



HAL
open science

Outlier filtering: a new method for improving the quality of surface measurements

Gaëtan Le Goïc, Charles A. Brown, Hugues Favrelière, Serge Samper, Fabien Formosa

► **To cite this version:**

Gaëtan Le Goïc, Charles A. Brown, Hugues Favrelière, Serge Samper, Fabien Formosa. Outlier filtering: a new method for improving the quality of surface measurements. *Measurement Science and Technology*, 2013, 24 (1), pp.015001. 10.1088/0957-0233/24/1/015001 . hal-00801636

HAL Id: hal-00801636

<https://hal.science/hal-00801636v1>

Submitted on 21 Nov 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License

Outlier filtering: a new method for improving the quality of surface measurements

G Le Goic¹, C A Brown², H Favreliere¹, S Samper¹ and F Formosa¹

¹ SYMME Laboratory (Systems and Materials for Mechatronics), Université de Savoie, Polytech Annecy-Chambery, F-74940 Annecy-le-Vieux, France

² Surface Metrology Laboratory, Worcester Polytechnic Institute, Worcester, MA, USA

E-mail: Gaetan.Le-Goic@univ-savoie.fr

Abstract

This paper introduces a new filtering technique that aims to improve the quality of measured surface data by removing measurement artefacts, such as spikes and batwings, that impact the data analysis. The methods currently available for measuring surfaces allow a large number of heights on a surface to be measured with high lateral and vertical resolution. However, measured data often contain outliers, which take the form of sharp peaks on the surface (Dirac type) and are particularly common in optical measuring methods. By nature, these peaks cannot belong to a ‘real’ surface; thus, an outlier filtering step is necessary and should be the initial preparation step for the measured data prior to any further analysis. The measurement artefacts can strongly influence the topographic characterization parameters and adversely influence quality control efforts as well as functional analyses for discrimination and correlation of the surfaces. The proposed method makes the filtering of such outliers easier and more effective with criteria linked to the standard deviation (Peirce method) and associated with a modal form-filtering method that is independent of the presence of these peaks. The filter is then applied to the surface at different scales; thus, the method is scale sensitive and improves the filter efficiency. This new method is applied to surface data that were measured by optical means from different surfaces. The surfaces were measured with a 3D-measuring laser microscope and laser scanning confocal microscopy. The results of testing these examples will determine the extent to which this method can improve the quality of measured data and thus influence the results of further analyses.

Keywords: outlier identification, surface analysis, surface topography, surface measurement quality, modal form filtering

1. Introduction

Surface metrology uses high-precision measurement machines that can acquire a set of statistical data on a surface. The data interpretation provides insight into the different levels of default orders, including the form, waviness, roughness and micro-roughness of the measured surface. This interpretation also helps to determine indicators corresponding to the surface characteristics or to perform more advanced analyses on these

data, such as discriminating between the surfaces (area-scale analysis).

In this context, the work presented in this paper is intended to improve the quality of the acquired data by identifying and, if necessary, excluding outliers from the surface measurements. These data are characterized by a distribution of points that does not verify the normality assumption, by multivariate and ordered data, and by the very nature of a surface. Furthermore, the current measurement methods often lead to obtaining

surfaces that are increasingly substantial, with a high number of observations and with potentially many outliers. Finally, many surface analyses and indicators calculated from such data are not robust with respect to the presence of outliers on these surfaces. The outliers are aberrant height samples that are often present in the form of sharp peaks on the measured surfaces (Dirac type). These outliers negatively affect the visualization of the surface in 3D representation (flattening or contracting the scale in false color display); in particular, they influence the analyses and post-treatment results that are applied to these data. Therefore, additional analyses are performed based on the measurement data, and these aberrant points must be considered. The new outlier filter proposed here enables the quick identification (and exclusion, if necessary) of outliers in measured surface datasets.

There are numerous methods for treating aberrant points in a dataset, but their direct applications are unsatisfactory for surfaces for the following reasons: the measured heights samples are arranged on two axes (x, y), and the height samples $z(x, y)$ are nonindependent and not normally distributed in the general case. The objective of this paper is to propose a new method for the detection of outliers, dedicated to surface metrology, i.e. the measurement and analysis of surface topographies. However, the principles and methods presented can be transposed to all field measurement data.

2. Principle, definition and methods to identify outliers

2.1. Historical and principle

Pearce (1852) was the first to establish a formal exclusionary rule, i.e. an objective criterion to identify outliers of a dataset. His work, based on calculations of probabilities from a normal distribution model have inspired most of current methods, including those of Chauvenet (1868), Wright (1884), and more recently Grubbs (1969), Stefansky (1972). The key concepts of this reference method, and how it can be implemented with the current means of computation are presented in section 2.3. However, the limitations of this approach in terms of computation time and for large data samples make its application to surface metrology issues difficult.

The proposed approach for surface measurements uses the method of Grubbs (also called Grubbs' test). Grubbs' test allows for a criterion related to the level of risk in identifying outliers with respect to points modelled by a normal distribution. The method then compares the absolute values of the reduced deviations to a limit value that is calculated from the critical values of Student's t distribution (Student 1908, Hotelling 1931). However, this test is effective only if the assumption of normality is verified, which is not the case of raw data measured on surfaces. Moreover, it does not consider that the measured heights are arranged in space. To better address the specific problem of outliers on measured surfaces, we propose to provide partial answers to the following fundamental questions in the study of outliers put forth by Barnett (1978):

- (1) What are the possible causes of outliers in statistical data?

- (2) In what way do outliers influence data analysis?
- (3) What probabilistic model might be employed to explain the presence of outliers?

To which we respond:

- (1) There are many potential causes for the presence of outliers on surface measurements; these causes are primarily related to the means of measurement (i.e. sensor or optical measurement) or the presence of alterations in the measured sample (i.e. dust, marks or dirt). This point will not be developed in this study. Instead, this study focuses on issues (2) and (3).
- (2) Before attempting to deal with aberrant points, the influence of these points on analyses from a measured surface must be emphasized. For example, many parameters of the standards ISO-8785 (1998) and ISO-25178-3 (2008) for surface parameters are not robust with respect to outliers, i.e. their values are sensitive to the outliers (see figures 10(e) and (f)). This feature can also be true of many more advanced analyses using surface measurements. One can cite the example of area-scale analysis (ASME-B46.1 2009, Brown *et al* 1994) of different measured surfaces to discriminate between them (Jordan and Brown 2006, Scott *et al* 2005), or to find correlations with behaviour (Brown and Siegmann 2001, Berglund *et al* 2010) or processing (Cantor and Brown 2009), since this analysis can normally best be performed after filtering the surface outliers, if they exist, although this is not always the case (Brown and Brown 2010, Vessot *et al* 2012).
- (3) The distribution assumptions used to detect the presence of supposed outliers is crucial; in the case of surface metrology, the assumption that the dataset is approximately normally distributed (normality assumption) does not hold in general; thus, the methods of outlier detection based on a criterion related to the standard deviation would not be optimal. Instead, a transformation is preferred to satisfy this hypothesis. We propose filtering by the modal method (see section 3.3.2) and then identifying the outliers using Grubbs' test.

Finally, we demonstrate in this work the importance of taking into account the scale of analysis in surface metrology. The proposed method for outliers identification uses the fact that for measured surfaces, an outlier is defined in relation to its immediate environment, i.e. it differs greatly from its neighbours, but not necessarily from all of the observations (measured heights). Thus, we vary the size and position of the neighbourhood, i.e. analysis window, recursively when identifying outliers. The results show that the scale of the analysis is fundamental to the study of surfaces.

2.2. Definition

The literature offers many definitions of outliers. Grubbs (1950) and Grubbs and Beck (1972) proposed a definition that is often used to define an outlier, considering that an outlying observation, or outlier, is *one that appears to deviate markedly from other members of the sample in which it occurs*. The

standard ISO-16269-4 (2010) on the detection and treatment of aberrant values defines an outlier as being *a member of a small subset of observations that appears to be inconsistent with the remainder of a given sample*.

In the specific case of surfaces, if we use these definitions, outliers must constitute a small subset of all of the studied surface's points. Furthermore, we propose to complete the definition of an outlier in the particular case of surfaces, emphasizing that, unlike a dataset in the usual sense, measured height data on a surface are not independent because they are ordered or linked by their respective positions on two axes (x, y) . Therefore, an outlier on a surface is defined in comparison with its environment on the surface and not necessarily in relation to all of the surface's measured heights.

