Li, T., Leach, Richard K., Jung, L, Jiang, Xiang and Blunt, Liam

Comparison of Type F2 Software Measurement Standards for Surface Texture

National Physical Laboratory

Report

NPL REPORT ENG 16

Comparison of Type F2 Software Measurement Standards for Surface Texture

T Li, R K Leach, X Jiang and L A Blunt

June 2009

National Physical Laboratory | Hampton Road | Teddington | Middlesex | United Kingdom | TW11 0LW

Switchboard 020 8977 3222 | **NPL Helpline** 020 8943 6880 | **Fax** 020 8614 0446 | **www.npl.co.uk**

Comparison of Type F2 Software Measurement Standards for Surface Texture

T Li[1], R K Leach[2], X Jiang[1] and L A Blunt[1]

[1.] Centre for Precision Technologies, University of Huddersfield
[2.] Industry & Innovation Division, National Physical Laboratory

ABSTRACT

This report describes the comparison of three type F2 software measurement standards, in the form of internet-based software, for testing surface texture software by NIST, NPL and PTB. A set of reference data is examined by the type F2 standards and three commercial software packages. Calculations of the parameters given in ISO 4287 (1996) are compared and the differences between the software packages are discussed.

ISSN 1754-2987

National Physical Laboratory
Hampton Road, Teddington, Middlesex, TW11 0LW

Approved on behalf of the Managing Director, NPL
by Susan Evans, Director, Industry and Innovation Division

Contents

1 Introduction

1.1 Background

Surface metrology is often a critical part of product quality control and has a close relationship to product design and manufacture. To maintain traceability for surface measurement, widely used measurement procedures and conditions are standardised, and many material artefacts are available to calibrate the vertical and horizontal characteristics of a probing system, the probe tip condition and the ability of the instrument to measure surface texture parameters (Figure 1 and Table 1). In the last two decades, there have emerged many advanced surface characterisation techniques to enhance the "quality" of the mathematical model to represent the geometrical properties of an engineered surface, such as areal characterisation, fractal parameters and various filtering methods [1, 2]. The use of digital techniques makes it possible to implement complex mathematical models and this has led to the need for more complex standardisation and calibration procedures. The assessment of the quality of a mathematical process includes two parts: the quality of the mathematical models and algorithms, and the quality of the numerical implementation, *i.e.* software. Compared with variations in measured surface texture caused by the surface inhomogeneity [3], the measurement environment [4], the choice of sampling interval [3], and different data collection methods (stylus, optical, AFM[5]), the variation contributed by software might seem at first sight to be insignificant. However, without a formal validation, this consideration remains intuitive and, when challenged, a convincing response is rarely provided.

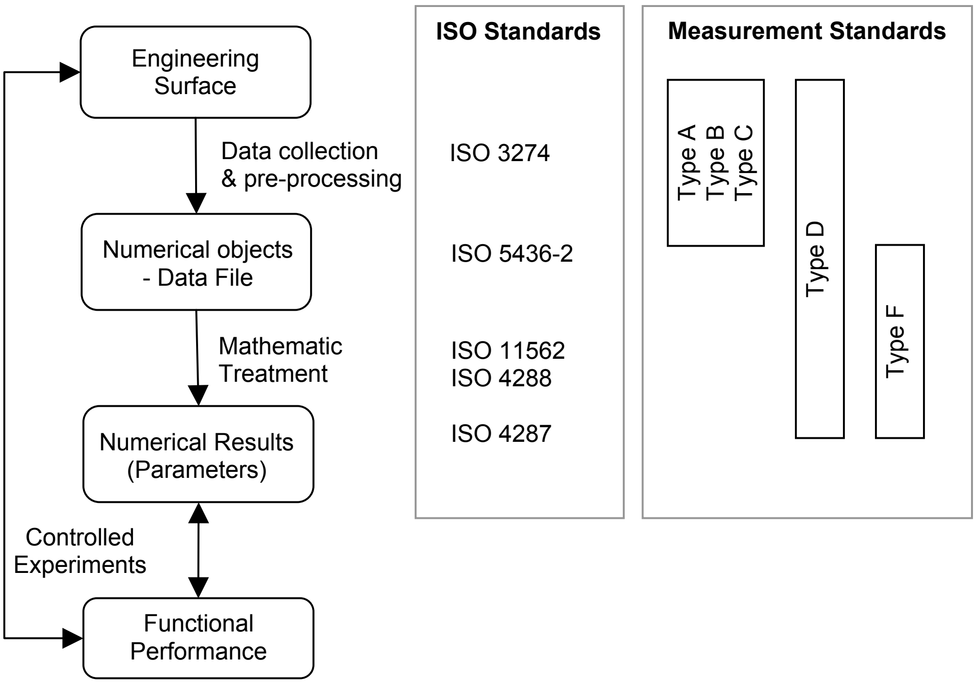


Figure 1: Surface measuring procedure, some relevant ISO standards and measurement standards[1]

Some related work on validation of software in surface texture can be found in the literature. In 1988, Scott introduced the concept of "the reference surface measuring instrument", a mathematically defined conceptual surface measuring instrument to provide the reference for a

[1] The term "*controlled experiments"* is defined as "*the activity to recognise parameters that are functionally important for each application and determine their control values*" [6]. Thus, the use of the measurement standards does not cover the assessment of controlled experiments.

specific instrument [7]. The checking of the algorithms for calculating parameters was proposed in calibration procedures of surface profile and areal instruments [6, 8]. Stout *et al.* used simulated specimens with known characteristics in the form of data files to carry out software verification [8]. Two surface parameter algorithm comparisons were undertaken by the National Institute of Standards and Technology (NIST) in the USA, and showed good agreement for most parameters and most software packages but with some disagreement for a few parameters [9, 10]. Another comparison among seventeen national metrology institutes in Europe reported large differences in some height and spacing parameters [11]. These references show that software is a primary contributor to the variation of the final results.

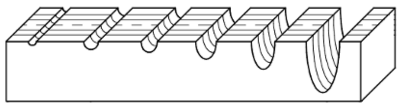
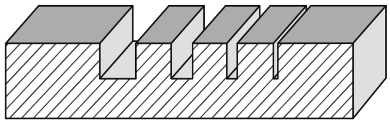
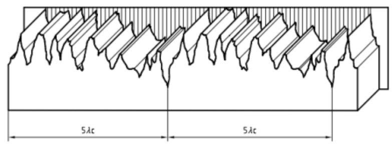
Type A: Depth measurement standard		Calibration of vertical displacement
Type B: Tip condition measurement standard		Calibration of the state of tip
Type C: Spacing measurement standard		Calibration of the horizontal displacement
Type D: Roughness measurement standard		Total calibration of the instrument (parameter *Ra* & *Rz*)
Type F: Software measurement standard	http://www.ptb.de/en/org/5/51/517/rptb_web/wizard/greeting.php http://syseng.nist.gov/VSC/jsp/index.jsp http://www.npl.co.uk/server.php?show=ConWebDoc.160	Calibration of software algorithms (filter and all parameters)

Table 1: Examples of surface texture calibration standards described by ISO 5436 [12, 13]

ISO 5436-2 (2000) introduced into international standardisation the concept of software measurement standards in the form of reference data (*type F1 or softgauges*) and reference software (*type F2*) to verify surface metrology software [2]. Physikalisch-Technische Bundesanstalt (PTB) in Germany has developed reference software to test software for roughness analysis [14]. NIST has developed a surface metrology algorithm testing system to serve as master algorithms to validate surface analysis software [15]. The National Physical Laboratory (NPL) in the UK, with University of Huddersfield and Taylor Hobson, has developed software measurement standards for surface topography software assessment [16]. In this report, these reference software packages are referred to as type F2 software measurement standards because they were developed to address the same requirement defined in ISO 5436-2 (2000) and are maintained by a national measurement institute (NMI) to serve as a metrological tool. All type F2 software measurement standards (later in this report referred to as simply type F2 standards) claim to have been developed to high standards and to have been thoroughly tested. However, some initial comparisons have already shown some disagreement among these type F2 standards [15]. Therefore, there are some essential questions that need to be addressed before these type F2 standards can safely and reliably be used. Some important questions are:

1) Is it safe to ignore calibration software?
2) Do the type F2 standards qualify to be used as calibration tools?

[2] In the context of surface texture, a *software measurement standard* is a metrological tool (Clause 5.1-Note 9, VIM: 2007), and is akin to a primary standard in measurement, such as a kilogram mass to which secondary standards are compared for calibration purposes.

3) How does one make a judgement when there is discrepancy between an industrial software package and a type F2 standard, or even between two type F2 standards?

To address these questions, a comparison of the three NMI's type F2 standards has been undertaken. Six reference data sets are used to compare the results obtained from these type F2 standards together with three widely used commercial software packages. The sources of variation, including the specification variation, computational errors and data uncertainty, are analysed.

1.2 Participants and test software[3]

The participants in this comparison are listed in Table 2. The detailed descriptions of the type F2 standards are available online and links are provided in Table 3. In addition, three commercial software packages were used in this comparison. They are named as CA, CB and CC for commercial protection.

1.	Physikalisch-Technische Bundesanstalt, Bundesallee 100, 38116 Braunschweig, Germany (PTB).
2.	National Institute of Standards and Technology, Metrology Building, Gaithersburg, Maryland 20899-0001, USA (NIST).
3.	National Physical Laboratory, Teddington, Middlesex TW11 0LW, United Kingdom (NPL).

Table 2: Participating laboratories

Institute	Software
PTB	Ref_soft_PTBIDL and Ref_soft_PTBweb[4] www.ptb.de/en/org/5/51/517/rptb_web/wizard/greeting.php
NIST	Internet Based Surface Metrology Algorithm Testing System syseng.nist.gov/VSC/jsp/index.jsp
NPL	nplsmd1.01 www.npl.co.uk/server.php?show=ConWebDoc.160
CA	Commercial software package A
CB	Commercial software package B
CC	Commercial software package C

Table 3: Type F2 standards and commercial packages

1.3 Scope

To address the major concern of metrologists, this comparison was mainly focused on the metrological traceability of the measurement results. Many software quality characteristics according to ISO/IEC 9126 [17] (*i.e.* usability, efficiency, maintainability, portability, *etc.*) have not been assessed in this report. The parameters to be compared here are those defined within ISO 4287 (1996), and the related standard documents are:

ISO 3274: 1996 and Cor 1: 1998, *Geometrical Product Specifications (GPS) — Surface texture: Profile method — Nominal characteristics of contact (stylus) instruments.*

ISO 4287: 1997, Cor 1: 1998 and Cor 2: 2005, *Geometrical product specifications (GPS) — Surface texture: Profile method — Terms, definitions and surface texture parameters.*

ISO 4288: 1996 and Cor 1: 1998, *Geometrical Product Specifications (GPS) — Surface texture: Profile method — Rules and procedures for the assessment of surface texture.*

---

[3] The term *test software* used throughout this report refers to the software under test, including the type F2 standards and commercial packages.

[4] PTB provides reference software in the form of a desktop version and web version.

ISO 11562: 1996 Cor 1: 1998, *Geometrical Product Specifications (GPS) — Surface texture: Profile method — Metrological characterization of phase correct filters.*

ISO 5436-1: 2000, *Geometrical Product Specifications (GPS) — Surface texture: Profile method; Measurement Standards — Part 1: Material measures.*

ISO 5436-2: 2000, Cor 1: 2006 and Cor 2: 2008, *Geometrical Product Specifications (GPS) — Surface texture: Profile method; Measurement Standards — Part 2: Software measurement standards.*

ISO 1302: 2002, *Geometrical Product Specifications (GPS) — Indication of surface texture in technical product documentation.*

ISO 14253-1: 1998, *Geometrical Product Specifications (GPS) — Inspection by measurement of workpieces and measuring equipment — Part 1: Decision rules for proving conformance or non-conformance with specification.*

ISO/IEC Guide 99: 2007, *International vocabulary of metrology — Basic and general concepts and associated terms (VIM).*

1.4 The definition of measurement uncertainty

Measurement uncertainty quantifies the dispersion of values attributed to a measurand. For decades, measurement uncertainty has been formulated in terms of probability theory. The most commonly used procedure for calculating measurement uncertainty is described in the *Guide to the Expression of Uncertainty in Measurement* (the GUM) [18]. The components of measurement uncertainty are grouped into two categories: type A and type B, according to whether they were evaluated by a statistical analysis of the values from a series of measurements or otherwise, respectively.

In the GUM, the uncertainty in communication and cognition level is considered to be negligible with respect to the other components of measurement uncertainty. However, ISO/TC 213 has recognised that the disagreement of the measurement results from two different parties is often a result of different interpretations of the specification, and/or different choices of influential conditions that are not pre-specified [19]. The concepts of method uncertainty and specification uncertainty have been introduced to quantify those uncertainties due to the "lack of information". In ISO/TS I7450-2, the uncertainty is divided into correlation uncertainty, specification uncertainty and measurement uncertainty [20]. Measurement uncertainty includes two components, method uncertainty and implementation uncertainty. It is noted that these terms and concepts are still evolving and are subject to modification and refinement as the work of ISO/TC 213 progresses.

In VIM (2007), the concept of definitional uncertainty is introduced by ISO and IEC as a component of measurement uncertainty arising from the finite amount of detail in the definition of a measurand [21]. The definition of metrological traceability and measurement uncertainty is defined as:

> *"**Metrological Traceability**: property of a measurement result whereby the result can be related to a reference through a documented unbroken chain of calibrations, each contributing to the measurement uncertainty."*
>
> *"**Measurement Uncertainty**: non-negative parameter characterizing the dispersion of the quantity values being attributed to a measurand, based on the information used."*
>
> *(ISO/IEC VIM (2007))*

According to VIM (2007), definitional uncertainty is related to the description detail of a measurand. Many description details are inside software packages. Software, therefore, is a contributor to measurement uncertainty.

In this report, measurement uncertainty refers to the definition given by VIM (2007).

1.5 Acknowledgements

The authors would like to thank the following participants in the comparison: Dr Ludger Koenders, Dr Rolf Krüger-Sehm and Dr Lena Jung (PTB), Dr Ted Vorburger and Dr Son Bui (NIST) and Prof. Paul Scott (Taylor Hobson).

2 Differences between national software packages

NIST, PTB and NPL have some differences in their interpretation of the ISO standard documents. Most of these differences are detailed in Section 5, and a summary of them is given below.

- **Continuous model or discrete model:** a significant difference is the type of mathematical model used to represent a surface in the ISO standard documents. Should this be a continuous model or a discrete model? Figure 2 shows the phases of applying the ISO standard documents a software implementation with the example of the *Ra* parameter. The measurement procedure and condition are standardised. Some calculation phases are described clearly in mathematical models (for example, the parameter *Ra*); some are introduced as concepts (for example, discrimination of a profile element); while others are provided in the discrete form directly (for example, the parameter *Rdq*). ASME B46 provides both analytical and digital definitions [22]. PTB and NIST use a discrete model while NPL uses a continuous model.

- ***P*-parameters and *W*-parameters:** ISO 5436 (1996) introduces *P*-parameters and *W*-parameters into ISO standards. There are differences about the sampling length of *P*-parameters and *W*-parameters (see Section 5.3).

- **Levelling:** there is a difference about the levelling operator (see Section 5.2).

- ***RSm*/*PSm*/*Pc*/*Rc* parameters:** the interpretations of these parameters vary among NIST, NPL and PTB. This is because the definition of a profile element is ambiguous in ISO 5436 (see Section 5.5).

- **SMD format:** ISO 5436-2 introduces the SMD file format as the protocol of software calibration. However, different interpretations of the SMD file format are provided by NPL and PTB [11].

This report discusses the measurement conditions and parameter specifications mainly based on *The Specification of Parameters* issued by NPL [23]. However, this does not mean that other interpretations are incorrect or improper. In most cases, the differences are due to the incomplete and imperfect nature of definitions given within ISO standard documents. From this point of view, all interpretations that are considered in ISO documents are valid. However, to be unambiguous, the definitions should be clarified in the future by improved ISO standards.

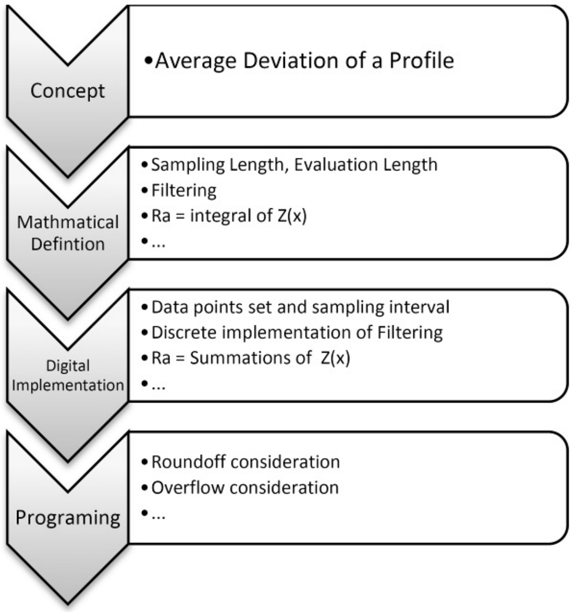


Figure 2: Phases of applying the ISO standards

3 Methodology

3.1 Classification of errors and uncertainties

In VIM (2007), the concept of measurement uncertainty is extended to cover definitional uncertainty. The current versions of the GUM and ISO/TS 17450 are based on the previous definition of measurement uncertainty. There is no guidance on how to express and estimate the definitional uncertainty[5]. In this report, we combine the concepts from the GUM and ISO/TS 17450 to describe the sources of errors and uncertainties from a software perspective. The variation in the final results obtained from different software comes from a variety of sources (see Figure 3) and is discussed below.

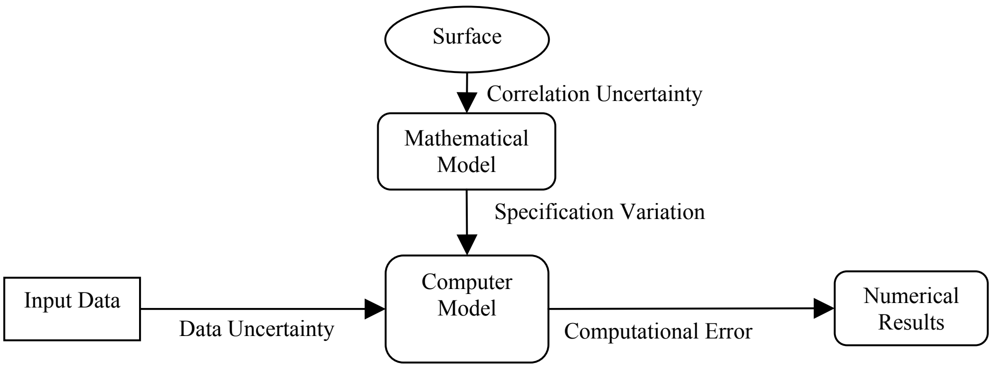


Figure 3: Sources of error and uncertainty

3.1.1 Correlation uncertainty

The correlation uncertainty is defined within ISO/TS 17450 as the difference between a functional requirement and an actual geometric specification. Correlation uncertainty varies from application to application and should be evaluated by experiment. Thus, software measurement standards do not take correlation uncertainty into consideration.

[5] The VIM(2007) does not define the concept of measurand definition, while it is still an open topic in the field of metrology [24-28].

3.1.2 Specification variation

To develop surface metrology software, a full mathematical specification is required. The term specification variation is used to describe the variation of the full mathematical specification between different software packages[6]. The main sources of specification variation are listed below.

1. **Incomplete definition in an ISO standard.** For the ISO standards documents, it is impossible and unnecessary to detail every measurement procedure and condition because standards need to achieve a balance between over-specification and lack of focus. Incompleteness of the definition can lead to ambiguity and implementation in different ways by different software developers based on their own knowledge and experience.
2. **Imperfect definition in an ISO standard.** We may never claim perfection in standards because the standards development process is in a continuous improvement mode[7]. Instrument manufacturers may adhere to the definitions or make an improvement.
3. **The transformation between different models.** Generally, the conceptual model and mathematical model are based on a continuous profile, while the implementation model is based on a discrete profile. There are errors and uncertainties when mapping one model to another.
4. **Mistakes.** There are human errors due to the misinterpretation and misunderstanding of ISO standards.

Many specification variations are negligible with respect to the other components of measurement uncertainty. For software validation, the main tasks are to evaluate the effect of the influencing conditions which are not standardised, and that of significant mistakes.

3.1.3 Computational error

A well-known source of error is contributed by the limitation of the computer. Due to the computer's binary representation of numbers with finite precision, there are two types of errors when engaging in numerical computation: rounding error and truncation error that are contributed by the computer hardware and software separately [29]. Another source of inaccuracy is the numerical algorithm itself. There can exist many different ways to compute a quantity, and some are better than others. Such errors are referred to as computational errors in this report.

3.1.4 Data uncertainty

In addition to the above, there are errors within the input data (the measurement data) due to noise, variation of the measurement environment, *etc.* These errors are referred as data uncertainty in this report. Investigating and quantifying such data uncertainty is the subject of uncertainty evaluation (GUM), based on a statistical model.

3.2 Evaluation of uncertainty

The evaluation of the data uncertainty would require knowledge of the uncertainties associated with the measured co-ordinates of the points. The type F2 standards of NPL and PTB do not provide the (measurement) uncertainties associated with the calculated reference values for the surface texture parameters. NIST's type F2 standards evaluate the uncertainty due to data uncertainty using Monte Carlo simulation (MCS) [15]. NIST simulates the measurement error

[6] We do not use the term "specification uncertainty" (defined in ISO/TS 17450-2) due to 1) specification uncertainty is used to quantify the communication uncertainty between designers and metrologists; 2) it is difficult to distinguish between the specification uncertainty, method uncertainty and implementation uncertainty in surface texture measurement.

[7] However, we could say that the standards are based on the best knowledge for surface texture at the time of writing.

by adding normal distributed random noise to the co-ordinates of each data point. In this comparison, MCS will be used to estimate the contribution of data uncertainty to the final results.

Due to the fact that software packages often do not disclose detailed information of the algorithm or methods of implementation used, the full specification is normally not readily apparent from documents within the software packages. Therefore, a comparison of output of the different software packages for the same input - black box testing - is the preferred method used in the testing of software implementations.

3.2.1 Evaluation of specification variation

The specification variation is generally fixed within a particular application. Thus, we evaluate it by a testing procedure that is divided into the following stages:

1. identifying the influential conditions which are ambiguous,
2. listing all possible interpretations of those conditions,
3. estimating the effect on the final results.

In stage 1, we use two methods to identify the source of specification variation, a) the top-down method, by analysing the definition given by ISO standards, and b) the reversing method, by analysing the unexpected variation of the final results obtained from test software.

In stage 2, the possible interpretations are listed by reviewing the existed publications, and available documents of software packages.

In stage 3, the effect on the final results is estimated using mathematically synthesised profiles and measured profiles. Measurement profiles could be used as the reference data for purposes as follows:

- To demonstrate reproducibility of a surface measurement from a software perspective; and it also shows the stability and robustness of a parameter definition from an ISO standards perspective;
- To estimate the contribution of variation from a software perspective in the final results by comparing with the contribution from data uncertainty.
- To evaluate the robustness of a software algorithm in practice by transforming or changing measured data.

3.2.2 Evaluation of computational errors

Another key issue that needs to be addressed is the “true” value of measurement results of the reference data sets. In a mathematically well-defined model, this can be achieved by using both a high precision processor and data. NIST implements this method to produce the *Statistical Reference Datasets* (StRD) to benchmark statistical software packages [30]. The reference results were obtained from a multiple precision FORTRAN pre-processor and reference data sets with 500 decimal digits of accuracy. It has been used to benchmark many well-known statistical software packages such as Microsoft Excel (version: 97, 2000, XP, 2003, 2007) [31-34]. Another method is to start with some reference results and produce the corresponding reference data set by a data generator through the null-space approach [35]. NPL has implemented this approach to test a range of software packages [36, 37]. These methods are too advanced and not suitable for this comparison due to the following:

- The methods are developed for investigating the fitness for purpose of a software implementation of an algorithm to solve a specified, and well-defined, mathematical model. Unfortunately, there are significant variations in mathematical models among surface metrology software.

