Uncertainty Determination for Dimensional Measurements with Computed Tomography

Kim Kiekens¹,², Tan Ye¹,², Frank Welkenhuyzen², Jean-Pierre Kruth², Wim Dewulf¹,²

¹ Group T – Leuven University College, KU Leuven Association
A. Vesaliusstraat 13, 3000 Leuven, Belgium, e-mail: wim.dewulf@group-t.be

² Katholieke Universiteit Leuven, Mechanical Engineering Department, Division PMA
Celestijnenlaan 300B, 3001 Heverlee, Belgium, e-mail: jean-pierre.kruth@mech.kuleuven.be

Abstract
Due to the availability of increased computational power combined with better X-ray source and detector technologies, the obtainable resolutions now allow to use CT-machines not only for material and medical science, but also for dimensional metrology. However, the transition from visualization to measuring requires calibration steps to ensure that measurements are traceable to the unit of length. There is still a lack of proper standards to calibrate CT machines and to check their performance for dimensional metrology. In addition, there is still room for a better systematic approach for measurement uncertainty determination. Therefore, this paper presents a method to determine the uncertainty of dimensional measurements with industrial CT machines based on the ISO-GUM. Subsequently, some of the uncertainty factors are quantified based on a series of systematic measurements.

Keywords: Uncertainty evaluation, computed tomography, dimensional metrology, traceability

Introduction

Uncertainty evaluation
Determining the uncertainty of dimensional CT measurements is challenging, due to the numerous factors influencing the CT accuracy and uncertainty; e.g. workpiece, measurement procedure, scanning parameters, user, … No correct statement of the measurement uncertainty of those measurements exists, since CT systems are not traceable to the unit of length. Some studies have estimated the uncertainty calculation of CT measurements, based on simulations [1], experiments [2, 3, 4], or a combination of both [5]. Schmitt et al. conclude that the procedure using calibrated workpieces is the most promising to investigate measurement uncertainty for CT. Analytical studies of the measurement uncertainty of CT systems are mostly based on the procedure described in the guideline ISO/TS 15530-3 [6]. The expanded uncertainty is divided in three main uncertainty contributors: the standard uncertainty due to the calibration, the standard uncertainty due to the measurement procedure and the standard uncertainty resulting from the workpiece. Additionally, a bias is added to take the systematic errors into account. Subcategories of those main uncertainty contributors are rarely made [7].

After a description of the basic measurement procedure, a first main part of this paper presents a theoretical study of the CT measurement uncertainty, based on the described measurement procedure and an analytical equation based on the ISO-GUM [8]. In a second part the main uncertainty contributors are identified and illustrated with some examples, based on experimental measurements with different test objects.

Measurement procedure including voxel size and edge correction calibration (Figure 1)
In the first step of a CT measurement, some hundreds or even thousands of 2D X-ray images of the object are taken from different angular positions in the CT machine. In the second step, those images are reconstructed into a 3D voxel model, e.g. using the software CTpro. Afterwards, the data analysis can be performed, and dimensional measurements can be extracted.

![Data captation Reconstruction Data analysis](image)

Figure 1: Different steps of a CT measurement

Before the resulting length $L$ of a CT measurement is known, the voxel model has to be rescaled for the correct voxel size. The resolution (voxel size) is mainly determined by the position of the workpiece between the X-ray source and the detector (i.e. magnification). The closer the object is to the source, the better is the resolution and the smaller are the voxels. Voxel size calibration can be carried out in different ways. A first, commonly used option is to use a calibration object. This calibration object is either scanned simultaneously with the measurement object, or scanned separately at the same position of the magnification axis. The rescale factor is calculated as the ratio of the CT and the reference measurement of the calibration length (e.g. the distance between the center points of two spheres). Another option is to use a known distance on the workpiece itself to rescale the other features of the CT model. After rescaling for the correct voxel size, an edge detection step is needed making a segmentation between material voxels and background voxels based on a locally or globally optimized grey value selection. In this paper, an advanced (locally adaptive) edge detection method has been used (VG Studio Max 2.1). The result of this step is a 3D model that allows to measure the different features of the object.