2.3. Identification and rejection of outliers : existing methods

In their book, Barnett and Lewis (1994) reviewed different methods for treating observations *surprisingly far away from the main group* of a dataset. They distinguish four methods for treating these observations:

- *Accommodation*: Protect further analysis against outliers.
- *Incorporation*: Replace the initial model by another model in relation to which observations appear discordant.
- *Identification*: Test if a point is formally aberrant or not.
- *Rejection*: Eliminate outliers from the surface measurement.

Labelling, or flagging potential outliers for further analysis, can also be added to these strategies (Iglewicz and Hoaglin 1993); this method may be appropriate when outliers are related to information useful about the surface. Outliers can sometimes indicate information that is scientifically useful for the analysed sample and do not necessarily represent a degradation of the data. For example, for a surface measured by optical means (without contact), outliers may be related to the heterogeneity of reflectivity. The strategy chosen in this work is to *identify* and *reject* (if appropriate) outliers detected in the data. There are numerous methods for identifying and eliminating outliers that are commonly used in academia and industry. The objective of this study is to remove the outliers while avoiding any changes to the other measured heights. The intent is to improve subsequent analyses. The removal of outliers should improve the ability to discriminate surfaces suspected of having different topographies and to correlate surfaces with measures of topographically related processing and performance phenomena. Skilful removal of outliers should also improve the use of surface metrology in quality control.

The standard ISO-16269-4 (2010) proposes a strategy for processing general data and details how to detect and process outliers in the general case. Before applying a rule for identifying outliers, this method calls for transforming the data to verify the assumption of normality. If the data are not normally distributed, then it is imperative that the data analysis procedures are robust with respect to outliers.

As for surface metrology, the typical methods cannot be directly applied for at least three reasons. First, the measurements do not follow a normal distribution (see

section 3.2), second, the data are linked to each other by their relative positioning on the axes (x, y) ; and third, there is a large number of observations (measured heights). It is interesting to examine the various methods and exclusion tests proposed in the literature to better understand the problem of outliers in 'general' data and observe whether these methods can provide answers in the context of data arising from measured surfaces.

In 1805, the French mathematician Legendre (1805) attempted to determine the orbit of comets from three observations of latitude and longitude. Faced with anomalous observations, he already proposed to exclude data that were 'too large to be admissible'; the exclusion criteria remained subjective.

In 1852, Peirce (1852) formalized the problem of identifying outliers, by establishing an exclusionary rule based on the use of a Gaussian distribution model, and a probability calculation from this model. In the field of outliers, these are seminal works. We describe in this section how to implement this exclusion criterion. The exclusionary rule of Peirce is based on the following principle:

'The proposed observations should be rejected when the probability of the system of errors obtained by retaining them is less than the probability of the system of errors obtained by their rejection multiplied by the probability of making so many, and no more, abnormal observations'

Gould (1855), in his correspondence with Professor A D Bache in 1855, takes up the work of Peirce to facilitate the implementation of the rule. In particular, he defines logarithmic calculation tables and special functions, allowing, with the means available in the late 19th century, to proceed with the identification of outliers by using the criterion of Peirce. We propose to examine the equations written by Gould and to show how they are used in practice, that is, how the equations establish the critical deviation x_{peirce} , from which an observation is identified as being an outlier.

Firstly, the critical deviation x_{peirce} is a function of N , the whole number of observations, k , the number of outliers to be detected and m , the number of unknown quantities contained in the observations. Equation (1), defined by Gould, gives a first definition of $R(x)$:

$$R(x) = e^{\frac{1}{2}(x^2-1)} \cdot \psi(x) \quad \text{with} : \psi(x) = \frac{2}{\sqrt{2\pi}} \cdot \int_x^\infty e^{-\frac{1}{2}t^2} dt \quad (1)$$

$\psi(x)$ is the probability that a standard normal distributed variable ($\mu = 0, \sigma = 1$) does not lie between $-x$ and $+x$. Thus, this function corresponds to the probability of having an abnormal observation in a dataset, as stated in Peirce criterion. Combining equations (2a) and (2b), given by Gould, yields a new equation (3), defining R as a function of x :

$$\lambda^2 = \frac{N - m - kx^2}{N - m - k} \quad (2a)$$

$$\lambda^{N-k} \cdot R^k = Q^N \quad (2b)$$

$$R(x) = Q^{N/k} \cdot \left(\frac{N - m - kx^2}{N - m - k} \right)^{\frac{-(N-k)}{2k}} \quad (3)$$

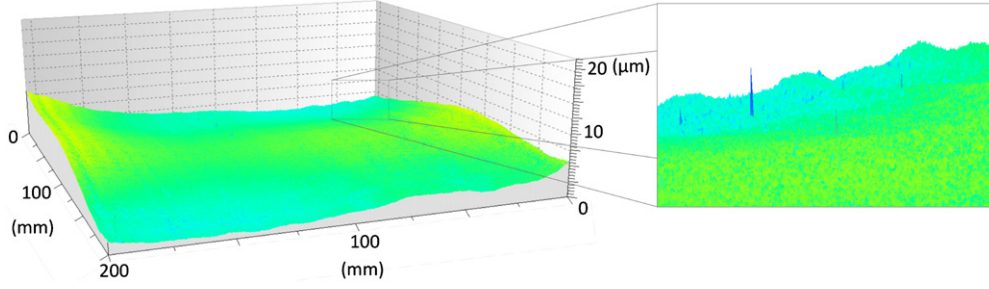


Figure 1. Rendering in 3D perspective of a measurement of Surf-1.

Then, it is possible to solve the system constituted by equations (1) and (3), and to determine the critical value of x (i.e. the critical deviation x_{peirce}), defined as a function of N (whole number of observations), k (number of outliers to be rejected) and m (number of unknown quantities). The criterion (4) is applied initially for $k = 1$, then the number of potential outliers is gradually increased until no aberrant points are identified according to the criterion of Peirce. Ross (2003) described the manner in which this principle can be implemented

$$|x_i - x_m|_{\max} = x_{\text{peirce}} \times \sigma \quad (4a)$$

$$|x_i - x_m| > x_{\text{peirce}} \times \sigma \Leftrightarrow x_i \text{ identified as an outlier} \quad (4b)$$

with	{	$ x_i - x_m _{\max}$	Maximum allowable deviation
		x_{peirce}	Peirce critical value
		σ	Standard deviation of the dataset
		x_i	Tested observation
		x_m	Mean of the dataset.

For the issue of outlier detection, the method of Peirce is a fundamental initial contribution: it defines the general principles of outlier identification and provides a formal and effective criterion. For a significant number of observations, the calculation of the critical deviation x_{peirce} (approximation by Newton's method), recursively, makes this method difficult to apply to more consistent data samples, and thus not suitable to deal with outliers on surface measurements. A few years later, Chauvenet (1868) proposed a new formalization for outlier identification. It is close to that proposed by Peirce but less general. The Chauvenet criterion is easier to implement and is still used. More recently, Grubbs' test (Grubbs 1969, Stefansky 1972), also referred to as the maximum normed residual test, proposed a general and effective method to identify outliers in the general case (unordered data), with the assumption of normality. In the same manner as Peirce, Grubbs' test processes one outlier at a time until no outliers are detected. Similarly, its generalized form (Tietjen and Moore 1972) allows the method to be applied in the case of a known number of outliers. Finally, Rosner (1983) and Paul and Fung (1991) proposed a method called the generalized extreme studentized deviate test (ESD test), which was adapted to the case of an unknown number of outliers in the data.

The existing methods for addressing outliers in a dataset are generally univariate and only apply in the case of normally distributed data, making their direct application difficult

and inefficient in the case of measured surfaces, especially when the surfaces have a large number of measured heights. However, we show how one of the most effective methods for a general set of data (Grubbs' test) can be applied to identify outliers on a measured surface, under the condition of previously assumed normality and considering the notion of a local outlier. In this work, we take the specific properties of the measured surfaces (nonnormal distribution, multivariate and nonindependent data) into account to identify outliers effectively in these types of data, using the general principles defined in the standard ISO-16269-4 (2010).

3. Dealing with outliers: a new approach for surface measurement data

3.1. Material and means of measurement

To illustrate the proposed method, we rely on two surface measurements.

