A Novel Method for Automatic Detection of Incipient Micropitting in Ground Surfaces

W.M. Britton, Clarke A, H.P. Evans

PII: S0301-679X(21)00107-9

DOI: https://doi.org/10.1016/j.triboint.2021.106959

Reference: JTRI 106959

To appear in: *Tribology International*

Received Date: 6 November 2020

Revised Date: 12 February 2021

Accepted Date: 26 February 2021

Please cite this article as: Britton WM, A C, Evans HP, A Novel Method for Automatic Detection of Incipient Micropitting in Ground Surfaces, *Tribology International*, https://doi.org/10.1016/j.triboint.2021.106959.

William Britton: Methodology, Software, Validation, Formal Analysis, Investigation, Writing – original draft, Visualisation

Dr Alastair Clarke: Conceptualisation**,** Validation, Writing – Review and Editing, Supervision

Prof H P Evans: Conceptualisation, Validation, Writing – Review and Editing, Supervision

A Novel Method for Automatic Detection of Incipient Micropitting in Ground Surfaces

Britton W. M.[a*], Clarke A[a]., Evans H.P[a].

[a] School of Engineering, Cardiff University, Cardiff, United Kingdom

***Corresponding Author: W M Britton, brittonw@cardiff.ac.uk**

Keywords: Micropitting, Mixed Lubrication, Surface Roughness

Abstract

Micropitting is a surface fatigue phenomenon, on the scale of the surface roughness, which can be a significant problem in heavily loaded mixed lubrication contacts in power transmission gearing. Modern surface profilometry equipment allows detailed areal measurements of micropitted surfaces to be obtained. However, the processing of these surfaces to determine the extent of micropitting is difficult and often involves manual or semi-manual evaluation. This paper presents a new algorithm to allow the automated detection of micropits in areal surface measurements of ground surfaces, and its effectiveness is demonstrated by examination of measurements taken from a test where the early stages of micropitting developed. Comparison with previous methods for evaluating the amount of micropitting based on linear surface profiles clearly illustrates the truly three-dimensional nature of incipient micropitting.

Nomenclature

a	Hertzian half dimension of contact in the entrainment direction
b	Hertzian half dimension of contact perpendicular to entrainment direction
Δ_n	Data point spacing in areal scan in given direction
f	Analysis surface after high-pass filtration
h	Areal scan surface with asperities straightened
Ra	Mean absolute roughness of surface
s	Analysis surface expressed as mean of gradients at each point in f
SRR	Slide-Roll Ratio
u_f	Fast surface speed

u_s Slow surface speed

x Entrainment / circumferential direction

y Axial direction

1. Introduction

Micropitting is a form of rolling contact fatigue characterized by small pits of up to 10µm depth, most commonly found in gears with a ground surface finish. Micropitting can develop in any mixed Elastohydrodynamic Lubricated (EHL) contact, where the film thickness is on a similar scale to the surface roughness, but is prominently associated with the highly loaded and comparatively slow-moving gear systems found in applications such as wind turbine drivetrains [1]. Left unchecked, accumulation of micropits can result in significant stress concentrations and catastrophic failure.

Due to the significant industrial implications of micropitting a large amount of research is dedicated to better understand its mechanisms and influencing factors, and it is of fundamental importance to have a reliable means of quantifying the extent of micropitting on a surface. The standard methods for evaluating micropitting in gears (for example, using FZG test rigs) are based upon profile deviations from the involute caused by pitting, however this analysis method is applicable only once significant pitting has already taken place to cause gross profile changes, and is not applicable to cases of incipient micropitting at the asperity level.

When attempting to detect incipient micropitting in areal scans it can prove to be a complicated process to distinguish micropits from valley features remaining from the manufacturing process or occurring due to other forms of wear. Many different approaches to this problem can be found in the literature.

An ideal approach is to develop a method that identifies micropits in a surface roughness measurement or image and designate locations as either pitted or not pitted. Several researchers have attempted to achieve this by simple height thresholding methods [2,3], however these methodologies have no protection against falsely detecting features from other wear mechanisms or valley features as micropits. Prajapati and Tiwari [4] also used a thresholding method but required the feature to have a diameter of 10-50µm, which can help to reduce these false detections. It does however assume that micropits are circular which is frequently not the case. Li and Kahraman [5] identified micropit features by eye under close microscope inspection, an accurate but prohibitively time-consuming method for all

but very small areas. Oila and Bull [6] used a combination of surface measurements and image processing software to process very small areal scans produced by an optical profiler. In many cases the percentage of the surface that is micropitted is recorded, but the method by which micropitted area is assessed is not disclosed [7–10].

Often it is not feasible to obtain an areal scan of the surface mid-test. Hutt et al [11] devised a methodology for realigning two-dimensional profiles collected in-situ between stages of twin disk tests using a portable profilometer. A cross-correlation method was used to align the profiles circumferentially, and valley features were used for vertical alignment. Micropits were detected as locations of high variation in surface height through the course of the test.

Some alternative methodologies have attempted to gauge the severity of micropitting by means other than micropit detection. One common methodology is to use the deviation from the original form of the test specimen – for example the decrease in roller diameter [12,13] or change in tooth involute profile[14]. This has the clear limitation of an inability to verify that all loss of diameter is due to micropitting. Even in that case, this is of limited effectiveness for mild or early stage micropitting where gross alterations to form have not yet occurred. Additionally, Kattelus et al [15] devised a means of monitoring micropitting progression through monitoring of debris in the oil.

This paper presents a new methodology for the evaluation of micropitting in surfaces. A process for obtaining replica surfaces is outlined and a new algorithm for evaluating micropitting in each surface based on detailed measurements of the replicas is devised. The effectiveness of the algorithm is demonstrated in evaluating the results of a test where the mild micropitting occurs.

2. Twin-Disk Test Rig

The experimental work was carried out using a twin-disk test rig, a means of simulating gear contact conditions without the complex kinematics of true gears. The twin-disk test rig serves as a test-bed for incipient micropitting, where the micropits are in their earliest stages affecting single asperities and have not yet induced sufficient material removal to cause deviation from the manufactured profile. The test rig in question has previously been used for investigations of acoustic emission (AE) during running in and fatigue [11], contact voltage (CV) across lubricant films [16], and the influence of material hardness on running in and micropitting fatigue [17], and the reader is referred to these papers for full details of the rig. A brief description is presented here for completeness.

A schematic of the test head is shown in in Figure 1. The fast shaft is driven by a three-phase electric motor, which is geared to the fast shaft via a speed increasing gear pair. The slow shaft is geared to the fast shaft by a pair of speed decreasing gears which impose a Slide Roll Ratio (SRR) in the contact of 0.50 calculated as shown in Equation 1, where u_f and u_s are the surface speeds of the fast and slow disks respectively.

$$SRR = \frac{2|u_f - u_s|}{u_f + u_s} \qquad \text{Equation 1}$$

Each shaft is supported by two bearings, and an arrangement of flexible couplings ensures that the shafts remain parallel. The bearings of the slow shaft are mounted in a pivoting housing which enables the centres distance between the two disks to vary as required and allows the disks to reach the required speed out of contact prior to loading. Load is applied by a hydraulic ram which is manually engaged and released by operation of a ball valve. The applied load is measured by a load cell, and at the start of each test stage the load was gradually ramped to the target value over the course of thirty seconds to minimise shock loading effects.

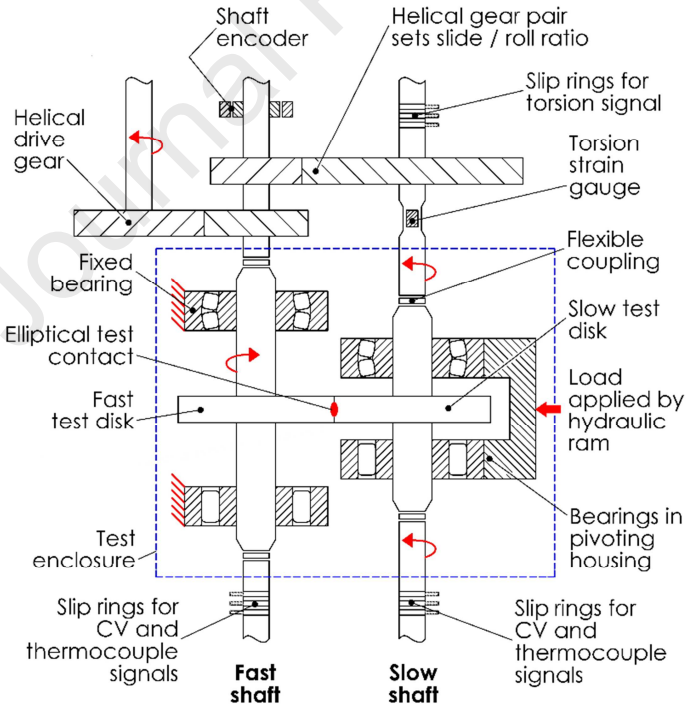

Figure 1 Schematic of the test head of the twin disk rig. figure adapted from [11].