The advanced edge detection algorithm used for the measurements in this paper varies the edge grey value locally to correct for artifacts that influence the measurement. Nevertheless, even for monomaterial objects, this advanced thresholding method is often inadequate to find the correct edge, thus introducing an uncertainty on the measurement result. As proposed by Kiekens et al. [9] an additional calibration step can be used to correct locally for this wrong edge determination, resulting in an edge correction step. Notice, however, that such edge correction step is not required for all distances.

A distinction has to be made between edge independent and edge dependent distances. For edge independent distances, the measurement result is independent of the chosen grey value for the segmentation between material and the surrounding air (e.g. Figure 2 left), hence no edge correction is required. For other distances, however, the measurement result is heavily dependent on the edge determination step (e.g. Figure 2 right), hence a potential edge offset needs to be corrected.

![Figure 2: Edge INdependent (left) and edge dependent (right) distances](image)

Analytically, this rescaling and edge correction can be formulated as follows:

$$L = \alpha \cdot L_{CT} + \Delta_{cal,CT}$$

(1)

where $L$ is the measurement result and $\alpha$ is a rescaling factor to correct the size of the voxels applied on $L_{CT}$, the measured length on the uncorrected CT model. The edge correction term is defined as $\Delta_{cal,CT}$. 

184
In Equation (1), the rescale factor $\alpha$ is the ratio of the calibration length, measured by a reference measurement instrument, e.g. a CMM, resulting in $L_{\text{cal,ref}}$ and the measurement of this length by computed tomography $L_{\text{cal,CT}}$, where this distance is measured for the same position of the magnification axis as $L_{\text{CT}}$.

$$L = \frac{L_{\text{cal,ref}}}{L_{\text{cal,CT}}} \cdot L_{\text{CT}} + \Delta_{\text{cal,CT}}$$  \hspace{1cm} (2)

Equation (2) represents the analytical equation for the main calibration steps performed on a CT measurement result. In order to minimize errors on the rescale factor, it is recommended to rescale on a distance that is as long as possible.

**PART 1: Theoretical evaluation of CT uncertainty**

This section presents a theoretical evaluation of the uncertainty of dimensional measurements with computed tomography, based on a GUM-based analytical uncertainty equation and on the measurement procedure described above, and more specifically, by Kiekens et al. [9].

In general, the combined standard uncertainty $u_L$ on the measured length $L$ can be calculated according to the ISO-GUM [8], based on Equation (3), where the input quantities $x_i$ with associated uncertainty $u_{x_i}$ are supposed to be uncorrelated:

$$u_L^2 = \sum_{i=1}^{N} \left( \frac{\partial L}{\partial x_i} \right)^2 \cdot u_{x_i}^2$$  \hspace{1cm} (3)

Applying Equation (3) on Equation (2) yields:

$$u_L^2 = \left( \frac{L_{\text{CT}}}{L_{\text{cal,CT}}} \right)^2 \cdot u_{L_{\text{cal,ref}}}^2 + \left( \frac{L_{\text{cal,ref}} \cdot L_{\text{CT}}}{L_{\text{cal,CT}}} \right)^2 \cdot u_{L_{\text{cal,CT}}}^2 + \left( \frac{L_{\text{cal,ref}}}{L_{\text{cal,CT}}} \right)^2 \cdot u_{\Delta_{\text{cal,CT}}}^2 + (1)^2 \cdot u_{\Delta_{\text{cal,CT}}}^2$$  \hspace{1cm} (4)

Basically, the uncertainty on a CT measurement consists of four main uncertainty contributors, the uncertainty on the reference measurement of the calibration length ($u_{L_{\text{cal,ref}}}$), the uncertainty on this length in the CT measurement ($u_{L_{\text{cal,CT}}}$), the uncertainty on the CT measurement of the workpiece ($u_{L_{\text{CT}}}$) and the uncertainty on the edge offset term ($u_{\Delta_{\text{cal,CT}}}$). Each term consists of several subterms.

Dependent on the characteristics of the measured distance, on the calibration method, and on the measurement procedure, some of the (sub)terms of Equation (4) will be very small or zero, as will be exemplified in Part 2.