Surf-1 is a measured surface of a glass plane with flatness calibrated at $\lambda/5$, i.e. a fifth of the wavelength of light used or a flatness deviation of about $0.1 \mu\text{m}$ on the surface. With the topography of this surface known at a high resolution (prior measurement of the glass plane by interferometry), the objective of these measurements is to determine their capability to potentially identify systematic surface defects (form-waviness) generated by surface measuring machines (Favrelière *et al* 2011). This identification makes it possible to apply a correction to the measurements made and thus improve their accuracy.

The measurement was performed with a wide-field confocal microscope (Altisurf 520), equipped with a confocal chromatic optical probe with a field depth of $300 \mu\text{m}$ and a Z -axial resolution of $0.06 \mu\text{m}$. The measuring region is 1001×1001 ($\approx 10^8$) measured heights, and the sampling intervals on the X - and Y -axes are 0.2 mm . After a quick visual assessment (see figure 1), some outliers appear that could be due to dust on the glass during the measurement. A 3D representation of the measurement and a close-up view of a region with outliers are shown in figure 1.

Surf-2 is a portion of the edge of a ski. The study of this surface was performed with the aim of improving the modelling of the ski-snow contact link on the roughness

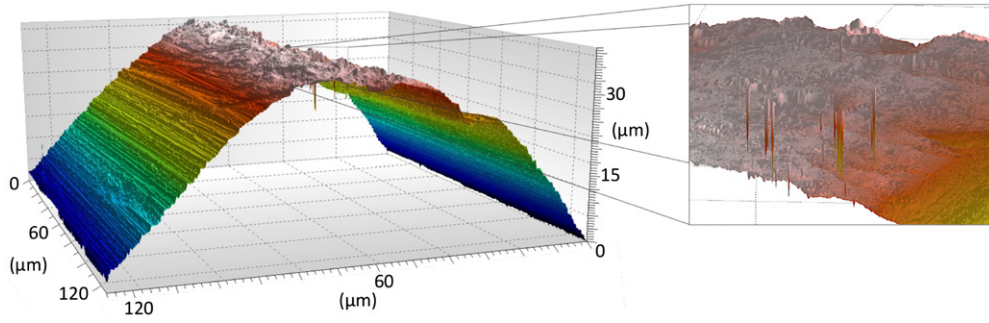


Figure 2. Rendering in 3D perspective of a measurement of Surf-2.

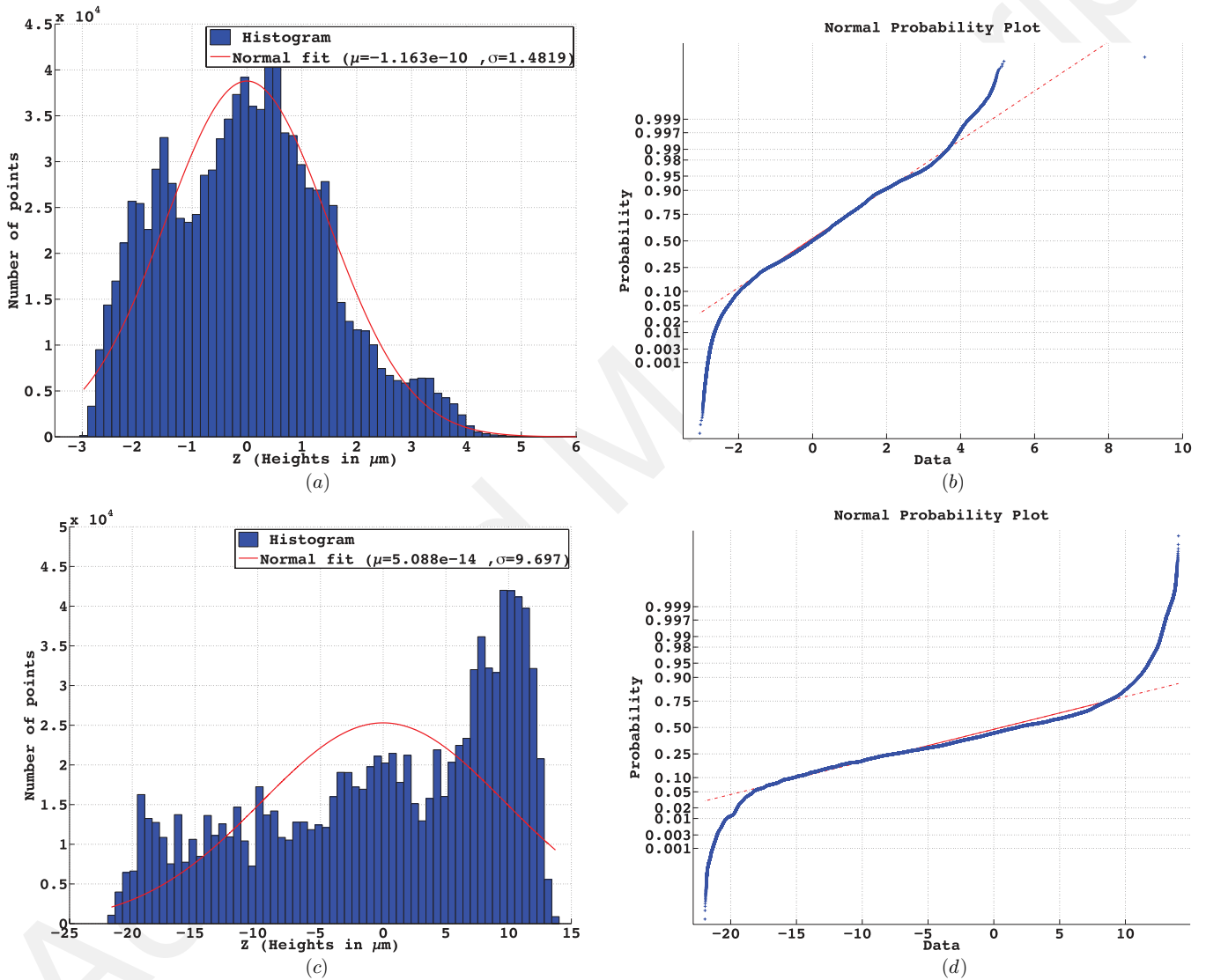


Figure 3. Normality test. (a) Surf-1: histogram and fitted normal. (b) Surf-1: Q-Q plot (Henry diagram). (c) Surf-2: histogram and fitted normal. (d) Surf-2: Q-Q plot (Henry diagram).

scale. The measurements were made with a 3D measuring laser microscope (Olympus Lext OLS4000), equipped with a 100× objective. This surface presents slopes and reflectivity changes that make it difficult to measure. This difficulty partly explains the presence of many outliers in

the obtained data. The measuring region is 1024×1024 measured heights, and the sampling intervals along the X- and Y-axes are $0.125 \mu\text{m}$. A 3D representation of the measurement and a close-up view of a region with outliers are shown in figure 2.

3.2. Height distributions

According to Estler (1999), the result of a measurement is a *probability distribution that provides an unambiguous encoding of one's state of knowledge about the measured quantity*. Metrology is thus defined as a statistical interpretation of the data obtained from the measurement. To reach this interpretation, it is thus useful to determine the nature of our data prior to any other analysis and to determine whether the measured heights verify the assumption of normality (see section 2.3).

Figures 3(a) and (c) show the height distributions from the two surfaces and a representation of the normal distribution nearest to that of the measured heights. In addition, figures 3(b) and (d) provide the Q–Q plot (Henry diagram), to more accurately assess the normality of the data (Wilk and Gnanadesikan 1968). For these two surfaces, the measured heights are far from a normal distribution. Furthermore, the fitted normal distributions are centred at zero because these measurements had been adjusted by the least squares method prior to analysis. At this stage, we cannot apply a criterion for identifying outliers related to the standard deviation. As a first step, we propose (section 3.3.2) the application of a transformation to these data to approximate a normal distribution.

3.3. Proposed method: filtering outliers on the surface measurement data

3.3.1. Compliance with the normality assumption. Following the finding of nonnormality made in the previous section (section 3.2), we describe a method for transforming the distributions of points close to that of the model used (normal distribution) for surface measurements.

The study of *flat* surfaces, i.e. those with surface defects of the order of waves and defects that arise solely from roughness or microroughness, shows that the data distributions obtained are almost normal. Thus, as a first step, we propose the application of a form filter to the measured surfaces. The outliers are then identified from the residuals, i.e. filtered heights; thus, we can easily exclude outliers identified on the original surface if necessary. The proposed method for filtering the form on the surfaces is detailed in section 3.3.2.