The test disks used in this work were made of EN36 (655M13) steel with a hardness of 714Hv. Each disk had a diameter of 76.20mm and a crown of radius of 304.8mm, resulting in an elliptical contact with a

4:1 aspect ratio. The disks had an approximately-transversely ground finish produced using the arrangement in Figure 2. In this arrangement the disk is ground using an internally conical grinding wheel, a process which also imparts a 305 mm crown radius onto the disk. This process is explained in greater detail in reference [18]. Two Vickers indents were then made in the disk surface outside of the contact area to serve as reference points.

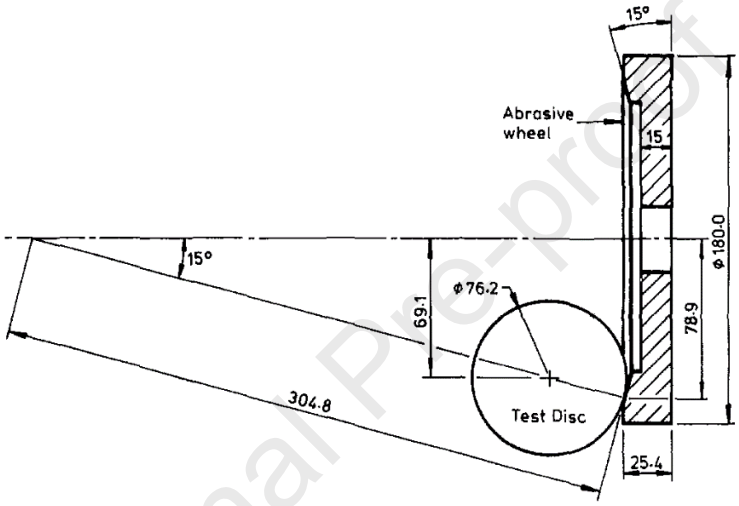

Figure 2 Schematic of the transverse grinding setup used to produce the crown and surface finish on the disks.

The test conditions are shown in Table 1. The test was interrupted between each load stage to allow the disk surface to be measured and replicas taken. The disks remained in-situ in the test rig during measurement and replication, to ensure that the detailed alignment between the surface roughness features on each disk was maintained. Load stages are shown in Table 2.

Table 1 – Conditions for micropitting test.

Max. Hertzian Pressure / GPa	1.20
Entrainment Velocity / ms^{-1}	2.00
Slide Roll Ratio	0.50
Fast Disk circumferential Ra / µm	0.41
Slow Disk circumferential Ra / µm	0.41
Oil Supply Temperature / °C	80

Hertzian Half Dimension a / mm	0.385
Hertzian Half Dimension b / mm	1.506
Mean Specific Film Thickness	0.21

Table 2 Load stages for micropitting test.

Load Stage	Fast disk Cycles this stage	Cumulative Fast Disk Cycles	Cumulative Slow Disk Cycles
0	0	0	0
1	3×10^3	3×10^3	1.8×10^3
2	3×10^3	6×10^3	3.6×10^3
3	14×10^3	20×10^3	12×10^3
4	30×10^3	50×10^3	30×10^3
5	50×10^3	100×10^3	60×10^3
6	100×10^3	200×10^3	120×10^3
7	100×10^3	300×10^3	180×10^3
8	100×10^3	400×10^3	240×10^3
9	200×10^3	600×10^3	360×10^3
10	400×10^3	1000×10^3	600×10^3
11	500×10^3	1500×10^3	900×10^3
12	500×10^3	2000×10^3	1200×10^3

3. **Replica Collection**

The twin-disk rig used in this work has the capability of taking two-dimensional line measurements of surface roughness in-situ, using a portable surface profilometer. This process is described in detail in [19]. However, it has no capability for collecting areal scans of the disks in-situ, and removal of the disks from the rig would be both prohibitively time consuming and would result in loss of alignment of the disks. Therefore, to enable areal scans to be collected between test stages, surface replicas were used.

To ensure uniform replicas, the components shown in Figure 3 were made for each disk. These were clamped to the edge of the disk using bolts, and acrylic endplates were placed in slots S1 and S2. The

areas behind each endplate and beneath each archway were packed with plasticine to prevent any leakage of the replica material prior to setting. The sides of the archway components in contact with the replica material are slightly tapered to facilitate clean removal of the replicas from the disks after setting. This creates a lip on the underside of each arch which makes contact with the top of the disk surface. The replica material used was Acrulite® Microtech Type A, supplied by Rubert & Co. A ratio of two parts liquid to five parts powder was used with a minimum curing time of one hour. This ratio was found to give best results within the wider range of acceptable ratios in the supplier's instructions.

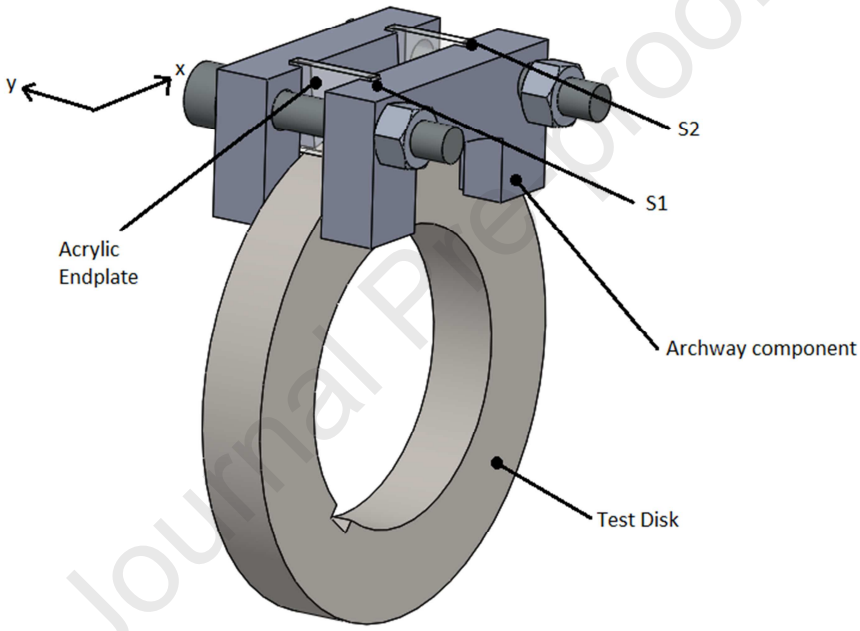

Figure 3 A CAD image of the assembly attached to a test disk.

Areal scans of replica surfaces were made using a Taylor Hobson Form Talysurf Series 2 with an automated stage. As these scans were of long duration, surfaces from each disk were scanned for load stages 0, 2, 4, 5, 6, 8, 10 and 12. The area of each scanned surface was 6mm long in the circumferential (x) direction (perpendicular to the lay of the roughness) and at least 7mm in the axial (y) direction such that the entire width of the contact area was captured, and that the hardness indentations in the unrun area were included. Data points in the x-direction were recorded at 0.5µm intervals, and the spacing between each profile in the y-direction was 3µm. The larger spacing in the y-direction was adopted to

reduce the profilometer scan time and ensure that surface scans are of a size that is practical to analyse. Due to the comparatively slow changing, extruded appearance of the asperities and valleys along the y-direction in the as-manufactured state and the fact that micropit features are an order of magnitude larger than the adopted measurement spacing in the y-direction, an appropriate resolution of the data was still maintained.

4. Micropit Detection

Once scanned, the surfaces were loaded into TalyMap Platinum 7.1. The line by line levelling operator was applied to subtract the least squares second order polynomial from each profile in the surface, and a gaussian filter was then applied to the surface with a cut off wavelength of 0.8mm. Replica surfaces were then inverted about the z (height) axis, and the surface heights, *h*, were exported as ASCII text files.

The method developed to detect micropits in the surface is reliant on two key observations:

1. The asperities in the ground surface surfaces are extruded along shallow arcs in the y-direction as a result of the grinding process and any changes along an asperity ridge are gradual in comparison with changes in the x-direction roughness profiles.
2. The sides of micropits are the steepest features found on the surface.

