

Detecting diamond breakouts of diamond impregnated tools for core drilling of concrete by force measurements

Christine H. Müller, Hendrik Dohme, Dennis Malcherczyk, Dirk Biermann, and Wolfgang Tillmann

Abstract Diamond impregnated tools for core drilling consist of segments in which synthetical diamonds are bounded in a metal matrix. The wear of these tools depends on the time points when active diamonds breakout and new diamonds from deeper layers of the metal matrix become active. Up to now, these time points were measured only by visual inspection at given inspection time points, a measurement which is very error-prone and labour-intensive. Hence the aim is to use the automatic force measurements during the drilling process for detecting the breakouts of the diamonds. These force measurements consist of three time series observed over about 75 minutes, each minute with over 300 000 measurements. At first, we present here an approach of an analysis of these time series in three steps: identification of the time periods of active drilling, identification of the rotation periods, and determination of differences between successive rotations. Based on the detected rotation periods, 147 features for classification of minutes with and without diamonds breakout are created. Some of these features are based on the differences between successive rotations and some on p-values for testing the independence of the detected rotation lengths. After a feature selection step, random forest and logistic regression are applied. This leads at least for one of two considered series of experiments to a classification error which is smaller than the trivial classification error.

Christine Müller (responsible for statistical methods)
Department of Statistics, TU University Dortmund, e-mail: cmueller@statistik.tu-dortmund.de

Hendrik Dohme (responsible for statistical methods)
Department of Statistics, TU University Dortmund, e-mail: hendrik.dohme@tu-dortmund.de

Dennis Malcherczyk (responsible for statistical methods)
Department of Statistics, TU University Dortmund, e-mail: dennis.malcherczyk@tu-dortmund.de

Dirk Biermann (responsible for experimental setup)
Institute of Machining Technology, TU University Dortmund, e-mail: biermann@isf.de

Wolfgang Tillmann (responsible for experimental setup)
Institute of Materials Engineering, TU University Dortmund, e-mail: wolfgang.tillmann@udo.edu

1 Introduction

The wear of diamond impregnated tools for core drilling of concrete depends on the wear of active diamonds visible on the surface of the segments of the tool. The segments consist of synthetical diamonds bounded in a metal matrix. As soon as an active diamond breaks out, usually a new diamond embedded in a deeper layer of the metal matrix becomes active. This so-called “self-sharpening property” means that new sharp diamonds are exposed at the tool surface at any time of the process.

Several authors already identified the main wear mechanisms of diamond impregnated tools, e.g. [7], [19]. Authors as [3], [11], [15] considered also some statistical analysis. However, these approaches mainly concern diamond impregnated tools for sawing applications of rock. Only a few authors are dealing with the diamond core drilling process, see e.g. [1], [8], [9], [14]. In particular, [14] showed how the size of the diamonds and the used concrete influence the lifetime of the active diamonds. This analysis was complicated by the fact that the breakout times of the active diamonds and the appearance times of new active diamonds were only measured by visual inspections at given inspection times. Thereby, the number of visible and active diamonds on the tool surface was determined by microscopical inspections of the tool at the given points in time. This leads to so-called “doubly interval-censored” data. Moreover, the intervals between the inspection times lasted always one minute which is not the best choice as indicated by [13]. However, more grave is the fact that the visual inspections are very labour-intensive and so error-prone that different inspectors provided different results.

Hence an important aim is to detect automatically the time points of the breakout of active diamonds and the appearance of new diamonds. The automatic measurements of the process forces during the drilling process are especially appropriate for this task. Since in the given experimental setup, the force measurements are given by the intervals between the visual inspections, we consider the task to identify the intervals with and without diamond breakout via the force measurements. Each interval consists of three time series in x , y , and z direction of drilling, each with about 300 000 observations. A first attempt of classifying these intervals with and without diamond breakout was done in [8] by using simple features like classical and robust measures of location and scale of the force measurements in the intervals. Additionally, the number of bivariate change points was used by applying the method of [6] to two of the three time series. In particular, the number of change points in the intervals looked promising for the classification problem in a first series of experiments with 25 intervals. However, this result could not be confirmed using further 25 intervals, see [9].

The change point analysis suffers from the fact that there are additional oscillations within each rotation, see Figure 5. These oscillations vary over time. Hence, we consider here the approach to identify at first the rotations and then to measure the differences of the oscillations between successive rotations. However, in a first step, the time periods of active drilling must be identified automatically. Although this was done already in [8], even this task is challenging. Further challenges appear by identifying the rotations and by calculating the differences between the rotations.

In particular, to test the quality of the identification of the rotations, we test for independence of the detected rotation lengths with the runs test of [18] (see [4] pp. 78-86) and a new test based on the recently proposed generalized sign test of [12]. The differences between rotations are calculated by applying the method of dynamic time warping of [5].