3.3.2. Form filtering using the modal method. This method is based on discrete modal decomposition (DMD), which is a mathematical tool for evaluating a discrete spectral representation of a surface measured by heights, such as the discrete Fourier transform or discrete wavelet transform applied to surface measurements. We have initiated this decomposition to describe the geometry of mechanical parts (Favrelière 2009). In particular, the decomposition describes the measured surface with a family of discrete functions, called modes or modal deformations. We can express the measured surface, represented by the vector ${}^{\text{meas}}V$, by the following linear combination:

$${}^{\text{meas}}V = \sum_{i=1}^{N_q} \lambda_i \cdot Q_i + \epsilon(N_q), \quad (5)$$

where λ_i defines the modal amplitude of the mode Q_i , N_q represents the number of modes chosen and ϵ the residue. The modes Q_i are obtained by the resolution of a classical vibratory mechanical problem of the type

$$M \cdot \ddot{q} + K \cdot q = 0 \text{ with } \begin{cases} M : \text{generalized mass matrix} \\ K : \text{generalized stiffness matrix} \\ q : \text{displacement vector.} \end{cases} \quad (6)$$

In special cases, this problem can be solved analytically (Leissa 1969), and in the general case, the problem can be solved numerically by the finite element method (Zienkiewicz *et al* 2002). In general, the resolution is numerical, which implies that the modal decomposition is discrete. With an approach similar to the Fourier transform, but applicable to all types of geometric elements, DMD is used to configure the measurement of surfaces in a space of discrete functions unique to each type of surface. These discrete functions (i.e. the set of modes Q_i) have many properties and form a *geometric* vector space.

The modal base Q naturally arranges the modes Q_i by increasing geometric complexity (arranged from long wavelength to short wavelength) because it uses the natural order of increasing frequencies. Due to this property, DMD can be used as a filtering method by reconstructing the measured surface in the space of modes Q_i . The reconstruction operation consists of realizing the linear combination of modal amplitudes λ_i and the associated modes Q_i . As an example, figure 4 shows modal filtering for a measured surface that allows the surface's various components to be extracted.

Figure 5 shows the surfaces obtained at different stages of filtering and the modal amplitude spectra associated with this filtering for both study samples. The visualization of the modal spectra (figures 6(a) and (b)) shows that selecting the number of filtration modes N_q equal to 50 is sufficient for filtering all of the high-amplitude modal contributions on the analysed surfaces. We note that the modal contributions decrease rapidly; this feature, which is detailed in the work of Formosa *et al* (2005), Le Goïc *et al* (2011) and Grandjean *et al* (2012), is intrinsic to the modal method and allows a surface with a small number of parameters to be accurately approximated without any *a priori* knowledge of the surface. This feature allows us to effectively filter the components of the form on the surfaces measured, using the modal method in a generic manner (i.e. without adapting the filter parameters for each surface). Figures 5(b) and (e) show a 3D representation of the filtered components during the form filtering for Surf-1 and Surf-2. These surfaces represent the contributions of the first 50 modes of the modal decomposition.

Figures 7(a) and (c) show the histograms of the measured data after form filtering and a normal distribution fit to these data. We found that the distributions are quasi-normal. Figures 7(b) and (d) provide the Q–Q plot diagrams (Henry diagram) of the data, to more accurately assess the normality of the data (figures 3(b) and (d)). Given the large number of measured heights present in close proximity to this line and the distributions after the form filtering, the analysed data are almost normal at this stage and allow us to consider the use of a criterion for identifying outliers associated with a standard deviation.

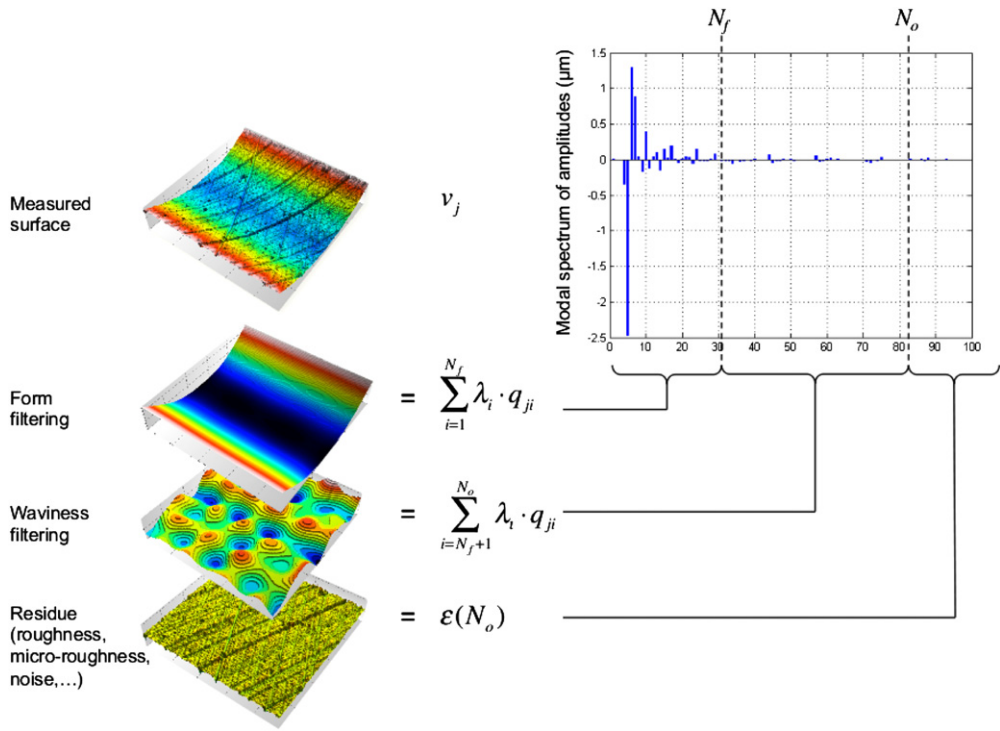


Figure 4. Modal filtering—principle and application.

