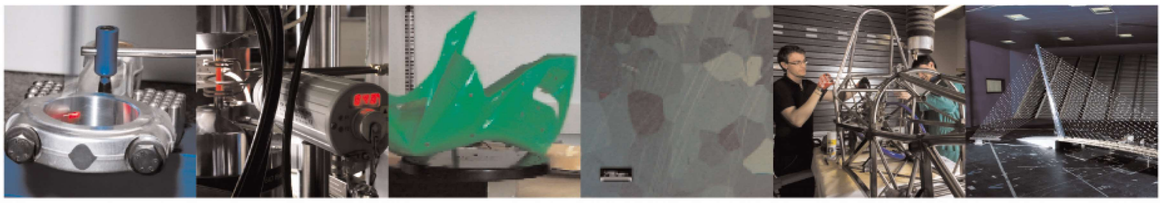
POLITECNICO
MILANO 1863

DIPARTIMENTO DI MECCANICA



A statistical point of view on the ISO 10360 series of standards for coordinate measuring systems verification

Stefano Petrò, Giovanni Moroni

This is a post-peer-review, pre-copyedit version of an article published in Measurement. The final authenticated version is available online at:
http://dx.doi.org/10.1016/j.measurement.2020.108937

DIPARTIMENTO DI MECCANICA ■ POLITECNICO DI MILANO
via G. La Masa, 1 ■ 20156 Milano ■ EMAIL (PEC): pecmecc@cert.polimi.it
http://www.mecc.polimi.it
Rev. 0

A Statistical Point of View on the ISO 10360 Serief of Standards for Coordinate Measuring Systems Verification

Stefano Petrò[a,*], Giovanni Moroni[a]

[a]*Department of Mechanical Engineering, Politecnico di Milano, Via La Masa 1, 20156, Milano, Italy*

Abstract

Coordinate measuring systems acceptance and reverification is based on standardized tests. This means that, although these tests yield a lot of useful information to diagnose any machine malfunction, the single test is just a go/no-go gauge. After conducting the test, the system may or may not conform performance statement. Test design aims for the maximum simplicity, involving repeated measurements of different measurement standards in specified conditions, and comparing the results with specification limits (maximum permissible errors). This makes the nature of these tests statistical and the test success probability depends on the actual behavior of the coordinate measuring system under test. The aim of this work is to analyze the statistical properties of tests for the performance verification of coordinate measuring systems. In particular, as they are the most widespread in industry, we will analyze the ISO 10360 series of standards tests in-depth, proposing operating characteristic curves for them.

Keywords:
Coordinate Measuring Systems, Performance Verification, Statistical Test, Operating Characteristic Curve

---

*Correspondig Author. Tel +39 02 23998530.
Email address: `stefano.petro@polimi.it` (Stefano Petrò)

1. Introduction

Coordinate Measuring Systems (CMS) are complex. Differing from conventional measuring systems, which in most cases can perform only a single kind of measurement (e.g. a ruler or caliper measures only size features, a goniometer measures only angles, etc.), CMSs are usually flexible. They perform a vast range of different measurement tasks, ranging from the simple thickness or diameter measurement to the geometric tolerance verification and even the free-form surface scan. Therefore, the uncertainty for CMS measurements is "task-specific" [1, 2], i.e. the uncertainty a CMS can yield will depend strongly on the measurement task the CMS itself is performing.

However, if the uncertainty changes as the measurement task changes, the uncertainty is not adequate to state the overall performance of the CMS. The uncertainty is no longer a parameter based on which a ranking of CMS can be proposed, nor it is possible to evaluate it once and for all through a calibration procedure. So it is often hard for the CMS buyer to identify whether the CMS meets his requirements. The availability of uncertainty only after the measurement has been performed is a major industrial issue. Furthermore, it is difficult to check, through an uncertainty evaluation only, whether the CMS is performing as expected and whether there is complete traceability to the meter.

Therefore, tests have been introduced for the "acceptance and reverification" of CMS, completely defined in national and international standards. Probably the most widespread standard of this kind is the ISO 10360 series, which at present consists of ten published parts and a few more presently being drafted. In this series, several tests have been developed which aim at verifying whether a CMS performs as stated by either the manufacturer (acceptance test) or the user (reverification test). These procedures act as "go/no-go" gauges and although much information is provided by the measurement operation involved in performing the test (and this is useful for the machine diagnostics), ultimately the test is passed or not passed, no intermediate possibility is considered. This limits the usefulness of the test in understanding what is going on with the machine, and where the machine itself can be improved.

In this paper, a statistical point of view on the tests proposed by the ISO 10360 series of standards is described. The aim is to evaluate the probability the tests state the CMS is behaving according to the stated performance, which is usually defined as the "operating characteristic" (OC) curve of the test. This helps us to understand what the test error probability is, i.e. the probability of either stating the CMS is misbehaving when it is behaving correctly or vice versa. The OC curve can be calculated also if the operator alters the test conditions by changing the number of sampling points or length measurements. Also the test uncertainty (related to the measuring equipment) influences the test result, and can be considered in the OC curve. The proposed results are useful for both the CMS manufacturer, helping him to correctly evaluate the performance of his machine (e.g. when declaring it in the system brochure) and the user when planning the machine test.

2. State of the art

Although performance verification of CMS is a widely discussed issue in literature, finding works that analyze in depth what conducting a performance test means is hard. Most of the proposed works, like those proposed by Hope and Blackshaw [3], Neuschaefer-Rube *et al.* [4], Acko *et al.* [5], Moroni *et al.* [6], and El Asmai *et al.* [7] deal with the definition of the artifacts to use for performance verification. Accurate artifacts are fundamental for effective performance verification. However, these papers neglect the impact that the specific procedure for the application of these artifacts can have on the performance test result. They just try to prove that they are "compatible" with the "standard test" results. In addition some authors, such as Curiel-Razo *et al.* [8], extended the application of the ISO 10360 series methodology to other systems which are not exactly CMSs (leap motion controller).

Swornowski [9] discusses the limitations of the ISO 10360-2:2009 approach to performance verification of coordinate measuring machines (CMM). In particular, he criticizes the fact that the standard describes the test as if it is conducted in ideal conditions, while in practice alignment and deflection of length measurement standards should be taken into account. Furthermore, the test is supposed to be conducted under a no-load condition, which is not coherent with the usual CMS working condition. But again, Swornowski does not analyze the test and its statistical aspects. Furthermore, some of the objections he raises are addressed in the ISO/TS 23165:2006 [10] standard.

Piratelli-Filho and Di Giacomo [11] apply factorial designs [12] to analyze the output of the CMS performance test described in the ASME B89.4.1 standard (now known as ASME B89.4.10360.2:2008 [13]). This is interesting, as the authors consider the results of the measurements involved in the test as random variables, and analyze them accordingly. Their aim is to evaluate the "uncertainty" of the CMS, considering a series of influencing factors. This is not completely correct, because the uncertainty in coordinate measurement is always task-specific [1]. What they evaluate is the influence on the measurement result of the orientation and length of the length measurement standard adopted in the test. Although this is surely relevant, it is interesting anyway as it describes a statistics-based method for the (on-line) performance verification of CMMs.

Franceschini *et al.* [14, 15, 16] discuss the possibility of setting up an on-line control for CMMs. In particular, the possibility of applying control charts was studied initially [14], and then improved through the definition of a standard "witness-part", i.e. a part to be periodically measured to guarantee the CMM is still behaving correctly [15, 16]. Although this approach is completely different from the one proposed in the verification tests described in the international standards, which is off-line and periodical, it is anyway interesting as it describes a statistics-based method for the (on-line) performance verification of CMMs.

Finally, part 5 of the ISO 10360 series of standards has been revised recently [17]. A few papers [18, 19] have been published in recent years discussing the novelties of the revision. However, this does not include any discussion on the statistical aspects of the procedure.

3. Tests in the ISO 10360 series of standards

The ISO 10360 series of standards consists of ten published parts, plus a few more parts still being drafted. Part 1 [20] is devoted to defining the various terms common to the other parts of the standard, including those related to the CMS as a device (probing system, rotary table, stylus system,. . .), the operation of the CMS, the artifacts to be used in the tests, and the performance of the CMS (although many more are defined within the other parts of the standard). It proposes a classification of CMSs based on their kinematic structure as well. The remaining parts describe the procedures for testing CMS. Each part describes the test(s) proposed for different kinds of CMSs, including CMSs equipped with contact (point-to-point and scanning) probes, optical distance sensors, imaging probing systems, multiple probes, articulated arm CMMs, and laser trackers. A few kinds of CMSs are not covered yet, including computed tomography scanners.

Every proposed test is related to a "performance indicator", which in most cases is a "maximum permissible error" (MPE) for the CMS operating in a specific way (e.g. scanning, probing point-to-point, etc.). The proposed performance indicators are listed in Tab. 1-3. It is worth noting that the symbols adopted in the various standard are not completely coherent. This is due to the different years in which the standards were published. Symbols before 2010 differed from those adopted later. The International Organization for Standardization is currently working on a homogenization of the symbols, so for instance ISO 10360-5:2010 was made consistent in 2020.