**PART 2: Identification of the uncertainty contributors**

This section elaborates on the four uncertainty contributors of Equation (4) and illustrates some of these with numerical values based on experimental measurements with various test objects.

1st term: Uncertainty on calibration length $L_{\text{cal,ref}}$ as measured by calibration device ($u_{L_{\text{cal,ref}}}$)

The uncertainty on the reference measurement $u_{L_{\text{cal,ref}}}$ is an uncertainty on a measurement with a conventional (reference) measurement instrument, such as a CMM or an Abbe comparator. It can therefore be determined based on the available standards for the reference measurement instrument used. Accurate manufacturing and reference measurement of the calibration object can allow reducing this term to very small values. Equation 4 allows appreciating that this is especially true when the calibration length is relatively large in comparison with the actual measurand.

2nd term: Uncertainty on calibration length $L_{\text{cal,CT}}$ as measured by CT device ($u_{L_{\text{cal,CT}}}$)
The uncertainty on the CT voxel size calibration length $L_{\text{cal,CT}}$ – which should be an edge independent distance – consists of both a random error and a systematic error component (see Equation 5). The former covers unpredictable variations in repeated observations of the measurand. If the latter is caused by an effect that is well understood and its size is significant relative to the required accuracy of the measurement, a correction $\Delta_{\text{cal,CT}}$ can be applied to $L_{\text{cal,CT}}$ to compensate for such systematic error, rather than accounting for such systematic error by adding an uncertainty component $u_{L_{\text{cal,CT}},\text{systematic}}$.

$$u_{L_{\text{cal,CT}},\text{random}}^2 = u_{L_{\text{cal,CT}},\text{random}}^2 + u_{L_{\text{cal,CT}},\text{systematic}}^2$$  

(5)

**Standard uncertainty due to random error $u_{L_{\text{cal,CT}},\text{random}}^2$**

The random error component $u_{L_{\text{cal,CT}},\text{random}}$ represents the repeatability of the relevant part of the CT measurement process for edge independent distances. When the X-ray CT data of the calibration length and of the measurement object originates from one single data capitation, hence also one single reconstruction step, the uncertainty on $L_{\text{cal,CT}}$ originates solely from the repeatability of the data analysis step. This is the case when the calibration object and the measurement object can be positioned together on the rotary table of the CT machine, or when an edge independent distance of the workpiece itself can be used as calibration distance. After reconstructing the voxel model, the measurement of the calibration length in the analysis software (here VG Studio Max 2.1) is repeated $n$ times.

Based on these $n$ independent measurements of the calibration length on the reconstructed CT model, the experimental variance of the mean of this distance can be calculated as the variance of those $n$ measurements $q_{\text{VG}}$:

$$s^2(q_{\text{VG}}) = \frac{1}{n} \cdot \frac{1}{n-1} \cdot \sum_{VG_{\text{cal}}}^n (q_{\text{VG}} - \overline{q_{\text{VG}}})^2$$  

(6)

where $\overline{q_{\text{VG}}}$ is the arithmetic mean or average of the $n$ independent observations $q_{\text{VG}}$ in the software VG on one single reconstructed voxel model. The experimental variance of the mean $s^2(q_{\text{VG}})$ can be used as a measure of the uncertainty on the calibration length in the CT measurement due to random error, when only one data captation is needed for both the measured and the reference length. This uncertainty is

$$u_{L_{\text{cal,random}}}(q_{\text{VG}}) = s(q_{\text{VG}})$$  

(7)

In contrast of doing a complete measurement, including data captation, performing different measurements in the software VG on one reconstruction only takes some minutes, and so is feasible to do for every measurement. However, this uncertainty also can be estimated based on previous measurements (according to the ISO-GUM [8] type B evaluation of standard uncertainty). When the standard deviation $s(q_{\text{VG}})$ is calculated based on a series of calibration measurements, this value can be used as an estimate of the standard uncertainty.

In case the calibration object and the measurement object cannot be positioned together on the rotary table of the CT device, two consecutive data captations with respective reconstruction steps are...
required. This implies that additional uncertainties are introduced, due to the repeatability of the data capitation step (CT device) and of the reconstruction algorithms: the voxel size of the CT model for the calibration object might not completely equal the voxel size of the CT model for the measurement object.