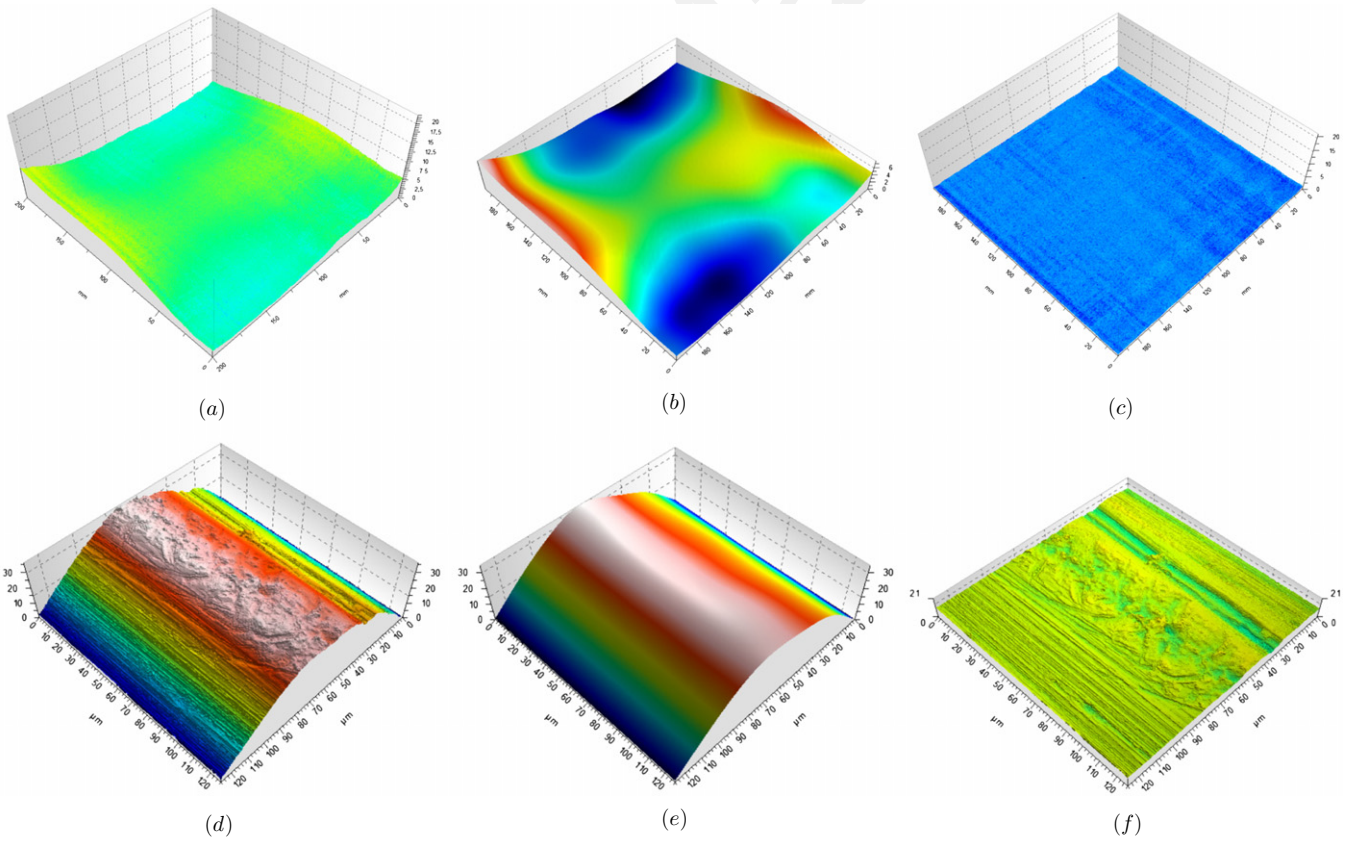


Figure 5. Modal form filtering: (a) Surf-1: measured surface (${}^{\text{meas}}V$), (b) Surf-1: form (${}^{\text{form}}V$), (c) Surf-1: residue (${}^{\text{form}}V - {}^{\text{meas}}V$), (d) Surf-2: measured surface (${}^{\text{meas}}V$), (e) Surf-2: form (${}^{\text{form}}V$), (f) Surf-2: residue (${}^{\text{form}}V - {}^{\text{meas}}V$).

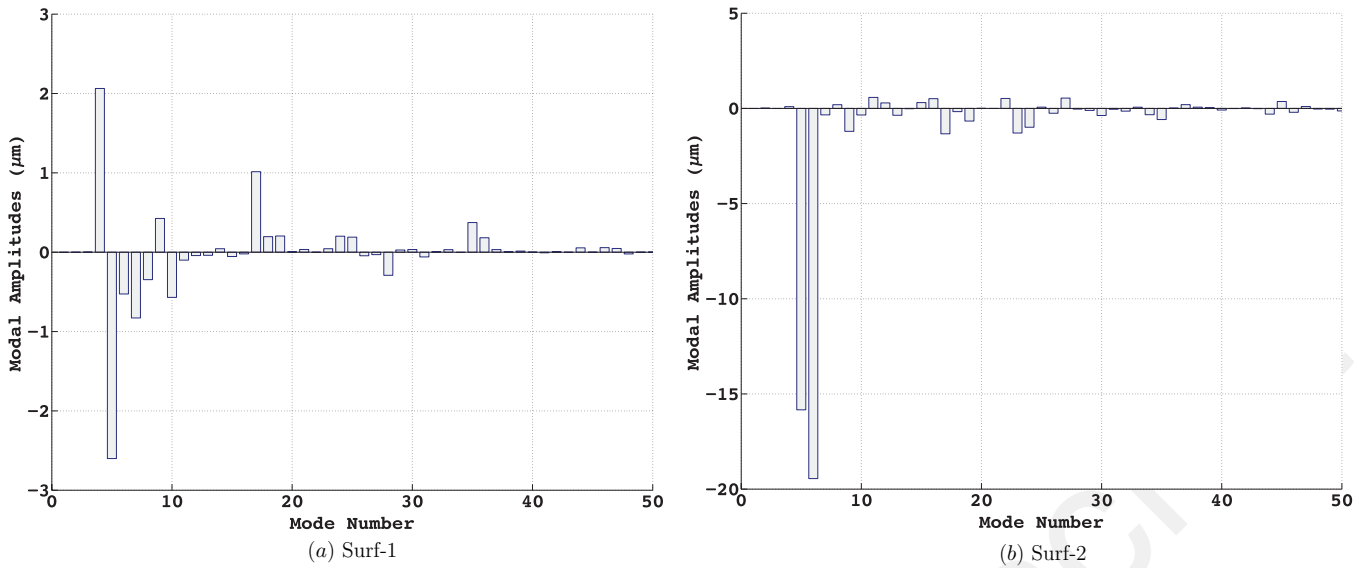


Figure 6. Modal spectrum of amplitudes.

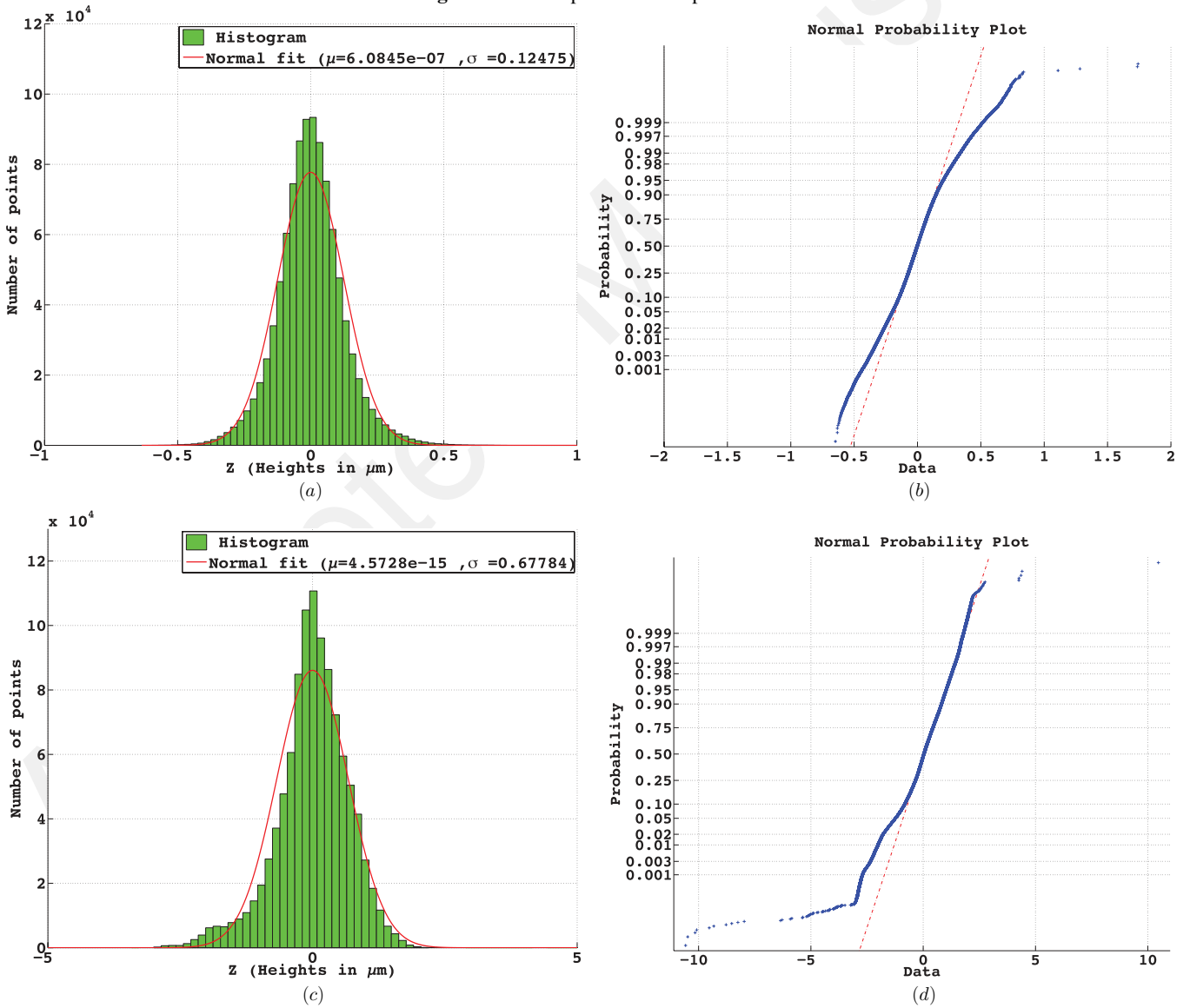


Figure 7. Normality test after modal form filtering. (a) Surf-1: histogram and fitted normal. (b) Surf-1: Q-Q plot (Henry diagram). (c) Surf-2: histogram and fitted normal. (d) Surf-2: Q-Q plot (Henry diagram).