As the reader may note, there are a total of 50 performance indicators defined in the current edition of the ISO 10360 series of standards. Although this may seem a very large number, it is worth noting that most of them can be grouped two of categories:

- Probing error indicators, in which the performance index is the range of a series of radii measured on a reference sphere ($P_{\text{Form.Sph.5}\times 25:j:\text{Tact,MPE}}$, $P_{\text{Form.Sph.1}\times 25:\text{SS:Tact,MPE}}$,$P_{\text{Form.Sph.Scan:}k\text{:Tact,MPE}}$, $P_{\text{FV2D,MPE}}$, $P_{\text{Form.Sph.1}\times 25:j:\text{ODS,MPE}}$,

Standard	Indicator	Symbol
ISO 10360-2:2009 CMMs used for measuring size [21]	Maximum permissible error of length measurement	$E_{\mathrm{L,MPE}}$
	Maximum permissible limit of the repeatability range	$R_{0,\mathrm{MPL}}$
ISO 10360-3:2000 CMMs with the axis of a rotary table as the fourth axis [22]	Maximum permissible radial four-axis error	MPE_{FR}
	Maximum permissible tangential four-axis error	MPE_{FT}
	Maximum permissible axial four-axis error	MPE_{FA}
ISO 10360-5:2020 CMMs using single and multiple stylus contacting probing systems [17]	Maximum permissible single-stylus form error	$P_{\mathrm{Form.Sph.1\times25:SS:Tact,MPE}}$
	Maximum permissible single-stylus size error	$P_{\mathrm{Size.Sph.1\times25:SS:Tact,MPE}}$
	Maximum permissible multi-stylus form error	$P_{\mathrm{Form.Sph.5\times25:}j\mathrm{:Tact,MPE}}$
	Maximum permissible multi-stylus size error	$P_{\mathrm{Size.Sph.5\times25:}j\mathrm{:Tact,MPE}}$
	Maximum permissible multi-stylus location error	$L_{\mathrm{Dia.5\times25:}j\mathrm{:Tact,MPE}}$
	Maximum permissible scanning mode form error on a sphere	$P_{\mathrm{Form.Sph.Scan:}k\mathrm{:Tact,MPE}}$
	Maximum permissible scanning mode size error on a sphere	$P_{\mathrm{Size.Sph.Scan:}k\mathrm{:Tact,MPE}}$
	Maximum permissible scanning mode form error on a ring gauge	$P_{\mathrm{Form.Cir.Scan:}lo\mathrm{:Tact,MPE}}$
	Maximum permissible scanning mode size error on a ring gauge	$P_{\mathrm{Size.Cir.Scan:}lo\mathrm{:Tact,MPE}}$
	Maximum permissible opposing-styli projected location error on a sphere	$L_{\mathrm{Dia.proj.Sph.2\times25:}j\mathrm{:Tact,MPE}}$
	Maximum permissible opposing-styli projected location error on a ring gauge	$L_{\mathrm{Dia.proj.Cir.Scan:}j\mathrm{:Tact,MPE}}$
ISO 10360-6:2001 Estimation of errors in computing Gaussian associated features [23]	Maximum permissible error	MPE_q

Table 1: Performance indicators in the ISO 10360 series of standards (part 1).

Standard	Indicator	Symbol
ISO 10360-7:2011 CMMs equipped with imaging probing systems [24]	Maximum permissible error of bidirectional length measurement	$E_{\mathrm{B,MPE}}$
	Maximum permissible limit of the bidirectional repeatability range	$R_{\mathrm{B,MPL}}$
	Maximum permissible error of unidirectional length measurement	$E_{\mathrm{U,MPE}}$
	Maximum permissible limit of the unidirectional repeatability range	$R_{\mathrm{U,MPL}}$
	Maximum permissible error of Z bidirectional length measurement	$E_{\mathrm{BZ,MPE}}$
	Maximum permissible error of Z unidirectional length measurement	$E_{\mathrm{UZ,MPE}}$
	Maximum permissible error of XY bidirectional length measurement	$E_{\mathrm{BXY,MPE}}$
	Maximum permissible error of XY unidirectional length measurement	$E_{\mathrm{UXY,MPE}}$
	Maximum permissible squareness error	$E_{\mathrm{SQ,MPE}}$
	Maximum permissible error of imaging probe bidirectional length measurement	$E_{\mathrm{BV,MPE}}$
	Maximum permissible error of imaging probe unidirectional length measurement	$E_{\mathrm{UV,MPL}}$
	Maximum permissible probing error	$P_{\mathrm{F2D,MPE}}$
	Maximum permissible probing error of the imaging probe	$P_{\mathrm{FV2D,MPE}}$
ISO 10360-8:2013 CMMs with optical distance sensors [25]	Maximum permissible probing form error	$P_{\mathrm{Form.Sph.1\times25:}j\mathrm{:ODS,MPE}}$
	Maximum permissible limit of probing dispersion	$P_{\mathrm{Form.Sph.D95\%:}j\mathrm{:ODS,MPE}}$
	Maximum permissible probing size error	$P_{\mathrm{Size.Sph.1\times25:}j\mathrm{:ODS,MPE}}$
	Maximum permissible probing size error All	$P_{\mathrm{Size.Sph.All:}j\mathrm{:ODS,MPE}}$
	Maximum permissible length measurement error	$E_{\mathrm{Bi:}j\mathrm{:ODS,MPE}}$, $E_{\mathrm{Uni:}j\mathrm{:ODS,MPE}}$
	Maximum permissible flat form measurement error	$P_{\mathrm{Form.Pla.D95\%:}j\mathrm{:ODS,MPE}}$
	Maximum permissible limit of the articulated location value	$L_{\mathrm{Dia.5\times25:Art:ODS,MPE}}$
ISO 10360-9:2013 CMMs with multiple probing systems [26]	Maximum permissible multiple probing system form error	$P_{\mathrm{Form.Sph.}n\mathrm{\times25::ODS,MPE}}$
	Maximum permissible multiple probing system size error	$P_{\mathrm{Size.Sph.}n\mathrm{\times25::ODS,MPE}}$
	Maximum permissible multiple probing system location error	$L_{\mathrm{Dia.}n\mathrm{\times25::ODS,MPE}}$

Table 2: Performance indicators in the ISO 10360 series of standards (part 2).

Standard	Indicator	Symbol
ISO 10360-10:2016 Laser trackers for measuring point-to-point distances [27]	Maximum permissible error of length measurement (unidirectional)	$E_{\mathrm{Uni:L:LT,MPE}}$
	Maximum permissible error of length measurement (bidirectional)	$E_{\mathrm{Bi:L:LT,MPE}}$
	Maximum permissible error of probing form	$P_{\mathrm{Form.Sph.1\times25::SMR.LT,MPE}}$
	Maximum permissible error of probing size	$P_{\mathrm{Size.Sph.1\times25::SMR.LT,MPE}}$
	Maximum permissible error of location	$L_{\mathrm{Dia.2\times1:P\&R:LT,MPE}}$
ISO 10360-12:2016 Articulated arm coordinate measurement machines (CMM) [28]	Maximum permissible error of articulated location error, tactile	$L_{\mathrm{Dia.5\times5:Art:Tact.AArm,MPE}}$
	Maximum permissible error of bidirectional length measurement	$E_{\mathrm{Bi:0:Tact.AArm,MPE}}$
	Maximum permissible error of unidirectional length measurement	$E_{\mathrm{Uni:0:Tact.AArm,MPE}}$
	Maximum permissible error of probing form, tactile	$P_{\mathrm{Form.Sph.1\times25::Tact.AArm,MPE}}$
	Maximum permissible error of probing size, tactile	$P_{\mathrm{Size.Sph.1\times25::Tact.AArm,MPE}}$

Table 3: Performance indicators in the ISO 10360 series of standards (part 3).

$P_{\mathrm{Form.Sph.}n\times25\mathrm{::ODS,MPE}}$, $P_{\mathrm{Size.Sph.1\times25::SMR.LT,MPE}}$, $P_{\mathrm{Form.Sph.1\times25::Tact.AArm,MPE}}$)

- Error of length measurement, in which the performance is given by the maximum measurement error yielded by the measurement of a set of calibrated standards of length, like a series of gauge blocks or a laser interferometer ($E_{\mathrm{L,MPE}}$, $E_{\mathrm{B,MPE}}$, $E_{\mathrm{U,MPE}}$, $E_{\mathrm{BZ,MPE}}$, $E_{\mathrm{UZ,MPE}}$, $E_{\mathrm{BXY,MPE}}$, $E_{\mathrm{UXY,MPE}}$, $E_{\mathrm{BV,MPE}}$, $E_{\mathrm{UV,MPE}}$, $E_{\mathrm{Bi:}j\mathrm{:ODS,MPE}}$, $E_{\mathrm{Uni:}j\mathrm{:ODS,MPE}}$, $E_{\mathrm{Uni:L:LT,MPE}}$, $E_{\mathrm{Bi:L:LT,MPE}}$, $E_{\mathrm{Bi:0:Tact.AArm,MPE}}$, $E_{\mathrm{Uni:0:Tact.AArm,MPE}}$)

These two kinds of indicators derive directly from those included in the first editions of the ISO 10360-2:2009 [29, 30]. In particular, $P_{\mathrm{Form.Sph.1\times25:SS:Tact,MPE}}$ and $E_{\mathrm{L,MPE}}$, which are geared toward a CMM equipped with a single-stylus touch-trigger probe, are probably the most widespread and applied, the other indicators being of more recent introduction and applicable in fewer cases. Moreover, from a statistical point of view most of the indicators can be reduced to these two by simply changing some parameters withing the models, e.g. the

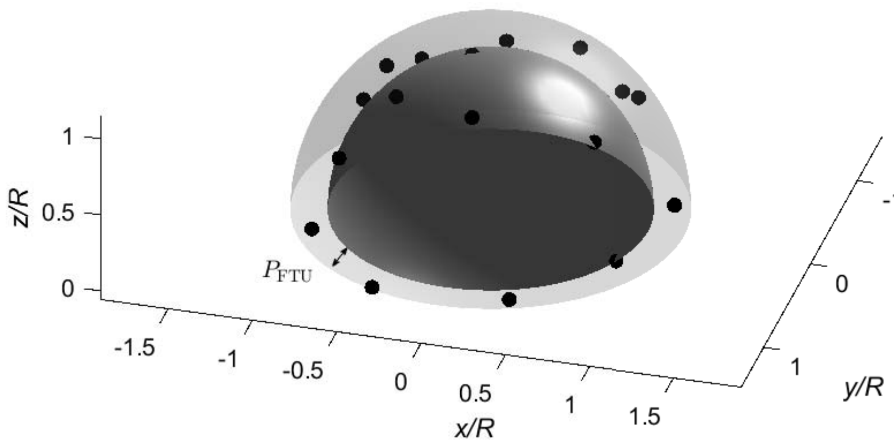


Figure 1: The concept of $P_{\mathrm{Form.Sph.1\times25:SS:Tact}}$. The coordinates are scaled to the sphere radius. Please also refer to Fig. 4 of ISO 10360-5:2020 [17].

sample size or the test uncertainty. Finally, the main CMM manufacturers currently still ignore many indicators. Only recently have some indicators, like $R_{0,\mathrm{MPL}}$, been introduced by a few manufacturers. Therefore, the remainder of this work will focus on $P_{\mathrm{Form.Sph.1\times25:SS:Tact,MPE}}$, $E_{\mathrm{L,MPE}}$, and the related acceptance/reverification tests.

4. Single-stylus form error

The ISO 10360-5:2020 defines [17] the "single-stylus form error" $P_{\mathrm{Form.Sph.1\times25:SS:Tact}}$ (Fig. 1) as

> observed form of a test sphere, the measurements being performed by a CMM with a single-stylus (SS), using the discrete-point probing mode taking 25 points on a single sphere (1 × 25).

This error is strictly related to the repeatability of the single point probed by the CMM and the probing system under test attached to it. The measurement of a test sphere requires only small displacements of the sensor itself (compared to the typical measurement volume of a CMM). It is reasonable that the volumetric error added in this case is then quite small. The error aims at characterizing the potential performance of the CMM when the mechanical structure is perfect.

The related performance index is $P_{\text{Form.Sph.1×25:SS:Tact,MPE}}$, which is defined [17] as

> extreme value of the single-stylus form error, $P_{\text{Form.Sph.1×25:SS:Tact}}$, permitted by specifications.

Please note that the definition of $P_{\text{Form.Sph.1×25:SS:Tact}}$ and $P_{\text{Form.Sph.1×25:SS:Tact,MPE}}$ are extremely generic: the number of sampling points is defined, but neither the distribution of these points on the test sphere nor the fitting principle are defined. Without the definition of these two parameters, $P_{\text{Form.Sph.1×25:SS:Tact,MPE}}$ cannot be adopted to state the performance of a CMM. Therefore, the standard [17] includes a specific test to verify this performance indicator. The test can be summarized as follows.

1. Take a reference sphere. The diameter of the sphere must be at least equal to 10 mm and at most equal to 51 mm. The geometric deviation of the sphere shall be "far lower" (specific limits are suggested in the standard) than the $P_{\text{Form.Sph.1×25:SS:Tact,MPE}}$ of the CMM under test.
2. Sample 25 points on the reference sphere. The pattern of the points can be defined by the operator (provided that the points are approximately evenly distributed over at least a hemisphere of the test sphere); a suggested pattern is given in the standard (Fig. 2).
3. Fit a Gaussian (unconstrained least-squares) sphere to the sampling points. To do this, considering a generic sphere with center $[x_0 \quad y_0 \quad z_0]$ and radius r, the distance of the i^{th} point $[x_i \quad y_i \quad z_i]$ from its surface is:

$$d_i = \sqrt{(x_i - x_0)^2 + (y_i - y_0)^2 + (z_i - z_0)^2} - r \tag{1}$$

   The Gaussian fitting sphere can be calculated by solving

$$\min_{x_0, y_0, z_0, r} \sum_{i=1}^{25} d_i^2 \tag{2}$$

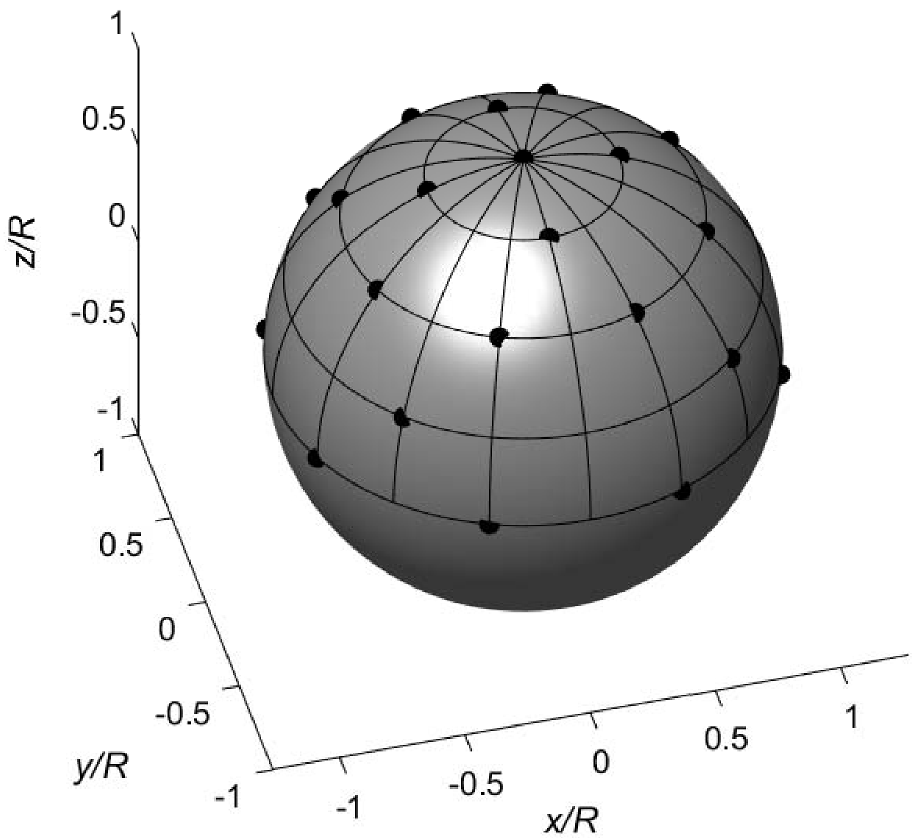


Figure 2: Suggested distribution of the sampling points for the test of the single-stylus form error. The coordinates are scaled to the sphere radius. Please also refer to Fig. 2 of ISO 10360-5:2020 [17].

4. Calculate the single-stylus form error as[1]

$$P_{\mathrm{FTU}} = \max_i d_i - \min_i d_i \tag{3}$$

[1]For the sake of compactness, from now on $P_{\mathrm{Form.Sph.1\times25:SS:Tact}}$ will be referred to as P_{FTU} and $P_{\mathrm{Form.Sph.1\times25:SS:Tact,MPE}}$ as $P_{\mathrm{FTU,MPE}}$, as in the former edition of ISO 10360-5:2010 [31].

5. The single-stylus probing performance is verified if

$$\begin{aligned} P_{\mathrm{FTU}} &\leq P_{\mathrm{FTU,MPE}} - U_{P_{\mathrm{FTU}}} \text{ Acceptance test} \\ P_{\mathrm{FTU}} &\leq P_{\mathrm{FTU,MPE}} + U_{P_{\mathrm{FTU}}} \text{ Reverification test} \end{aligned} \tag{4}$$

where $U_{P_{\mathrm{FTU}}}$ is the uncertainty of the measurement of P_{FTU}.

Please note that $U_{P_{\mathrm{FTU}}}$ is not, in general, influenced by the CMM under test, but only by the equipment (e.g. the reference sphere, gauge block) adopted for the test. The ISO/TS 23165:2006 standard [10] gives details on how to estimate this uncertainty. The presence of the uncertainty in the formula is dictated by the ISO 14253-1:2017 standard [32]. This standard requires that uncertainty always "plays against" who is performing the test. Therefore, when a supplier performs an acceptance test, the uncertainty is subtracted from the limit, so the odds of the test passing are reduced. Instead, if a customer performs a reverification test the uncertainty is added, so the odds of passing the test are increased.

From the definition of the test, it is clear that the P_{FTU} is the difference of the maximum minus the minimum, i.e. the range, of the set of 25 values. If this set of values can be seen as a set of random variables following some specific statistical distribution, then P_{FTU} will follow some (different) statistical distribution. The knowledge of this statistical distribution would allow the calculation of the probability the test is passed as $P_{\mathrm{FTU,MPE}}$ varies, which is the aim of the present work. Unfortunately, the distribution of the range of a finite set of random variables is seldom known. But if the variables are normally and independently distributed, then the distribution of the range is known [33, 34]. Measurement results, here including residuals from a Gaussian sphere, have often proven to be normal. Therefore, in the following it will be assumed the terms d_i are distributed according to a normal distribution and independent, i.e. $d_i \sim NIID\left(\mu, \sigma^2\right)$. The latter hypothesis is never verified. The Gaussian sphere is fitted on the same sampling points for which the distances d_i are calculated. This generates some correlation between the d_i seen as random variables. However, this correlation, as three parameters only are fitted, is usually very small.

Consider a set of n random variables $d_i \sim NIID\left(\mu, \sigma^2\right)$. Define the random variable

$$w = \frac{\max_i d_i - \min_i d_i}{\sigma} \tag{5}$$

Hartley has demonstrated [33] that the cumulative probability of this random variable is

$$\begin{aligned} P_n\left(W\right) = &\left(\int_{-\frac{1}{2}W}^{+\frac{1}{2}W} \phi\left(x\right) dx\right)^n + \\ &+ 2n \int_{\frac{1}{2}W}^{\infty} \phi\left(u\right) \left(\int_{u-W}^{u} \phi\left(x\right) dx\right)^{n-1} du \end{aligned} \tag{6}$$

where $\phi\left(x\right) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}$ is the probability density of the standard normal distribution. When Hartley obtained this result, integrating Eq. (6) was rather difficult, but at present any computer can numerically solve it.

Eq. (6) allows the calculation of the probability the single-stylus probing performance test is passed for given values of σ^2, $P_{\mathrm{FTU,MPE}}$, and $U_{P_{\mathrm{FTU}}}$. This probability can be summarized in a single graph expressing this probability as a function of $\frac{P_{\mathrm{FTU,MPE}}}{\sigma}$ when $U_{P_{\mathrm{FTU}}} = 0$ (Fig. 3), as often considered in common (and incorrect) industrial practice. In the context of statistical inference a curve which defines the probability the null hypothesis of the test is not rejected given some specific condition (usually a specific violation of the null hypothesis itself) is called the "operating characteristic" (OC) curve of the test, and then in the following the curve of the probability the single-stylus probing performance is verified will be referred to as the OC curve of the test. In Fig. 3 two points have been highlighted: the one for which $\frac{P_{\mathrm{FTU,MPE}}}{\sigma} = 4$, and the one for which $\frac{P_{\mathrm{FTU,MPE}}}{\sigma} = 6$. The first one is linked to the classical value of the expansion factor, $K = 2$, for the expanded uncertainty U evaluation. In this case, it is commonly assumed that $P\left(x - U \leq y \leq x + U\right) \cong 0.95$, where x is the reference

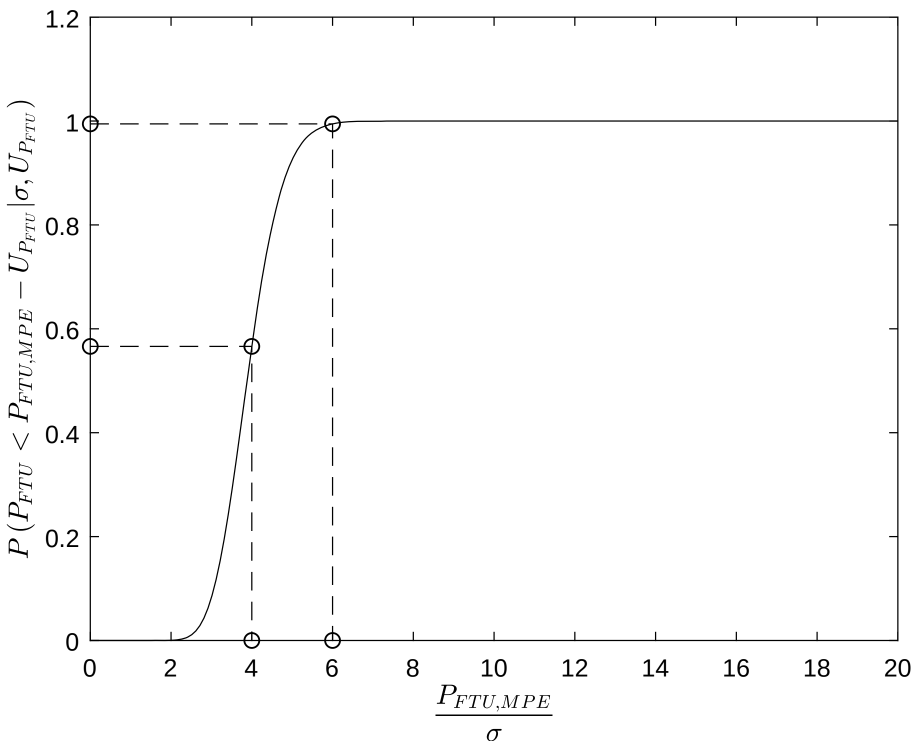


Figure 3: Operating characteristic curve of the single-stylus probing performance test.

value of the measurand and y is the measurement result [35]. It is worth noting that $P\left(P_{\mathrm{FTU}} < P_{\mathrm{FTU,MPE}} | \frac{P_{\mathrm{FTU,MPE}}}{\sigma} = 4\right) = 0.57$: in practice, if $P_{\mathrm{FTU,MPE}} = 4\sigma$, nearly once every two tests the CMM will be declared "not behaving correctly". It is apparent that, supposing σ is known, stating $P_{\mathrm{FTU,MPE}} = 4\sigma$ is not a good choice. A better choice could be stating $P_{\mathrm{FTU,MPE}} = 6\sigma$. In fact, in this case, $P\left(P_{\mathrm{FTU}} < P_{\mathrm{FTU,MPE}} | \frac{P_{\mathrm{FTU,MPE}}}{\sigma} = 6\right) = 0.995$ that means the test is (wrongly) failed only once every 182 tests on average, that is probably acceptable, considering that CMMs are usually verified once a year. Please note this can be considered a "six-sigma" approach.

4.1. Influence of the sample size (number of sampling points)

As stated in §4, the number of sampling points in the single-stylus probing error test is fixed and equal to 25. Therefore, a discussion on the influence

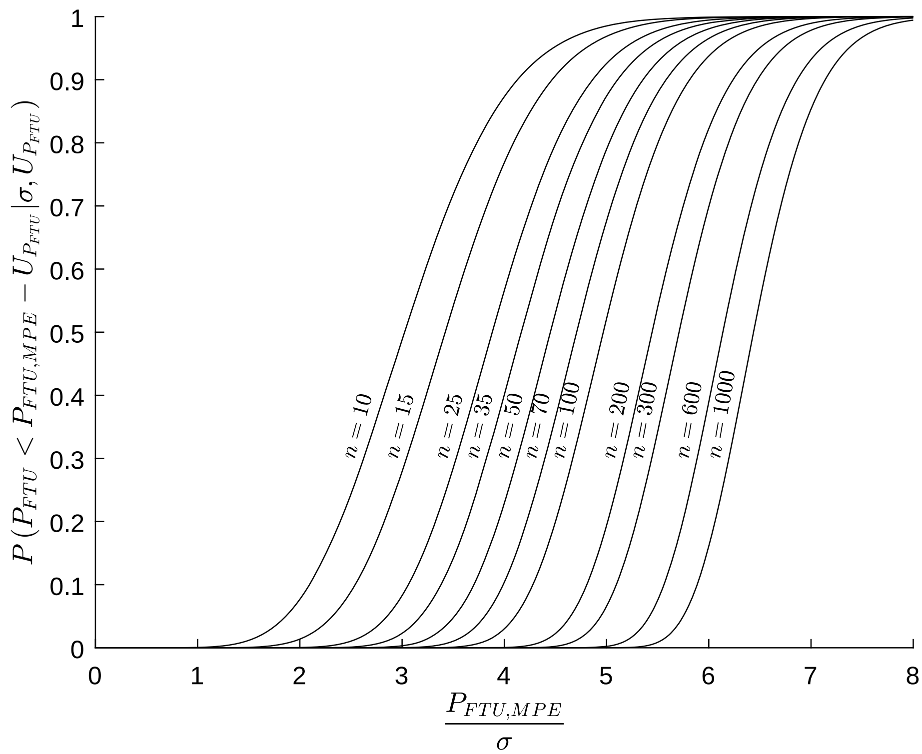


Figure 4: Operating characteristic curve when the sample size changes.

of the number of sampling points could appear superfluous. However, other performance parameters of CMSs, e.g. MPE_{Tij} or $P_{\mathrm{FT}j,\mathrm{MPE}}$, are based on tests requiring more than 25 sampling points. A discussion on them can be useful to understand how the probability of failing the test could influence the statement of the performance index. Fig. 4 shows how the OC curve changes as the sample size changes.

As one might expect, as the sample size increases the probability of passing the test decreases. As the sample size increases the probability of encountering points with an abnormally high distance from the Gaussian fitting sphere increases. Since the range is influenced only by the maximum and minimum distance, the range itself is likely to increase. Even if the increase of the sample size can increase the accuracy in the definition of Gaussian sphere, the same structure of the test is such that this will not alter its result, as it always includes the single point dispersion. In fact, the probing error test is, in the end, based on single probing points (only the maximum and minimum deviating points influence the range). These results suggest that great care should be taken in the evaluation of performance indices when many sampling points are involved in the test. In fact, in most cases, the indices requiring many points are calculated in situations in which the CMM accuracy is limited by its operating mode (e.g. in MPE_{Tij} the accuracy is limited by the dynamic distortion of the machine structure, as the CMM operates in scanning mode), and this calls for lower performance if compared to the single-stylus probing error. The increase in the sample size also calls for a reduction of the performance indexes to allow the CMM to pass the test. This makes hard to understand whether low performance is due to inaccuracy or due to large test sample size.

4.2. Influence of the test uncertainty $U_{P_{FTU}}$

The ISO 14253-1:2017 standard [32] requires the measurement uncertainty to be considered in any statement about the conformance to tolerances based on measurement results. In the field of CMS testing, this means a reduction (acceptance test) or an increase (reverification test) of the limit value for the performance index considered, as highlighted by Eq. (4) for the single-stylus probing error. The discussion proposed so far has completely neglected this issue, concentrating on the pure test. Now, if uncertainty is considered, the situation changes. Fig. 5 plots how the OC curve changes as the test uncertainty changes. The test uncertainty has been expressed as a fraction of the $P_{\text{FTU,MPE}}$, i.e. $U_{P_{\text{FTU}}}/P_{\text{FTU,MPE}}$. An acceptance test has been considered for the verification, i.e. $P_{\text{FTU}} \leq P_{\text{FTU,MPE}} + U_{P_{\text{FTU}}}$, but both positive and negative values have been considered for $U_{P_{\text{FTU}}}/P_{\text{FTU,MPE}}$. Therefore, positive values of $U_{P_{\text{FTU}}}/P_{\text{FTU,MPE}}$ refer to the acceptance test and negative values $U_{P_{\text{FTU}}}/P_{\text{FTU,MPE}}$ to the reverification test. The reverification test does not present any particular issue (the probability of passing the test simply increases) but, in the case of the acceptance test, the situation is more complicated. The

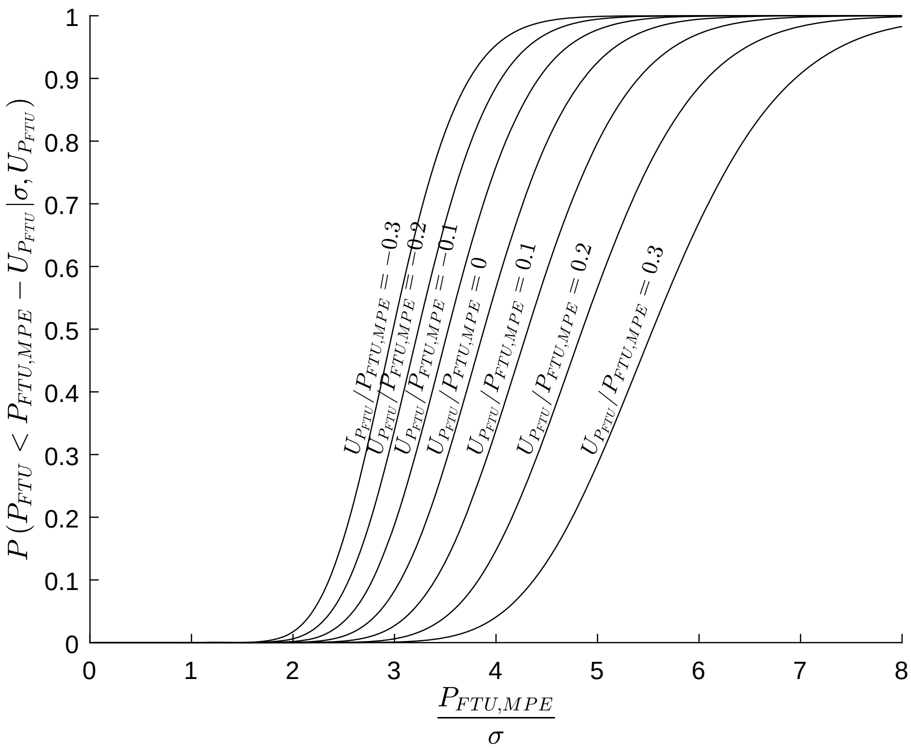


Figure 5: Operating characteristic curve when the test uncertainty is considered.

test uncertainty depends on the form deviation and calibration uncertainty of the test artifact. Even if ultra-precision spheres, with a nanometric form deviation can be manufactured, yielding calibration uncertainties smaller than 0.1 µm is difficult. Therefore, the minimum test ucertainty is of the same order of magnitude. This means that testing CMMs, for which a $P_{\mathrm{FTU,MPE}} \leq 1$ µm is stated, is very difficult because the uncertainty strongly increases the probability of not passing the test. At present the same CMM manufacturer states that probably their highest accuracy machines are characterized by an accuracy that is better than their $P_{\mathrm{FTU,MPE}}$ suggests, but they cannot state this because no reference artifact with an adequately low uncertainty is available

4.3. Method validation

Validating the method requires the hypotheses to be verified. The cumulative probability in (6) is based on the normality and independence of the values in the sample. As stated in §4, values are not independent so testing their independence would be improper.

Normality can be verified instead. The $P_{\mathrm{Form.Sph.1\times25:SS:Tact,MPE}}$ test has been run 100 times on a "Zeiss Prismo 5 VAST HTG" tactile CMM. The value stated by the manufacturer for the $P_{\mathrm{Form.Sph.1\times25:SS:Tact,MPE}}$ of this machine is equal to 2 µm. The data are made available [36]. After fitting the Gaussian sphere and calculating the radius of each probing point, the Anderson-Darling test [37] was applied to each set of radii. The distribution of the resulting p-values is shown in Fig. 6, the minimum p-value being 0.007 (the unique value below 0.1). As there is no statistical evidence of non-normality in any dataset, it is proved that in general the radii probed on the reference sphere are normally distributed.

5. Length measurement error E_{L}

The ISO 10360-2:2009 defines [21] the "length measurement error" E_{L} as

> error of indication when measuring a calibrated test length using a CMM with a ram axis stylus tip offset of L, using a single probing point (or equivalent) at each end of the calibrated test length.

The length measurement test is dual with the single-stylus probing error test. The latter aims at testing the machine within a limited volume, to extract the sensor contribution of the error as far as possible. The first instead measures a length. If the length is adequately large compared to the measuring volume the volumetric error of the machine will be necessarily involved. However, the length measurement error includes *de facto* also the probing error. A single probing point or equivalent for each end of the calibrated test length is taken. Therefore, any purely random and non-volumetric error is retained in its assessment. If the

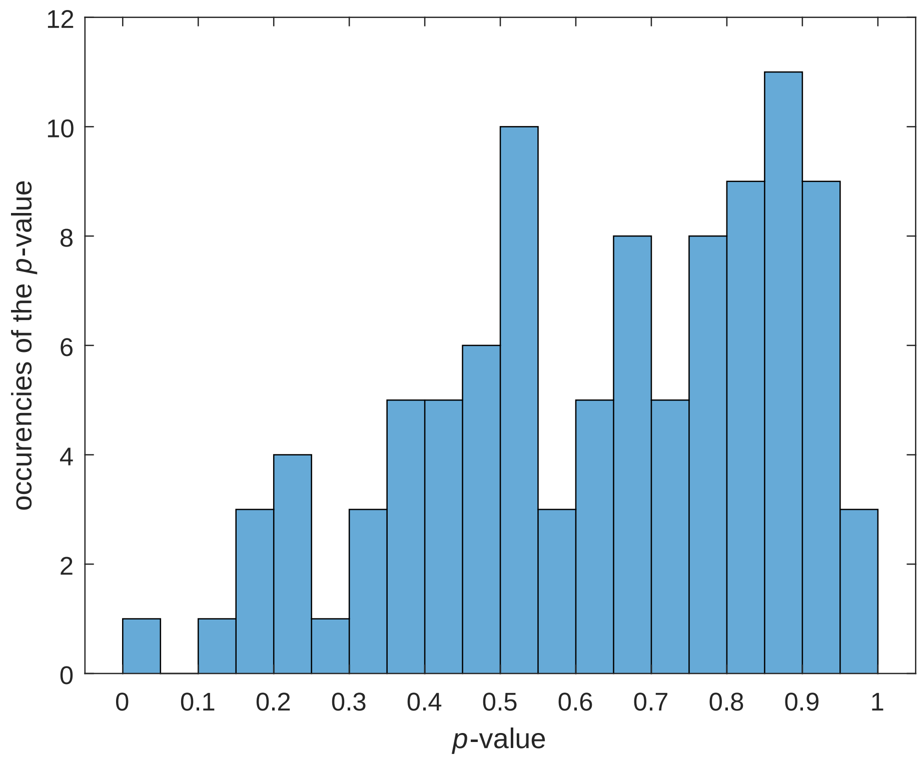


Figure 6: Histogram of the p-values obtained from 100 Anderson-Darling normality test.

length measurement was based for example on the measurement of the center of the spheres of a ball bar through Gaussian fitting of the two spheres of the ball bar itself, the averaging effect of Gaussian fitting would reduce or even eliminate the random error contribution.

The performance index related to the length measurement error is the "maximum permissible error of length measurement" $E_{\mathrm{L,MPE}}$, defined [21] as

> extreme value of the length measurement error, E_{L}, permitted by specifications.

The L symbol in E_{L} stands for the perpendicular distance of the stylus tip to the ram axis of the CMM. Of course, this is meaningful only in case of CMMs, and not for a generic CMSs. A particular case is $L = 0$, i.e. the tip coincides with the ram axis of the CMM. The test in this case is more complete, and is usually considered by CMM manufacturers. Furthermore, tests for optical CMS are usually inspired by this test for E_0. Therefore, in the following only this specific case will be considered.

The procedure for testing E_0 is more complex and time consuming than the one for the single-stylus form error because the whole measuring volume of the CMS must be investigated. The overall procedure can be summarized as follows.

1. Select seven positions (locations and orientations) in the measurement volume along which the test will be conducted[2].
2. Select five length measurement standards (e.g. gauge blocks, step gauge, laser interferometer,...) to be measured (Fig. 7). The length measurement standards can be different for every considered direction. The longest length measurement standard should cover at least $0.66 \cdot l$, where l is the machine travel in the considered position. Each calibrated test length must differ significantly from the others in length. Their lengths must be well distributed over the measurement line.
3. Measure each length measurement standard three times for each position (total 105 length measurements).
4. For each measurement, calculate the length measurement error, E_0, by calculating the difference between the indicated value and the calibrated value of each test length (where the calibrated value is taken as the conventional true value of the length).
5. The length measurement performance is verified if

$$\begin{aligned} &E_0 \leq E_{0,\text{MPE}} - U_{E_0} \text{ Acceptance test} \\ &E_0 \leq E_{0,\text{MPE}} + U_{E_0} \text{ Reverification test} \end{aligned} \tag{7}$$

where U_{E_0} is the uncertainty of the measurement of E_0 (which in ge-

[2] The most recent edition of the ISO 10360-2:2009 standard [21] requires four of these directions to be selected along the volumetric diagonals of the measurement volume, and suggests that the remaining three be parallel to the three Cartesian axes of the machine and passing through the center of the measurement volume.

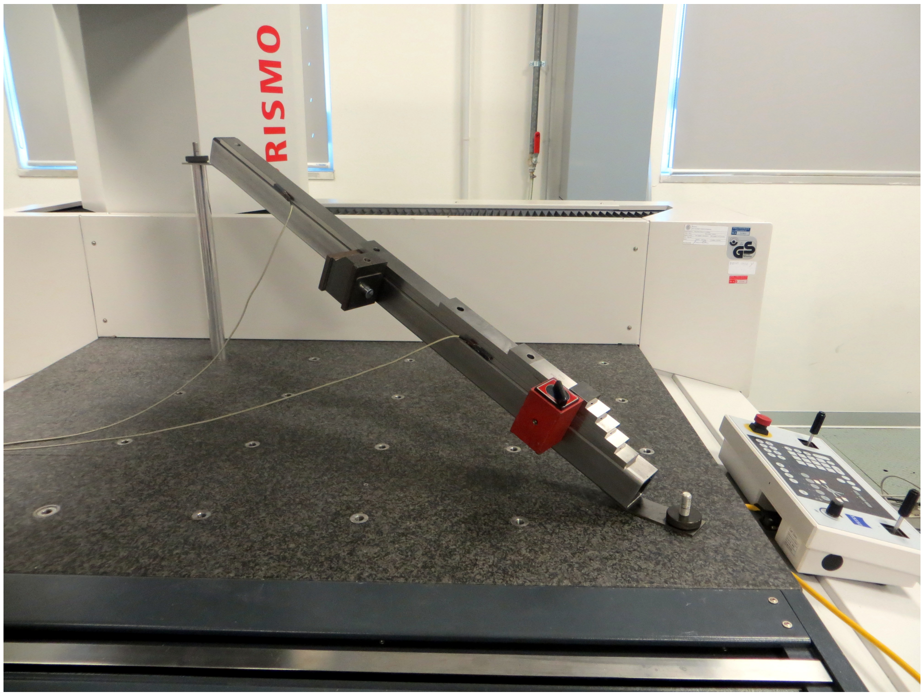


Figure 7: An example of the use of gauge blocks as artifacts for the execution of the length measurement error test (courtesy of Politecnico di Milano).

neral depends only on the equipment adopted for the test). The ISO/TS 23165:2006 standard [10] gives details on how to estimate this uncertainty. Again, the presence of the uncertainty is dictated by the ISO 14253-1:2017 standard.

In most cases $E_{0,\mathrm{MPE}}$ is not expressed as a constant, but as a function of the nominal or calibrated length l of the measured length standard, e.g.

$$E_{0,\mathrm{MPE}} = a + b \cdot l \tag{8}$$

where a and b are constants specific for the considered CMS. This suggests that in general $E_{0,\mathrm{MPE}}$ is proportional to the measured length, i.e. as the measured length increases the absolute measurement accuracy is likely to reduce.

From this description, it is apparent that the length measurement error test is a set of 105 separated tests, which are considered together. Now, consider a single length measurement error evaluation, and the acceptance test (in case of reverification test the solution is similar). Suppose that $E_0 \sim N\left(0, \sigma_j^2\right)$, where $j \in \{1, 2, \ldots, 5\}$ is an index denoting a length standard. In practice, it has been supposed that the length measurement error is distributed according a Gaussian distribution, with a null expected value, and a variance which depends on the considered length standard. It is worth nothing that this normality assumption is probably a severe simplification of the reality. The presence of residual systematic errors, scale errors, and calibration errors of the reference standards can easily make E_0 non-normal and with an expected value differing from zero. However, this simplified model is the only one allowing an easy calculation of the operating characteristic curves. To allow the statistical discussion of the test, it is then accepted. It is easy to prove (see e.g. Montgomery and Runger [38]) that $P\left(|E_0| \le E_{0,\mathrm{MPE}} - U_{E_0}\right)$ can be calculated as

$$\begin{aligned}
&P\left(|E_0| \le E_{0,\mathrm{MPE}} - U_{E_0}\right) = \\
=&P\left(E_0 \le E_{0,\mathrm{MPE}} - U_{E_0}\right) - P\left(E_0 < U_{E_0} - E_{0,\mathrm{MPE}}\right) = \\
=&\Phi\left(\frac{E_{0,\mathrm{MPE}} - U_{E_0}}{\sigma_j}\right) - \Phi\left(\frac{U_{E_0} - E_{0,\mathrm{MPE}}}{\sigma_j}\right) = \\
=&2\Phi\left(\frac{E_{0,\mathrm{MPE}} - U_{E_0}}{\sigma_j}\right) - 1
\end{aligned} \tag{9}$$

where $\Phi(x)$ denotes the cumulative distribution function of the standard normal distribution calculated at x. This is the probability that a single measurement conforms to the stated $E_{0,\mathrm{MPE}}$. To consider all the measurement results together, independence of each measurement result from any other must be supposed. The probability that two independent events happen together is the product of the probability of the two events [38]. Supposing that $\frac{E_{0,\mathrm{MPE}} - U_{E_0}}{\sigma_j} = \frac{E_{0,\mathrm{MPE}} - U_{E_0}}{\sigma} \quad \forall j$ (i.e. it does not depend on j or the size of the length standard), the probability that the test is passed is equal to[3]

$$P\left(\text{“test passed”}\right) = \left[2\Phi\left(\frac{E_{0,\mathrm{MPE}} - U_{E_0}}{\sigma}\right) - 1\right]^{105} \tag{10}$$

Please note that supposing that $\frac{E_{0,\mathrm{MPE}} - U_{E_0}}{\sigma} = k$, where k is a constant, if $E_{0,\mathrm{MPE}}$ is expressed as in Eq. (8), means to state that $\sigma = \frac{1}{k}\left(a + bl_j\right)$, that is, the dispersion of E_0 is proportional to the size of the considered length measurement standard, which is reasonable.

Similarly to what has already been done for the single-stylus form error test in §4, operating characteristic curves based on Eq. (10) can be proposed, plotting the probability the test is passed as a function of $\frac{E_{0,\mathrm{MPE}}}{\sigma}$ when $U_{E_0} = 0$ (Fig. 8). A few points have been highlighted. The point for which for which $\frac{E_{0,\mathrm{MPE}}}{\sigma} = 2$ is linked to the classical value of the expansion factor, $K = 2$, for the expanded uncertainty U evaluation. It is worth noting that $P\left(E_0 < E_{0,\mathrm{MPE}} \middle| \frac{P_{\mathrm{FTU,MPE}}}{\sigma} = 4\right) = 0.0075$: in practice, if $E_{0,\mathrm{MPE}} = 2\sigma$, the test is seldom passed. It is apparent that, supposing σ is known, stating $E_{0,\mathrm{MPE}} = 2\sigma$ is not a good choice. Even the point for which $E_{0,\mathrm{MPE}} = 3\sigma$ is not good. In fact, in this case $P\left(E_0 < E_{0,\mathrm{MPE}} \middle| \frac{E_{0,\mathrm{MPE}}}{\sigma} = 3\right) = 0.75$ that means the test is (incorrectly) failed only once every 4 tests on average, which is still not acceptable. Although this is a sort of “six-sigma” approach, the CMS manufacturers should consider larger values when stating the performance indices of their systems. For example, choosing 3.5σ leads to a probability that the test is passed approximately equal to 0.95, which still signifies a fail once every 20 tests on average.

5.1. Influence of the sample size (number of sampling points)

As stated in §5, the number of sampling points in the length measurement error test is fixed and equal to 105. Therefore, a discussion on the influence

[3]The ISO 10360-2:2009 standard allows ten additional measurements to be performed in case a single measurement fails the test: if all these ten measurements pass the test, then the test is passed. Here, for sake of simplicity, this possibility is neglected.

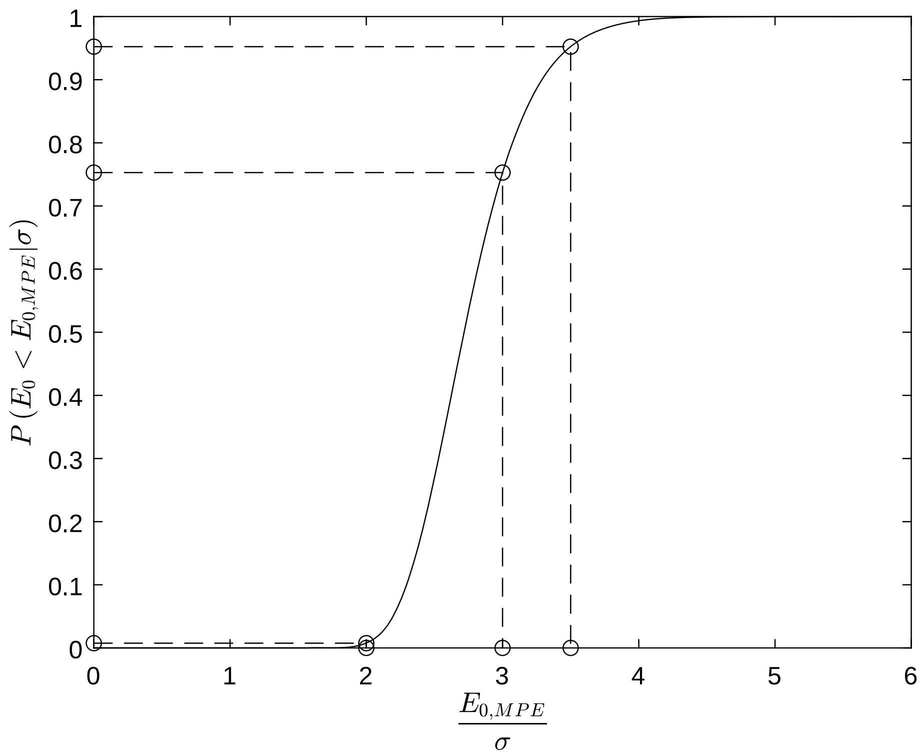


Figure 8: Operating characteristic curve of the length measurement error test ($n = 105$).

of the number of sampling points could appear superfluous. However, this can be deemed useful in the case a company decides to increase the number of replicates, position or length standards to increase their knowledge of machine behavior, or vice versa to reduce it to shorten the time required by the test. Fig. 9 shows how the OC curve changes as the sample size changes (here n denotes the overall number of measurements taken).

As expected, as the sample size increases the probability of passing the test decreases. This is obvious, considering the formula generating the OC curve is

$$P\left(\text{“test passed”}\right) = \left[2\Phi\left(\frac{E_{0,\mathrm{MPE}} - U_{E_0}}{\sigma}\right) - 1\right]^n \tag{11}$$

These results suggests that manufacturers should not vary the number of measurements when stating $E_{0,\mathrm{MPE}}$. Increasing it could lead to an overestimation of

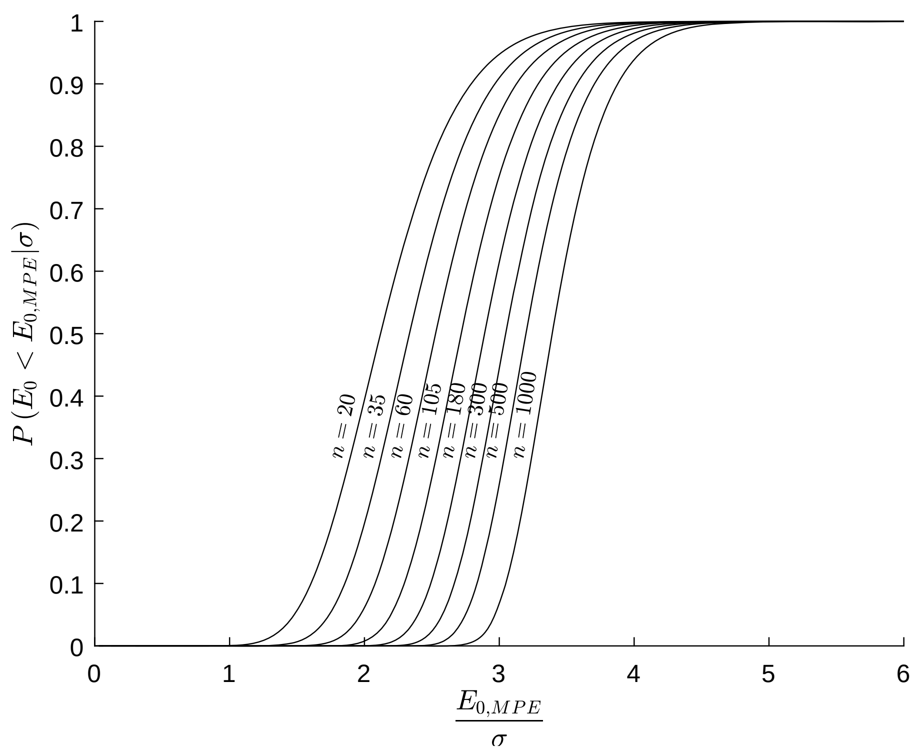


Figure 9: Operating characteristic curve when the sample size changes.

the performance parameter, and conversely reducing it to an underestimation.

5.2. Influence of the test uncertainty U_{E_0}

The ISO 14253-1:2017 standard [32] requires the measurement uncertainty to be considered in any statement about conformance to tolerances based on measurement results. In the field of CMSs testing, this means a reduction (acceptance test) or an increase (reverification test) of the limit value for the performance index considered, as highlighted by Eq. (7) for the length measurement error test. The discussion proposed so far has completely neglected this issue, concentrating on the pure test. Now, if uncertainty is considered, the situation changes. Fig. 10 plots how the OC curve changes as the test uncertainty changes. The test uncertainty has been expressed as a fraction of the $E_{0,\mathrm{MPE}}$, i.e. $U_{E_0}/E_{0,\mathrm{MPE}}$. An acceptance test has been considered, i.e. the

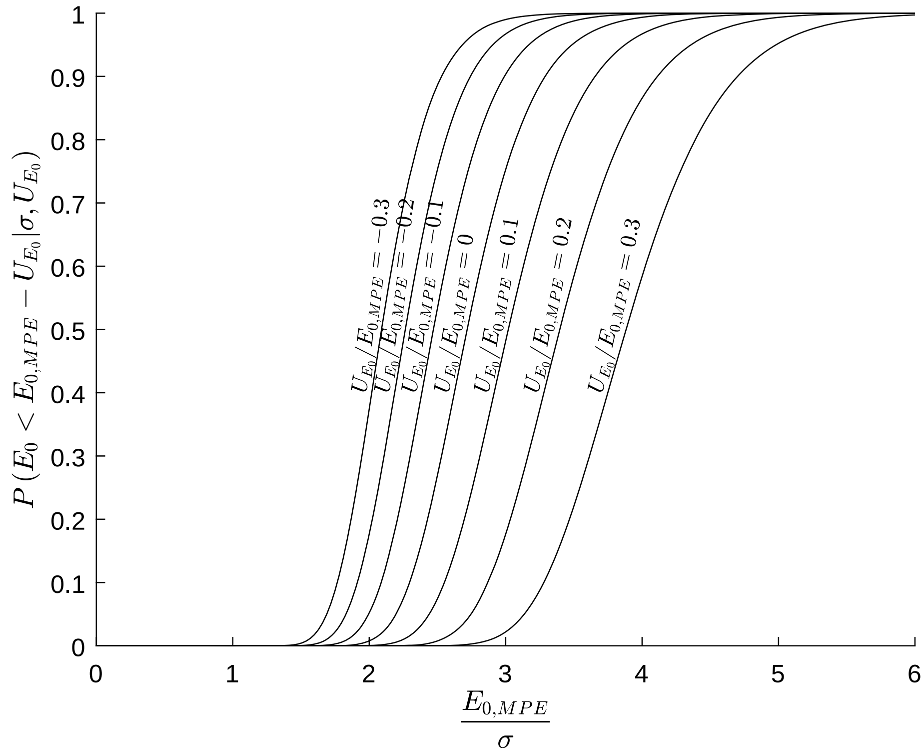


Figure 10: Operating characteristic curve when the test uncertainty is considered.

test is verified if $E_0 \leq E_{0,\mathrm{MPE}} + U_{E_0}$, but both positive and negative values have been considered for $U_{E_0}/E_{0,\mathrm{MPE}}$. Therefore, positive values of $U_{E_0}/E_{0,\mathrm{MPE}}$ refer to the acceptance test, and negative values $U_{E_0}/E_{0,\mathrm{MPE}}$ to the reverification test.

Considerations similar to those of the probing error test can be drawn. In the case of the reverification test, a large uncertainty increases the probability the test is passed. Usually this is not an issue (unless one suspects the CMS is not behaving correctly). In the case of the acceptance test instead, the problem of testing high accuracy CMSs arises. It is difficult to have a test uncertainty smaller than 0.1 µm, especially when lengths above 100 mm are involved. Testing CMMs for which $E_{0,\mathrm{MPE}} \leq 1$ µm can therefore be very difficult. In some cases CMSs manufacturers cannot state the real performance of their systems because they cannot test them.

5.3. Method validation

The results proposed in §5-5.1-5.2 are based on several assumptions:

- E_0 values are normally distributed;
- E_0 values are zero-mean;
- E_0 values are independent.

The assumption according to which $\frac{E_{0,\mathrm{MPE}}-U_{E_0}}{\sigma_j} = \frac{E_{0,\mathrm{MPE}}-U_{E_0}}{\sigma} \quad \forall j$ is actually required only to yield (10) and the related OC curves, but is not fundamental in the model.

Differing from the case of the $P_{\mathrm{Form.Sph.1\times 25:SS:Tact,MPE}}$ test, the hypotheses in this case are not verified. The $E_{0,\mathrm{MPE}}$ test was run 12 times on a "Zeiss Prismo 5 VAST HTG" tactile CMM. The value of $E_{0,\mathrm{MPE}}$ stated by the manufacturer is $2+\frac{L}{300}$ µm, where L is the measured length in [mm]. The data is made available [36]. Again normality has been verified by applying the Anderson-Darling test. The test has shown that the data are not normally distributed. In addition, Fig. 11 shows that often the E_0 values are not zero-mean. Independence was not tested. This was actually expected, as residual systematic errors, scale errors, and calibration errors of the reference standard are always present in CMM tests.

It is worth noting that the problem of non-normality can be solved by applying the Johnson transformation [39]. This has been tested: all datasets can be normalized. The non-zero mean makes the calculation of (10) more complicated, but still possible, as (9) changes into:

$$\begin{aligned} &P\left(|E_0| \leq E_{0,\mathrm{MPE}} - U_{E_0}\right) = \\ =&\Phi\left(\frac{E_{0,\mathrm{MPE}} - U_{E_0} - \mu_j}{\sigma_j}\right) - \Phi\left(\frac{U_{E_0} - E_{0,\mathrm{MPE}} - \mu_j}{\sigma_j}\right) \end{aligned} \tag{12}$$

With this in mind, OC curves are still a useful tool to evaluate the applicability of the $E_{0,\mathrm{MPE}}$ test.

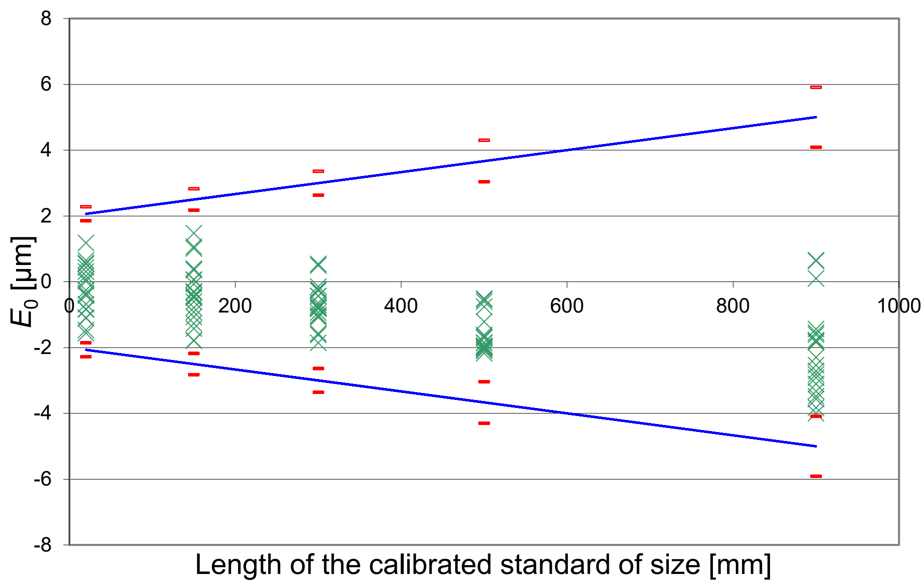


Figure 11: Example of a typical result of the $E_{0,\mathrm{MPE}}$ test. Green crosses indicate the measured values for E_0, red lines the test uncertainty, and blue lines the actual value of $E_{0,\mathrm{MPE}}$

6. Conclusions

The problem of verifying the performance of coordinate measuring systems is usually considered a black-box. In practice the test consists of a simple series of measurements, whose results must comply with some tolerance interval or, with a more proper term, maximum permissible error. But, as measurement results are always characterized by some variability, which is by itself strictly linked to the system performance, there is always the possibility of the test failing, despite the actual performance of the machine.

By analyzing the performance verification tests defined in the ISO 10360 series of international standards from a statistical point of view, the present work characterizes the probability that the tests fail under specific assumptions. The analysis relies in particular on the assumption that the measurement results are distributed according to a normal distribution, as commonly found in practical applications of measurement. The influence of the sample size and of the test uncertainty on the probability of the test failing has been assessed as well.

These results may help the coordinate measuring system experts to better understand the meaning of the results of the tests they are performing, by giving them a clearer look at the probabilities of the test failing. Moreover, they can help the coordinate measuring system manufacturers to state the value of the performance indices of their machines correctly, avoiding that under- and over-estimation. Under-estimation could lead to seeming non-competitive with other products; over-estimation could lead to frequent test failures, damaging the corporate image at the customer.

Finally, the ISO 10360 series test series tests should be discussed again considering their statistical implications. As any test, they are subject to "type I" (stating a well-functioning system malfunctioning) and "type II" (stating a malfunctioning system well-functioning) errors. The current formulation of the tests neglects this. Acceptance and reverification tests differ in the respective impact of these two kind of errors: therefore, a different consideration should be taken into account.

References

[1] R. G. Wilhelm, R. Hocken, H. Schwenke, Task Specific Uncertainty in Coordinate Measurement, CIRP Annals - Manufacturing Technology 50 (2) (2001) 553–563. doi:10.1016/S0007-8506(07)62995-3.

[2] International Organization for Standardization, ISO/TS 15530-1: Geometrical Product Specifications (GPS) - Coordinate measuring machines (CMM): Technique for determining the uncertainty of measurement - Part 1: Overview and metrological characteristics (2013).

[3] A. D. Hope, D. M. S. Blackshaw, The application of artefacts and lasers to performance verification of co-ordinate measuring machines, in: G. N. Peggs (Ed.), Laser Metrology and Machine Performance V, Vol. 34 of Engineering Sciences, WIT press, Birmingam, UK, 2001, pp. 117–126.

[4] U. Neuschaefer-Rube, W. Ehrig, M. Neugebauer, K. Wendt, Test procedures and artefacts for optical coordinate metrology, in: Tan, J. and Wen, X. (Ed.), Proceedings of the 5th International Symposium on Instrumentation Science and Technology, Vol. 7133 of Proceedings of SPIE, Bellingham, WA, 2009. doi:10.1117/12.810474.

[5] B. Acko, M. McCarthy, F. Haertig, B. Buchmeister, Standards for testing freeform measurement capability of optical and tactile coordinate measuring machines, Measurement Science & Technology 23 (9) (2012) 94013. doi:10.1088/0957-0233/23/9/094013.

[6] G. Moroni, W. Syam, S. Petrò, Performance Verification of a 4-Axis Focus Variation Co-Ordinate Measuring System, IEEE Transactions on Instrumentation and Measurement 66 (1). doi:10.1109/TIM.2016.2614753.

[7] S. El Asmai, F. Hennebelle, T. Coorevits, R. Vincent, J. F. Fontaine, Proposition of a periodic verification test for Articulated Arm Coordinate Measuring Machines using a small 3D artefact, Measurement: Journal of the International Measurement Confederation 154 (2020) 107472. doi:10.1016/j.measurement.2020.107472.

[8] Y. I. Curiel-Razo, O. Icasio-Hernández, G. Sepúlveda-Cervantes, J. B. Hurtado-Ramos, J. J. González-Barbosa, Leap motion controller three dimensional verification and polynomial correction, Measurement: Journal of the International Measurement Confederation 93 (2016) 258–264. doi:10.1016/j.measurement.2016.07.017.

[9] P. J. Swornowski, A critical look at the coordinate measuring technique, Mechatronics 23 (1) (2013) 80–93. doi:10.1016/j.mechatronics.2012.11.002.

[10] International Organization for Standardization, ISO/TS 23165: Geometrical product specifications (GPS) – Guidelines for the evaluation of coordinate measuring machine (CMM) test uncertainty (2006).

[11] A. Piratelli-Filho, B. Di Giacomo, CMM uncertainty analysis with factorial design, Precision Engineering 27 (3) (2003) 283–288. doi:10.1016/S0141-6359(03)00035-7.

[12] D. C. Montgomery, Design and Analysis of Experiments, 6th Edition, Wiley, New York, 2004.

[13] The American Society of Mechanical Engineers, ASME B89.4.10360.2: Acceptance Test and Reverification Test for Coordinate Measuring Machines (CMMs) Part 2: CMMs Used for Measuring Linear Dimensions (Technical Report) (2006).

[14] F. Franceschini, L. Settineri, Control charts for the on-line diagnostics of {CMM} performances, Int J Comput Integrated Manuf 13 (2) (2000) 148–156. doi:10.1080/095119200129993.

[15] F. Franceschini, M. Galetto, L. Settineri, On-Line Diagnostic Tools for {CMM} Performance, The International Journal of Advanced Manufacturing Technology 19 (2) (2002) 125–130. doi:10.1007/s001700200005.

[16] F. Franceschini, M. Galetto, A taxonomy of model-based redundancy methods for CMM online performance verification, International Journal of Technology Management 37 (1) (2007) 104–124. doi:10.1504/IJTM.2007.011806.

[17] International Organization for standadization, ISO 10360-5: Geometrical product specifications (GPS) — Acceptance and reverification tests for coordinate measuring systems (CMS) - Part 5: Coordinate measuring machines (CMMs) using single and multiple stylus contacting probing systems using (2020).

[18] B. Boeckmans, G. Probst, W. Dewulf, J. P. Kruth, Performance evaluation of a CMM tactile scanning probing system according to the revised ISO 10360 standard, in: European Society for Precision Engineering and Na-

notechnology, Conference Proceedings - 18th International Conference and Exhibition, EUSPEN 2018, 2018, pp. 479–480.

[19] B. Boeckmans, M. Vlaeyen, G. Probst, W. Dewulf, H. Haitjema, J. P. Kruth, CMM tactile probe performance evaluation according to ISO/DIS 10360-5:2018, in: Laser Metrology and Machine Performance XIII - 13th International Conference and Exhibition on Laser Metrology, Machine Tool, CMM and Robotic Performance, LAMDAMAP 2019, euspen, 2019, pp. 111–119.

[20] International Organization for Standardization, ISO 10360-1: Geometrical Product Specifications (GPS) - Acceptance and reverification tests for coordinate measuring machines (CMM) - Part 1: Vocabulary (2000).

[21] International Organization for Standardization, ISO 10360-2: Geometrical Product Specifications (GPS) - Acceptance and reverification tests for coordinate measuring machines (CMM) – Part 2: CMMs used for measuring linear dimensions (2009).

[22] International Organization for Standardization, ISO 10360-3: Geometrical Product Specifications (GPS) - Acceptance and reverification tests for coordinate measuring machines (CMM) – Part 3: CMMs with the axis of a rotary table as the fourth axis (2000).

[23] International Organization for Standardization, ISO 10360-6: Geometrical Product Specifications (GPS) - Acceptance and reverification tests for coordinate measuring machines (CMM) – Part 6: Estimation of errors in computing Gaussian associated features (2001).

[24] International Organization for Standardization, ISO 10360-7: Geometrical Product Specifications (GPS) - Acceptance and reverification tests for coordinate measuring machines (CMM) - Part 7: CMMs equipped with imaging probing systems (2011).

[25] International Organization for Standardization, ISO 10360-8: Geometrical product specifications (GPS) - Acceptance and reverification tests for coordinate measuring systems (CMS) -Part 8: CMMs with optical distance sensors (2013).

[26] International Organization for Standardization, ISO 10360-9: Geometrical product specifications (GPS) - Acceptance and reverification tests for coordinate measuring systems (CMS) - Part 9: CMMs with multiple probing systems (2013).

[27] International Organization for Standardization, ISO 10360-10: Geometrical product specifications (GPS) - Acceptance and reverification tests for coordinate measuring systems (CMS) - Part 10: Laser trackers for measuring point-to-point distances (2016).

[28] International Organization for Standardization, ISO 10360-12: Geometrical product specifications (GPS) - Acceptance and reverification tests for coordinate measuring systems (CMS) - Part 12: Articulated arm coordinate measurement machines (CMM) (2016).

[29] International Organization for Standardization, ISO 10360-2: Coordinate metrology - Part 2: Performance assessment of coordinate measuring machines (1994).

[30] International Organization for Standardization, ISO 10360-2: Geometrical Product Specifications (GPS) - Acceptance and reverification tests for coordinate measuring machines (CMM) – Part 2: CMMs used for measuring size (2001).

[31] International Organization for Standardization, ISO 10360-5: Geometrical Product Specifications (GPS) - Acceptance and reverification tests for coordinate measuring machines (CMM) – Part 5: CMMs using single and multiple stylus contacting probing systems (2010).

[32] International Organization for Standardization, ISO 14253-1: Geometrical product specifications (GPS) - Inspection by measurement of workpieces and measuring equipment - Part 1: Decision rules for proving conformance or nonconformance with specifications (2017).

[33] H. O. Hartley, The Range in Random Samples, Biometrika 32 (3/4) (1942) 334–348.

[34] E. S. Pearson, H. O. Hartley, The Probability Integral of the Range in Samples of n Observations From a Normal Population, Biometrika 32 (3/4) (1942) 301–310.

[35] Joint Committee for Guides in Metrology, ISO/IEC GUIDE 99:2007(E/F): International vocabulary of metrology - Basic and general concepts and associated terms (VIM) (2007).

[36] S. Petrò, Support data for the "A Statistical Point of View on the ISO 10360 Standard for Coordinate Measuring Systems Verification" Article (2020). doi:10.17632/wr447vnw8x.

[37] T. W. Anderson, D. A. Darling, Asymptotic Theory of Certain "Goodness of Fit" Criteria Based on Stochastic Processes, The Annals of Mathematical Statistics 23 (2) (1952) 193–212. doi:10.1214/aoms/1177729437.

[38] D. C. Montgomery, G. C. Runger, Applied statistics and probability for engineers, 3rd Edition, John Wiley & Sons, New York, 2003.

[39] S. Kotz, N. L. Johnson, Process Capability Indices, Chapman and Hall/CRC, Boca Raton, Florida, 1993.