![Diagram](data-captation-reconstruction-data-analysis)

**Figure 4:** repeatability of the complete measurement process

The best estimate of the variance of the arithmetic mean $\bar{q}$ of $p$ different data captations (with the same measurement settings) is the experimental variance of the mean given by

$$s^2(q_{\text{meas}}) = \frac{1}{p} \cdot \frac{1}{p-1} \cdot \sum_{i=1}^{p} (q_{\text{meas}} - q_{\text{meas}})^2$$

where $q_{\text{meas}}$ represents the different measurements (including data captation). The associated standard uncertainty $u$ is the experimental standard deviation of the mean $s(q_{\text{meas}})$, here referred to as $u_{\text{cal, random}}(q_{\text{meas}})$, representing the uncertainty due to the random error on the calibration length in the CT measurement where different measurements are needed.

$$u_{\text{cal, random}}(q_{\text{meas}}) = s(q_{\text{meas}})$$

In practice, one will not perform multiple measurements to estimate the standard uncertainty due to different data captations for each and every measurement. We propose to base the uncertainty on an a priori estimation of $s(q_{\text{meas}})$ instead of on $s(q_{\text{meas}})$ (i.e. use a so-called type B evaluation, according to the ISO-GUM [8]). In situations where a type A evaluation (evaluation based on statistical methods) is based on a comparatively small number of statistically independent observations, a type B evaluation of the standard uncertainty can be as reliable as a type A evaluation.

The standard uncertainty on the calibration length is simply this standard deviation, calculated based on a series of calibration measurements $u_{\text{cal, random}}(q_{\text{meas}}) = s(q_{\text{meas}})$.

The uncertainty on the calibration length due to random errors always is smaller than 1 µm.

**Standard uncertainty due to systematic error $u^2_{\text{cal, CT, systematic}}$**

The second term in Equation (5) covers all uncompensated systematic errors, i.e. all uncompensated repeatable errors that cause the voxel size to be non-constant throughout the voxel model.

For example, a misalignment between the detector and the rotation axis in a CT machine can lead to voxel sizes that are dependent on the position of the measurand. Figure 5 illustrates this effect on the measurement of a test object that was designed with three rows of 18 steel spheres with a diameter of 4 mm ± 1 µm. The different spheres were in contact, allowing to measure an edge independent distance, i.e. the distance between the centre points of two adjacent spheres. Plastic is used to hold the spheres, without introducing additional artefacts, except for the first and the last sphere, which were excluded from the analysis. Figure 5 shows that the distance between the sphere center points are 7 µm
larger at the top than at the bottom, for each of the three columns. This is significantly more than the variations due to random errors, discussed above.

Figure 5: Illustration of systematic error on calibration length – influence of position (top-bottom)

Due to the reproducibility of the effect, it is clear that this concerns a systematic error. Nevertheless, straightforward compensation is difficult since the magnitude of the observed deviations depends also on the position along the magnification axis. Therefore, it is advisable to perform a machine calibration step at different magnifications on beforehand, in order to estimate the related $u_{L_{\text{cal,CT,systematic}}}$.

However, without knowing the exact relationship, often one can estimate the bounds (upper and lower limits $a_-$ and $a_+$) of the interval including all possible measurements in a particular case. If there is no specific knowledge about the possible values within the interval, one can assume that it is equally probable for a measured value to lie anywhere within it (a uniform or rectangular distribution of possible values). The expected value $x_i$ is the midpoint of this interval and a first estimate of the associated standard uncertainty is (according to the ISO-GUM 4.3.7): $u(x_i) = \sqrt{\frac{a^2}{3}}$. When a component of uncertainty determined in this manner contributes significantly to the uncertainty of a measurement result, it is prudent to obtain additional data for its further evaluation.

Another example of a systematic error can be a temperature change, causing volumetric expansion of the measurement object, hence a measurement error. However, if the coefficient of thermal expansion and the temperature difference are known, compensation is possible. Moreover, if the calibration object and measurement object are made from the same material, this uncertainty contributor automatically reduces to zero.

