4-6-2024

On the Feasibility of Ultrasonic Full Waveform Evaluation with Changing Testing Conditions for the Quality Control of Manufacturing Parts

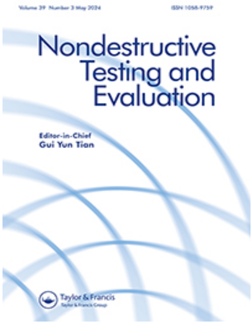
Simon Schmid
Technical University of Munich, Munich, Germany

Thomas Schumacher
Portland State University, thomas.schumacher@pdx.edu

Christian U. Grosse
Technical University of Munich, Munich, Germany

Citation Details

Schmid, S., Schumacher, T., & Grosse, C. U. (2024). On the feasibility of ultrasonic full waveform evaluation with changing testing conditions for the quality control of manufacturing parts. Nondestructive Testing and Evaluation, 1–26.



Taylor & Francis
Taylor & Francis Group

Nondestructive Testing and Evaluation

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/gnte20

On the feasibility of ultrasonic full waveform evaluation with changing testing conditions for the quality control of manufacturing parts

Simon Schmid, Thomas Schumacher & Christian U. Grosse

To cite this article: Simon Schmid, Thomas Schumacher & Christian U. Grosse (04 Apr 2024): On the feasibility of ultrasonic full waveform evaluation with changing testing conditions for the quality control of manufacturing parts, Nondestructive Testing and Evaluation, DOI: 10.1080/10589759.2024.2337064

To link to this article: https://doi.org/10.1080/10589759.2024.2337064

NONDESTRUCTIVE TESTING AND EVALUATION
https://doi.org/10.1080/10589759.2024.2337064

On the feasibility of ultrasonic full waveform evaluation with changing testing conditions for the quality control of manufacturing parts

Simon Schmid[a], Thomas Schumacher[b] and Christian U. Grosse[a]

[a]TUM School of Engineering and Design, Department of Materials Engineering, Technical University of Munich, Munich, Germany; [b]Department of Civil and Environmental Engineering, Portland State University, Portland, USA

ABSTRACT

Fast volumetric non-destructive testing methods are needed, especially for quality control in manufacturing lines. Ultrasonic testing with full waveform evaluation is a promising method for this. However, changes in coupling conditions or environmental factors can significantly alter the ultrasound signal, sometimes more than actual defects. This study investigates the effect of various factors on the ultrasound signal based on a Monte Carlo study with wavefield simulations. The test specimens comprise aluminium plates with holes of varying sizes and positions. Using both experimental as well as simulated data, the performance of two commonly used comparison metrics, namely the R^2 score and the magnitude-squared coherence integral, for detecting defects in manufactured parts is evaluated. It was found that the magnitude-squared coherence integral is more robust against random influences than the R^2 score. Additionally, factors influencing the entire plate exhibit the most significant impact on the signals. The hole positions and dimensions change the signals and the value of the comparison metrics significantly and are difficult to distinguish by one metric. A deep learning model, however, is capable of performing this task and it outperforms the comparison metrics in defect detection. The performance of the approaches is assessed with probability of detection curves.

ARTICLE HISTORY
Received 25 January 2024
Accepted 21 March 2024

KEYWORDS
Full waveform comparison; automation; ultrasound; environmental influences; deep learning; simulation; probability of detection

1. Introduction

In large-scale manufacturing facilities, a main objective is the efficient production of each component, ensuring cost-effectiveness. To achieve this goal, all processes are tailored to minimise time consumption, including the integration of quality control into the production line. Common quality assurance methods are non-destructive testing (NDT) methods. NDT methods that enable volumetric inspections, like X-ray computed tomography (CT), are often time-consuming and not suitable choices for a fast-paced manufacturing environment. Further, in such use cases, it is beneficial if the testing method

CONTACT Thomas Schumacher thomas.schumacher@pdx.edu

and the evaluation can be integrated into the production and automated [1]. In this study, the feasibility of a specific type of ultrasonic testing is studied for use in such applications.