- The indication of a surface texture parameter is normally only needed two or three decimal digits of accuracy. Thus, using an extra high precision method (both processor and data) is too advanced for this comparison.

In this comparison, we use a simplified method to produce the reference pair, the reference data set and corresponding "true" result, as illustrated in Figure 4. This method uses mathematically designed synthetic reference data, by sampling mathematically defined functions for which the profile parameter values are know *a priori*, whose certified results can be calculated directly. This approach has the following advantages:

- The basic profile, such as a sinusoidal profile, represents the fundamental concept of the mathematical treatment for a surface profile. A reference pair is based on an unambiguous concept and mathematical model (which should be stated and agreed) with the smallest specification variations. Thus, the computing errors can be estimated and F2 software measurement standards can be certified directly and straightforwardly.
- The "true" results can be calculated by algebraic calculation. The results can then be used to evaluate the numerical errors.
- Furthermore, synthetic reference data can often be designed that discriminates between known perturbations from the specified parameter definition.

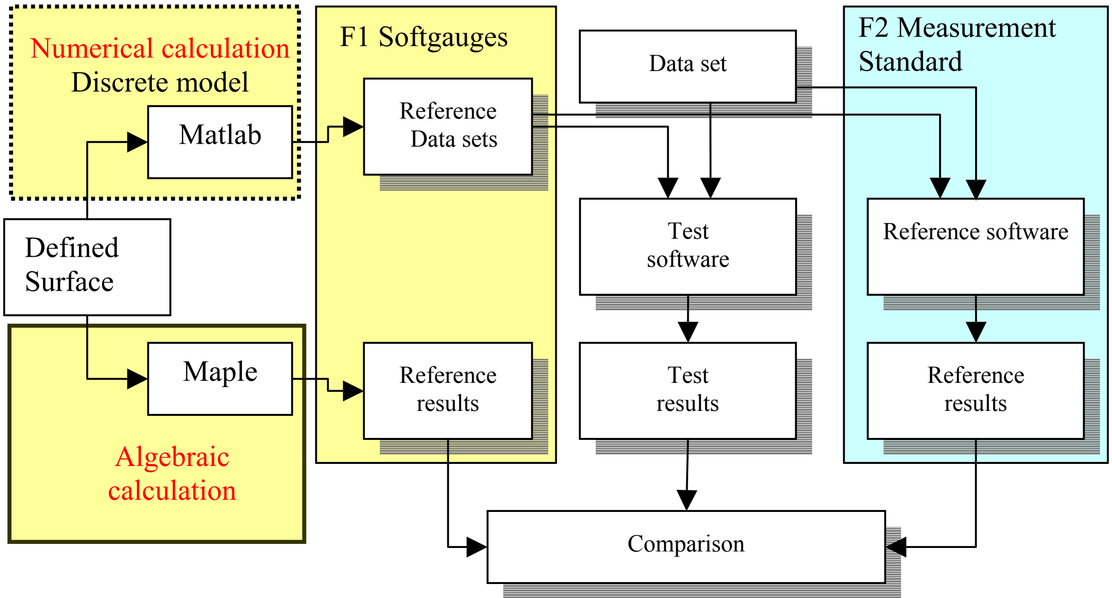


Figure 4: Procedure of using algebraic calculation and numerical calculation

3.3 Decision rules and uncertainty management

A key question concerns the comparison of the results delivered by the type F2 standards and commercial packages. The comparison should be objective and address the requirements of the application. The result of the comparison is the means by which a decision is made about the fitness-for-purpose of the F2 software standards and commercial packages.

If y^{test} and y^{ref} denote, respectively, the test[8] and reference results, then

$$d_A\left(y^{\text{test}}, y^{\text{ref}}\right) = \left|y^{\text{test}} - y^{\text{ref}}\right|,$$

and, for $y^{\text{ref}} \neq 0$,

$$d_R\left(y^{\text{test}}, y^{\text{ref}}\right) = \frac{\left|y^{\text{test}} - y^{\text{ref}}\right|}{\left|y^{\text{ref}}\right|},$$

are metrics for the numerical correctness of the test result that measure, respectively, the absolute and relative differences between the test and reference results.

[8] Test result is the result obtained from one of type F2 standards or commercial packages.

It is unnecessary (and perhaps unreasonable) to expect that the absolute difference between the test and reference results is comparable to the computational precision of the arithmetic used to deliver the test result. (For 16-digit arithmetic, for example, the computational precision is of the order of 10^{-16}.) If the developer of the software has made a claim about the numerical correctness of the results returned by the software, then this can be used as the basis for setting a tolerance against which to compare the calculated value of the absolute difference. If the user of the software has documented a requirement on the numerical correctness of the result, then this can also be used as a basis of the comparison. If the uncertainty associated with the test result is available (evaluated in terms of the uncertainties associated with the measured data defining the surface profile), then it may be sufficient to require that the calculated value of the absolute difference is smaller (by several orders of magnitude, say) than this uncertainty.

Fitness for purpose can also mean that the effects arising from the use of the approximate mathematical model, approximate algorithm, *etc.* are quantitatively small compared to those effects arising from the data, the latter being described by uncertainty. In cases that the use of an approximate mathematical model, approximate algorithm, *etc.*, are shown not to be fit for purpose it is necessary either to correct for these effects, for example, to use a more sophisticated mathematical model, algorithm, *etc.*, or to quantify the effects and include them as additional contributions in the uncertainty evaluation.

4 Specification of reference pairs

4.1 Reference data sets

This comparison used six reference data sets as listed in Table 4. The cosine wave has a wavelength of 160 µm and amplitude of 2 µm. To minimise the effect of sampling conditions on evaluation, the same sampling interval, number of sampling lengths and number of points are consistent over the six profiles (see Table 5). Based on the information gained from a consultation exercise [16], four measured surface profiles, a milled, a polished, an EDM and a ground surface were used to address industrial requirements. Use of measured profiles as the reference data sets can assess the variation caused by different software packages. Using a reference data pair can demonstrate the robustness of parameter definitions in ISO standards.

Data files	Description
Cos.smd	A cosinusoidal profile
EDM.smd	Measured profile of an EDM surface
Mill.smd	Measured profile of a milled surface
Ground.smd	Measured profile of a ground surface
Ground2.smd	Same data set as Ground.smd with the order of the data points reversed to simulate the opposite measuring direction.
Polish.smd	Measured profile of a polished surface

Table 4: Reference data sets

Condition	Setup
Sampling interval	0.25 µm
Sampling length	0.8 mm
Evaluation length for P-parameter	7 × 0.8 mm = 5.6 mm
Evaluation length for R-parameter	5 × 0.8 mm = 4.0 mm
Total number of points	22401

Table 5: Profile measurement condition

4.2 Reference result

For the simulated profile Cos.smd, the reference result is obtained from algebraic calculation in Maple 10.03. There are discretisation errors built into the synthetic reference data, therefore, the term *expected result* is used to refer the “true” value of reference results. For the measured profiles using in the comparison, *the reference result* is the non-weighted mean of results obtained from the three F2 standards.

4.3 The performance metrics for Cos.smd

Sinusoidal artefacts have been widely used in the calibration of surface measuring instruments since they were first introduced by Sharmen in 1967 [38]. Sinusoids are insensitive to many measurement conditions. Some research comparison results are listed in Table 6. To reproduce these research results, the effect of software is assumed to be insignificant. The metric for *Ra* is set based on the reproducibility of these research results from the respect of software.

No	Reference[1]	Metric for *Ra*[2]
1	H. Haitjema (1998) estimated the uncertainty of roughness parameters using styles instrument [4]. It was shown that the uncertainty of *Ra* and *RSm* for a sinusoidal artefact (nominal *Ra*: 2.9 µm and *RSm*: 100 µm) are 0.25 % and 0.03 % (at 95 % confidence).	0.04 %
2	NIST F2 standard is able to simulate the measurement error by adding the normal distributed random noise to each data point. The uncertainty of *Ra* and *RSm* for Cos.smd (nominal Ra: 636.62 nm and *RSm*: 160 µm) is ± 2.68 nm and ± 6.97 µm (at 95 % confidence).	0.07 %
3	T. Vorburger *et. al.* (2007) undertook a comparison between optical and stylus methods [5]. For a sinusoidal specimen (nominal *Ra*: 500 nm and *RSm*: 50 µm), the *Sa* - *Ra* differences was 6 nm obtained from difference type of instruments.	0.1 %
4	T Thomas (1982) investigated the (in)homogeneity of some typical manufactured surfaces [3]. The variation of 1.8~3 % for *Ra* were found on RTH reference standards (Two-dimension sinusoidal surfaces, nominal *Ra*: 0.27 µm).	0.6 %

Note: 1. Based on the assumption that software used in this research work is qualified.

2. The pass margin set as 1/3 of value of uncertainty and 1/10 of absolute difference.

Table 6: The performance metrics for Cos.smd

4.4 The performance metrics for measured profiles

Table 7 lists the percentage coefficients of variation from place to place on a manufactured surface studied by Thomas [3]. ISO 4288 introduces the "The 16 %-rule" and "The max.-rule" for comparison of the measured values within tolerance limits. For the measured profiles used in this comparison, we provide six significant digits that include false precision and guard digits. For measured profiles, the software effect is considered as insignificant when the relative difference of results obtains from the test software and a reference result is less than 0.5 %. Therefore, we set the "pass margin" as 0.5 % in this comparison.

	Milled	Ground
Ra /%	17 ~ 65	7~80
Rq /%	15 ~ 61	9 ~ 56
*Rsk**	0.35 ~ 0.75	0.22 ~ 0.73

(Except* which is an absolute value)

Table 7: Percentage coefficients of variation from place to place on a manufactured surface [3]

5 Specification variation

In the following sections the square brackets refer to the associated software packages or ISO documents.

5.1 Pre-process operator

[CC]

Deletes the last point of the data set when inputting a data file.

[CB]

Adds an extra point when opening a data file. The method seems to (by analysing the output file):

- Add an extra point in the middle of the profile.
- Change the height value of the last point to be equal to the penultimate point.
- Adjust the value of the spacing to keep the sampling length consistent.

Comments:

The behaviour of [CB] and [CC] suggests that they implement a conversion between a point-based length definition and an interval-based length definition (see Section 5.6.1).

5.2 Levelling

[PTB]

The least-squares method is a mandatory operator.

[NPL]

The start point of a software measurement standard is the primary profile that does not contain form errors.

[Others]

The least-squares method is an optional operator.

Comments:

Levelling is the operator to remove tilt from a profile. The least-squares method is widely used in industrial practice. However, conventional least-squares is not an appropriate method for a sine wave [39], and the least-squares normal method is more appropriate [40].

5.3 Evaluation length and sampling length

5.3.1 R-parameters

5.3.1.1 End effect of filtering

See Appendix C.2.

5.3.1.2 RSm

[ISO]

RSm is defined within a sampling length.

[NPL][PTB]

RSm is assessed over the evaluation length.

Comments:

There is no mandatory method to reduce the end effects of the filtering operators[9], thus different types of filters could cause different parts of a measured profile to be evaluated.

Defining the *RSm* parameter over the evaluation length will deliver a more stable result (see Section 5.5).

5.3.2 P-parameters

[ISO]

lp is equal to the length of the feature being measured.

[PTB][CA]

lp(default): The remaining profile after removing one *λc* cut-off at each end of profile

[NPL][NIST][PTB][CB][CC]

lp(default): All the measured points in a data file.

Comments:

For interpretation of CA, it should be noticed that:

- Currently, there is no standardised method to reduce the end effect of a filtering operator.
- The value of a *P*-parameter will depend on the selection of the *λc* value.

Thus, to avoid ambiguity, this interpretation requires specifying the *λc* value when stating the *P*-parameters (it does not follow ISO 1302: 2000).

5.3.3 W-parameters

There is no common understanding of the meaning and use of waviness parameters [39]. ISO 4287 defines the sampling length of *W*-parameters based on the cut-off of the profile filter *λf*. Some industrial practice ignores this filter step and uses a sampling length *λw* equal to the cut-off wavelength *λc*.

5.4 Filtering

5.4.1 *λs* filtering

[CC]

The *λs* profile filtering is a mandatory part of the *λc* profile filtering (following ISO 3274 - Table 1).

[Others]

The *λs* profile filtering is an optional operator (ISO 4287).

5.5 Profile element

5.5.1 Joint direction

[ISO]

A profile element is defined as a peak with a following valley (ISO 4287: 1997 - figure 3), or a valley with a following peak (ISO 4287:1997 - figure 10).

[NIST]

---

[9] There is a standard under development, *ISO/CD TS 16610-28, Geometrical product specifications (GPS) -- Filtration -- Part 28: Profile filters: End effects.*

A profile element is defined as a peak with a following valley.

Comments:

The ambiguity of ISO's definition is shown in Figure 5. In this case, there are two valid profile elements with different height and width value. Profile element should be defined as a concept only. All the calculations should be defined based on features (a peak or a valley).

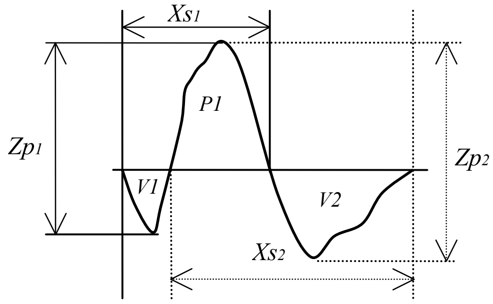


Figure 5: Ambiguity of the definition of profile element, the height and width of the profile element could be 1) Zp_2 and Xs_2; 2) Zp_1 and Xs_1; 3) $(Zp_2 + Zp_1)/2$ and $(Xs_2 + Xs_1)/2$

5.5.2 Incomplete portion

[ISO]

The *incomplete portion* is the feature at the beginning or end of a sample length (e.g. the gray areas in Figure 6), and a handling method is provided as:

> *"The positive or negative portion of the assessed profile at the beginning or end of the sampling length should always be considered as a profile peak or as a profile valley. When determining a number of profile elements over several successive sampling lengths, the peaks and valleys of the assessed profile at the beginning or end of each sampling length are taken into account once only at the beginning of each sampling length."* (*ISO 4287: 1997 Clause 3.2.7)*

[PTB][NIST][NPL]

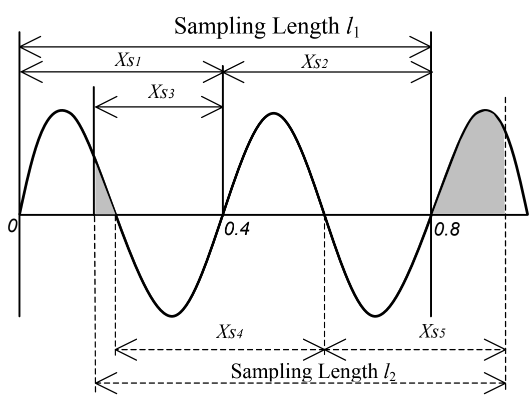
The NMIs do not follow the definition according to ISO due to its ambiguity.

[NPL][PTB]

Assess the *RSm* and *PSm* parameters within the evaluation length and discard the incomplete portion at each end of the evaluation length.

Comments:

Figure 6 illustrates the imperfection of the ISO definition.



Sampling Length	Measuring Direction	*RSm* /mm
l_1	⇨	0.4
l_1	⇦	0.4
l_2	⇦	0.375
l_2	⇨	0.325

Figure 6: Ambiguity of definition of the incomplete portion: results vary from 0.325 mm to 0.4 mm while the true value is the 0.4 mm in this case

5.5.3 Insignificant features

5.5.3.1 Identification of insignificant feature

[ISO]

Height discrimination (*Hd*) for profile elements: 10 % of the value of an amplitude parameter. Spacing discrimination (*Wd*) for profile elements: 1 % of the sampling length.

[NPL]

2H method (see Figure 7):

Height discrimination (h_d) for profile feature: ± 5 % of *Rz*.

Spacing discrimination (w_d) for profile feature: 0.5 % of sampling length.

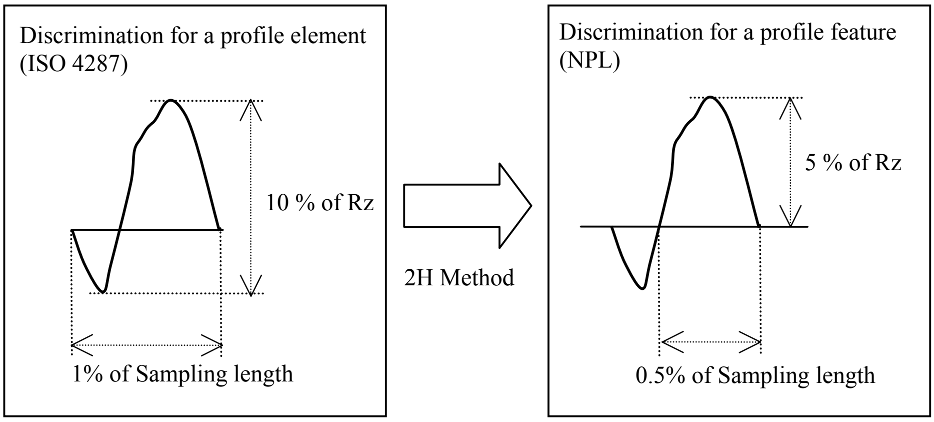


Figure 7: NPL's implementation of 2H method for height and spacing discrimination

[PTB]

2H method:

Height discrimination (h_d) for profile feature: ± 5 % of *Rq* on the Internet portal.

The desktop version of the software (Ref_soft_PTBIDL) allows the setting of the vertical and horizontal discrimination. In this comparison the height discrimination of ± 5 % of *Rz* was used.

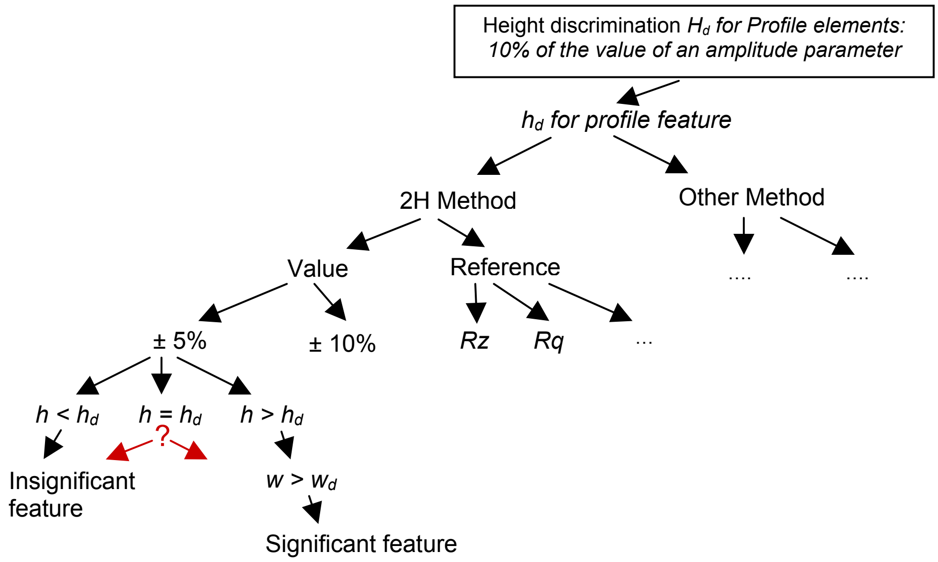
[NIST]

2H method:

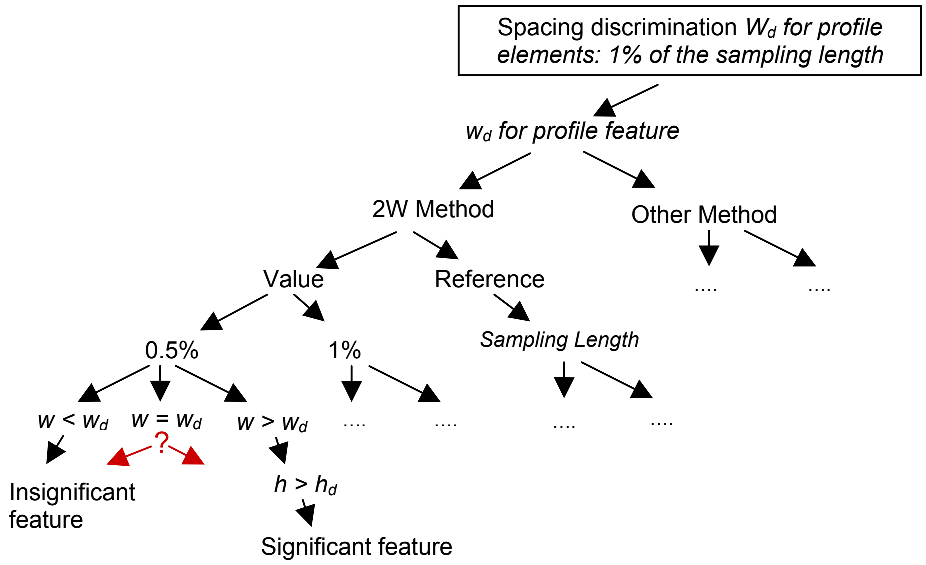
Height discrimination (h_d) for profile feature: ± 5 % of *Rz*.

Comments:

To identify the insignificant features, the concept of discrimination of a profile element is introduced into ISO standards. The definition of profile element is ambiguous (see section 5.5.1). All the software implementations, thus, identify the insignificant peaks/valleys directly. There is a significant specification variation for applying discrimination to peak and valley as illustrated in Figure 8.



(a) Height Discrimination



(b) Spacing Discrimination

Figure 8: Ambiguity of discrimination

5.5.3.2 Handling method for insignificant feature

After being identified, the insignificant features can be classified in to five types by their position, that of:

1) insignificant features at each end of the sampling length,
2) insignificant features at each end of the evaluation length,
3) incomplete insignificant features at each end of the sampling length,
4) incomplete insignificant features at each end of the evaluation length and
5) insignificant features in the middle of the sampling length.

According to its type, an insignificant feature can be discarded or merged into its neighbouring feature.

[ISO]

There is no specified method for handling an insignificant feature.

[NPL]

1) *RSm* is calculated over the evaluation length to reduce the number of type 1 and type 3 features.
2) Discard type 2 and type 4 features.
3) For type 5 features, start with the smallest segments and combine with two adjacent segments.

[PTB]

See the description of the calculation of *XSm* and *Xc* parameters on the Internet portal.

Comments:

Figure 9 illustrates five combination algorithms for removing the type 5 insignificant features. The method 1), 2), 3), and 4) were studied by Leach and Harris and showed the difference of the results of *RSm* obtained by different methods could up to 12 % [41]. Scott proposed method 5 and proved its stability by the representation theory of measurement [42].

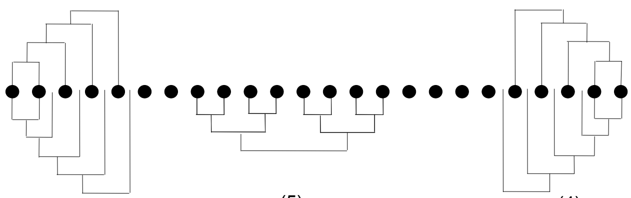


Figure 9: Ambiguity of combination method (the black points represent the crossing points of a profile)

5.6 Discrete interpretation

5.6.1 Length definition

The total profile is defined in the digital form as a discrete profile (ISO 3274), while most definitions of further operations within ISO 4287 and ISO 11562 still take the continuous form. There are two typical discrete algorithms used to calculate the distance between two points in the horizontal direction (see Table 8).