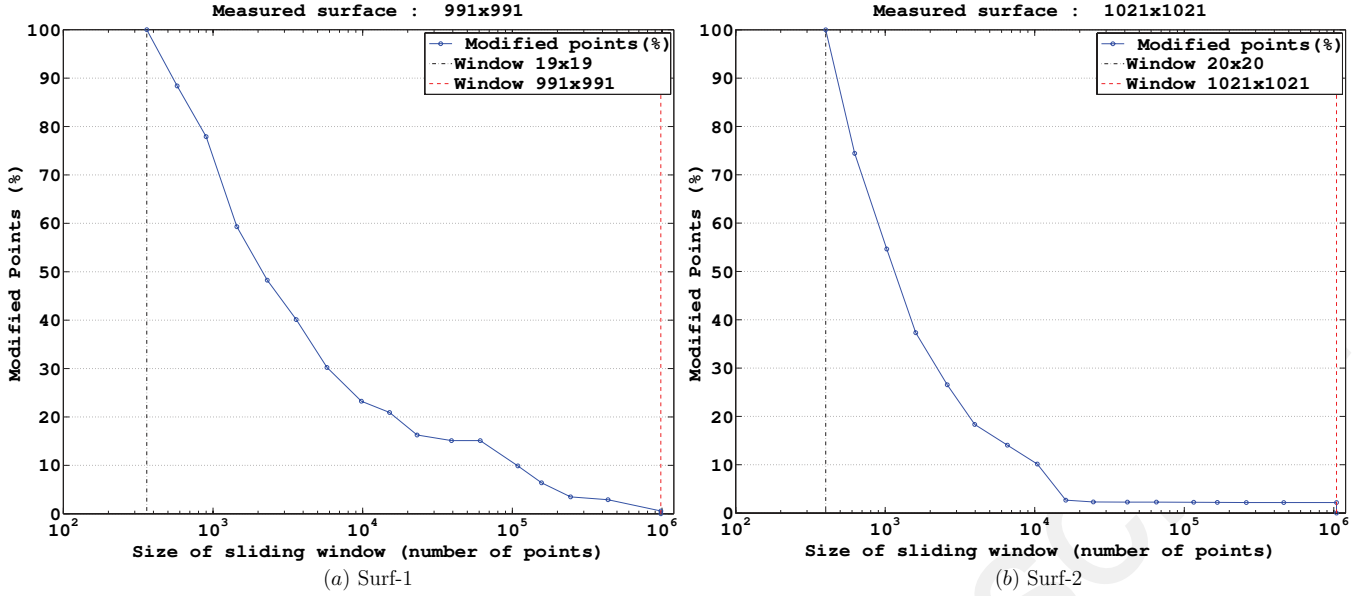


Figure 8. Evolution of the percentage of modified points at each analysis window size.

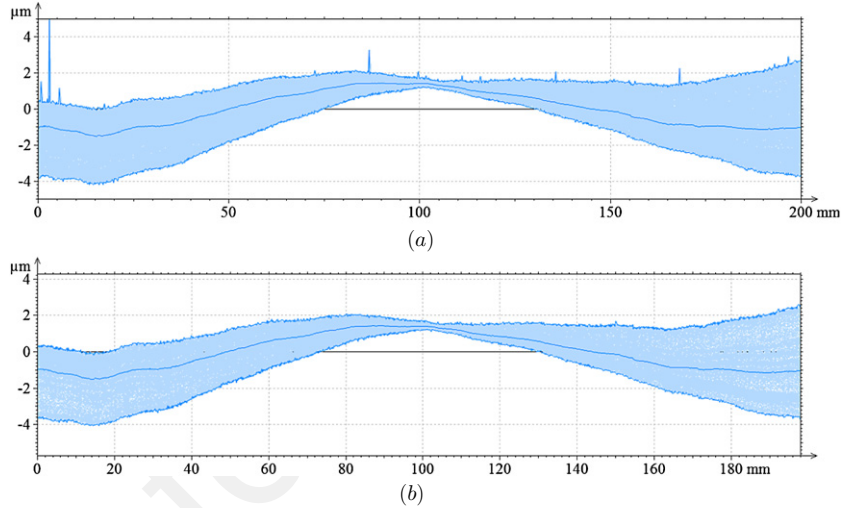


Figure 9. Series of profiles extracted from Surf-1: (a) original surface, (b) after outlier filtering.

3.3.3. *Identify outliers using a scale-sensitive standard deviation criterion.* (a) *Standard deviation criterion.* Because the transformation performed in the previous stage allows us to arrange the data close to a normal distribution, we can consider the identification of outliers using a criterion based on the standard deviation. As discussed in section 2.3, there are many criteria that, based on the probability calculations, are more or less complex; the criterion of Peirce is a good example of these criteria. With the current methods of measurement, the number of observations (measured heights) is often large ($>10^6$), which sometimes makes it difficult to apply criteria based on recursive operations to measured surfaces; however, other criteria are more suitable for smaller batches of data. We chose to use a criterion based on Grubbs' test because it is known for its speed and efficiency in the case of general data; however, we applied it to take specificities of the measured surfaces into account. The test is applied on an analysis window of variable size that slides (to cover the entire

surface) and is recursive over the surface (see section 3.3.3.b). We show that this method allows outliers on a surface to be effectively and quickly identified, while minimizing the risk of misidentification of outliers.

The Grubbs (1969) test is a statistical test based on the ratio of the standard deviation to the distribution of the data. This test assumes a normal or near normal distribution and calculates a numeric value G_i (called the *sample criterion*, see equation (7a)) for each observation (measured point). This value is compared to a value G_{lim} , calculated from the critical values of Student's t distribution (see equation (7b)), to determine whether the observation in question must be identified as an outlier. G_{lim} is a limited value, which can be exceeded by the sample criterion with a probability called the *significance level*, or as α risk (given the fact that we model the data distribution with a normal distribution). This risk can also be observed as the risk of erroneously rejecting an improbable observation. Rather than setting an arbitrary standard deviation

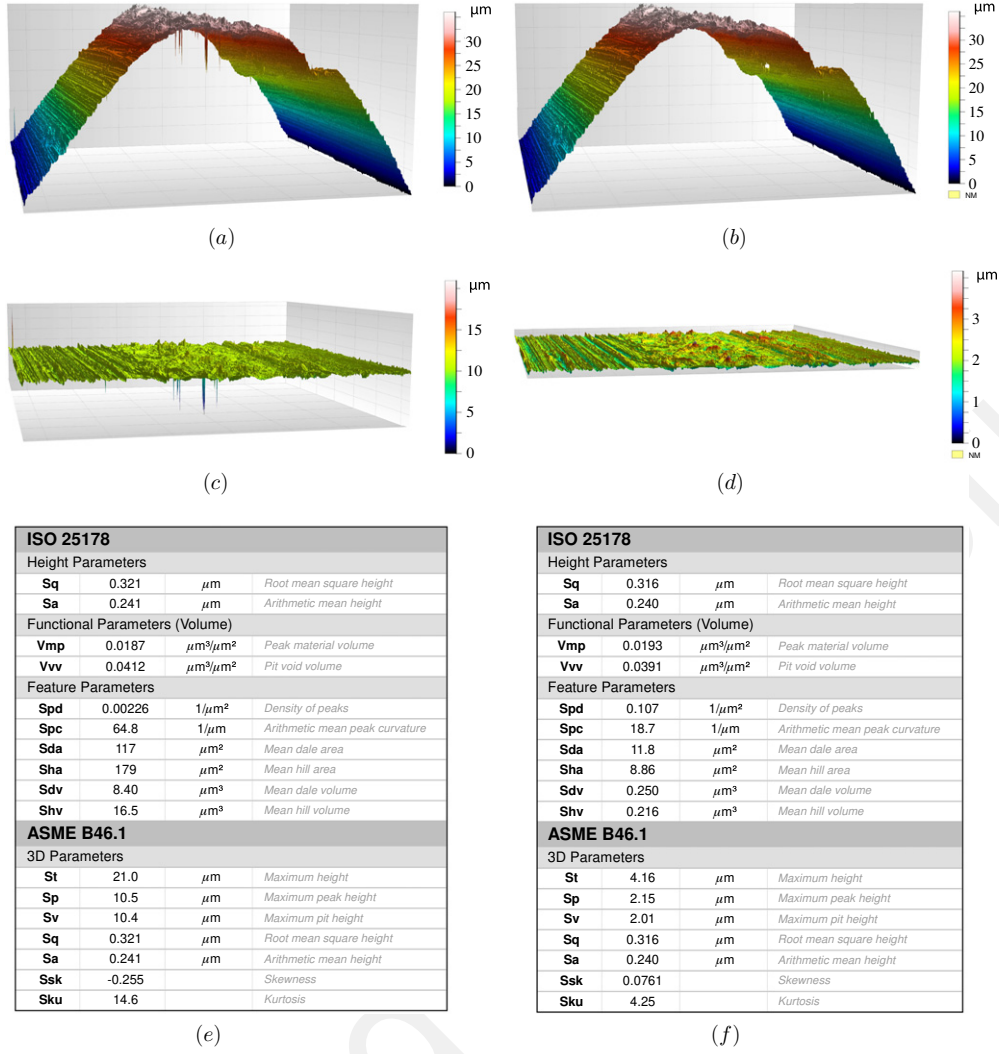


Figure 10. Outlier filtering on Surf-2: (a) original surface, (b) after outlier filtering, (c) original surface—roughness, (d) after outlier filtering—roughness, (e) after outlier filtering—surface parameters, (f) original surface—surface parameters.