In order to exploit the first of these observations it must be possible to conduct analyses along the length of an asperity ridge, but as can be seen in Figure 4a the asperity ridges in the disk surface do not run perfectly straight across the disk width, but have a “swept” appearance. This is a feature of the grinding process illustrated in Figure 2. A cross-correlation procedure was used to “straighten” the asperities in the surface. Beginning at y=0 a standard cross correlation process was performed between each profile and its neighbour, where the current profile was held stationary and its neighbour was offset in data point increments. The maximum value of cross correlation indicates the offset in the x-direction that must be applied to the neighbouring profile to achieve the closest alignment between the profiles. This process is illustrated in Figure 4, where an offset of 8.5µm is required to achieve the best alignment between the profiles. The profiles chosen for this illustration were a distance of 102µm apart in the y direction, so that a clear offset was obtained for illustrative purposes.

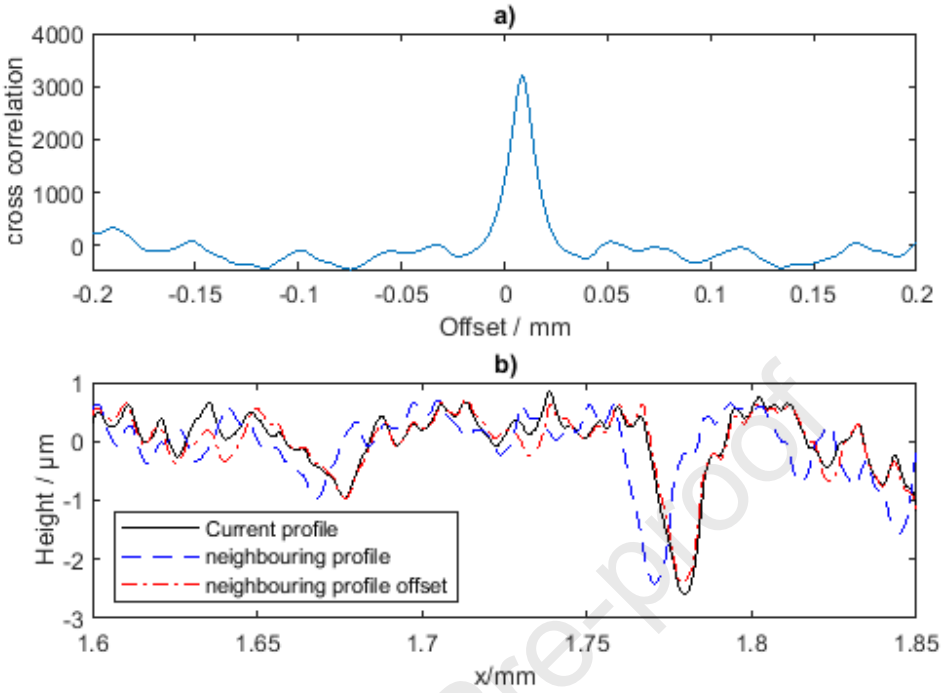

Figure 4 a) The cross correlation at different offsets of the neighbouring profile; b) The current and neighbouring profiles before and after applying the offset

The lag required to offset each profile to the location of maximum cross-correlation with its neighbour was recorded, and applying the lag cumulatively results in a straightened surface as shown in Figure 5. A small loss of area is incurred as partial asperities and valleys (i.e. those which do not span the entire range of y values) are discarded. A typical 6 x 7mm surface scan lost less than 0.3mm in x-direction profile length – although scans with a larger y axis range (such as the full disk width scan below) lost up to 0.5mm due to the increasing sweep angle. The two hardness marks can also be seen in the lower part of the area shown in Figure 5.

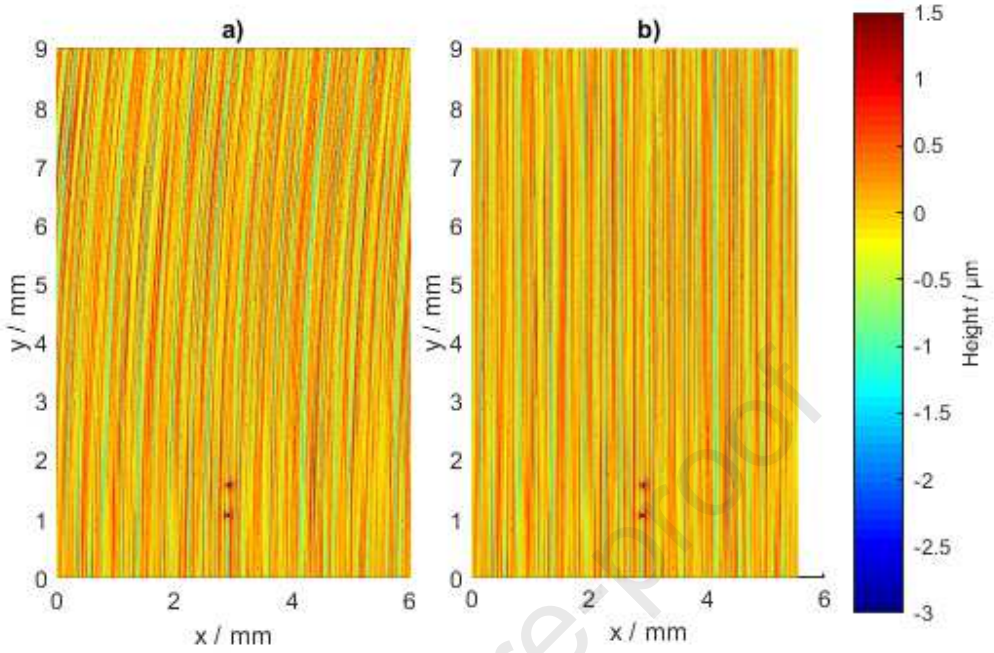

Figure 5 The same unrun surface before a) and after straightening b). The area of the surface is slightly reduced during straightening as a function of the lag.

Following straightening, the directions x and y were taken as for the straightened surfaces in which asperity ridges are parallel to the y axis. Initial attempts were made to detect micropits by both height thresholds and gradients after straightening, but it was found difficult to distinguish between pits and valley features remaining in the surface. To overcome this a filter was applied along each line of constant x (each asperity or valley) in the straightened surface to leave only the very short wavelength changes in the surface due to wear or fatigue processes, while removing valley and asperity features resulting from the manufacturing process. A zero-phase digital filter was applied using the "filtfilt" function in MATLAB to ensure the locations of features remained unchanged. The filtered surface was subtracted from the unfiltered surface to leave only short wavelength features, effectively a high-pass filter. As the largest micropits in these surfaces have diameters of approximately 50µm a cutoff wavelength of 60µm was used in the work reported here to ensure all micropit features would be retained. The heights of this filtered surface are denoted *f*.

The steep gradients at the walls of micropits were used to detect micropit boundaries in the filtered surface. For each internal point in the filtered surface the mean absolute slope to the point from its eight neighbours was calculated according to Equation 2

$$s = \sum_{neighbours} \frac{|f - f_n|}{8\Delta_n} \qquad \text{Equation 2}$$

where f_n is a neighbouring point f value and Δ_n is the separation distance, Δ_x, Δ_y or $\sqrt{\Delta_x^2 + \Delta_y^2}$. Utilising the absolute gradient negates any vertical offsets introduced by the high-pass filter. The use of the absolute height difference reflects the fact that the direction of the gradient relative to the current point is unimportant, and it is simply the magnitude of the gradient that is required to detect micropit boundaries.

Examples of pits are illustrated in Figure 6, where the straightened original surface (*h*), the high-pass filtered surface (*f*) and the mean slope (*s*) are shown alongside corresponding profiles in the y-direction.