	Point-based definition	**Interval-based definition**
Where n is the number of points between its ends, and i is the sampling interval.	The length l calculated as $l=(n-1)\times i$,	The length l calculated as $l=n\times i$,
In mathematics, in a particular geometry, a distance function satisfies the following conditions: 1) $d(x,y)\geq 0$, and $d(x,y)=0$ if and only if $x=y$. 2) $d(x,y)=d(y,x)$ 3) $d(x,z)\leq d(x,y)+d(y,z)$	It satisfies condition 1, 2 and 3.	It satisfies condition 1 only.

Table 8: Length algorithm

Table 9 lists the influence of the length definition in this testing. If the last point is a significant point such as a crossing point or the highest/lowest point, the effect will be significant. PTB and NIST use the interval-based definition in sampling length and evaluation length while NPL use a point-based definition. The two definitions are used in various ways in many software implementations.

	Data file : Cos.smd Evaluation length: 5.6 mm sampling interval: 0.25 µm Lc: 0.8 mm Ls: 2.5 um	
	Point-based definition	**Interval-based definition**
The number of data points:		
within a data profile	22401	22400
within a evaluation length	16001	16000
within a sampling length	3201	3200
The number of data points used to implement Gaussian weight convolution.		
Ls filtering	11	10
Lc filtering	3201	3200

Table 9: The effect of the different length definition

Mean line crossing-points

[NPL]

Use a natural cubic spline to interpolate through the discrete data values.

Comments:

The mean line is a base to which feature parameters are referred. Unfortunately, most of the mean line crossing-points are excluded in the measured data point set, and the position of a mean line crossing-point is generally estimated from its neighbouring points. Thus, there are many different algorithms to estimate a crossing-point and the method uncertainty is introduced to the final results. In addition, there is error when calculation based on the measuring data point set without those mean line crossing-points [43]. Brennan recommended the need to include implied mean line crossing points simply by interpolating the data where these occur and provide each profile peak or valley element with calculated boundary values.

5.7 Two examples: *Ra* and *RSm*

Table 10 and Table 11 estimate the effect of specification variation in the final result of test software.

	Operator	Estimation of Effect (Weight in the final result[1])
5.1	Pre-process operator	[CB] *** [CC]
5.2	Levelling	[PTB] *
5.3.1.1	End effect of Lc filtering	[CC] *
5.4.1	Ls filtering	[CC] *
5.6.1	Length Definition	
5.6.2	Mean Line Crossing-point	

[1] ***** More than 100 % variation
**** More than 10 % variation
*** More than 1 % variation
** More than 0.1 % variation
* More than 0.01 % variation

Table 10: The estimation of effect of specification variation for *Ra*

	Operator	Estimation of Effect (Weight in the final result[1])
5.1	Pre-process operator	[CB] *** [CC]
5.2	Levelling	
5.3	Evaluation Length	
5.4.1	Ls filtering	[CC]*
5.5	Profile element	
5.5.1	Joint Direction	[NIST] ***
5.5.2	Incomplete portion	[ALL] ***
5.5.3.1	Identifying the insignificant feature	[ALL] *****
	Handling of insignificant feature	[ALL] ****
5.6.1	Length Definition	
5.6.2	Mean Line Crossing-point	

[1] ***** More than 100 % variation
**** More than 10 % variation
*** More than 1 % variation
** More than 0.1 % variation
* More than 0.01 % variation

Table 11: The estimation of effect of specification variation for *RSm*

6 Evaluation and discussion

6.1 Evaluation of Cos.smd

Table 18 and Table 19 in Appendix B present results for Cos.smd obtained from the test software, together with the mean and standard deviation of the results.

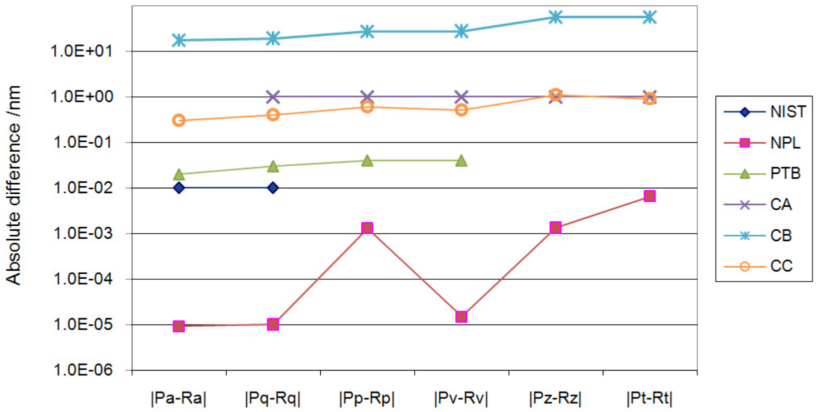
6.1.1 Effect of levelling operator

For PTB's type F2 standard and CB, the effect of the levelling operator is analysed in Appendix C.2. Due to the different length definition, the last point of Cos.smd is not taken into account by PTB's type F2 standard and CB. Thus *Pp* is 0.04 nm less than the expected value while *Pv* is 0.04 nm greater.

6.1.2 Effect of filtering

	Cos |*Pa-Ra*| /nm	Cos |*Pq-Rq*| /nm	Cos |*Psk-Rsk*|	Cos |*Pku-Rku*|	Cos |*Pp-Rp*| /nm	Cos |*Pv-Rv*| /nm	Cos |*Pz-Rz*| /nm	Cos |*Pt-Rt*| /nm	Cos |*Pc-Rc*| /nm	Cos |*PSm-RSm*| /µm	Cos |*Pdq-Rdq*|
NIST	1.0E-02	1.0E-02	1.3E-04	2.0E-05	0.0E+00	0.0E+00	0.0E+00	0.0E+00	0.0E+00	0.0E+00	2.0E-05
NPL	9.1E-06	1.0E-05	2.2E-09	1.0E-09	1.3E-03	1.5E-05	1.3E-03	6.5E-03	2.9E-05	0.0E+00	*
PTB	2.0E-02	3.0E-02	1.8E-04	0.0E+00	4.0E-02	4.0E-02	0.0E+00	0.0E+00	0.0E+00	0.0E+00	4.9E-05
CA	0.0E+00	1.0E+00	0.0E+00	0.0E+00	1.0E+00	1.0E+00	1.0E+00	1.0E+00	*	*	1.9E-05
CB	1.8E+01	1.9E+01	8.8E-04	1.4E-04	2.7E+01	2.7E+01	5.5E+01	5.5E+01	8.5E+01	2.9E+00	8.0E-04
CC	3.0E-01	4.0E-01	1.0E-04	0.0E+00	6.0E-01	5.0E-01	1.1E+00	9.0E-01	1.1E+00	4.6E-01	1.7E-05

Table 32 in Appendix C.3 and Figure 10 present the effect of filtering by comparing the resulting *P*-parameter and *R*-parameter obtained for Cos.smd. No visible difference would be expected (see Appendix C.3). The three type F2 standards perform well with less than 0.1 nm absolute difference. For CA and CC, the absolute difference is between 0.1 nm to 1 nm. CB performs poorly with more than 10 nm difference..



1) There are some missing points due to zero values cannot be plotted in this log chart.
2) Some results may be overstated or understated due to rounding effect.

Figure 10: Effect of filtering

6.1.3 *RSm/PSm*

The value of *PSm/RSm* should be 160 µm, which is the "true" value for this cosine wave. If we strictly adhere to ISO 4287: 1997, to evaluate within every sampling length and discard incomplete portions at the end of sampling length,

RSm = 152 µm and *PSm* = 158.857 µm.

If, following ASME B46.1-2002 [22], we evaluate within the evaluation length and discard incomplete portions at the end of evaluation length,

RSm = 158.4 µm.

If we use the interval-based length definition to define the sampling length, and the point-based length definition to calculate the width of a profile element within each sampling length, and discard the incomplete portions at each ends,

RSm = 159.95 µm.

If we use the interval-based length definition to define the evaluation length, and the point-based length definition to calculate the width of a profile element within the evaluation length, and discard the incomplete portions at each ends,

RSm = 159.99 µm.

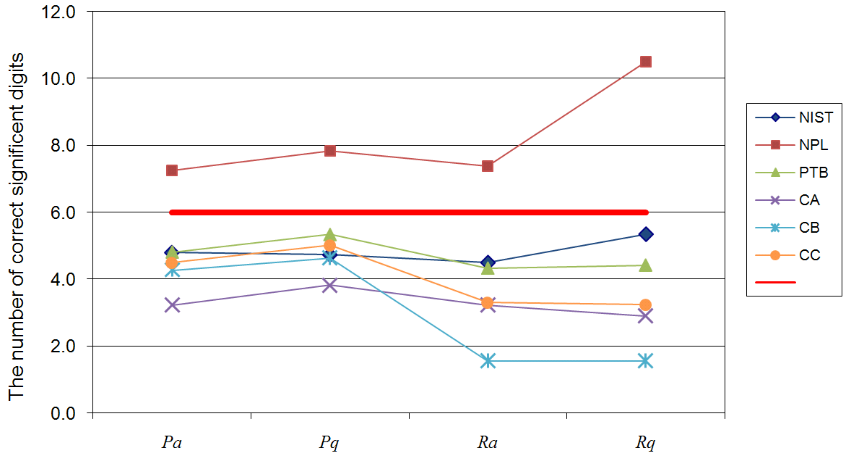
Table 12 presents the *PSm/RSm* results for Cos.smd. It shows that NIST, NPL, PTB and CA have fixed the distortion introduced by ISO 4287. Commercial package CC adheres to ASME B46.1-2002. CB delivers significantly different results. PTB's and NPL's type F2 standards perform well in this test. PTB's type F2 standard delivers a small error due to a different length definition.

	PSm /µm	*RSm /µm*
NIST	160.00	160.00
NPL	160.00	160.00
PTB	159.99	159.99
CA	-	160.00
CB	**158.89**	**156.01**
CC	**158.86**	**158.40**

Table 12: Influence of incomplete portion for *RSm* and *PSm*

6.1.4 Computational error

Table 33 in Appendix C.4 and Figure 11 show the number of correct significant digit of the results obtained from all software implementations. PTB's and NIST's type F2 standards give results where only last digits are inaccurate. NPL's type F2 standard delivers seven to ten accurate significant digits in this case, even higher than the precision of measuring data. For commercial packages, the level of precision reduces significantly with the range of LRE values falling between 1.6 to 4.6. Some of the results provide two to four false significant digits. Commercial package CB provides less than two accurate significant digits.



(The red line indicates the significant digits of the measuring data within data file.)

Figure 11: The number of correct significant digits

6.1.5 Software effect

Table 13 presents the software effect by comparing variation due to different software and data errors. The mean standard deviations of the results for Cos.smd obtained from different software have been used. These are compared to the measurement uncertainty due to data errors calculated by the NIST software. The software variation among the three software measurement standards is small compared with the data uncertainty. The exception to this is the *RSm* parameter and *Rc* parameter due to their ambiguous definitions within standards. When commercial packages are compared , the software variations are significant.

Parameter	*Ra*	*Rq*	*Rku*	*Rp*	*Rv*	*Rz*	*Rt*	*RSm*	*Rdq* [1]
Uncertainty calculated by NIST software [2]									
Mean value R x	636.59	706.93	1.5009	1012.62	1012.94	2024.96	2035.69	160.33	0.0288
Uc(Rx)	1.34	1.09	0.003	1.72	1.74	2.53	4.16	3.49	0.00063
Deviation by 3 software measurement standards									
ΔRx	0.02	1.3E-02	4.6E-08	6.2E-04	6.9E-06	6.2E-04	3.1E-03	4.7E-03	5.0E-06
$\Delta Rx/uc(Rx)/\%$	1.53	1.23	0.00	0.04	0.00	0.02	0.07	0.14	0.79
Deviation by all software (3 software measurement standards and 3 industrial packages)									
ΔRx	6.56	7.20	6.0E-05	10.21	10.21	20.34	20.34	1.49	0.00
$\Delta Rx/uc(Rx)/\%$	489.33	660.39	1.99	593.82	586.89	803.90	489.06	42.61	49.34
Deviation by all software without industrial package CB									
ΔRx	0.22	0.43	3.9E-08	0.52	0.49	0.66	0.60	0.64	0.00
$\Delta Rx/uc(Rx)/\%$	16.61	39.38	1.29E-03	29.96	28.16	26.27	14.45	18.31	2.92

Note: [1] Parameter *Rdq* without NPL
[2] See Table 6.

Table 13: Software effect vs. data effect

6.1.6 Fitness for purpose

Table 14 presents the performance of test software for *Ra* of Cos.smd. It indicated that the three type F2 standards could be used to reproduce all the measurement tasks listed in Table 6, while CA and CC could do task 2, 3 and 4. CB should not used to reproduce all four tasks.

	NIST	NPL	PTB	CA	CB	CC
Task	1,2,3,4	1,2,3,4	1,2,3,4	2,3,4	**None**	2,3,4

Table 14: The performance of software implementations for *Ra* of Cos.smd (assessed by the effect of the reproducibility of measurement tasks listed in Table 6[10])

6.2 Evaluation of measured profiles

Table 20 to 24 in Appendix B present results for measured profiles obtained from test software, together with the mean and standard deviation of the results. Figure 17 to 38 in Appendix C show the relative differences between test results and reference results. Table 15 and 16 present the percentage of coefficients of variation among the three type F2 standards and three commercial packages. For the three type F2 standards, most of the relative differences are less than 0.5 %. NPL's type F2 standard delivers slightly greater values of *Rp*, *Rv*, *Rz*, *Rt*, *Pp*, *Pv*, *Pz* and *Pt* due to its interpolating method, and only one result is greater than 0.5 % (*Rp* for Polish.smd). For *PSm*, *RSm*, *Pc* and *Rc*, variations are significant as the result of ambiguous definition (see Figure 12-15). Together with the three commercial packages, most of relative differences for *R*-parameters are more than 0.5 %.

	Ra	*Rq*	*Rsk*	*Rku*	*Rp*	*Rv*	*Rz*	*Rt*	*Rc*	*RSm*	*Rdq*
Cos.smd	0.00	0.00	-	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.02
EDM.smd	0.00	0.14	0.19	0.06	0.05	0.06	0.06	0.05	**3.18**	**6.95**	0.02
Mill.smd	0.00	0.15	0.12	0.01	0.14	0.18	0.16	0.13	**3.78**	**7.10**	0.07
Polish.smd	0.01	0.13	0.03	0.02	0.38	0.13	0.17	0.15	**10.76**	**24.98**	0.04
Ground.smd	0.00	0.25	0.13	0.00	0.05	0.03	0.04	0.02	**6.39**	**16.66**	0.01

Table 15: Percentage of coefficients of variation among three type F2 standards

	Ra	*Rq*	*Rsk*	*Rku*	*Rp*	*Rv*	*Rz*	*Rt*	*Rc*	*RSm*	*Rdq*
Cos.smd	**1.03**	**1.02**	-	0.00	**1.03**	**1.03**	**1.02**	**1.02**	**1.72**	0.93	**1.13**
EDM.smd	**1.03**	**1.36**	**30.17**	0.89	**7.10**	**5.65**	**2.71**	0.60	**6.10**	**9.08**	**2.31**
Mill.smd	**2.60**	**2.32**	**8.19**	**2.85**	**5.59**	**12.28**	**1.69**	**1.19**	**6.25**	**45.41**	**17.17**
Polish.smd	**1.03**	**1.21**	**1.29**	**1.98**	**7.30**	**7.95**	**3.11**	**1.24**	**13.66**	**33.77**	**9.12**
Ground.smd	0.80	0.84	**15.90**	**2.12**	**13.90**	**7.03**	**1.24**	**0.62**	**9.91**	**22.30**	**4.37**

Table 16: Percentage of coefficients of variation among three type F2 standards and commercial packages

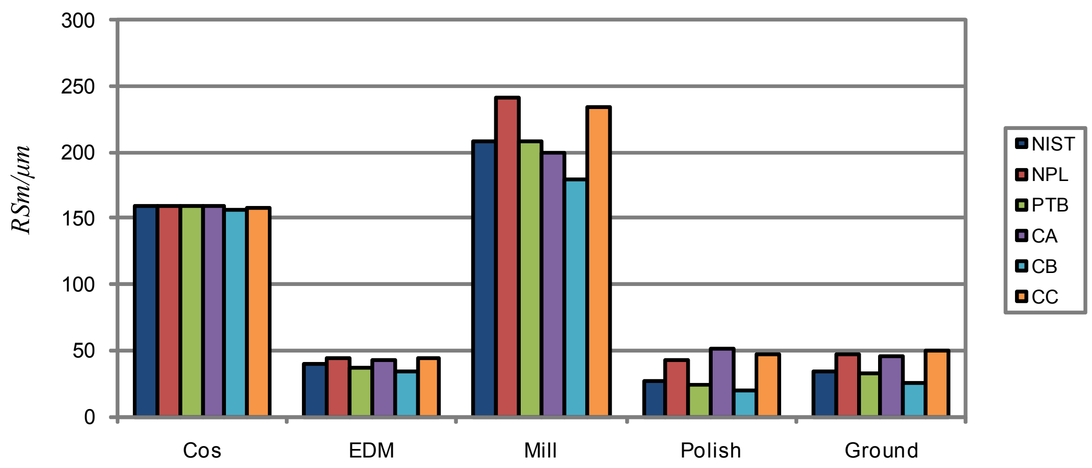


Figure 12: Results of *RSm* parameter

[10] It is based on the assumption that software used to produce those research results is qualified.

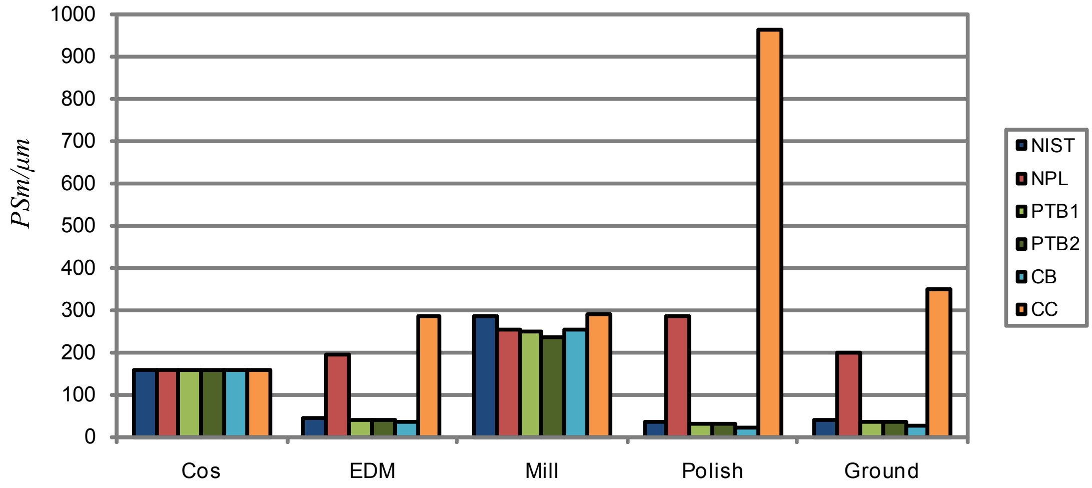


Figure 13: Results of *PSm* parameter

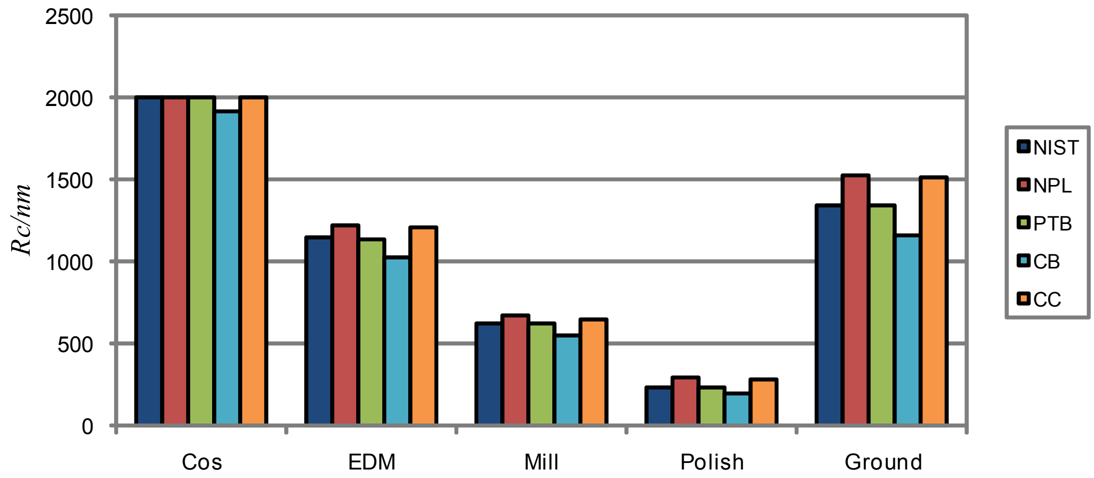


Figure 14: Results of *Rc* parameter

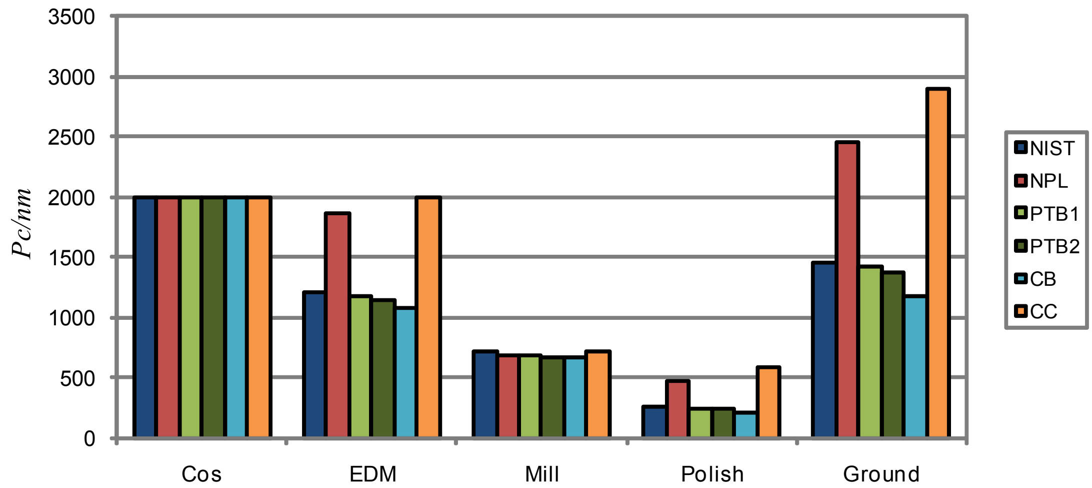
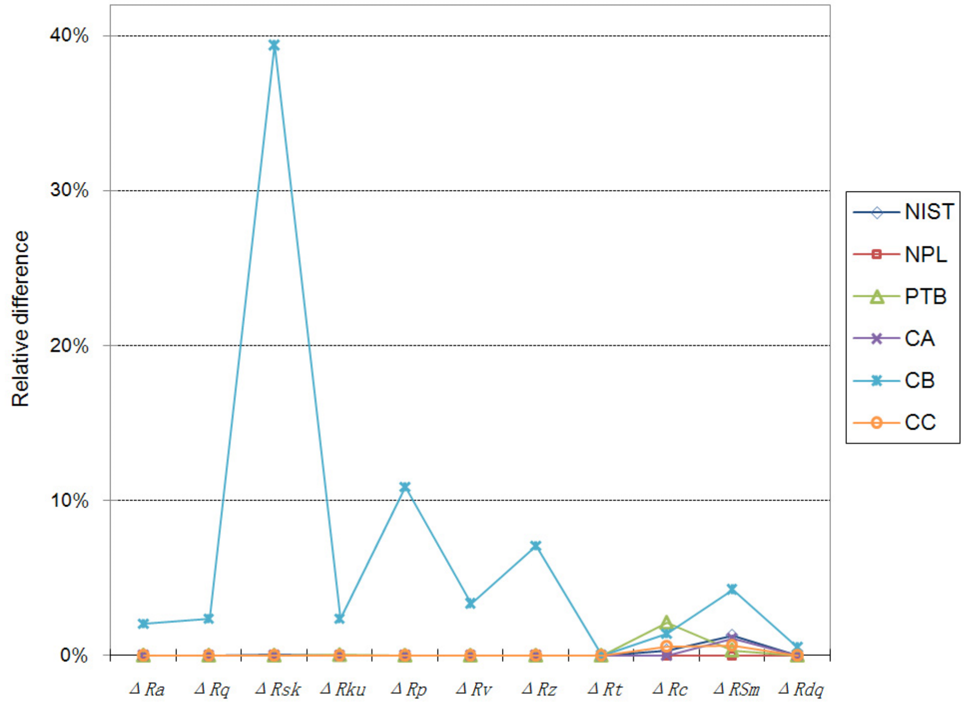


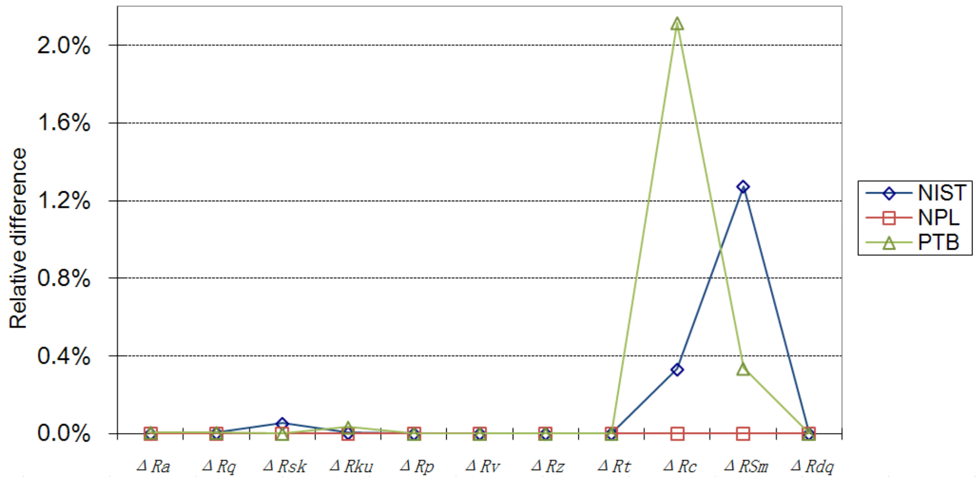
Figure 15: Results of *Pc* parameter

Table 34 in Appendix C.5 and Figure 16 present the reproducibility and stability of parameters by comparing the results obtained from Ground.smd and Ground2.smd that are same data set with different order to simulate difference measurement directions. For the three F2 software standards, the relative difference of *Ra*, *Rq*, *Rsk*, *Rp*, *Rv*, *Rz*, *Rt* and *Rdq* fall with the range of

0.05 %, while the difference of *Rc* and *RSm* fall with the range of 2.1 % and 0.76 %. For commercial package CA and CC, the relative difference of *RSm* is 1 % and 0.64 %. Commercial package CB delivers significant errors for all parameters, the relative difference is up to 40 %.



