

## Full length article

# One-fault-shot learning for fault severity estimation of gears that addresses differences between simulation and experimental signals and transfer function effects



Omri Matania <sup>a</sup>, Lior Bachar <sup>a</sup>, Varun Khemani <sup>b</sup>, Diganta Das <sup>b</sup>, Michael H. Azarian <sup>b</sup>, Jacob Bortman <sup>a,\*</sup>

<sup>a</sup> PHM Laboratory, Department of Mechanical Engineering, Ben-Gurion University of the Negev, P.O. Box 653, Beer Sheva 8410501, Israel

<sup>b</sup> Center for Advanced Life Cycle Engineering (CALCE), University of Maryland, College Park, MD 20742, United States

## ARTICLE INFO

**Keywords:**

Simulated signals  
Physical domain adaptation  
Fault size estimation  
Gear vibration signals  
Transfer function  
Transfer across different machines  
Simulation to reality

## ABSTRACT

Gearboxes are integral elements in rotating machines and have a high tendency to fail due to their operation in harsh conditions. A robust method to estimate the fault size of gears is desirable for a successful prognostic process, which is, to date, still unavailable in the literature. The fault size can be estimated by vibration analysis, using signal processing and machine-learning tools. However, the availability of labeled or unlabeled vibration signals from faulty rotating machinery components is rare, making it challenging to apply machine-learning algorithms. Therefore, some physical pre-knowledge should be incorporated in the model for a successful learning process. This can be done by exposing the learning model to simulated data, and by a physical pre-processing procedure. This paper suggests a novel algorithm to overcome the lack of faulty data (labeled and unlabeled), and it is trained on a combination of simulated data and some real data. The algorithm tunes the differences between simulation and experiment using one faulty experimental example, and transfers knowledge from simulation to reality by addressing the transfer function effects. It addresses the transfer function by spectrum background estimation and minimum phase estimation while also selecting features that are invariant to the unmitigated effects of the transfer function. The new algorithm is demonstrated on simulated signals and measured transfer function, and on experimental signals with known fault sizes.

The codes and the data of the study are available via the link: [https://github.com/omriMatania/one\\_fault\\_shot\\_learning\\_for\\_gears\\_fault\\_severity\\_estimation](https://github.com/omriMatania/one_fault_shot_learning_for_gears_fault_severity_estimation).

## 1. Introduction

Prognostics and health management by vibration analysis is a very common approach for predictive maintenance, aspiring to extend the useful lifespan of rotating machinery [1,2]. The vibration signals are measured via vibration sensors and are analyzed by unique signal processing tools and machine-learning algorithms [3–5]. Condition monitoring of rotating components (such as gears [6–9], bearings [10–12], and shafts [13,14]) enables early identification of developing faults or wear that requires maintenance actions [15–17], thus, preventing catastrophic failures [1,18].

Health monitoring of gears by vibration analysis has resulted in several achievements, including realistic physical models [19–23],

unique signal processing procedures such as synchronous averaging (SA) and difference signal analysis [24–26], and machine-learning algorithms [18,27]. Machine-learning algorithms for prognostics and health management, also known as intelligent fault diagnosis (IFD) algorithms [28,29], seem like a promising improvement for gear diagnostics [30,31].

IFD of gears by machine-learning algorithms has been widely investigated. Many of the published papers apply traditional machine-learning methods for IFD, e.g., support vector machine [32–34], nearest neighbor [35–38], artificial shallow neural network [39–41], and decision trees [42]. However, since the revival of neural networks in recent years via the incarnation of deep-learning models, many studies have focused on applying deep-learning algorithms for IFD, e.g., deep belief networks [43], convolutional neural networks [44–48], and

\* Corresponding author.

E-mail addresses: [omrimatania@gmail.com](mailto:omrimatania@gmail.com) (O. Matania), [liorbac@post.bgu.ac.il](mailto:liorbac@post.bgu.ac.il) (L. Bachar), [vkheman@terpmail.umd.edu](mailto:vkheman@terpmail.umd.edu) (V. Khemani), [diganta@umd.edu](mailto:diganta@umd.edu) (D. Das), [mazarian@umd.edu](mailto:mazarian@umd.edu) (M.H. Azarian), [jacobt@bgu.ac.il](mailto:jacobt@bgu.ac.il) (J. Bortman).

Nomenclature	
Acronyms	
ACS	Adaptive clutter separation
FTFF	Full tooth face fault
IFD	Intelligent fault diagnosis
KNN	K nearest neighbor
RMS	Root mean square
RPS	Revolutions per second
SA	Synchronous average
TIM	Transfer in the identical machine
TDM	Transfer across different machines

learning [62], and pseudo labeling [59], the TDM algorithms transfer the knowledge from the source machine to the tested machine. However, TIM and TDM algorithms cannot be implemented in critical rotating machinery such as airplanes and helicopters, because the assumption that enough representative **faulty data (labeled or unlabeled)** can be collected from the tested machinery is impractical, as only a very small amount (e.g., less than ten examples) of faulty data is available due to safety considerations.

Except for the above-mentioned problem, many of the studies in the field of gear diagnosis by machine-learning suffer from a lack of pre-processing stages, as they process the signal directly in the time domain and do not include common signal processing techniques, e.g., angular resampling and SA calculation. Signal processing techniques make the learning process robust and are therefore essential for diagnosing a real gearbox, where the speed fluctuations and the interference

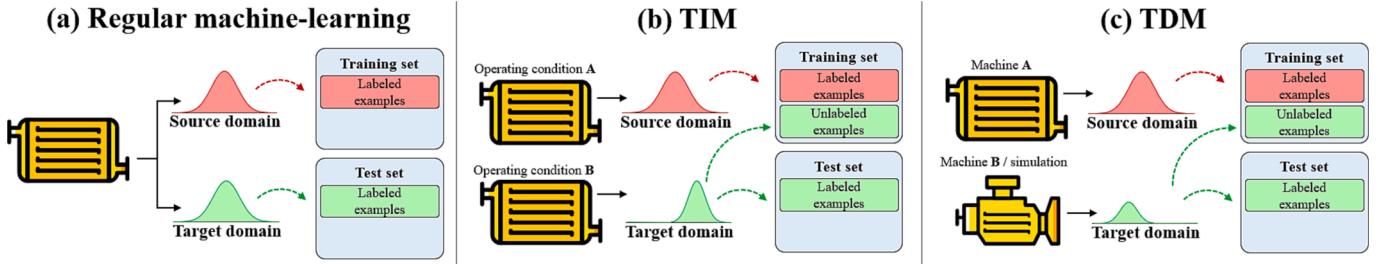


Fig. 1. Training and test sets of: (a) regular machine-learning, (b) TIM, and (c) TDM. Machine B in (c) can be a simulation.

recurrent-based neural networks [4,49].

Techniques for fault severity estimation of gears have been suggested over the years [1,50–52], where features that are correlated with the fault's severity level are extracted, while a model predicts the severity based on the learned correlation from former cases of the same machine (e.g., a model that predicts the severity based on the root mean square (RMS) of the vibration signal [1]). However, these learning techniques are relevant only if the source and target domains share the same correlation between the features and the severity, as presented in Fig. 1 (a), where a “regular” machine-learning is depicted. However, as explained by Lei et al. [53], the assumption that the source and target domains are identical is problematic, as it is difficult to collect enough **labeled faulty data**, for two main reasons: (1) machines work most of the time under a healthy state, in which faults are rare; hence, healthy data are much more common than faulty data; and (2) it is hard to label data, because it requires time and cost to stop a machine, inspect its rotating component status, and label the data. The process of disassembly entails expensive manpower and long shutdowns, while inevitable re-assembly errors might generate new faults.

To mitigate this shortcoming, several transfer learning algorithms have been developed [25]. Transfer-learning algorithms transfer accumulated “knowledge” from one task to another related task [54,55]. As explained in [53], transfer learning techniques for health monitoring in general can be divided into two groups: transfer in the identical machine (TIM) and transfer across different machines (TDM). For TIM algorithms, domain adaptation methods [56] are used to transfer knowledge from one condition to a new condition in which it is assumed that there are labeled data from the first condition and unlabeled data from the new condition in the training set as depicted in Fig. 1 (b) [57,58]. TDM algorithms also use domain adaptation methods to transfer knowledge from the source machine to a new tested machine [59,60]. These algorithms assume that the training set contains unlabeled examples of faulty signals from the target domain, as depicted in Fig. 1 (c). Hence, by solving joint discrepancy distribution problem using several techniques such as minimization of the discrepancy measurement [61], adversarial

## One-fault-shot learning

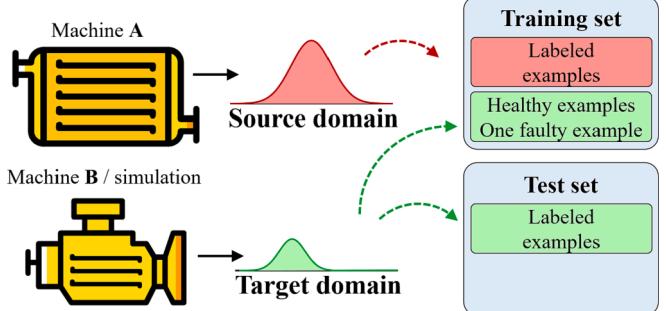
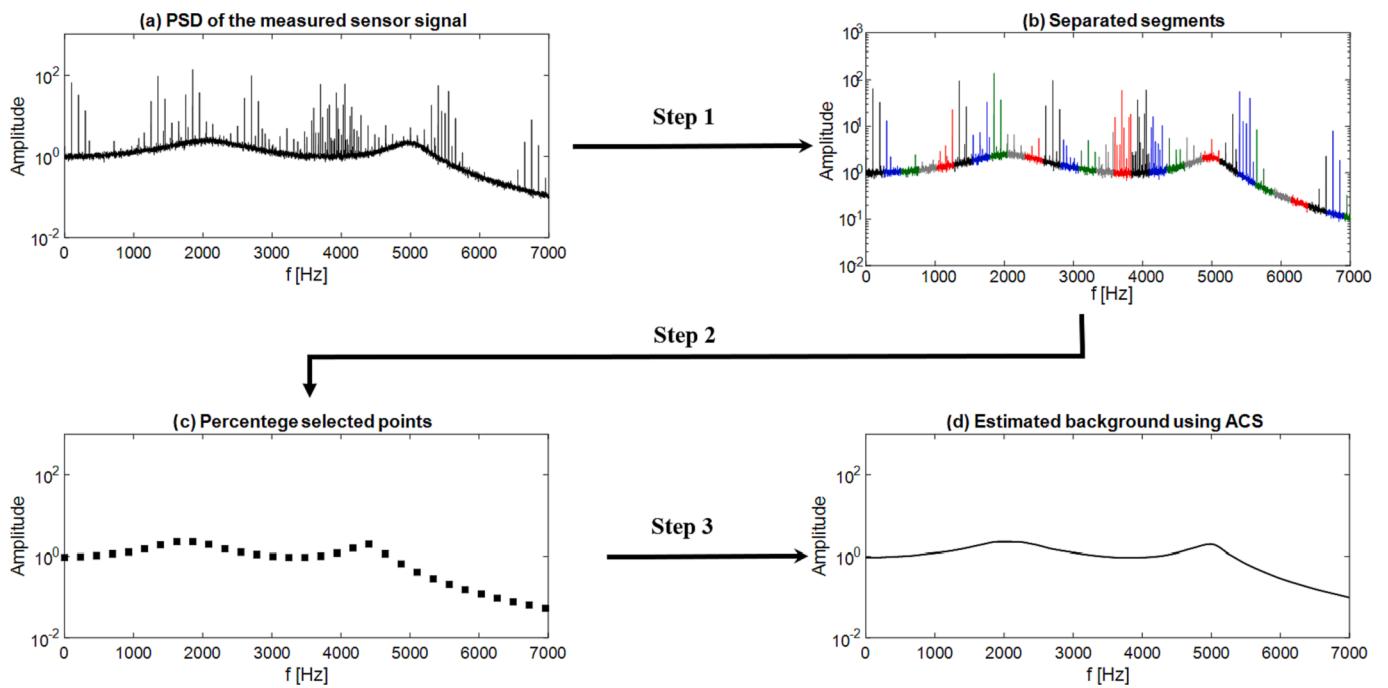


Fig. 2. Training and test sets of one-fault-shot learning. Machine B in can be a simulation.

of other rotating components are much more significant in comparison to the laboratory test rigs.

