Akmal, Jan Sher; Salmi, Mika; Hemming, Björn; Teir, Linus; Suomalainen, Anni; Kortesniemi, Mika; Partanen, Jouni; Lassila, Antti

Cumulative Inaccuracies in Implementation of Additive Manufacturing Through Medical Imaging, 3D Thresholding, and 3D Modeling: A Case Study for an End-Use Implant

Published in:
Applied Sciences (Switzerland)

DOI:
10.3390/app10082968

Published: 24/04/2020

Document Version
Publisher's PDF, also known as Version of record

Published under the following license:
CC BY

Please cite the original version:
Cumulative Inaccuracies in Implementation of Additive Manufacturing Through Medical Imaging, 3D Thresholding, and 3D Modeling: A Case Study for an End-Use Implant

Jan Sher Akmal 1,2,*, Mika Salmi 1,2, Björn Hemming 2, Linus Teir 2, Anni Suomalainen 3, Mika Kortesniemi 3, Jouni Partanen 1 and Antti Lassila 2

1 Department of Mechanical Engineering, Aalto University, Otakaari 4, 02150 Espoo, Finland
2 MIKES-National Metrology Institute of Finland, VTT Technical Research Centre of Finland, Tekniikantie 1, 02150 Espoo, Finland
3 Department of Radiology, Medical Imaging Center, University of Helsinki and Helsinki University Hospital, P.O.Box 263, HUS, 00029 Helsinki, Finland

* Correspondence: jan.akmal@aalto.fi; Tel.: +358-50-566-2207

Received: 25 March 2020; Accepted: 21 April 2020; Published: 24 April 2020

Featured Application: Accuracy of additively manufactured implants for clinical surgery.

Abstract: In craniomaxillofacial surgical procedures, an emerging practice adopts the preoperative virtual planning that uses medical imaging (computed tomography), 3D thresholding (segmentation), 3D modeling (digital design), and additive manufacturing (3D printing) for the procurement of an end-use implant. The objective of this case study was to evaluate the cumulative spatial inaccuracies arising from each step of the process chain when various computed tomography protocols and thresholding values were independently changed. A custom-made quality assurance instrument (Phantom) was used to evaluate the medical imaging error. A sus domesticus (domestic pig) head was analyzed to determine the 3D thresholding error. The 3D modeling error was estimated from the computer-aided design software. Finally, the end-use implant was used to evaluate the additive manufacturing error. The results were verified using accurate measurement instruments and techniques. A worst-case cumulative error of 1.7 mm (3.0%) was estimated for one boundary condition and 2.3 mm (4.1%) for two boundary conditions considering the maximum length (56.9 mm) of the end-use implant. Uncertainty from the clinical imaging to the end-use implant was 0.8 mm (1.4%). This study helps practitioners establish and corroborate surgical practices that are within the bounds of an appropriate accuracy for clinical treatment and restoration.

Keywords: 3D printing; craniomaxillofacial surgery; medical imaging; computed tomography (CT); segmentation; digital design; CAD; implants; cumulative error; error propagation; standard uncertainty; spatial accuracy; quality control; quality assurance

1. Introduction

Additive manufacturing (AM), non-technically known as three-dimensional (3D) printing, is a process in which material is added and joined typically on a layer-by-layer principle to make products using digital data of a 3D model [1]. AM is a rapidly evolving technology [2], which originated in early 1980s [3]. To this end, it differs from the conventional manufacturing methodologies, i.e., subtractive and formative manufacturing. AM has been widely implemented in the medical industry [4], and it has gained clearance for more than 80 AM medical devices such as maxillofacial implants and a knee joint...
replacement system from the United States Food and Drug Administration (FDA) [2]. Substantial AM growth is enabled by its economical lot-size-one that allows for unprecedented geometrical freedom [5] in one build cycle using minimal set-up because it is a general-purpose digital technology which is not constrained by patient-specific tooling [6–8]. Further, it enables a unique supply chain with increased communication and responsiveness [9] that complements the state of urgency in the medical field.

The powder bed fusion (PBF) technique is one of the seven AM processes in which thermal energy is used to selectively fuse regions of the powder bed [1]. When this method is applied to fabricate metallic parts, it is typically known as direct metal laser sintering (DMLS), selective laser melting (SLM), and electron beam melting (EBM) depending on the power source and the machine manufacturer. This method has been widely implemented to manufacture custom-made and patient-specific implants from metal alloys, particularly titanium [10–17]. Implants seem to be one of the most utilized applications of additive manufacturing in the head and neck area in hospitals [18]. Further, it is increasingly fulfilling the demand for porous titanium that can facilitate osseointegration through the independent adjustment of the mechanical properties and porosity in terms of size and distribution [19–23].

A classification for medical applications of AM was established by Tuomi et al. [24,25], which comprises medical models for planning, medical aids, tools for medical devices, inert implants, and bio-manufacturing. Akmal et al. [26] have extended the classification of custom-made inert implants to intelligent implants containing patient-specific digital data and drug delivery systems. The advent of preoperative virtual planning, based on medical imaging, segmentation, implant modeling, and manufacturing [27] in particular with AM techniques, has enabled excellent functional and aesthetic reconstruction of different craniofacial surgical procedures including the reduction of intraoperative time [28–33].

The craniofacial, in particular the mandibular, deformities can be caused by tumor, injuries, and/or infections [34]. The reconstruction of mandibles followed by an ablative tumor surgery poses a challenge for practitioners [31]. This is due to the nature of complex 3D patient-specific geometry that is difficult to replicate. Any signs of an aberration in its structural alignment may lead to functional disturbance due to the malocclusion or temporomandibular disorders [33]. Further, difficulties may arise due to extensive bone resection, temporomandibular joint involvement, or severely distorted bony contour [33].

Evidently, the development of medical imaging, in particular computed tomography (CT) techniques, has changed the dynamics of the medical industry because it has enabled data capture from the interior and exterior of the patient-specific anatomy. This captured data is volumetric in nature comprising 3D block of smaller cuboid structures called voxels (volume element). Each voxel represents a relative X-ray attenuation [35] that is characterized by grey scale values in Hounsfield unit (HU). More specifically, the Hounsfield units represent material X-ray attenuation value relative to water. In principle, this relationship enables a calibration of CT images by means of the grey scale range. The slice-by-slice captured voxels are interpreted by stacking the slices to obtain multiple 2D representations [35]. The CT images are generally stored in 16 bits according to the international standard of digital imaging and communications in medicine (DICOM) format.

CT imaging can essentially be performed using either a fan-beam multi-slice CT (MSCT) or a cone-beam CT (CBCT) system depending on the geometry of the X-ray acquisition [35]. Further, a dual-energy CT (DECT) system provides more possibilities for the clinical applications by acquiring CT images using two different X-ray spectra [36]. Though a MSCT system may provide a higher low-contrast resolution, a reduced image noise [37], and a rapid scan time [35], a CBCT system is increasingly used in the maxillofacial surgeries owing to an acceptable image quality [38], a similar scanning time [35,37], a lower cost, and a reduced radiation dose [35,37–39].

Regarding the quality assurance (QA) of the CT systems, QA programs for the MSCT system are well-established. However, the QA programs for the CBCT system are still evolving and present only a few established sources, i.e., programs for high-end and large-volume CBCT.
systems [40], the QA program proposed by the SEDENTEXCT Project [41], and the EFOMP CBCT QA Protocol [42]. The standardized QA Phantoms including the European project-based [41], commercial [43], and low-cost alternatives [44] enable the assessment of accuracy and quality of the CT systems [40]. The composition of a Phantom varies according to the application and involves several properties, i.e., radiation type, energy range, linear attenuation coefficient, and density, which mimic water or body tissues by means of attenuating photons [45].