(a)



(b)

Figure 16: The effect of direction

7 Conclusions

In general the results for *R*-parameters obtained from the three type F2 standards are in good agreement. The exceptions are the *RSm* and *Rc* parameters. The reason for this disagreement is the ambiguous and unstable definitions given within ISO standards. The three software measurement standards performed better than the three commercial packages by giving high precision results and their specifications adhere closely to ISO standards.

For commercial packages, the results indicate that software is a primary contributor to variability in the results of surface profile measurement. One commercial software package delivered significantly different results. The variation of the results obtained from these software packages is even greater than the variation caused by the surface inhomogeneity, variation of measurement environment and different data collection methods. Therefore, it is not safe to ignore the calibration of software embedded within a surface instrument. Some particular conclusions are as follows:

- The current specifications of parameters *Ra, Rq, Rsk, Rp, Rt and Rz* are clearly defined and stable. The three type F2 software standards are qualified to provide accredited results for those parameters for commercial packages.
- The specifications of parameters *RSm* and *Rc* are ambiguous and unstable. The variation of *RSm* is significant. The revised specification of *RSm* proposed by Scott [42] is mathematically stable in this test.
- The specifications of *P*-parameters are unambiguous in standard documents. However, there are different understandings of the meaning of *P*-parameters, which leads to different interpretations.
- The effect of rounding error is insignificant in the test. The major contributor to the variation is the specification variation.

In addition, there are significant variations on the results of W-parameter as well[11].

8 Reference:

1. Jiang, X., et al., *Paradigm shifts in surface metrology. Part I. Historical philosophy.* Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences, 2007. **463**(2085): p. 2049-2070.
2. Jiang, X., et al., *Paradigm shifts in surface metrology. Part II. The current shift.* Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences, 2007. **463**(2085): p. 2071-2099.
3. Thomas, T.R. and G. Charlton, *Variation of roughness parameters on some typical manufactured surfaces.* Precision Engineering, 1981. **3**(2): p. 91-96.
4. Haitjema, H., *Uncertainty analysis of roughness standard calibration using stylus instruments.* Precision Engineering, 1998. **22**(2): p. 110-119.
5. Vorburger, T., et al., *Comparison of optical and stylus methods for measurement of surface texture.* International journal of advanced manufacturing technology, 2007. **33**: p. 110-118.
6. Song, J.F. and T.V. Vorburger, *Standard reference specimens in quality control of engineering surfaces.* J. Res. Natl. Inst. Stand. Technol., 1991. **96**: p. 271-289.
7. Scott, P.J., *Developments in surface texture measurement.* Surface Topography, 1988. **1**: p. 153-163.
8. Stout, K.J., et al., *The Development of Methods for the Characterization of Roughness in Three Dimensions*. 1993, Luxembourg: Publication No. EUR 15178 EN of the Commission of the European Communities.
9. Bui, S.H., et al. *Surface metrology software variability in two-dimensional measurement*. in *Pro. ASPE 2003*. 2003.
10. Bui, S.H., et al., *Internet-based surface metrology algorithm testing system.* Wear, 2004. **257**(12): p. 1213-1218.
11. Euromet600 - Final Report. *Euromet Project 600 – comparison of surface roughness standards*. 2004 Available from: www.bipm.org/utils/common/pdf/final_reports/L/S11/EUROMET.L-S11.pdf.

[11] We do not present the results of W-parameters in this report because they are seldom used and have same "nature" as R-parameters and P-parameters. Moreover, the specifications of *W*-parameters defined within ISO standards do not accepted by industry.

12. ISO 5436-1:2000, *Geometrical Product Specifications (GPS) -- Surface texture: Profile method; Measurement standards -- Part 1: Material measures*: International Organization for Standardization.
13. ISO 5436-2:2001, *Geometrical Product Specifications (GPS) -- Surface texture: Profile method; Measurement standards -- Part 2: Software measurement standards*: International Organization for Standardization.
14. Jung, L., et al. *Reference software for roughness analysis-features and results*. in *The XI international colloquium on surfaces*. 2004. Shaker Verlag.
15. Bui, S.H. and T.V. Vorburger, *Surface metrology algorithm testing system.* Precision Engineering, 2007. **31**(3): p. 218-225.
16. Blunt, L., et al., *The development of user-friendly software measurement standards for surface topography software assessment.* Wear, 2008. **264**(5-6): p. 389-393.
17. ISO/IEC TR 9126-2:2003, *Software engineering – Production Quality – Part 2: External metrics*: International Organization for Standardization.
18. IEC BIPM ISO IFCC I IUPAC, *Guide to the expression of uncertainty in measurement*. 1995, Geneva.
19. ISO TC213. *ISO 2004 Business Plan of ISO/TC213*. 2004 [cited; Available from: http://isotc213.ds.dk.
20. 17450-2:2002, I.T., *Geometrical product specifications (GPS) -- General concepts -- Part 2: Basic tenets, specifications, operators and uncertainties.*
21. ISO/IEC Guide 99:2007, *International vocabulary of metrology -- Basic and general concepts and associated terms (VIM)*: International Organization for Standardization.
22. ASME B46.1-2002, *Surface texture (surface roughness, waviness, and lay)*, New York: Am Soc Mech. Eng.
23. NPL. *Specification of parameters*. Softgauges for Surface Topography [cited; Available from: http://161.112.232.32/softgauges/pdf/Specification.pdf.
24. Mari, L. *On the measurand definition*. in *XVIII IMEKO world congress*. 2006. Rio de Janeiro, Brazil
25. Phillips, S.D., et al., *A Careful Consideration of the Calibration Concept.* J. Res. Natl. Inst. Stand. Technol., 2001. **106**: p. 371-379.
26. Ehrlich, C., R. Dybkaer, and W. Wöger, *Evolution of philosophy and description of measurement (preliminary rationale for VIM3).* Accreditation and Quality Assurance: Journal for Quality, Comparability and Reliability in Chemical Measurement, 2007. **12**(3): p. 201-218.
27. Baratto, A.C., *Measurand: a cornerstone concept in metrology.* Metrologia, 2008. **45**(3): p. 299-307.
28. Pavese, F., *The definition of the measurand in key comparisons: lessons learnt with thermal standards.* Metrologia, 2007. **44**(5): p. 327-339.
29. McCullough, B.D., *Assessing the Reliability of Statistical Software: Part I.* The American Statistician, 1998. **52**(4): p. 358-366.
30. NIST. *STRD- Statistical Reference Datasets*Available from: http://www.nist.gov/itl/div898/strd.
31. McCullough, B.D., *Special section on Microsoft Excel 2007.* Computational Statistics & Data Analysis, 2008. **52**(10): p. 4568-4569.
32. McCullough, B.D. and D.A. Heiser, *On the accuracy of statistical procedures in Microsoft Excel 2007.* Computational Statistics & Data Analysis, 2008. **52**(10): p. 4570-4578.
33. McCullough, B.D. and B. Wilson, *On the accuracy of statistical procedures in Microsoft Excel 97.* Computational Statistics & Data Analysis, 1999. **31**(1): p. 27-37.
34. McCullough, B.D. and B. Wilson, *On the accuracy of statistical procedures in Microsoft Excel 2003.* Computational Statistics & Data Analysis, 2005. **49**(4): p. 1244-1252.
35. Cox, M.G. and P.M. Harris, *Design and use of reference data sets for testing scientific software.* Analytica Chinmica Acta, 1999. **380**(2-3): p. 339-351.
36. Harris, P.M., K.J. Lines, and I.M. Smith, *Testing the numerical correctness of software*, in *NPL Report*. 2006, National Physical Laboratory: London.

37. Lines, K.J., F.O. Onakunle, and I.M. Smith, *Testing functions for calculation the discrete Fourier transform and its inverse*, in *NPL Report*. 2007, National Physical Laboratory: London.
38. Sharman, H.B., *Calibration of surface texture measuring instruments*, in *Proceedings of the Institution of Mechanical Engineers*. 1967, Journal of engineering manufacture. p. 319.
39. Whitehouse, D.J., *Surfaces and Their Measurement*. 2002: Taylor & Francis.
40. Murthy, T., G. Reddy, and V. Radhakrishnan, *Different functions and computations for surface topography.* Wear, 1982. **83**: p. 203-214.
41. Leach, R.K. and P.M. Harris, *Ambiguities in the definition of spacing parameters for surface-texture characterization.* Measurement Science and Technology, 2002(12): p. 1924.
42. Scott, P.J., *The case of surface texture parameter RSm.* Measurement Science and Technology, 2006(3): p. 559.
43. Brennan, J.K., et al. *Reconstruction of continuous surface profiles from discretely sampled data*. in *Proc. 5th Int. Conf. euspen*. 2005. Montpellier, France.

9 Appendix A: Measurement conditions

Form Removed

The form removal operation is set as standard in PTB's F2 standard, is an option in NIST's F2 standard, and is not used in NPL's F2 standard. Therefore, in this test, all measured reference data sets were levelled by the least-squares straight line method on NIST F2 standards before input into all F2 software standards and commercial packages.

Filtering

At the filtration stage, we used only a Gaussian filter with long-wavelength cut-off λc of 0.8 mm calculated by the convolution method.

Sampling Length and Evaluation Length

To minimise the distortion due to the convolution filter, one cut-off at each end of the roughness profile is normally removed. All data sets include 7 cut-offs and the evaluation of the R-parameters is based on the middle five cut-offs. P-parameters are calculated based on all data points in files and, therefore the evaluation length of P-parameters is equal to 5.6 mm in these profiles.

Parameters

The parameters to be compared here are defined by ISO 4287: 1996. The waviness profile is not well defined in current ISO standards due to there is no common understanding of the meaning and use of waviness parameters. Thus, the comparison of W-parameters calculation is not very meaningful and is not discussed in the report.

File format

It should be noted that some test used different data file format due to some software packages do not support SMD data format. In some case, data type is converted and precision is reduced.

The variation of measuring condition

Variation of measuring conditions is listed in Table 17.

	PTB	NIST	NPL	CA	CB	CC
Levelling	LSQ				LSQ	
Profile filtering λs						2.5 µm
The number of data points	22400	22400	22401	22400	22400	22400[a] 19200[b]
The number of removed cut-offs at each end of the profile (End effect of λc profile filtering)	1	1	1	1	1	0.5

Note: [a]. The number of points of *P*-profile, [b]. The number of points of *R*-profile

Table 17: The variation of measuring condition

10 Appendix B: Tables of the Results

			Cos	Cos	Cos	Cos	Cos	Cos	Cos	Cos	Cos	Cos	Cos
	λc	*λs*	*Ra*	*Rq*	*Rsk*	*Rku*	*Rp*	*Rv*	*Rz*	*Rt*	*Rc*	*RSm*	*Rdq*
	/mm	/μm	/nm	/nm			/nm	/nm	/nm	/nm	/nm	/μm	
NIST	0.8	-	636.64	707.11	0	1.5	1000	1000	2000	2000	2000	160	0.02779
NPL	0.8	-	636.6197	707.1068	-3.9E-09	1.5000	1000.00	1000.00	2000.00	2000.01	2000.00	160	*
PTB	0.8	-	636.59	707.08	-0.00018	1.50	1000.00	1000.00	2000.00	2000.00	2000.00	159.99	0.0278
CA	0.8	-	637	708	0	1.5	1001	1001	2001	2001	*	160	0.02778
CB	0.8	-	619.044	687.913	-9.4E-04	1.4998	972.70	972.73	1945.43	1945.45	1914.58	156.01	0.02700
CC	0.8	2.5	636.3	706.7	-0.0001	1.5	999.4	999.5	1998.9	1999.1	1998.9	158.4	0.02775
Deviation by three NMI's software measurement standards													
Mean			636.6166	707.0989	-0.00006	1.5000	1000.00	1000.00	2000.00	2000.00	2000.00	160.00	0.02780
Stdev			0.0205	0.0134	-	0.0000	0.00	0.00	0.00	0.00	0.00	0.00	0.00000
Deviation by all six software packages													
Mean			633.6990	703.9850	-0.00020	1.5000	995.52	995.54	1990.89	1990.93	1982.70	159.07	0.02763
Stdev			6.5570	7.1983	-	0.0001	10.21	10.21	20.34	20.34	34.06	1.49	0.00031
Expected Result[b]													
Ref			636.6198	707.1068	0.00000	1.5000	1000.00	1000.00	2000.00	2000.00	2000.00	160.00	0.02777
Absolute Difference: |*Rx-Rx(expected)*|													
NIST			2.0E-02	3.2E-03	0.0E+00	0.0E+00	0.0E+00	0.0E+00	0.0E+00	0.0E+00	0.0E+00	0.0E+00	2.2E-05
NPL			2.6E-05	2.2E-08	3.9E-09	9.7E-08	1.3E-03	1.5E-05	1.3E-03	6.5E-03	2.9E-05	0.0E+00	*
PTB			3.0E-02	2.7E-02	1.8E-04	0.0E+00	0.0E+00	0.0E+00	0.0E+00	0.0E+00	0.0E+00	1.0E-02	3.2E-05
CA			3.8E-01	8.9E-01	0.0E+00	0.0E+00	1.0E+00	1.0E+00	1.0E+00	1.0E+00	*	0.0E+00	1.2E-05
CB			1.8E+01	1.9E+01	9.4E-04	1.6E-04	2.7E+01	2.7E+01	5.5E+01	5.5E+01	8.5E+01	4.0E+00	7.6E-04
CC			3.2E-01	4.1E-01	1.0E-04	0.0E+00	6.0E-01	5.0E-01	1.1E+00	9.0E-01	1.1E+00	1.6E+00	1.7E-05
Relative Difference: |*Rx-Rx(expected)*|/|*Rx(expected)*|													
NIST			0.0032%	0.0005%	-	0.000%	0.000%	0.000%	0.000%	0.000%	0.000%	0.000%	0.079%
NPL			0.0000%	0.0000%	-	0.000%	0.000%	0.000%	0.000%	0.000%	0.000%	0.000%	*
PTB			0.0047%	0.0038%	-	0.000%	0.000%	0.000%	0.000%	0.000%	0.000%	0.006%	0.115%
CA			0.0597%	0.1263%	-	0.000%	0.100%	0.100%	0.050%	0.050%	*	0.000%	0.045%
CB			2.7608%	2.7144%	-	0.011%	2.730%	2.728%	2.729%	2.728%	4.271%	2.495%	2.751%
CC			0.0502%	0.0575%	-	0.000%	0.060%	0.050%	0.055%	0.045%	0.055%	1.000%	0.062%

[a] Some results are rounded to make them easier to read. All calculations used the original values.
[b] The expected result is obtained from algebraic calculation by Maple 10.03.
* It is not available.

Table 18: Results of *R*-parameter for data file Cos.smd [a]

				Cos	Cos	Cos	Cos	Cos	Cos	Cos	Cos	Cos	Cos	Cos
	λc	*λs*	*lp*	*Pa*	*Pq*	*Psk*	*Pku*	*Pp*	*Pv*	*Pz*	*Pt*	*Pc*	*PSm*	*Pdq*
	/mm	/µm	/mm	/nm	/nm			/nm	/nm	/nm	/nm	/nm	/µm	
NIST	-	-	5.6	636.63	707.12	0.00013	1.49998	1000	1000	2000	2000	2000	160	0.02777
NPL	-	-	5.6	636.6197	707.1068	-1.7E-09	1.5000	1000.00	1000.00	2000.00	2000.00	2000.00	160	*
PTB	-	-	5.6	636.61	707.11	0.000	1.50	999.96	1000.04	2000.00	2000.00	2000.00	159.99	0.02775
CA	-	-	4	637	707	0	1.5	1000	1000	2000	2000	*	*	0.02776
CB	-	-	5.6	636.586	707.123	-6.3E-05	1.5000	999.96	1000.04	2000.00	2000.00	2000.00	158.89	0.02780
CC	-	-	5.6	636.6	707.1	0	1.5	1000	1000	2000	2000	2000	158.86	0.02777
Deviation by three NMI's software measurement standards														
Mean				636.6199	707.1123	0.00004	1.5000	999.99	1000.01	2000.00	2000.00	2000.00	160.00	0.02776
Stdev				0.0082	0.0056	0.00006	0.0000	0.02	0.02	0.00	0.00	0.00	0.00	0.00001
Deviation by all six software packages														
Mean				636.6743	707.0933	0.00001	1.5000	999.99	1000.01	2000.00	2000.00	2000.00	159.55	0.02777
Stdev				0.1463	0.0424	0.00006	0.0000	0.02	0.02	0.00	0.00	0.00	0.55	0.00002
Expected Result[b]														
Ref				636.6198	707.1068	0.00000	1.5000	1000.00	1000.00	2000.00	2000.00	2000.00	160.00	0.02777
Absolute Difference: *|Px-Px(expected)|*														
NIST				1.0E-02	1.3E-02	1.3E-04	2.0E-05	0.0E+00	0.0E+00	0.0E+00	0.0E+00	0.0E+00	0.0E+00	2.0E-06
NPL				3.5E-05	1.0E-05	1.7E-09	9.8E-08	0.0E+00	0.0E+00	0.0E+00	0.0E+00	0.0E+00	0.0E+00	*
PTB				9.8E-03	3.2E-03	0.0E+00	0.0E+00	4.0E-02	4.0E-02	0.0E+00	0.0E+00	0.0E+00	1.0E-02	1.7E-05
CA				3.8E-01	1.1E-01	0.0E+00	0.0E+00	0.0E+00	0.0E+00	0.0E+00	0.0E+00	*	*	6.8E-06
CB				3.4E-02	1.6E-02	6.3E-05	2.0E-05	4.4E-02	4.0E-02	0.0E+00	0.0E+00	0.0E+00	1.1E+00	3.2E-05
CC				2.0E-02	6.8E-03	0.0E+00	0.0E+00	0.0E+00	0.0E+00	0.0E+00	0.0E+00	0.0E+00	1.1E+00	7.0E-07
Relative Difference: *|Px-Px(expected)|/|Px(expected)|*														
NIST				0.0016%	0.0019%	-	0.001%	0.000%	0.000%	0.000%	0.000%	0.000%	0.00%	0.007%
NPL				0.0000%	0.0000%	-	0.000%	0.000%	0.000%	0.000%	0.000%	0.000%	0.00%	*
PTB				0.0015%	0.0005%	-	0.000%	0.004%	0.004%	0.000%	0.000%	0.000%	0.01%	0.062%
CA				0.0597%	0.0151%	-	0.000%	0.000%	0.000%	0.000%	0.000%	*	*	0.025%
CB				0.0053%	0.0023%	-	0.001%	0.004%	0.004%	0.000%	0.000%	0.000%	0.69%	0.115%
CC				0.0031%	0.0010%	-	0.000%	0.000%	0.000%	0.000%	0.000%	0.000%	0.71%	0.003%

[a] Some results are rounded to make them easier to read. All calculations used the original values.
[b] The expected result is obtained from algebraic calculation by Maple 10.03.
* It is not available.