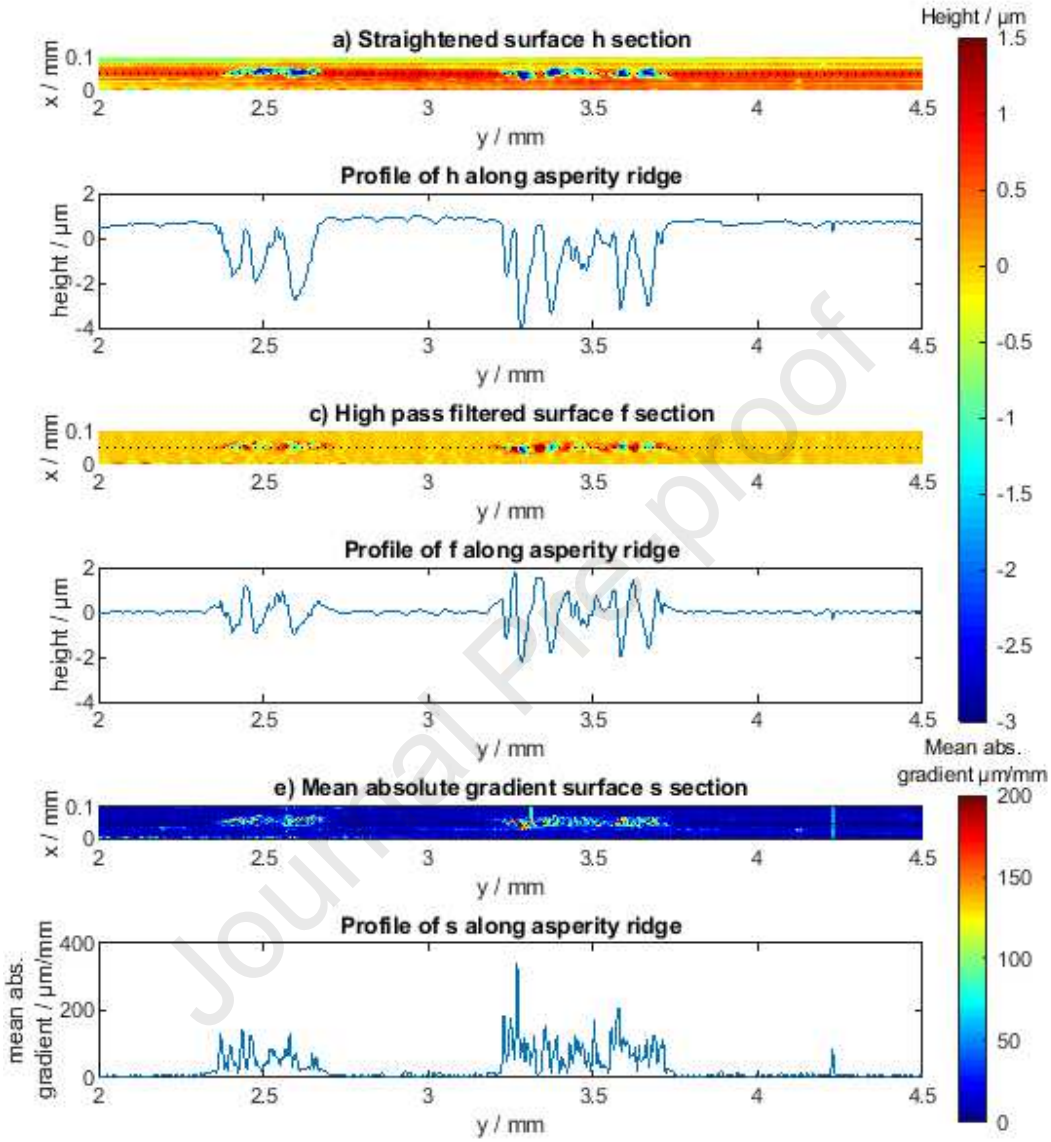

Figure 6 Straightened surface sections and corresponding profiles for surfaces h (a,b), f (c,d) and s (e,f). The profile sections are taken in the y direction (axially), and correspond to the dotted line on each surface section.

Each point of the surface is considered to be part of a pitted region if its *s* value exceeds a specified threshold value, and its *h* value is less than zero. Initially a detection threshold for surface *s* was selected manually by systematic variation of the threshold and comparing the results obtained to manual evaluation of the surfaces. It was found that for the range of surfaces and conditions tested (initial

Ra=0.36 - 0.46μm), the optimal thresholds were consistently close to three times the median *s* for the surface. Typical median values of *s* for the surfaces considered were between 30-45 μm/mm, but for some examples values of up to 70 μm/mm were obtained). Using this definition for the threshold allowed the detection algorithm to naturally adapt to the surfaces measured without additional user input. This threshold is not claimed to be universally applicable and may require further adaptation when a wider range of roughnesses and conditions are considered. The points identified in this way were at, or close to, the boundaries of the micropits, but did not necessarily form closed curves.

The actual pitted areas were then determined by a pit filling process. This considered each point that had a pitted neighbour and added it to the pitted region if its *f* value was less than a set threshold value. This step was applied by sweeping through each surface point repeatedly until there were no further additions to the pitted area. In order to set a consistent threshold that scales with surface roughness, the threshold was determined as the mean absolute height of the surface in the contact area multiplied by a coefficient A. A secondary pit filling process was then applied using the same approach for points neighbouring the pit where the height *h* was less than the mean absolute height multiplied by a coefficient B. For the secondary pit filling stage, only neighbours in the *x* direction were considered, to safeguard against filling of valley points.

Values for coefficients A and B were determined by systematically varying these coefficients to achieve the most complete pit filling without excessive false detection when tested across a large number of surfaces. False and under-detection were identified by manual inspection of a number of analysed surfaces. The values of A and B used in this work were -0.28 and -2.8 respectively. It may be noted that at some points inside pits, *f* becomes positive which would prevent detection at the pit filling stage. However, positive points in *f* tend to occur at pit edges which are detected through their high gradients in surface *s*. Table 3 shows a summary of each stage.

Table 3: Conditions for each stage of micropit detection

	Surface *s* detection	Surface *f* pit-filling	Surface *h* pit filling
Pitted neighbour required	No	Yes – any neighbour	Yes – x direction neighbours only
Criteria	$s \geq 3 \times median(s)$ $h < 0$	$f < -0.28Ra$	$h < -2.8Ra$

An example of the method is shown in Figure 7 which illustrates the process on a small section of the surface. The initial surface gradient step tends to detect the outer edges of pits only, effectively identifying the perimeters of pits as can be seen in Figure 6b. This is a perfectly logical outcome, as the gradients will lessen and reach a turning point in the centre of the pits. Almost all of the pits are then successfully filled by the first pit-filling stage, as the pits are small enough for their centres to be strongly negative throughout in surface *f.* A very small number of larger micropits are sufficiently long and flat-bottomed in the axial direction for their centres not to be captured in surface *f*, and hence are missed by the first pit-filling stage. As these micropits are invariably among the deepest in the surface the second pit-filling stage based on surface *h* depth alone fills these effectively. This second filling stage often applies to only four or five pits, and hence only contributes a comparatively small number of detections to the evaluation. Figure 6c shows the detected pits after both pit-filling processes have been applied.

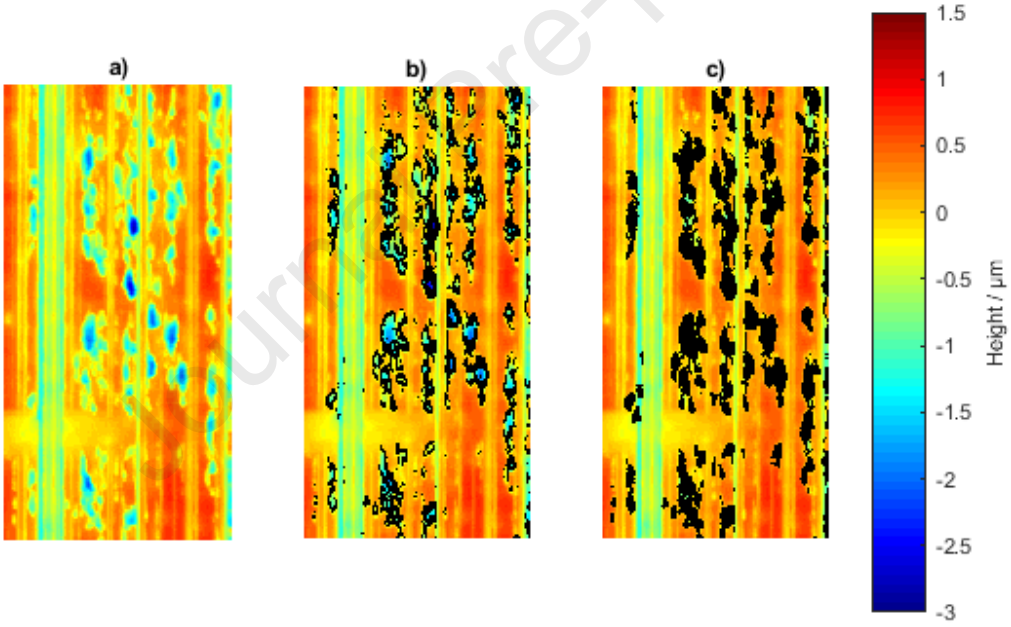

Figure 7 a) A 0.3mm x 0.6mm section of straightened surface; b) Gradient-based detection only overlaid in black; c) Gradient-based detection and pit-filling overlaid in black.

As this point, any score marks resulting from debris entrained through the contact may have been detected. These can be easily discounted as the length of score marks in the entrainment direction is much greater than the length of micropits, while their width perpendicular to the entrainment direction is much less than the width of micropits. These features are small in number and identified manually.

Knowing the lag that was applied to the original surface during the straightening operation, it is then possible to overlay the binary matrix of pitted locations over the original surface. The percentage of the surface that is micropitted can be evaluated over the area to which the detection was applied.