One of the benefits of ultrasound is that it can provide volumetric information. However, the conventional analysis approach involves evaluating only the direct reflections from defects, resulting in a time-consuming process when an entire volume should be measured [2]. Alternatively, ultrasound offers also an indirect method to assess deviations across the entire volume of a part. By introducing an ultrasonic wave into a part and measuring it after a sufficiently long propagation time, deviations in the entire part can be examined. Depending on the size of a part, more than one sensor may be required to cover the entire volume due to wave attenuation. The recorded waveform carries information about any defects within the material. Extracting this information is challenging, as it is not always evident how a waveform is affected in response to the presence of a defect. One method for extracting this information involves comparing the full waveforms measured from a reference sample (deemed defect-free) and the sample to be examined. For the comparison, several similarity metrics can be used (see, e.g [3–5]).

The same technique is employed to assess subtle variations in scattering materials such as concrete [6–9], glass fibre-reinforced plastic (GFRP) [10] or aluminium with multiple circular holes [11]. In this context, a notable portion of the energy from the ultrasound wave arrives after the first arrival, forming what is known as the coda, which is the tail of the signal. The comparison of this tail to a reference is called coda wave interferometry (CWI) [12,13]. Hereby, the stretching of one signal with respect to another is estimated allowing for the determination of the relative velocity change dv/v [14]. When the material remains unchanged, the path of a propagating ultrasonic wave remains constant, resulting in an unaltered captured waveform. The coda wave proves highly sensitive to minor changes in the material [6]. Typically, this method is employed to monitor material changes over time, such as fluctuations in stress levels, which is related to acousto-elastic theory.

In Wang et al. [10], ultrasound full waveform comparison based on the squared correlation coefficient R^2 is reported. Here, a transmitter and receiver are attached to a glass fiber-reinforced polymer (GFRP) plate with dimensions, 102.0 × 80.5 × 19.8 mm. The transmitter excites an ultrasound wave with a center frequency of 100 kHz and holes of increasing diameter are drilled into the plate during measurements. It was found that the R^2 score decreases as the hole diameter increases. Further, the hole diameter and the R^2 score are linearly related, and holes significantly smaller than the wavelength could be detected. A similar study was conducted by Chen and Schumacher [15] for steel plates with an increasing notch length. As a similarity metric, the magnitude squared coherence integral (MSCI) was used (see Subsection 2.1). It was found that the MSCI value decreases for an increasing notch size. It should be noted that in both of these studies, the transducers were glued to the test specimen, to ensure constant coupling conditions. Also, temperature variations were not significantly large enough to affect the measurements.

Lu and Michaels [5] conducted a study examining the impact of temperature on ultrasound signals. Here measurements on an aluminium plate (50.8 × 152.4 × 6.35 mm) with a hole and a single transmitter-receiver setup were conducted. This is a very similar setup as used in the presented study. Notably, in Lu and Michaels [5] the transmitter and receiver remained stationary while the holes were progressively drilled to larger dimensions during the measurements. Moreover, the temperature was systematically adjusted between 8.9 and 37.8 °C throughout the measurement process. This temperature variation allowed the generation of a comprehensive baseline dataset. By incorporating a reference signal measured at a similar temperature, they could effectively compensate for the temperature effects observed in the ultrasound signals.

Besides metrics used for evaluating changes in ultrasound signals, neural networks can be employed to automatically estimate suitable features to detect subtle differences between signals. In Ebrahimkhanlou et al. [16] it was possible to estimate the source locations of acoustic emission (AE) events, with a single sensor at a fixed position. The AE events were introduced through an HSU Nielsen test at 13 different positions. The specimen was an aluminium plate with rivets of the dimensions 914.4 × 914.4 × 3.2 mm. The captured signals were transformed into an image using the wavelet transform, on which a stacked autoencoder and a convolutional neural network were trained. Both models were able to accurately estimate the positions of the sources. It shall be noted that the presented approach was a classification problem, which inherently constrains the range of possible source positions to be estimated.

Another study using machine learning methods for defect detection with ultrasound is presented in Zhao et al. [17] They used an adapted gcForest algorithm (AWGA-gcForest) for defect detection on a phased array full matrix capture dataset. The proposed algorithm outperforms other benchmark models in terms of accuracy and computation time.