Table 19: Results of *P*-parameter for data file Cos.smd [a]

			EDM	EDM	EDM	EDM	EDM	EDM	EDM	EDM	EDM	EDM	EDM
	λc	*λs*	*Ra*	*Rq*	*Rsk*	*Rku*	*Rp*	*Rv*	*Rz*	*Rt*	*Rc*	*RSm*	*Rdq*
	/mm	/µm	/nm	/nm			/nm	/nm	/nm	/nm	/nm	/µm	
NIST	0.8	-	449.65	539.02	-0.11255	2.29302	1126.11	1217.02	2343.13	2673.15	1147.25	39.92	0.10644
NPL	0.8	-	449.660	539.019	-0.112546	2.29304	1127.32	1218.70	2346.01	2675.94	1215.96	44.75	*
PTB	0.8	-	449.67	540.64	-0.113	2.29	1126.11	1217.03	2343.14	2673.15	1130.4	38.00	0.1064
CA	0.8	-	450	541	-0.119	2.325	1309	1367	2345	2677	*	42.6	0.10516
CB	0.8	-	437.089	520.228	-0.033794	2.33586	1040.69	1133.24	2173.93	2631.05	1023.45	34.81	0.10233
CC	0.8	2.5	447	538	-0.1227	2.3359	1119.5	1212.7	2332.3	2668.7	1211.4	45.20	0.10036
Deviation by three NMI's software measurement standards													
Mean			449.660	539.560	-0.11270	2.29202	1126.51	1217.58	2344.09	2674.08	1164.54	40.89	0.10642
Stdev			0.008	0.764	0.00021	0.00143	0.57	0.79	1.36	1.32	37.01	2.84	0.00002
Deviation by all six software packages													
Mean			447.178	536.318	-0.10226	2.31214	1141.45	1227.61	2313.92	2666.50	1145.69	40.88	0.10414
Stdev			4.624	7.268	0.03086	0.02047	81.01	69.36	62.77	16.07	69.91	3.71	0.00241
Reference Result[b]													
Ref			449.660	539.560	-0.11270	2.29202	1126.51	1217.58	2344.09	2674.08	1164.54	40.89	0.10642
Absolute Difference: *|Rx-Rx(Ref)|*													
NIST			0.010	0.540	0.00015	0.00100	0.40	0.56	0.96	0.93	17.29	0.97	0.00002
NPL			0.000	0.541	0.00015	0.00102	0.80	1.12	1.92	1.86	51.42	3.86	*
PTB			0.010	1.080	0.00030	0.00202	0.40	0.55	0.95	0.93	34.14	2.89	0.00002
CA			0.340	1.440	0.00630	0.03298	182.49	149.42	0.91	2.92	*	1.71	0.00126
CB			12.571	19.332	0.07890	0.04384	85.82	84.34	170.16	43.03	141.09	6.08	0.00409
CC			2.660	1.560	0.01000	0.04388	7.01	4.88	11.79	5.38	46.86	4.31	0.00606
Relative Difference: *|Rx-Rx(Ref)|/|Rx(Ref)|*													
NIST			0.002%	0.100%	0.132%	0.044%	0.04%	0.05%	0.041%	0.035%	1.5%	2.4%	0.019%
NPL			0.000%	0.100%	0.136%	0.045%	0.07%	0.09%	0.082%	0.070%	4.4%	9.4%	*
PTB			0.002%	0.200%	0.267%	0.088%	0.04%	0.05%	0.041%	0.035%	2.9%	7.1%	0.019%
CA			0.076%	0.267%	5.591%	1.439%	16.20%	12.27%	0.039%	0.109%	*	4.2%	1.188%
CB			2.796%	3.583%	70.014%	1.913%	7.62%	6.93%	7.259%	1.609%	12.1%	14.9%	3.842%
CC			0.592%	0.289%	8.875%	1.914%	0.62%	0.40%	0.503%	0.201%	4.0%	10.5%	5.698%

[a]. Some results are rounded to make them easier to read. All calculations used the original values.
[b]. The reference result is the non-weighted mean of results obtained from NIST, NPL and PTB.
*. It is not available.

Table 20: Results of *R*-parameter for data file EDM.smd [a]

				EDM	EDM	EDM	EDM	EDM	EDM	EDM	EDM	EDM	EDM	EDM
	λc	*λs*	*lp*	*Pa*	*Pq*	*Psk*	*Pku*	*Pp*	*Pv*	*Pz*	*Pt*	*Pc*	*PSm*	*Pdq*
	/mm	/μm	/mm	/nm	/nm			/nm	/nm	/nm	/nm	/nm	/μm	
NIST	-	-	5.6	464.34	559.94	-0.13265	2.36461	1394.59	1467.7	2862.3	2862.3	1210.12	46.10	0.10536
NPL	-	-	5.6	464.360	559.953	-0.13264	2.36453	1396.06	1469.77	2865.83	2865.83	1868.86	194.17	*
PTB	-	-	5.6	464.37	559.97	-0.133	2.364	1394.59	1467.7	2862.3	2862.3	1169.41	41.62	0.1054
CB	-	-	5.6	464.335	559.944	-0.13265	2.36461	1394.59	1467.7	2862.3	2862.3	1078.56	36.975	0.1054
CC	-	-	5.6	464.4	560	-0.1327	2.3645	1394.6	1467.7	2862.3	2862.3	1994.9	284.22	0.10616
Deviation by three NMI's software measurement standards (*lp* : 5.6mm)														
Mean			5.6	464.357	559.954	-0.13276	2.36438	1395.08	1468.39	2863.48	2863.48	1416.13	93.96	0.10538
Stdev			5.6	0.012	0.012	0.00017	0.00027	0.69	0.97	1.66	1.66	320.56	70.88	0.00002
Deviation by five software packages (*lp* : 5.6mm)														
Mean			5.6	464.361	559.961	-0.13273	2.36445	1394.89	1468.11	2863.01	2863.01	1464.37	120.62	0.10558
Stdev			5.6	0.023	0.022	0.00014	0.00023	0.59	0.83	1.41	1.41	386.15	100.96	0.00034
Reference Result[b]														
Ref				464.3567	559.954	-0.133	2.36438	1395.08	1468.39	2863.48	2863.48	1416.13	93.96	0.10538
Absolute Difference: *|Px-Px(Ref)|*														
NIST				0.017	0.014	0.00011	0.00023	0.49	0.69	1.18	1.18	206.01	47.87	0.0000
NPL				0.003	0.001	0.00012	0.00015	0.98	1.38	2.35	2.35	452.73	100.21	*
PTB				0.013	0.016	0.00024	0.00038	0.49	0.69	1.18	1.18	246.72	52.34	0.0000
CB				0.022	0.010	0.00012	0.00023	0.49	0.69	1.18	1.18	337.57	56.99	0.0000
CC				0.043	0.046	0.00006	0.00012	0.48	0.69	1.18	1.18	578.77	190.26	0.0008
Relative Difference: *|Px-Rx(Ref)|/|Px(Ref)|*														
NIST				0.004%	0.003%	0.087%	0.010%	0.035%	0.047%	0.041%	0.041%	15%	51%	FALSE
NPL				0.001%	0.000%	0.091%	0.006%	0.070%	0.094%	0.082%	0.082%	32%	107%	*
PTB				0.003%	0.003%	0.177%	0.016%	0.035%	0.047%	0.041%	0.041%	17%	56%	0.019%
CB				0.005%	0.002%	0.087%	0.010%	0.035%	0.047%	0.041%	0.041%	24%	61%	0.019%
CC				0.009%	0.008%	0.049%	0.005%	0.034%	0.047%	0.041%	0.041%	41%	202%	0.741%
Results of P-parameters (*lp* : 4mm)[c]														
	λc	*λs*	*lp*	*Pa*	*Pq*	*Psk*	*Pku*	*Pp*	*Pv*	*Pz*	*Pt*	*Pc*	*PSm*	*Pdq*
	mm	μm	mm	nm	nm			nm	nm	nm	nm	μm	μm	
PTB	-	-	4	462.19	557.06	-0.088	2.34	1316.59	1301.27	2617.87	2617.87	1152.21	40.31	0.107
CA	-	-	4	462	557	-0.177	2.359	1300	1318	2591	2618	*	*	0.10516

[a] Some results are rounded to make them easier to read. All calculations used the original values.
[b] The reference result is the non-weighted mean of results obtained from NIST, NPL and PTB.
[c] The evaluation length is middle-five λc cut-off.
*. It is not available.

Table 21: Results of *P*-parameter for data file EDM.smd [a]

	λc /mm	*λs* /μm	Mill *Ra* /nm	Mill *Rq* /nm	Mill *Rsk*	Mill *Rku*	Mill *Rp* /nm	Mill *Rv* /nm	Mill *Rz* /nm	Mill *Rt* /nm	Mill *Rc* /nm	Mill *RSm* /μm	Mill *Rdq*
NIST	0.8	-	167.64	204.79	0.13861	2.37947	466.72	431.61	898.33	1094.08	617.91	208.00	0.02287
NPL	0.8	-	167.647	204.789	0.1387	2.37946	468.107	433.22	901.328	1097.15	668.759	240.97	*
PTB	0.8	-	167.65	205.43	0.139	2.379	466.72	431.61	898.33	1094.08	617.91	207.99	0.0229
CA	0.8	-	169	208	0.135	2.449	533	570	906	1104	*	199.9	0.02312
CB	0.8	-	156.352	193.292	0.16959	2.57352	449.313	413.352	862.665	1064.89	551.503	0.18	0.020822
CC	0.8	2.5	167.2	205.1	0.1503	2.4525	469.2	411.1	880.3	1079.6	640.3	233.98	0.013788
Deviation by three NMI's software measurement standards													
Mean			167.646	205.003	0.13877	2.37931	467.18	432.15	899.33	1095.10	634.86	218.99	0.02289
Stdev			0.004	0.302	0.00017	0.00022	0.65	0.76	1.41	1.45	23.97	15.55	0.00001
Deviation by all six software packages													
Mean			165.915	203.567	0.14520	2.43549	475.51	448.48	891.16	1088.97	619.28	181.84	0.02070
Stdev			4.313	4.726	0.01190	0.06951	26.58	55.08	15.04	12.99	38.70	82.57	0.00356
Reference Result[b]													
Ref			167.646	205.003	0.13877	2.37931	467.18	432.15	899.33	1095.10	634.86	218.99	0.02289
Absolute Difference: |*Rx-Rx(Ref)*|													
NIST			0.006	0.213	0.00016	0.00016	0.46	0.54	1.00	1.02	16.95	10.99	0.00002
NPL			0.001	0.214	0.00007	0.00015	0.92	1.07	2.00	2.05	33.90	21.99	*
PTB			0.004	0.427	0.00023	0.00031	0.46	0.54	1.00	1.02	16.95	11.00	0.00001
CA			1.354	2.997	0.00377	0.06969	65.82	137.85	6.67	8.90	*	19.09	0.00024
CB			11.294	11.711	0.03082	0.19421	17.87	18.80	36.66	30.21	83.36	218.81	0.00206
CC			0.446	0.097	0.01153	0.07319	2.02	21.05	19.03	15.50	5.44	14.99	0.00910
Relative Difference: |*Rx-Rx(Ref)*|/|*Rx(Ref)*|													
NIST			0.003%	0.104%	0.115%	0.007%	0.10%	0.12%	0.111%	0.093%	2.7%	5.0%	0.066%
NPL			0.001%	0.104%	0.051%	0.006%	0.20%	0.25%	0.222%	0.187%	5.3%	10.0%	*
PTB			0.003%	0.208%	0.166%	0.013%	0.10%	0.12%	0.111%	0.093%	2.7%	5.0%	0.066%
CA			0.808%	1.462%	2.717%	2.929%	14.09%	31.90%	0.742%	0.812%	*	8.7%	1.029%
CB			6.737%	5.713%	22.212%	8.162%	3.82%	4.35%	4.077%	2.759%	13.1%	99.9%	9.014%
CC			0.266%	0.047%	8.309%	3.076%	0.43%	4.87%	2.116%	1.416%	0.86%	6.8%	39.750%

[a]. Some results are rounded to make them easier to read. All calculations used the original values.
[b]. The reference result is the non-weighted mean of results obtained from NIST, NPL and PTB.
*. It is not available.

Table 22: Results of *R*-parameter for data file Mill.smd [a]

				Mill	Mill	Mill	Mill	Mill	Mill	Mill	Mill	Mill	Mill	Mill
	λc	*λs*	*lp*	*Pa*	*Pq*	*Psk*	*Pku*	*Pp*	*Pv*	*Pz*	*Pt*	*Pc*	*PSm*	*Pdq*
	/mm	/μm	/mm	/nm	/nm			/nm	/nm	/nm	/nm	/nm	/μm	
NIST	-	-	5.6	199.46	249.46	-0.11733	2.99552	682.72	729.44	1412.15	1412.15	715.34	287.22	0.023
NPL	-	-	5.6	199.457	249.451	-0.11745	2.99564	684.18	730.88	1415.06	1415.06	684.94	252.40	*
PTB	-	-	5.6	199.46	249.46	-0.117	2.996	682.72	729.44	1412.15	1412.15	685.38	248.2	0.0230
CB	-	-	5.6	199.455	249.458	-0.11733	2.99552	682.717	729.437	1412.15	1412.15	669.372	252.94	0.023
CC	-	-	5.6	199.5	249.5	-0.1174	2.9955	682.7	729.4	1412.2	1412.2	715.3	292.8	0.02522
Deviation by three NMI's software measurement standards (*lp* : 5.6mm)														
Mean			5.6	199.459	249.457	-0.11726	2.99572	683.21	729.92	1413.12	1413.12	695.22	262.60	0.02300
Stdev			5.6	0.001	0.004	0.00019	0.00020	0.69	0.68	1.37	1.37	14.23	17.49	0.00000
Deviation by five software packages (*lp* : 5.6mm)														
Mean			5.6	199.466	249.466	-0.11730	2.99564	683.01	729.72	1412.74	1412.74	694.07	266.71	0.02355
Stdev			5.6	0.017	0.017	0.00016	0.00019	0.59	0.58	1.16	1.16	18.29	19.17	0.00096
Reference Result[b]														
Ref				199.4592	249.457	-0.117	2.99572	683.21	729.92	1413.12	1413.12	695.22	262.60	0.02300
Absolute Difference: *|Px-Px(Ref)|*														
NIST				0.001	0.003	0.00007	0.00020	0.49	0.48	0.97	0.97	20.12	24.61	0.0000
NPL				0.002	0.006	0.00019	0.00008	0.98	0.96	1.94	1.94	10.28	10.21	*
PTB				0.001	0.003	0.00026	0.00028	0.49	0.48	0.97	0.97	9.84	14.40	0.0000
CB				0.004	0.001	0.00007	0.00020	0.49	0.48	0.97	0.97	25.85	9.66	0.0000
CC				0.041	0.043	0.00014	0.00022	0.51	0.52	0.92	0.92	20.08	30.20	0.0022
Relative Difference: *|Px-Px(Ref)|/|Px(Ref)|*														
NIST				0.000%	0.001%	0.061%	0.007%	0.071%	0.066%	0.069%	0.069%	2.9%	9.4%	0.000%
NPL				0.001%	0.002%	0.160%	0.003%	0.143%	0.131%	0.137%	0.137%	1.5%	3.9%	*
PTB				0.000%	0.001%	0.221%	0.009%	0.071%	0.066%	0.069%	0.069%	1.4%	5.5%	0.000%
CB				0.002%	0.000%	0.057%	0.007%	0.072%	0.066%	0.069%	0.069%	3.7%	3.7%	0.000%
CC				0.020%	0.017%	0.121%	0.007%	0.074%	0.071%	0.065%	0.065%	2.9%	11.5%	9.631%
Results of P-parameters (*lp* : 4mm)[c]														
	λc	*λs*	*lp*	*Pa*	*Pq*	*Psk*	*Pku*	*Pp*	*Pv*	*Pz*	*Pt*	*Pc*	*PSm*	*Pdq*
	mm	μm	mm	nm	nm			nm	nm	nm	nm	μm	μm	
PTB	-	-	4	207.35	258.34	-0.375	2.988	682.72	729.44	1412.15	1412.15	678.33	237.49	0.0232
CA	-	-	4	206	258	-0.139	2.945	703	709	1326	1412	*	*	0.02313

[a] Some results are rounded to make them easier to read. All calculations used the original values.
[b] The reference result is the non-weighted mean of results obtained from NIST, NPL and PTB.
[c] The evaluation length is middle-five λc cut-off.
*. It is not available.

Table 23: Results of *P*-parameter for data file Mill.smd [a]

			Polish	Polish	Polish	Polish	Polish	Polish	Polish	Polish	Polish	Polish	Polish
	λc	*λs*	*Ra*	*Rq*	*Rsk*	*Rku*	*Rp*	*Rv*	*Rz*	*Rt*	*Rc*	*RSm*	*Rdq*
	/mm	/µm	/nm	/nm			/nm	/nm	/nm	/nm	/nm	/µm	
NIST	0.8	-	63.25	89.64	-2.24439	10.35517	121.93	555.94	677.88	794.72	232.97	26.58	0.04446
NPL	0.8	-	63.236	89.626	-2.24535	10.35957	122.932	557.42	680.356	797.283	289.3394	42.41	*
PTB	0.8	-	63.25	89.88	-2.244	10.354	121.94	555.94	677.88	794.72	231.11	25.10	0.0445
CA	0.8	-	63.487	90.007	-2.194	9.965	143.944	652.135	678.718	796.079	*	51.2	0.044338
CB	0.8	-	61.8035	87.0938	-2.2534	10.1455	117.408	506.272	623.681	779.153	199.166	19.82	0.041962
CC	0.8	2.5	62.1	88.1	-2.1799	9.8577	117.4	537	654.5	771.9	280.9	46.66	0.034558
Deviation by three NMI's software measurement standards													
Mean			63.245	89.715	-2.24458	10.35625	122.27	556.43	678.71	795.57	251.14	31.36	0.04448
Stdev			0.007	0.117	0.00057	0.00240	0.47	0.70	1.17	1.21	27.02	7.83	0.00002
Deviation by all six software packages													
Mean			62.854	89.058	-2.22684	10.17282	124.26	560.79	665.50	788.98	246.70	35.29	0.04196
Stdev			0.650	1.081	0.02867	0.20174	9.07	44.61	20.70	9.78	33.70	11.92	0.00383
Reference Result[b]													
Ref			63.245	89.715	-2.24458	10.35625	122.27	556.43	678.71	795.57	251.14	31.36	0.04448
Absolute Difference: |*Rx-Rx(Ref)*|													
NIST			0.005	0.075	0.00019	0.00108	0.34	0.49	0.83	0.85	18.17	4.78	0.00002
NPL			0.009	0.089	0.00077	0.00332	0.66	0.99	1.65	1.71	38.20	11.04	*
PTB			0.005	0.165	0.00058	0.00225	0.33	0.49	0.83	0.85	20.03	6.26	0.00002
CA			0.242	0.292	0.05058	0.39125	21.68	95.70	0.01	0.50	*	19.84	0.00014
CB			1.442	2.622	0.00882	0.21075	4.86	50.16	55.02	16.42	51.97	11.55	0.00252
CC			1.145	1.615	0.06468	0.49855	4.87	19.43	24.21	23.67	29.76	15.30	0.00992
Relative Difference: |*Rx-Rx(Ref)*|/|*Rx(Ref)*|													
NIST			0.007%	0.084%	0.009%	0.010%	0.28%	0.09%	0.122%	0.107%	7.2%	15.2%	0.045%
NPL			0.015%	0.099%	0.034%	0.032%	0.54%	0.18%	0.243%	0.215%	15.2%	35.2%	*
PTB			0.007%	0.183%	0.026%	0.022%	0.27%	0.09%	0.122%	0.107%	8.0%	20.0%	0.045%
CA			0.382%	0.325%	2.253%	3.778%	17.73%	17.20%	0.002%	0.063%	*	63.3%	0.318%
CB			2.280%	2.922%	0.393%	2.035%	3.97%	9.02%	8.107%	2.064%	20.7%	36.8%	5.662%
CC			1.811%	1.801%	2.882%	4.814%	3.98%	3.49%	3.566%	2.976%	11.9%	48.8%	22.308%

[a.] Some results are rounded to make them easier to read. All calculations used the original values.
[b]. The reference result is the non-weighted mean of results obtained from NIST, NPL and PTB.
*. It is not available.

Table 24: Results of *R*-parameter for data file Polish.smd [a]

	λc /mm	λs /µm	lp /mm	Polish *Pa* /nm	Polish *Pq* /nm	Polish *Psk*	Polish *Pku*	Polish *Pp* /nm	Polish *Pv* /nm	Polish *Pz* /nm	Polish *Pt* /nm	Polish *Pc* /nm	Polish *PSm* /µm	Polish *Pdq*
NIST	-	-	5.6	63.02	89.53	-2.18916	10.32558	155.06	682.31	837.37	837.37	259.92	34.40	0.04449
NPL	-	-	5.6	63.073	89.517	-2.19008	10.32942	155.06	682.42	837.49	837.49	474.97	286.21	*
PTB	-	-	5.6	63.02	89.53	-2.189	10.326	155.06	682.31	837.37	837.37	250.74	30.75	0.0445
CB	-	-	5.6	63.0137	89.5284	-2.18916	10.3256	155.063	682.306	837.37	837.37	210.77	22.67	0.0445
CC	-	-	5.6	63	89.5	-2.1891	10.3255	155.1	682.3	837.4	837.4	590.30	963.40	0.04668
Deviation by three NMI's software measurement standards (*lp* : 5.6mm)														
Mean			5.6	63.038	89.526	-2.18941	10.32700	155.06	682.35	837.41	837.41	328.54	117.12	0.04450
Stdev			5.6	0.025	0.006	0.00048	0.00172	0.00	0.05	0.06	0.06	103.61	119.58	0.00000
Deviation by five software packages (*lp* : 5.6mm)														
Mean			5.6	63.025	89.521	-2.18930	10.32642	155.07	682.33	837.40	837.40	357.34	267.49	0.04504
Stdev			5.6	0.025	0.012	0.00039	0.00151	0.02	0.05	0.05	0.05	148.62	361.93	0.00094
Reference Result[b]														
Ref				63.0376	89.526	-2.189	10.32700	155.06	682.35	837.41	837.41	328.54	117.12	0.04450
Absolute Difference: $\|Px-Px(Ref)\|$														
NIST				0.018	0.004	0.00025	0.00142	0.00	0.04	0.04	0.04	68.62	82.72	0.0000
NPL				0.035	0.009	0.00067	0.00242	0.00	0.08	0.08	0.08	146.43	169.09	*
PTB				0.018	0.004	0.00041	0.00100	0.00	0.04	0.04	0.04	77.80	86.37	0.0000
CB				0.024	0.003	0.00025	0.00140	0.00	0.04	0.04	0.04	117.77	94.45	0.0000
CC				0.038	0.026	0.00031	0.00150	0.04	0.05	0.01	0.01	261.76	846.28	0.0022
Relative Difference: $\|Px-Px(Ref)\|/\|Px(Ref)\|$														
NIST				0.028%	0.005%	0.012%	0.014%	0.001%	0.006%	0.005%	0.0%	21%	71%	0.011%
NPL				0.056%	0.010%	0.030%	0.023%	0.001%	0.011%	0.009%	0.009%	45%	144%	*
PTB				0.028%	0.005%	0.019%	0.010%	0.001%	0.006%	0.005%	0.005%	24%	74%	0.011%
CB				0.038%	0.003%	0.012%	0.014%	0.001%	0.006%	0.005%	0.005%	36%	81%	0.011%
CC				0.060%	0.029%	0.014%	0.015%	0.025%	0.007%	0.001%	0.001%	80%	723%	4.906%
Results of P-parameters (*lp* : 4mm)[d]														
	λc mm	λs µm	lp mm	*Pa* nm	*Pq* nm	*Psk*	*Pku*	*Pp* nm	*Pv* nm	*Pz* nm	*Pt* nm	*Pc* µm	*PSm* µm	*Pdq*
PTB[e]	-	-	4	65.07	92.51	-2.154	9.703	155.06	682.31	837.37	837.37	253.23	30.46	0.0445
CA	-	-	4	65.693	92.472	-2.084	9.512	157.283	680.087	786.431	837.37	*	*	0.04434

[a] Some results are rounded to make them easier to read. All calculations used the original values.
[b] Height discriminations: ±5 % *Pq*.
[c] The reference result is the non-weighted mean of results obtained from NIST, NPL and PTB.