5. Two-dimensional pit detection

To evaluate the effectiveness of the surface-based micropit detection method, it was compared to a two-dimensional analysis building upon the method developed by Hutt *et al.* [11]. Two-dimensional profiles of the disk surfaces were made at nine axial locations on each disk between every load stage using a Taylor Hobson Intra 2 portable profilometer mounted on a manually controlled stage. Measuring positions were axially relocated at each load stage by the following procedure:

- The stylus was placed in contact with the disk at the approximate centre by eye
- The profilometer stage was moved axially towards the slip rings by use of a lead screw until the profilometer reading decreased by 60µm, indicating the edge of the disk. As this work considers early stage micropitting, no significant geometry loss occurs during testing to influence this relocation stage
- The position gauge on the stage was zeroed at this location
- The required offset was applied to the stage for each measuring position

Measurements were made at the centreline and four increments of $\frac{2}{9}b$ each side, as illustrated in Figure 8. Approximate circumferential alignment was achieved using a mark on the side of each disk. Precise circumferential and vertical profile alignment between load stages were performed using cross correlation and reference heights of unchanged valley features respectively. All two-dimensional scans were 12mm in length and filtered using a 0.8mm gaussian filter to remove the form and waviness features.

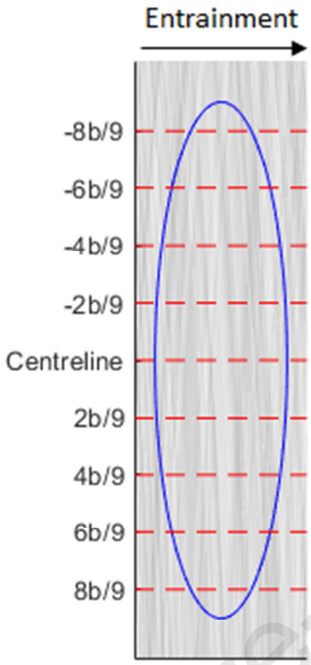

Figure 8 Scale illustration of the contact ellipse with measuring positions for two dimensional profiles indicated.

In order to obtain values for percentage of points pitted at each stage, and to reduce biasing towards locations that pit early, a windowing method was used. Beginning with the fully run in profiles (Load stage 2 onwards) a window of five load stages was considered at a time. The standard deviation was calculated for the five height values at each relocated data point and locations with a high standard deviation were declared pitted. The profile for the earliest load stage in the window was then discarded and the profile from the next load stage in time was added to the analysis window. Pitting was said to occur during the final load stage in a window where a pit was detected. Where a pit was detected during the first window of load stages, the height history was examined to locate the earliest significant height decrease. Figure 9 shows an example section of relocated profiles for the test.

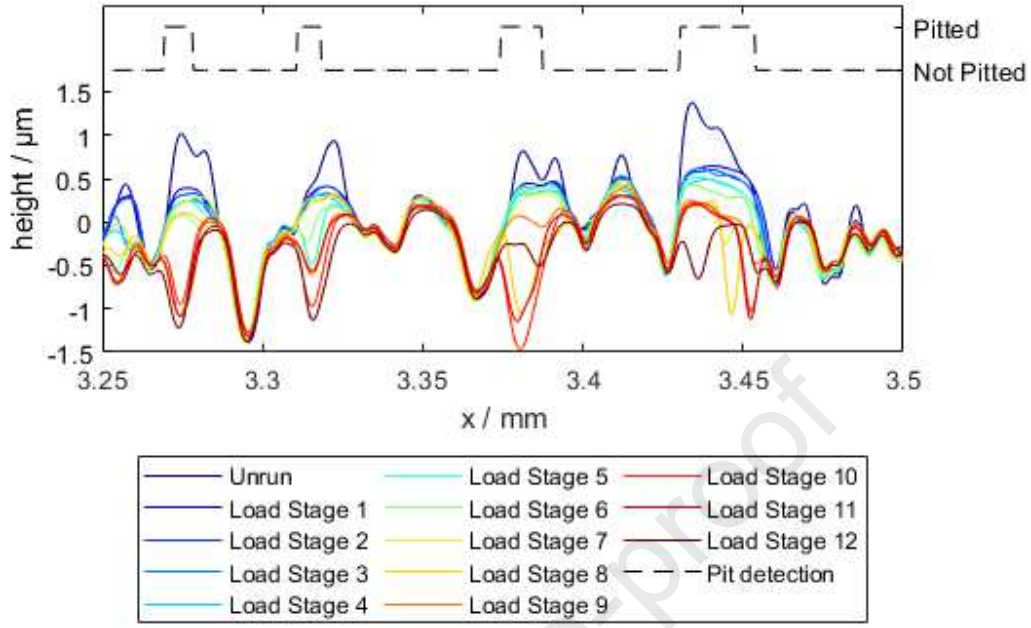

Figure 9 An example section of the realigned fast disk centreline profile and the pits detected by the 2D pitting analysis.

6. Application to a micropitting test

In this section the surface analysis is applied to a micropitting test, following the experimental procedure already described, and the results are compared to those found when using the two-dimensional analysis. The effectiveness of the surface analysis method is discussed.

Figure 10 shows a 1mm^2 area on each disk after test stages 5, 8 and 12, measured using a replica. It is clear from these surfaces that micropitting fatigue was firmly established on the surface prior to load stage 5, however the rate of increase in micropitted area between load stages 5 and 8 was much reduced in comparison to earlier stages of the test. The sections of surface at the end of testing show that asperities on the same surface could exhibit markedly different micropitting behaviour. In the fast surface image e) particularly, very heavily micropitted asperities are neighboured by asperities which show no micropitting at all in that region. Both fast and slow surfaces show areas along an asperity which remain non-micropitted while significant portions of that same asperity exhibit severe micropitting.

On both fast and slow surfaces a handful of spurious points can be seen to come and go from the detected regions as the test progresses. This arises due to the replica method, which makes a largely

accurate but imperfect model of the surface, with subtle differences in the way pits are captured by successive replications. The vast majority of micropits present on the surface are detected successfully and the growth and development of initiated pits can be clearly followed through each subsequent stage.

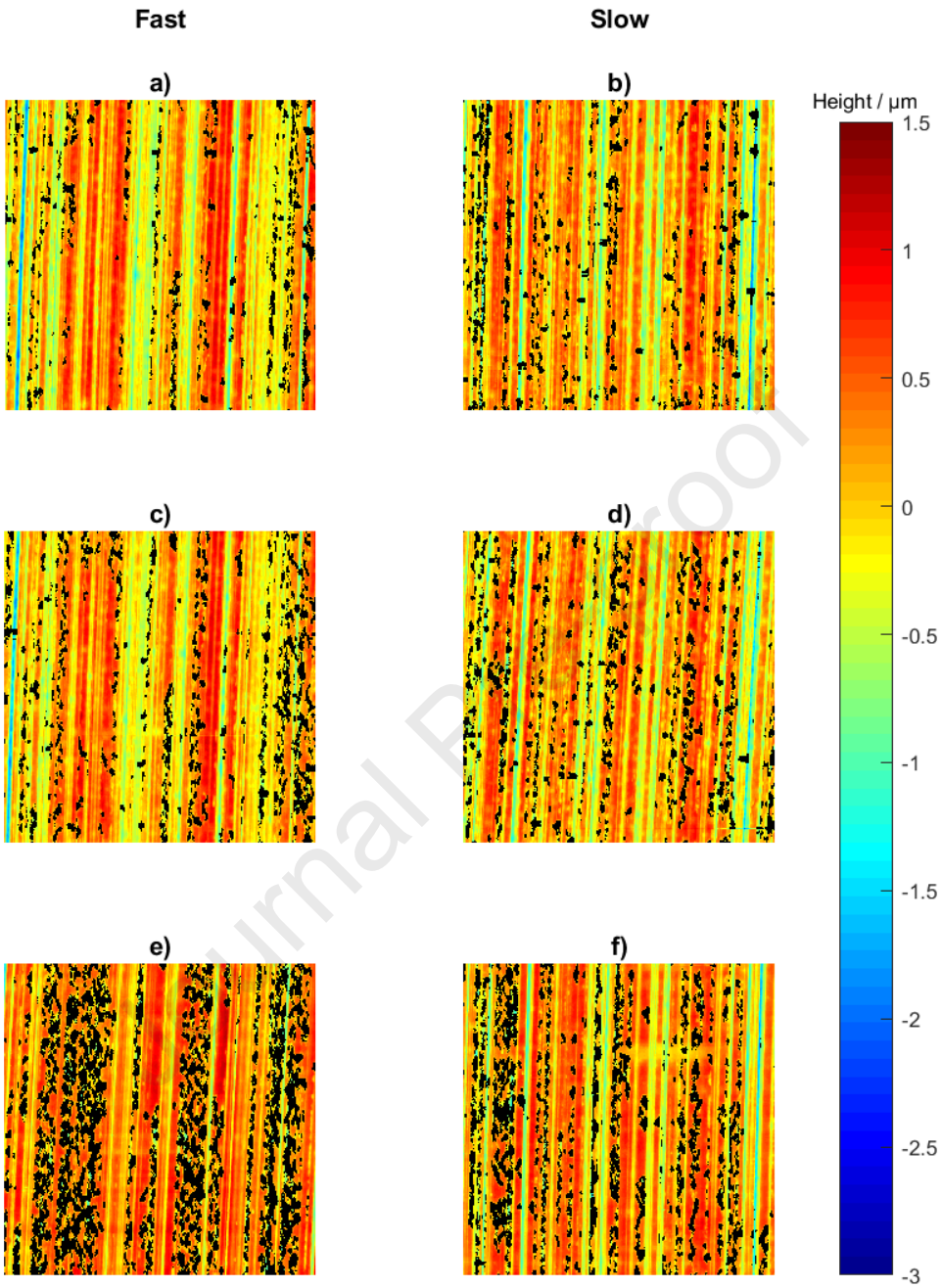

Figure 10 Matching 1mm x 1mm sections of the fast (a,c,e) and slow (b,d,f) surfaces at load stages 5, 8 and 12 (top to bottom). Detected pits are overlaid in black. Heights are relative to the mean height of the surface, which decreases as more pits develop.