In this paper, a new methodology of simulation to reality for gear diagnosis is suggested. The learning model is trained on realistic simulations and then generalizes to experimental data using healthy experimental signals and one faulty experimental signal, as depicted in Fig. 2. (“one-fault-shot learning”, other examples from other fields of few-shot learning can be found in ref. [54]). The new algorithm has two main stages: (1) tuning the simulation to mitigate its differences from the reality using one faulty experimental signal, and (2) suppressing the transfer function effects by estimating them while also selecting features which are invariant (i.e., not affected by) to the unmitigated effects of the transfer function.

The gained pre-knowledge is incorporated into the learning algorithm by three means, enabling generalization to a new system with only



**Fig. 3.** Estimation of the spectrum background using ACS. (a) The spectrum, (b) separation of the spectrum into different segments, (c) the p percentage value for every section, and (d) the estimated spectrum background using ACS. Reproduced from ref. [66].

one faulty signal: (1) The simulated data reflect the physical phenomena and, thus, include pre-knowledge. For example, the simulated data can incorporate the knowledge that, in most cases, the energy of the signal is expected to increase when a fault occurs. (2) Common signal processing techniques for gear diagnosis are applied to the data, including angular resampling, synchronous average, and calculation of the difference signal. (3) The pre-processing of the data includes suppression of changes generated by the transfer function. The transfer function is a unique characteristic of every dynamic system. Therefore, suppressing its effects will help eliminate the influence of irrelevant components and generalize insights across different machines. This study makes three important contributions to the current knowledge of gear diagnosis:

1. A new algorithm for fault severity estimation of gears: The new hybrid algorithm combines machine-learning with signal processing methods and enables fault severity estimation with only one faulty example, as presented in Fig. 2. The current paper demonstrates the advantages of this hybrid approach which makes use of the two methods.
2. Better generalization from simulation to reality by addressing the inevitable differences from the reality (resulted from many assumptions such as point contact, rigid bodies, fully elastic/plastic collision, etc.) using one-fault-shot learning.
3. A novel process of feature selection for addressing the unmitigated effects of the transfer function.

The rest of this paper is organized as follows: The theoretical background is described in Section 2, including description of techniques for transfer function estimation, extraction of physical features and nearest neighbor algorithm. The simulated and the experimental databases are described in Section 3. The paper includes two types of datasets described in Section 4: (1) simulated signals combined with measured transfer functions, and (2) simulated signals combined with experimental signals. The new algorithm is described in Section 5, including the description of the feature selection process and addressing of differences between simulation to reality. The performance of the new

algorithm is demonstrated in Section 6.

## 2. Theoretical background

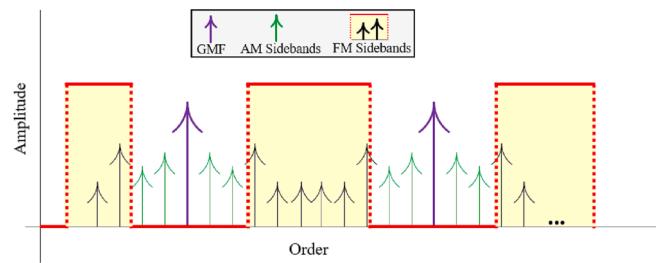
This section describes the required background for the study, including a description of the algorithm for the transfer function estimation, analysis of the synchronous average signal and feature extraction methodology, and introduction to the nearest neighbor algorithm.

### 2.1. Transfer function estimation

The transfer function of a dynamic system is a complex function that describes the relationship between the output response and the input signal. There are several methods to estimate the transfer function [1,63]. For example, during impact hammer modal testing, both the input and output signals are measured and, thus, allow an approximated calculation of the transfer function. However, in most real-world applications, the input signal is not measured, and the transfer function ought to be estimated from the measured output signal only. Here, we use the adaptive clutter separation (ACS) for spectrum background estimation; i.e., a close approximation of the magnitude of the transfer function [64], and the minimum phase assumption for phase estimation [1].

#### 2.1.1. Adaptive clutter separation (ACS)

ACS is described in [65] and was broadly investigated in [66,67]. The algorithm estimates the background of the vibration spectrum in three steps, as presented in Fig. 3: First, the spectrum is separated into overlapping segments. Then the value of the pre-defined percentile is calculated for each segment (usually percentile is between 15% and 50%), and this value is assigned to the middle point of the segment as depicted in Fig. 3 (c). Finally, the spectrum background is estimated by interpolating the points from the former step to match the original dimension of the spectrum. ACS suppresses peaks in the original spectrum as their values are in the 80%–90% percentiles or higher, thus, preserving the shape of the spectrum background which is attributed to



**Fig. 4.** Illustration of the SA spectrum in the order domain. Reproduced from ref. [69]. GMF. garmesh frequencies.

the structure.

### 2.1.2. Minimum phase estimation

The idea of minimum phase assumption using cepstral analysis was introduced by [68]. A minimum phase system is characterized by having no zeros or poles outside of the unit circle. It is shown in [1] that by zeroing the negative quefrequencies and doubling positive quefrequencies in the real cepstrum, a minimum phase is obtained. The estimated transfer function is obtained by applying the minimum phase assumption to the estimated spectrum background (see Equations (1)–(4)).

$$|tf| = \text{ACS}(\text{PSD}(x)) \quad (1)$$

$$s^{\text{cepst}} = \text{IDFT}(\ln(|tf|)) \quad (2)$$

$$s_{\text{mp}}^{\text{cepst}}[q] = \begin{cases} s^{\text{cepst}}[q], & q = 0 \\ 2 \bullet s^{\text{cepst}}[q], & q > 0 \\ 0, & q < 0 \end{cases} \quad (3)$$

$$tf_{\text{mp}} = e^{\text{DFT}(s_{\text{mp}}^{\text{cepst}})} \quad (4)$$

where  $x$  is the measured signal,  $|tf|$  is the estimated magnitude of the transfer function, ACS is the ACS algorithm described in Section 2.1.1, PSD is the power spectral density,  $\ln()$  is the logarithmic operation,  $q$  is the quefrency coordinate in the cepstrum domain, IDFT is the inverse discrete Fourier transform, DFT is the discrete Fourier transform and  $tf_{\text{mp}}$  is the estimated minimum phase transfer function.

### 2.2. Analysis of the synchronous average signal in the cycle domain

Synchronous average (SA) is a signal processing algorithm, very commonly used for gear diagnosis [24–26] because gears are synchronous elements, meaning that there is no sliding between them. According to the SA algorithm, the vibration signal after angular resampling (described in Appendix A.1) is separated into segments with a length of one cycle, and then all the segments are averaged together. The periodic pattern of the elements synchronous to the corresponding shaft's speed is preserved, while reducing noise and eliminating other components (such as other gearboxes in the system). The SA spectrum includes high peaks at harmonics of the garmesh frequency, surrounded by sidebands representing modulation of the rotational shaft speed (see Fig. 4) [26]. Previous study shows that local tooth faults are emphasized in the modulation sidebands that are far from the garmesh harmonics, while the garmesh harmonics and their associated close sidebands are more affected by the operational conditions, shaft phenomena, and distributed faults [26]. The harmonic signal is the SA signal after spectral band filtering around the garmesh harmonics and their associated close pair of modulation sidebands. The difference signal is extracted from the SA signal by suppressing the garmesh harmonics and their associated close pair of sidebands, i.e., subtracting the harmonic signal from the SA signal.

### 2.3. Feature extraction

Several features can be extracted from the difference signal [70]. Four of these features are tested in this paper: RMS and kurtosis of the difference signal, and RMS and skewness of the envelope of the difference signal [69] (see Equations 5–7). These features are expected to indicate energy increase and/or appearance of random sharp peaks in the difference signal, which are usually associated with local tooth faults.

$$\text{rms} = \sqrt{\frac{1}{N} \sum_{n=1}^N x[n]^2} \quad (5)$$

$$\text{skewness} = \frac{\frac{1}{N} \sum_{n=1}^N (x[n] - \hat{\mu})^3}{\sigma^3} \quad (6)$$

$$\text{kurtosis} = \frac{\frac{1}{N} \sum_{n=1}^N (x[n] - \hat{\mu})^4}{\hat{\sigma}^4} \quad (7)$$

where  $\hat{\mu}, \hat{\sigma}$  are the estimated mean and variance of the signal  $x = \{x[n]\}_{n=1}^N$ .

### 2.4. Nearest neighbor algorithm

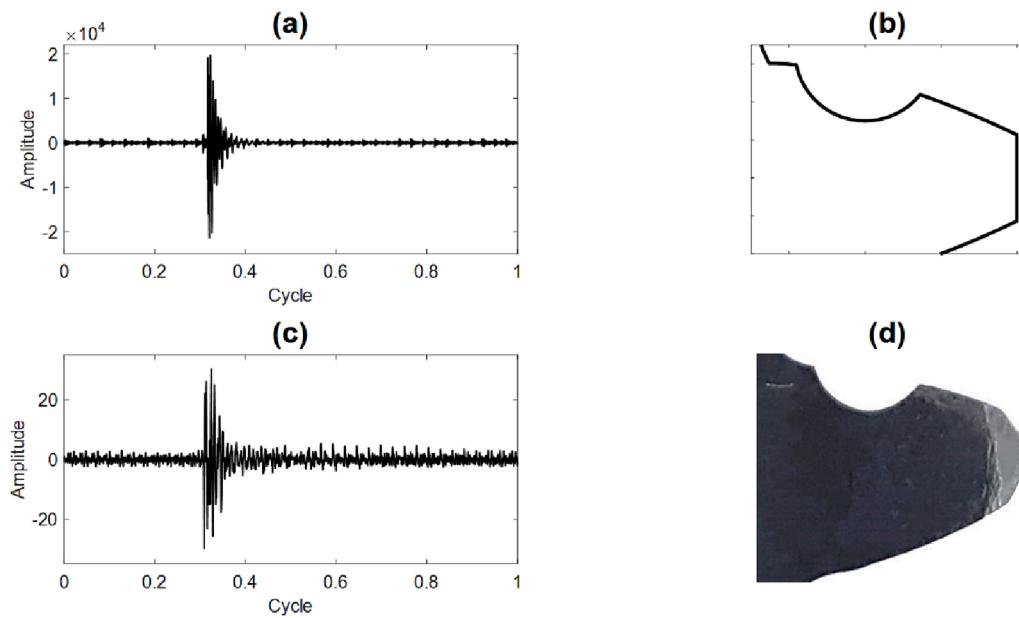
The nearest neighbor is a classic machine-learning algorithm for regression and classification problems, described widely in [71,72]. The algorithm predicts the label of a new example by the label of its the nearest training example under metric  $\rho$ . A common augmentation of the nearest neighbor algorithm is called the K nearest neighbors (KNN), in which the prediction label is determined by averaging/majority vote of K neighbors instead of one, for reducing the generalization error [71]. The hyper parameter K is selected from different values as the K corresponding to the minimal error of the validation set after applying KNN to the training and validation sets. In terms of fault diagnosis, the KNN algorithm could be used to predict the fault size of a new example in the test set, based on the nearest samples in the training set.

## 3. Databases

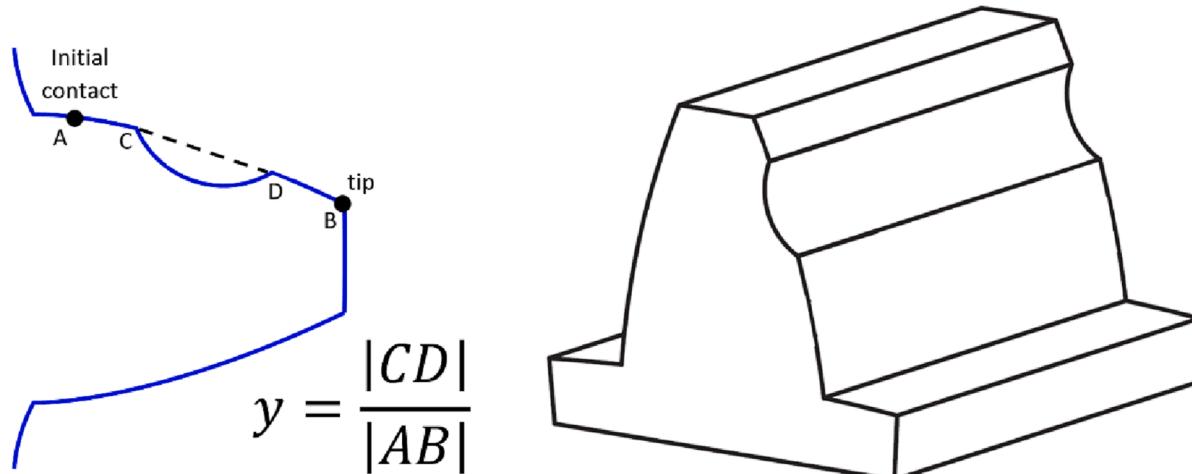
This section describes the three databases used in this study: an experimental database of vibration signals measured from a designated test rig, a simulated database of vibration signals, and a database of measured transfer functions obtained from impact hammer testing. As described in Section 4, these databases were used to create two different datasets to demonstrate the suggested new algorithm described in Section 5. In this paper, each dataset is composed from a collection of data from different databases.

### 3.1. The simulated & experimental databases

The inspected gearbox is a single-stage spur gear transmission. The driving pinion wheel contains 17 teeth, while the driven gear wheel



**Fig. 5.** Examples of simulated and measured signals. (a) SA of a simulated signal with its FTFF tooth depicted in (b). (c) SA of a measured signal with its FTFF tooth depicted in (d).



**Fig. 6.** An illustration of a full tooth face fault (FTFF). Left. qualitative definition of the fault size and prediction label  $y$ . Right. an isometric view of the pitted tooth.

**Table 1**  
Database Properties.

Parameter	Simulated database	Experimental database
Signal duration	38 cycles of the motor	60 sec
Number of healthy signals	100	6
Number of faulty signals	700 faulty signals with FTFF in range of $0 < y$ less than 1 (distributed uniformly)	3 cases of FTFF $y=\{0.32, 0.56, 0.66\}$ six repetitions for each case
Rotational speed [rps]	{10,11, ..., 35}	{15,30}
Load [Nm]	{5,10,15,20}	{5,10}

contains 38 teeth (Module 3). Both wheels are manufactured with a high precision grade of DIN7, according to a DIN-3692 standard [73]. Two different databases of vibration signals of the inspected gearbox were used in this study: an experimental database [69] and a simulated database. The simulated database was generated by the realistic dynamic model described in Appendix A.2. The database included both healthy and faulty signals. The inspected fault in this study was a form of pitted gear, designated in the paper as a “full tooth face fault” (FTFF), expressed by removal of material from the entire tooth width (see Fig. 6) [26]. All faults were seeded on a single tooth of the driven gear wheel. Fig. 5 compares the vibration signal between the simulation and the experiment, for the same case of FTFF. It can be noticed that the interaction with the defected tooth generated a sharp impulse in the vibration signal. However, the amplitude of the simulated signal was significantly higher than the experiment, as will be discussed later.

Both the experimental and simulation setups are detailed in Table 1. The prediction label  $y$  is a geometric ratio referred as the fault size (see

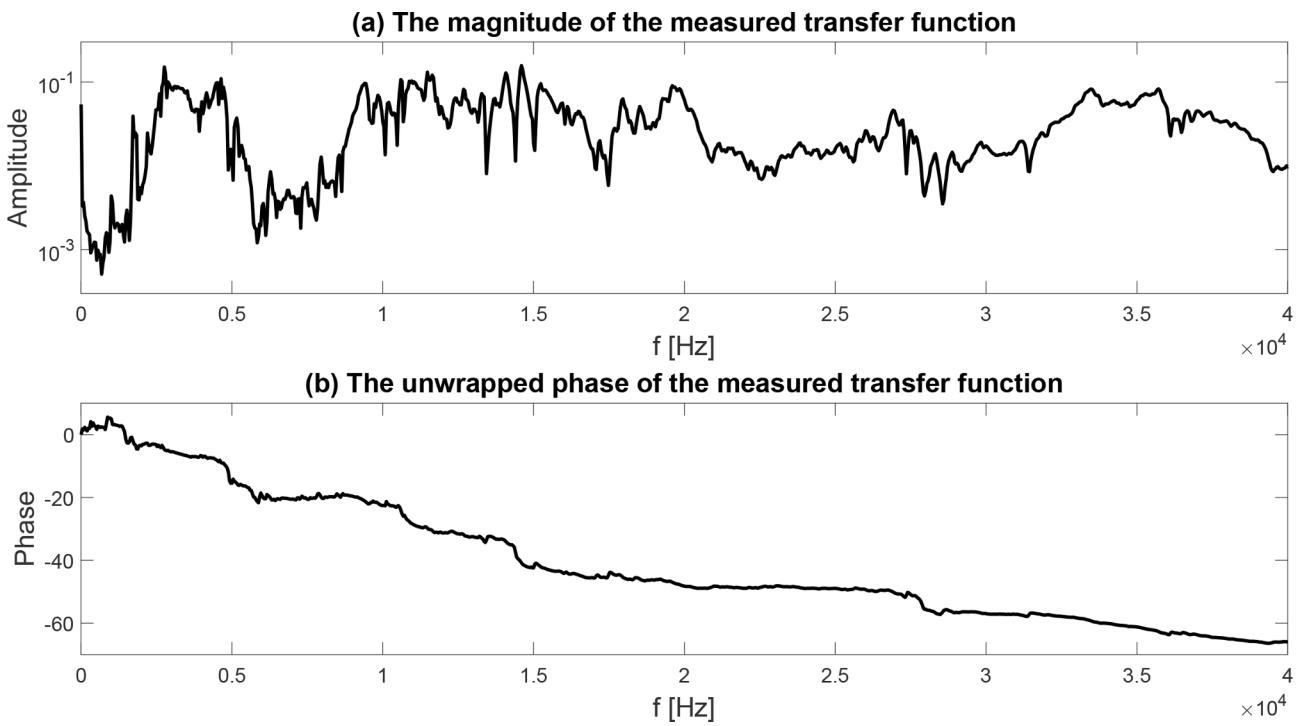


Fig. 7. A representative measured transfer function: (a) gain; (b) phase.

**Fig. 6.** A value of  $y = 0$  indicates a healthy gear, whereas  $y = 1$  is the worst case, where the entire involute profile is removed. In this study, the error was defined by Eq. (8) as the absolute difference between the real ( $y$ ) and the estimated ( $\hat{y}$ ) size of the fault.

$$\text{error} = |y - \hat{y}| \times 100\% \quad (8)$$

A degenerated algorithm that constantly predicts  $y = 0.5$  for every input of vibration signal has an expected error of 25% (assuming uniform distribution fault sizes in the range of  $[0,1]$ ). Thus, an effective algorithm should have an averaged error below 25%.

### 3.2. Measured transfer function database

Impact hammer testing is a common technique that was used to

estimate the transfer function in the current study. According to technique, a hammer gives short strikes to the desired system, creating impulse responses. The transfer function is estimated using the measured response in the vibration sensor mounted on the system, and the input signal measured in a load cell attached to the hammer. A set of 24 transfer functions are evaluated by impact hammer testing. The transfer functions are measured on different test rigs that include rotating components such as bearings, shafts, joints, etc. An example of a gain and phase representation of a measured transfer function in the frequency domain is depicted in Fig. 7. The attenuation of the transfer function is at least 10 dB for all frequencies, and there is a variation of the gain and phase as function of the frequency. The measurement of the transfer functions is discussed with more details in Appendix A.3. The attenuation and variation have significant effects on transferring

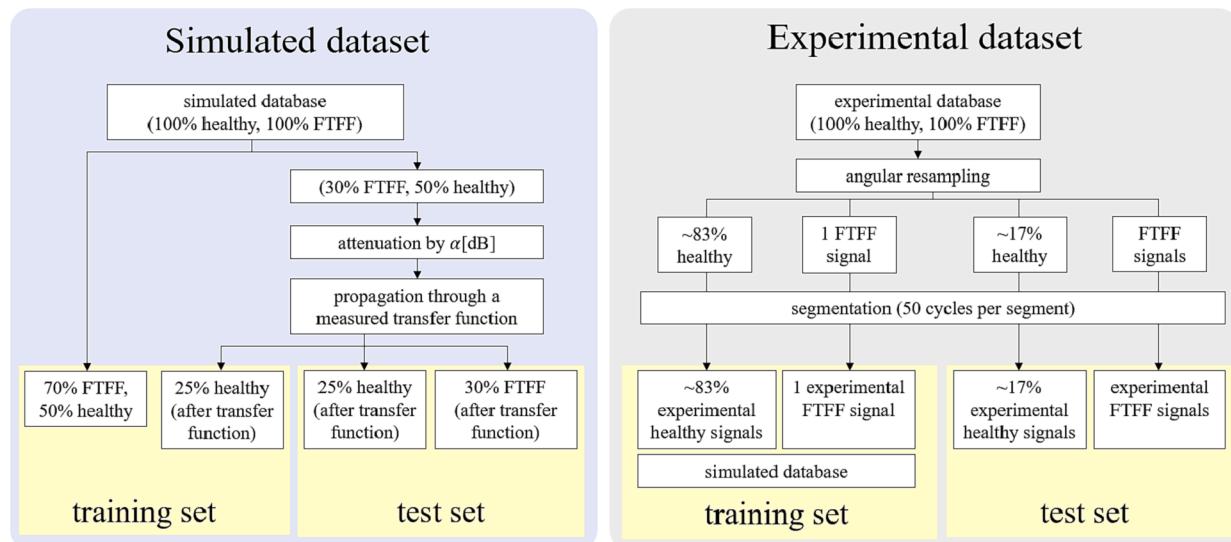
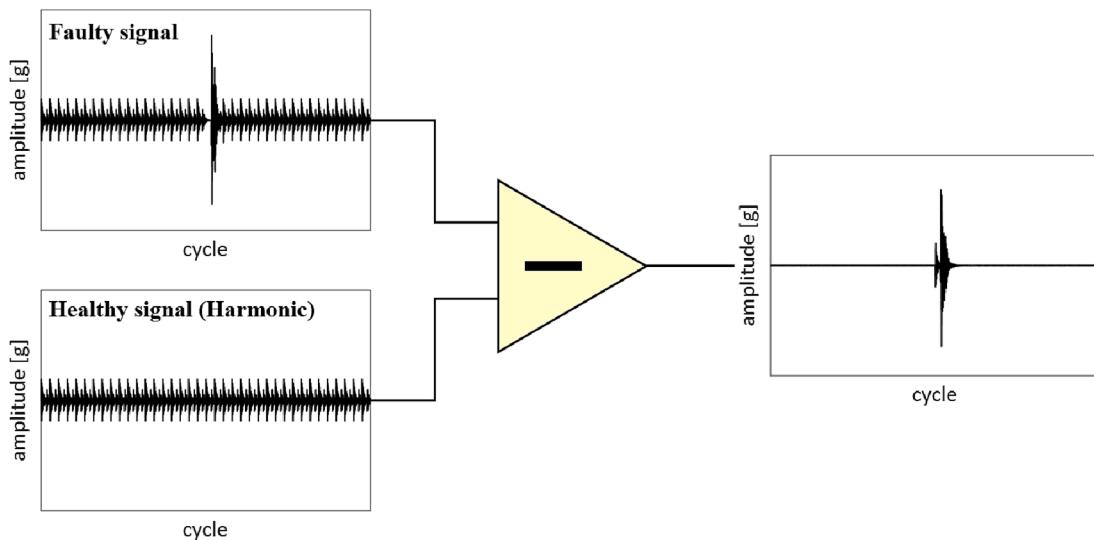


Fig. 8. The datasets used for the new algorithm: Left. simulated dataset. Right. experimental dataset. Angular resampling process is described in Appendix A.1.



**Fig. 9.** Subtraction of a healthy signal from a faulty signal.

knowledge from the simulation to reality as explained and demonstrated in the following sections.

It is worth noting that there are differences between the estimated transfer functions obtained through impact hammer testing and the real transfer functions. However, these differences do not affect the comparison process in [Section 6.1](#), as the performance of the algorithm is evaluated using the estimated transfer function as if it were the real transfer function. What is important is that these transfer functions represent real cases.

#### 4. The simulated & experimental datasets

The new algorithm for estimating the FTFF size is demonstrated using two types of datasets collected from the databases described in [Section 3](#): (1) A simulated dataset composed of simulated signals and measured transfer functions (from [Section 3.2](#)), and (2) an experimental dataset composed of simulated and measured experimental signals. In the current section, these two types of datasets are described.

##### 4.1. The simulated dataset

A block diagram of the simulated dataset is presented in [Fig. 8](#) and described as following: The healthy signals are divided equally between the training and test sets, while the faulty signals are split 70%–30% randomly between the training and test sets, respectively. The simulated test signals are attenuated by a factor of  $\alpha$  and then are propagated through a transfer function from the measured transfer functions database described in [Section 3.2](#), which is a typical transfer function of a laboratory test-rig system, as presented in Appendix A.3. Instead of limiting the simulated dataset only for the measured transfer functions, we added the parameter  $\alpha$  to control the attenuation, while the measured transfer function controls the varying gain and phase. In the next step, 25% of the healthy signals in the test set are moved to the training set (see [Fig. 8](#)), serving as the healthy signals of the target domain as part of the training set, as depicted in [Fig. 2](#). To generate the simulated datasets, 23 out of the 24 transfer functions in the database of [Section 3.2](#) are used (the remaining transfer function will be used for the feature selection process, described later). The generation of the dataset is repeated for each of the 23 measured transfer functions and for each combination of speed and load, resulting in 2,392 simulated datasets (see [Table 1](#)).

#### 4.2. The experimental dataset

A block diagram of the experimental dataset is presented in [Fig. 8](#) and described as following: Each measured vibration signal is divided into consecutive segments that correspond to 50 cycles. Five out of six healthy signals are added to the training set, as well as one segment of the FTFF signal, selected randomly. The rest of the signals are added to the test set. In addition, all the simulated faulty signals and 50% of the simulated healthy signals are added to the training set (see [Table 1](#)). The generation of the dataset is repeated for each combination of speed and load, resulting in four experimental datasets (see [Table 1](#)).

### 5. Algorithm description

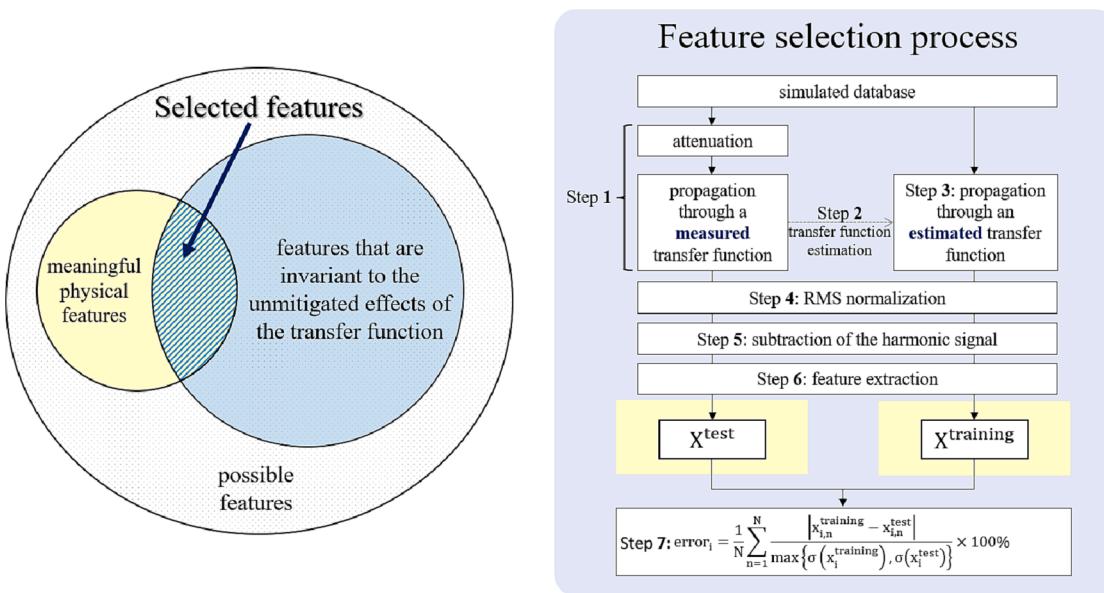
In this section the new algorithm is described. The new algorithm overcomes three challenges: (1) differences between the physical model and the real world resulting in differences between the simulated signals and the measured signals, (2) suppression of the transfer function effects, and (3) selection of features that are invariant to the unmitigated effects of the transfer function. The physical pre-processing methods are described in [Section 5.1](#), and are incorporated in the new algorithm, described in [Section 5.2](#).

#### 5.1. Physical pre-processing

This Section describes physical pre-processing methods that are applied to the dataset.

##### 5.1.1. Physical domain adaptation

The effect of the transfer function can be separated into attenuation effects and dispersion/distortion effects. The attenuation of the transfer function directly affects the RMS of the signal, while the gear signal's frequencies have serious distortion as a result of phase shifting and gain. The term “domain adaptation” in this paper refers to the following two physical procedures applied to the data in order to suppress the effects of the transfer function. The first procedure is called “RMS normalization”, according to which the vibration signal is divided by the average RMS of the healthy signals, thus, mitigating the attenuation effects. The second procedure is propagation through an estimated normalized transfer function that is estimated based on the methodology described in [Section 2.1](#), namely ACS and minimum phase estimation. The approximation of the gain of the transfer function using the estimated spectrum background of the vibrations signal is relevant for gear signals, as the



**Fig. 10.** Description of the feature selection algorithm: Left. qualitative Venn diagram to explain the concept of the suggested algorithm. Right. block diagram.

surface shape [9] and the transmission errors of the gears [1,63] generate wideband noise that propagates through the same transfer function as the gear signal. The estimated spectrum background follows the slow variations of the spectrum signal as illustrated in Fig. 3, which represents the spectrum signal of the wideband noise multiplied by the transfer function magnitude.

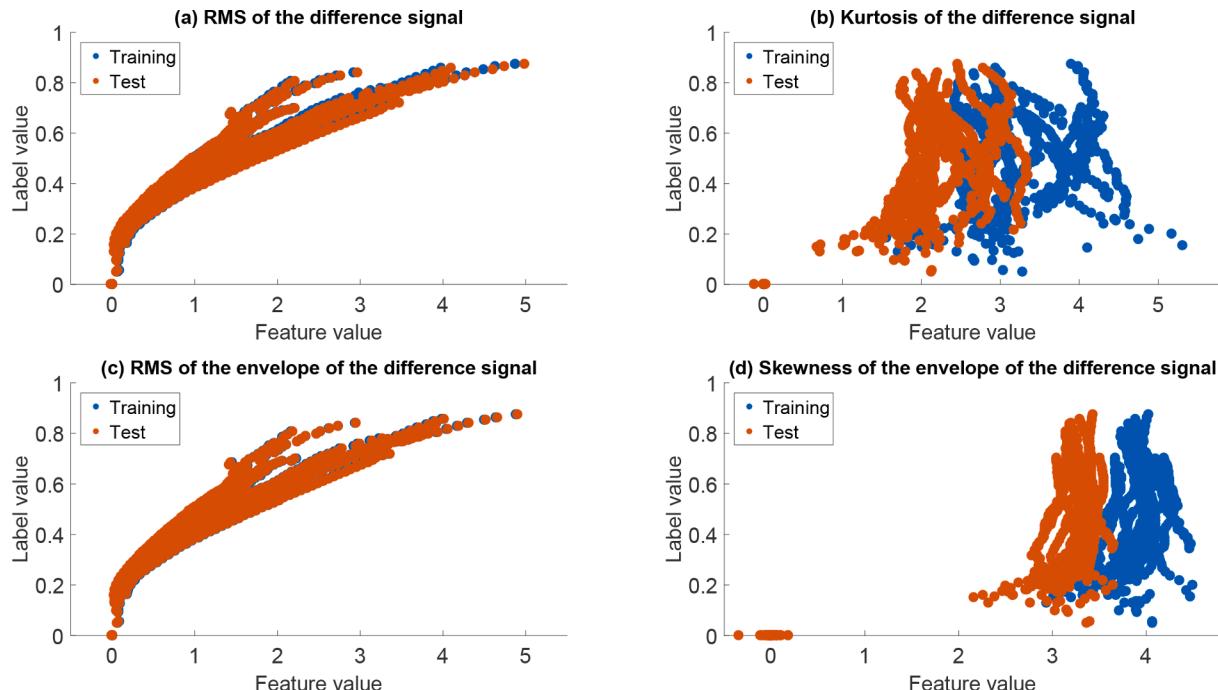
#### 5.1.2. Suppression of the harmonic signal

The harmonic signal is less affected by local tooth faults, as described in Section 2.2. The subtraction of the harmonic signal helps to suppress effects attributed to the tooth profile and other elements uncorrelated to the fault. However, the harmonic signal contains some information about the fault that is removed after the subtraction. Therefore, we

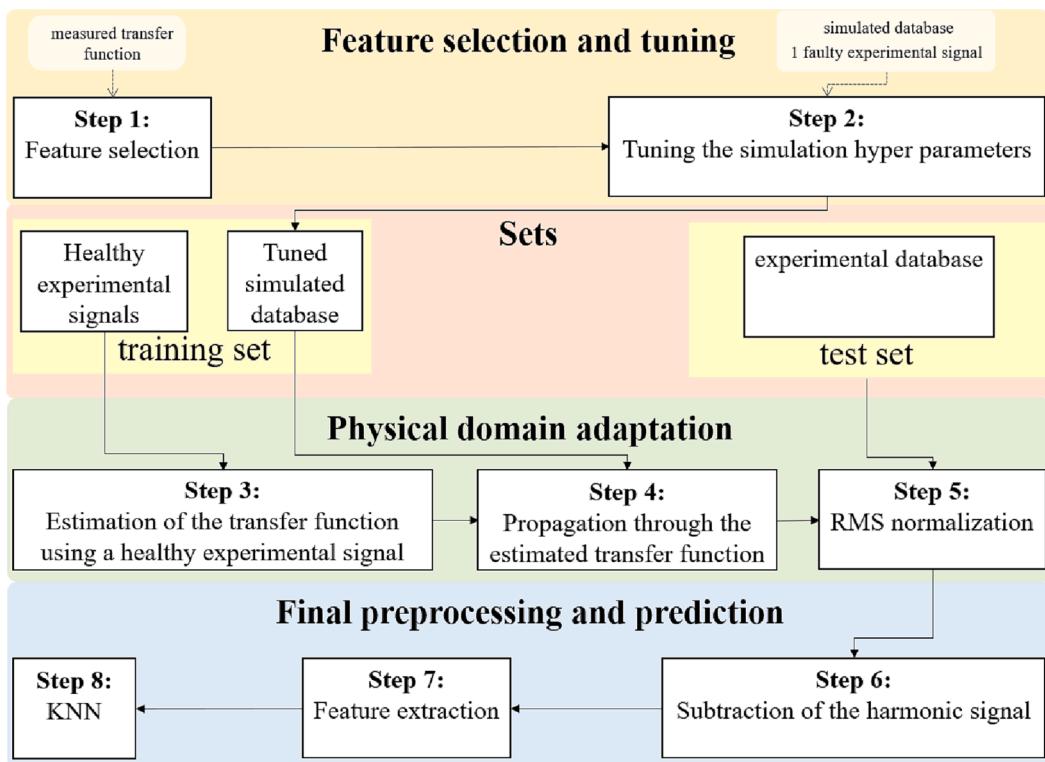
suggest suppressing the harmonic signal by basically subtracting a healthy signal from its calculated SA, as demonstrated in Fig. 9.

#### 5.1.3. Feature selection

As the estimation of the transfer function of the test set suffers errors, a complementary process is needed for addressing the unmitigated effects of the transfer function. Therefore, this study suggests a novel feature selection process, in which features that are invariant to the unmitigated effects are selected. The concept of the suggested feature selection algorithm is described in Fig. 10, including a Venn diagram and a detailed block diagram. The feature selection process selects the invariant features by calculating the error between the feature distributions of the training set  $X^{\text{training}}$  and the test set  $X^{\text{test}}$ , as presented in the



**Fig. 11.** Comparison of the four physical features from Section 2.3 using the feature selection process presented in Fig. 10. The features of the training and the test sets are presented after the Step 5 of the feature selection process.



**Fig. 12.** A block diagram of the new suggested algorithm for fault severity estimation of a gearbox by one-fault-shot learning.

block diagram in Fig. 10. The feature selection process consists of seven steps:

1. The test set  $X^{\text{test}}$  is generated by propagating the simulated data through a measured transfer function from the database described in Section 3.2 (see Fig. 7); i.e., the simulated signals are multiplied by an attenuation factor (e.g., if the attenuation is 10 dB the signals are multiplied by a factor of 0.1) and then propagated through the measured transfer function by convolution in the time domain.
2. The transfer function is estimated by ACS and minimum phase estimation by using a healthy signal from the test set  $X^{\text{test}}$ , as described in Section 5.1.1. The estimated transfer function is not identical to the measured transfer function from Step 1 due to errors in estimation process.
3. The training set  $X^{\text{training}}$  is propagated through the estimated transfer function of Step 2.
4. The RMS of  $X^{\text{test}}$  and  $X^{\text{training}}$  is normalized using the RMS normalization process described in Section 5.1.1.
5. The harmonic signal is subtracted from the calculated SA as explained in Section 5.1.2.
6. The features are extracted from  $X^{\text{test}}$  and  $X^{\text{training}}$ .
7. The error of the  $i^{\text{th}}$  feature between  $X^{\text{test}}$  and  $X^{\text{training}}$  is calculated based on the mean absolute differences among  $N$  examples, normalized by a scaling factor of the maximal standard deviation, as described in Fig. 10. In the current study, features with an error of less than 10% between  $X^{\text{test}}$  and  $X^{\text{training}}$  were selected.

The results are not sensitive to the threshold of 10% in Step 7 (e.g., all the range between 7% and 50% works well), and it is set an order of magnitude below a significant error of 100%, which represents a high deviation that makes the feature not relevant. Furthermore, this threshold creates a good separation between RMS of the difference and RMS of the envelope of the difference which have errors below 6% and to the kurtosis of the difference signal and the skewness of the envelope

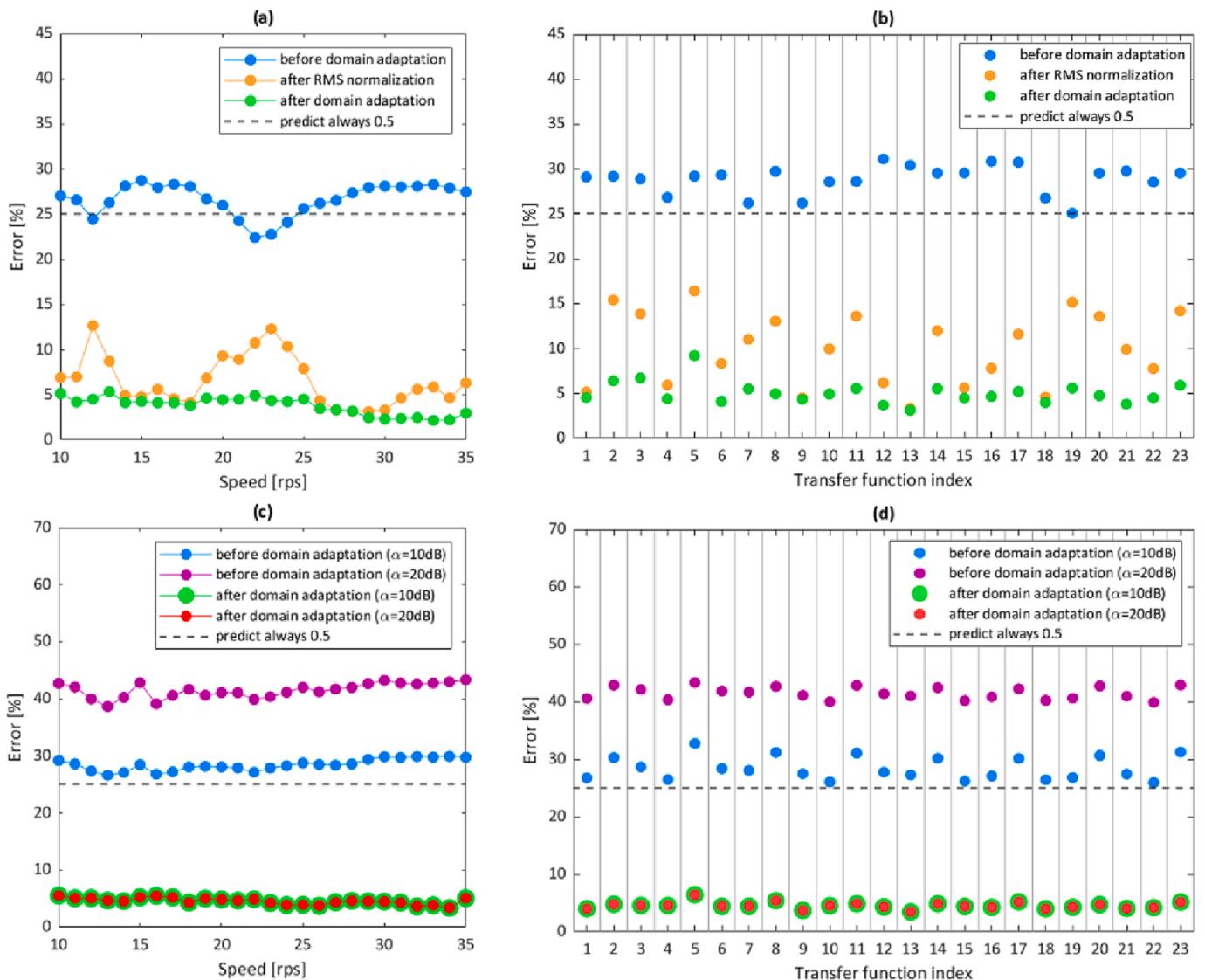
of the difference with errors above 100%.

An example of the examination of four features standardized according to the training set (i.e., subtraction of the mean and division by standard deviation) is presented in Fig. 11. For each feature in Fig. 11 the data are presented after Step 6 of the feature selection process, before the error calculation in Step 7 (*error*; in Fig. 10). The differences between the values of  $X^{\text{test}}$  and  $X^{\text{training}}$  are small for the features of the RMS of the difference [Fig. 11 (a), *err* = 2.0%] and RMS of the envelope of the difference [Fig. 11 (c), *err* = 4.3%] and are large for the features of the kurtosis of the difference signal [Fig. 11 (b) *err* = 132.3%] and the skewness of the envelope of the difference [Fig. 11 (d) *err* = 122.8%]. Thus, the RMS of the difference and the RMS of the envelope are selected because their deviation error is less than 10%.

In the current study, the four features were selected based on Bachar et al. [69] because they are all correlated with the fault size and have clear physical meaning. In cases where the features are not necessarily correlated with the fault size or in the case of too many features (may result in curse of dimensionality [71]), the feature selection process can be extended by applying a standard process for feature selection [74–76] (such as greedy selection approach, filters, and sparsity-induced norms such as L1 [71] or compression process [71], after the features with the high error between  $X^{\text{test}}$  and  $X^{\text{training}}$  have been eliminated).

#### 5.1.4. Simulation tuning

Most of the physical models are based on many assumptions and limitations that inevitably create differences between the simulated signals and real measured signals. A generalization from simulations to experimental data is improved when minimizing these differences to some extent. However, each simulation has unique properties and is constructed differently. Therefore, the process of mitigating the differences between the simulation and the reality relies on the researcher's knowledge about the model limitations. Hence, the simulation tuning described in the current section should be managed differently for each model.



**Fig. 13.** Results of simulated datasets. (a) Simulated datasets with different speeds for transfer function number 1 (from Section 3.2), 5Nm load, and attenuation of 10 dB. (b) Simulated datasets with different transfer functions of Section 3.2 for speed 15rps, 10Nm load, and attenuation of 10 dB. (c) Simulated datasets with different speeds, averaged over all transfer functions of Section 3.2 and loads. (d) Simulated datasets with different transfer functions of Section 3.2, averaged over all speeds and loads.

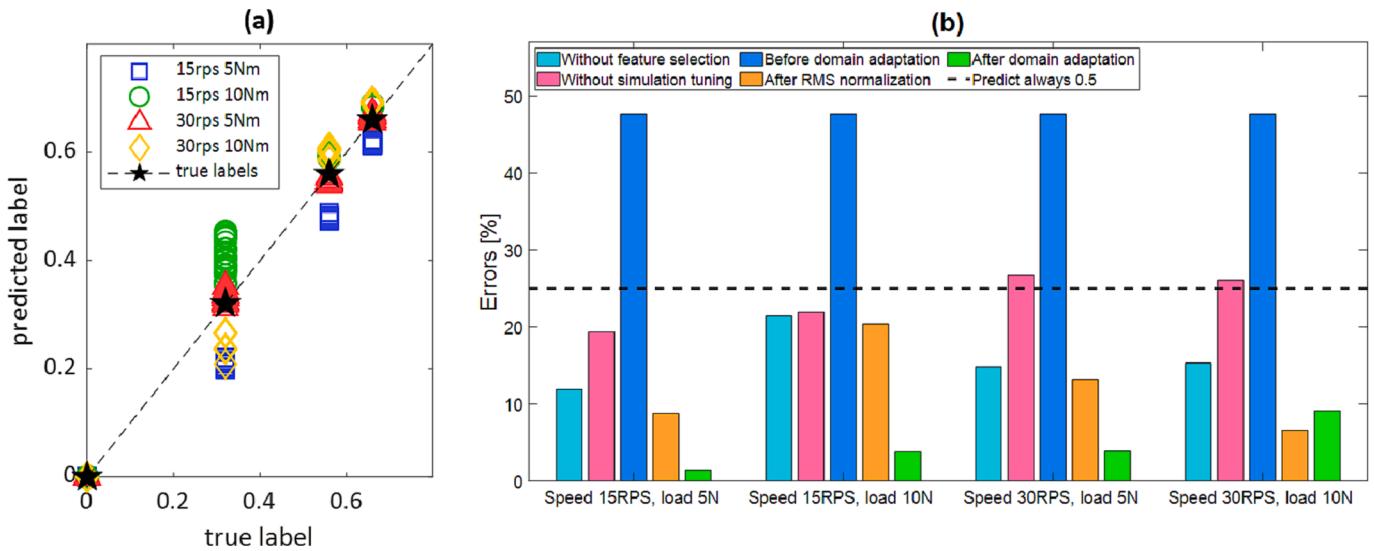
In the current study, the most significant problem of the model is the ratio between the signal stemming from the fault and the signal stemming from the tooth profile (see Fig. 5). This difference is mitigated by tuning the surface roughness grade of the simulated data, in the range between DIN1 (finest) and DIN12 (coarsest). The surface roughness grade is tuned by applying the new algorithm (described in Section 5.2) to the dataset in which the one FTFF experimental signal in the training set is used. The size of the FTFF example from the training set is estimated by the algorithm for each surface roughness grade. The surface roughness grade corresponding to the estimated value nearest to the actual FTFF size is selected.

## 5.2. The new algorithm

A block diagram of the new algorithm is presented in Fig. 12. The new algorithm has two main stages: (1) tuning the simulation to mitigate its differences from the reality using one faulty experimental signal as implemented in Step 2. (2) Suppressing the transfer function effects by estimating the transfer function and addressing its attenuation by

propagating the simulated data through the estimated transfer function, as implemented in Step 3 and Step 4, and RMS normalization in Step 5, while also selecting features which are invariant (i.e., not affected by) to the unmitigated effects of the transfer function as implemented in Step 1.

The new algorithm consists of the following steps: In Step 1, the feature selection process is applied according to the procedure described in Section 5.1.3 with a threshold of 10%. Then, in Step 2, the simulation is tuned according to the procedure described in Section 5.1.4. Steps 3, 4 and 5 suppress the effects of the transfer function, as described in Section 5.1.1. In Step 3, the transfer function is estimated by ACS and minimum phase using a healthy experimental signal, described in Section 2.1, while in Step 4 the simulated data in the training set is propagated through the estimated transfer function (as suggested in ref. [64]). In Step 5, the RMS normalization is applied (the test set is normalized according to healthy experimental signals in the training set). Then the harmonic signal is subtracted from each signal in Step 6, as described in Section 5.1.2. In Step 7, the selected features from Step 1 are extracted from the training and test sets. Finally, in Step 8, the KNN algorithm (described in Section 2.4) gets as a training the extracted features of the



**Fig. 14.** Algorithm results for the experimental dataset: (a) Predicted sizes of the experimental signals. (b) Errors of the estimated fault sizes without feature selection, without simulation tuning, before domain adaptation, after RMS normalization, and after full domain adaptation (RMS normalization + full transfer function suppression). The results of before domain adaptation, RMS normalization and full domain adaptation include the feature selection and simulation tuning steps.

training set with their corresponding fault sizes and predicts the fault sizes of the test set based on the extracted features of the test set. The training set is split 70%–30% into training and validation sets to determine the optimal number of nearest neighbors ( $K=\{1,2,\dots,100\}$ ) and the signals of the test set are standardized according to the healthy experimental signals in the training set.

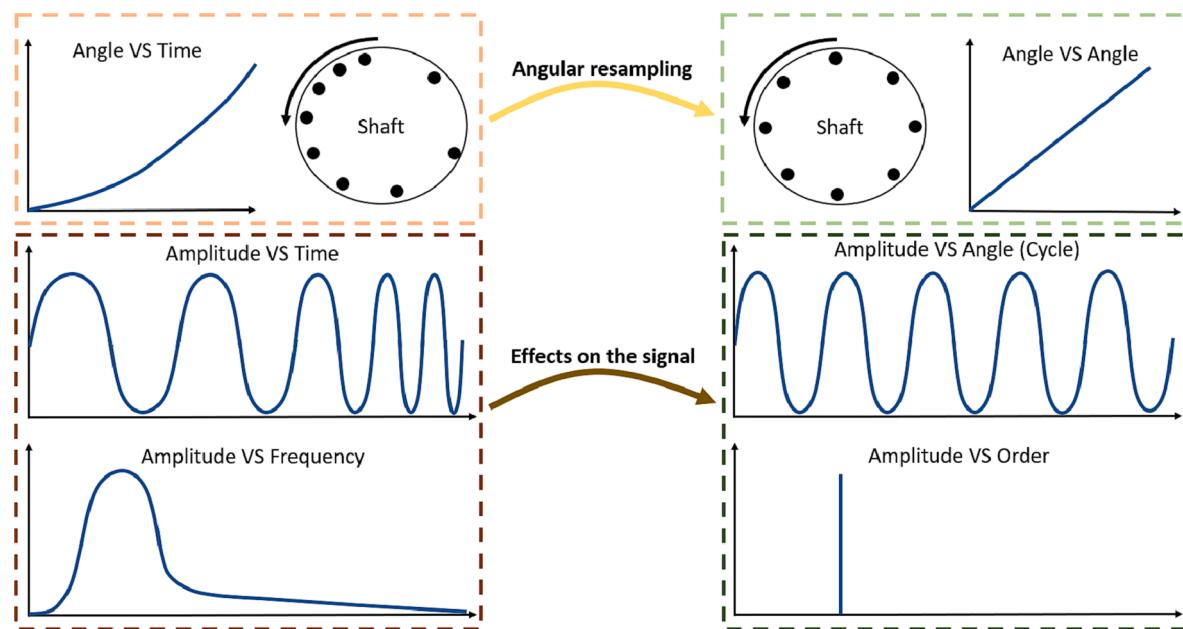
## 6. Results of the new algorithm

In this section the new algorithm is demonstrated in on the simulated datasets and the experimental datasets (see Fig. 8). The results were analyzed for datasets with different operating conditions, i.e., speed and load.

### 6.1. Demonstration of the new algorithm on the simulated datasets

The new algorithm (presented in Fig. 12) was applied to the simulated datasets described in Section 4.1 for each combination of speed, load, attenuation, and measured transfer function of Section 3.2 (see Table 1). Fig. 13 presents the results over the simulated datasets. For each dataset, the results are presented: (1) before the domain adaption process (i.e., before addressing the transfer function effects – attenuation and varied gain and phase), (2) after the RMS normalization which addresses the attenuation effect and (3) after the full domain adaption process (Steps 3, 4 and 5 in Fig. 12).

From Fig. 13 (a) and (b), we can notice significant improvement in the results after RMS normalization, and even more improvement after the full domain adaption (overall from an error of over 25% to less than 10%). Recall that the energy levels of the simulated signals in the



**Fig. 15.** An illustration of angular resampling, based on ref. [78].

**Table 2**  
Model Components' Parameters.

Parameter	Pinion	Gear
Teeth number	17	38
Module (mm)	3	3
Tooth width (m)	0.015	0.015
Pressure angle	20°	20°
Base helix angle	0°	0°
Young modulus ( $N/m^2$ )	$210 \times 10^9$	$210 \times 10^9$
Poisson's ratio	0.3	0.3
Components' masses (Kg)	0.8	3.2
Modal damping ratios	$\zeta_n = 0.05$ (for all modes)	$\zeta_n = 0.05$ (for all modes)

training set were significantly higher than the attenuated simulated signals in the test set, thus, decreasing the performance of the KNN dramatically when applied without domain adaptation. This result demonstrates the first contribution of this study, i.e., the advantages of a hybrid approach which makes use of machine-learning and signal processing methods, as the physical domain adaptation significantly improves the performance of KNN.

We can notice the sensitivity of the performance after only the RMS normalization to different rotational speeds and transfer functions. Because gear signals with different rotational speeds have different spectral positions, the effects of the transfer function are different. Different transfer functions have different gain and phase; hence, they have different effects on the gear signal. From Fig. 13 (c) and (d), we can notice that the attenuation level does not change the prediction error. This observation can be easily explained by the RMS normalization process which eliminates the effects of the attenuation.

## 6.2. Demonstration on the experimental datasets

The new algorithm was applied to the four experimental datasets described in Section 4.2 for each combination of speed and load,

separately (see Table 1). Similar to Section 6.1, for each dataset, the results are presented in Fig. 14 (b): (1) before the domain adaption process; (2) after the RMS normalization; and (3) after the full domain adaptation process (Steps 3, 4 and 5 in Fig. 12). Furthermore, the results without feature selection process and simulation tuning (Step 1 and Step 2 in Fig. 12) are also presented for demonstrating their contribution.

Fig. 14 presents the results of the new algorithm. According to the error analysis presented in Fig. 14 (b), the feature selection process, simulation tuning, RMS normalization and the full domain adaptation process have a significant contribution to the estimation of the fault size. We can notice that the errors of the algorithm without applying domain adaptation are high and similar for all the four datasets. Recall that the values of the simulated data are much higher than the values of the measured data, resulting in predicting a fault size of zero ( $y = 0$ ) for all of the test sets. The RMS normalization suppresses some of the effects of the transfer function, while the mitigation of the gain and the phase of the transfer function helped to suppress its effects even more (an error of less than 10%) for all four datasets. In addition, there were differences in the results for different speeds, as discussed previously in Section 6.1, considering Fig. 13 (a). There were also differences in the results for different loads. Different loads in the experimental system generate different transfer functions, and different transfer functions may result in different results, as discussed previously in Section 6.1 considering Fig. 13 (b).

The improvement in the results through simulation tuning demonstrates the second contribution of the study: better generalization from simulation to reality through the use of one-fault-shot learning to address the inevitable differences between simulation and reality. The improvement in the results through the feature selection process demonstrates the third contribution of the paper: a novel method for selecting features that are invariant to the unmitigated effects of the transfer function.

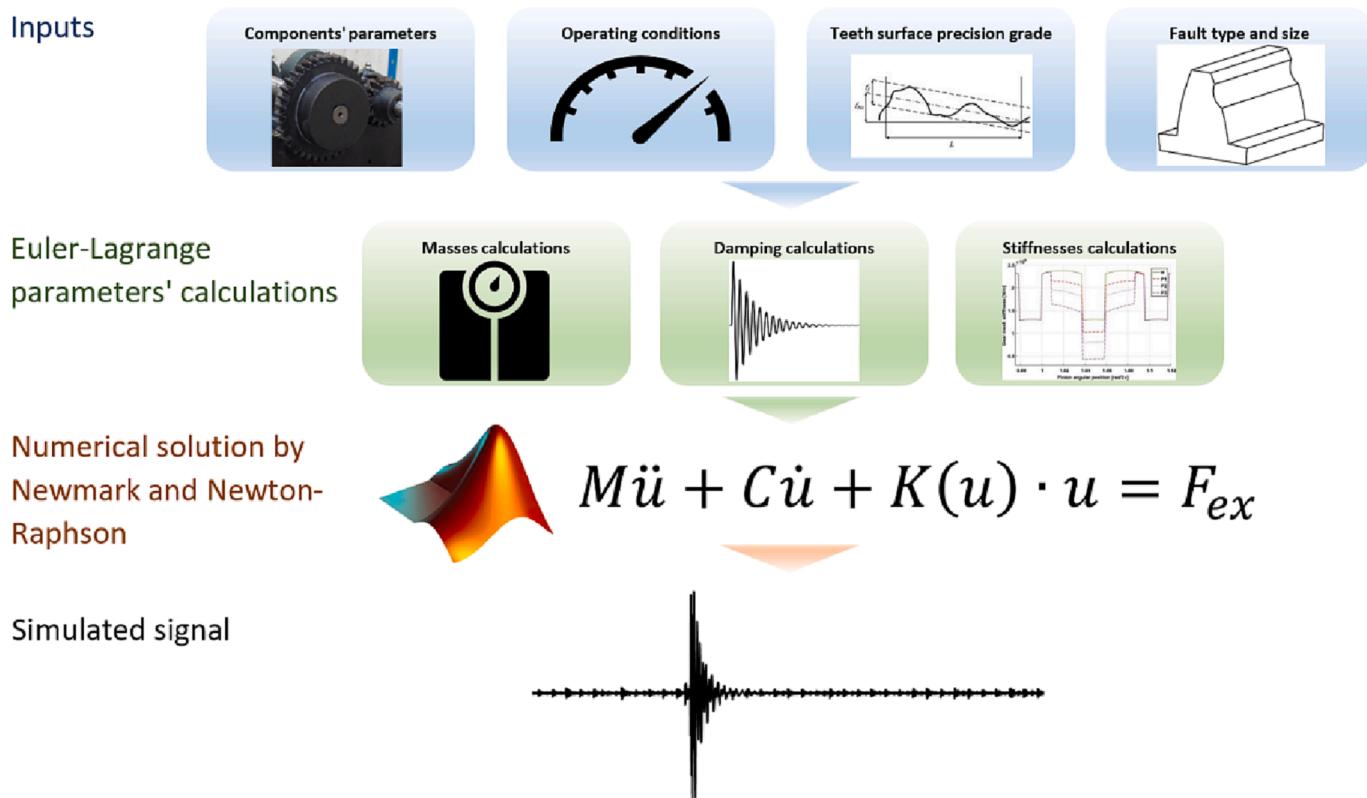
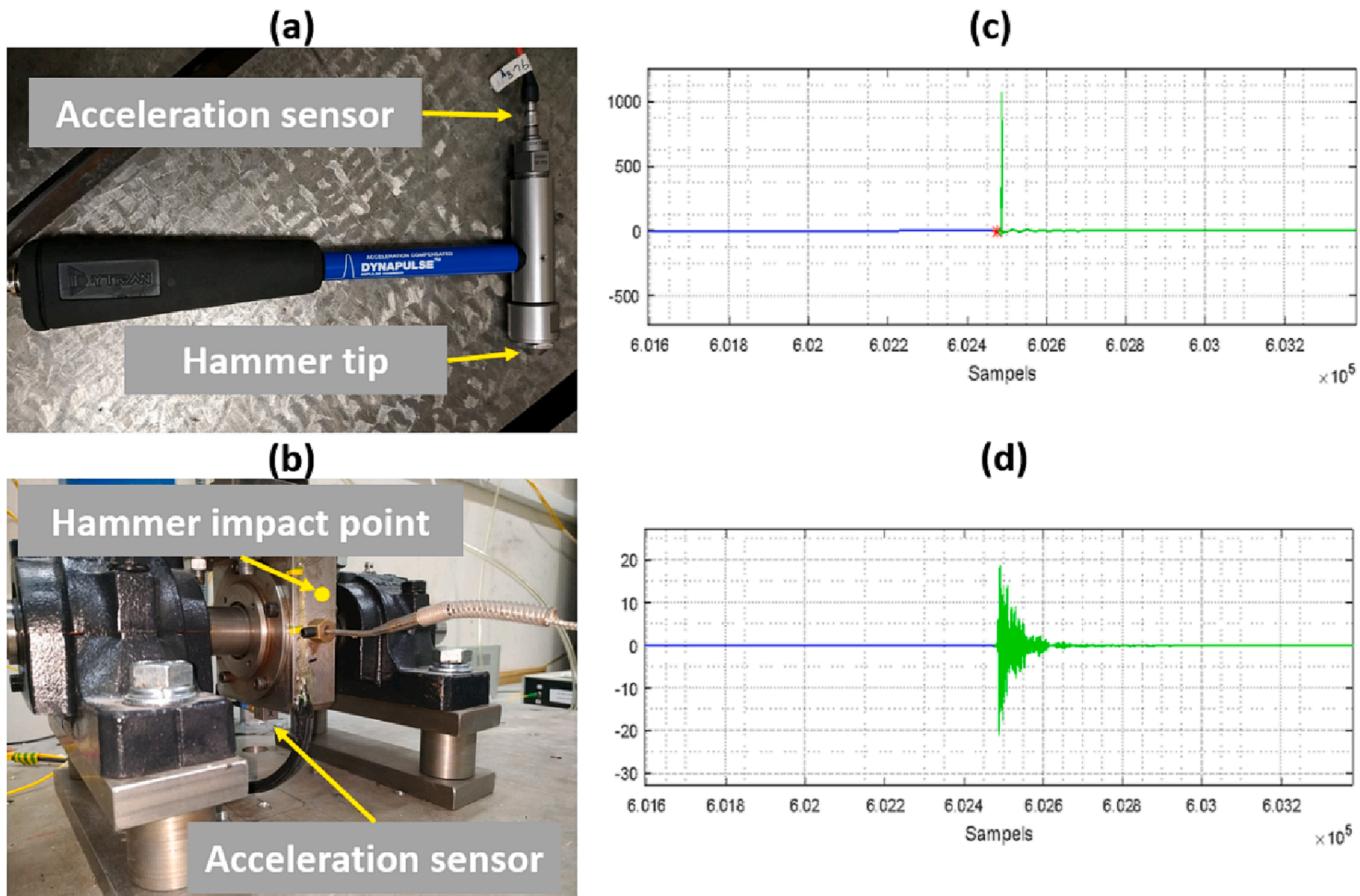


Fig. 16. Block diagram of the realistic dynamic model of Dadon et al. [79] that was used to generate the simulated signals in the current study.



**Fig. 17.** Measurement of a transfer function by impact hammer testing. The transfer function is measured by processing the measured (c) impulse signal which is generated by (a) the hammer and the measured (d) response signals which is measured by the acceleration sensor mounted on the (b) test rig. The hammer type is Dytan 5800B2 and the acceleration sensors type is Dytan 3053B2.

## 7. Summary and conclusions

Estimation of the fault size of gears is not a typical machine-learning task. In many rotating machines, such as helicopters and airplanes, regular machine-learning or deep-learning algorithms cannot be successfully applied directly to the tested system because faulty data – labeled or unlabeled – is sparse. This study suggested a novel algorithm for fault severity estimation that overcomes this challenge by using simulated signals for training, and physical domain adaptation process for suppressing the differences between the source and target domains. The new algorithm reduces the differences between the simulated and measured signals by using one measured faulty signal for tuning the simulation, and by the suppression of the transfer function effects. The algorithm learns patterns in the tuned simulated signals using the selected physical features that are invariant to the unmitigated effects of the transfer function.

The new algorithm was demonstrated on a large, simulated database including different combinations of speed, load, and measured transfer function, as well as through experimentation. It obtained good results on the simulated and experimental datasets – achieving an error of less than 10%. Parts of the new algorithm can be applied separately in other applications, for example, by applying only the domain adaptation and the feature selection processes in the case where a large dataset of experimental signals is available. This paves the way for one-fault-shot learning algorithms and, in the future, for more sophisticated algorithms for zero-fault-shot learning.

## Declaration of Competing Interest

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

## Data availability

Data will be made available on request.

## Appendix A

### 1. Angular resampling

In almost every rotating machinery, even under stationary conditions, the rotational shaft's speed fluctuates, and as a result the spectral analysis is deteriorated [1]. Angular resampling is a common process for addressing this problem [77], according to which the signal is converted from time representation to cycle representation. I.E., the vibration signal is synchronous to the cumulative angle of the shaft, as illustrated in Fig. 15. Angular resampling is an integral stage in the processing of vibration signals from rotating elements such as gears, bearing, etc. [65].

The implementation of angular resampling is discussed in several references, including refs. [1,77]. Basically, angular resampling can be implemented by calculating the cumulative angle of the shaft as function of the time, based on the measured speed, and then the vibration signal

can be resampled according to time intervals that correspond to even-angle intervals; Hence the name - angular resampling - as the signal is resampled according to the angle.

According to the aforementioned implementation, angular resampling has two parameters that need to be set: interpolation method and up-sampling rate, both for the resampling stage. In this study, the interpolation method is cubic Spline, and the up-sampling rate is set to 2.

## 2. Realistic dynamic model

The simulated signals were generated by a realistic dynamic model of Dadon et al. [79] that was examined by several studies [9, 26, 80]. The model gets as input the properties of the gear transmission, the operating condition (velocity and load), and the fault properties. The properties of the simulated signals for this study are presented in Table 1, while the component parameters are presented in Table 2.

The dynamic model is based on Euler-Lagrange equations of motion. The time variant gearmesh stiffness is calculated based on beam theory, in addition to the calculation of the constant mass and damping matrices. Then, the differential equations are solved numerically, using Newmark [81] and Newton-Raphson methods. The main components of the realistic dynamic model are depicted in Fig. 16.

## 3. Measurement of the transfer functions

The transfer functions of Section 3.2 were measured by impact hammer testing on several laboratory rest rigs and on several impact points. Overall, there were 8 different impact points, where for each point 3 transfer functions were measured in all three axes, namely X, Y and Z. Thus, overall, 24 transfer functions were measured. Each transfer function was measured by using the measured acceleration signal of the sensor mounted on the hammer (Fig. 17 (a)) and the acceleration sensor on the test rig. An example of the location of an acceleration sensor and of an impact point is depicted in Fig. 17 (b). Each transfer function was measured by averaging several different impulse-response signals. An example of an impulse signals is depicted in Fig. 17 (c) and of a response signal in Fig. 17 (d).

## References

- [1] Randall RB. Vibration-based Condition Monitoring - Industrial, Aerospace and Automotive Applications. 1st ed. Chichester, West Sussex, United Kingdom: WILEY; 2010. Doi: 10.1002/9780470977668.
- [2] E.P. Carden, P. Fanning, Vibration based condition monitoring: a review Struct Heal Monit 2004;3:355-77. Doi: 10.1177/1475921704047500.
- [3] R.B. Randall, *State of the art in monitoring rotating machinery - Part 1*, Sound Vib 38 (2004) 14–21.
- [4] Z. Tian, M.J. Zuo, Health condition prediction of gears using a recurrent neural network approach, IEEE Trans. Reliab. 59 (2010) 700–705, <https://doi.org/10.1109/TR.2010.2083231>.
- [5] P. Lyu, K. Zhang, W. Yu, B. Wang, C. Liu, A novel RSG-based intelligent bearing fault diagnosis method for motors in high-noise industrial environment, Adv Eng Informat. (2022) 52, <https://doi.org/10.1016/J.AEI.2022.101564>.
- [6] P. Kundu, A.K. Darpe, M.S. Kulkarni, A review on diagnostic and prognostic approaches for gears, Struct Heal Monit (2020) 1–41, <https://doi.org/10.1177/1475921720972926>.
- [7] P.J. Dempsey, D.G. Lewicki, D.D. Le, Investigation of current methods to identify helicopter gear health, IEEE Aerosp. Conf. Proc. (2007) 1–13, <https://doi.org/10.1109/AERO.2007.352844>.
- [8] El Badaoui M, Cahouet V, Guillet F, Danie're J, Velex P. Modeling and Detection of Localized Tooth Defects in Geared Systems. J Mech Des 2001;123:422-30. Doi: 10.1115/1.1349420.
- [9] I. Dadon, N. Koren, R. Klein, M.G. Lipsett, J. Bortman, Impact of gear tooth surface quality on detection of local faults, Eng. Fail. Anal. (2020) 108, <https://doi.org/10.1016/j.englfailanal.2019.104291>.
- [10] N. Tandon, A. Choudhury, Review of vibration and acoustic measurement methods for the detection of defects in rolling element bearings, Tribol. Int. 32 (1999) 469–480, [https://doi.org/10.1016/S0301-679X\(99\)00077-8](https://doi.org/10.1016/S0301-679X(99)00077-8).
- [11] C. Peeters, P. Guillaume, J. Helsen, A comparison of cepstral editing methods as signal pre-processing techniques for vibration-based bearing fault detection, Mech. Syst. Sig. Process. 91 (2017) 354–381, <https://doi.org/10.1016/j.ymssp.2016.12.036>.
- [12] G. Kogan, R. Klein, J. Bortman, A physics-based algorithm for the estimation of bearing spall width using vibrations, Mech. Syst. Sig. Process. 104 (2018) 398–414, <https://doi.org/10.1016/j.ymssp.2017.11.011>.
- [13] D.L. Dewell, L.D. Mitchell, Detection of a misaligned disk coupling using spectrum analysis, J. Vib. Acoust. Stress. Reliab. Des. 106 (1984) 9–16, <https://doi.org/10.1115/1.3269161>.
- [14] A.D. Dimarogonas, C.A. Papadopoulos, Vibration of cracked shafts in bending, J. Sound Vib. 91 (1983) 583–593, [https://doi.org/10.1016/0022-460X\(83\)90834-9](https://doi.org/10.1016/0022-460X(83)90834-9).
- [15] P. Raharjo, An investigation of surface vibration, airbourne sound and acoustic emission characteristics of a journal bearing for early fault detection and diagnosis, University of Huddersfield, 2013.
- [16] N. Mokhtari, M. Grzeszkowski, C. Gühmann, Vibration signal analysis for the lifetime-prediction and failure detection of future turbofan components, Tech Mech - Eur J Eng Mech 37 (2017) 422–431, <https://doi.org/10.24352/UB.OVGU-2017-118>.
- [17] P. Borghesani, P. Pennacchi, R.B. Randall, N. Sawalhi, R. Ricci, Application of cepstrum pre-whitening for the diagnosis of bearing faults under variable speed conditions, Mech. Syst. Sig. Process. 36 (2013) 370–384, <https://doi.org/10.1016/j.ymssp.2012.11.001>.
- [18] Y. Lei, Intelligent fault diagnosis and remaining useful life prediction of rotating machinery. 1st ed. Oxford: Butterworth-Heinemann; 2017. Doi: 10.1016/C2016-0-00367-4.
- [19] H.N. Özgür, D.R. Houser, Dynamic analysis of high speed gears by using loaded static transmission error, J. Sound Vib. 125 (1988) 71–83, [https://doi.org/10.1016/0022-460X\(88\)90416-6](https://doi.org/10.1016/0022-460X(88)90416-6).
- [20] F. Chaari, T. Fakhfakh, M. Haddar, Analytical modelling of spur gear tooth crack and influence on gearmesh stiffness, Eur J Mech - A/Solids 28 (2009) 461–468, <https://doi.org/10.1016/J.EUROMECHSOL.2008.07.007>.
- [21] R. Klein, E. Rudyk, E. Masad, M. Issacharoff, Model based approach for identification of gears and bearings failure modes, Int J Progn Heal Manag 2 (2011) 1–10.
- [22] E. Madar, R. Klein, J. Bortman, Contribution of dynamic modeling to prognostics of rotating machinery, Mech. Syst. Sig. Process. 123 (2019) 496–512, <https://doi.org/10.1016/j.ymssp.2019.01.003>.
- [23] Y.A. Yucesan, A. Dourado, F.A.C. Viana, A survey of modeling for prognosis and health management of industrial equipment, Adv Eng Informatics (2021) 50, <https://doi.org/10.1016/J.AEI.2021.101404>.
- [24] W.D. Mark, Time-synchronous-averaging of gear-meshing-vibration transducer responses for elimination of harmonic contributions from the mating gear and the gear pair, Mech. Syst. Sig. Process. 62–63 (2015) 21–29, <https://doi.org/10.1016/J.YMSSP.2015.03.006>.
- [25] E. Bechhoefer, M. Kingsley, A review of time synchronous average algorithms, Annu. Conf. Progn. Heal. Manag. Soc. PHM (2009, 2009,) 1–10.
- [26] L. Bachar, I. Dadon, R. Klein, J. Bortman, The effects of the operating conditions and tooth fault on gear vibration signature, Mech. Syst. Sig. Process. (2021) 154, <https://doi.org/10.1016/j.ymssp.2020.107508>.
- [27] A. Stetco, F. Dinmohammadi, X. Zhao, V. Robu, D. Flynn, M. Barnes, et al., Machine learning methods for wind turbine condition monitoring: A review, Renew. Energy 133 (2019) 620–635, <https://doi.org/10.1016/J.RENENE.2018.10.047>.
- [28] A. Zhou, D. Yu, W. Zhang, A research on intelligent fault diagnosis of wind turbines based on ontology and FMECA, Adv Eng Informatics 29 (2015) 115–125, <https://doi.org/10.1016/J.AEI.2014.10.001>.
- [29] H. Qiu, J. Lee, J. Lin, G. Yu, Robust performance degradation assessment methods for enhanced rolling element bearing prognostics, Adv Eng Informatics 17 (2003) 127–140, <https://doi.org/10.1016/J.AEI.2004.08.001>.
- [30] A. Kumar, C.P. Gandhi, Y. Zhou, R. Kumar, J. Xiang, Latest developments in gear defect diagnosis and prognosis: A review, Meas J Int Meas Confed (2020) 158, <https://doi.org/10.1016/j.measurement.2020.107735>.
- [31] W. Li, X. Zhong, H. Shao, B. Cai, X. Yang, Multi-mode data augmentation and fault diagnosis of rotating machinery using modified ACGAN designed with new framework, Adv Eng Informatics 52 (2022), 101552, <https://doi.org/10.1016/J.AEI.2022.101552>.
- [32] Z. Liu, M.J. Zuo, H. Xu, Feature ranking for support vector machine classification and its application to machinery fault diagnosis, Proc Inst Mech Eng Part C J Mech Eng Sci 227 (2013) 2077–2089, <https://doi.org/10.1177/09546212469757>.
- [33] C. Li, R.V. Sanchez, G. Zurita, M. Cerrada, D. Cabrera, R.E. Vásquez, Multimodal deep support vector classification with homologous features and its application to gearbox fault diagnosis, Neurocomputing 168 (2015) 119–127, <https://doi.org/10.1016/j.neucom.2015.06.008>.
- [34] W. Lu, W. Jiang, G. Yuan, L. Yan, A gearbox fault diagnosis scheme based on near-field acoustic holography and spatial distribution features of sound field, J. Sound Vib. 332 (2013) 2593–2610, <https://doi.org/10.1016/j.jsv.2012.12.018>.
- [35] M.H. Gharavani, F. Almas Ganj, A.R. Ohadi, B.H. Heidari, Comparison of FDA-based and PCA-based features in fault diagnosis of automobile gearboxes, Neurocomputing 121 (2013) 150–159, <https://doi.org/10.1016/j.neucom.2013.04.033>.
- [36] Z. Li, X. Yan, Z. Tian, C. Yuan, Z. Peng, L. Li, Blind vibration component separation and nonlinear feature extraction applied to the nonstationary vibration signals for the gearbox multi-fault diagnosis, Meas J Int Meas Confed 46 (2013) 259–271, <https://doi.org/10.1016/j.measurement.2012.06.013>.
- [37] S. Park, S. Kim, J.H. Choi, Gear fault diagnosis using transmission error and ensemble empirical mode decomposition, Mech. Syst. Sig. Process. 108 (2018) 58–72, <https://doi.org/10.1016/j.ymssp.2018.02.028>.