Table 25: Results of *P*-parameter for data file Polish.smd [a]

Results obtained from Ground.smd													
	λc /mm	*λs* /µm	*Ra* /nm	*Rq* /nm	*Rsk*	*Rku*	*Rp* /nm	*Rv* /nm	*Rz* /nm	*Rt* /nm	*Rc* /nm	*RSm* /µm	*Rdq*
NIST	0.8	-	443.03	555.27	-0.17149	3.03608	1522.79	1706.8	3229.59	4095.96	1335.35	34.12	0.17303
NPL	0.8	-	443.005	555.259	-0.1714456	3.03614	1524.342	1707.74	3232.09	4097.82	1527.908	46.77	*
PTB	0.8	-	443.01	558.18	-0.171	3.036	1522.79	1706.81	3229.60	4095.97	1340.55	32.75	0.173
CA	0.8	-	444	559	-0.168	3.163	2125	1982	3238	4107	*	45.7	0.167042
CB	0.8	-	434.327	544.887	-0.102592	3.17181	1549.29	1588.37	3137.65	4081.71	1157.83	26.06	0.171176
CC	0.8	2.5	438.1	552.3	-0.1701	3.1675	1489.8	1672.4	3162.2	4030.4	1515	50.17	0.153589
Deviation by three NMI's software measurement standards													
Mean			443.015	556.236	-0.17131	3.03607	1523.31	1707.12	3230.43	4096.58	1401.27	37.88	0.17302
Stdev			0.011	1.374	0.00022	0.00006	0.73	0.44	1.17	0.87	89.57	6.31	0.00002
Deviation by all six software packages													
Mean			440.912	554.149	-0.15910	3.10175	1622.34	1727.35	3204.85	4084.81	1375.33	39.26	0.16757
Stdev			3.514	4.681	0.02530	0.06573	225.46	121.39	39.58	25.43	136.31	8.76	0.00732
Reference Result[b]													
Ref			443.015	556.236	-0.17131	3.03607	1523.31	1707.12	3230.43	4096.58	1401.27	37.88	0.17302
Absolute Difference: $\lvert Rx-Rx(Ref)\rvert$													
NIST			0.015	0.966	0.00018	0.00001	0.52	0.32	0.84	0.62	65.92	3.76	0.00002
NPL			0.010	0.977	0.00013	0.00007	1.03	0.63	1.66	1.24	126.64	8.89	*
PTB			0.005	1.944	0.00031	0.00007	0.52	0.31	0.83	0.61	60.72	5.13	0.00001
CA			0.985	2.764	0.00331	0.12693	601.69	274.88	7.57	10.42	*	7.82	0.00597
CB			8.688	11.349	0.06872	0.13574	25.98	118.75	92.78	14.87	243.44	11.82	0.00184
CC			4.915	3.936	0.00121	0.13143	33.51	34.72	68.23	66.18	113.73	12.29	0.01943
Relative Difference: $\lvert Rx-Rx(Ref)\rvert/\lvert Rx(Ref)\rvert$													
NIST			0.003%	0.174%	0.104%	0.000%	0.03%	0.02%	0.026%	0.015%	4.7%	9.9%	0.009%
NPL			0.002%	0.176%	0.078%	0.002%	0.07%	0.04%	0.051%	0.030%	9.0%	23.5%	*
PTB			0.001%	0.349%	0.182%	0.002%	0.03%	0.02%	0.026%	0.015%	4.3%	13.5%	0.009%
CA			0.222%	0.497%	1.933%	4.181%	39.50%	16.10%	0.234%	0.254%	*	20.6%	3.452%
CB			1.961%	2.040%	40.114%	4.471%	1.71%	6.96%	2.872%	0.363%	17.4%	31.2%	1.063%
CC			1.109%	0.708%	0.707%	4.329%	2.20%	2.03%	2.112%	1.616%	8.1%	32.4%	11.228%
Results obtained from Ground2.smd													
	λc /mm	*λs* /µm	*Ra* /nm	*Rq* /nm	*Rsk*	*Rku*	*Rp* /nm	*Rv* /nm	*Rz* /nm	*Rt* /nm	*Rc* /nm	*RSm* /µm	*Rdq*
NIST	0.8	-	443.03	555.25	-0.1714	3.03619	1522.79	1706.8	3229.59	4095.96	1339.78	34.56	0.17303
NPL	0.8	-	443.005	555.259	-0.1714456	3.03614	1524.342	1707.74	3232.09	4097.82	1527.908	46.77	*
PTB	0.8	-	442.98	558.15	-0.171	3.037	1522.79	1706.81	3229.60	4095.97	1312.20	32.86	0.173
CA	0.8	-	444	559	-0.168	3.163	2125	1982	3238	4107	*	46.2	0.16704
CB	0.8	-	443.258	557.775	-0.143042	3.24715	1717.97	1641.79	3359.76	4081.71	1173.95	27.17	0.17028
CC	0.8	2.5	438.1	552.4	-0.1701	3.1672	1489.9	1672.5	3162.3	4030.4	1523.6	50.49	0.153589
Absolute Difference of results obtained from ground.smd and ground2.smd													
NIST			0.000	0.020	0.00009	0.00011	0.00	0.00	0.00	0.00	4.43	0.44	0.00000
NPL			0.000	0.000	0.00000	0.00000	0.00	0.00	0.00	0.00	0.00	0.00	*
PTB			0.030	0.030	0.00000	0.00100	0.00	0.00	0.00	0.00	28.35	0.11	0.00000
CA			0.000	0.000	0.00000	0.00000	0.00	0.00	0.00	0.00	*	0.50	0.00000
CB			8.931	12.888	0.04045	0.07534	168.68	53.42	222.11	0.00	16.12	1.11	0.00090
CC			0.000	0.100	0.00000	0.00030	0.10	0.10	0.10	0.00	8.60	0.32	0.00000

[a.] Some results are rounded to make them easier to read. All calculations used the original values.

[b]. The reference result is the non-weighted mean of results obtained from NIST, NPL and PTB.

*. It is not available.

Table 26: Results of *R*-parameter for data file Ground.smd and Ground2.smd[a]

Results obtained from Ground.smd														
	λc /mm	*λs* /µm	*lp* /mm	*Pa* /nm	*Pq* /nm	*Psk*	*Pku*	*Pp* /nm	*Pv* /nm	*Pz* /nm	*Pt* /nm	*Pc* /nm	*PSm* /µm	*Pdq*
NIST	-	-	5.6	464.3	586.96	-0.10044	3.14597	2253.07	2036.71	4289.78	4289.78	1458.09	41.70	0.17434
NPL	-	-	5.6	464.309	586.967	-0.10047	3.14595	2253.39	2038.59	4291.98	4291.98	2461.00	199.58	*
PTB[b]	-	-	5.6	464.33	586.99	-0.100	3.146	2253.07	2036.71	4289.78	4289.78	1414.76	37.30	0.1743
CB	-	-	5.6	464.293	586.964	-0.10044	3.14597	2253.07	2036.71	4289.78	4289.78	1174.64	28.1843	0.1743
CC	-	-	5.6	464.3	587	-0.1004	3.1458	2253.1	2036.7	4289.8	4289.8	2897.1	350	0.17449
Deviation by three NMI's software measurement standards (*lp* : 5.6mm)														
Mean			5.6	464.313	586.972	-0.10030	3.14597	2253.18	2037.34	4290.51	4290.51	1777.95	92.86	0.17432
Stdev			5.6	0.013	0.013	0.00021	0.00002	0.15	0.89	1.04	1.04	483.31	75.48	0.00002
Deviation by five software packages (*lp* : 5.6mm)														
Mean			5.6	464.306	586.976	-0.10035	3.14594	2253.14	2037.08	4290.22	4290.22	1881.12	131.35	0.17436
Stdev			5.6	0.013	0.016	0.00018	0.00007	0.13	0.75	0.88	0.88	672.91	126.48	0.00008
Reference Result[c]														
Ref				464.313	586.972	-0.10030	3.14597	2253.18	2037.34	4290.51	4290.51	1777.95	92.86	0.17432
Absolute Difference: |*Px-Px(Ref)*|														
NIST				0.013	0.012	0.00014	0.00000	0.11	0.63	0.73	0.73	319.86	51.16	0.00002
NPL				0.004	0.005	0.00016	0.00002	0.21	1.25	1.47	1.47	683.05	106.72	*
PTB				0.017	0.018	0.00030	0.00003	0.11	0.63	0.73	0.73	363.19	55.56	0.00002
CB				0.020	0.008	0.00014	0.00000	0.11	0.63	0.73	0.73	603.31	64.68	0.00002
CC				0.013	0.028	0.00010	0.00017	0.08	0.64	0.71	0.71	1119.15	257.14	0.00017
Relative Difference: |*Px-Px(Ref)*|/|*Px(Ref)*|														
NIST				0.003%	0.002%	0.138%	0.000%	0.005%	0.031%	0.017%	0.017%	18%	55%	0.011%
NPL				0.001%	0.001%	0.163%	0.001%	0.009%	0.061%	0.034%	0.034%	38%	115%	*
PTB				0.004%	0.003%	0.301%	0.001%	0.005%	0.031%	0.017%	0.017%	20%	60%	0.011%
CB				0.004%	0.001%	0.139%	0.000%	0.005%	0.031%	0.017%	0.017%	34%	70%	0.011%
CC				0.003%	0.005%	0.098%	0.006%	0.003%	0.031%	0.017%	0.017%	63%	277%	0.096%
Results of P-parameters (*lp* : 4mm)[d]														
	λc /mm	*λs* /µm	*lp* /mm	*Pa* /nm	*Pq* /nm	*Psk*	*Pku*	*Pp* /nm	*Pv* /nm	*Pz* /nm	*Pt* /nm	*Pc* /nm	*PSm* /µm	*Pdq*
PTB[e]	-	-	4	455.72	573.96	-0.279	3.185	2138.09	2036.71	4174.80	4174.80	1370.84	34.89	0.1732
CA	-	-	4	456	574	-0.196	3.163	2154	2021	3675	4175	*	*	0.16704
Results obtained from Ground2.smd														
	λc /mm	*λs* /µm	*lp* /mm	*Pa* /nm	*Pq* /nm	*Psk*	*Pku*	*Pp* /nm	*Pv* /nm	*Pz* /nm	*Pt* /nm	*Pc* /nm	*PSm* /µm	*Pdq*
NIST	-	-	5.6	464.3	586.96	-0.10044	3.14597	2253.07	2036.71	4289.78	4289.78	1447.3	40.61	0.17434
NPL	-	-	5.6	464.309	586.967	-0.10047	3.14595	2253.39	2038.59	4291.98	4291.98	2461.00	199.58	*
PTB[b]	-	-	5.6	464.29	586.97	-0.101	3.146	2253.07	2036.71	4289.78	4289.78	1398.89	37.76	0.1743
CB	-	-	5.6	464.27	586.964	-0.10044	3.14597	2253.07	2036.71	4289.78	4289.78	1167.81	28.1256	0.1743
CC	-	-	5.6	464.3	587	-0.1004	3.146	2253.1	2036.7	4289.8	4289.8	2845.1	350	0.17447
PTB	-	-	4	455.68	573.92	-0.279	3.185	2138.09	2036.71	4174.80	4174.80	1312.2	32.26	0.1732
CA	-	-	4	456	574	-0.196	3.163	2154	2021	3675	4175	*	*	0.16704
Absolute Difference of results obtained from ground.smd and ground2.smd														
NIST			5.6	0.000	0.000	0.00000	0.00000	0.00	0.00	0.00	0.00	*	1.0933	0.00000
NPL			5.6	0.000	0.000	0.00000	0.00000	0.00	0.00	0.00	0.00	0.00	0.0000	*
PTB			5.6	0.040	0.020	0.00100	0.00000	0.00	0.00	0.00	0.00	15.87	0.4600	0.00000
CB			5.6	0.023	0.000	0.00000	0.00000	0.00	0.00	0.00	0.00	6.83	0.0587	0.00000
CC			5.6	0.000	0.000	0.00000	0.00020	0.00	0.00	0.00	0.00	52.00	0.0000	0.00002
PTB	-	-	4	0.040	0.040	0.00000	0.00000	0.00	0.00	0.00	0.00	58.64	2.6300	0.00000
CA	-	-	4	0.000	0.000	0.00000	0.00000	0.00	0.00	0.00	0.00	*	*	0.00000

[a]. Some results are rounded to make them easier to read. All calculations used the original values.
[b]. The reference result is the non-weighted mean of results obtained from NIST, NPL and PTB.
*. It is not available.

Table 27: Results of *P*-parameter for data file Ground.smd and Ground2.smd [a]

	λc /mm	λs /µm	Discrimination Height	Spacing	*Rc* Cos /nm	*Rc* EDM /nm	*Rc* Mill /nm	*Rc* Polish /nm	*Rc* Ground /nm
NIST	0.8	-	±5 % *Rz*	1% *lr*	2000	1147	618	233	1336
NPL	0.8	-	±5 % *Rz*	1% *lr*	2000	1216	669	289	1528
PTB	0.8	-	±5 % *Rz*	1% *lr*	2000	1130	618	231	1341
CB	0.8	-	±5 % *Rz*	1% *lr*	1915	1023	552	199	1158
CC	0.8	2.5	±5 % *Rz*	1% *lr*	1999	1211	640	281	1515

Table 28: Results of *Rc* parameter

	lp /mm	Discrimination Height	Spacing	*Pc* Cos /nm	*Pc* EDM /nm	*Pc* Mill /nm	*Pc* Polish /nm	*Pc* Ground /nm
NIST	5.6	±5 % *Pz*	1% *lp*	2000	1210	715	260	1458
NPL	5.6	±5 % *Pz*	1% *lp*	2000	1869	685	475	2461
PTB1	5.6	±5 % *Pz*	1% *lp*	2000	1169	685	251	1415
PTB2	4.0[a]	±5 % *Pz*	1% *lp*	2000	1152	678	253	1371
CB	5.6	±5 % *Pz*	1% *lp*	2000	1079	669	211	1175
CC	5.6	±5 % *Pz*	1% *lp*	2000	1995	715	590	2897

[a] The evaluation length is middle-five λc cut-off.

Table 29: Results of *Pc* parameter

	λc /mm	λs /µm	Discrimination Height	Spacing	*RSm* Cos /µm	*RSm* EDM /µm	*RSm* Mill /µm	*RSm* Polish /µm	*RSm* Ground /µm
NIST	0.8	-	±5 % *Rz*	1% *lr*	160.0	39.9	208.0	26.6	34.1
NPL	0.8	-	±5 % *Rz*	1% *lr*	160.0	44.7	241.0	42.4	46.8
PTB	0.8	-	±5 % *Rz*	1% *lr*	160.0	38.0	208.0	25.1	32.8
CA	0.8	-	±5 % *Rz*	1% *lr*	160.0	42.6	199.9	51.2	45.7
CB	0.8	-	±5 % *Rz*	1% *lr*	156.0	34.8	179.7	19.8	26.1
CC	0.8	2.5	±5 % *Rz*	1% *lr*	158.4	45.2	234.0	46.7	50.2

Table 30: Results of *RSm* parameter

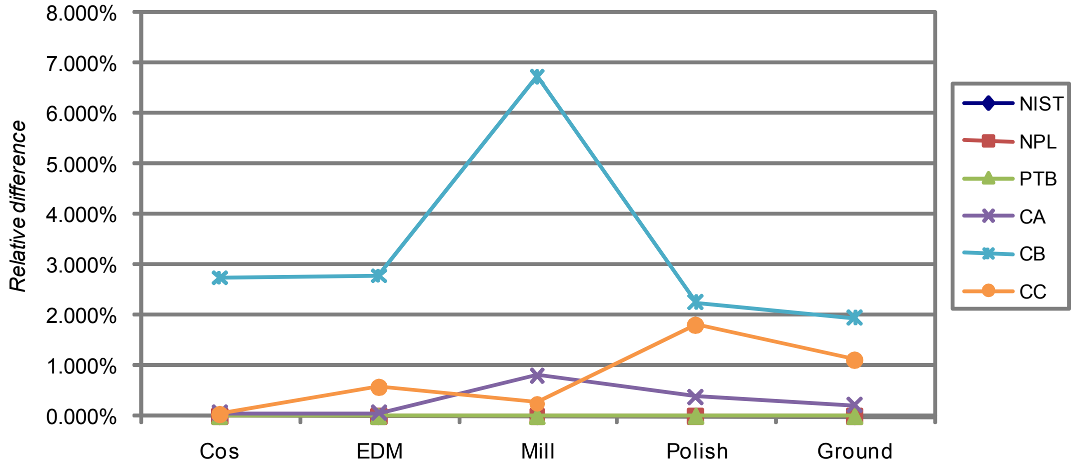
	lp /mm	Discrimination Height	Spacing	*PSm* Cos /µm	*PSm* EDM /µm	*PSm* Mill /µm	*PSm* Polish /µm	*PSm* Ground /µm
NIST	5.6	±5 % *Pz*	1% *lp*	160.0	46.1	287.2	34.4	41.7
NPL	5.6	±5 % *Pz*	1% *lp*	160.0	194.2	252.4	286.2	199.6
PTB1	5.6	±5 % *Pz*	1% *lp*	160.0	41.6	248.2	30.8	37.3
PTB2	4.0[a]	±5 % *Pz*	1% *lp*	160.0	40.3	237.5	30.5	34.9
CB	5.6	±5 % *Pz*	1% *lp*	158.9	37.0	252.9	22.7	28.2
CC	5.6	±5 % *Pz*	1% *lp*	158.9	284.2	292.8	963.4	350.0

[a] The evaluation length is middle-five λc cut-off.

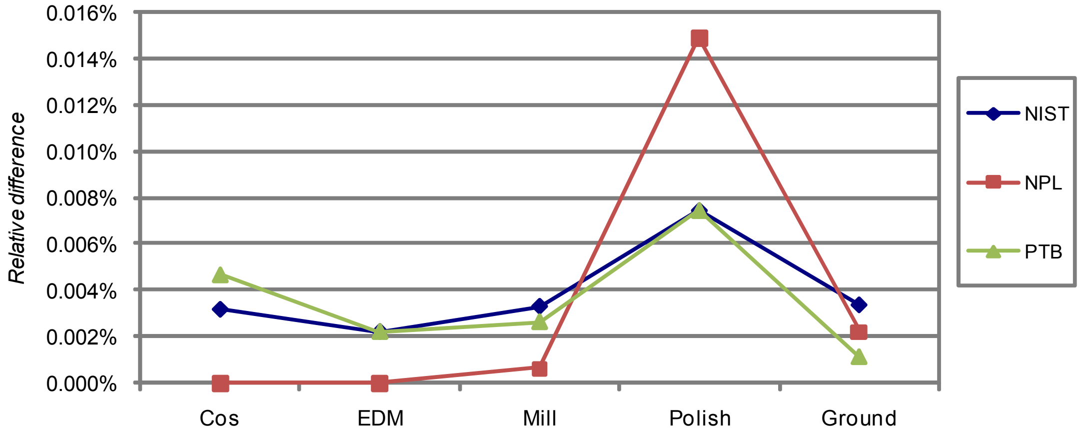
Table 31: Results of *PSm* parameter

11 Appendix C: Result Analysis

C.1 Parameter Analysis



(a)



(b)

Figure 17: *Ra* parameter analysis

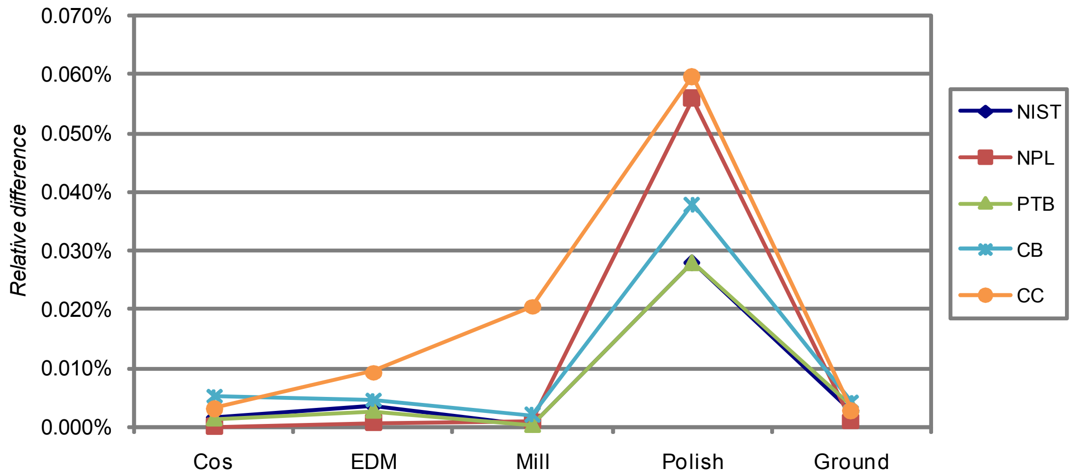
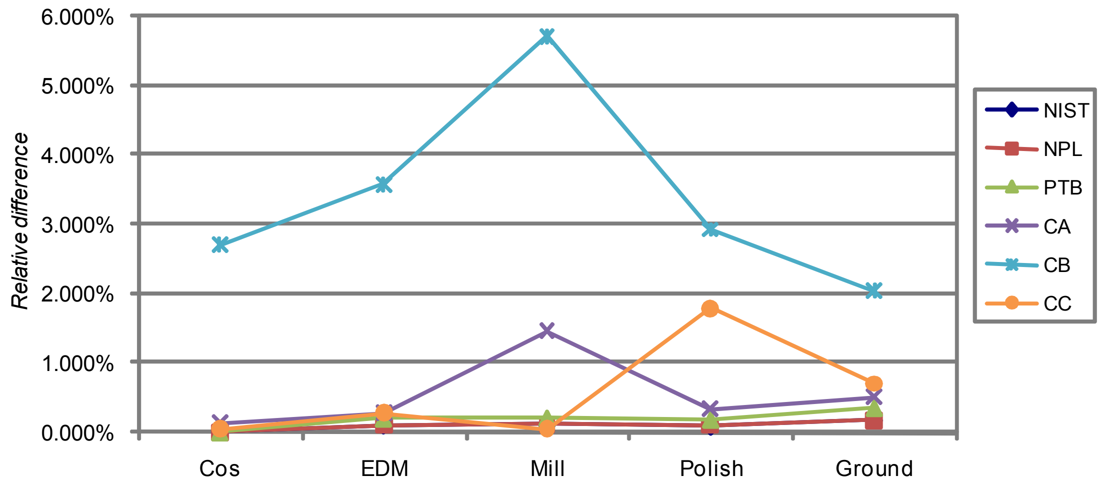
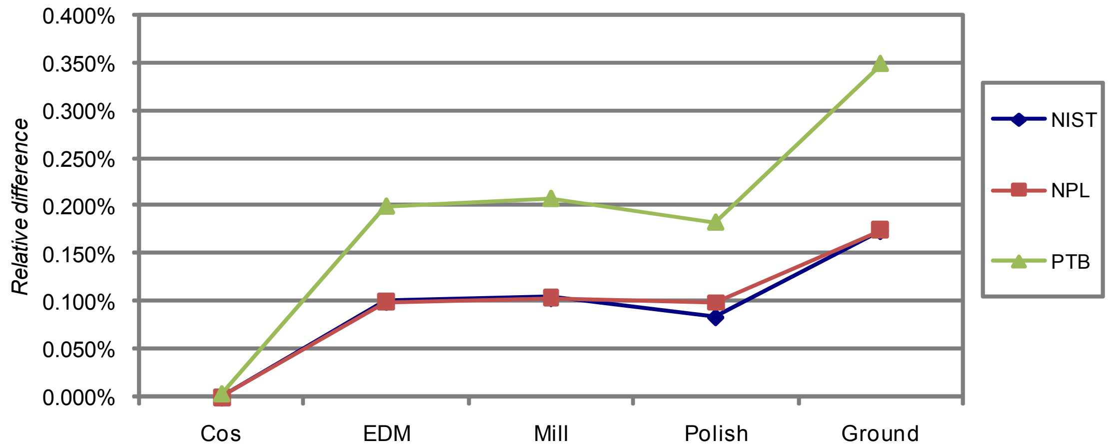


Figure 18: *Pa* parameter analysis



(a)



(b)