Looking at a small section of an asperity ridge that did pit, as in Figure 11, it is possible to follow the development of individual micropits or groups of micropits through the test. In Figure 11a the asperity ridge is largely intact except for two very small, shallow micropits in their earliest stages in the top left of the figure, which have been successfully detected by the algorithm. These have grown and slightly changed shape in Figure 11b, although the shape may be influenced by axial offsets of up to 1.5µm in the traverse of the stylus. Some additional features of approximately -0.2µm depth can be seen but are undetected - at this point these features are not sufficiently developed to be detected by the algorithm and likely indicate material lost at the mouth of a surface crack. In Figure 11c the micropits have expanded in size, and the small surface features have developed into full micropits, which continue to grow and amalgamate with existing and new pits until the end of the test as seen in Figure 11f.

Factors such as the changing mean height of the surface influence the apparent depth of features in each scan. The figure clearly illustrates that the detection algorithm is able to adjust well to this change and maintain effective detection despite this variation.

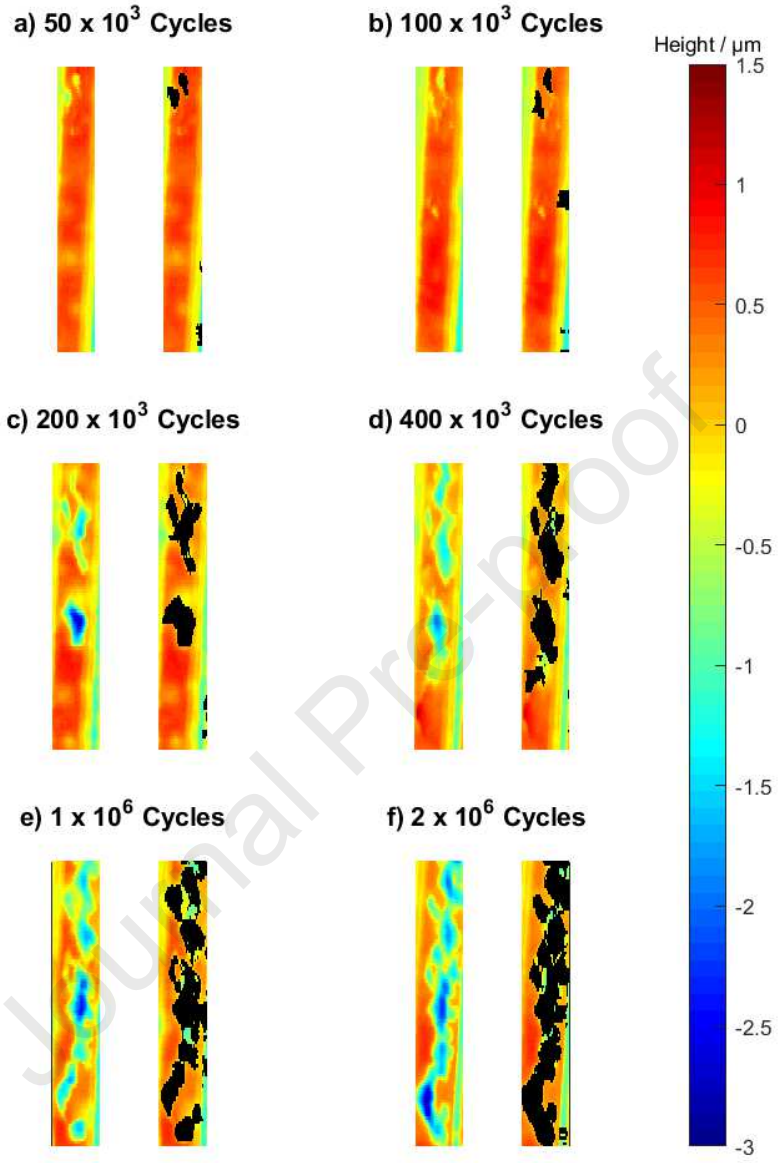

Figure 11 40µm x 300µm sections of an asperity ridge shown at different stages in the test, both as scanned (Left) and with black overlay (right) indicating detected micropits. This illustrates the initial formation of micropits followed by their growth and amalgamation into larger networks of connected micropits.

To accurately evaluate the pitted area on the surface, the relationship between Acrulite replicas and the surfaces from which they are cast must be considered. Areal scans of disks could not be obtained in-test, however areal scans of both replicas and disks were made for comparison at the end of testing. Scans of

the replicas were comparable to the disk surface scans, with the shapes and locations of surface features accurately captured. The height distribution in the Acrulite replica differs from the disk however, with an apparent stretching of negative values. This is assumed to be the result of either a non-uniform distortion of the replica (which is an inverse model of the surface) or stretching of valley and pit features as the replica is removed from the disk surface after curing.

Unsurprisingly, there was a consequent effect on the detection of pitted points using the algorithm described in this paper. Testing of twelve pairs of corresponding disks and replicas at the end of tests found that the same pits were captured in all cases, but that detected pit features on replica surfaces were artificially enlarged as a result of the replication method. An example of this can be seen in Figure 12. The extent to which the detected pitted area was enlarged was consistent however, with the true disk value ranging from 70-76% of the value determined for the replica across twelve surfaces tested, where micropitted area ranged from 3.75% to 16.2% micropitted. The values obtained for each of the surfaces were plotted on a graph of percentage pitted evaluated on the replica surface versus the value found for the corresponding disk surface. A least squares line passing through the origin was fitted to the data. The resulting relationship showed that the disk value was 72.3% of the Acrulite value with R^2=0.9985. This can be seen in Figure 13.

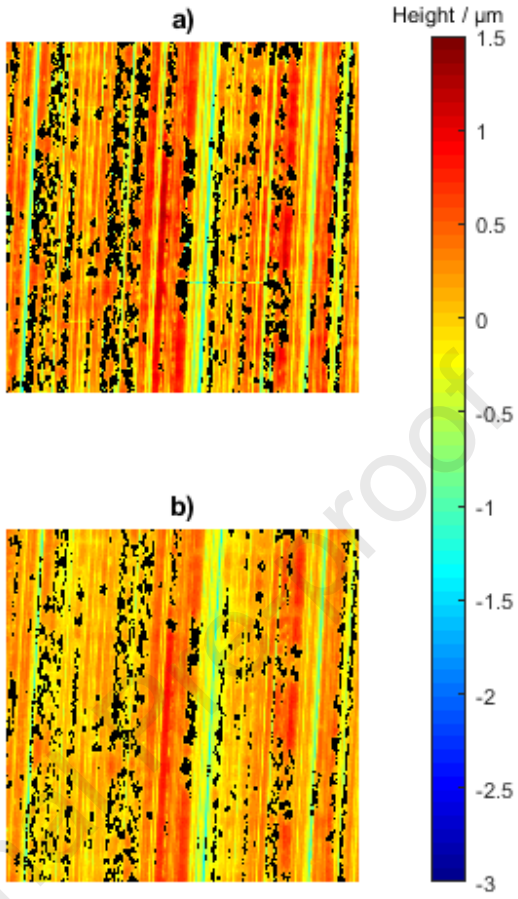

Figure 12 Micropit detection on matching 1mm x 1mm sections of the a) replica and b) disk for the fast surface at the end of test. Detected micropits are overlaid In black.