We create 147 features from the identified rotations, the differences between rotations, the p-values of the independence tests, and additionally from more simple quantities per interval. After a feature selection step, we use the random forest and the logistic regression for classifying minutes with and without diamond breakout. We apply this for two series of experiments with two types of concrete. The experiments at the more homogeneous concrete provide a low breakout rate of 0.173 so that the trivial classification method which classifies all minutes as "no breakout" could not be beaten. However, the experiments at the inhomogeneous concrete show a breakout rate of 0.342 while the leave-one-out misclassification rate was 0.260 for the logistic regression and 0.329 for the random forest.

The paper is organized as follows. The experimental setup is given in Section 2. Section 3 deals with identification of periods of active drilling and Section 4 with the identification of the rotations. Section 5 shows how time warping can be used to calculate differences between rotations. In Section 6, the 147 features are given and used for the classification problem. Finally, a discussion of the results is given in Section 7.

2 Experimental setup and the data

Four sequences of experiments, each with 75 sequential drilling phases of (approximately) equal length, were conducted. In each drilling experiment, automatic force measurements were obtained in time intervals of length of 61 up to 83 seconds where each time interval should contain active drilling of about one minute length. During each interval, the process forces F_z , F_x , F_y in z , x , and y direction were measured with measurement frequency of about 5 000 Hz so that each process time series consists in average about 300 000 measurements per time interval of active drilling. The circumferential speed was 3.225 m/s leading to ca. 616 rotations per minute with ca. 487 observations per rotation. After each experiment, the number of diamonds which have been broken out and which newly appeared were recorded by visual inspections of photos obtained by a microscope. For more details of the experiments, see [8].

The four sequences of experiments differ by the size of diamonds (small diamonds from grid size of $d_k = 40/50$ mesh and large diamonds from grid size of $d_k = 20/30$ mesh) and two types of concrete (conventional concrete with compressive strength of C20/25 and homogeneous concrete with high strength of C100/115). However, only the sequences of experiments with the small diamonds provided enough intervals with diamond breakouts. These two sequences of experiments are called B28 and B29, where B28 concerns the drilling in the C20/25 concrete and

B29 the drilling in the C100/115 concrete. In the sequence B28, 22 diamonds were visible at the beginning and a diamond breakout was observed in 25 intervals while 35 diamonds were visible in the beginning of the series B29 and here 13 intervals showed a diamond breakout.

All calculations were done in R [16] and all cited packages are R packages available on the website given by [16].

3 Identification of periods of active drilling

Each single drilling experiment in the sequence of experiments should last about one minute before it is interrupted for the visual inspection. Because of these interruptions, there are phases of no active drilling at the beginning and at the end of each of the 75 time series so that the time intervals for each experiment are longer than one minute. Hence, identifying the phase of active drilling is a necessary first step. For this, the examination of the feed forces F_z in the time intervals is the most appropriate.

Figure 1 depicts the different phases of F_z in a time slot of 66 seconds, which are separated by the dashed, red lines. In the beginning of the drilling no forces are acting. As soon as the drilling tool comes into contact with the concrete workpiece, the feed forces start to rise until a stationary main phase is reached. At the end of the drilling, the segment is drawn back and the acting forces decrease rapidly.

The beginning and end phase of each drilling are not relevant for the analysis of diamond breakouts and should not be considered for statistical analysis. Nevertheless, breakouts might possibly occur in the phase of rising forces. But since the length and form of this phase varies a lot over the different experiments (especially in B28), including this phase would lead to a distortion of statistical properties of the drilling. Hence these phases also were excluded from further analysis, concentrating merely on the stationary phase.

An automatic detection of the phase of interest for each experiment was achieved by calculating the standard deviation in running windows of 100 observations throughout the time series of the feed forces F_z (see Figure 2) using the function `rollapply` from the `zoo`-package [20]. The mean of these windowed standard deviations was defined as μ_w and their standard deviation as σ_w . At first, the end of the stationary phase was detected by simply identifying the point in time when the windowed standard deviation of the feed forces takes a value above a threshold T_{End} for the last time. Here, T_{End} was set as $T_{End} = \mu_w + 0.15 \cdot \sigma_w$, where the factor 0.15 appeared to be most appropriate for the experiments B28 and B29.

Next, a threshold for the beginning of the stationary phase was defined. For this purpose, the standard deviation of 10 000 preceding observations before reaching the detected endpoint of the stationary phase was calculated and defined as σ_{main} . This represents the standard deviation during the main phase of the drilling. The threshold T_{Start} was set as $T_{Start} = \sigma_{main} + 0.5 \cdot \sigma_w$. Defining the beginning of this phase as the first point in time when the windowed standard deviations exceed this

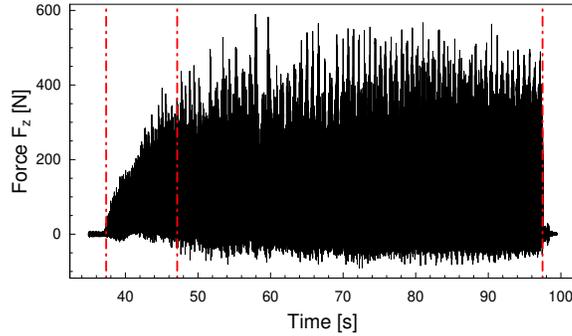


Fig. 1 F_z time series of the 25th minute of the B28 data with different phases separated by red lines.

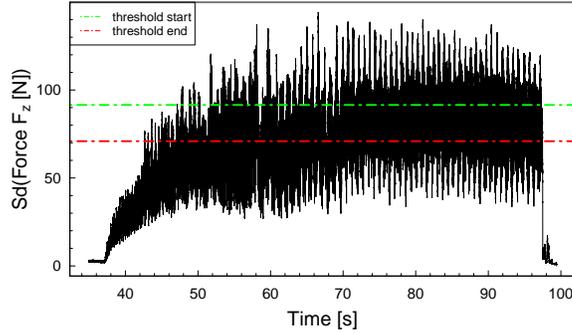


Fig. 2 Windowed standard deviations of F_z time series of Figure 1 with thresholds for the beginning and the end of the stationary phase.

threshold leads to satisfactory results for each experiment in the B29 as well as in the B28 data. However, some minutes of the B28 data provide very bad results. An ensuing inspection of those minutes shows that those minutes were badly affected by measuring disturbances so that they were removed ending up with 73 minutes.

In other drilling experiments the thresholds might require some adjusting which can be achieved by modifying the factors of the standard deviations σ_w and σ_{main} . In choosing the thresholds, it is particularly difficult to classify the beginning of the stationary phase since it varies a lot over the different experiments and minutes. Thus, it was important that the corresponding threshold is a function of the actual variance of this phase (here σ_{main}). Furthermore, if the threshold was defined too small some drilling periods contained big parts of the phase of the rising forces leading to distortion of statistical properties. Too large thresholds on the other hand

resulted in later onset points and a loss of information due to unnecessary short active drilling periods that were to be considered for further analysis. In extreme cases, no onset points can be detected.

4 Identification of the rotations

Since statistical methods will be applied for the single rotations of the drilling tool (eg. dynamic time warping), the second step is to identify the starting and endpoints of each rotation. To accomplish this, a local polynomial regression (loess) (see [2]) was run on the time series of the tangential forces F_x of the experiments since the periodic structure, which can be attributed to the rotations, is clearly visible in these time series (see Figure 3). For reasons of symmetry, using F_y would lead to similar results.

The loess method is used to smooth a time series by fitting a polynomial function to a neighbourhood $N(x_0, h) = [x_0 - h(x_0), x_0 + h(x_0)]$ of an observation x_0 from a time series $(x_t)_{t \in T}$, where h is a span function. The function `loess` in R uses a so called `span` parameter α , which defines the relative amount $p = \lfloor \alpha n \rfloor$ of n observations in the neighbourhood $N(x_0, h)$ of x_0 . The degree of the polynomial can be specified using the parameter `degree` and fitting is accomplished by using weighted least squares with a tricube weight function, see [2].

Here, `loess` was used to smooth F_x by using quadratic polynomials and a final `span` parameter of 0.0015 for the B28 experiments and 0.00125 for B29. The starting and endpoints of each rotation were then calculated by applying the differences operator Δ on the smoothed time series. Then the sign changes from negative to positive of these differences represent the minima of the smoothed curve. These minima are used as the onset points of a new rotation.

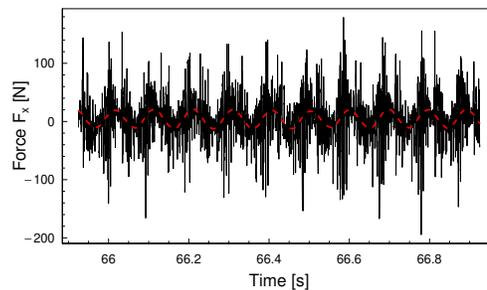


Fig. 3 Close up of the time series of tangential force F_x together with the smoothed time series by `loess` (red dashed line).

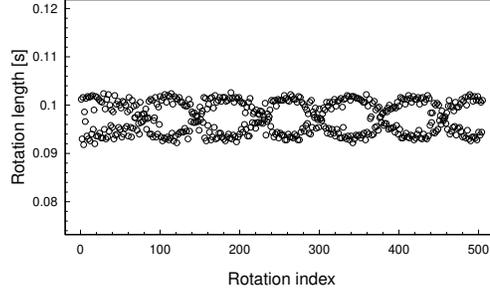


Fig. 4 Detected rotation lengths in the first minute of Experiment B29 using the span parameter $\alpha = 0.002$.

Note that too small span parameters lead to unrealistic short rotations which result in strong distortions of the computed features for classification. Too large parameters on the other hand result in alternating lengths of rotations. In particular, if a rotation length is determined as slightly too long by a too large parameter then the following identified rotation length is too short and vice versa. This leads to the chain structure of rotation lengths shown in Figure 4 and means that the rotation lengths are negatively correlated. Hence in choosing appropriate span parameters, it is necessary to compromise between correlated rotation lengths and the amount of unrealistic short rotations.

To test for correlation, we applied the runs test of [18] (see [4] pp. 78-86) and the generalized sign test of [12]. This generalized sign test is based on the 3-sign depth (3-depth) of residuals r_1, \dots, r_N defined by

$$d_3(r_1, \dots, r_N) := \frac{1}{\binom{N}{3}} \sum_{1 \leq n_1 < n_2 < n_3 \leq N} \left(\mathbb{1}\{r_{n_1} > 0, r_{n_2} < 0, r_{n_3} > 0\} + \mathbb{1}\{r_{n_1} < 0, r_{n_2} > 0, r_{n_3} < 0\} \right)$$

where $\mathbb{1}\{\cdot\}$ denotes the indicator function. Hence, the 3-depth is the relative number of 3-tuples with alternating residuals. Here, the residuals are the deviations of the rotation lengths from their median in the considered minute. The null hypothesis

$$H_0 : \text{The residuals are independent}$$

is rejected if the 3-depth is too small or too large. A too small 3-depth indicates positive correlation while a too large 3-depth indicates negative correlation.

Table 1 provides the rejection rates of H_0 by the 3-sign depth tests (3-depth tests) and the runs tests, as well as the 1%-quantiles of the rotation length, the number of unrealistic short rotations and the minimal rotation length over all minutes using different values for the span parameter α . A rotation is regarded as too short, when

B28					
Span parameter α	Rejection rate 3-depth test	Rejection rate runs test	1%-quantile of rotation lengths	Minimal rotation length	Number of short rotations
.00075	0.0000	0.8082	0.0048	0.0004	0
.001	0.0274	0.7945	0.0042	0.0008	7912
.00125	0.0685	0.6712	0.0918	0.0024	92
.0015	0.1096	0.6986	0.0918	0.0836	1
.002	0.4110	0.7397	0.0904	0.0854	0

B29					
Span parameter α	Rejection rate 3-depth test	Rejection rate runs test	1%-quantile of rotation lengths	Minimal rotation length	Number of short rotations
.00075	0.0267	0.6267	0.0046	0.0008	5817
.001	0.1200	0.5067	0.0052	0.0006	930
.00125	0.2933	0.7600	0.0934	0.0890	0
.0015	0.3067	0.7867	0.0928	0.0890	0
.002	0.4933	0.8000	0.0908	0.0888	0

Table 1 Rejection rates of the 3-depth test and the runs test of H_0 , 1%-quantiles of rotation lengths, number of short rotations and minimal rotation length for the experiments B28 and B29 over all minutes using α as span parameter for detecting the rotations.

its shorter than the median of rotation-lengths minus 3 times their IQR (interquartile range). Table 1 shows that the runs test always rejects the independence assumption in more than 50% of the minutes. The rejection rates of the 3-depth test are smaller but also increases with growing span parameter. These high rejection rates also occur when two successive rotations were put together. Hence, these rejection rates indicate to choose the span parameter as small as possible as soon as the number of short rotations is small enough. Based on Table 1, the choice of the span parameters (0.0015 for B28 and 0.00125 for B29) is comprehensible. Note, that the number of short rotations for the B28 experiment and span parameter 0.00075 is zero, because almost all rotations are predicted as too short. Furthermore for $\alpha = 0.0015$, only one rotation is considered as too short (with a rotation length of 0.0836). Nevertheless, this circumstance is negligible since its length is not unrealistic like it is the case for smaller span parameters.

An explanation for the different optimal values of the span parameters may be the different kind of concrete used. Here, the time series of the more inhomogenous concrete (B28) might require more smoothing. This results in a higher optimal span parameters because a wider window of local polynomial regression reduces the variance of the smoothed time series. The experiments with larger diamonds leads to the same relationship of the appropriate span parameters.

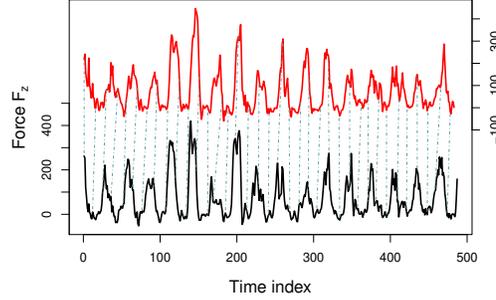


Fig. 5 Example of time warping applied to two successive rotations.

5 Calculation of differences between rotations

Differences between rotations can be calculated with the method of dynamic time warping as given by the R package `dtw` of [5], see also [17]. This method can be applied for time series $x = (x_1, \dots, x_N)$ and $y = (y_1, \dots, y_L)$ of eventually different lengths N and L as this can be the case for successive rotations. The idea of time warping is to find a warping of the time axis so that the distance between x and y becomes as small as possible. The minimized distance between x and y is then the so-called time warping distance. To define the distance between x and y , consider the matrix $M = (d(x_n, y_l))_{n=1, \dots, N, l=1, \dots, L}$ of pointwise distances between x and y where d is a given metric. Usually d is the Euclidean distance. Additionally, define a path through the indices of the matrix M by

$$\Phi : \{1, \dots, T\} \ni k \rightarrow \Phi(k) = (\Phi_x(k), \Phi_y(k)) \in \{1, \dots, N\} \times \{1, \dots, L\}$$

with $(\Phi_x(1), \Phi_y(1)) = (1, 1)$, $(\Phi_x(T), \Phi_y(T)) = (N, L)$ and $\Phi_x(k+1) \geq \Phi_x(k)$, $\Phi_y(k+1) \geq \Phi_y(k)$ for $k = 1, \dots, T$, where $T \in \{\max\{N, L\}, \dots, N+L-1\}$. The distance between x and y with respect to Φ is then given by

$$d_\Phi(x, y) := \frac{1}{T} \sum_{k=1}^T d(x_{\Phi_x(k)}, y_{\Phi_y(k)}).$$

The path Φ^* with

$$d_{\Phi^*}(x, y) := \min_{\Phi} d_\Phi(x, y)$$

provides then the smallest distance between x and y and $d_{\Phi^*}(x, y)$ is called the dynamic time warping (DTW) distance between x and y . Figure 5 demonstrate the principle of time warping using two successive rotations. It also shows the similarity of successive rotations.

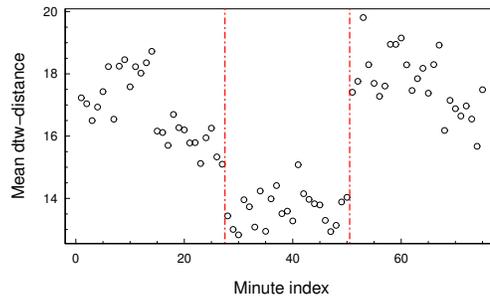


Fig. 6 Means of the dynamic time warping distances in the B29 data over all minutes with suspected changes of the workpiece marked by red lines.

Here, dynamic time warping was applied on the F_z time series of successive rotations in the periods of active drilling. In the case of a diamond breakout, it might be reasonably assumed that the drilling performance changes rapidly. Such events should be reflected in high time warping distances. Unfortunately, high peaks in the dynamic time warping distances appear several times in multiple minutes of the B28 and the B29 experiments and do not seem to relate directly with breakouts.

However, an interesting observation can be made when investigating the means and standard deviations of the DTW distances for each minute. In both experiments these features show obvious time dependent clusters. The minutes that mark the beginning of a new cluster can be associated with exchanges of the concrete workpiece which occurred at minute 25 and 51 (see Figure 6). These effects within the DTW distances are more pronounced in the B29 experiments. This might result from the more homogeneous concrete used which leads to more noticeable differences between workpieces.

Since some successive rotations show a different number of peaks (compare Figure 5), one might assume that the identification of rotations as described in Section 4 might not be optimal. Thus, it seems convenient to relax the constraint that beginnings and ends of two rotations have to match allowing a subsequence finding procedure. Applying this modification, which can be computed by setting the parameters `open.end=TRUE` and `open.beginning=TRUE` in the `dtw` function, should eliminate distortions of the DTW distances by suboptimally detected rotation onset- and offset-points. However, the results using this modification are hardly distinguishable from the ones obtained from the regular DTW method so that the classical DTW method is used hereinafter.

Table 2 Designations and descriptions of the generated features based on the single rotations of each minute which are only associated with one acting force.

Featurename	Description
DTWMeans	Mean value of the DTW-distances based on F_z
DTWSd	Standard deviation of the DTW-distances based on F_z
OutDTW	Number of outliers of the DTW-Distances based on F_z
CptsDTW	Number of change points in the DTW-distances between rotations based on F_z
pValueVZT	The p-value of the 3-depth test based on F_x
pValueRuns	The p-value of the runs test based on F_x
DifpValues	The differences of p-values of the correlation tests based on F_x

6 Feature generation and classification

In the following, various features for each minute of the B28 and B29 data were collected which should serve for the task of classifying a diamond breakout. The single rotations of the drilling tool, whose detection has been discussed in Section 4, form the basis of all these features. An overview of the 147 generated features can be found in Tables 2 and 3. Note, that location and scale parameters are given only by the mean and standard deviation in these tables although their robust counterparts (median and MAD, the median of the absolute deviation from the median) were calculated as well. Several features concern location and scale parameters as well as the number of outliers of parameters calculated from the measurements of single rotations. The parameters calculated for single rotations are again location and scale parameters and additionally the surface under the curve of a single rotation. The surface under the curve was approximated by trapezoids like it is usually done in numerical integration. A rotation has been classified as an outlier in a minute if its feature value is either bigger than the median of the feature plus 3 times its IQR or smaller than the median minus 3 times its IQR. This was done for the acting forces F_y , F_x , and F_z separately, but also two or three of the acting forces were treated simultaneously by calculating the euclidean distances from the spatial median. Since we in particular expected changes in the dynamic time warping distances (DTW distances) of rotations, especially outliers and change points of the DTW distances are considered. All features of change points were computed with the function `cpt.mean` of the `changepoint` package using the PELT method of [10] for the detection of change points. Since the independence assumption of the rotation lengths was often rejected by the runs test and the 3-depth test, their p-values and the differences of their p-values are considered as features as well. Features concerning all three forces F_y , F_x , F_z are given in Table 3. Table 2 contains features which make sense only for one force. These are features based on the time warping distances of rotations and the p-values of the independence tests.

Before the actual classification task, all of the 147 mentioned features in Tables 2 and 3 were put through a feature selection. In a first step, the LASSO method for logistic regression, which can be carried out with the function `cv.glmnet` of the `glmnet`-package, was applied for all observations. Here, the number of folds

Table 3 Designations and descriptions of the generated features based on the single rotations of each minute which are used for all acting forces F_x , F_y , F_z where the terms in brackets (X,Y,Z) are associated with the forces F_x , F_y , F_z .

Featurename	Description
MeanAbsMax(X,Y,Z)	Mean of the absolute maximal force values of each rotation
MeanMean(X,Y,Z)	Mean of the means of force value of each rotation
MeanFl(X,Y,Z)	Mean of the surface under the graph of each rotation
MeanSd(X,Y,Z)	Mean of the standard deviation of each rotation
SdAbsMax(X,Y,Z)	Standard deviation of the absolute maximal force values of each rotation
SdMean(X,Y,Z)	Standard deviation the means of force values of each rotation
SdFl(X,Y,Z)	Standard deviation of the surface under the graph of each rotation
SdSd(X,Y,Z)	Standard deviation of standard deviations of force values of each rotation
OutSd(X,Y,Z)	Number of outliers of the standard deviations of rotations
OutMean(X,Y,Z)	Number of outliers of the mean values of rotations
CptsMean(X,Y,Z)	Number of change points in the mean values of rotations
CptsSd(X,Y,Z)	Number of change points in the standard deviation of rotations
Max(XY,XZ,YZ)Mean	Maximal euclidean distance of the two-dimensional means of rotations from the spatial median
Mean(XY,XZ,YZ)Mean	Mean euclidean distance of the two-dimensional means of rotations from the spatial median
Out(XY,XZ,YZ)Mean	Number of outliers of the euclidean distances of the two-dimensional means of rotations from the spatial median
Max(XY,XZ,YZ)Sd	Maximal euclidean distance of the two-dimensional standard deviations of rotations from the spatial median
Mean(XY,XZ,YZ)Sd	Mean euclidean distance of the two-dimensional standard deviations of rotations from the spatial median
Out(XY,XZ,YZ)Sd	Number of outliers of the euclidean distances of the two-dimensional standard deviations of rotations from the spatial median
Max(X,Y,Z)SdMean	Maximal euclidean distance of the two-dimensional feature of the standard deviation and mean of rotations from the spatial median
Mean(X,Y,Z)SdMean	Mean of the euclidean distances of the two-dimensional feature of the standard deviation and mean of rotations from the spatial median
Out(X,Y,Z)SdMean	Number of outliers of the euclidean distances of the two-dimensional feature of the standard deviation and mean of rotations from the spatial median
MaxMeanXYZ	Maximal euclidean distance of the three-dimensional means of rotations from the spatial median
MeanMeanXYZ	Mean of the euclidean distances of the three-dimensional means of rotations from the spatial median
OutMeanXYZ	Number of outliers of the euclidean distances of the three-dimensional means of rotations from the spatial median
MaxSdXYZ	Maximal euclidean distance of the three-dimensional standard deviations of rotations from the spatial median
MeanSdXYZ	Mean of the euclidean distances of the three-dimensional standard deviations of rotations from the spatial median
OutSdXYZ	Number of outliers of the euclidean distances of the three-dimensional standard deviations of rotations from the spatial median

used in the crossvalidation was set to the number of observations which leads to a leave-one-out procedure.

For the B29-data only 4 important features were detected, namely `MaxXZMean`, `CptsMADZ`, `MADFIZ` and `MADMedianY`. It should be noted that, in the case of highly correlated features (which may be expected here due to the fact that all features were also computed with their robust counterparts), the LASSO algorithm only chooses one of those features which may lead to the small number selected. A useful tool to judge whether the choice of only four features is justified is the so called cross-validation curve which can be plotted for a `cv.glmnet` object. This curve shows that the choice of more than 4 features significantly increases the binomial deviance of the model and thus reduces its predictive power which can be seen by an obvious minimum in the curve. Using only those four identified features leads to a misclassification error of 0.186 which is above the "trivial" error of 0.173, that arises, when all observations are classified as "no breakout" and thus is no satisfactory result.

Using the same approach for the B28 data leads to the choice of the features `OutMADX`, `OutSdX` and `MADAbsMaxY` and a misclassification error of 0.369 when the trivial error is 0.342. Studying the shape of the cross-validation curve reveals that for the B28 data the selected number of features is not so clear. Here the binomial deviance does not significantly increase up to a number of 14 features. For that reason, it seems to make sense to identify more features which can contribute to the predictive performance of the classification.

For this purpose, a first approach was to use the integrated methods of the random forest. The `randomforest` package provides the possibility to compute the mean decreased Gini index and the mean decreased accuracy for each feature in the random forest based on its bootstrapping approach. This can be carried out by using the `VarImpPlot` function which plots the 30 most important features with corresponding importance values.

These plots show clearly visible gradations in the importances for both datasets. However, when trying to reproduce these results, which means constructing a new random forest, the important variables look quite different. For that reason it is not possible to identify a set of features which has systematically high importance so that the feature selection methods of the random forest do not seem to be very suitable for this classification task. A possible reason for this behavior might be the small sample size which could lead to problems concerning the bootstrapping approach of the random forest, paired with the high amount of features.

Therefore, to identify further important features, the crossvalidation-curves of the LASSO were considered once again. This time the maximal amount of variables that does not lead to a significant increase in the binomial deviance was used for the classification task. For B29, there is an obvious minimum in the binomial deviance curve for the mentioned 4 features so that this procedure does not lead to additional important features for the classification task. However, for B28, as already mentioned before, 14 features can be selected this way. These features are `MeanFIY`, `MeanAbsMaxZ`, `MADAbsMaxX`, `SdMeanY`, `SdMeanZ`, `MADAbsMaxY`, `MaxZSdMean`, `OutMedianXYZ`, `OutSdX`, `OutMADX`, `OutMADY`, `OutMeanX`, `OutMedianX`, `Out-`

MedianZ. Note, that these features include the three features identified in the first approach.

Based on the 4 and 14 detected features of the B29 and B28 data, respectively, a random forest and a logistic regression were constructed and the corresponding misclassification errors were computed by performing a leave-one-out cross-validation. Using the 4 features for B29 which are the same as before, the misclassification error for the logistic regression cannot be improved while it is even worse for the random forest with 0.213. Using the 14 features of the B28 data on the other hand, the misclassification error in the logistic regression managed to outperform the trivial rule by 6 correctly classified observations with a misclassification error of 0.260 while the random forest with a relative error of 0.329 also provides a slight improvement compared to the trivial error of 0.342.

7 Discussion

It was tried to classify minutes with and without diamond breakout by generating 147 features based on the ca. 616 rotations per minute. After a feature selection step, random forest and logistic regression were used for the classification. It turned out that the logistic regression is superior to the random forest for this problem. However, only for one of the two considered series of experiments, the misclassification error was smaller than the trivial classification. One problem could be that the diamond breakout happens before the used stationary part of active drilling which was identified in a first step of the analysis. Another problem was the questionable identification of the start and end points of the rotations which could cause features based on differences between rotations which achieve no good classification power. Surprisingly, these features were able to detect differences in the three concrete samples used in the series of experiments which were expected to behave similarly. This showed that in fact, the three concrete samples behaved differently. As a consequence, the three different concrete samples complicated the classification problem. Since only 25 drilling experiment were performed at each of the three concrete samples, the resulting 75 minutes are not enough to find a good classification rule so that more experiments are necessary. Moreover, a better identification of the start and end points of the rotations could improve the classification results. However, it may be too difficult to detect diamond breakouts via force measurements because of the inhomogeneity of the concrete.

Acknowledgments

The authors gratefully acknowledge support from the Collaborative Research Center “Statistical Modelling of Nonlinear Dynamic Processes” (SFB 823, B4) of the German Research Foundation (DFG).

References

1. Carpinteri, A., Dimastrogiovanni, L., Pugno, N.: Fractal coupled theory of drilling and wear. *International Journal of Fracture* 131:131–142, (2005)
2. Cleveland, W. S., Grosse, E.: Computational methods for local regression. *Statistics and Computing* 1:47–62, (1991)
3. Ersoy, A., Buyuksagic, S., Atici, U.: Wear characteristics of circular diamond saws in the cutting of different hard abrasive rocks. *Wear* 258:1422–1436, (2005)
4. Gibbons, J., Chakraborti, S.: Nonparametric statistical inference. Statistics, textbooks and monographs. Marcel Dekker Incorporated, (2003)
5. Giorgino, T.: Computing and visualizing dynamic time warping alignments in R: The dtw package. *Journal of Statistical Software*, 31:1–24, (2009)
6. Holmes, M., Kojadinovic, I., Quessy, J.-F.: Nonparametric tests for change-point detection à la Gombay and Horváth. *Journal of Multivariate Analysis* 115:16–32, (2013)
7. Hu, Y.N., Wang, C.Y., Ding, H.N., Wang, Z.W.: Wear mechanism of diamond saw blades for dry cutting concrete. *Key Engineering Materials* 304–305:315–319, (2006)
8. Kansteiner, M., Biermann, D., Dagege, M., Müller, C., Ferreira, M., Tillmann, W.: Statistical evaluation of the wear behaviour of diamond impregnated tools used for the core drilling of concrete. In: *Proceedings of the EURO PM 2017, Europe’s annual powder metallurgy congress and exhibition, Milan, 1st - 5th October 2017*, (2017)
9. Kansteiner, M., Biermann, D., Malevich, N., Horn, M., Müller, C., Ferreira, M., Tillmann, W.: Analysis of the wear behaviour of diamond impregnated tools used for the core drilling of concrete with statistical lifetime prediction. In: *Proceedings of Euro PM 2018, Europe’s annual powder metallurgy congress and exhibition, Bilbao, 14 - 18 October 2018*, (2018)
10. Killick, R., Fearnhead, P., Eckley, I.A.: Optimal detection of changepoints with a linear computational cost. *Journal of the American Statistical Association*, 107:500, 1590-1598, (2012)
11. Konstanty, J.S., Tyrala, D.: Wear mechanism of iron-base diamond-impregnated tool composites. *Wear* 303:533–540, (2013)
12. Leckey, K., Malcherczyk, D., Müller, C.H.: Powerful generalized sign tests based on sign depth. *SFB Discussion Paper 12/20*. <https://www.statistik.tu-dortmund.de/2630.html>. (2020)
13. Malevich, N., Müller, C.H.: Optimal design of inspection times for interval censoring. *Statistical Papers* 60:99–114, (2019)
14. Malevich, N., Müller, C.H., Kansteiner, M., Biermann, D., Ferreira, M. and Tillmann, W.: Statistical analysis of the lifetime of diamond impregnated tools for core drilling of concrete. In: *Applications in Statistical Computing - From Music Data Analysis to Industrial Quality Improvement*. Eds. K. Ickstadt, H. Trautmann, G. Szepannek, N. Bauer, K. Lbke, M. Vichi, Springer, 233–249, (2019).
15. Özçelik, Y.: Multivariate statistical analysis of the wear on diamond beads in the cutting of andesitic rocks. In: Xipeng, X. (Eds.): *Key Engineering Materials. Machining of Natural Stone Materials* 250:118–130, (2003)
16. R Development Core Team: R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria, <http://cran.r-project.org/>, (2019)
17. Tormene, P., Giorgino, T., Quaglini, S., Stefanelli, M.: Matching incomplete time series with dynamic time warping: An algorithm and an application to post-stroke rehabilitation. *Artificial Intelligence in Medicine* 45: 11–34, (2008)
18. Wald, A., Wolfowitz, J.: On a test whether two samples are from the same population. *The Annals of Mathematical Statistics*, 11(2):147–162, (1940)
19. Yu, Y.Q., Xu, X.P.: Improvement of the performance of diamond segments for rock sawing, Part 1: Effects of segment components. In: Xipeng, X. (Eds.): *Key Engineering Materials. Machining of Natural Stone Materials* 250:46–53, (2003)
20. Zeileis, A., Grothendieck, G.: zoo: S3 infrastructure for regular and irregular time series. *Journal of Statistical Software* 14:1–27, (2005)