Figure 19: *Rq* parameter analysis

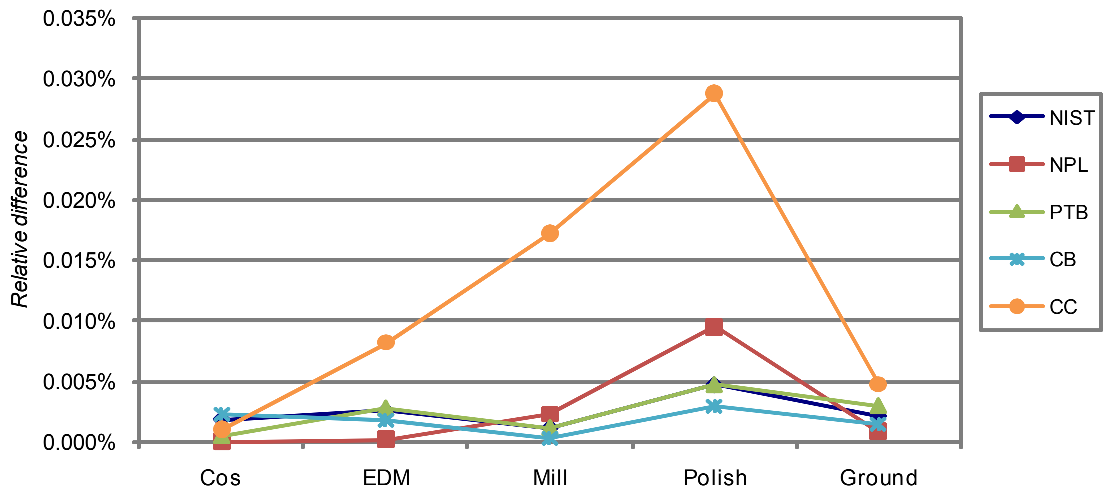
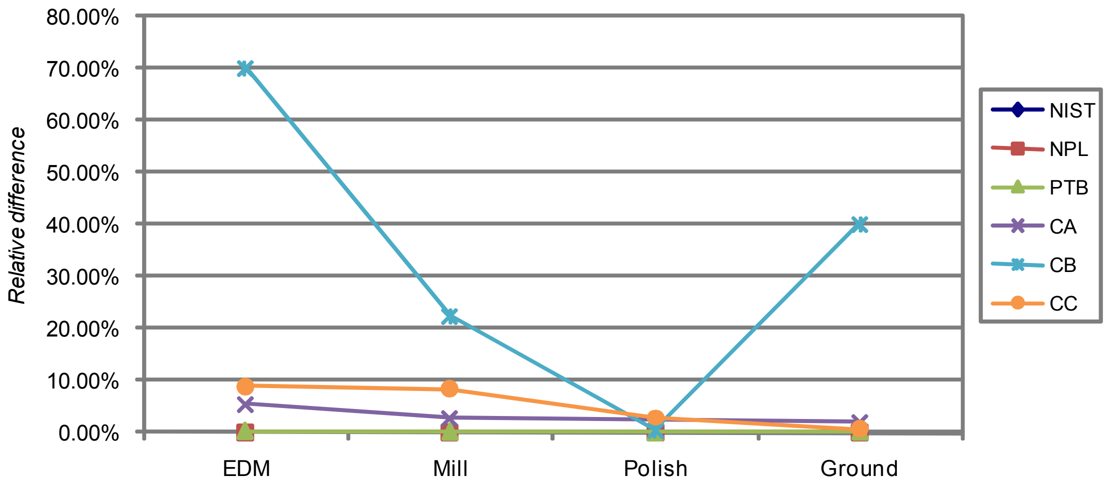
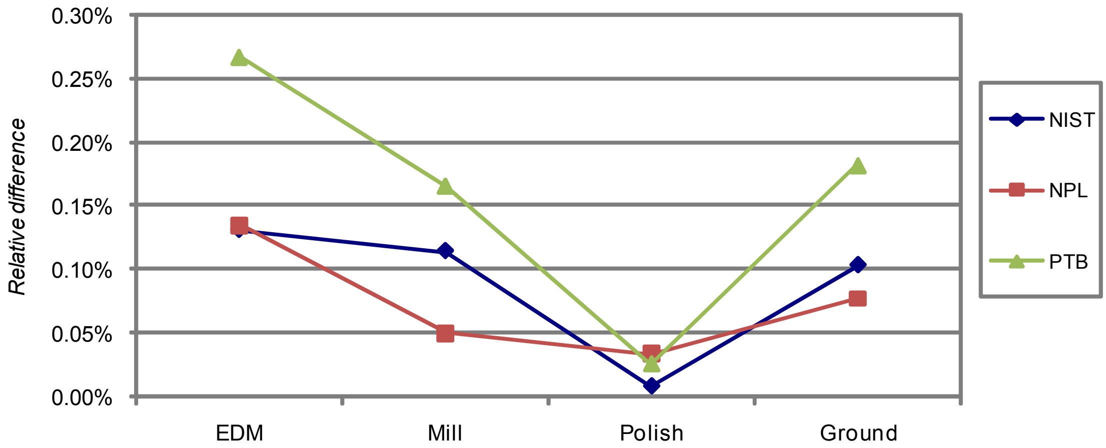


Figure 20: *Pq* parameter analysis



(a)



(b)

Figure 21: *Rsk* parameter analysis

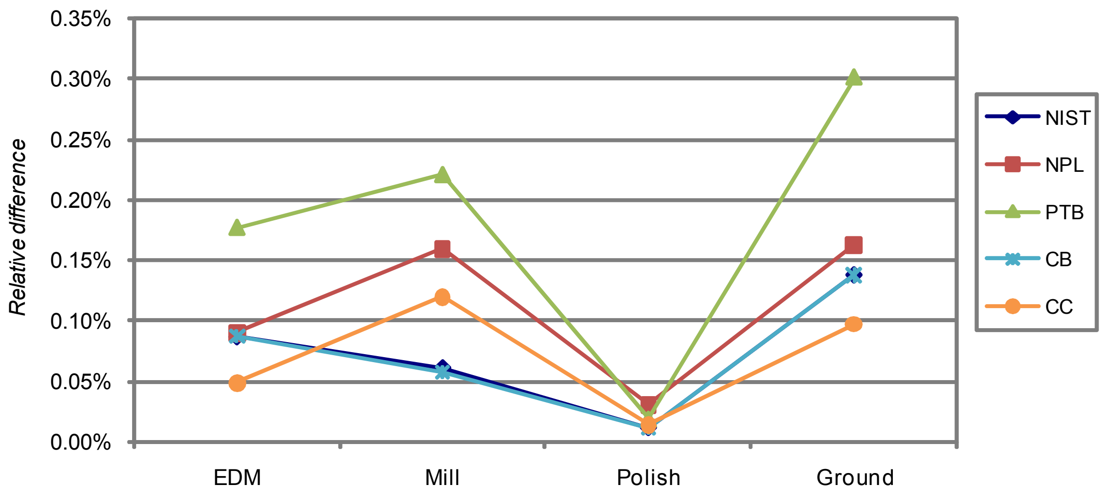
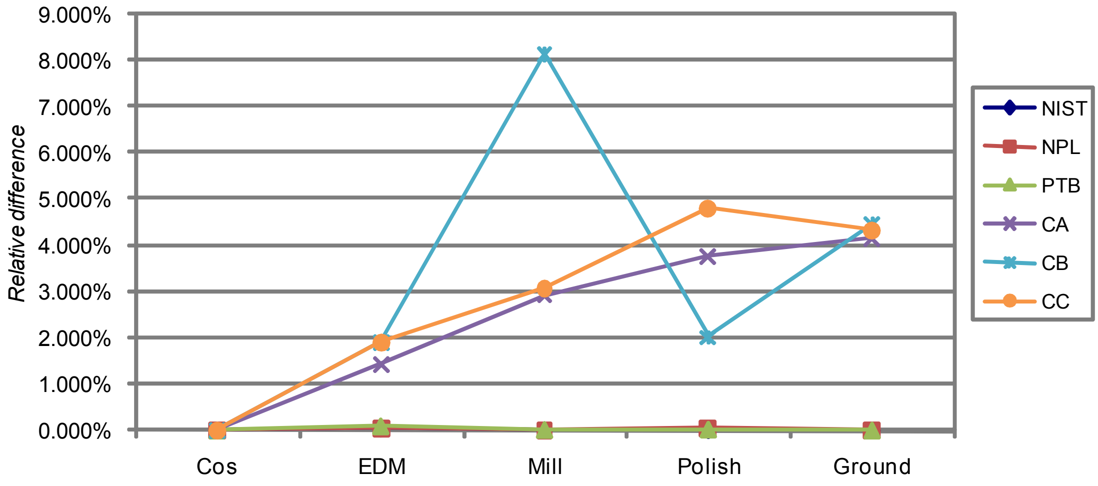
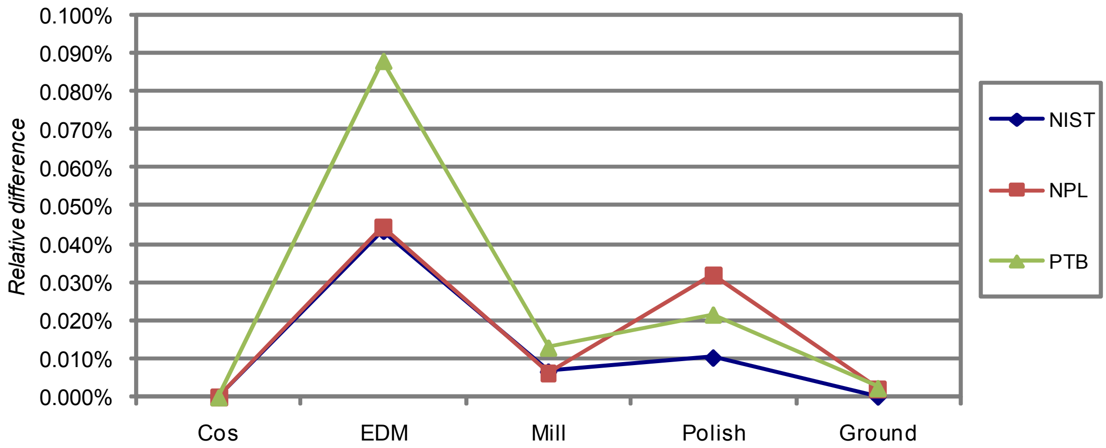


Figure 22: *Psk* parameter analysis



(a)



(b)

Figure 23: *Rku* parameter analysis

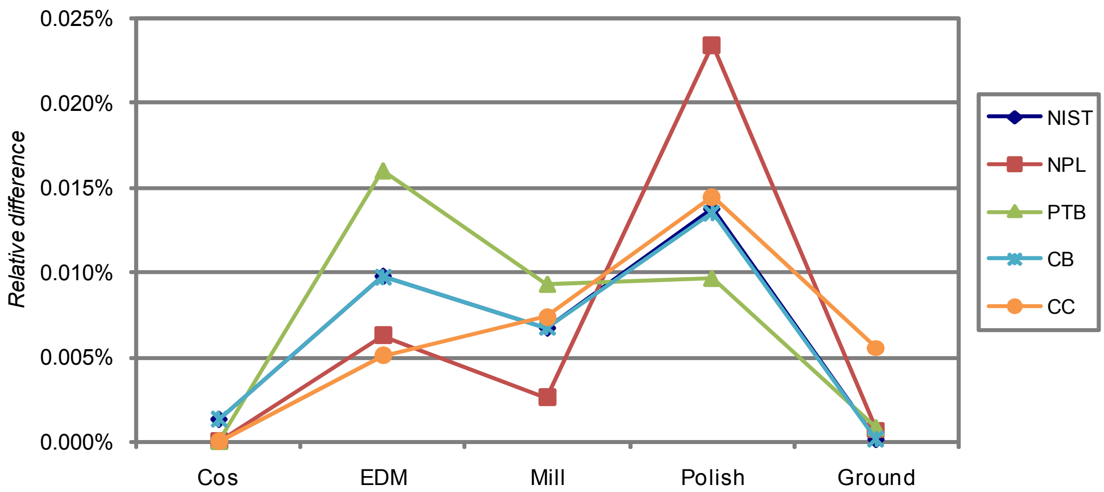
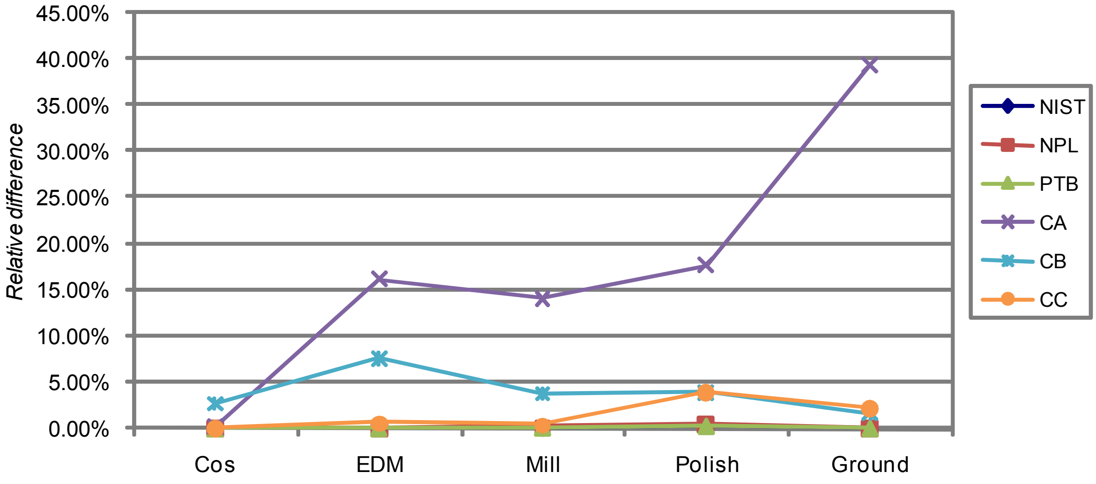
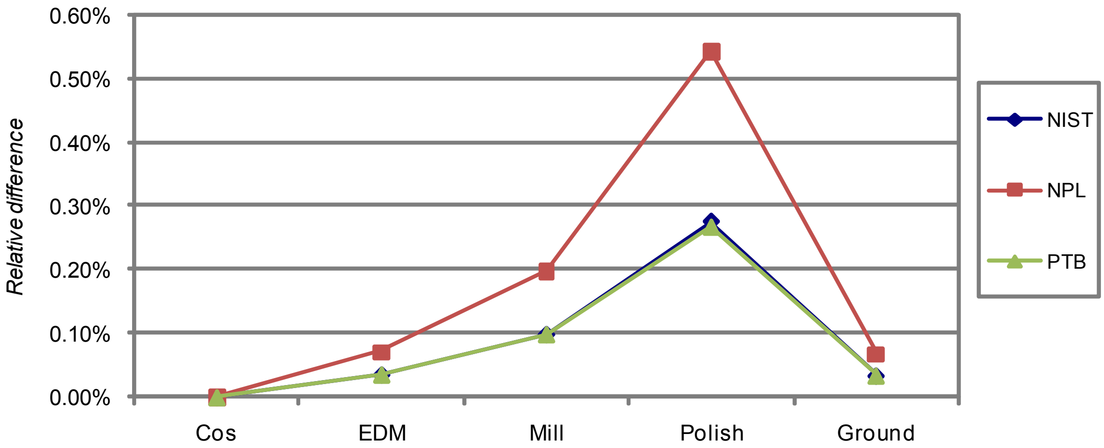


Figure 24: *Pku* parameter analysis



(a)



(b)

Figure 25: *Rp* parameter analysis

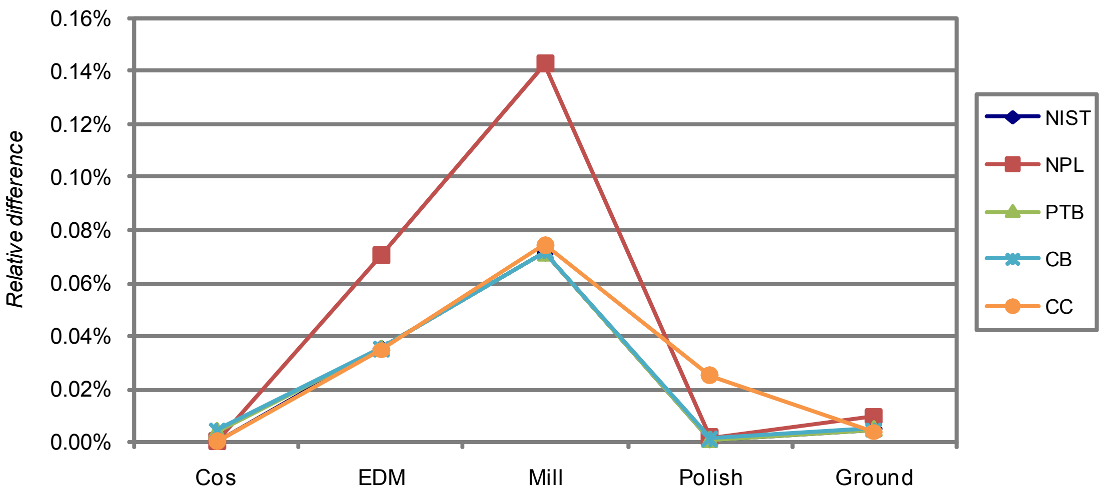
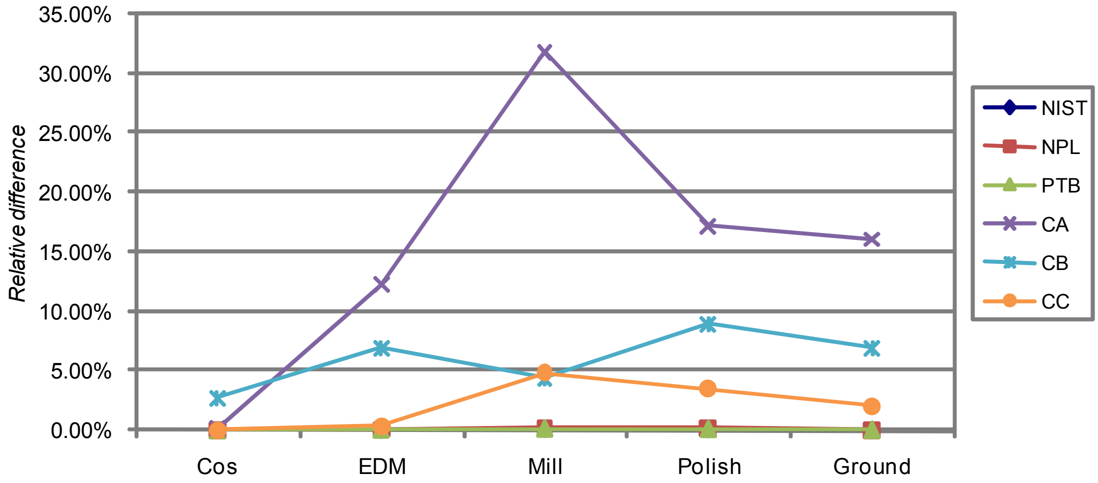
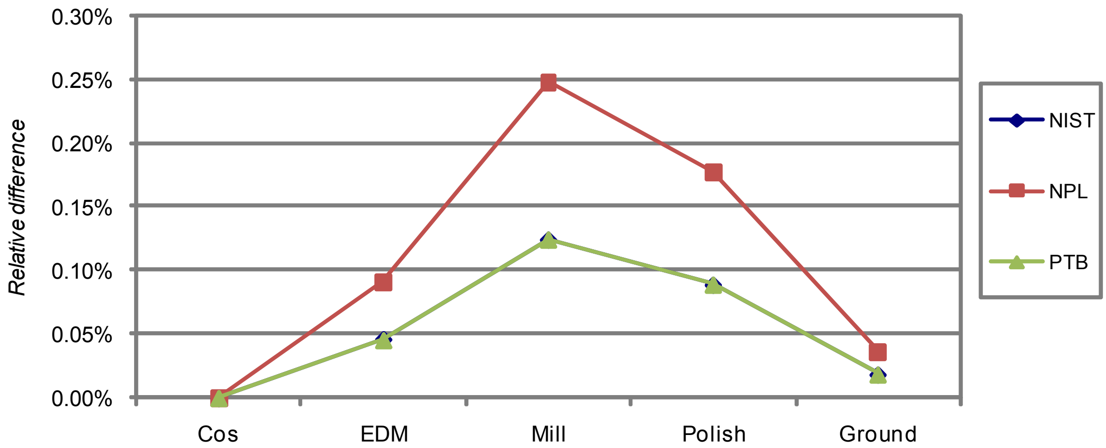


Figure 26: *Pp* parameter analysis



(a)



(b)

Figure 27: *Rv* parameter analysis

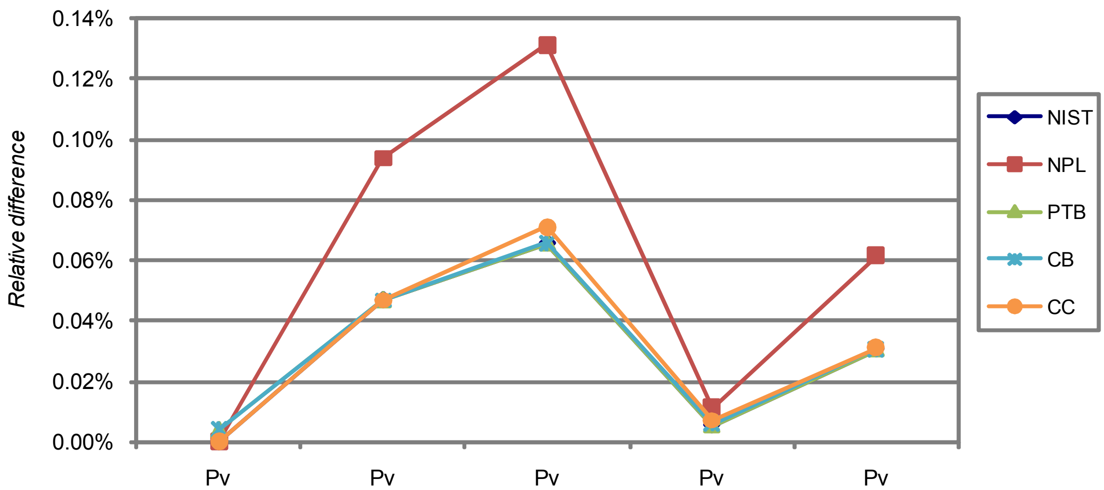
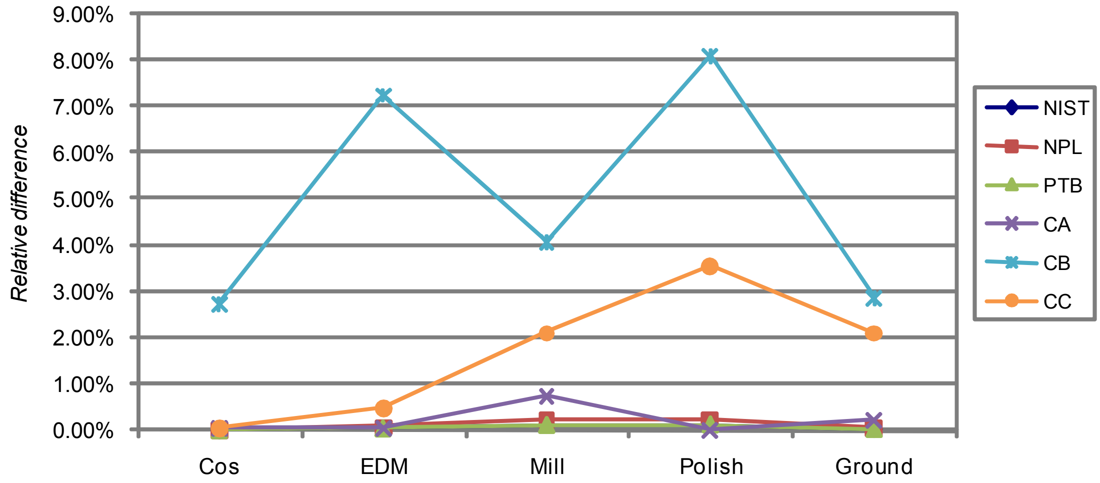
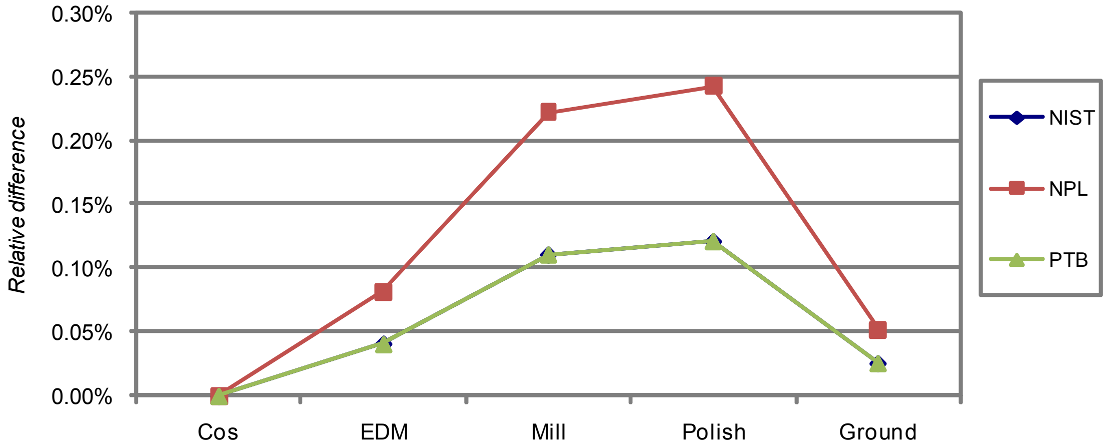