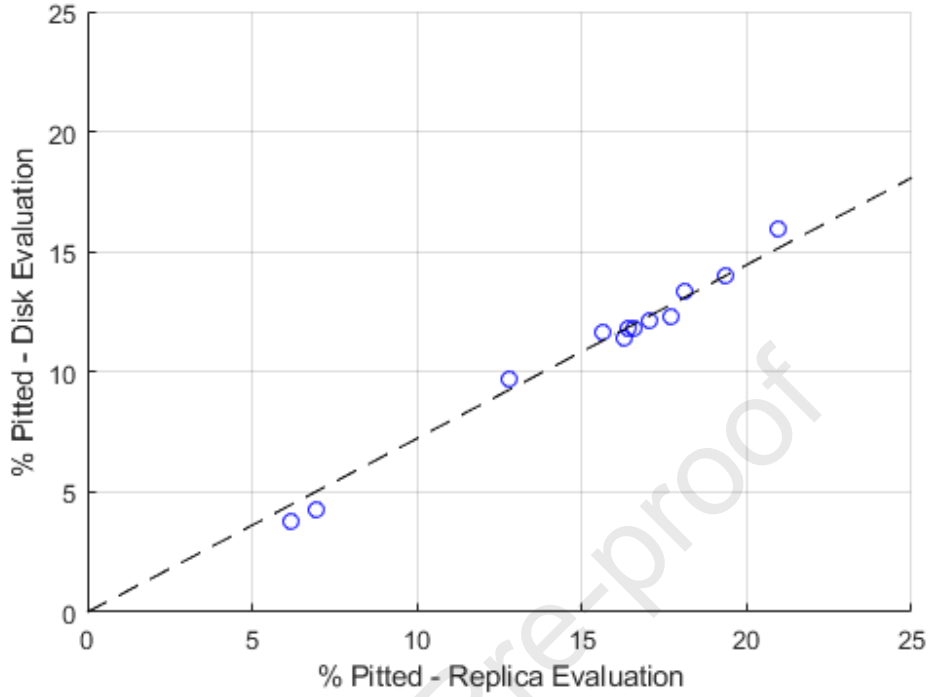

Figure 13 Percentage pitted evaluated on corresponding pairs of disks and replicas

Figure 14 shows the percentage of surface micropitted as evaluated by both the surface analysis and two-dimensional profile analysis. A rapid onset of micropitting was seen for both disks, each exceeding 5% micropitted surface within one hundred thousand fast disk cycles as evaluated by the surface analysis. In both cases the two-dimensional analysis detected less micropitting in the earliest stages of the test. For both fast and slow surfaces both analysis methods determined an approximately equal extent of pitting at the end of load stage 6. The two-dimensional analysis then went on to exceed the percentage pitted value determined by the areal analysis significantly. At the end of the test the two-dimensional method calculated 20.8% surface micropitted for the fast surface, compared to 11.9% for the areal surface analysis. For the slow surface, the two-dimensional method calculated 17.5% micropitted and the areal method 11.8% micropitted. This corresponds to the two-dimensional analysis overestimating the amount of micropitting by a factor of 1.7 for the fast disk and 1.5 for the slow disk.

Differences between the analyses in the earlier stages of the test are most likely the result of underestimation on the part of the two-dimensional analysis, and this can be argued as follows. While it can be guaranteed that the areal scan will capture all micropits in the contact area due to the very small x and y spacing of measured points in comparison to pit size, the same cannot be said for the two-dimensional profiles measured from the disks in-situ which are 335 µm apart in the y direction. On any

asperity over which a two-dimensional trace passes there may either be a pit present in the region represented by that trace, or no pits at all. In the case where a pit is present the pit may be missed entirely by falling between traces, or the trace may pass through the pit. Due to axial realignment uncertainty the trace may or may not pass through the pit at any of the measurement stages. Realignment error is caused by inconsistency in locating the disk edge between load stages and traverse errors in reaching the profile position from the establish edge position. This error may be up to ±80µm from the nominal measuring position.

Where there are fewer, and smaller, micropits on a surface the likelihood of the two-dimensional profile passing through a micropit is reduced. This is mitigated somewhat by the fact that the calculated percentage of surface pitted for each detection is overestimated, as each detected location effectively represents a 335µm wide pit – far in excess of the true width of a pit, as evidenced in Figure 11. In addition to the appearance of new pits, previously missed pits can also be added to the results with each stage.

In the later stages of the test the two-dimensional analysis tends to exaggerate the extent of micropitting. This is due to a tendency for the two-dimensional analysis to exaggerate the extent of micropitting as the surface history increases, arising from a combination of the axial relocation error inherent in the method and the large profile spacing. As the number of micropits increases the probability of a profile passing through a micropit also increases, and the relocation error causes slightly different locations to be measured each time. The analysis technique dictates that any previously detected micropitted point must necessarily remain as such. Referring back to Figure 10, it is clear that there are sections of asperities that are not micropitted, and that even on densely pitted asperities there are numerous gaps between micropits where the original asperity remains. Only one micropit on the same asperity need be encountered by a profile measurement for the two-dimensional method to classify a 335µm length of asperity as irrevocably and completely micropitted.

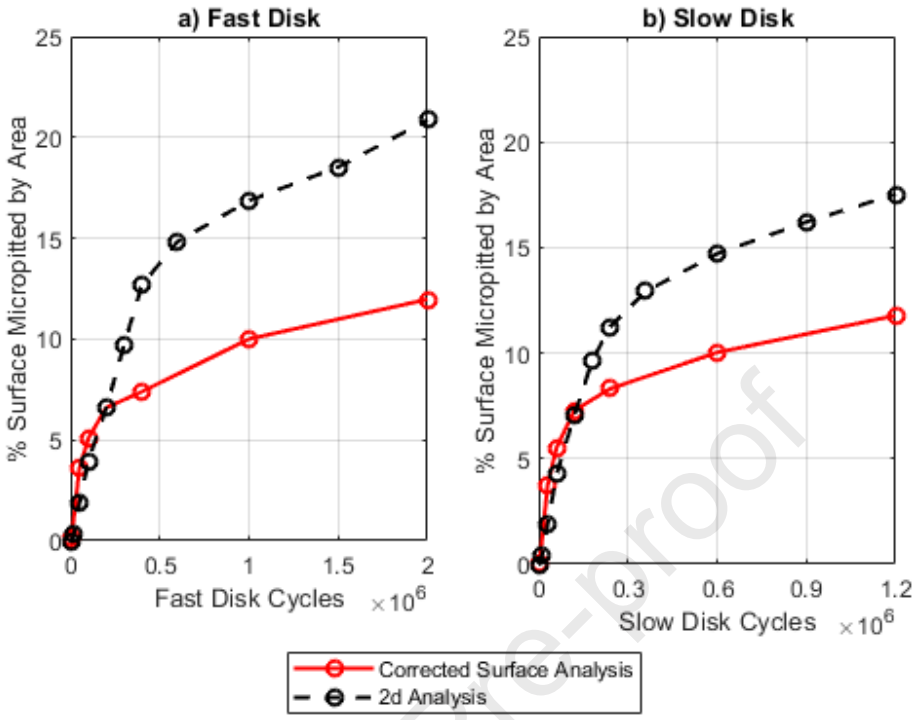

Figure 14 Progression of micropitting on both disks as evaluated for both the two-dimensional and surface analyses.

Additionally, the rate at which pitting progresses during any test stage could be assessed. The difference between the percentage of points pitted at the start and end of a test stage were divided by the number of loading cycles comprising that stage to give the average increase in percentage of points pitted for each cycle through the contact. This analysis can be seen in Figure 15, where the average rate of pitting is plotted at the number of cycles corresponding to the mid-point between two measurements.

On both disks the highest rate of pitting occurred in the early stages of the test before decreasing to a lower rate of pitting that was maintained until the end of test. The peak rate of pitting in the two-dimensional analysis was attributed to load stage 4, which is in agreement with the areal analysis for which the coarser sampling intervals attribute the peak to load stages 3 and 4. In both analyses the peak rate of pitting was greater on the slow disk.

In the areal analysis the pitting rate on the fast disk decayed more rapidly than on the slow disk, but this effect is much less pronounced in the two-dimensional analysis. While the surface analysis showed a rapid decay from the peak value, the two-dimensional analysis for both disks showed a second peak occurring across load stages 7 and 8. The most likely cause of this is the relocation error and small sample size of the two-dimensional analysis. As the relocation error causes the profilometer trace to

pass through pits which were formed in previous stages, these pits are then attributed to the current stage – effectively resulting in a time-lag in the detection of these pits. This also accounts for the much lower peak rates of pitting in the two-dimensional analysis, as the detection of pits formed in the earliest stages is spread across a wider range of load stages.

