Article

Using Multivariate Quality Statistic for Maintenance Decision Support in a Bearing Ring Grinder

Muhammad Ahmer [1,3,*], Fredrik Sandin [2], Pär Marklund [3], Martin Gustafsson [1] and Kim Berglund [3]

[1] Manufacturing and process development, AB SKF, Gothenburg, Sweden; muhammad.ahmer@skf.com; martin.gustafsson@skf.com
[2] EISLAB, Computer Science, Electrical and Space Engineering, Luleå University of Technology, Luleå, Sweden; fredrik.sandin@ltu.se
[3] Engineering Sciences and Mathematics, Luleå University of Technology, Luleå, Sweden; par.marklund@ltu.se; kim.berglund@ltu.se
* Correspondence: muhammad.ahmer@skf.com

Abstract: Grinding processes' stochastic nature poses a challenge in predicting the quality of the resulting surfaces. Post-production measurements for form, surface roughness, and circumferential waviness are commonly performed due to infeasibility in measuring all quality parameters during the grinding operation. Therefore, it is challenging to diagnose the root cause of quality deviations in real-time resulting from variations in the machine's operating condition. This paper introduces a novel approach to predicting the overall quality of the individual parts. The grinder is equipped with sensors to implement condition-based maintenance and is induced with five frequently occurring failure conditions for the experimental test runs. The crucial quality parameters are measured for the produced parts. Fuzzy c-means (FCM) and Hotelling's T-squared (T^2) have been evaluated to generate quality labels from the multi-variate quality data. Benchmarked random forest regression models are trained using fault diagnosis feature set and quality labels. Quality labels from the T^2 statistic of quality parameters are preferred over FCM approach for their repeatability. The model, trained from T^2 labels achieves more than 94% accuracy when compared to the measured ring disposition. The predicted overall quality using the sensors' feature set is compared against the threshold to reach a trustworthy maintenance decision.

Keywords: grinding; multivariate statistics; maintenance decision; condition-based maintenance; condition monitoring; health management; prognostics; fault diagnosis

check for updates

Citation: Ahmer, M.; Sandin, F.; Marklund, P.; Gustafsson, M.; Berglund, K. Using Multivariate Quality Statistic for Maintenance Decision Support in a Bearing Ring Grinder. *Preprints* **2022**, *1*, 0. https://doi.org/

1. Introduction

Grinding is a key process in bearing production. Being at the end of the process chain, it is crucial to avoid quality variations that can constitute producing scrap. The high demand for output productivity and fulfillment of various surface quality parameters makes the area of grinders and grinding process an active research field [1]. The changing machine conditions of the bearing ring grinder make it challenging to achieve a predictable process [2]. Despite the integration of several process monitoring techniques based on measurement of in-situ cutting forces, power, vibrations, etc, today's grinding process and machines struggle to produce parts with desired quality without manual intervention in setting up the process for the first time [3–5]. This variability of the process, in addition to the machine's maintenance condition dependency, requires an in-depth understanding and knowledge of the influence of the involved parameters and how the deviation in one affects the other [1]. This is especially valid when it comes to bearing production where the tolerances on the produced quality are kept very tight.

In any production system, apart from the operational process impacts, the machines and subsystems are subject to physical degradation [6,7]. To avoid unplanned downtime the industry focuses on predicting behaviors in equipment that can affect the process and undertaking actions to prevent failures [8,9]. The idea of machine fault diagnosis is to determine and classify the severity of an asset or its subsystem failure to achieve higher productivity and avoid catastrophic breakdowns which have a significant effect

 doi:10.20944/preprints202208.0283.v1

on maintenance costs [10]. Sophisticated maintenance strategies are thus considered and practiced for complex and advanced machines in today's manufacturing. Significant expenditure goes into maintenance programs where one-third to one-half is wasted due to ineffective maintenance [11]. To improve the maintenance effectiveness of machine systems affected by the stochastic nature of machining operations, a well-consulted fault diagnosis strategy with a maintenance decision support system is needed [12].

Condition-based maintenance (CBM) is the maintenance strategy of using sensors in machines for the purpose of monitoring, diagnosis, and prognostics to effectively achieve and plan cost-efficient maintenance while maintaining the uptime of the monitored assets [13]. The primary challenge is to predict the health state of the equipment through the use of sensor data with a level of certainty to accurately determine maintenance action points through effective reasoning on the remaining useful life (RUL) [14]. To achieve the level of certainty where the action can be taken, a perception has to be developed for the current state that can lead to the understanding of the failure as part of condition-based maintenance [11]. To anticipate the manifestation of the failure, as soon as possible, complex analysis methodologies have to be adapted to quantify the chance of the machine's operation without fault [15]. Despite that Machine learning (ML) approaches and methodologies in failure prediction through collected data for predictive maintenance (PdM) have been assessed several times [16], failure prognostics is still considered a less explored task due to its specific nature in relation to the process and equipment [17]. As a result maintenance decision-making becomes challenging where accuracy and robustness are crucial in making decisions [18]. Due to limitations in run-to-failure data that can be used in extrapolating machine conditions, the PdM is approached by obtaining labeled quality data and interpreting it. The use of these methodologies as maintenance decision support is an open issue due to the lack of annotations in such data [19,20].

The challenge faced in achieving CBM leading towards PdM for a bearing ring grinder is addressed in this work. In the previous works [2], the effective use of sensor data, belonging to both process control and condition monitoring, has been demonstrated for the purpose of failure diagnostics in the grinding machine. The proposed approach is to use predicted quality information in addition to failure mode classification to trigger maintenance action. From the intelligent fault diagnosis of the condition-based maintenance setup [21], the measured output quality is considered as the evidence to identify if the failure impacts the operational performance of the machine or subsystem. To prepare quality data annotations, two approaches have been explored for which respective regression learners are trained to predict the produced quality using the feature set extracted from sensor data. Repeatability and reliability, in terms of implementation, of explored approaches are considered to propose the preferred model of choice. A quality criterion, based on measured quality parameters, is also developed to verify and validate the overall quality prediction and performance quantification of the proposed approach. The combination of failure diagnostics and classification of produced quality gives a complete CBM setup through reliable and actionable maintenance decision support that is fundamental to adapting PdM strategy in a bearing ring grinder.

2. Method

The modeling of machine degradation, being a stochastic phenomenon, is extremely important for failure diagnostics and maintenance planning. Taking advantage of all the available information from health monitoring data is advantageous to precisely describe the extent of degradation. In this work, the Lidköping SGB55 grinder, shown in Figure 1, is equipped with a state-of-the-art real-time data acquisition and health monitoring system [2]. To enable early fault detection in the bearing production process, rough grinding is chosen as the process to be monitored and analyzed. In the CBM context, the maintenance strategy has to follow the implementation steps of *data acquisition*, *data processing*, and *maintenance decision-making*. The maintenance decision-making presented in this paper builds on the previous work on the development of intelligent fault diagnosis [21] and follows the steps as depicted in Figure 2. As shown in Figure 3, this work focuses on the severity estimation model and its support in maintenance decision-making for the bearing ring grinder.

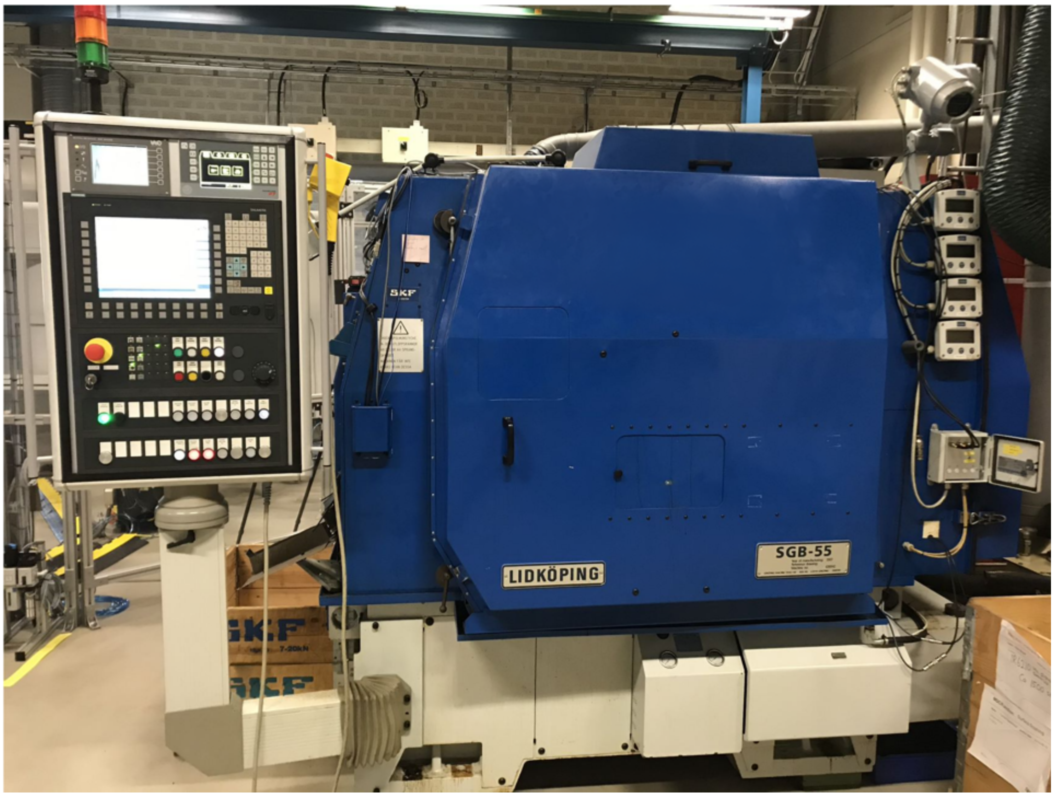

Figure 1. The Lidköping SGB55 bearing ring grinder was used in this investigation.

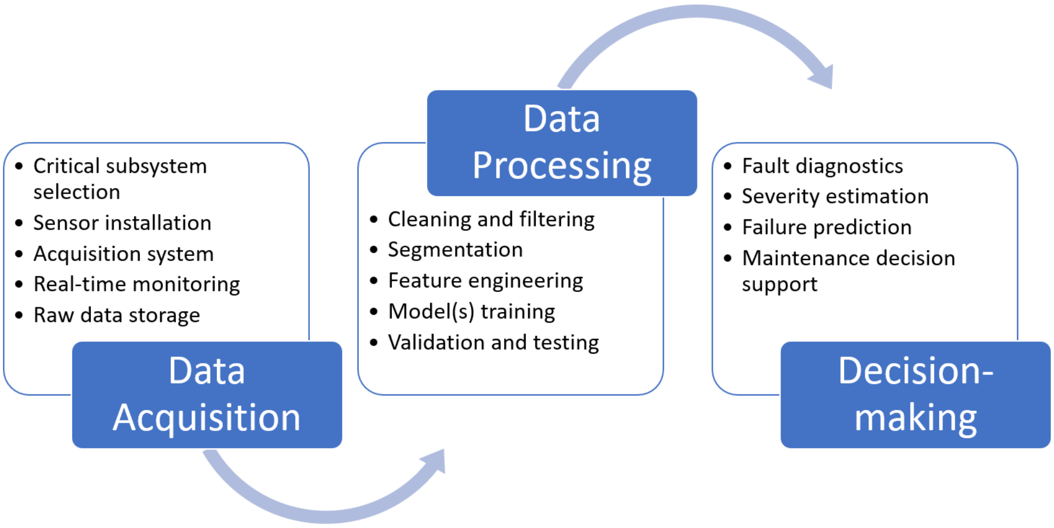

Figure 2. Implementation steps of CBM process for failure prognostics. Adapted from [10,13].

2.1. Data Acquisition

Knowing the maintenance history of the SGB55 grinder, critical subsystems i.e. *Grinding Slide assembly* and *Workhead assembly* are monitored with sensors installed on strategic locations [2]. The Figure 4 shows the schematics of the SGB55 grinder and its subsystems. The machine is equipped with sensors, listed in Table 1, for process control as well as additional sensors for condition monitoring. The sensor data is acquired using National Instruments Data Acquisition hardware and the LabView system. The data acquisition

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 16 August 2022 doi:10.20944/preprints202208.0283.v1

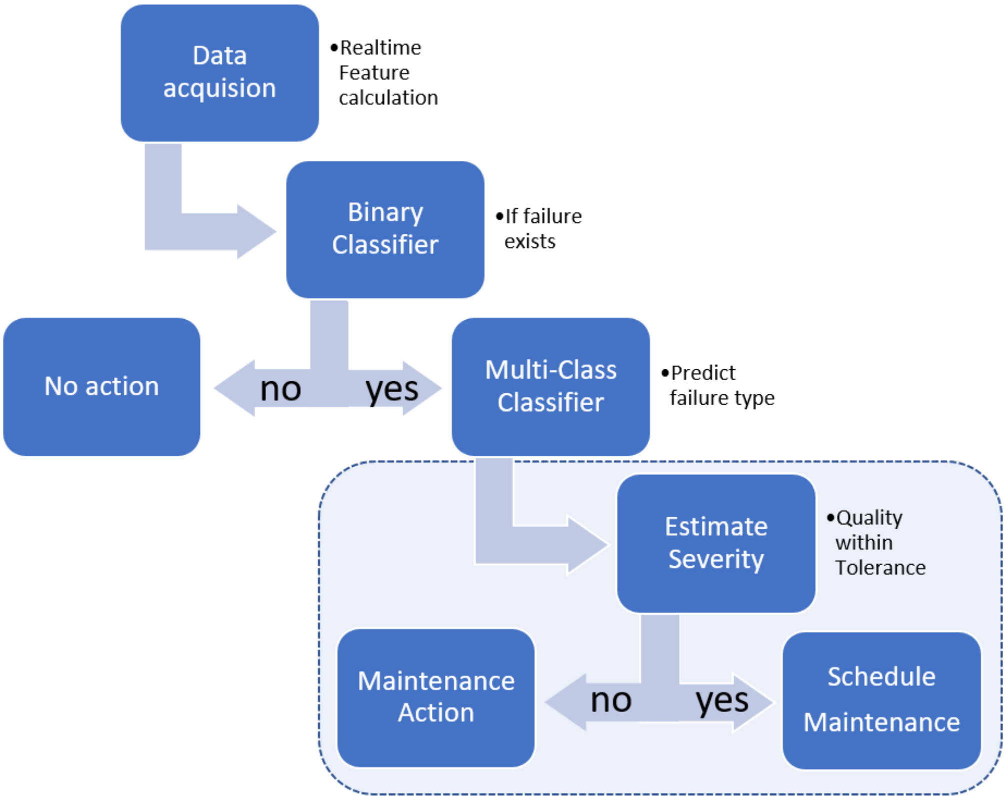

Figure 3. Failure prediction framework utilizing classification models for failure diagnostics and regression models as severity estimation for the prediction of produced quality. The bounding box represents the scope of this article.

system has the capability to simultaneously acquire and store sensor data in sync with the machine's cyclic operation. For each grinding cycle, the operational parameters are also stored in a database for each grinding cycle.

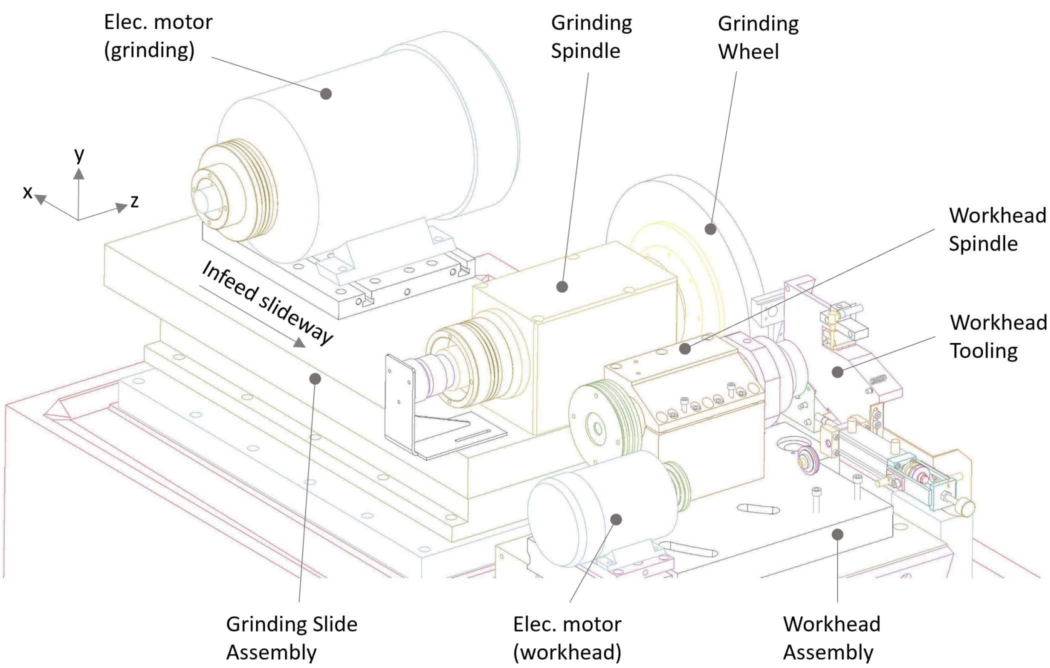

Figure 4. SGB55's critical subsystems for sensor installations as part of CBM implementation.

Table 1. List of sensors installed in SGB55 grinder.

Measured quantity	Sensor	Target subsystem
Force	Kistler 9105C	Workhead Assembly
Acoustic emission	Dittel m6000	Grinding spindle
Power	Montronix PS100	Elec. motor (grinding)
Strain	Kistler 9238B	Workhead Assembly
Acoustic emission	Parker U247	Workhead Tooling
Vibration	PCB triax A45	Workhead Tooling
Vibration	SKF CMSS2200	Elec. motor (grinding)
Vibration	SKF CMSS2200	Elec. motor (workhead)
Vibration	IMI601A01	Grinding Spindle
Temperature	NTCALUG02A103F	Grinding Spindle
Temperature	NTCALUG02A103F	Workhead Spindle
Temperature	NTCALUG02A103F	Workhead Tooling

2.1.1. Test and Measurement Criteria

A grinding cycle [22] consisting of a *roughing stage* and *spark-out stage*, shown in Figure 5, is programmed to grind the rings. The *Roughing stage* removes the material to reach the desired dimension of the rings and the *Spark-out stage* influences the final quality parameters that are being measured at the end. Since the data is acquired w.r.t. the individual grinding cycle, all the ground workpieces for the tests are saved as well. Failure modes are introduced in the selected subsystem components to simulate failures during production. To collect enough statistical data, each test is run for 7 dressing intervals where the grinding wheel is refreshed after each interval. 15 rings are ground in each dressing interval which gives a total of 105 rings for each test and produces 735 rings for all the tests. Figure 6 maps the operating conditions for each type of test and the corresponding rings produced in each test interval. It is infeasible to measure every ring produced using standard equipment. Hence a subset of the produced rings is chosen to be measured for the quality parameters, e.g. form, surface roughness, and waviness, as listed in Table 2. As shown in Figure 6, ring numbers 1, 3, 7, and 15 from each dressing interval are measured for the quality parameters to evaluate the quality being produced during each test run. In addition to the measured quality disparity between different tests, the choice of rings allows capturing the quality variations not only between dressing intervals but also within the dressing interval of a test.

Table 2. Measured Quality Parameters.

Quality Parameter	Equipment Used
Relative Diameter	Dial gauge (referenced)
Roundness Waviness	SKF MWA 160D
Surface Roughness Form	Taylor Hobson - Form Talysurf

2.2. Data Processing

MATLAB® is used for data and signal processing where the data is accessed from the network storage and databases and is cleaned and filtered before further processing. Each cycle is divided into segments as shown in the Figure 5 where the *Idle segment*, the *Steady Grinding segment*, and the *Spark-out segment* are isolated from each sensor signal for further processing in feature extraction. To be able to estimate overall ring quality using grinding cycle data, the *Spark-out segment* is selected for feature engineering. The table 2 lists the selected main quality parameters derived from quality measurements. Extreme data points resulting from measurement errors can skew the analysis, therefore the quality parameters are cleaned and verified before further processing.

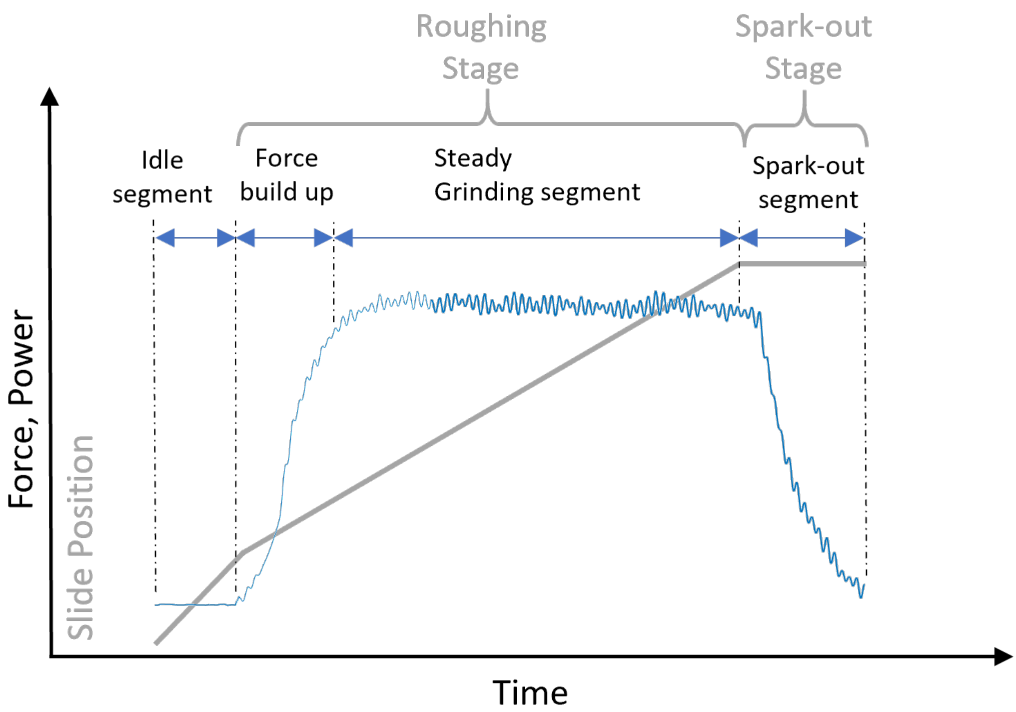

Figure 5. Grinding cycle to produce parts during tests. The grey line represents the grinding slide position as it moves into the workpiece. Blue represents the resulting in typical force/power signals and the segments that are extracted from each sensor signal for every cycle.

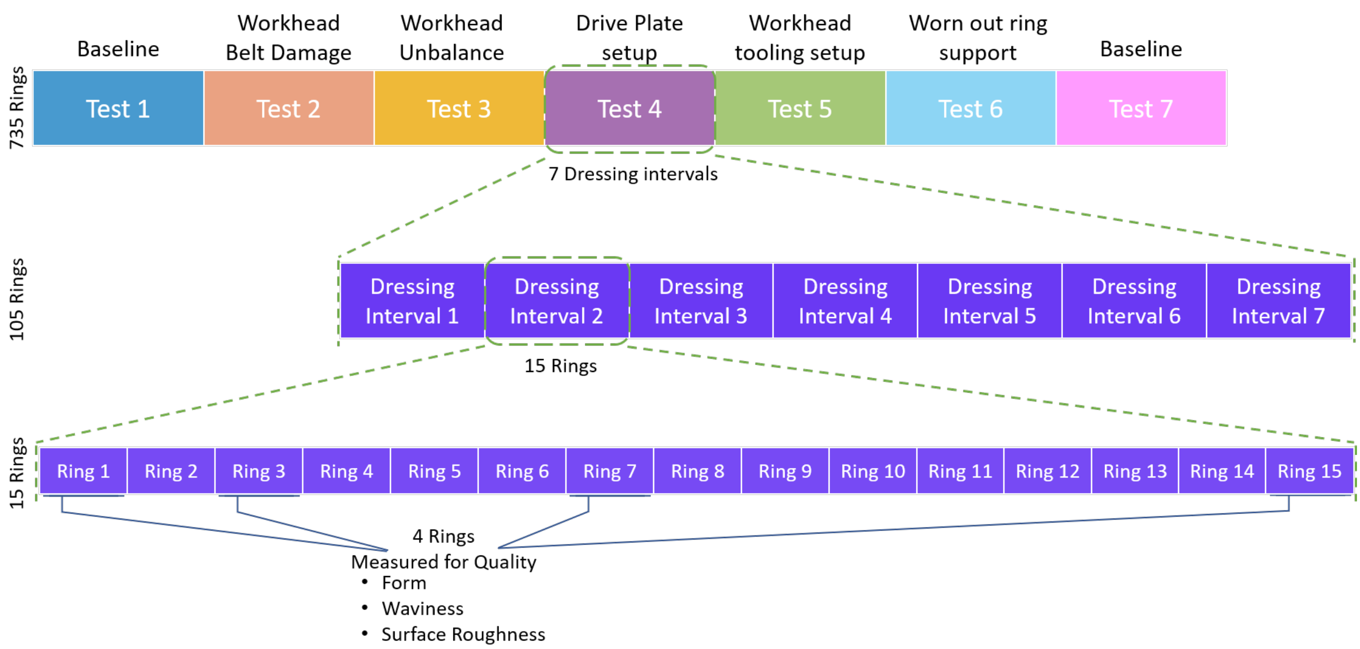

Figure 6. The figure presents the failure mode tests where the *Test* 4, as an example, is expanded to depict the test procedure with dressing intervals and the rings in each of the dressing intervals. This gives a total of 105 rings produced for each test.

2.2.1. Feature Engineering

After initial data processing, statistical features [23] are extracted from time and frequency domain components of the sensor data for the selected segment. The 9 main quality

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 16 August 2022 doi:10.20944/preprints202208.0283.v1

parameters from measured *Form*, *Surface roughness*, and multiple bands of *Circumferential waviness* are selected as quality features. The quality data is mean normalized per feature for the measured rings according to

$$z = \frac{X - \mu}{\sigma}, \tag{1}$$

where X is the set of *observations*, μ is the *mean* of X and σ is the *standard deviation* of X. MATLAB's Principal Component Analysis (PCA) algorithm is used to calculate principal components through singular value decomposition (SVD).

2.2.2. Model Development

To predict the overall quality as part of the quality severity model, two annotation approaches are used to prepare labeled data from measured quality parameters.

Approach 1

In the first approach, top 6 principal components are used where the data from 9 quality parameters of measured parts are transformed using the principal component score of the PCA method in MATLAB. The transformed data is then used to perform fuzzy c-means (FCM) clustering, which is an unsupervised clustering approach, using the parameters given in Table 3. The parameters are modified from default values to account for the possible clusters in the quality data with the reduced provision of overlap of the learned clusters. The FCM clustering method allows each data point to belong to multiple clusters with varying degrees of membership. The cluster where the baseline tests, 1 and 7 get higher membership probability is used as the reference cluster for the acceptable quality and a label for the training data set. The regression model is then trained using the sensor signal feature set to estimate the probability of membership of the output quality to the reference cluster.

Table 3. Optional parameters used for the fuzzy c-means clustering algorithm in MATLAB.

Option	Value
Nr. of clusters	4
Fuzzy overlap	1.1
Max. Iterations	50
Min. obj. improvement	0.001
Info Disp. flag	False

Approach 2

In the second approach, the T^2 statistic given by the PCA in MATLAB provides the statistical measure of the multivariate distance of each observation, i.e. ring quality data, from the center of the dataset. The PCA function also supports an output method of Hotelling's T-Squared Statistic (T^2) for the input data according to

$$T^2 = n(x - m)'(cov)^{-1}(x - m), \tag{2}$$

where x belong to feature set of observations X, m is the *distribution mean* of X, $(x - m)$ is the *vector distance* of an observation point x from m and $(cov)^{-1}$ is the *inverse covariance matrix* of X. The PCA method uses all the principal components to compute the T-squared statistic such that it is computed in full feature space. The T^2 statistic received for the measured rings becomes the label and is used in training a regression model using the feature set from sensor signal data. The trained model estimates and thus populates a T^2 control chart for the ring produced in each grinding cycle given the feature set from the cycle data.

To select the regression learner, MATLAB's regression learner app is used to train different models ranging from linear and support vector machines to regression trees and random forest. The feature set used to benchmark models originates from the sensor data as mentioned previously in this section. The models are trained separately with quality labels from both approaches including the labels for baseline tests. The random forest regression learner with default hyper-parameters, listed in Table 4, came out to be the top performer in this bench-marking. Hence the random forest regression model is trained further to be the selected overall ring quality estimator.

Table 4. Hyperparameters used for training of "Fit ensemble of learners for regression" in MATLAB.

Parameter	Value
Method	Least-squares boosting
Number of ensemble learning cycles	30
Weak learners to use in ensemble	Decision tree template
Minimum observations per leaf	8
Learning rate for shrinkage	0.1

2.3. Decision-making

The significance of any maintenance strategy is reflected through the accuracy and reliability of the maintenance decision-making. The random forest regression learners estimate the overall quality output of individual grinding cycles in terms of predicting the produced quality parameters as a multivariate statistical measure. Failure diagnostic is an important first step which is achieved from random forest classifiers, trained on the feature set from sensor signal data, to predict if the failure exists and the type of failure mode in the acquired data. Once the existing failure mode has been identified, the overall produced quality is predicted.

The random forest regression model from the first approach, trained using data from the FCM clustering method, estimates the probability of the output quality belonging to the reference quality cluster. A pre-selected threshold is used to trigger the maintenance action if the probability falls short of the threshold indicating the failure mode to be severe as the quality reaches the unacceptable limit. This method relies on the accuracy of the learned cluster membership used as a label to train the regression model as well as the selection of the threshold. The FCM clustering, being an unsupervised learning methodology, adds uncertainty to the decision-making as the training of the regression model relies on the learned clusters being representative of the failure and reference classes. The threshold itself adds another dimension that needs optimization based on knowledge and validation through quality measurements.

As for the second approach where the random forest regression model is trained using Hotelling's T-squared statistic, it estimates the T^2 statistic by taking feature set input from sensor data of individual grinding cycles. The T^2 statistic can be used to populate the control chart where the quality deviation can also be visually monitored against an upper control limit (UCL). The estimated T^2 statistic is compared against the UCL to trigger the maintenance action if the T^2 value exceeds the limit. The UCL is calculated based on the data from the baseline tests 1 and 7 according to

$$UCL = \left(\frac{p(k-1)(n-1)}{kn-k-p+1} F_\alpha [p, (kn-k-p+1)] \right), \tag{3}$$

where $(1-\alpha)100\%$ is the *confidence level*, n is the *size of the sample set*, k is the *size of the subgroup*, p is the *degrees of freedom* and F_α is the *F-statistic* at α. Note that the UCL does not depend on the T^2 values calculated for the sample set. Keeping the confidence level $(1-\alpha)$ less than 100% reduces the chances of false positives in the control chart. This comparison for making the maintenance decision has a dependency on the measurement data itself for the calculation of UCL which acts as the threshold for the predicted quality. Thus it is more reliable and repeatable than the clustering approach where the probability of the learned clusters differing in every iteration is higher. Also, the variation in the incoming data will

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 16 August 2022 doi:10.20944/preprints202208.0283.v1

have a different distribution of quality parameters which will influence the cluster learning significantly.

2.3.1. Predicted quality validation criteria

The rough grinding considered in this work is the intermediate step in the bearing ring production. Therefore, the produced parts are measured from the tests according to Figure 6. To verify the performance of the proposed model that estimates the overall ring quality, criteria to classify produced rings to be within specifications are defined. Since the T^2 multivariate statistic does not signify variations in the individual quality parameters, entire quality data from the baseline test is taken as a reference. Individual quality parameters for the measured rings are categorized as within specifications if they fall in the 20% of the mean of the parameter of their respective ring number in the baseline test. For a ring to be considered of acceptable quality, at least 4 of the 9 quality parameters are to be within specifications. The pseudo-code for this criteria calculation is presented in the algorithm 1. This results in the individual quality parameter to be quantified as $Rpq = 0$ if within the range and $Rpq = 1$ if outside the 20% range. Hence, the rings get classified as either accepted $(Rq \leq \frac{4}{9} \mid Rq \rightarrow 0)$ or rejected $(Rq > \frac{4}{9} \mid Rq \rightarrow 1)$ as per the proposed quality criteria. Thus the ground truth of the measured rings allows the validation of the output of the severity model predictions from the test dataset.

Algorithm 1 Calculating acceptable quality criteria based on the measured quality parameters.

Input:
Q : Measurement data
P_q : List of measured quality parameters
T : List of Tests
D : List of Dressing cycles in T
R : List of ring numbers measured in D
ref : Baseline test

```
for i ∈ R do
    for j ← 1 to length of P_q do
        q = (1/n) Σ_{k=1}^{n} Q_{(ref)kj}  where k ∈ D
        for k ← 1 to length of D do
            for l ← 1 to length of T do
                if q − 0.2 ∗ q < Q_{ijk} < q + 0.2 ∗ q then
                    Rpq_{ijkl} = 0
                else
                    Rpq_{ijkl} = 1
                end if
            end for
        end for
    end for
end for
Rq = (1/n) Σ_{k=1}^{n} Rpq_k  where k ∈ P_q
```

Output:
Rq : Classified rings as per measured quality

3. Results and discussion

This paper is the extension of the previously proposed failure classification framework [2] to include quality as depicted in Figure 3. As explained in Section 2.1, data from the installed sensors and machine's operating parameters are acquired simultaneously for each grinding cycle from which the statistical features are extracted after filtering and segmentation. The selected feature set for the failure classification results in greater than 98% accuracy, for both the binary and the multi-class failure mode classifier. The intelligent

 doi:10.20944/preprints202208.0283.v1

fault diagnosis along with published dataset and feature and sensor selection for failure mode classification are covered in the implementation of the CBM setup for the bearing ring grinder [21]. To include quality data in the analysis, the produced rings from the experimental test runs are measured as described in the Section 2.1.1.

The Figure 7 shows the box plot of surface roughness and circumferential form measurement of all the measured rings. It is to be noted that different types of failure modes will affect the measured parameters differently as evident for the two quality parameters, *Surface Roughness (Ra)* and *form*, as shown in the Figure 7. The PCA from the 9 measured quality parameters of the produced parts is shown in Figure 8. Although more than 60% of the variance is explained by the first two principal components, it is not enough to separate all the test classes. From the Figure 8, the tests 1, 2, 3, and 7 seem to overlap. Since tests 1 and 7 are reference baseline tests, they are bound to be closer to each other in the hyperspace. Due to less severity of the failure modes 2 and 3, statistically, it is possible to produce parts within tolerance. At this early stage of material removal in production, it will not be possible to identify quality variations resulting from these failure modes due to the limited number of in-line quality parameters being measured.

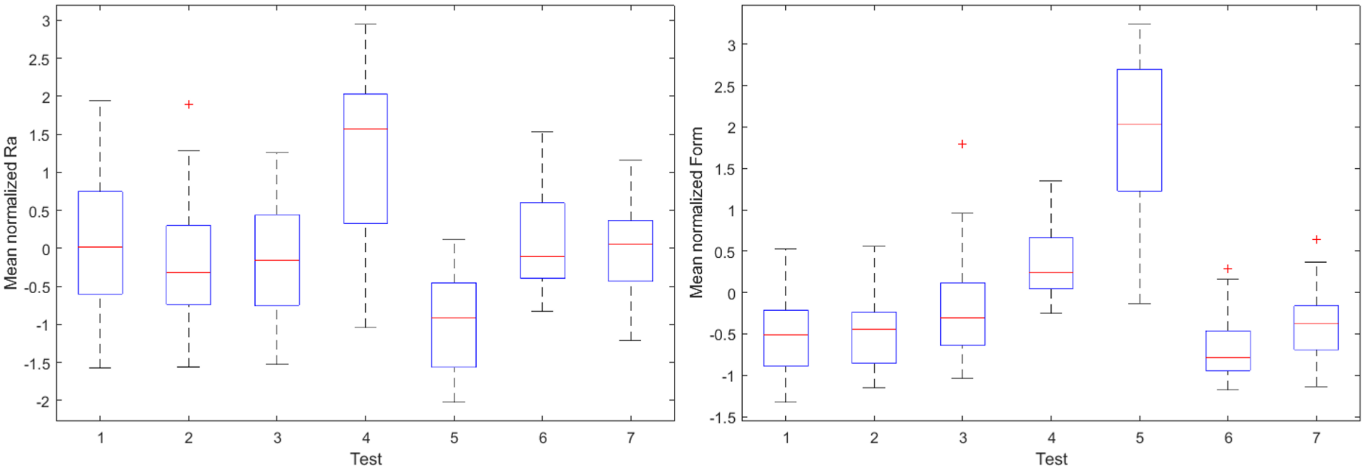

Figure 7. Box plot of Mean Normalized Surface Roughness on the left and circumferential form measurement on the right. The box plot for each test $1-7$ includes measured rings from all the dress cycles

As explained in Section 2.2.2 for the Approach 1 using the fuzzy c-means clustering algorithm, the top 6 principal components are used to learn 4 quality clusters in MATLAB using the parameters listed in the table 3. The 4 number of clusters better explained the variations before the clusters within a class start to appear. Examining the raw quality data from a domain expert perspective also suggests the close existence of the test classes as shown in Figure 8. The learned cluster centers, in Figure 9, are represented in the first 3 principal component dimensions as learned from the FCM clustering algorithm. The resulting allocation of each test class for the 4 centers is depicted in a stacked bar plot in the Figure 10. Since FCM is based on optimization, the cluster allocation varies for each run of the algorithm which affects the repeatability of the results in this approach. From the Figure 10, it is evident that the cluster center 4 is the reference cluster and is used as the quality label to train the regression model.

The selected regression model is the random forest as per MATLAB's regression learner app benchmark figures presented in the table 5. Thus the center 4 data, representing the degree of membership for each measured observation, is used as a label to train the random forest regression model with an achieved Root Mean Square Error (RMSE) of 0.28 out of the max scale of 1. The low RMSE gives a good predictor which is evident from the Figure 11 where the test data from tests 1 and 7 are estimated to be belonging to

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 16 August 2022 doi:10.20944/preprints202208.0283.v1

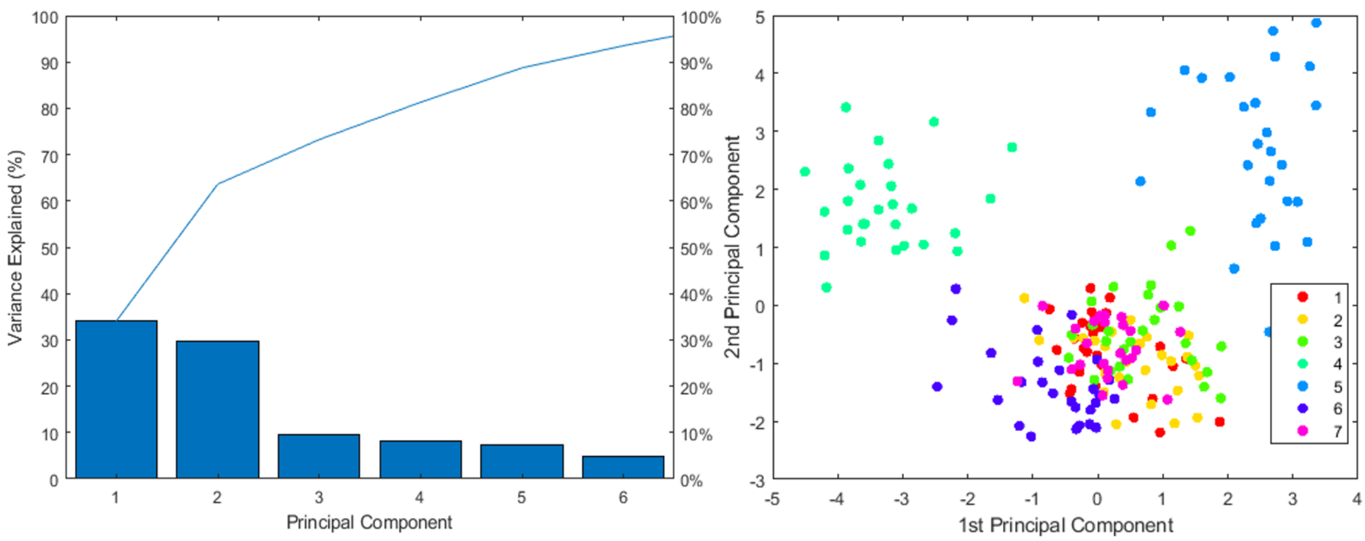

Figure 8. The Pareto plot on the left shows more than 95% of the variance explained by 6 principal components. On the right, the scatter plot for the first two principal components shows the separation of the test classes $1-7$ as per Figure 6.

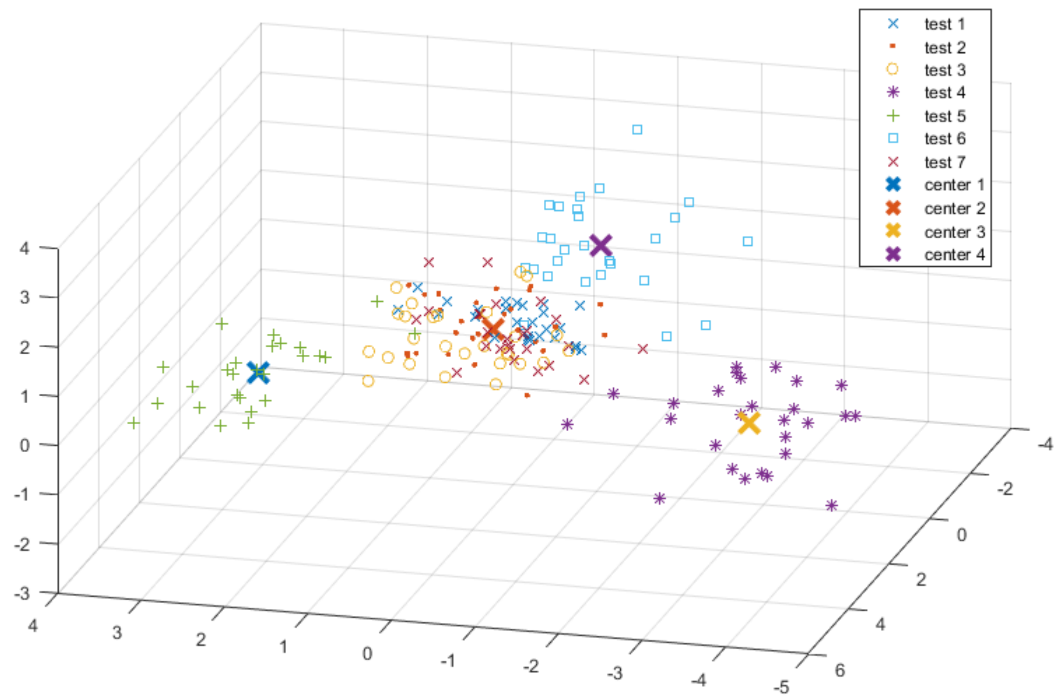

Figure 9. The 3-D plot represents the cluster centers where axes are the top principal component axes from PCA.

the reference cluster. The uncertainty in cluster learning and cluster center membership allocation makes it difficult to repeatedly use the method for reliable predictability. The significant difference between the benchmark and the presented first approach's RMSE results from FCM-based labeled data is evidence of the inherent fuzzy behavior of the algorithm directly influencing the performance accuracy.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 16 August 2022 doi:10.20944/preprints202208.0283.v1

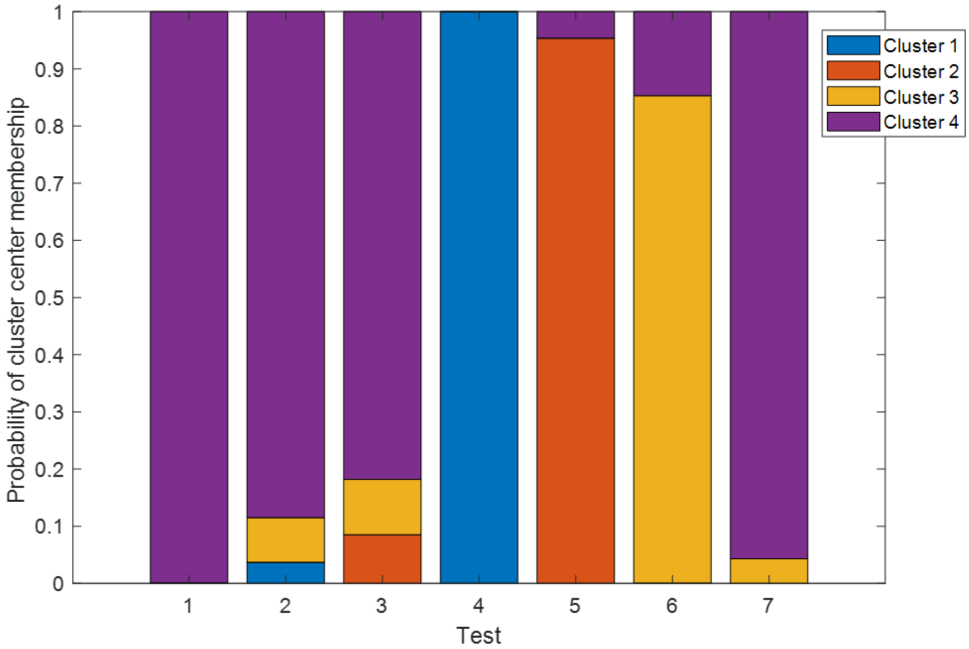

Figure 10. Probability or degree of cluster center membership to the individual quality observation of the respective test classes in 6 PCA dimensions.

Table 5. Bench-marking of regression models in MATLAB's regression learner app. Quality Label columns represent the achieved Root Mean Square Error (RMSE) for the labels used from the two approaches in training the models.

Regression Model	Quality Labels 1	Quality Labels 2
Random Forest	0.358	4.960
Regression Tree	0.414	5.743
Support Vector Machine	0.544	6.184
Linear Regression	1.755	11.065

In contrast, the Approach 2 of Hotelling's T-squared statistic is less uncertain due to repeatability and reliability based on the available data. The T^2 statistic values achieved from the output of the PCA algorithm in MATLAB are used as labels to train the random forest regression learner using default hyper-parameters and only modifying the ones listed in the table 4. The training of the regression learner results in an RMSE of 5.19 out of max scale T^2 value of 26.62 which, in terms of error rate, is lower than the RMSE of Approach 1, i.e. 19% $<$ 28%. The UCL calculated from the training quality dataset using the equation 3 becomes 9.1 with the confidence level of 97% obtained from $\alpha = 0.027$. The training dataset when compared against the UCL is shown in Figure 12.

The regression model estimations on the test set of sensor feature set are shown in the Figure 13. The UCL is shown as a dashed line and serves as a threshold above which the quality becomes unacceptable. Even though the test set, comprising of feature set from the selected segment of the sensor data, is larger than the entire quality dataset, the results follow the trend of the training set of measured quality. It is only fitting to verify the regression model predictions using absolute truth which is the measured ring quality itself. The Box plot in the Figure 14 depicts the classification of rings as accepted or rejected based on the criteria defined in Section 2.3.1. Note that the criteria result in the classification of individual rings based on all the measured quality parameters. Hence, the measured rings that end up out of spec according to quality criteria are represented as red circle markers in the Figure 13. Most of the markers being above the UCL line of the plot verify

 doi:10.20944/preprints202208.0283.v1

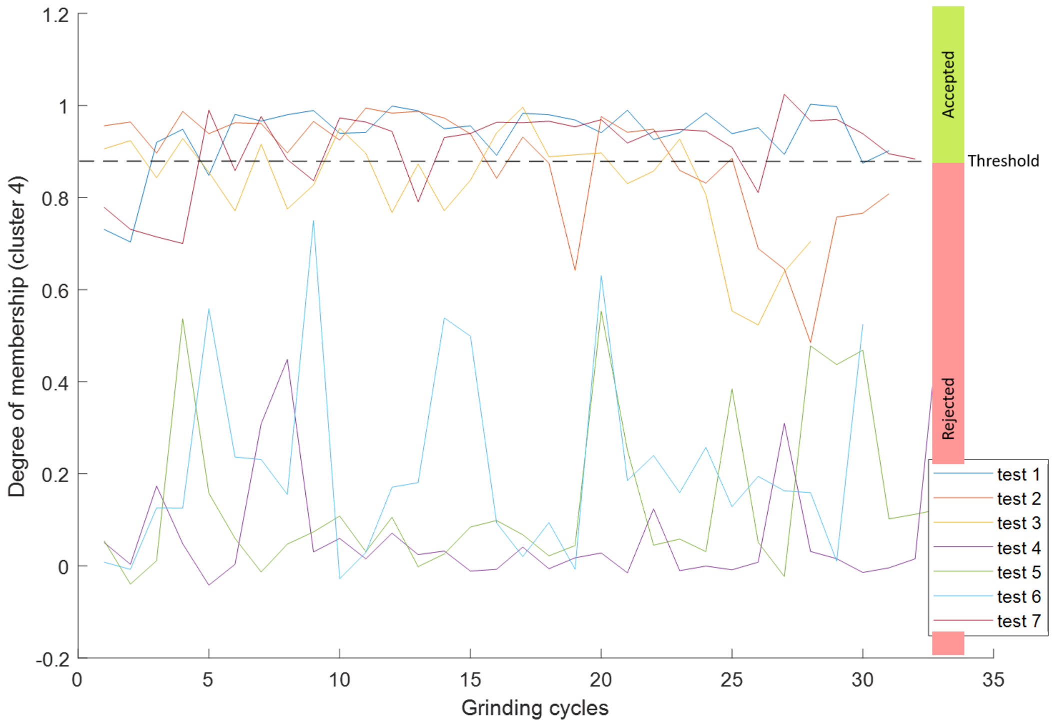

Figure 11. Predicting the overall ring quality produced in test data. The dashed line represents the threshold above which the quality is acceptable.

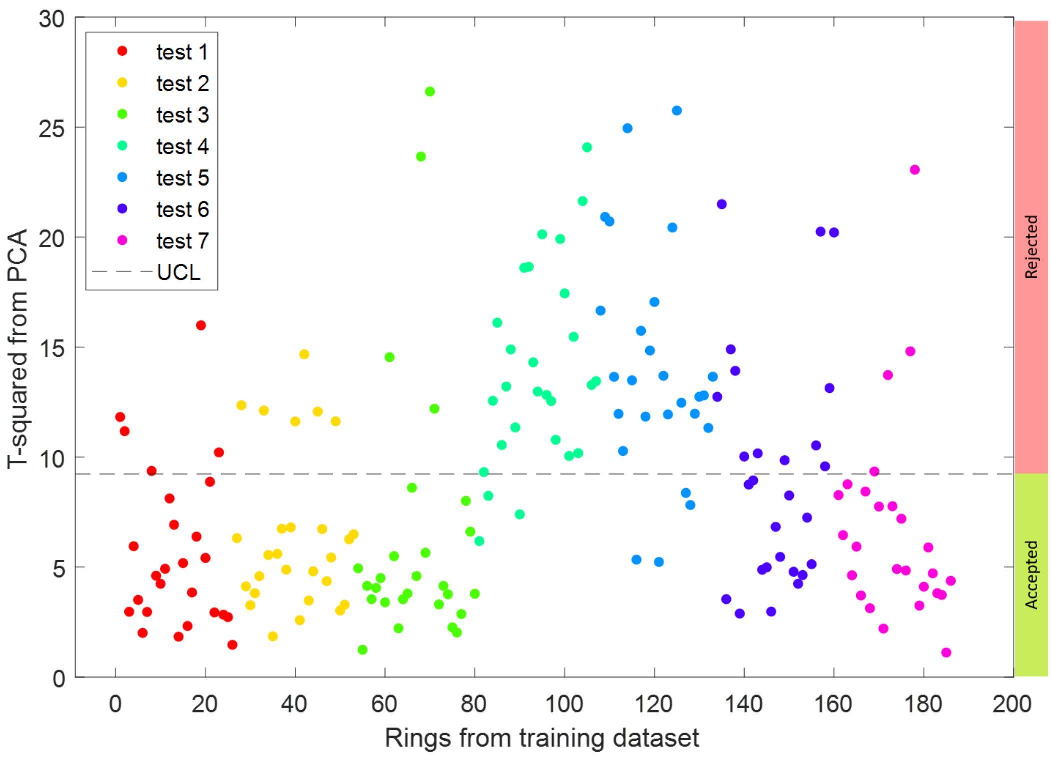

Figure 12. Calculated T^2 statistic from training quality dataset using MATLAB's Principal Component Analysis (PCA) algorithm. The dashed line represents Upper Control Limit (UCL).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 16 August 2022 doi:10.20944/preprints202208.0283.v1

the performance of the model on the test set with calculated accuracy of more than 93% on the corresponding rings measured for quality. In comparison, the FCM approach has less absolute differentiation between the quality data from different test classes which confirms the overlapping quality clusters in the hyper-space. However, the results from the T^2 statistic are repeatable which is desirable in any failure prediction model.

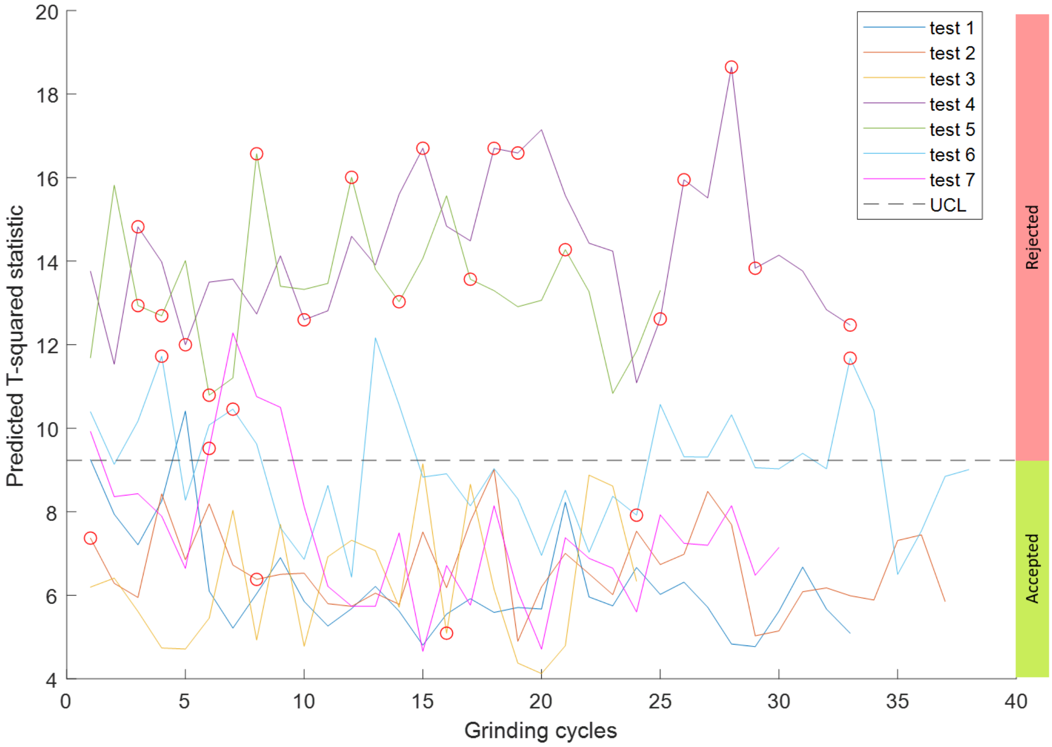

Figure 13. Predicting the overall produced quality in test data through estimating T^2 statistic. The dashed line represents Upper Control Limit threshold above which the quality is unacceptable. The red circle markers indicate the rejected rings according to quality criteria.

The results presented here demonstrate the potential of using data from the sensors, e.g. acoustic emission, vibration, force, power, and temperature, installed for the purpose of process control and condition monitoring to predict quality in a bearing ring grinder. In this work, the class balance has been ensured in setting up the experimental tests to avoid over-representation of any failure mode in the dataset [2]. Although different failure modes affect the measured quality parameters differently, the defined quality criteria account for the individual parameter in comparison to the reference quality test. The high accuracy prediction results achieved on the dataset give the confidence to use the presented approach in regular production to estimate overall quality for rings using the sensor data only.

4. Conclusion

This paper presents an approach to predicting the overall quality of ground bearing rings using a feature set from the sensor data. The grinder is equipped with additional acoustic emission, vibration, force (strain), and temperature sensors for machine health monitoring purposes. The feature set, resulting from the failure diagnostics using sensor data, is also used to benchmark the random forest as a top-performing regression model to estimate the quality of the produced rings. Using Hotelling's T-squared statistic to generate quality labels is presented as the preferred choice over fuzzy c-means clustering, for its repeatable results. The model prediction is compared against a threshold value for ring quality disposition to trigger a maintenance action. The quality criteria based on individual

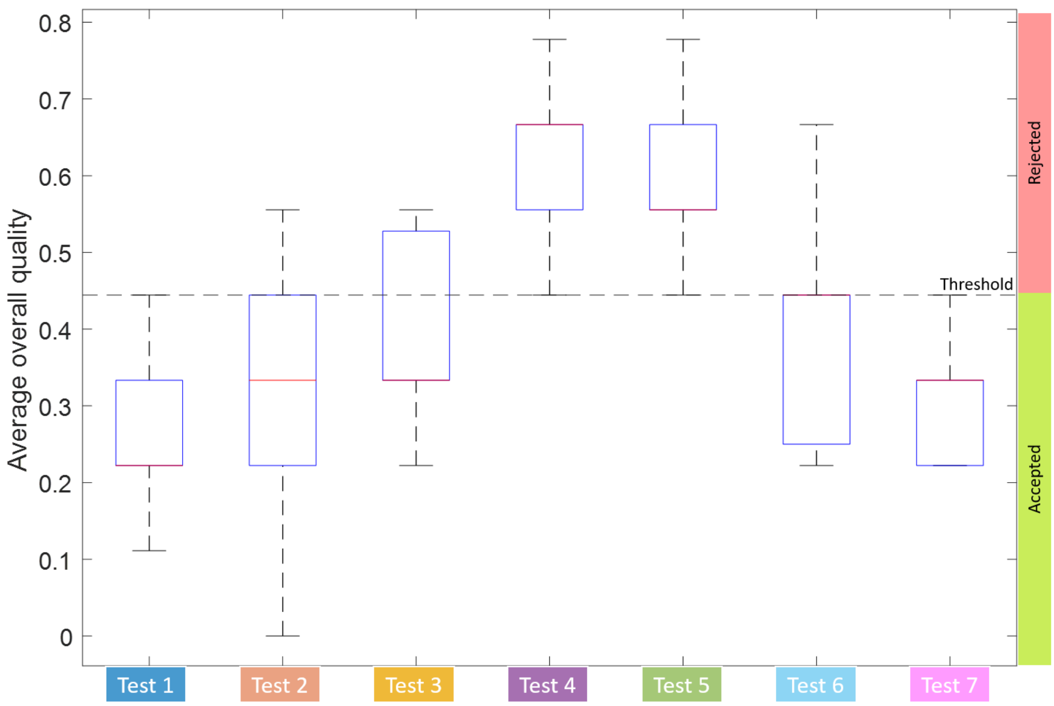

Figure 14. Box plot showing the quantification of the quality produced in each test after the acceptance criteria have been applied to the measured parts.

quality parameters validate the proposed model performance with 94.2% accuracy on the test set of measured rings. The use of multivariate quality statistic ensures the consideration of all quality parameter variations that are otherwise infeasible to measure in-line during the grinding operation. Thus, this work successfully demonstrates the possibility to use the data from the installed sensors to not only estimate the condition and performance but also to predict the produced quality variations of the production grinder. The high accuracy achieved in predicting the overall quality evidently shows the effectiveness of such decision support in triggering maintenance action. The potential to improve the performance through enhanced quality classification criteria needs to be verified through extended testing and measurement of the produced parts. Additionally, individual quality parameters can be predicted to take specific remedial actions. However, for the scope of this work, the herein presented approach demonstrates an efficient and effective implementation of a maintenance decision support system for a bearing ring grinder. With the availability of multiple sensor data from the entire grinding cycle, using more sophisticated data processing and model development can be considered to improve failure prognostics as part of predictive maintenance.

Author Contributions: All authors contributed to the study's conception and design. Instrumentation, software, testing, formal analysis, and data curation were performed by Muhammad Ahmer. Data collection and analysis was supervised and reviewed by Fredrik Sandin, Pär Marklund and Martin Gustafsson. The first draft of the manuscript was written by Muhammad Ahmer and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable

Informed Consent Statement: Not applicable

Data Availability Statement: The dataset will be made available along with the acceptance and publication of this manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Urgoiti, L.; Barrenetxea, D.; Sánchez, J.A.; Godino, L. Experimental study of thermal behaviour of face grinding with alumina angular wheels considering the effect of wheel wear. *CIRP Journal of Manufacturing Science and Technology* **2021**, *35*, 691–700. https://doi.org/10.1016/j.cirpj.2021.07.010.
2. Ahmer, M.; Marklund, P.; Gustafsson, M.; Berglund, K. An implementation framework for condition-based maintenance in a bearing ring grinder. *Procedia CIRP* **2022**, *107*, 746–751. Leading manufacturing systems transformation – Proceedings of the 55th CIRP Conference on Manufacturing Systems 2022, https://doi.org/10.1016/j.procir.2022.05.056.
3. Tönshoff, H.; Friemuth, T.; Becker, J. Process Monitoring in Grinding. *CIRP Annals* **2002**, *51*, 551–571. https://doi.org/10.1016/S0007-8506(07)61700-4.
4. Teti, R.; Jemielniak, K.; O'Donnell, G.; Dornfeld, D. Advanced monitoring of machining operations. *CIRP Annals* **2010**, *59*, 717–739. https://doi.org/10.1016/j.cirp.2010.05.010.
5. Wegener, K.; Bleicher, F.; Krajnik, P.; Hoffmeister, H.W.; Brecher, C. Recent developments in grinding machines. *CIRP Annals* **2017**, *66*, 779–802. https://doi.org/10.1016/j.cirp.2017.05.006.
6. Fedele, L. *Methodologies and Techniques for Advanced Maintenance*; Springer, 2011.
7. Vogl, G.W.; Weiss, B.A.; Helu, M. A review of diagnostic and prognostic capabilities and best practices for manufacturing. *Journal of Intelligent Manufacturing* **2016**, *30*, 79–95. https://doi.org/10.1007/s10845-016-1228-8.
8. Diez-Olivan, A.; Del Ser, J.; Galar, D.; Sierra, B. Data fusion and machine learning for industrial prognosis: Trends and perspectives towards Industry 4.0. *Information Fusion* **2019**, *50*, 92–111. https://doi.org/10.1016/j.inffus.2018.10.005.
9. Zhang, P.; Gao, Z.; Cao, L.; Dong, F.; Zou, Y.; Wang, K.; Zhang, Y.; Sun, P. Marine Systems and Equipment Prognostics and Health Management: A Systematic Review from Health Condition Monitoring to Maintenance Strategy. *Machines* **2022**, *10*. https://doi.org/10.3390/machines10020072.
10. Teixeira, H.N.; Lopes, I.; Braga, A.C. Condition-based maintenance implementation: a literature review. *Procedia Manufacturing* **2020**, *51*, 228–235. 30th International Conference on Flexible Automation and Intelligent Manufacturing (FAIM2021), https://doi.org/10.1016/j.promfg.2020.10.033.
11. Heng, A.; Zhang, S.; Tan, A.C.; Mathew, J. Rotating machinery prognostics: State of the art, challenges and opportunities. *Mechanical Systems and Signal Processing* **2009**, *23*, 724–739. https://doi.org/10.1016/j.ymssp.2008.06.009.
12. Goyal, D.; Pabla, B. Condition based maintenance of machine tools—A review. *CIRP Journal of Manufacturing Science and Technology* **2015**, *10*, 24–35. https://doi.org/10.1016/j.cirpj.2015.05.004.
13. Jardine, A.K.; Lin, D.; Banjevic, D. A review on machinery diagnostics and prognostics implementing condition-based maintenance. *Mechanical Systems and Signal Processing* **2006**, *20*, 1483–1510. https://doi.org/10.1016/j.ymssp.2005.09.012.
14. Camci, F. Autonomous Diagnostics and Prognostic Framework. PhD thesis, 2005.
15. Dragomir, O.E.; Gouriveau, R.; Dragomir, F.; Minca, E.; Zerhouni, N. Review of prognostic problem in condition-based maintenance. In Proceedings of the 2009 European Control Conference (ECC), 2009, pp. 1587–1592. https://doi.org/10.23919/ECC.2009.7074633.
16. Leukel, J.; González, J.; Riekert, M. Adoption of machine learning technology for failure prediction in industrial maintenance: A systematic review. *Journal of Manufacturing Systems* **2021**, *61*, 87–96. https://doi.org/10.1016/j.jmsy.2021.08.012.
17. Lei, Y.; Yang, B.; Jiang, X.; Jia, F.; Li, N.; Nandi, A.K. Applications of machine learning to machine fault diagnosis: A review and roadmap. *Mechanical Systems and Signal Processing* **2020**, *138*, 106587. https://doi.org/10.1016/j.ymssp.2019.106587.
18. Orth, P.; Yacout, S.; Adjengue, L. Accuracy and robustness of decision making techniques in condition based maintenance. *Journal of Intelligent Manufacturing* **2009**, *23*, 255–264. https://doi.org/10.1007/s10845-009-0347-x.
19. Vollert, S.; Atzmueller, M.; Theissler, A. Interpretable Machine Learning: A brief survey from the predictive maintenance perspective. In Proceedings of the 2021 26th IEEE International Conference on Emerging Technologies and Factory Automation (ETFA). IEEE, 2021. https://doi.org/10.1109/etfa45728.2021.9613467.
20. Rosati, R.; Romeo, L.; Cecchini, G.; Tonetto, F.; Viti, P.; Mancini, A.; Frontoni, E. From knowledge-based to big data analytic model: a novel IoT and machine learning based decision support system for predictive maintenance in Industry 4.0. *Journal of Intelligent Manufacturing* **2022**. https://doi.org/10.1007/s10845-022-01960-x.
21. Ahmer, M.; Sandin, F.; Marklund, P.; Gustafsson, M.; Berglund, K. Failure mode classification for condition-based maintenance in a bearing ring grinding machine **2022**. https://doi.org/10.21203/rs.3.rs-1521317/v1.
22. Shore, P.; Billing, O.; Puhasmagi, V. A Standard Grinding Wheel Assessment Method to Support a Sophisticated Grinding Knowledge Based System. *Key Engineering Materials - KEY ENG MAT* **2004**, *257-258*, 285–290. https://doi.org/10.4028/www.scientific.net/KEM.257-258.285.
23. Ahmer, M.; Marklund, P.; Gustafsson, M.; Berglund, K. Integration of process monitoring and machine condition diagnostics to improve quality prediction in grinding. *Procedia CIRP* **2021**, *101*, 170–173. 9th CIRP Conference on High Performance Cutting, https://doi.org/10.1016/j.procir.2021.02.019.