threshold to identify outliers, Grubbs' method allows a risk level to be used as the basis for setting the threshold of identification

$$G_i = \frac{\text{Max}(|X_i - \bar{X}|)}{\sigma} \quad (7a)$$

$$G_{\text{lim}} = \frac{(N-1)}{\sqrt{N}} \sqrt{\frac{t_{(\alpha/N, N-2)}^2}{N-2 + t_{(\alpha/N, N-2)}^2}} \quad (7b)$$

where

- G_i is the sample criterion
- G_{lim} is the Grubbs' test critical value, defined for a risk α
- $t_{(\alpha/N, N-2)}$ is the Student's critical value t for a sample of size N at $N-2$ degrees of freedom.

For each observation, we verify whether G_i is greater than G_{lim} . If G_i is greater than G_{lim} , X_i is considered an outlier; the point is removed from the sample, and the procedure is repeated until no additional outliers are identified on the surface or the portion of the surface analysed.

Given the type of data and potential outliers arising from the surface measurements, a low level of risk should be chosen for identifying outliers. We chose a α risk value $\alpha = 0.001$, which corresponds to a confidence level of 99.9%. The user can also set this parameter; in the case of particular surfaces or requirements, the user can select the appropriate value.

(b) *Scale-sensitive method for data measurements.* As stated in section 2.2, an outlier on a surface is defined primarily according to its close neighbours and not necessarily in relation to all of the surface's measured points. The notion of a *close neighbour* is actually a type of scale-based analysis. We propose performing the identification of outliers, according to the aforementioned criterion, over an analysis region of variable size to account for this specificity of the measured surfaces. The criterion (see section 3.3.3(a)) is first applied on a global scale (analysis window covering the entire surface); the size of the analysis window is then gradually reduced, and the window moves over the entire surface (sliding window). At each window size, a new identification is performed on the subsurface that is rectified by the least squares method (to recentre distributions), and the identified outliers are replaced

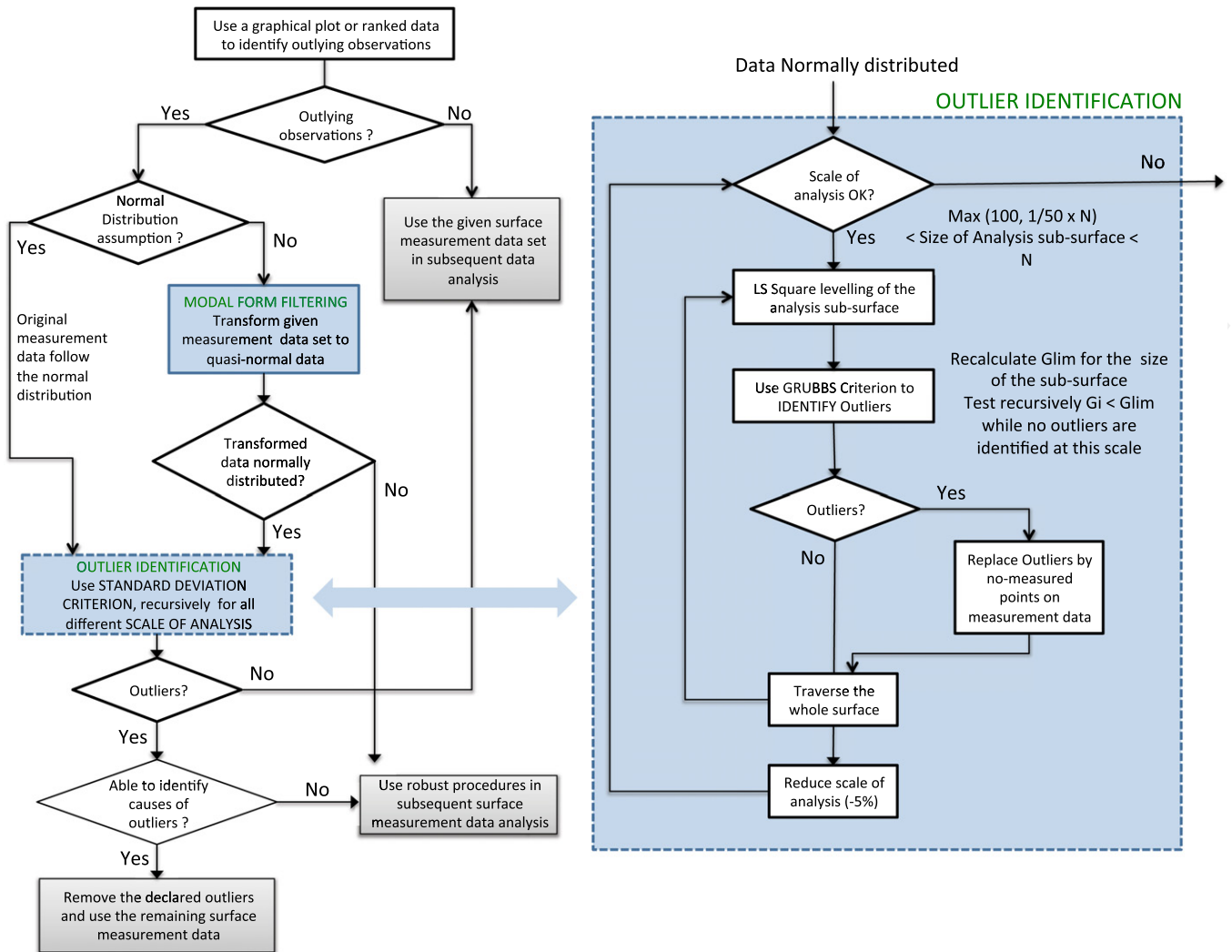


Figure 11. Flow chart—detection and treatment of outliers in the surface measurement data.

by nonmeasured points. Figure 8 shows the evolution of the number of outliers removed (as a percentage of the total points identified) as a function of the size of the analysis window for samples Surf-1 and Surf-2. These results confirm the importance of varying the scale of analysis for detecting outliers for this type of data.

The parameters for reducing and displacing the sliding window were chosen using the filtering time, with a constant filter efficiency as the selection criterion. A reduction factor of 5% (calculated with respect to the previous window size) was applied to each of the iterations, providing a nonlinear evolution of the sliding window size (more rapid decreases at 'global' levels and more progressive decreases at local levels, i.e. smaller window sizes). The plots of the number of points identified as outliers (figures 8(a) and (b)) show that this method for reducing the scaling factor allows a good distribution of heights to be identified at different scales of analysis. The second factor related to the sliding and variable-size window is its collection rate; for the same reasons mentioned previously, we chose a factor proportional to the scale of the study, setting the displacement of the window for each of the iterations to 50% of its dimension on the X- and Y-axes.

It is also necessary to always have a sufficiently representative region of the analysis window study area. Therefore, we have experimentally fixed a minimum size of the analysis window at 100 height samples. In addition, to avoid lengthy computation times in the case of surfaces with large sizes, the window value is set at 1/50 of the initial surface size. Furthermore, because the heights identified as outliers are transformed during the execution of the filter to nonmeasured points, the percentage of valid heights at each iteration should be great enough to establish a reasonable mean and standard deviation. This percentage was set at 95%.

Finally, in some cases, outliers may significantly change the value of the standard deviation, calculated for the considered analysis window. Thus, the criterion derived from Grubbs' test (see section 3.3.3) is applied iteratively on each subsurface until no outliers are identified. This step ensures that the standard deviations calculated before and after filtering do not differ.