It is important to notice that the effect obtained from a Type B evaluation (the systematic error), is included as an independent component of the total uncertainty to prevent “double-counting” of uncertainty components. The portion of the uncertainty that contributes to the observed variability (random error) is already included in the component of the uncertainty obtained from the statistical analysis of the observations (ISO-GUM - 4.3.10).

3rd term: Uncertainty on the CT measurement of the workpiece $u_{L_{CT}}$

Basically, the uncertainty on the CT measurement of the workpiece consists of the same subterms as the CT measurement of the calibration object:

$$u^2_{L_{CT}} = u^2_{L_{CT,random}} + u^2_{L_{CT,systematic}} \quad (10)$$

Nevertheless, this term should be included separately, since the quantitative values can be very different. For example, if the measured object contains a distance between two flat surfaces, the repeatability of this distance can be an order of magnitude bigger than the repeatability of measuring
well conditioned sphere centre distances, traditionally used for voxel size calibration. Other type of
distances (cylinders, planes, ...) also can introduce systematic measurements, included in $u_{L,\text{CT,syst}}$.

4th term: Uncertainty on the edge offset term $u_{\Delta\text{cal,CT}}$

For edge dependent distances an additional term should be included, representing the uncertainty on
the edge offset term. Even after an edge detection, based on an advanced algorithm, numerous
parameters are influencing the uncertainty on this edge. The position of the correct edge is dependent
on a number of parameters, which are not included in the available software algorithms.

Different systematic measurements indicate that some of the parameters having an influence here are
surrounding material, feature size, orientation, filtration, beam hardening correction and settings.
Since some of these influencing factor are clearly correlated, a summation of the different uncertainties
caused by those different influences would overestimate the total uncertainty. But, despite of this
correlation, it is useful to give some numerical examples of the influence of some of these terms,
indicating the importance of this uncertainty contributor.

Dewulf et. al. [10] investigated the influence of surrounding material on the measurement of a steel
cylinder. A systematic error arises due to the surrounding material around one part of the object. To
illustrate this influence on the measurement uncertainty, an accurate aluminium cylinder with a
nominal diameter of 8 mm ± 1 µm was surrounded partly by a thick aluminium hollow cylinder ($d_{\text{out}} =
50 \text{ mm}; d_{\text{in}} = 40 \text{ mm}$).

This big aluminium cylinder introduces beam hardening problems, since the beam has other properties
entering the inner cylinder at the top compared to the bottom, hindering a correct edge detection. The
measurement was done without any hardware or software filter. On Figure 6, representing the diameter
of the inner cylinder over different slices, an obvious change in diameter can be observed were the
inner cylinder leaves the outer one, introducing an uncertainty on the edge offset term.

![Figure 6: Accurate aluminium pin surrounded by a thick aluminium ring, without filter, BH 1](image)

**Conclusion**

This paper presents a method to calculate the measurement uncertainty of dimensional measurements
with computed tomography, based on the analytical equation that describes the calibration steps
needed to calibrate a CT measurement (voxel size calibration and edge correction).

Four main uncertainty contributors are identified. Besides the uncertainty on the reference
measurement and the CT measurement of the calibration length, the uncertainty on the CT
measurement of the workpiece and, for edge dependent distances, the uncertainty on the edge
correction term contribute to the total uncertainty. Each of those terms can be subdivided into different
subterms. Based on systematic measurements, some of them have already been identified, and have
been included into the analytical equation.

Some (sub)terms need to be included for every CT measurement, but depending on the measured
distance and the measurement procedure used, some of the subterms cancel out in some situations.
A series of systematic measurements was performed to quantify some of the subterms, allowing to
indicate the importance of these terms. In this paper, an illustration was given for the systematic errors
on edge independent distances as well as for the uncertainty on the edge offset term for edge dependent distances.

At this moment, more systematic measurements are ongoing, based on different calibration objects with the objective to identify and quantify more of the different subterms, for edge independent as well as for edge dependent distances.

References

[1] Wenig, P., Kasperl, S., Examination of the measurement uncertainty on dimensional measurements by X-ray computed tomography, ECNDT, We.3.3.1, 2006.