Though conventional 2D representations of medical images benefit from the localized analysis of different anatomical structures, a reconstruction of medical images to a 3D model enables enhanced visualization of craniomaxillofacial structures, deformities, and implant sites [37,46,47]. Further, it assists in 3D modeling of patient-specific implants, saw guides, and drill guides [34,48]. A conversion of 2D medical images (DICOM) to 3D data (standard tessellation language: STL) requires 3D thresholding techniques for segmenting different forms of tissues. There are several 3D thresholding techniques for the 3D reconstruction of medical images, namely a volume-based approach, contour-based approach, and point-cloud-based approach [49]. However, the most commonly used method is the volume-based [39] that uses a marching cubes algorithm [50]. The volume-based approach involves an elimination of all voxels consisting of the grey scale values in HU that are lower than the selected threshold value \( t \). It includes all the voxels that are equal to or greater than \( t \). The following binary mask \( M_{x,y} \) (Equation (1)) denotes these relationships where \( I_{x,y} \) depicts the grey scale values of medical images in HU [51]:

\[
M_{x,y} = \begin{cases} 
0 & I_{x,y} < t \\
1 & I_{x,y} \geq t 
\end{cases}
\]  

(1)

Though the collective processes of CT imaging, 3D thresholding, 3D modeling, and AM yield beneficial outcomes of surgical procedures [15,16,28–33], both academics and practitioners have noticeably understudied the potential sources of their cumulative error [52]. Each step of the process can contribute geometric error to the resulting implant [53]. Previous efforts have been made to evaluate inaccuracies through different CT systems and their protocols [37]; different CT system voltage and current [54]; various CT systems and thresholding values [39,55]; different thresholding values [54,56]; various thresholding values and software [57]; different AM processes [53,58]; and various metal AM machines [59], all of which have prompted additional verification studies that account for the whole process chain.

The aim of this study is to evaluate cumulative spatial inaccuracies [52] in the process chain of implementing AM from medical imaging involving different CT imaging protocols, 3D thresholding involving various Hounsfield units, and 3D modeling involving patient-specific geometry. The case study follows the process chain of a pre-operative virtual planning for the procurement of an end-use implant to aid practitioners in establishing appropriate clinical practices that can allow treatment and restoration. The study designs and evaluates a QA Phantom in line with sound metrological expertise to validate a MSCT system. In addition, accurate measurement techniques traceable to the SI unit [60] are adopted for the reference measurements. These include a tactile-probing coordinate measuring machine (CMM) [61] and an optical 3D coordinates measuring system which comprises a combination of the structured light [62] and stereo vision [63] techniques. The procedure offers an innovative approach for determining the inaccuracy of an end-use implant involving the whole process chain as opposed to the partial analyses in the literature involving one or few process steps.

2. Materials and Methods

This study was conducted in compliance with relevant laws and institutional guidelines of the EU Directive 2010/63/EU. No human subjects or living animals were used in this study. Figure 1 shows the process of this case study.
2.1. The Sus Domesticus

This study involves a sus domesticus (SD) with an age of approximately nine months. The SD head was acquired from a conventional supplier in the food industry. The subsequent weight of the head was 9.553 kg. The head, intact with its soft tissue, was CT-scanned. After medical imaging, dermestid maculatus beetles were used to remove the soft tissue from the head to preserve a dry skull. The dry skull was covered with a plastic bag, and it was stored in a container at the room temperature to prevent dimensional changes, i.e., shrinkage, from external environmental factors such as exposure to humidity and UV. Subsequently, the dry skull was scanned and referenced using an optical 3D coordinates measuring system and a CMM, as shown in Figure 1.

2.2. The Quality Assurance Phantom

A QA Phantom was designed to verify the accuracy of the CT system in line with metrological expertise. The specific features of the Phantom allowed for measuring volumetric references of linear and angular dimensions as opposed to freeform anatomical features of the SD head. The Phantom was manufactured using conventional and precise techniques. The material of the Phantom was Polyoxymethylene (POM) due to its close representation of the bone density [64] and the bone linear attenuation coefficient [40,44]. Several studies have used POM as an appropriate Phantom material for the QA of the CT systems in the literature [40,41,65,66]. Subsequently, the Phantom was CT-scanned and referenced using the CMM, as presented in Figure 1.
2.3. Medical Imaging

A Siemens Somatom Definition Edge MSCT system was used to acquire medical images. Four different protocols of the CT system generated DICOM images of both the SD head and the Phantom, as shown in Figure 1 and Table 1. The CT imaging reconstruction kernel and image thickness were critical variables for evaluating the accuracy of the CT system. Consequently, these independent variables led to a variation in the pixels, current, and number of cuts. The CT images of the SD head are made available through an open access repository [67].

**Table 1.** Parameters of multi-slice computed tomography (MSCT) imaging for the sus domesticus (SD) head and the Phantom.

<table>
<thead>
<tr>
<th>Input</th>
<th>Protocol [No.]</th>
<th>Reconstruction Kernel</th>
<th>Image Thickness [mm]</th>
<th>Cuts [No.]</th>
<th>Voxel [mm]</th>
<th>Pixels</th>
<th>Field of View [cm × cm]</th>
<th>Voltage [kV]</th>
<th>Current [mA]</th>
</tr>
</thead>
<tbody>
<tr>
<td>SD head</td>
<td>1</td>
<td>J30S</td>
<td>0.5</td>
<td>870</td>
<td>0.625</td>
<td>512 × 512 × 870</td>
<td>32 × 43.5</td>
<td>180.00</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>J70H</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>J30S</td>
<td>1.0</td>
<td>435</td>
<td>0.488</td>
<td>512 × 512 × 435</td>
<td>25 × 18.2</td>
<td>180.00</td>
<td></td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>J70H</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>43.00</td>
<td></td>
</tr>
<tr>
<td>Phantom</td>
<td>1</td>
<td>J30S</td>
<td>0.5</td>
<td>364</td>
<td>0.488</td>
<td>512 × 512 × 364</td>
<td>25 × 18.2</td>
<td>180.00</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>J70H</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>J30S</td>
<td>1.0</td>
<td>182</td>
<td>0.488</td>
<td>512 × 512 × 182</td>
<td></td>
<td>54.00</td>
<td></td>
</tr>
</tbody>
</table>

2.4. 3D Thresholding

The volume-based approach of OsiriX MD (8.0.2, Pixmeo SARL, Geneva, Switzerland) software, which has been approved by the United States FDA, reconstructed the CT DICOM images to the STL format. Three individual HU values were used to segment the medical images of both the SD head and the Phantom, as shown in Figure 1. Table 2 presents the resolution, decimate resolution, and smooth iterations in addition to the HU values.

**Table 2.** Three-dimensional (3D) thresholding parameters for the SD head and the Phantom.

<table>
<thead>
<tr>
<th>Description</th>
<th>Thresholding Values [HU]</th>
<th>Resolution</th>
<th>Decimate Resolution</th>
<th>Smooth Iteration</th>
</tr>
</thead>
<tbody>
<tr>
<td>SD head</td>
<td>250 300 500</td>
<td>Normal</td>
<td>0.50</td>
<td>20</td>
</tr>
<tr>
<td>Phantom</td>
<td>0 −300 −500</td>
<td>Normal</td>
<td>0.50</td>
<td>20</td>
</tr>
</tbody>
</table>

2.5. 3D Modeling of the Implant

Assuming a case where a patient had to undergo a resection of a zygomatic bone tumor, an implant was digitally modeled to replace the deformity using 3Data Expert (11.0.0.32, Deskartes Oy, Espoo, Finland) software. The implant was designed using the right zygomatic bone projections of the SD head that were obtained from the optimal CT imaging and thresholding parameters. The resultant implant was stored in the STL format, and it was also used as a reference model, as shown in Figure 1.