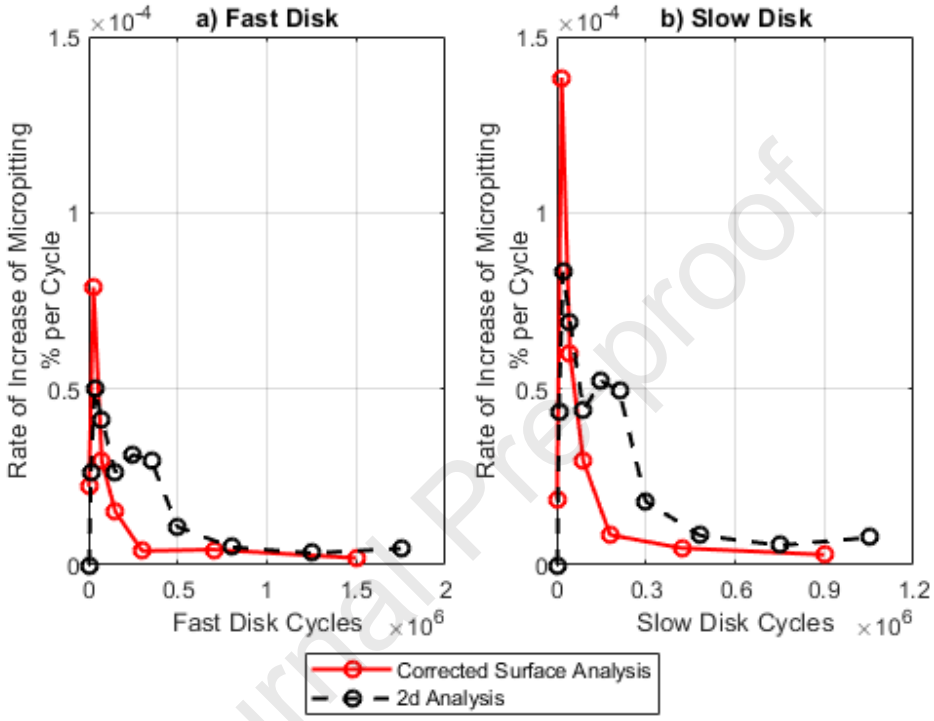

Figure 15 The rate of increase of micropitting on both disks as evaluated by both the surface and two-dimensional analyses.

Conclusion

- A new method has been developed for detecting the pitted area in micropitting test surface samples, representative of typical ground gear tooth surfaces, displaying the early stages of micropitting.
- The method has been applied to both directly measured steel surfaces, and to replicas taken from the same surfaces, and has successfully identified pits via an automated method.
- The method was used to determine the progression of pitted area during a micropitting experiment using a twin-disk rig operating under mixed lubrication conditions.
- When used to process both direct measurements, and ones obtained via replicas, the results show comparable trends in terms of the progression of micropitting, although there is a

consistent scaling factor difference between replica-based measurements and direct measurements, indicating a potential distortion of micropit geometry by the replication process.

- This new method has been compared with a previously-published method based on two-dimensional line measurements, and this has highlighted the potential for two dimensional measurements to mis-represent the extent of pitting, which is a truly three-dimensional phenomenon.
- The new method offers the ability to automatically process the large amounts of areal surface data which can be produced using modern profilometers in order to analyse very subtle surface geometry changes during the initiation and early stages of micropitting.

Acknowledgements

The authors acknowledge the financial support for William Britton's studentship under the EPSRC Doctoral Training Partnership, and the practical advice provided by Design Unit, Newcastle University regarding replica materials.

References

[1] A. Clarke, H.P. Evans, R.W. Snidle, Understanding micropitting in gears, Proc. Inst. Mech. Eng. Part C J. Mech. Eng. Sci. 230 (2016) 1276–1289. https://doi.org/10.1177/0954406215606934.

[2] A. Vrcek, T. Hultqvist, Y. Baubet, M. Björling, P. Marklund, R. Larsson, Micro-pitting and wear assessment of engine oils operating under boundary lubrication conditions, Tribol. Int. 129 (2019) 338–346. https://doi.org/10.1016/j.triboint.2018.08.032.

[3] S. Roy, D. White, S. Sundararajan, Correlation between evolution of surface roughness parameters and micropitting of carburized steel under boundary lubrication condition, Surf. Coat. Technol. 350 (2018) 445–452. https://doi.org/10.1016/j.surfcoat.2018.05.083.

[4] D.K. Prajapati, M. Tiwari, Experimental analysis of contact fatigue damage using fractal methodologies, Wear. 450–451 (2020) 203262. https://doi.org/10.1016/j.wear.2020.203262.

[5] S. Li, A. Kahraman, Micro-pitting fatigue lives of lubricated point contacts: Experiments and model validation, Int. J. Fatigue. 48 (2013) 9–18. https://doi.org/10.1016/j.ijfatigue.2012.12.003.

[6] A. Oila, S.J. Bull, Assessment of the factors influencing micropitting in rolling/sliding contacts, Wear. 258 (2005) 1510–1524. https://doi.org/10.1016/j.wear.2004.10.012.

[7] P. Rabaso, T. Gauthier, M. Diaby, F. Ville, Rolling Contact Fatigue: Experimental Study of the Influence of Sliding, Load, and Material Properties on the Resistance to Micropitting of Steel Discs, Tribol. Trans. 56 (2013) 203–214. https://doi.org/10.1080/10402004.2012.737504.

[8] G.E. Morales-Espejel, P. Rycerz, A. Kadiric, Prediction of micropitting damage in gear teeth contacts considering the concurrent effects of surface fatigue and mild wear, Wear. 398–399 (2018) 99–115. https://doi.org/10.1016/j.wear.2017.11.016.

[9] R. Martins, C. Locatelli, J. Seabra, Evolution of tooth flank roughness during gear micropitting tests, Ind. Lubr. Tribol. 63 (2011) 34–45. https://doi.org/10.1108/00368791111101821.

[10] V. Brizmer, H.R. Pasaribu, G.E. Morales-Espejel, Micropitting Performance of Oil Additives in Lubricated Rolling Contacts, Tribol. Trans. 56 (2013) 739–748. https://doi.org/10.1080/10402004.2013.790097.

[11] S. Hutt, A. Clarke, H.P. Evans, Generation of Acoustic Emission from the running-in and subsequent micropitting of a mixed-elastohydrodynamic contact, Tribol. Int. 119 (2018) 270–280. https://doi.org/10.1016/j.triboint.2017.11.011.

[12] A. Kadiric, P. Rycerz, Influence of Contact Conditions on the Onset of Micropitting in Rolling-Sliding Contacts Pertinent to Gear Applications, AGMA Tech. Pap. 16FTM21 AGMA Va. USA. (2016).

[13] E. Lainé, A.V. Olver, T.A. Beveridge, Effect of lubricants on micropitting and wear, Tribol. Int. 41 (2008) 1049–1055. https://doi.org/10.1016/j.triboint.2008.03.016.

[14] I.S. Al-Tubi, H. Long, J. Zhang, B. Shaw, Experimental and analytical study of gear micropitting initiation and propagation under varying loading conditions, Wear. 328–329 (2015) 8–16. https://doi.org/10.1016/j.wear.2014.12.050.

[15] J. Kattelus, J. Miettinen, A. Lehtovaara, Detection of gear pitting failure progression with on-line particle monitoring, Tribol. Int. 118 (2018) 458–464. https://doi.org/10.1016/j.triboint.2017.02.045.

[16] I.J. Weeks, An Experimental Investigation into the Mixed Lubrication of Steel Surfaces, Cardiff University, 2015.

[17] A. Clarke, I.J.J. Weeks, R.W. Snidle, H.P. Evans, Running-in and micropitting behaviour of steel surfaces under mixed lubrication conditions, Tribol. Int. 101 (2016) 59–68. https://doi.org/10.1016/j.triboint.2016.03.007.

[18] R.W. Snidle, H.P. Evans, M.P. Alanou, Determination of gear and bearing material scuffing limits using high-speed disk machines, in Bench testing of industrial fluid lubrication and wear properties used in machinery applications, edited by Totten, G., Wedeven, L., Anderson, M., and Dickey, J.

(West Conshohocken, PA: ASTM International, 10.1520/STP10503S), 124-2001. https://doi.org/978-0-8031-5453-7

[19] M.F. AL-Mayali, S. Hutt, K.J. Sharif, A. Clarke, H.P. Evans, Experimental and Numerical Study of Micropitting Initiation in Real Rough Surfaces in a Micro-elastohydrodynamic Lubrication Regime, Tribol. Lett. 66 (2018). https://doi.org/10.1007/s11249-018-1110-2.

A Novel Method for Micropit Detection in Ground Surfaces – Highlights

- Replica materials can be used to obtain areal scans of surfaces mid-test
- A novel algorithm is developed to detect micropits in areal scans
- Areal micropit detection algorithm shows improvement against established 2D methods

Declaration of interests

☒ The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

☐The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: