Statistical analysis of damage evolution with a new image tool

C.H. Müller[1], C. Gunkel[2], and L. Denecke[3]
[13]*University of Technology Dortmund, Faculty of Statistics, Vogelpothsweg 87, D-44227 Dortmund, Germany,* [2]*University of Kassel, Department of Mathematics, Heinrich-Plett-Str. 40, D-34109 Kassel, Germany,*
[1]*Corresponding Author*: E-mail: cmueller@statistik.tu-dortmund.de

ABSTRACT

The surface damage evolution under stress is often analysed by images of long-distance microscopes. Usually hundreds of images are obtained during the fatigue process. To analyse this huge amount of images automatically, a new image tool is presented. This new image tool is included in free statistic software so that a statistical analysis of the damage evolution is easily possible. In particular several specific damage parameters can be calculated during the fatigue process. Some of these specific damage parameters are compared statistically here with simple damage parameters using images of two specimens under different stress levels at different time points of the fatigue process. It is shown that the specific damage parameters discriminate between the two different damage evolutions in an earlier stage than the simple parameters. They are also less influenced by different brightness and scales of the images and show other desirable properties of a damage parameter.

Keywords: micro crack growth, damage parameter, automatic analysis, image analysis, statistical analysis, discrimination between stress levels

NOMENCLATURE

CumL	=	cumulated length of detected cracks [µm]
FGV.Auto	=	fraction of grey values between 0 and an automatic threshold [%]
FGV.140	=	fraction of grey values between 0 and 140 [%]
MaxL	=	maximum length of detected cracks [µm]
MeanL	=	mean length of detected cracks [µm]
MGV	=	mean grey value [-]
NoC	=	number of detected cracks [-]

INTRODUCTION

Although damage evolution is often studied by images of the surface of specimens using a long-distance microscope, these studies are mainly of qualitative nature. In particular special chosen cracks were analysed in more detail as the zigzag micro crack in Besel et al.[1] Other studies of damage evolution used simple quantitative approaches as analysing crack numbers, crack densities, or crack orientations. [2,3] However, these studies were restricted to few cracks since the cracks had to be marked by hand.

Since more and more complex probabilistic models are proposed for crack initiation and crack propagation[4-10], these models must be verified and specified by statistical methods. The more complex the probabilistic model is the more data are needed for the statistical analysis. Hence, cracks and related quantities as plastic deformations should be determined automatically. A simple automatic method for measuring damage evolution is the fraction of grey values below a given threshold value proposed by Besel and Brueckner-Foit[3]. However, this fraction of grey values below a given threshold value is rather unspecific since it takes only into account that the images become darker with increasing load cycles due to more cracks and plastic deformations. Moreover, this parameter needs a method to determine an appropriate threshold value.

More specific damage parameters can be obtained by tools of image analysis. However the existing image tools have several disadvantages. Some tools are only able to detect few cracks[11-13] or they can only highlight cracks and damage areas[14,15]. Damage areas can be easily highlighted by a threshold value, i.e. everything below a threshold value is highlighted. These highlighted areas can be reduced to some lines by skeletonization and thinning.[16-18] But simple skeletonization methods do not take into account the different grey levels in the damage areas. Hence, they do not result in the darkest parts of the damage area which are usually cracks. Moreover, every skeletonization method is very sensitive to small changes of the highlighted area. And since micro cracks are usually surrounded by more or less wide damage areas, the usual edge detection methods[17-19] also cannot be used.

The connected components of highlighted damage areas can be used for calculating circumscribing ellipses and rectangles. This is used for example in the free software UTHSCSA Image Tool developed by Wilcox et al.[20] The lengths of the main axes of the ellipses can be used as the lengths of the cracks.[13] This is of course only a bad approximation of true crack paths, i.e. the darkest parts of the damage area. Moreover cracks are kinked and curved so that the true length of a crack is much longer than the length of the longest axis of a circumscribing ellipse or rectangle. For a detailed crack analysis it is also important to know what the kinks and curves of a crack are.

Gunkel et al.[21] proposed a new crack recognition method which follows the most dark parts of a damage area. Although it is based on the damage areas, it has the ability to provide all kinks and curves of a crack. Since it uses an automatic threshold value, this automatic threshold value can also be used for determining the fraction of grey values. Additionally, this image tool provides more specific damage parameters based on detected micro cracks and damage areas. It can be used in situations where hundreds up to thousands micro cracks or damage areas, respectively, are present, and it can be applied automatically to hundreds of images. Since the new method is included in the free statistical software *R*[22], statistical methods can be easily applied to the detected micro cracks and damage areas as well as to the related damage parameters.

Here it is shown that this image tool leads to several more specific damage parameters. Thereby a good damage parameter should not only satisfy that it is strictly increasing with an increasing number of load cycles as proposed by Besel and Brueckner-Foit[3]. It also should provide a statistically significant discrimination between damage evolution at different stress levels at an early stage of the fatigue process. Moreover, its variation should increase during the fatigue process and it should not be influenced by different scales and brightness of the photos.

The paper is organized as follows. At first, the experimental procedure for getting the images of the surface of two specimens at different stress is described. After that the new image tool is explained. Finally, the image tool is used to define five more specific damage parameters and these five damage parameters are in particular analysed with respect to their discrimination property between stress levels. The behaviour of these damage parameters are compared with two less specific damage parameters which can be automatically obtained, namely the fraction of grey values below a fixed threshold value and the mean grey value.

EXPERIMENTAL PROCEDURE

Standard fatigue experiments were performed on a servohydraulic testing machine with two symmetrical-notched round specimens with a minimum diameter of 6 mm. The material of the specimens was a low carbon steel. The specimens were grounded and polished before the experiments in order to facilitate surface observation. The surface observation was carried out with a long-distance microscope (Hirox, Japan) and a 1.14 MPig digital b/w camera. During the fatigue experiments, the surface was scanned after a predefined number of load cycles. For more details of the experimental procedure see the description of Besel and Brueckner-Foit[3].

The two specimens were exposed to two different stress levels. The nominal stress level for the first specimen, Specimen 31, was 400 MPa and for the second specimen, Specimen 10, was 360 MPa. The images were taken from an area of approximately 4x5 mm^2 size. A linear elastic FE analysis of the specimen under test conditions showed that the stress within this area varies about 50 MPa. This area was too large to cover it with one photo so that 9x6=54 image segments were taken in the case of Specimen 31 and 9x5=45 image segments in the case of Specimen 10. Each image segment has 696x512 pixels, where 80 pixels correspond to 100 μm. The segments have different quality. Some have shadows at the border and some are blurred. Since the segments overlap, they could be joined leading to images of the whole surface with 3337x4165 pixels for Specimen 31 and images with 2659x4221 pixels for Specimen 10. These images were compressed for a faster analysis to images with 1669x2083 and 1330x2111 pixels, respectively. The compressed total image of Specimen 31 after 18 000 load cycles is shown in Figure 1.

Fig. 1 Composed total image of Specimen 31 after 18 000 load cycles.

The images of Specimen 31 were taken at the beginning and after 1 000, 2 000, 3 000, 4 000, 5 000, 6 000, 7 000, 8 000, 9 000, 10 000, 12 000, 14 000, 16 000, 18 000 load cycles, i.e. at 15 time points. The life time of this specimen was approximately 20 000 load cycles. The images of Specimen 10 were obtained at the beginning and in steps of 1 000 load cycles up to 20 000 load cycles and then after 25 000, 30 000, 35 000, 37 000, 39 000, 40 000, 42 000, and 44 000 load cycles, i.e. at 29 time points. This specimen had a life time of 44 300 load cycles. The life time of both specimens was much lower than expected from other fatigue experiments where no images were taken.

The new crack recognition method

The new crack recognition method is provided by the *R* package *crackrec* which can be downloaded from the homepage of the corresponding author. The main part is written in *C* to be as fast as possible. It is included in the statistical software *R*[22] to allow easy applications of statistical methods. *R* is free software with a huge amount of statistical methods, which is meanwhile widespread in the statistical community. Hence cracks and damage areas detected inside *R* can be easily analysed with many statistical methods provided by *R*. Since *R* is also a programming language, additional tools for analysing the cracks and the damage areas can be easily implemented as well. Hence damage analysis and statistical analysis can be carried out with the same software.

The damage analysis with the package *crackrec* proceeds in four steps. In the first step the image must be converted to an *R* data set. *R* provides several methods for this. But *crackrec* provides with the function *rbmp* a procedure to convert BMP files to *R* data sets. This was used in the applications regarded here. In a second step shadows are removed. This is necessary since many image segments have a strip of shadow at the border (see Figure 1). It can be removed by the median filter[17] in the function *median.filter* or by the more special function *shadow.remove*. This special function was developed for the Specimen 31, but it turned out that it also works well for Specimen 10. In the results presented below, the median filter with a 51x51 window was used, but the results are similar if other windows or *shadow.remove* are used. Since the image segments have different light exposures (see Figure 1), it is not possible to use the same threshold value for all segments to determine the damage areas. Therefore the threshold value is determined for each image segment with the function *threshold.msi*. A simple method to determine the threshold could be the mean grey level. But it turned out that this is overestimating the threshold. Alternatively, a method of steepest increase of a kernel density estimator of the grey values is used in *threshold.msi*. The kernel density estimator depends on a bandwidth so that the thresholds can be influenced by the bandwidth. The results presented below were obtained with a bandwidth of 30. If grey levels from 0 to 255 were used, then this leads for several image segments to a threshold of approximately 140, i.e. the same threshold as Besel and Brueckner-Foit[3] used to determine the damage areas. But *threshold.msi* provides also different thresholds for many other image segments.

Fig. 2 *Crackcluster* with different grey values.

Fig. 3 *Crackpath* through the *crackcluste*r.

In the last step, the function *crackrec* provides the connected components of the damage areas, called *crackclusters*, and longest shortest paths through these crack clusters, called *crackpaths*. The crack paths are found with Dijkstra's shortest path algorithm.[23] In this algorithm, a distance measure is used which forces the paths to follow the darkest parts of the crack clusters. See Figures 2 and 3 which concern the *crackcluster* in the middle of the image given in Figure 4. More details of the method and the output are given in Gunkel et al.[21] But not all detected *crackclusters* and *crackpaths* are related to real cracks. They include all dark areas and these dark areas can also consist of pits and scratches caused by grinding and polishing the surface. But for simplicity all these phenomena are called *crackclusters* and lines through these are called *crackpaths*. The crack paths can be plotted in the same image or separately with the function *crackplot*. A result of *crackrec* using a separate plot with *crackplot* is shown in Figure 5 for the image segment in Figure 4. Thereby, the start and end points of a crack path are connected with a grey line so that different crack paths are better visible.

Fig. 4 Image segment of Specimen 31 after 18 000 load cycles.

Fig. 5 Detected c*rackpaths* in the image segment of Fig. 4.

RESULTS AND DISCUSSIONS

Results of the image tool

Figure 6 shows the detected crack paths in the composed total image of Specimen 31 which is shown in Figure 1. This composed image consists of the 54 image segments taken after

18 000 load cycles, i.e. shortly before failure. Since 12 449 crack clusters were found, only crack paths with a distance between start and end point larger than 30 pixels, i.e. 75 μm, are shown (note that the total image is compressed by 50%). Most of these cracks have a horizontal orientation, i.e. a perpendicular direction to the loading axis which is given by the vertical axis. Moreover, the longest cracks are concentrated in the middle of the image where the strain is up to 450 MPa while the strain at the border is approximately 400 MPa. The same observation was made for Specimen 10 shortly before failure, i.e. after 44 000 load cycles. To compare the two specimens with the crack recognition method, Figure 7 provides the detected cracks in the composed image of Specimen 10 also after 18 000 load cycles. It consists of much less long cracks compared with the image of Specimen 31 in Figure 6 and scratches are still visible as vertical lines.

Fig. 6 Detected cracks longer than 75 μm of the composed total image of Specimen 31 after 18 000 load cycles.

Fig. 7 Detected cracks longer than 75 μm of the composed total image of Specimen 10 after 18 000 load cycles.

Comparison of several damage parameters

For a statistical comparison between Specimen 31 and Specimen 10, the single image segments were analysed with the crack recognition program. In each image segment, the number of detected crack clusters (NoC), the mean of the lengths of the crack paths (MeanL), the maximum length (MaxL) and the cumulative length (CumL) of the detected crack paths, and the fraction of grey levels below the automatic threshold value of the image tool (FGV.Auto) were used as damage parameters. These parameters were compared with two simple parameters, namely the mean grey value (MGV) and the fraction of grey levels between 0 and 140 (FGV.140), which were obtained without using the median filter or other special features of the image tool. The threshold 140 was used for the FGV.140 parameter since in several cases it is the appropriate threshold, the reason why it was also used in Besel and Brueckner-Foit[3]. The automatic threshold in FGV.Auto was always determined with *threshold.msi* of *crackrec* with bandwidth 30 after the use of the median filter with a 51x51 window. For the more specific length based parameters MeanL, MaxL, and CumL, the total lengths of the crack paths were used, instead of the distances between the start and the end points, although this is also possible. These seven quantities were obtained at each time point. Hence 54 observations exist for each quantity and each time point for Specimen 31, and 45 observations for each quantity and each time point for Specimen 10.

Although the image segments overlap and a crack cluster may influence the damage behaviour in its neighbourhood, these influences are so small that stochastic independence between the observations of each quantity at a given time point can be assumed. But the observations of different quantities and at different time points are stochastically dependent. Besides independence, the identical distribution of observations is an important condition for statistical methods. Since the damage is more concentrated in the inner part of the whole image, this property could be put into question. However, for Specimen 31 the boxplots in Figure 8 show that the distribution of the determined maximum lengths using all 54 image segments does not differ very much from that using only the 7x4=28 interior image segments. Similar results hold for the other six quantities and for Specimen 10. Sometimes the variation is smaller using only the interior segments. But the mean values and the appearance of outliers are the same. Hence all image segments can be used. But it cannot be assumed that the distribution of the observed quantities is normal, since the distribution is given by a mixture of distributions.

Fig. 8 Boxplots of the calculated maximum lengths for Specimen 31 after 18 000 load cycles.

For this reason only the distribution-free Wilcoxon test[24] can be used for calculating confidence intervals and for testing whether the quantities behave differently at the two specimens. Figures 9 to 15 show the 95% Wilcoxon confidence intervals and the medians for the quantities obtained at the two specimens. They all increase monotonously with increasing number of load cycles. Some small deviations from this behaviour are caused by different image quality. The only significant exceptions are the parameters MGV, FGV.140, and MeanL which decrease at the beginning. However, the MGV and FGV.140 show this behaviour only for Specimen 10 and this is caused by a change of scaling of the photos after 3 000 load cycles. But the decreasing behaviour of MeanL is not an artefact. The reason is that pits and scratches caused by grinding and polishing the surface are not filtered out at time 0. They were not filtered out because it should be tested whether the surfaces of Specimen 31 and 10 differ with respect to these phenomena. Because of these pits and scratches, some long “cracks” are detected at time 0 which are no real cracks. After 1 000 and 2 000 load cycles many short micro cracks have appeared and are detected. These high amounts of short micro cracks reduce the mean length. But after 2 000 load cycles for Specimen 31 and after 3 000 load cycles for Specimen 10, the mean length increases steadily. Hence all seven quantities satisfy the main property of a reasonable damage parameter, namely increasing amount during the fatigue process. In this sense, the seven damage parameters behave similar.

Fig. 9 95% Wilcoxon confidence intervals and medians for the mean grey value (MGV).

Fig. 10 95% Wilcoxon confidence intervals and medians for the fraction of grey levels between 0 and 140 (FGV.140).

Fig. 11 95% Wilcoxon confidence intervals and medians for the fraction of grey levels below the automatic threshold (FGV.Auto).

Fig. 12 95% Wilcoxon confidence intervals and medians for the number of detected cracks (NoC).

Fig. 13 95% Wilcoxon confidence intervals and medians for the mean length of detected cracks (MeanL).

Fig. 14 95% Wilcoxon confidence intervals and medians for the maximum length of detected cracks (MaxL).

Fig. 15 95% Wilcoxon confidence intervals and medians for the cumulated length of detected cracks (CumL).

But a good damage parameter should not be influenced by the scaling or brightness of the images. However, this is not the case for the parameters MGV and FGV.140 for Specimen 10, as can be seen from Figures 9 and 10. All other five damage parameters are not influenced by this different scaling of photos of Specimen 10, see the Figures 11 to 15. Additionally, the distributions of the FGV.140 parameters are so much asymmetric that the median is often not included in the Wilcoxon confidence intervals.

Moreover, a good damage parameter should show increasing variability, and thus increasing length of the confidence intervals, during the fatigue process since the stress at different image segments varies about 50 MPa. This property is in particular satisfied by the five parameters based on the image tool and not satisfied by MGV and FGV.140. It is surprising that the behaviour of the fraction of grey levels below a threshold becomes much better using the automatic threshold value from the image tool than the fixed threshold value 140.

Table 1 P-values of the two-sample Wilcoxon test

Finally, a good damage parameter should discriminate between different levels of stress. Figures 9-15 show that indeed the behaviour of all seven automatic damage parameters is different for the two specimens. In particular, the growth of the parameters is much more slowly at Specimen 10, the specimen with less stress. The parameters achieve sometimes but not always the same level shortly before failure. It seems that a smaller stress provides more cracks, but with shorter length. To analyse whether the observed differences are statistically significant, the seven parameters were tested with the Wilcoxon test[24]. Since Figures 9 to 15 show that the Wilcoxon confidence intervals are not overlapping after 3 000 load cycles, there is always a statistically significant difference between the two specimens for all seven damage parameters after 3 000 load cycles. Before that time point, the seven damage parameters behave differently. As soon as the Wilcoxon confidence intervals overlap – which sometimes is the case here -, the two-sample Wilcoxon test must be used. Therefore, Table 1 shows the p-values of the two-sample Wilcoxon test at the beginning, after 1 000, 2 000, and 3 000 load cycles. It shows that FGV.Auto provides a significant difference already after 1 000 load cycles, while NoC, MeanL, and CumL provide the significant difference after 2 000 load cycles. I.e. these four damage parameters are able to discriminate between the two specimens already at 1 000 and 2 000 load cycles. The parameters MGV, FGV.140, and MaxL do not have this property. They only discriminate from 3 000 load cycles on. Since the scaling changes from 3 000 load cycles for Specimen 10, the discrimination property of MGV and FGV.140 may be even worse. It is remarkable that the five damage parameters based on the image tool are not influenced by this change of scaling. That MaxL discriminates only after 3 000 load cycles, may be caused by the fact that the growth of cracks starts for Specimen 31 at this time point. Before this time point, crack initiation is dominating in both specimens.

Except MaxL, all considered damage parameters do not show significant difference at the beginning. This indicates that the two specimens provide the same starting conditions. That the maximum length is an exception is only caused by some long scratches on Specimen 10 which are still visible after 18 000 load cycles in the image in Figure 7.

CONCLUSION

Two specimens of standard steel (51CrV4) were subjected to fatigue loading with different stress, one with nominal stress of 400 MPa and one with nominal stress of 360 MPa. The surfaces of these two specimens were scanned with a long-distance microscope at different time points given by load cycles of the fatigue process. The resulting hundreds of images were analysed with a new image tool which provides automatically several specific damage parameters. Five of these damage parameters were studied in more detail, namely the fraction of grey values below an automatic threshold value given by the image tool (FGV.Auto), the number of detected crack clusters (NoC), the mean length (MeanL), the maximum length (MaxL), and the cumulated length (CumL) of detected cracks. These five parameters show all desirable properties of a damage parameter, i.e.

- the parameter values increase monotonously with increasing number of load cycles,
- they are not influenced by different scales of the photos,
- the variability increases during the fatigue process,
- the parameters discriminate between the two specimens with different stress in an early stage of the fatigue process.

These properties were not satisfied by two simple automatic parameters, namely the mean grey level (MGV) and the fraction of grey values below a fixed threshold value (FGV.140). It was surprising how much the varying automatic threshold value improves the parameter based on the fraction of grey values.

The discrimination was statistically significant from 2 000 load cycles on whereas the life time of the specimens were about 20 000 and 44 000 load cycles, respectively. The discrimination property of FGV.Auto, NoC, MeanL and CumL was slightly better than that of MaxL. The discrimination property of MaxL was similar to that of the two simple parameters MGV and FGV.140. But the good discrimination property of the two simple parameters may be caused only by a jump of the scales after 3 000 load cycles for one of the two specimens.

The new image tool is not restricted to provide the considered damage parameters. Also other quantities like median crack lengths, crack lengths projected to the axis perpendicular to the loading axis, crack orientations and orientations of linear parts of a crack[21,25] can be considered and shall be analysed in future in more detail. Thereby an interesting question is to quantify the dependence of such damage parameters on the stress level and to find a relationship to the life time.

Acknowledgements

The results of this paper base on a common project with the group of Prof. Dr. Angelika Brückner-Foit from the Institute for Materials Engineering of the University of Kassel. This project was supported by the transregional collaborative research centre SFB/TR TRR 30, which was kindly supported by the German Research Foundation (DFG).

REFERENCES

1 Besel, M., Brueckner-Foit, A., Zeismann, F., Gruening, A. and Mannel, J. (2010) Damage Accumulation of Graded Steel. *Engineering Failure Analysis* **17**, 633–640.

2 Brueckner-Foit, A. and Huang, X. (2008) On the determination of material parameters in crack initiation laws. *Fatigue Fract. Eng. Mater. Struct.* **31**, 980–988.

3 Besel, M. and Brueckner-Foit, A. (2008) Surface damage evolution of engineering steel. *Fatigue Fract. Eng. Mater. Struct.* **31**, 885–891.

4 Ihara, C. and Tanaka, T. (2000). A stochastic damage accumulation model for crack initiation in high-cycle fatigue. *Fatigue Fract. Eng. Mater. Struct.* **23**, 375–280.

5 Nicholson, D. W., Ni, P. and Ahn, Y. (2000) Probabilistic theory for mixed mode fatigue crack growth in brittle plazes with random cracks. *Engineering Fracture Mechanics* **66**, 305–320.

6 Meyer, S., Brückner-Foit, A., Möslang, A. and Diegele, E. (2002) Stochastic simulation of fatigue damage accumulation in a martensitic steel. *Mater. wiss. Werkstofftech.* **33**, 275–279.
7 Brückner-Foit, A., Meyer, S. and Möslang, A. (2003). A stochastic simulation model for microcracks in a martensitic steel. *Comp. Mater. Sci.* **26**, 102–110.
8 Heron, E.A. and Walsh, C.D. (2008) A continuous latent spatial model for crack initiation in bone cement. *Appl. Statist*. **57**, 25–42.
9 Jha, S.K., Millwater, H. R. and Larsen, J. M. (2009) Probabilistic sensitivity analysis in life-prediction of an α+β titanium alloy. *Fatigue Fract. Eng. Mater. Struct.* **32**, 493–504**.**
10 Chiquet, J., Limnios, N. and Eid, M. (2009) Piecewise deterministic Markov processes applied to fatigue crack growth modelling. *J. Statist. Plann. Inference* **139**, 1657–1667.
11 Purcell, D. (1983) Automatic crack detection. *Sensor Review* **3**, 130–131.
12 Cheu, Y. F. (1984) Automatic crack detection with computer vision and pattern recognition of magnetic particle indications. *Materials Evaluation* **42**, 1506–1510, 1514.
13 Fletcher, D.I., Franklin, F.J. and Kapoor, A. (2003) Image analysis to reveal crack development using a computer simulation of wear and rolling contact fatigue. *Fatigue Fract. Eng. Mater. Struct*. **26**, 957–967.
14 Iyer, S. and Sinha, S. K. (2005) A robust approach for automatic detection and segmentation of cracks in underground pipeline images. *Image and Vision Computing* **23**, 921–933.
15 Fujita, Y., Mitani, Y. and Hamamoto, Y. (2006) A method for crack detection on a concrete structure. *Proceedings - International Conference on Pattern Recognition* **3**, 901–904.
16 Russ, J. C. (2006) *The Image Processing Handbook*. Taylor \& Francis Ltd, London.
17 Burger, W. and Burge, M.J. (2007) *Digital Image Processing: An Algorithmic Introduction Using Java.* Springer, New York.
18 O'Gorman, L., Sammon, M. J. and Seul, M. (2008) *Practical Algorithms for Image Analysis with CD-ROM*. Cambridge University Press, Cambridge.
19 Jain, A. K. (1989) *Fundamentals of Digital Image Processing.* Prentice-Hall, Upper Saddle River, NJ.
20 Wilcox, C.D., Dove, S.B., McDavid, W.D. and Greer, D.B. (2009) UTHSCSA Image Tool. Department of Dental Diagnostic Science at The University of Texas Health Science Center, San Antonio, Texas, http://ddsdx.uthscsa.edu/dig/itdesc.html.
21 Gunkel, C., Stepper, A., Müller, A.C., and Müller, Ch.H. (2009) Micro crack detection with Dijkstra's shortest path algorithm. *Submitted.*
22 R Development Core Team (2009) R: A language and environment for statistical computing. *R Foundation for Statistical Computing*, Vienna, Austria, http://cran.r-project.org/.
23 Cormen, T.H., Leiserson, C.E., Rivest, R.L., and Stein, C. (2001) *Introduction to Algorithms*. MIT Press, Cambridge.
24 Lehmann, E. (2006) *Nonparametrics. Statistical Methods Based on Ranks*. Springer, New York.
25 Müller, Ch.H. (2009) Data depth for simple orthogonal regression with application to crack orientation. To appear in *Metrika*. DOI 10.1007/s00184-009-0294-8.

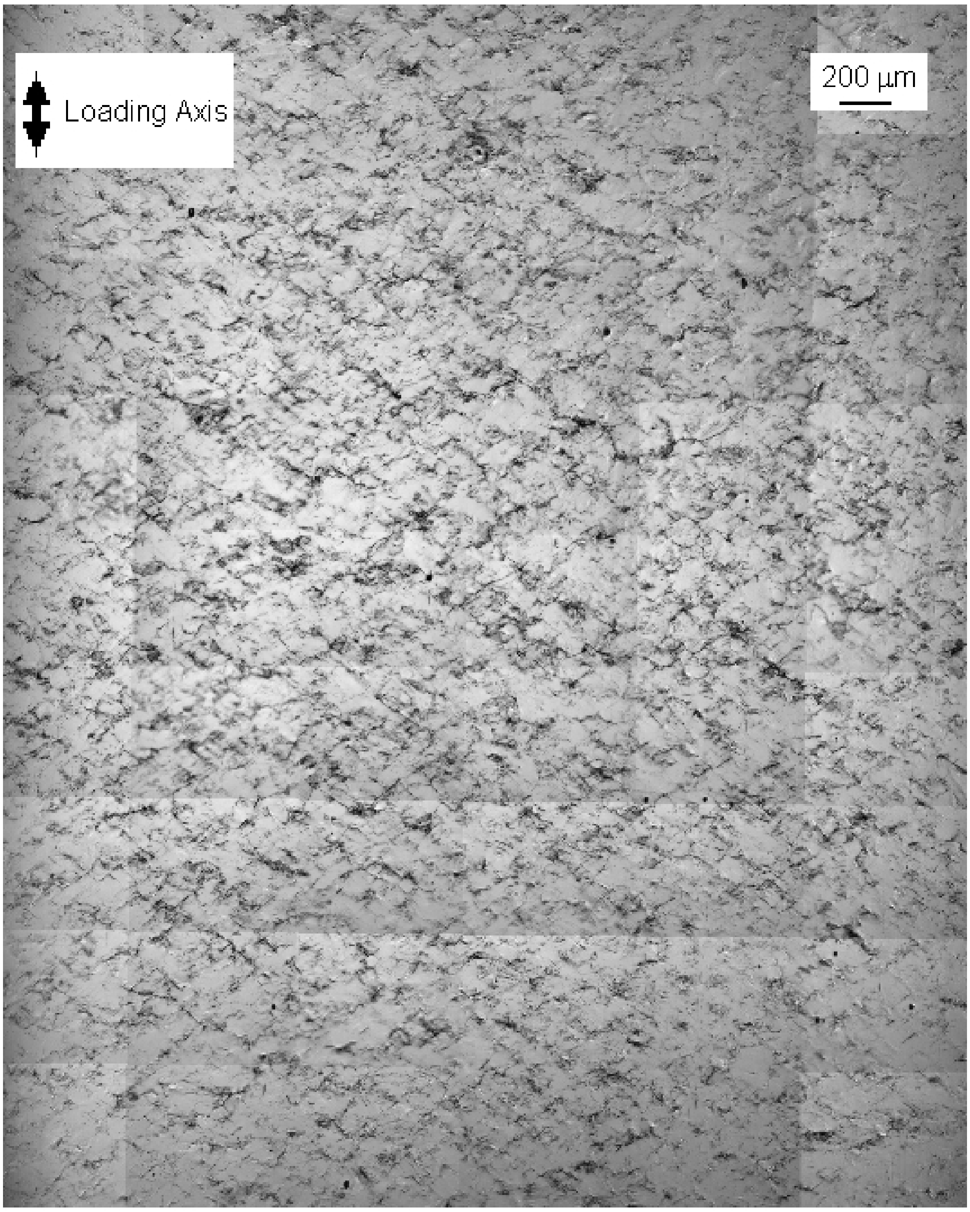

Fig. 1 Composed total image of Specimen 31 after 18 000 load cycles.

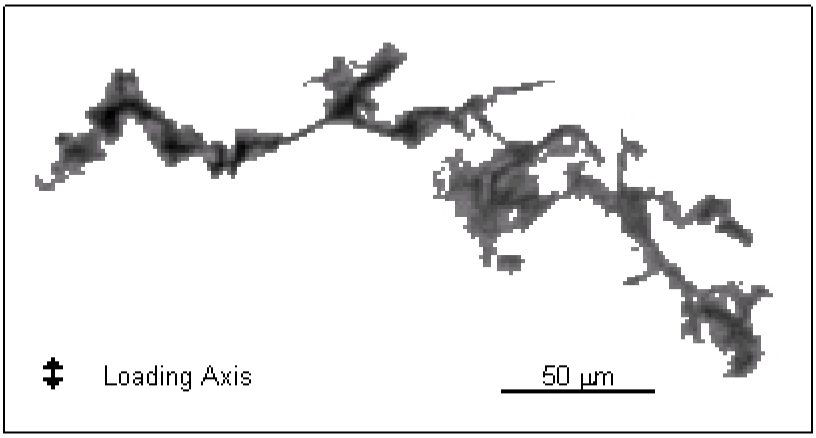

Fig. 2 *Crackcluster* with different grey values.

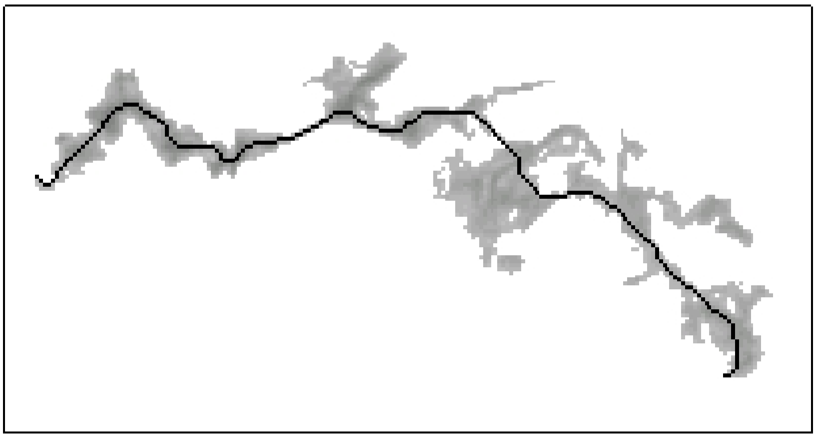

Fig. 3 *Crackpath* through the *crackcluste*r.

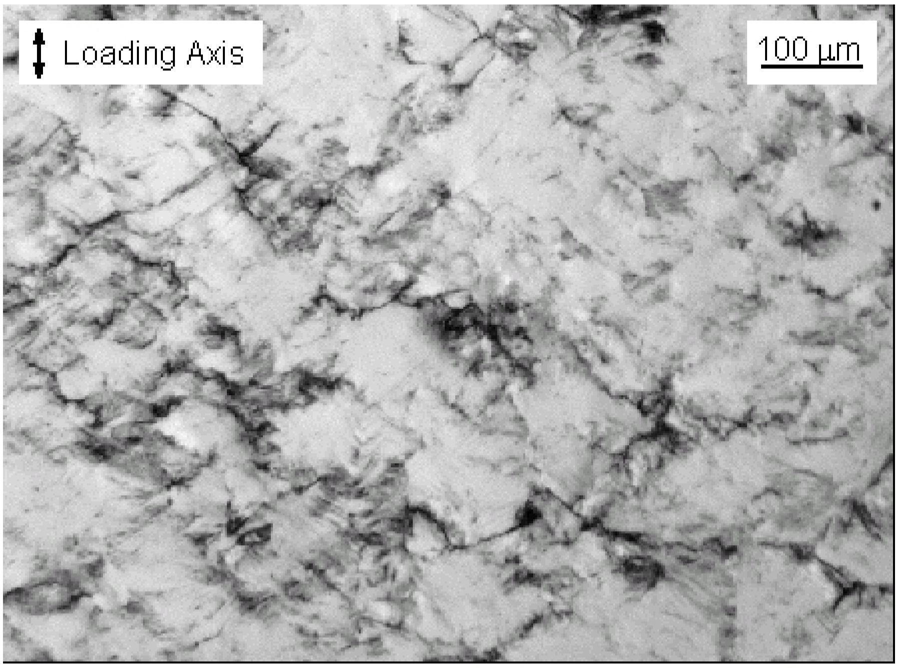

Fig. 4 Image segment of Specimen 31 after 18 000 load cycles.

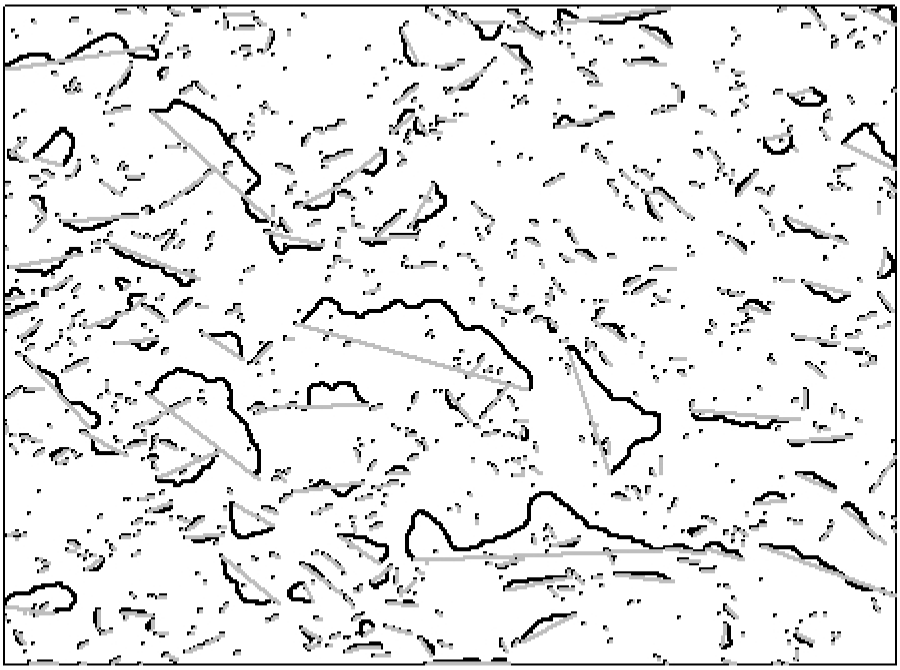

Fig. 5 Detected c*rackpaths* in the image segment of Fig. 4.

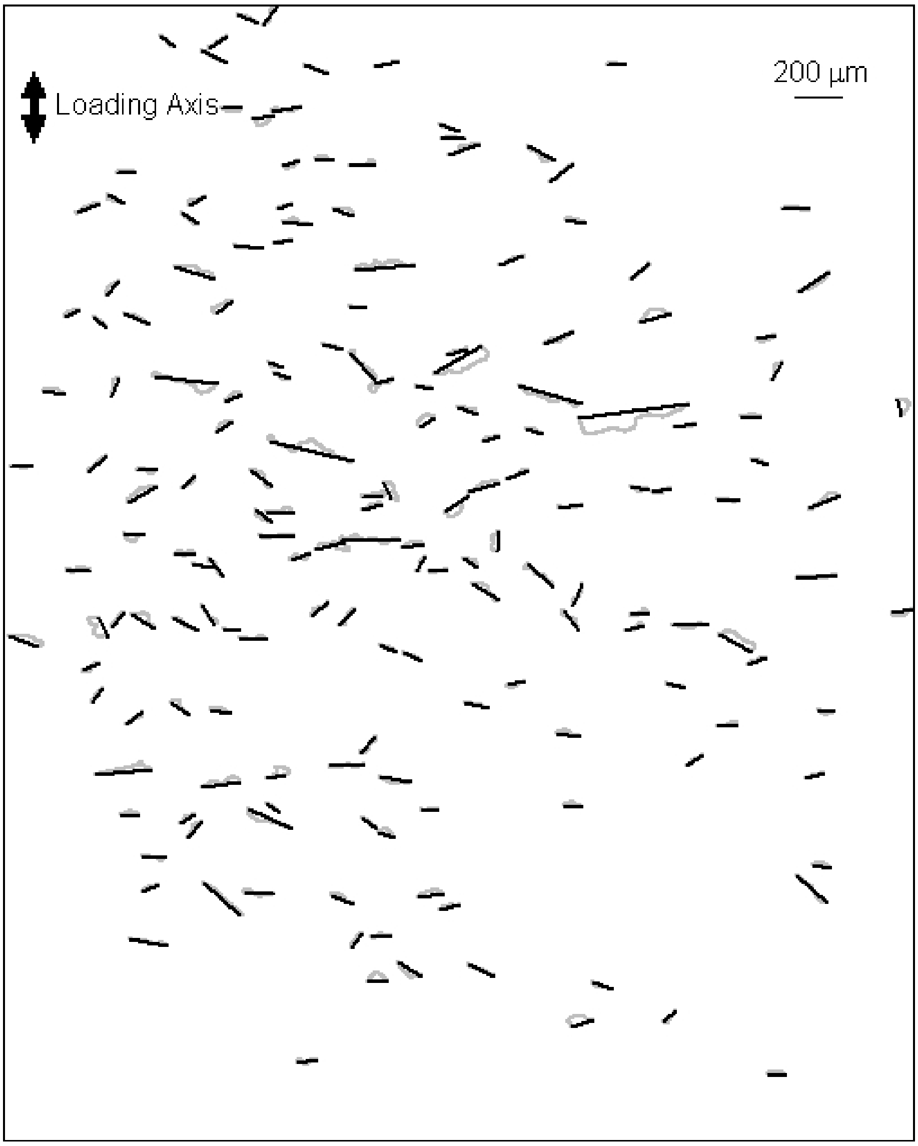

Fig. 6 Detected cracks longer than 75 µm of the composed total image of Specimen 31 after 18 000 load cycles.

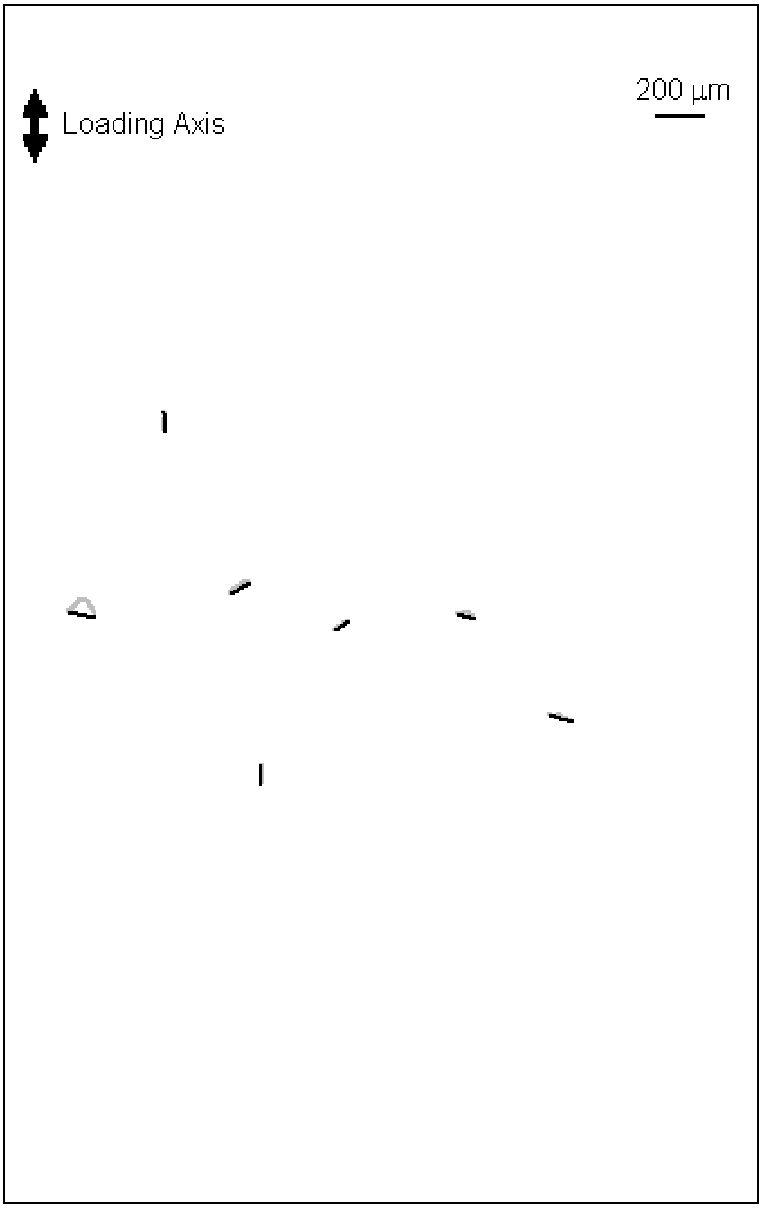

Fig. 7 Detected cracks longer than 75 µm of the composed total image of Specimen 10 after 18 000 load cycles.

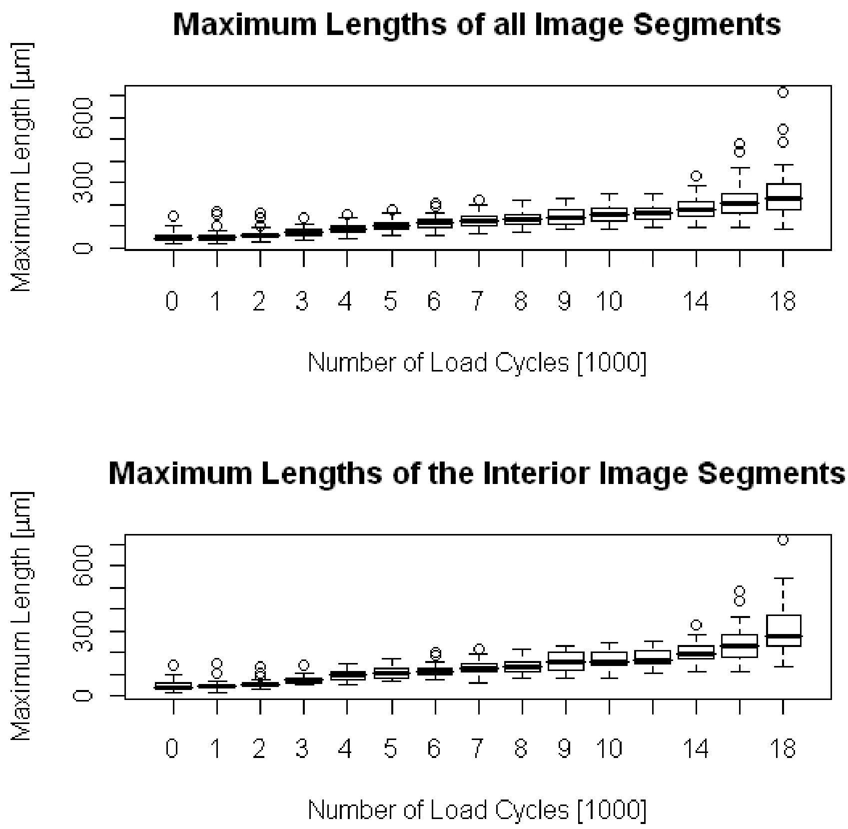

Fig. 8 Boxplots of the calculated maximum lengths (MaxL) for Specimen 31 after 18 000 load cycles.

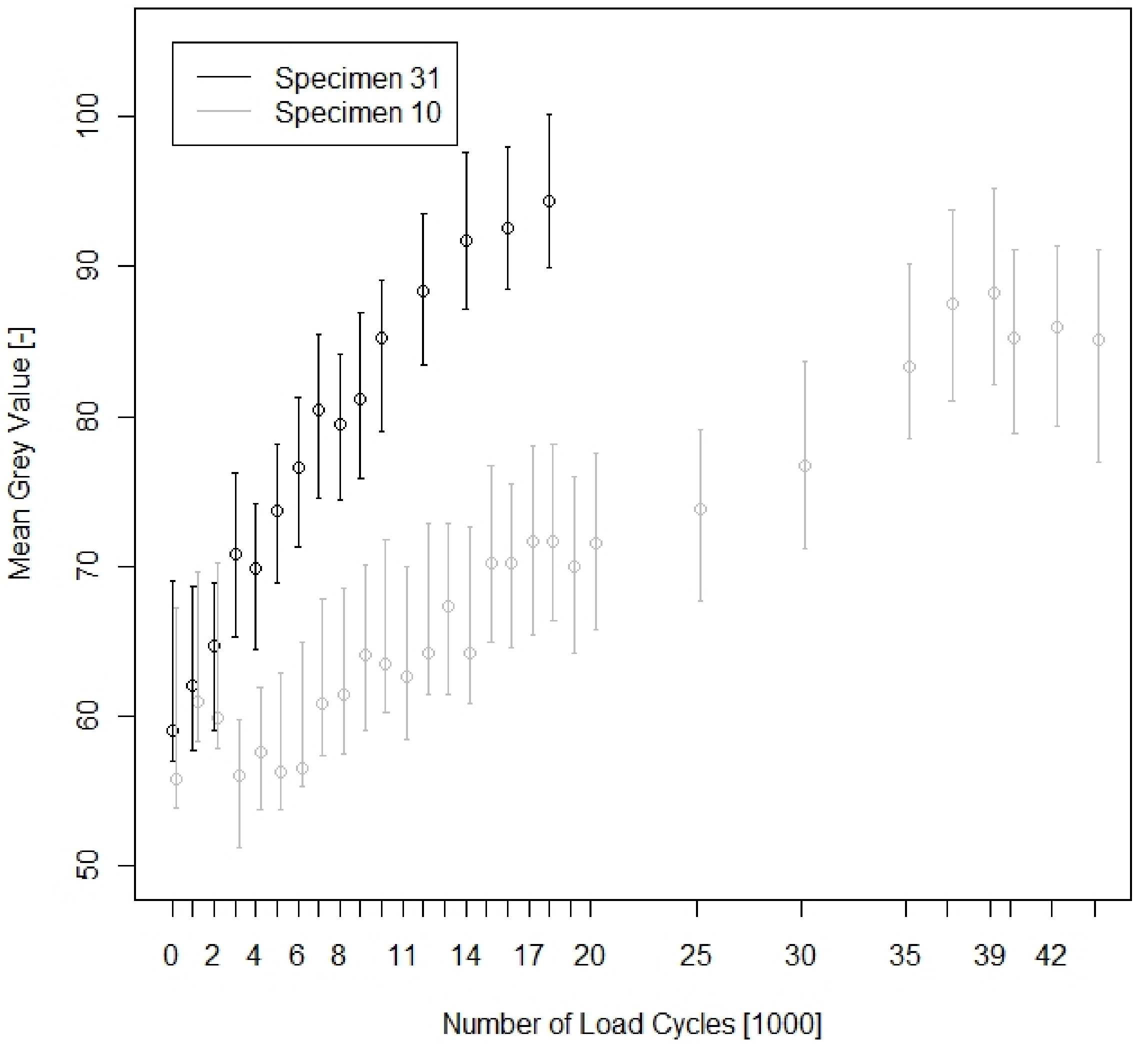

Fig. 9 95% Wilcoxon confidence intervals and medians for the mean grey value (MGV).

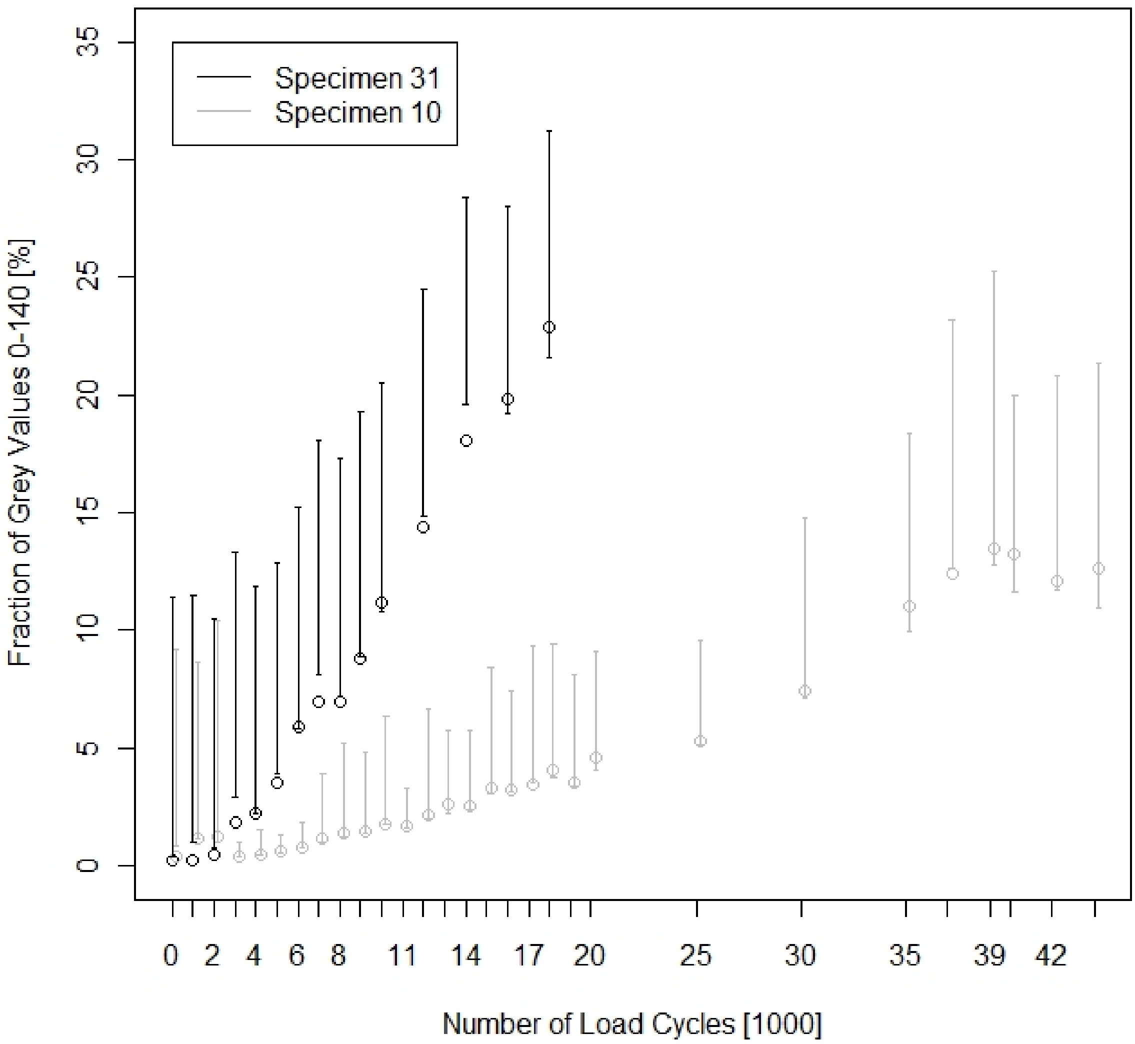

Fig. 10 95% Wilcoxon confidence intervals and medians for the fraction of grey levels between 0 and 140 (FGV.140).

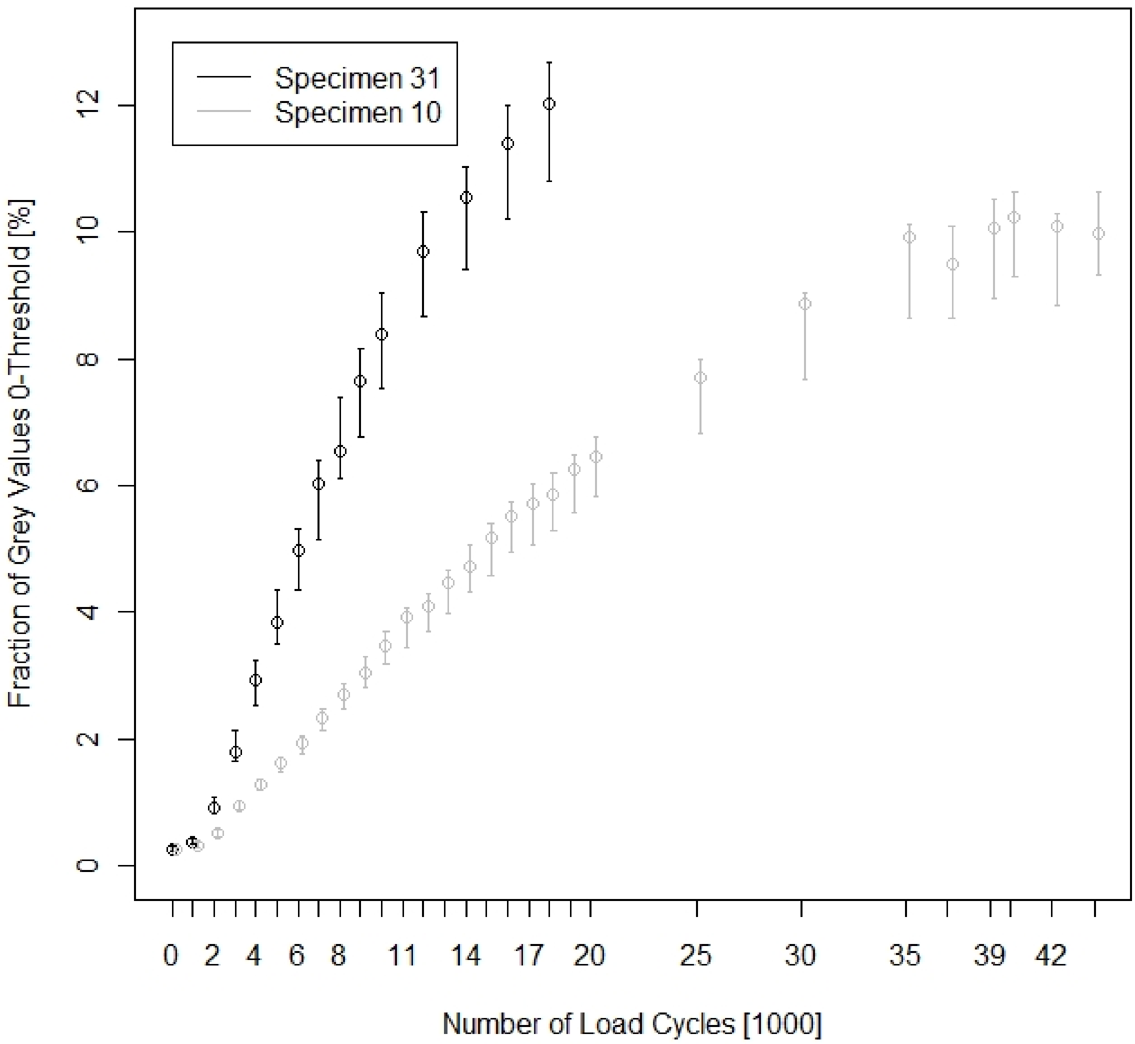

Fig. 11 95% Wilcoxon confidence intervals and medians for the fraction of grey levels below the automatic threshold (FGV.Auto).

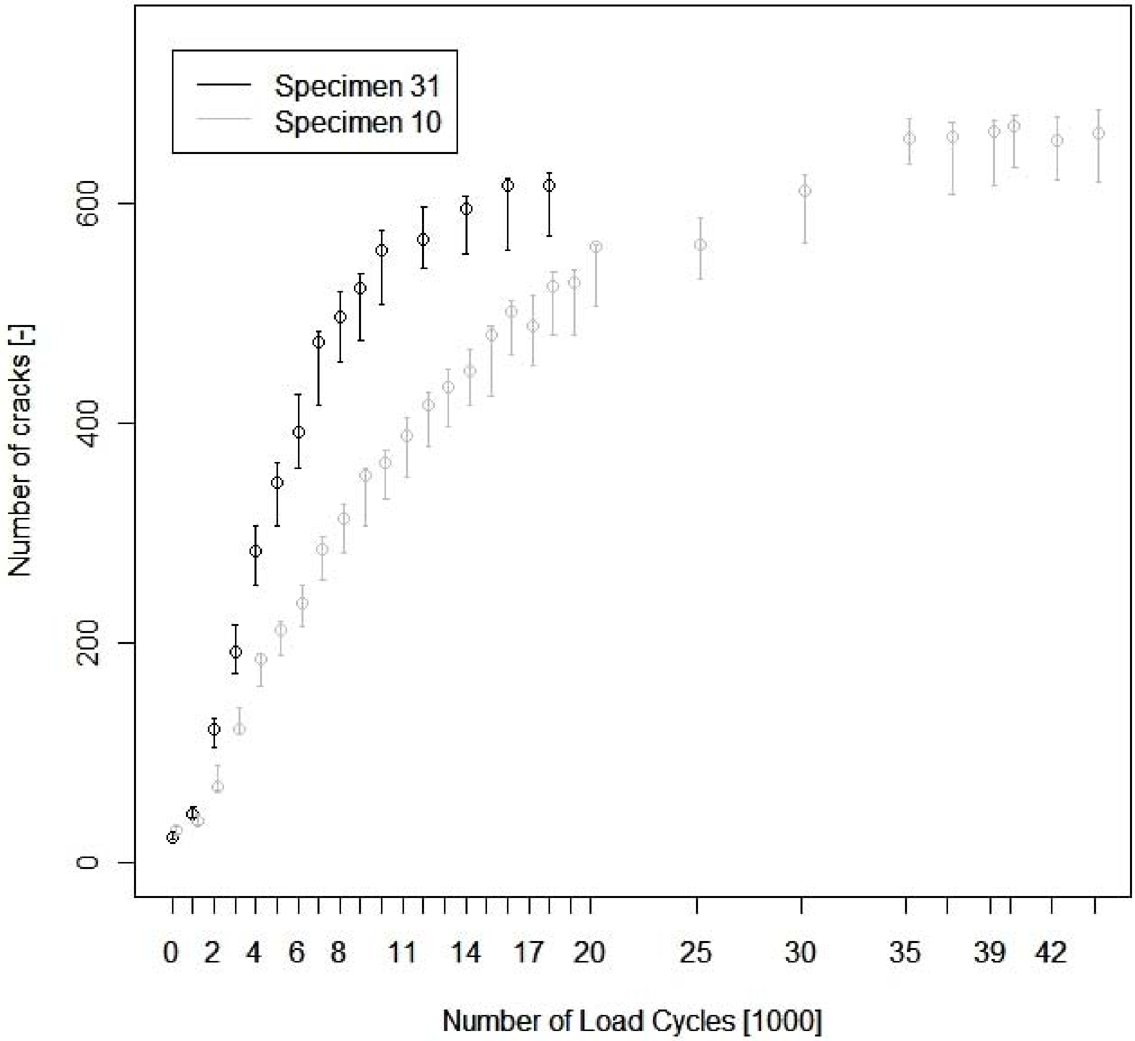

Fig. 12 95% Wilcoxon confidence intervals and medians for the number of detected cracks (NoC).

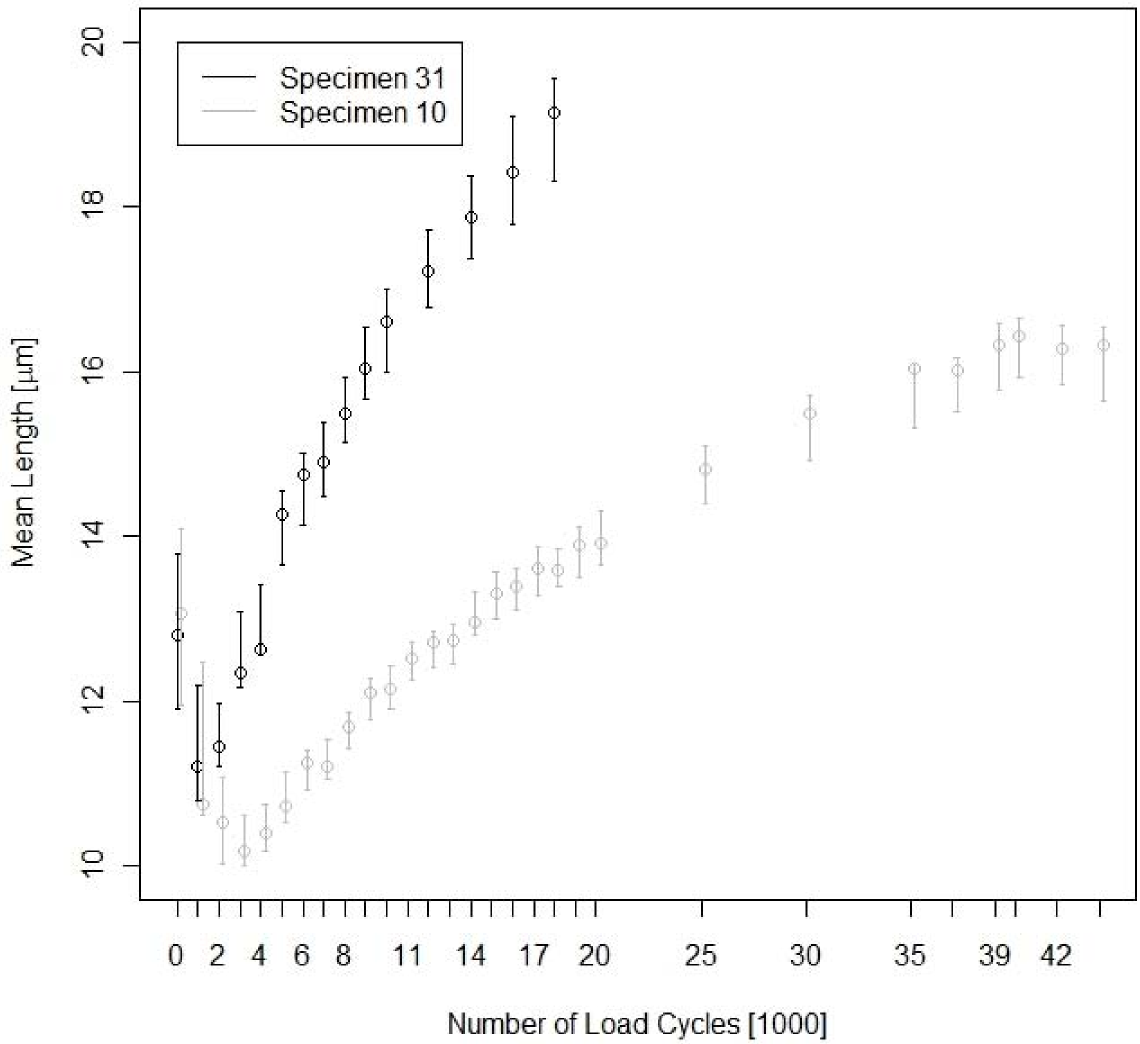

Fig. 13 95% Wilcoxon confidence intervals and medians for the mean length of detected cracks (MeanL).

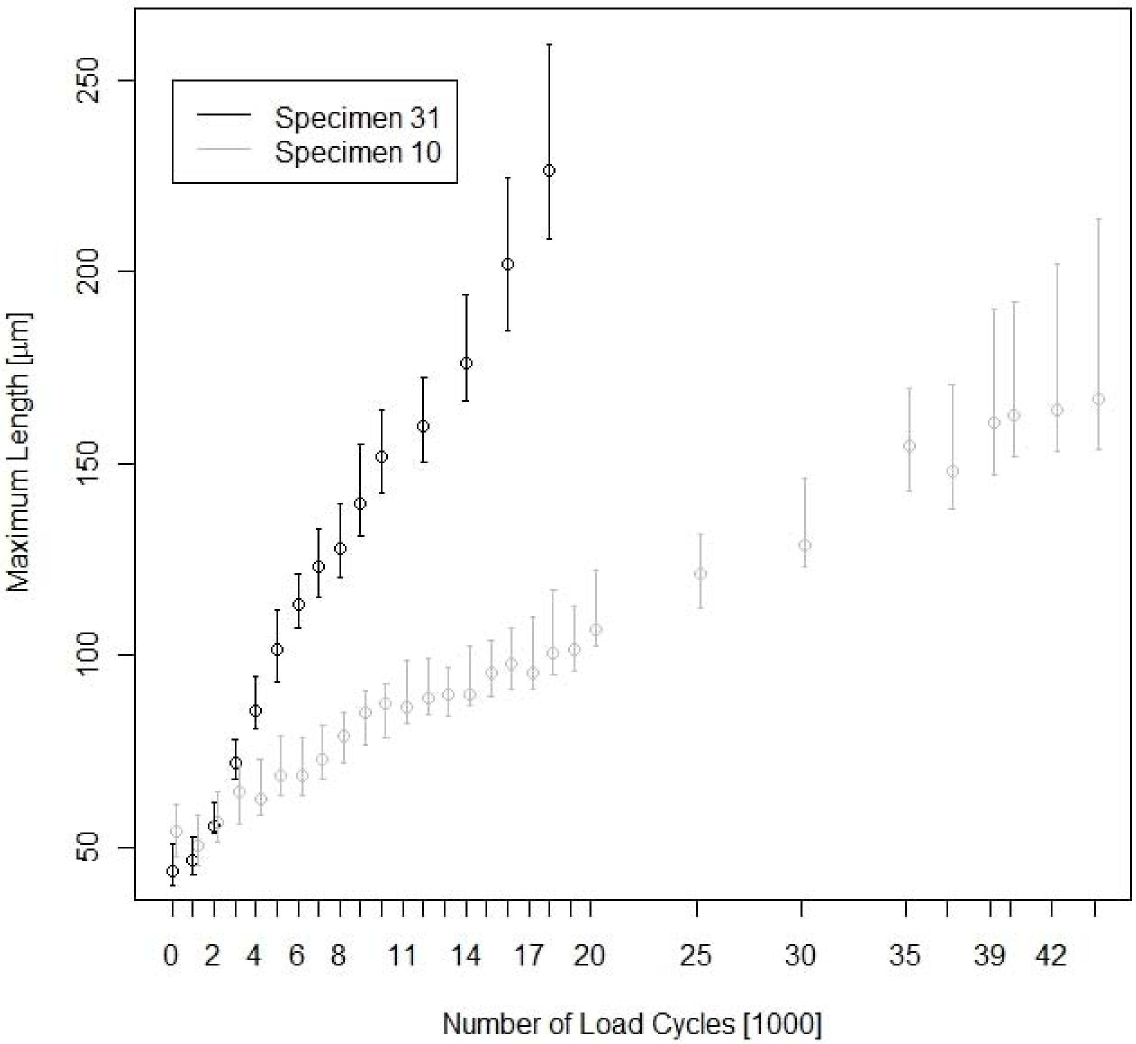

Fig. 14 95% Wilcoxon confidence intervals and medians for the maximum length of detected cracks (MaxL).

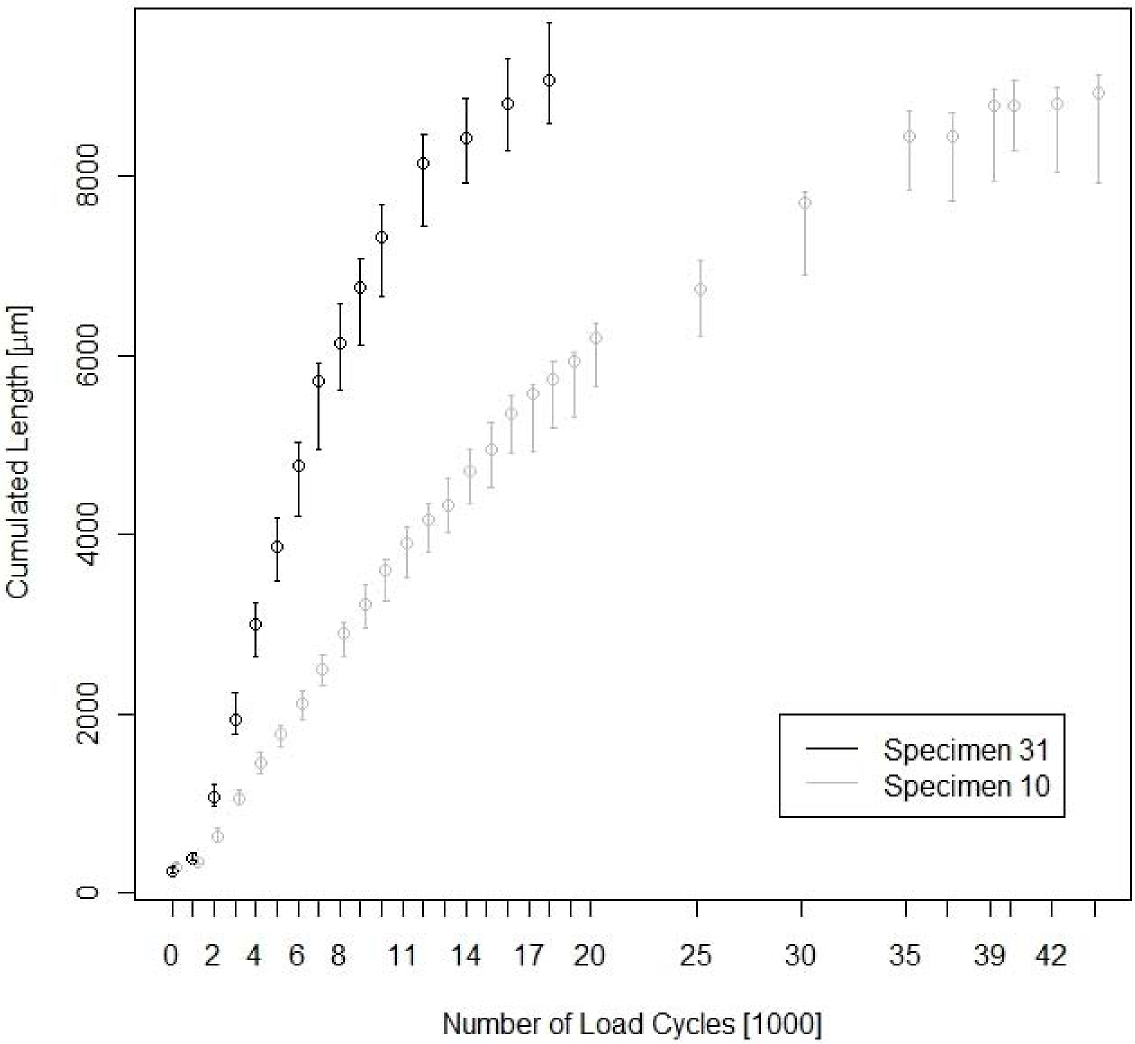

Fig. 15 95% Wilcoxon confidence intervals and medians for the cumulated length of detected cracks (CumL).

Table 1 P-values of the two-sample Wilcoxon test

	At the beginning	After 1 000 load cycles	After 2 000 load cycles	After 3 000 load cycles
Number of crack clusters (NoC)	0.0633	0.1145	<0.0001	<0.0001
Mean Length (MeanL)	0.7759	0.8578	0.0008	<0.0001
Maximum Length (MaxL)	0.0475	0.4500	0.9105	0.0179
Cumulative Length (CumL)	0.0537	0.1317	<0.0001	<0.0001
Fraction of Grey Values (FGV.Auto)	0.9411	0.0249	<0.0001	<0.0001
Fraction of Grey Values ≤ 140 (FGV.140)	0.8057	0.8688	0.5066	<0.0001
Mean Grey Value (MGV)	0.3893	0.7227	0.9860	<0.0001