2.6. Additive Manufacturing of the Implant

The computer-aided design (CAD) implant in the STL format was used to configure pre-processing parameters of AM using Materialise Magics (22.0, Materialise, Leuven, Belgium) software. During the pre-processing phase, four replicas of the implant were placed on the platform in the build chamber. The implants were oriented to create support structures for an adequate heat sink. Further, they were sliced to create contours including a laser hatch pattern and vector scanning for selectively melting the layers.
The pre-processed data was used to initiate additive manufacturing, in particular for the titanium PBF using SLM Solutions 125HL (SLM Solutions Group AG, Lübeck, Germany) machine. Table 3 lists the printing parameters, and Figure 2 illustrates the printing orientation, printing location, and support structures of the implants. Subsequently, the additively manufactured implants underwent heat treatment in an argon atmosphere for stress relief. During the heat treatment process, the temperature of the chamber was increased from 20 to 800 °C at a rate of 5 °C/min, it was held at 800 °C for four hours, and eventually, it was cooled to the room temperature at the same rate. The support structures of the implants were removed by an expert using manual instruments, i.e., pliers, a saw, and an angle grinder, to yield the end-use implants. Finally, one of the implants was scanned using the optical 3D coordinates measuring system to capture the resultant geometry, as presented in Figure 1.

**Table 3.** Additive manufacturing (AM) parameters for the end-use implant.

<table>
<thead>
<tr>
<th>ISO/ASTM AM Method</th>
<th>Machine</th>
<th>Software</th>
<th>Material</th>
<th>Layer Thickness</th>
<th>Other Parameters</th>
<th>Post Processing</th>
</tr>
</thead>
<tbody>
<tr>
<td>Powder bed fusion</td>
<td>SLM Solutions 125HL</td>
<td>Materialise Magics 22.0</td>
<td>Ti-6Al-4V (Grade 5)</td>
<td>0.03 mm</td>
<td>Standard process parameters</td>
<td>Heat Treatment and manual support removal</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Platform heating: 200 °C</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Particle size: Dv(10) = 32.5 μm, Dv(50) = 44.3 μm, Dv(90) = 59.9 μm</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Number of layers: 1460</td>
<td></td>
</tr>
</tbody>
</table>

**Figure 2.** The print orientation and location of four identical implants on the build platform illustrating supports in blue.

### 2.7. The Reference Measurements and Cumulative Error

A GOM ATOS II Triple Scan (GOM GmbH, Braunschweig, Germany) 3D coordinates measuring system and a Mitutoyo Legex 9106 (Mitutoyo Corporation, Kawasaki, Japan) CMM with tactile-probing systems were used to conduct reference measurements, as shown in Figure 1.

The optical 3D coordinates measuring system follows the structured light and stereo vision techniques built on the principle of camera-camera and camera-projector triangulation. Two calibrated reference gauge blocks were used to establish a traceability to the SI unit during the optical measurements. The measurements were performed in a room with a temperature range of 22.0–23.3 °C. The results of the measurements were compensated with the subsequent thermal expansion.

The CMM consists of a fixed bridge structure and uses the scanning and touch-trigger probing principles. The CMM is periodically verified with the interferometrically calibrated gauge blocks ensuring a maximum permissible error, \( E_{0,\text{MPE}} \) value of 0.35 μm + L/1000 μm where \( L \) is length in...
mm. The CMM measurements were performed in an automated climate-controlled laboratory with a temperature stability of $20 \pm 0.2 \, ^\circ C$ [68].

In this study, uncertainty was calculated according to the principles of Guide to the expression of uncertainty in measurement (GUM) [52]. Independent measurements were compared with the reference measurement of the specimens to calculate geometric variation using GOM Inspect 2018 (v2.0.1, GOM GmbH, Braunschweig, Germany) software, which is tested and certified by the National Institute of Standards and Technology (NIST) and the National Metrology Institute of Germany (PTB).

Inaccuracy of the CT system was calculated in terms of linear error in the $x$-, $y$-, and $z$-axis and orthogonal error in the $x$-$y$, $x$-$z$, and $y$-$z$ planes. The linearity and orthogonality are estimated by establishing a coherent local coordinate system for the CT-scanned Phantom and the tactile-probed Phantom. The coherent local coordinate system enabled measurements of their deviations. The linear error in the $x$- and $y$-axis is estimated by linear regression analyses. These are conducted by measuring distance deviations from hole-to-hole center points with respect to positions in the local coordinate system. The linear error in $z$-axis is calculated by evaluating the distance deviation of the hole-to-hole center points situated in the $x$-$z$ and $y$-$z$ planes. It is presumed that the effects of thresholding the CT-scanned Phantom are negligible due to the custom-designed geometrical features and the selected measuring principles.

The 3D thresholding error is analyzed by relying on a mean distance deviation which is estimated by comparing surfaces of the segmented SD head and the optically scanned dry skull.

The 3D modeling error is investigated by evaluating the absolute accuracy of the modeling software. Potential sources of error in AM are analyzed in terms of a mean distance deviation and a magnification error. The mean distance deviation is estimated by comparing the surfaces of the optically scanned implant and the nominal CAD implant. The magnification error is obtained by rescaling the optically scanned implant in the $x$-, $y$-, and $z$-axis scales to fit the nominal CAD implant.

In order to verify the results of the surface comparisons, mesh sensitivity analyses were performed to confirm an optimal mesh size for reference models of both the SD skull and the implant.

Finally, the combined standard uncertainty of medical imaging, 3D thresholding, 3D modeling, and AM is evaluated in terms of length-dependent (LD) and non-length-dependent (NLD) components. The worst-case error of the respective process is calculated by multiplying the LD error with the maximum length of the implant and summing the NLD error. Subsequently, the error of each process is summed to evaluate a cumulative error for the end-use implant.

3. Results

3.1. The CT Imaging Error

The custom-designed QA Phantom is shown in Figure 3. Table 4 shows the linear and orthogonal inaccuracies of the MSCT system measured with the QA Phantom. The magnification error represents the slope of the regression line, which is a dimensionless scale factor. The standard deviation of regression (Standard error) denotes the average distance between the measured deviations and the predicted deviations (regression line) in millimeters. The magnification error of the $z$-axis depicts the distance deviation of the hole centers with respect to the position in the local co-ordinate system. The standard deviation of the $z$-axis signifies the spread of the distance deviations from their average. The orthogonal error is described by the mean and standard deviations in degrees.
Table 4. Linear and orthogonal inaccuracies of the MSCT system.

<table>
<thead>
<tr>
<th>CT Protocol [No.]</th>
<th>Thresholding Value [HU]</th>
<th>Linearity</th>
<th>Orthogonality</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>X</td>
<td>Y</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>−0.000017</td>
<td>0.024945</td>
</tr>
<tr>
<td>−300</td>
<td>−0.000025</td>
<td>0.020638</td>
<td>−0.000005</td>
</tr>
<tr>
<td>−500</td>
<td>−0.000008</td>
<td>0.016195</td>
<td>−0.000053</td>
</tr>
<tr>
<td>1</td>
<td>−0.000032</td>
<td>0.149389</td>
<td>−0.000282</td>
</tr>
<tr>
<td>−300</td>
<td>−0.000100</td>
<td>0.144068</td>
<td>−0.000045</td>
</tr>
<tr>
<td>−500</td>
<td>−0.000089</td>
<td>0.130743</td>
<td>−0.000026</td>
</tr>
<tr>
<td>2</td>
<td>−300</td>
<td>−0.000022</td>
<td>0.026015</td>
</tr>
<tr>
<td>−500</td>
<td>−0.000030</td>
<td>0.022180</td>
<td>−0.000012</td>
</tr>
<tr>
<td>3</td>
<td>−300</td>
<td>−0.000145</td>
<td>0.115172</td>
</tr>
<tr>
<td>−500</td>
<td>−0.000214</td>
<td>0.129425</td>
<td>0.000051</td>
</tr>
<tr>
<td>4</td>
<td>−300</td>
<td>−0.000205</td>
<td>0.133742</td>
</tr>
<tr>
<td>Min.</td>
<td>−500</td>
<td>0.000100</td>
<td>0.016195</td>
</tr>
<tr>
<td>Max.</td>
<td>0</td>
<td>0.000008</td>
<td>0.149389</td>
</tr>
</tbody>
</table>
3.2. The 3D Thresholding Error

Verification results of the optically scanned dry skull are presented in Table 5. The results verified that the standard STL surface, generated by GOM ATOS (Professional 2016, GOM GmbH, Braunschweig, Germany) software, contained an appropriate mesh size. The results reached steady state values from 3rd to 6th iterations. The results of the standard STL mesh size are within 0.166 % of the mean distance and 1.094 % of the standard deviation for the final iteration.

Table 5. The mesh sensitivity analysis of the SD skull surface with respect to the segmented SD head surface.

<table>
<thead>
<tr>
<th>Iteration</th>
<th>Optically Scanned Dry Skull Mesh (STL) Properties</th>
<th>Mean Distance Deviation [mm]</th>
<th>Standard Deviation [mm]</th>
<th>Difference [%]</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Max. Edge Length [mm]</td>
<td>Number of Triangles [-]</td>
<td>Number of Points [-]</td>
<td>Mean</td>
</tr>
<tr>
<td>Standard</td>
<td>-</td>
<td>988577</td>
<td>502772</td>
<td>0.602</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>1147107</td>
<td>582037</td>
<td>0.620</td>
</tr>
<tr>
<td>2</td>
<td>0.5</td>
<td>1915921</td>
<td>966444</td>
<td>0.621</td>
</tr>
<tr>
<td>3</td>
<td>0.1</td>
<td>2940303</td>
<td>1478635</td>
<td>0.602</td>
</tr>
<tr>
<td>4</td>
<td>0.05</td>
<td>8809479</td>
<td>4413223</td>
<td>0.602</td>
</tr>
<tr>
<td>5</td>
<td>0.01</td>
<td>26428255</td>
<td>13222611</td>
<td>0.601</td>
</tr>
<tr>
<td>6</td>
<td>0.005</td>
<td>79284523</td>
<td>39650745</td>
<td>0.603</td>
</tr>
</tbody>
</table>

The thresholding error, including relatively negligible CT imaging error, of the verified standard mesh size is shown in Figure 4. The legend contains a range of the distance deviations. The red color code depicts the segmented surface above the optically scanned surface, the green color code denotes the segmented surface with no deviations, and the blue color code represents the segmented surface below the optically scanned surface. In addition, the legend contains a histogram to graphically represent the frequency of the deviations. In this case, the histogram shows asymmetric deviations that skew to the left (negatively skewed).
The thresholding error of the SD head for all protocols and thresholding values is shown in Figure 5. The vertical axis shows the mean distance between the segmented SD head as a function of the CT protocols and Hounsfield units compared to the optically scanned dry skull. The horizontal axis depicts four CT protocols. The legend contains a shade plot of the thresholding values. The error bars represent a standard deviation. A maximum allowable distance of 4 mm ensured that all the data points were within the computational range of the surface comparison. The mean distance represents a central arithmetic average of all points of the surface comparison. The standard deviation evaluates the spread of the mean distance.

According to the results shown in Figure 5, as the thresholding value increases, the mean distance decreases and the standard deviation increases. Estimating the optimal CT protocol and thresholding

**Figure 4.** The thresholding error of the segmented SD head on the dry skull with the verified mesh size.

**Figure 5.** The thresholding error of the segmented SD head on the dry skull with the verified mesh size.
value based on these relationships alone was deemed trivial. To this end, Table 6 lists the additional parameters of the surface comparison results. The area of valid distance is the reference surface area that is used for the comparison. The integrated distance is the inscribed volume of the surface comparison. Finally, the integrated ABS distance describes an absolute value of the integrated distance.

Table 6. Three-dimensional (3D) thresholding inaccuracies of the SD head intact with its soft tissue.

<table>
<thead>
<tr>
<th>CT Protocol [No.]</th>
<th>Thresholding Value [HU]</th>
<th>Mean Distance [mm]</th>
<th>Distance STD Deviation [mm]</th>
<th>Area of Valid Distance [mm²]</th>
<th>Integrated ABS Distance Per Protocol [mm³]</th>
<th>Integrated Distance Per Protocol [mm³]</th>
<th>Integrated ABS Distance [mm³]</th>
<th>Integrated [mm³]</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>250</td>
<td>0.725</td>
<td>0.901</td>
<td>222723.74</td>
<td>175534.97</td>
<td>210469.72</td>
<td>190894.02</td>
<td>541060.77</td>
</tr>
<tr>
<td>2</td>
<td>300</td>
<td>0.602</td>
<td>0.914</td>
<td>221392.80</td>
<td>148337.84</td>
<td>372419.02</td>
<td>190894.02</td>
<td>541060.77</td>
</tr>
<tr>
<td>3</td>
<td>500</td>
<td>0.144</td>
<td>1.045</td>
<td>207069.17</td>
<td>48542.21</td>
<td>139697.03</td>
<td>139697.03</td>
<td>541060.77</td>
</tr>
<tr>
<td>4</td>
<td>300</td>
<td>0.200</td>
<td>0.884</td>
<td>222958.92</td>
<td>52302.18</td>
<td>149661.33</td>
<td>434771.85</td>
<td>634867.67</td>
</tr>
<tr>
<td>5</td>
<td>500</td>
<td>~0.060</td>
<td>1.026</td>
<td>210867.91</td>
<td>2941.39</td>
<td>137802.07</td>
<td>137802.07</td>
<td>541060.77</td>
</tr>
<tr>
<td>Min.</td>
<td>250</td>
<td>~0.084</td>
<td>0.834</td>
<td>202369.41</td>
<td>864.35</td>
<td>129962.70</td>
<td>129962.70</td>
<td>543835.75</td>
</tr>
<tr>
<td>Max.</td>
<td>500</td>
<td>0.753</td>
<td>1.082</td>
<td>224349.47</td>
<td>658176.30</td>
<td>433835.75</td>
<td>433835.75</td>
<td>546576.67</td>
</tr>
</tbody>
</table>

The results of the qualitative and quantitative analyses deduced that CT Protocol 1 and 300 HU were the optimal independent parameters for imaging the SD head because they created surfaces that were closest to the optically scanned dry skull. Subsequent thresholding error is recorded as 0.602 mm.

3.3. The 3D Modeling Error

Figure 6 shows a render of the digitally designed implant attached to the segmental resection of the right zygomatic bone. In this case, the 3D modeling error is represented by the numerical resolution of the modeling software, which is recorded as 0.0001 mm.

Figure 6. The computer-aided design (CAD) implant attached to the segmental resection of the zygomatic bone.
3.4. The Additive Manufacturing Error

Figure 7 shows the additively manufactured implant with and without the support structures. Verification results of the reference (nominal CAD) implant are shown in Table 7.

**Figure 7.** The additively manufactured titanium implant with the supports (right) and without the supports (left).

**Table 7.** The mesh sensitivity analysis of the nominal CAD implant surface with respect to the optically scanned implant surface.

<table>
<thead>
<tr>
<th>Iteration</th>
<th>CAD Implant Mesh (STL) Properties</th>
<th>Mean Distance [mm]</th>
<th>Standard Deviation [mm]</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Max. Edge Length [mm]</td>
<td>Number of Triangles [-]</td>
<td>Number of Points [-]</td>
</tr>
<tr>
<td>Original</td>
<td>-</td>
<td>13368</td>
<td>6584</td>
</tr>
<tr>
<td>1</td>
<td>0.5</td>
<td>13672</td>
<td>6736</td>
</tr>
<tr>
<td>2</td>
<td>0.4</td>
<td>371204</td>
<td>185490</td>
</tr>
<tr>
<td>3</td>
<td>0.3</td>
<td>526942</td>
<td>263351</td>
</tr>
<tr>
<td>4</td>
<td>0.2</td>
<td>1015062</td>
<td>507408</td>
</tr>
<tr>
<td>5</td>
<td>0.1</td>
<td>3660090</td>
<td>1829738</td>
</tr>
</tbody>
</table>

The mesh sensitivity analysis verified that the fourth iteration of the analysis reached a steady state value of 0.042 mm and 0.062 mm for the mean distance and standard deviation of the AM error, respectively. Figure 8 shows the AM error color plot of the end-use implant on the nominal CAD implant including a histogram, which shows asymmetric and positively skewed deviations.

A maximum allowable distance of 0.5 mm ensured that all the data points of the surfaces were within the computational range. The grey regions of the color plot depict the surface data of the end-use implant that were not captured by the optical 3D coordinates measuring system. These regions represent locations where the reference points were placed for triangulating the surface geometry. The removal of the reference point surfaces ensured that only the implant surfaces were considered for the surface comparison.
Figure 8. The additive manufacturing error of the end-use implant on the nominal CAD implant with the verified mesh size.

The magnification error of the additively manufactured implant was evaluated through a series of steps. Firstly, the lattices of the optically scanned implant were removed which eliminated small dimensional features. Then, the base implant was re-oriented to axially align its longest dimension with the global co-ordinate system. Finally, the absolute integrated distance of the base implant was minimized. The removal of lattices decreased the computational time and outlier contribution. The resultant magnification error is presented in Table 8. The compensation of the magnification error decreased the absolute integrated distance by a factor of 3.5 while the mean distance and standard deviations were also decreased.

Table 8. The magnification error of the additively manufactured end-use implant with reference to the nominal CAD implant.

<table>
<thead>
<tr>
<th>Base Implant Orientation</th>
<th>Axis</th>
<th>Magnification Error [-]</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>X</td>
<td>0.0005</td>
</tr>
<tr>
<td></td>
<td>Y</td>
<td>0.0010</td>
</tr>
<tr>
<td></td>
<td>Z</td>
<td>0.0150</td>
</tr>
</tbody>
</table>

3.5. Uncertainty and Cumulative Error of the Implant

Uncertainty and error budget of the end-use implant is presented in Table 9.
Figure 9 shows the way in which cumulative error increases at each process step. Uncertainty of the reference measurements was significantly smaller than the uncertainty of the process chain.

The maximum magnification error of CT imaging is 0.000359 which resulted from CT Protocol 1 and the z-axis of the MSCT system, as shown in Table 4. CT Protocol 1 was also used to image the SD head which was used for the implant modeling process. To calculate uncertainty of this error, it is estimated to have a rectangular probability distribution. Therefore, it contributes to the LD uncertainty of CT imaging by a value of 0.000359/√3 = 0.0000289 mm. Likewise, the maximum mean orthogonal error of
CT imaging is 0.114° (0.00199 rad.) which resulted from CT Protocol 1 and the x-z planes, as shown in Table 4. Uncertainty for the orthogonal error is 0.114°/√3 = 0.066° (0.00115 rad.). Considering that the angle is small, the effect of the orthogonal error for a length is the length multiplied by the angle in radians. Thus, it contributes to the LD uncertainty with a value of 0.00115. The combined LD uncertainty for CT imaging is calculated to be 0.00117 by combining the LD uncertainty of the magnification and orthogonal error.

Similarly, the maximum standard deviation is 0.0779 mm for the magnification error in Table 4. It also occurs in the z-axis of the MSCT system. The standard deviation of the orthogonal error was neglected due to its insignificant contribution to the combined standard uncertainty. Thus, the NLD standard uncertainty of CT imaging is 0.0779 mm.

Finally, the worst-case CT imaging error is 0.191 mm for the 56.852 mm long end-use implant.

The average thresholding error of the volume-based segmentation is 0.602 mm, which resulted from CT Protocol 1 and 300 HU, as shown in Table 6. In addition to CT Protocol 1, this thresholding value was used to segment the SD head which was used for the implant modeling process. If the length measurement of the implant involves two boundary surfaces, then the thresholding error is 2 × 0.602 mm = 1.2 mm. Uncertainty of the thresholding error is 0.602 mm assuming that it is not larger with a probability of 67%. Thus, it contributes to the NLD uncertainty with a value of 0.602 mm.

Consequently, the thresholding error is 0.602 mm for one boundary surface and 1.2 mm for two boundary surfaces of the implant.

The worst-case error of 3D modeling is 0.0001 mm/2 = 0.00005 mm that has a NLD uncertainty of 0.00005 mm/√3 = 0.0000289 mm.

The maximum magnification error of additive manufacturing is 0.015 that occurred in the z-axis of the implant, as shown in Table 8. It contributes to the LD uncertainty with a value of 0.015/√3 = 0.00866. The NLD uncertainty is 0.061 mm, which resulted from the standard deviation of the AM error.

Finally, the worst-case error of additive manufacturing is evaluated to be 0.914 mm for the maximum length of the end-use implant.

The combined uncertainty represents the positive square root of the combined variances that is expressed by the law of propagation of uncertainty. The implant error for each step of the process chain is calculated according to the following Equation (2):

$$IE_{\text{process}} = \sqrt{u_{P_{\text{NLD}}}^2 + (u_{P_{\text{LD}}} \times L)^2}$$

(2)

where $IE_{\text{process}}$ depicts the implant error as combined uncertainty of a process in millimeters, $u_{P_{\text{NLD}}}$ and $u_{P_{\text{LD}}}$ denote the NLD [mm] and LD uncertainty associated with the process, and $L$ represents the maximum length of the end-use implant in millimeters.

Finally, the worst-case cumulative error of the implant totals to a value of 1.71 mm for one boundary surface and 2.31 mm for two boundary surfaces. The combined standard uncertainty ($u_c$) of the end-use implant encompassing all processes is calculated for one boundary surface condition according to the following Equation (3):

$$u_c(IE_{\text{process}}) = \sqrt{U_{\text{NLD}}^2 + (U_{\text{LD}} \times L)^2} = \sqrt{0.610^2 + (0.00874 \times 56.852)^2} = 0.787 \text{ mm}$$

(3)

4. Discussion

This study investigated inaccuracies involved in the process chain of procuring an end-use implant. The increase in cumulative error at each step in the process chain must be considered for preplanning a surgical procedure to avoid ill-fitting implants that may cause complications during surgery.

4.1. CT Imaging

The worst-case CT imaging error emerged in the z-axis of the MSCT system and generated the second least error value compared to the other processes. The LD uncertainty of all four CT protocols
ranged at an insignificant scale. The evaluation of the hole-to-hole center points of the QA Phantom ensured that the thresholding effects were minimized when estimating the CT imaging error. This is because the center points of the Phantom holes remain unchanged regardless of the thresholding effect on the hole diameter. The homogeneous material of the Phantom also complemented the minimization of the thresholding effects to compute the CT imaging error. Moreover, the NLD uncertainty of CT imaging represented noise associated with the hole surfaces that were used to predict the center points. It consisted of components such as the resolution and non-linearity. This estimated error can also be negligible for most craniomaxillofacial surgical procedures.

Though the overall CT imaging error can be negligible, the impact of CT reconstruction kernel, image thickness, and Hounsfield units on the QA Phantom was evaluated quantitatively for practitioners seeking optimal results. Following a conservative conclusion, the J30S kernel with the 0.5 mm image thickness and –500 HU yielded the least amount of uncertainty for CT imaging. This is because the J70H kernel was more prone to the image noise than the J30S kernel.

The results indicated that the used MSCT system is a good candidate for AM applications, in particular for the development of an end-use implant, owing to its high accuracy. Though the calibration of the HU number is standardized in MSCT systems [40], a MSCT system from a different vendor does not have identical CT protocols and applied x-ray spectra due to different x-ray tubes, varying beam filtration, beam shaping filters, etc. in different CT scanner models. Furthermore, the soft tissue and bone have HU numbers which are outside of the calibration point of water (especially bone) and typically span a wide HU range even with a fixed scan protocol. To this end, linear and orthogonal errors calculated in this study can vary to a reasonably small extent [69]. Further, the reported error may deviate to a much larger extent when using a different type of CT system, e.g., a CBCT system, because of a lack of standardization in HU number and different CT protocols [40]. Nevertheless, a CBCT system might be an alternative when the manufacturer’s recommended protocol is used [70], but more research is needed. In principle, the developed methodology for error analysis can be extended to any type of CT system and its protocols using the tailor-made Phantom.

4.2. 3D Thresholding

The thresholding error was one of the most significant errors compared to the other processes. The thresholding error represents the edge location error considering that the LD and NLD uncertainty of the MSCT system were relatively low. This means that the volume-based segmentation overestimated the bone surfaces of the DICOM images to the STL format with a value of 0.602 mm outside the true surface. Assuming a craniomaxillofacial implant case that involves a hole into which an implant is intended to fit, then the thresholding error for the size of the implant would be $2 \times 0.602 \text{ mm} = 1.2 \text{ mm}$. Practitioners should consider that this error enables the correction of a systematic effect for the implant. The thresholding error propagates to the implant because the implant was modeled as a negative of the segmented SD head.

The variations in the thresholding value generated significantly altered SD skull geometry. The results revealed that the data loss of the skull surface was inversely proportional to the area of valid distance. Similarly, the staircase effect or the blocky coarse surfaces of the skull were directly proportional to the standard deviation, and the artefacts were directly proportional to the negative values of the integrated distance.

When excluding the J70H kernel that included the coarse skull surface comprising of data loss, the staircase effect (blocky data) with sharp edges, and a higher number of artefacts, the selection of an optimal CT reconstruction kernel delimited to the J30S kernel. Since CT Protocol 1 involved lower image thickness of 0.5 mm, it captured a slightly higher amount of the segmented SD head data than CT Protocol 3.

Though the 300 HU generated a similar and uniformly fine surface data than the 250 HU, it also contained a lower amount of inscribed volume. Comparatively, the 500 HU generated excessive artefacts.
Even though the higher sharpness enabled by the J70H kernel can be beneficial for bony structures that require a higher spatial resolution, it is also associated with accentuated noise. The optically scanned dry skull was mainly composed of smooth outer bone that was within the line of sight of the optical 3D coordinates measuring system. Thus, the results favored the J30S reconstruction kernel that generated smoother bone surfaces than the J70H kernel. The J70H kernel also exhibited properties that were more prone to the partial volume effect [39] than the J30S kernel. Though the attenuation coefficient and density of distinct matter, i.e., bone (+1000 HU), water (0 HU), and air (-1000 HU), are identifiable, the partial volume effect allocates voxels as soft tissue instead of bone at the bone-to-soft tissue boundary which results in voids and a loss of data.

The results indicated that the essential step of 3D thresholding remains as one of the critical factors that limit the accuracy of additively manufactured medical constructs. More specifically, the thresholding error was close to the imaging voxel size which has also been reported in the past using four commercially available software packages [71]. Though the use of different software can generally yield accurate geometry, discrepancies can be significant between software packages mainly due to a lack of identical array of control settings and the embedded algorithm for DICOM-to-STL conversion [71]. For example, Osirix MD, which was used in this study, is most probably related to an exact geodesic path algorithm, Dijkstra algorithm, and the fast marching method [71], whereas the Mimics software (Materialise NV, Leuven, Belgium), which is also approved by the United States FDA [72], is embedded with the adapted marching cubes algorithm with the partial volume effect for 3D imaging [73].

4.3. 3D Modeling

Uncertainty of 3D modeling was estimated to be the most insignificant in the process chain. It was derived from the arithmetic decimal places (resolution) of the software which represented the STL chordal tolerance [74]. Likewise, the rounding error that contributed to the implant budget was still the most insignificant even after assuming a worst-case scenario. The resolution of several CAD software is frequently characterized as the absolute or relative accuracy, which can be changed to a certain extent for most CAD packages.

As expected, the results indicated that the error originating from the modeling process was minimal. This error is expected to remain minimal when using different CAD software. Though the error of the process is minor, it still propagates error originating from the imaging and thresholding processes because the implant was modelled using the geometric surfaces and projections of the imaged and segmented bone. Top and bottom surfaces of the implant base are identical to the surfaces of the zygomatic bone.

4.4. Additive Manufacturing

The results indicated that uncertainty of the AM process can be the greatest in the process chain of procuring an end-use implant. Though the NLD uncertainty was reasonably low, the LD uncertainty can be the main contributor to the error budget.

The NLD uncertainty was close to two-layer thicknesses of the manufacturing process. This is most likely caused by the conglomeration of semi-fused particles (size: ~32–60 um) to the outer surface of the implant that was in contact with the surrounding powder owing to a high platform temperature of 200 °C during the build. Further, the stair-case effect [75] induced by the freeform geometry containing angled, curved, or lofted shapes (complex shapes) of the implant can also contribute to the NLD uncertainty. Considering that the machine was calibrated, and standard printing parameters were used for the standard material profile (i.e., particle size distribution and shape), this error can describe the capability of the used machine to manufacture shapes as designed when followed by a standard heat treatment process and manual support removal process.

Microstructural and mechanical anisotropy of additively manufactured metals, particularly titanium, is well understood [76–80]. The results of this study showed that there is also a significant
amount of dimensional anisotropy. Though the maximum magnification error was generally low (1.5%) considering also the heat treatment and manual support removal, there was a significant difference of 97% and 93% for the x- and y-scale error compared to the maximum z-scale error of the implant. This is most probably caused by the cumulative effect of localized microstructural anisotropy of the freeform geometry resulting from the layer-thickness-dependent z-scale and the build-mechanism-dependent x- and y-scales of the build.

The metal PBF process is inherent to high temperature distributions as the impinging laser energy melts the metal powder. Re-solidification is highly dependent on the heat transfer capability of the melt pool, and the material is prone to residual stresses and shrinkage due to contractions caused by the physical processes. To this end, support structures are necessary for heat dissipation and anchoring the part to the platform.

Shrinkage can vary as a function of print inclination [75]. Freeform surfaces of the implant can have a fluctuating print inclination resulting in non-linear shrinkage because the microstructural morphology shifts dependency between the melt pool overlaps and the layer overlaps for resisting contractions. Since the magnification error observed in the 3D scales of the implant was positive, it means that the pre-processing software overestimated the shrinkage compensation for the standard printing parameters and material. In addition to automatic shrinkage compensation, most pre-processing software allow users to manually compensate shrinkage/expansion in the 3D scale for each individual part before manufacturing.

The maximum magnification error was evaluated through multiple sources of evidence, i.e., conducting caliper measurements digitally and using a length measurement machine, however the additional methods were not reported to simplify the study.

The LD uncertainty describes the maximum magnification error of the end-use implant including the manufacturing, heat treatment, and the support removal processes. The heat treatment process is a standard procedure to cure parts and remove residual stresses. Though the manual support removal process was conducted by an expert in the field using standard tools, it may benefit from a decrease in uncertainty of human error when automated. Further, if the critical geometries are not directed on to the z-axis of the build platform that generally contain down-skin surfaces requiring support structures, the LD uncertainty can be significantly reduced by a factor of 15.

The results confirmed that the error in the metal AM process was dictated by the z-axis of the build which was also reported in a previous study using five different metal AM machines [59]. In the present study, differences of up to 80% were observed for x- and y-scales compared to the z-scale of the build. The z-scale of an AM machine is governed by the layer thickness, whereas the x- and y-scales are dependent on the size of the laser/electron beam spot and the resolution of its steering mechanism, e.g., galvanometer-based mirrors or deflection coils. Though different metal AM machines can generally create accurate geometries [59], inconsistencies remain mainly due to a lack of identical build mechanisms (i.e., laser beam, electron beam, build environment, scanning system, etc.) and process parameters (i.e., source power, layer thickness, scanning speed, scanning strategy, inert gas flow, substrate temperature, powder particle profile, etc.). When a different metal AM machine is used with the manufacturer’s recommended process parameters, the reported error may deviate with a fairly good precision of <0.14 mm, excluding the geometries that are prone to support removal error situated on the z-axis [59].

4.5. Uncertainty and Cumulative Error

Table 10 lists accuracy measurement studies [81–92] of additively manufactured constructs using various anatomical models, imaging medium and modality, measurement principles, AM technologies, and measurement instruments. Most studies have used CT imaging technology to image a dry skull. Though imaging a dry skull can account for an approximated error analysis, this is almost never the case in clinical practice and can yield unrealistic values. Several studies have imaged the skull in water to simulate close to realistic cases. However, the tissue surrounding the bone can certainly vary in
density. Due to ethical concerns, only few studies have used human skull with and without soft tissue for accuracy analysis, but they can still be limited to manual contact measurements. In fact, majority of the studies have employed a contact measurement principle of repeated measurements between anatomical landmarks using calipers to estimate the overall error of the process chain. However, this principle can easily induce error in pinpointing the exact location of the freeform anatomical surface points in multiple models for linear measurements. This error was mitigated by introducing artificial reference features, i.e., grooved holes [84] or markers [58,90], into the 3D model of the specimens before manufacturing. Though this has provided an ease of pinpointing the regions of interest on the freeform geometry, the measurement principle is often laborious and is limited in terms of number of measurements. Further in literature, human error has been mitigated using CMM for data collection.

In the present case study, the prescribed methodology aims to overcome the limitations. Principles of GUM are exploited to develop a systematic approach to evaluate overall cumulative error propagating from each process step. The use of animal model simulates a realistic case that also enables tissue removal and measurements of the dry skull to obtain a gold standard. Contact measurements are conducted on a tailor-made QA Phantom through controlled experiments that take advantage of the impeccable accuracy of the tactile-CMM. Optical measurements are conducted on freeform geometry of the dry skull and the additively manufactured implant without geometric manipulation. Subsequent mean difference is evaluated through an adequate number of measurements in the order of 500,000. The method also traces the measurements to the SI units using calibrated instruments. The prescribed methodology has potential of being a standard quality control and assurance protocol that can be applicable to various imaging, thresholding, modeling, and AM processes.

The LD and NLD components of combined uncertainty can yield a reliable measure of possible error of the end-use implant including resolution, non-linearity, and magnification error of CT imaging, segmentation error of 3D thresholding, resolution of 3D modeling, and resolution, non-linearity, and magnification error of AM including heat treatment and support removal. Further, uncertainty in terms of LD and/or NLD can also be calculated for each individual process step. The LD component allows for evaluating error as a function of length for the implant.

Worst-case cumulative error is calculated for the maximum length of the freeform implant which takes into account intricate features of the shape as a result of each process step. Evidently, a smaller dimension would yield a smaller error. Further, boundary surface plays a critical role to the length measurement of the implant which is dictated by the NLD component of thresholding.

Uncertainty results for both one and two boundary surfaces are comparable to those reported in the literature. However, when considering the worst-case cumulative error, in particular for two boundary surfaces, the error peaks compared to the reported mean differences.

4.6. Verification

The measurement instruments were calibrated using gauge blocks. Because uncertainty of the calibrated measurement instruments was insignificant compared to the combined uncertainty of the implant, it confirmed that the required accuracy for the measurements was adequate. The mesh sensitivity analyses verified the correct mesh size and the subsequent vertices that were essential to the surface comparison algorithm. This was confirmed by decreasing the maximum edge length of the mesh until steady state results were achieved. The reduction in the mesh size increased the number of triangles and their subsequent vertices. The number of mesh triangles and vertices was directly proportional to the number of measurements that were conducted for the surface comparison analysis. In this case, though the chordal length of the mesh controlled the accuracy of the intended geometry compared to the original geometry, the mesh size controlled the results of the surface comparison analysis.
Table 10. Cumulative error analysis of this study compared with literature.

<table>
<thead>
<tr>
<th>Source</th>
<th>Anatomical Model</th>
<th>Imaging Medium</th>
<th>Imaging Modality</th>
<th>Measurement Principle</th>
<th>ISO/ASTM Additive Manufacturing Technology</th>
<th>Absolute Mean Difference (mm) *</th>
<th>Relative Mean Difference (%) *</th>
<th>Measurement Instrument(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Present study</td>
<td>Skull and end-use</td>
<td>Skull</td>
<td>CT and optical 3D coordinates measuring system</td>
<td>Principles of GUM applied to a tactile-probed Phantom for CT imaging, surface comparison for 3D thresholding, software for 3D modeling, and surface comparison for additive manufacturing</td>
<td>Titanium PBF</td>
<td>$u_c = 0.610 + (0.0087) L$</td>
<td>$u_c (L = 56.852; 1 BS) = 0.787$</td>
<td>CMM and optical 3D coordinates measuring system</td>
</tr>
<tr>
<td>Barker et al. [81]</td>
<td>Skull</td>
<td>Water</td>
<td>CT</td>
<td>Repeated measurements between anatomical landmarks</td>
<td>Vat photopolymerization</td>
<td>$1.8 (0.10-4.62)$</td>
<td>0.6-3.7</td>
<td>Caliper</td>
</tr>
<tr>
<td>Chang et al. [82]</td>
<td>Skull</td>
<td>With and</td>
<td>CT and caliper</td>
<td>Repeated measurements between anatomical landmarks</td>
<td>Binder jetting</td>
<td>$1.54 (0.00-8.70)$</td>
<td>3.4 (0.00-23.6)</td>
<td>Caliper</td>
</tr>
<tr>
<td>Choi et al. [83]</td>
<td>Skull and mandible</td>
<td>Air</td>
<td>CT</td>
<td>Repeated measurements between anatomical landmarks</td>
<td>Vat photopolymerization</td>
<td>$0.62 \pm 0.35 (0.04-1.22)$</td>
<td>0.9 ± 0.5 (0.0-1.9)</td>
<td>Caliper</td>
</tr>
<tr>
<td>El-Katatny et al. [84]</td>
<td>Skull</td>
<td>Mandible</td>
<td>With soft tissue</td>
<td>Repeated measurements between artificial grooved holes on anatomical landmarks</td>
<td>Material extrusion</td>
<td>$0.11 \pm 0.05$</td>
<td>$0.24 \pm 0.16$</td>
<td>Digital caliper</td>
</tr>
<tr>
<td>Ibrahim et al. [85]</td>
<td>Mandible</td>
<td>Air</td>
<td>CT</td>
<td>Repeated measurements between anatomical landmarks</td>
<td>Binder jetting Material jetting Polymer PBF</td>
<td>$0.91 \pm 0.59 (0.44-2.52)$</td>
<td>$1.8 \pm 1.3 (0.4-4.4)$</td>
<td>Digital caliper and a dial test indicator attached to an electric milling machine</td>
</tr>
<tr>
<td>Nizam et al. [86]</td>
<td>Skull</td>
<td>Air</td>
<td>CT</td>
<td>Repeated measurements between anatomical landmarks</td>
<td>Vat photopolymerization</td>
<td>$0.23 \pm 1.37$</td>
<td>$0.08 \pm 1.25$</td>
<td>Digital caliper</td>
</tr>
<tr>
<td>Ono et al. [87]</td>
<td>Skull</td>
<td>Water</td>
<td>CT</td>
<td>Repeated measurements between anatomical landmarks</td>
<td>Vat photopolymerization</td>
<td>-</td>
<td>$&lt;3$</td>
<td>Digital caliper</td>
</tr>
<tr>
<td>Osman et al. [88]</td>
<td>Dental implant</td>
<td>-</td>
<td>-</td>
<td>Surface comparison</td>
<td>Ceramic vat photopolymerization</td>
<td>$0.1 \text{†}$</td>
<td>-</td>
<td>Optical scanner</td>
</tr>
<tr>
<td>Pinto et al. [54]</td>
<td>Phalanx</td>
<td>Air</td>
<td>CT</td>
<td>Voxel and surface comparison for each modified building parameter, i.e., acquisition, segmentation, triangulation, printing, and infiltration</td>
<td>Binder jetting</td>
<td>$-2.71$ and $3.03 \text{†}$</td>
<td>-</td>
<td>CT</td>
</tr>
<tr>
<td>Salmi et al. [58]</td>
<td>Skull</td>
<td>Air</td>
<td>CT</td>
<td>Repeated measurements between artificial reference points on anatomical landmarks</td>
<td>Polymer PBF Material jetting Binder jetting</td>
<td>-</td>
<td>$0.79 \pm 0.26$ and $0.80 \pm 0.32$</td>
<td>CMM</td>
</tr>
<tr>
<td>Silva et al. [89]</td>
<td>Skull</td>
<td>Air</td>
<td>CT</td>
<td>Repeated measurements between anatomical landmarks</td>
<td>Polymide PBF Binder jetting</td>
<td>$1.25 \pm 0.83 (0.19-2.94)$</td>
<td>2.1 ± 1.4 (0.2-3.9)</td>
<td>Digital caliper</td>
</tr>
</tbody>
</table>
Table 10. Cont.

<table>
<thead>
<tr>
<th>Source</th>
<th>Anatomical Model</th>
<th>Imaging Medium</th>
<th>Imaging Modality</th>
<th>Measurement Principle</th>
<th>ISO/ASTM Additive Manufacturing Technology</th>
<th>Absolute Mean Difference (mm) *</th>
<th>Relative Mean Difference (%) *</th>
<th>Measurement Instrument(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Taft et al. [90]</td>
<td>Skull</td>
<td>Air</td>
<td>CT</td>
<td>Repeated measurements between artificial reference points on anatomical landmarks</td>
<td>Vat photopolymerization</td>
<td>0.09 ± 0.07 (0.00–0.29)</td>
<td>-</td>
<td>CMM</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>x-axis</td>
<td>0.12 ± 0.10 (0.00–0.33)</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>y-axis</td>
<td>0.41 ± 0.15 (0.09–0.70)</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>van Eijnatten et al. [91]</td>
<td>Mandible</td>
<td>Air</td>
<td>CT</td>
<td>Repeated measurements between artificial reference points on anatomical landmarks</td>
<td>Binder jetting</td>
<td>0.88 ‡‡</td>
<td>-</td>
<td>Optical 3D scanner</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>CT-derived</td>
<td>1.73 ‡‡</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Magnetic resonance-derived</td>
<td>0.46 ‡‡</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Optical scanning-derived</td>
<td></td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Waitzman et al. [92]</td>
<td>Skull</td>
<td>Air</td>
<td>CT</td>
<td>Repeated measurements between artificial reference points on anatomical landmarks</td>
<td>-</td>
<td>-</td>
<td>0.9 (0.1–3.0)</td>
<td>Caliper</td>
</tr>
</tbody>
</table>

Note.—uc = combined standard uncertainty, L = length, BS = boundary surface; * Cited as mean or mean ± standard deviation with ranges in parentheses, unless otherwise stated by a function; † Root mean square estimate; ‡ Max. and min observed; § Reported differences in two models; ‡‡ Reported difference in two measurement series; ‡‡‡ 95th percentiles of the measured values with positive and negative differences.
4.7. Craniomaxillofacial Implants

The cumulative error pertinent to the process chain of an end-use implant can prevent patient-specific treatment and restoration owing to ill-fitting implants and complications in surgeries [93].

Though the cumulative error evaluated in the present study can be adequate for most craniomaxillofacial surgical procedures [72,94,95], a safety margin of 2.31 mm (4.1% relative) should be respected considering the worst-case error. Practitioners should also consider that the thresholding error and AM magnification error can be corrected through redesign of the implant to significantly improve the results in the future studies.

For example, the thresholding error (0.602 mm) can be compensated in the implant dimensions during the modeling process. The dominating AM magnification error in the z-axis can be manually compensated in the implant model using the pre-processing software before manufacturing, which may reduce the magnification error by a factor of 15.

Further, practitioners should avoid placing critical geometries of the implant on the z-axis, particularly to the down-skin surfaces that require addition and removal of support structures.

In addition to using standard AM process parameters and heat treatments, practitioners need to consider automating the support removal process in case it unavoidably affects critical geometries. This step should be considered before the AM process to ensure that the path of the machining tool is not obstructed by the implant material.

Through an evaluation of an increasing number of segmented anatomical parts with respect to a calibrated CT system, the thresholding error can be mitigated using regression and neural network models as a function of the HU number and software. Apart from thresholding that was performed in this study to standardize the results, practitioners should also consider manual thresholding alternatives. Though subjective and time consuming, manual thresholding may yield better STL models than the default thresholding values of the segmentation software [39]. Though not yet commercially available, advanced thresholding methods [72] are also expected to decrease the leading significance of the segmentation error. The CT images of the SD head are made available to promote future research [67].

Practitioners that require a high accuracy for the implants, e.g., the orbital floor implants [96], should respect the recommended safety margin. The prescribed correction measures can be taken if the intended processes and equipment are identical to this study or after implementing the applied methodology with the relevant processes and equipment as a quality control and assurance protocol. A lack of lower uncertainty (higher accuracy) can also be compensated with precise bone resection using additively manufactured cutting guides [97].

Moreover, it should be considered that it may not always be absolutely imperative to design the implant to fit perfectly at the initial phase. The perforation of the lingual mucosa is one of the complications related to the patient-specific implant. If the lingual foil of the implant scaffold is too high, it is prone to perforate the lingual mucosa during the postoperative healing [98]. For example, the extraction of teeth in the osteotomy line will lead to a resorption and further, remodeling of the marginal alveolar crest, which has to be considered during the planning of the implant.

A successful implementation of AM from radiologic images is a multidisciplinary task [95]. Achieving the most benefit from this technique requires a very close co-operation between all experts involved in the process.

5. Conclusions

Implementing additive manufacturing (AM) in a clinical setting, particularly for the craniomaxillofacial implants, requires tools that are sufficiently precise and accurate to conform to the patient-specific and freeform anatomical geometries considering the entire process chain. To this end, this study develops and evaluates a novel procedure for determining the cumulative spatial inaccuracies in the whole process chain of procuring an end-use implant from medical imaging, 3D
thresholding, 3D modeling, and AM, as opposed to partial analyses of the process chain found in the literature. The focus on each process step enables a quality control and assurance protocol that allows for correction measures to further enhance the accuracy of an implant that may prevent treatment and restoration. The study enables practitioners and clinical researchers to establish and substantiate medical practices within the scope of an appropriate accuracy for clinical use.


**Funding:** This research was funded by the Metrology for Additively Manufactured Medical Implants (MetAMMI) project, which received funding from the EMPIR program co-financed by the Participating States and from the European Union’s Horizon 2020 research and innovation program. The work is also part of the Academy of Finland Flagship Program, Photonics Research and Innovation (PREIN), decision 320168.

**Conflicts of Interest:** The authors declare no conflict of interest.

**References**


51. BIPM; IEC; IFCC; ILAC; ISO; IUPAC; IUPAP; OIML. *Evaluation of Measurement Data—Guide to the Expression of Uncertainty in Measurement; Joint Committee for Guides in Metrology: Sèvres Cedex, France, 2008.*