Figure 28: *Pv* parameter analysis



(a)



(b)

Figure 29: *Rz* parameter analysis

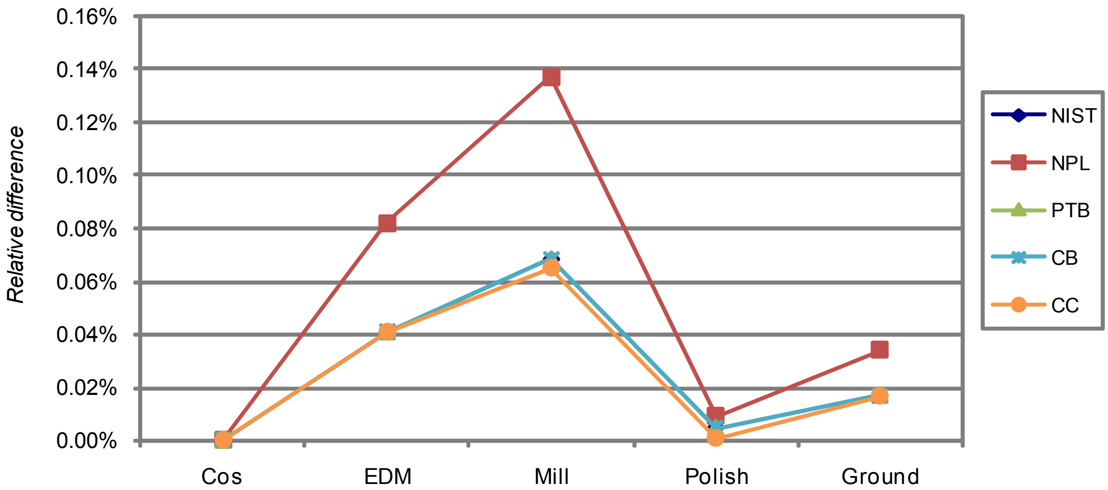
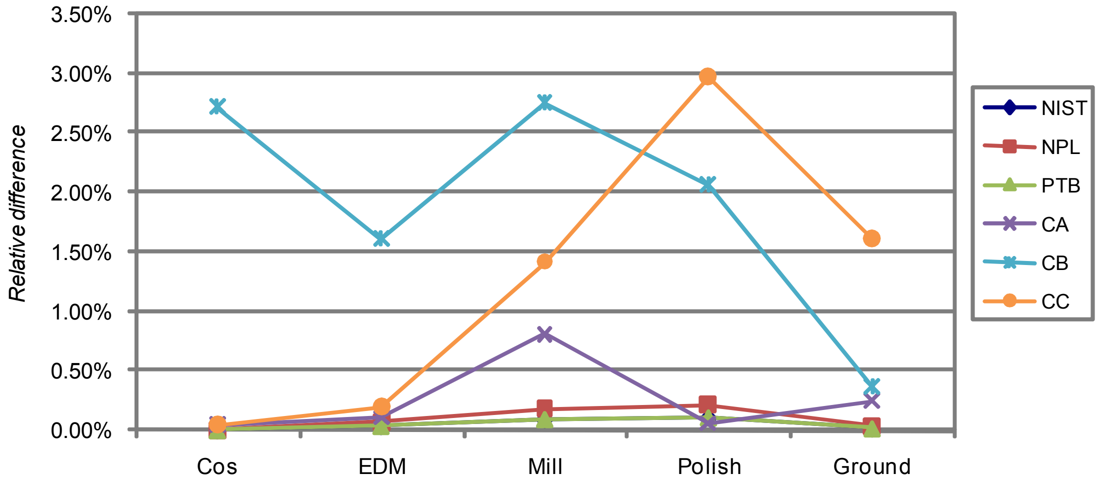
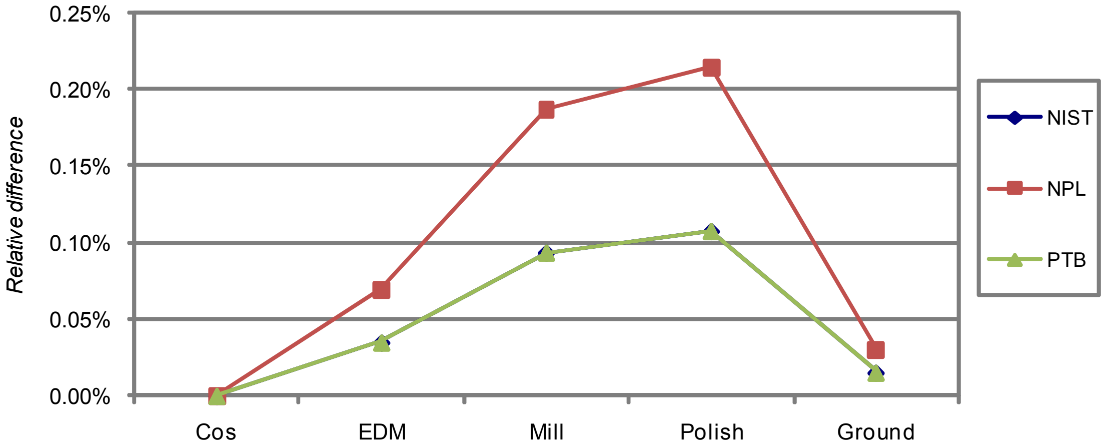


Figure 30: *Pz* parameter analysis



(a)



(b)

Figure 31: *Rt* parameter analysis

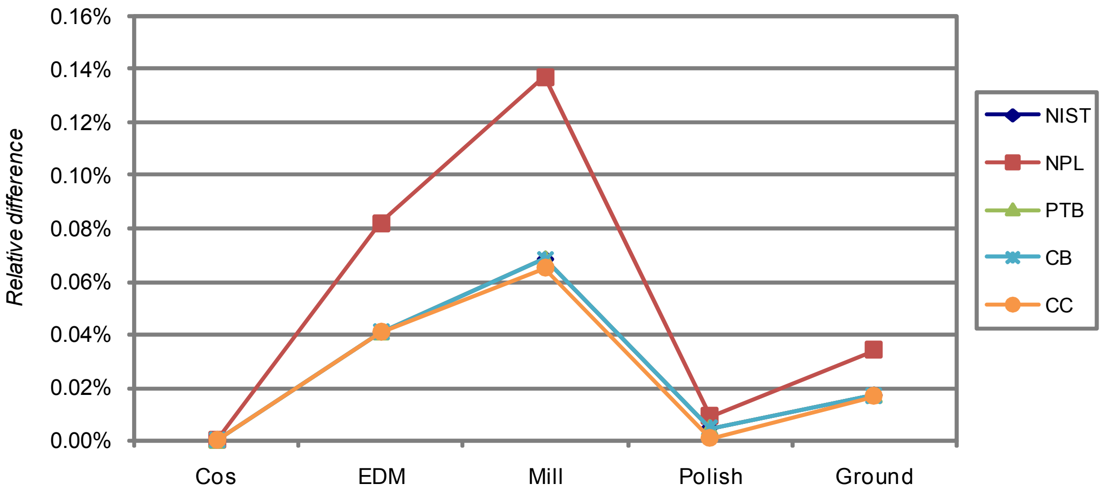
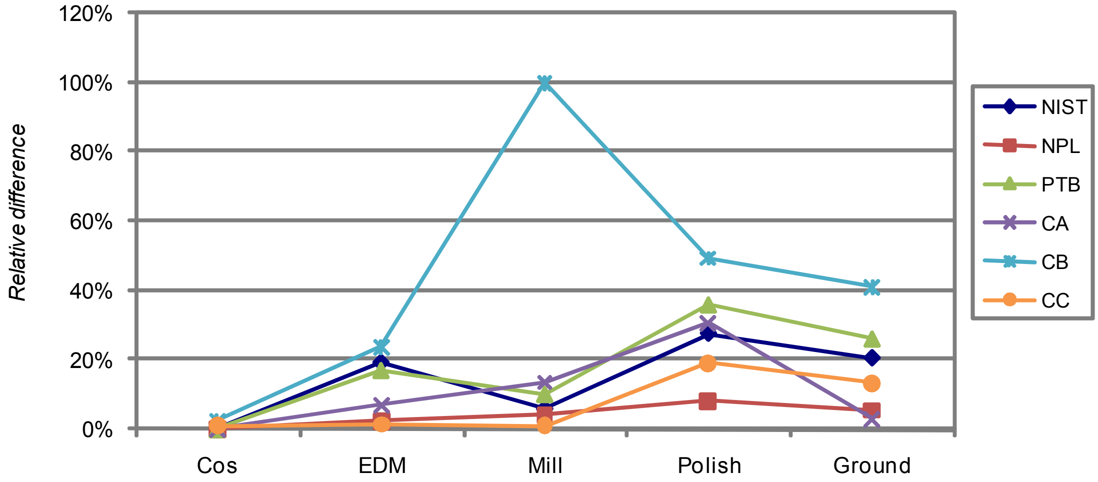
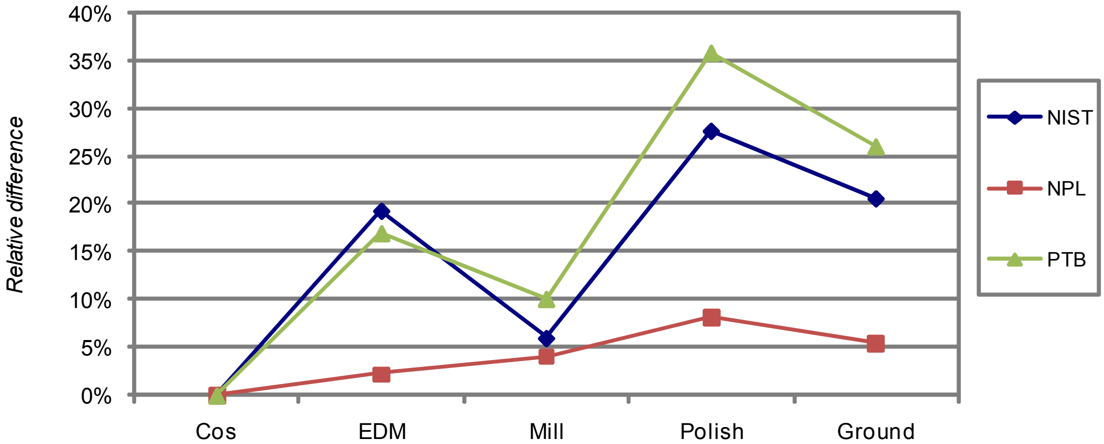


Figure 32: *Pt* parameter analysis



(a)



(b)

Figure 33: *RSm* parameter analysis

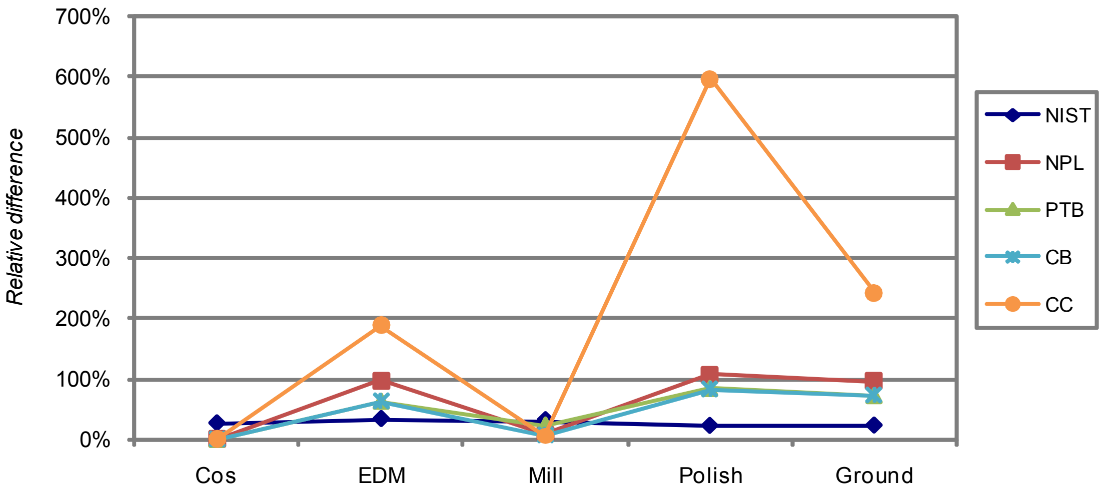


Figure 34: *PSm* parameter analysis

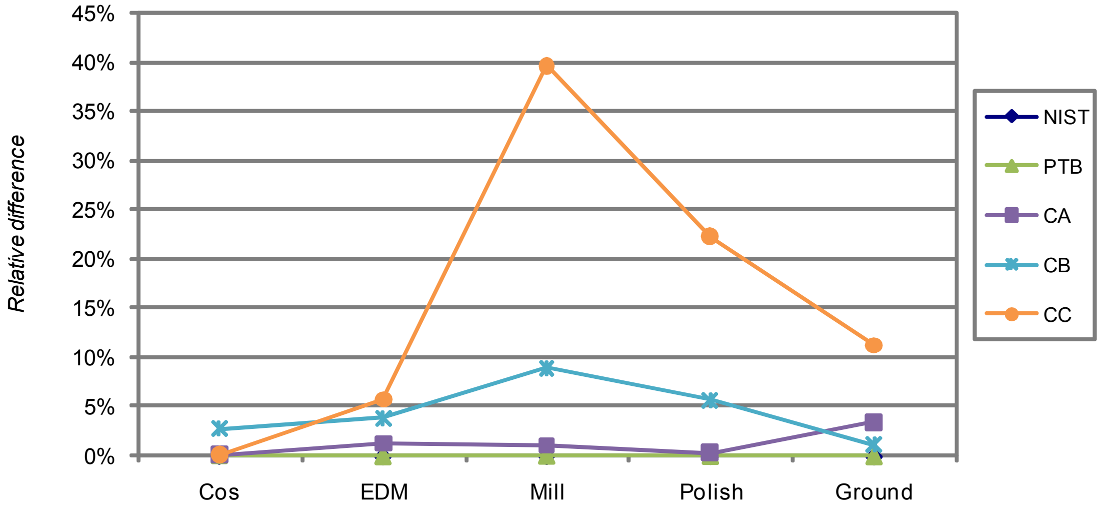


Figure 35: *Rdq* parameter analysis

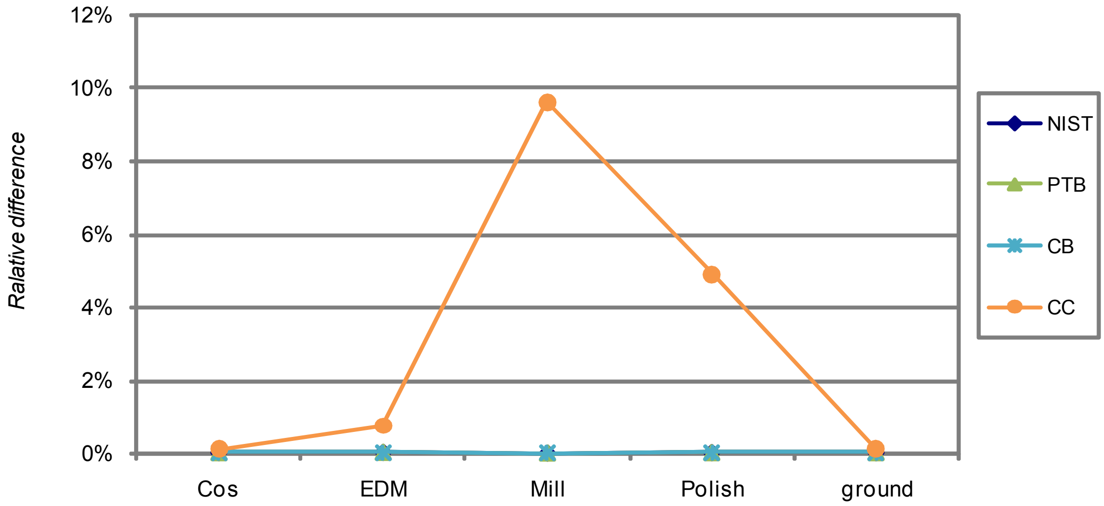


Figure 36: *Pdq* parameter analysis

C.2 The effect of levelling operator

Reference data set:

Cos.smd, Ground.smd and Ground2.smd

Expected Results:

1) No difference for the results obtained from data file Ground.smd and Ground2.smd are expected.

Source of variation:

- Rounding error
- The last point of data file Cos.smd is not taken into account by PTB's F2 standard because a different length definition is used.
- Some software packages use interval-based length definition. Thus, the last point of the data does not use in profile evaluation. For those packages, there is a point shift for assessing data files Ground.smd and Ground2.smd (see Figure 37).

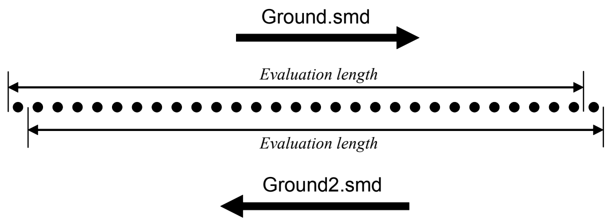


Figure 37: A point shift

Results:

- PTB

1) For data file Cos.smd, *Pp* is 999.96 nm (expected result: 1000 nm) and *Pv* is 1000.04 nm (expected result: 1000 nm) (see Table 19).
2) The relative difference of *Ra* and *Rq* are 0.007 % and 0.005 %. (see Table 26-27)

- CB

1) For data file Cos.smd, *Pp* is 999.96 nm (expected result: 1000 nm) and *Pv* is 1000.04 nm (expected result: 1000 nm) (see Table 19).
2) The relative difference of *Ra*, *Rq* and *Rsk* are 2.056 %, 2.365 % and 39.4 %. (see Table 26 -27)

C.3 The effect of filtering

Reference data set:

Cos.smd

Expected Results:

The transmission characteristic of a filter indicates the amount by which the amplitude of a sinusoidal profile is attenuated as a function of wavelength. According to the ISO 11562, the filter characteristic of data file Cos.smd is calculated as

$$\frac{a_2}{a_0} = 1 - e^{-\pi\left(\frac{\alpha\lambda co}{\lambda}\right)^2} = 1 - e^{-\pi\left(\frac{0.4697\times 0.8mm}{0.16mm}\right)^2} = 0.999999970157$$

where a_2 is the amplitude of the filtered profile and a_o is the amplitude of the cosine wave profile before filtering. The measured data only provides six significant digits. Thus, there is no significant difference between the result of *P*-parameter and *R*-parameter obtained from data file Cos.smd.

Source of Variation:

- Rounding error
- Length definition
- Levelling (PTB and CB)
- Filtering algorithm
- Profile filtering λs (CC)

Results:

	Cos |*Pa-Ra*| /nm	Cos |*Pq-Rq*| /nm	Cos |*Psk-Rsk*|	Cos |*Pku-Rku*|	Cos |*Pp-Rp*| /nm	Cos |*Pv-Rv*| /nm	Cos |*Pz-Rz*| /nm	Cos |*Pt-Rt*| /nm	Cos |*Pc-Rc*| /nm	Cos |*PSm-RSm*| /µm	Cos |*Pdq-Rdq*|
NIST	1.0E-02	1.0E-02	1.3E-04	2.0E-05	0.0E+00	0.0E+00	0.0E+00	0.0E+00	0.0E+00	0.0E+00	2.0E-05
NPL	9.1E-06	1.0E-05	2.2E-09	1.0E-09	1.3E-03	1.5E-05	1.3E-03	6.5E-03	2.9E-05	0.0E+00	*
PTB	2.0E-02	3.0E-02	1.8E-04	0.0E+00	4.0E-02	4.0E-02	0.0E+00	0.0E+00	0.0E+00	0.0E+00	4.9E-05
CA	0.0E+00	1.0E+00	0.0E+00	0.0E+00	1.0E+00	1.0E+00	1.0E+00	1.0E+00	*	*	1.9E-05
CB	1.8E+01	1.9E+01	8.8E-04	1.4E-04	2.7E+01	2.7E+01	5.5E+01	5.5E+01	8.5E+01	2.9E+00	8.0E-04
CC	3.0E-01	4.0E-01	1.0E-04	0.0E+00	6.0E-01	5.0E-01	1.1E+00	9.0E-01	1.1E+00	4.6E-01	1.7E-05

Table 32: Effect of filtering, illustrated by comparing results of *P*-parameters and *R*-parameter

C.4 On the accuracy of test software

Reference data set:

Cos.smd

Expected Results:

The *expected result* of Cos.smd is obtained from algebraic calculation by Maple 10.03.

Performance metric:

The number of correct significant digits obtained from the test software can be calculated by the *log relative error* (LRE)[12] as

$$LRE = -\log_{10}\left(\frac{y_{test} - y_{expected}}{y_{expected}}\right)$$

For example, if $y_{expected}$ = 0.636619 and y_{test} = 0.63663, then LRE= 4.8.

Source of Variation:

- Rounding error
- Length definition
- Levelling (PTB and CB)
- Filtering algorithm
- Profile filtering λs (CC)
- Parameter algorithm

Results:

	LRE				*Significant Digits of its result*[b]
	Pa	*Pq*	*Ra*	*Rq*	
NIST	4.8	4.7	4.5	5.3	5
NPL	7.3	7.8	7.4	10.5	15
PTB	4.8	5.3	4.3	4.4	5
CA	3.2	3.8	3.2	2.9	5
CB	4.3	4.6	1.6	1.6	6
CC	4.5	5.0	3.3	3.2	4

[a] The measured data in Cos.smd provided to six significant digits.
[b] The significant digits of the results delivered by software.

Table 33: The number of correct significant digits[a]

[12] McCullough, B.D., (1998), Assessing the reliability of statistical software: part I. Amer. Statist. 52, 358–366.

C.5 On the direction of numerical processing

Reference data set:

Ground.smd and Ground2.smd

Expected results:

No visible difference of results are expected from two profiles.

Results:

	ΔRa	ΔRq	ΔRsk	ΔRku	ΔRp	ΔRv	ΔRz	ΔRt	ΔRc	ΔRSm	ΔRdq
NIST	0.000%	0.004%	0.052%	0.004%	0.000%	0.000%	0.000%	0.000%	0.332%	1.275%	0.000%
NPL	0.000%	0.000%	0.000%	0.000%	0.000%	0.000%	0.000%	0.000%	0.000%	0.000%	*
PTB	0.007%	0.005%	0.000%	0.033%	0.000%	0.000%	0.000%	0.000%	2.115%	0.336%	0.000%
CA	0.000%	0.000%	0.000%	0.000%	0.000%	0.000%	0.000%	0.000%	*	1.094%	0.001%
CB	2.056%	2.365%	39.428%	2.375%	10.888%	3.363%	7.079%	0.000%	1.392%	4.260%	0.524%
CC	0.000%	0.018%	0.000%	0.009%	0.007%	0.006%	0.003%	0.000%	0.568%	0.638%	0.000%

[a] Simulated by same data set with different order. It illustrates here by the relative difference of results obtained from Ground.smd and Ground2.smd.

Table 34: Effect of direction