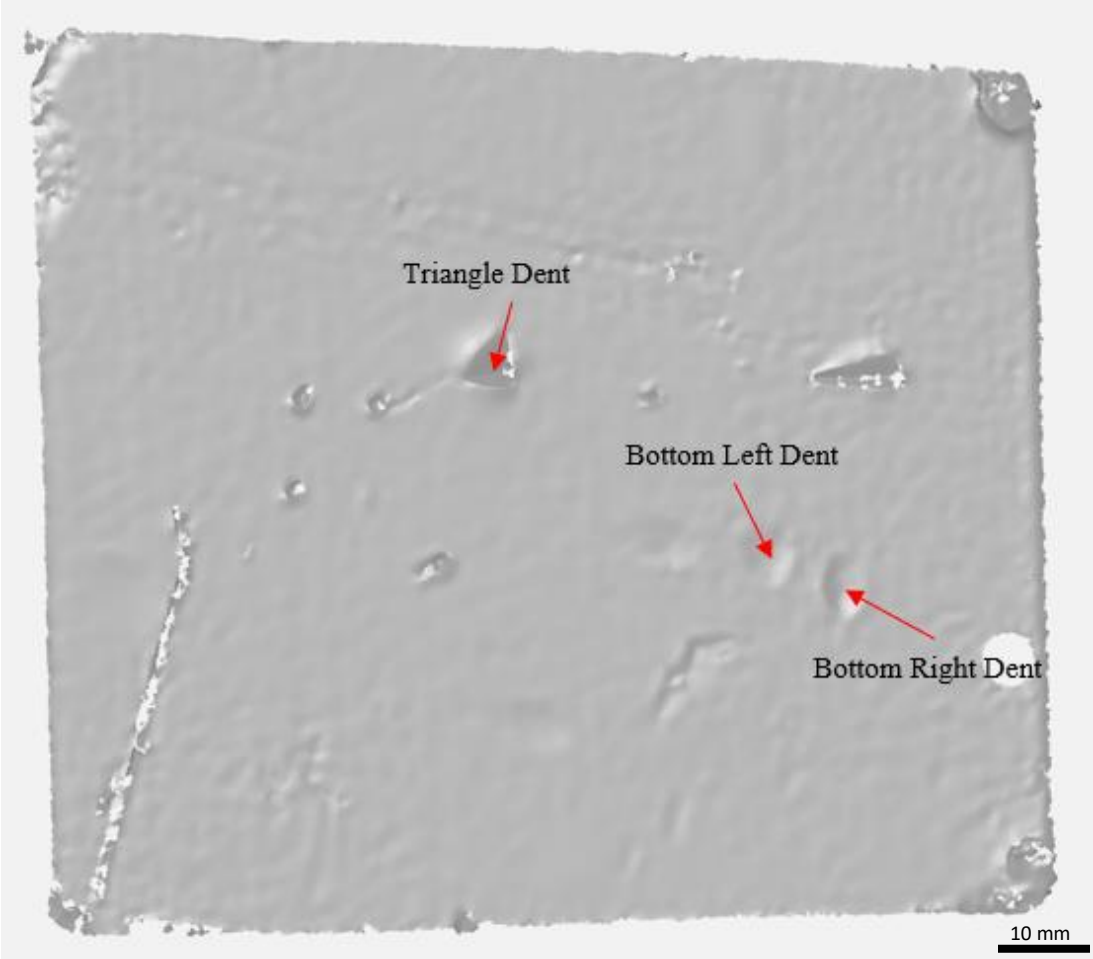

FIGURE B-4: Defect names and locations for Sample 5