Figure 9 shows a representation of a series of 120 east-west profiles taken on the surface Surf-1 before and after filtering. Figure 10 shows the results obtained by filtering outliers on the surface Surf-2. The various stages of implementing the filtering outliers are detailed in figure 11. The

total percentage of modified points on the samples is very low: 0.005 89% for the Surf-1 and 0.085 16% for Surf-2. However, the figures show that the outliers were filtered effectively by the proposed method. Furthermore, the removed heights have the characteristics of outliers on a surface, i.e. often isolated (peak shape) and clearly remote from the shape of the local neighbouring surface.

3.3.4. Proposed method for surface data measurements: summary. We sought to provide an appropriate response to the particular case of measured surfaces (multivariate data, many observations, nonnormal distributions and taking the scale into account in the identification of outliers) for the steps defined by the standard. Figure 11 summarizes the proposed method, incorporating the strategy from standard ISO-16269-4 (2010).

The filter uses form filtering by the DMD method to bring the distribution close to a normal distribution. Outliers are identified using a criterion based on Grubbs' test, the G_{lim} value of which varies, depending on the number of heights considered, applied to a variable analysis scale, which makes it particularly effective for this type of data. Finally, a user interface was developed to facilitate the use of the filter. This interface allows the results and indicators (distributions, Henry diagram, evolution of the percentage of modified points, total percentage of modified points, 3D representations of the surfaces at different stages of filtering) to be visualized at different stages of the process and allows us to save the resulting surfaces in the desired format, after identifying and excluding outliers. Outliers are replaced by nonmeasured points in the filtered surfaces. Subsequently, the nonmeasured points can be replaced or interpolated.

4. Conclusion

This study presents and demonstrates a method for the identification of outliers in measured surface data. The method is applied to two surface measurements, on which we have detailed the use of graphical and statistical indicators of the value and effectiveness of steps for this new method.

- (1) The method transforms data to approximate the normally distributed model used to assess whether an observation is an outlier. This step is performed by form filtering, using the modal method, which has the advantage of being relatively insensitive to outliers.
- (2) Outliers are identified using Grubbs' test, which allows the risk taken when identifying outliers in the data to be controlled. The application of this criterion is only supposed to be effective when considering a local outlier on a measured surface.
- (3) The identification is performed at different scales on the measured surface, using an analysis window of varying size that moves recursively over the entire measured surface.
- (4) Finally, this method fits into the overall framework of the methodology for detecting outliers, as presented in standard ISO-16269-4 (2010).

Acknowledgments

This research effort is supported by the research program INTERREG IV on the Surface Appearance of High Value-Added Products. This work has been performed in the Symme lab (Annecy) and in the Surface Metrology Lab at Worcester Polytechnic Institute (Worcester, MA, USA). We gratefully acknowledge the use of an Olympus LEXT OLS4000 confocal laser microscope through the generosity of *Industrial Microscopes, Olympus America Inc.*

References

- American National Standard 2009 *Surface Texture (Surface Roughness, Waviness and Lay)* ASME-B46.1
- Barnett V 1978 The study of outliers: purpose and model *Appl. Stat.* **27** 242–50
- Barnett V and Lewis T 1994 *Outliers in Statistical Data (Wiley Series in Probability & Statistics)* 3rd edn (New York: Wiley)
- Berglund J, Brown C A, Rosèn B G and Bay N 2010 Milled die steel surface roughness correlation with steel sheet friction *CIRP Ann. Manuf. Technol.* **59/1** 577–80
- Brown C A, Johnsen W A and Charles P D 1994 Method of quantifying the topographic structure of a surface *US Patent* 5307292
- Brown C A and Siegmans S 2001 Fundamental scales of adhesion and area-scale fractal analysis *Int. J. Mach. Tools Manuf.* **41** 1927–33
- Brown D R S and Brown C A 2010 Investigation of surface topography differences in native and exotic invertebrates in the St Lawrence River *Scanning* **32** 250–5
- Cantor G J and Brown C A 2009 Scale-based correlations of relative areas with fracture of chocolate *Wear* **266** 609–12
- Chauvenet W 1868 *A Treatise on the Method of Least Squares or, The Application of the Theory of Probabilities in the Combination of Observations* (Philadelphia, PA: Lippincott)
- Estler W T 1999 Measurement as inference: fundamental ideas *CIRP Ann. Manuf. Technol.* **48** 611–31
- Favrelière H 2009 Modal tolerancing : from metrology to specifications *PhD Thesis* Université de Savoie
- Favrelière H, Le Goïc G, Samper S and Belin P 2011 Flatness calibration of XY translation stages through modal parameterization *Proc. 15th Int. Congress of Metrology* pp 1–7
- Formosa F, Samper S and Perpoli I 2005 Modal expression of form defects *Proc. 9th CIRP Seminar on Models for Computer Aided Tolerancing* pp 1–9
- Gould B A 1855 On Peirce's criterion for the rejection of doubtful observations, with tables for facilitating its application *Astron. J.* **4** 81–7
- Grandjean J, Le Goïc G, Favrelière H, Ledoux Y, Samper S, Formosa F, Devun L and Gradel T 2012 Multi-scalar analysis of hip implant components using modal decomposition *Meas. Sci. Technol.* **23** 125702
- Grubbs F E 1950 Sample criteria for testing outlying observations *Ann. Math. Stat.* **21** 27–58
- Grubbs F E 1969 Procedures for detecting outlying observations in samples *Technometrics* **11** 1–21
- Grubbs F E and Beck G 1972 Extension of sample sizes and percentage points for significance tests of outlying observations *Am. Stat. Assoc. Am. Soc. Qual.* **14** 847–54
- Hotelling H 1931 The generalization of student's ratio *Ann. Math. Stat.* **2** 360–78
- Iglewicz B and Hoaglin D C 1993 How to detect and handle outliers *The ASQC Basic References in Quality Control* vol 16 (Milwaukee, WI: American Society for Quality Control)

- ISO-16269-4 2010 Statistical interpretation of data—Part 4: Detection and treatment of outliers (International Organization for Standardization)
- ISO-25178-3 2008 Geometrical product specifications (GPS)—Surface texture: Areal—Part 3: Specification operators (International Organization for Standardization)
- ISO-8785 1998 Geometrical Product Specification (GPS)—Surface Imperfections—Terms, definitions and parameters (International Organization for Standardization)
- Jordan S and Brown C 2006 Comparing texture characterization parameters on their ability to differentiate ground polyethylene ski bases *Wear* **261** 398–409
- Le Goïc G, Favrelière H, Samper S and Formosa F 2011 Multi scale modal decomposition of primary form, waviness and roughness of surfaces *Scanning* **33** 332–41
- Legendre A M 1805 Nouvelles méthodes pour la détermination des orbites des comètes *Firmin Didot, Librairie pour les Mathématique et les Editions stéréotypes*
- Leissa A 1969 Vibration of plates *NASA SP* vol 160 pp 1–316
- Paul S R and Fung K Y 1991 A generalized extreme studentized residual multiple-outlier-detection procedure in linear-regression *Technometrics* **33** 339–48
- Peirce B 1852 Criterion for the rejection of doubtful observations *Astron. J.* **2** 161–3
- Rosner B 1983 Percentage points for a generalized ESD many-outlier procedure *Technometrics* **25** 165–72
- Ross S M 2003 Peirce's criterion for the elimination of suspect experimental data *J. Eng. Technol.* **20** (2) 1–12
- Scott R S, Ungar P S, Bergstrom T S, Brown C A, Grine F E, Teaford M F and Walker A 2005 Dental microwear texture analysis shows within-species diet variability in fossil hominins *Nature* **436** 693–5
- Stefansky W 1972 Rejecting outliers in factorial designs *Technometrics* **14** 469–79
- Student 1908 The probable error of a mean *Biometrika* **6** 1–24
- Tietjen G L and Moore R H 1972 Some Grubbs-type statistics for the detection of several outliers *Technometrics* **14** 583–97
- Vessot K, Messier P and Brown C A 2012 Correlation between gloss reflectance and surface texture in photographic paper *Scanning* (submitted)
- Wilk M and Gnanadesikan R 1968 Probability plotting methods for the analysis of data *Biometrika* **55** 1–18
- Wright T W 1884 *A Treatise on the Adjustment of Observations: With Applications to Geodetic Work and Other Measures of Precision* (Ithaca, NY: Cornell University Library)
- Zienkiewicz O, Taylor R and Nithiarasu P 2002 *The Finite Element Method for Solid and Structural Mechanics* (Amsterdam: Elsevier)