A more established approach for volumetric defect localisation in plate-like structures involves utilising vibration measurements to assess Eigenfrequencies and mode shapes. Nonetheless, this method faces challenges such as symmetry issues, ill-posedness of the inverse problem, and limited sensitivity to small defects, arising from minor alterations in the mass or stiffness of the structure. Moreover, in the case of bulk materials, these constraints are increased, since the relative change of mass and stiffness are even smaller in that case. For studies about the application of modal analyses for defect detection or localisation, please refer to, e.g. Tufoi et al. [18], Le et al. [19], and Cheng et al. [20].

In the literature, the quantification of a damage extent based on full waveform comparison has already been studied. However, in these applications, the transmitter and receiver positions were kept constant, either glued onto the specimen [10,15] or embedded inside the material [6,7] to minimise the effect of varying coupling conditions. This is not feasible when different parts need to be investigated, as the transmitter and receiver have to be deployed in-situ to each part to be examined. In such cases, other factors come into play that can distort a recorded ultrasound waveform and affect the reproducibility of measurements. A significant factor among these is variations in coupling [21], including the force and position of the transmitter. These influences can be more pronounced than the changes due to defects in the material. Temperature variation, which can also notably affect the measurements (see, e.g [5]), and changes in the dimensions of the specimens need to be considered as well.

In this study, the influence of changing coupling conditions, transducer placement, specimen dimensions, and temperature variations on the ability to detect holes in a simple manufactured component using ultrasonic full waveform comparison are discussed. To achieve this, two waveform comparison metrics and deep learning models are evaluated. To estimate which hole size is still detectable under the given setup, probability of detection (POD) curves are utilised. The specimens are aluminium plates with dimensions of 140 × 100 × 3 mm, featuring varying hole diameters and positions. The same measurement setup is also simulated with randomised environmental and coupling influences. To the best of the author's knowledge, it is the first study systematically investigating the applicability of the use of full waveform comparison when different specimens are investigated.

This paper is structured as follows: Section 2 gives the relevant background of the comparison metrics, time series-based deep learning, and POD curves. The investigated specimens together with the measurement and simulation setup are described in Section 3. The results are presented and discussed in Section 4. Here, first, the optimal signal length for the comparison metrics is estimated, along with the influence of the hole position on the comparison metrics. Subsequently, the response of the two metrics R^2 score and MSCI to an increase in the hole diameter is evaluated, and the best-performing metric, which is the MSCI, is chosen for all further evaluations. Following that, the effect of each influence factor is estimated based on simulated data. Furthermore, a deep learning model time series transformer (TST) is trained on the simulated and measured dataset. The performance in estimating the hole diameter is evaluated with POD curves and compared to the POD curves based on the evaluation with MSCI. Concluding remarks are provided in Section 5.

2. Background

This section introduces the full waveform comparison metrics and the deep learning model. Furthermore, it provides relevant background on the $\hat{a}$ vs. a POD evaluation.

2.1. Full waveform comparison metrics

In the literature, multiple metrics exist for the comparison of the full waveform or the coda of ultrasound signals [7,14,22]. All of them compare a reference waveform $u(t)$ to a waveform $\tilde{u}(t)$ that was recorded after a permutation had been introduced in the material [6]. One approach to detect changes in the material is to determine the relative velocity change dv/v, which is commonly referred to as coda wave interferometry (CWI). The process involves cross-correlating the two waveforms, of which the one from the second waveform is stretched in time. dv/v is found when the two signals have the highest correlation coefficient. Relative velocity change was found to be not useful in this study for several reasons. First, the permutation is influenced also by other factors than the material alteration, which makes it not possible to analyze the coda. Second, the introduced damage (a hole) is likely not to induce a velocity change. Conventionally, stress changes or the resulting microstructure changes due to stress change are investigated. Lastly, the velocity change is a metric tailored for scattering materials such as concrete, and for the investigated aluminium plates in this study, the measurements are influenced by multiple reflections on boundaries rather than scattering due to the microstructure.

Besides metrics that estimate a velocity change, other metrics exist that aim to capture the differences in the amplitude and frequency characteristics of two waveforms. These are the squared correlation coefficient R^2 or the MSCI. The R^2 score, also known as the coefficient of determination, is the proportion of the variation in the signal $u(t)$ explained by the reference signal $\tilde{u}(t)$ [23]. It is the square of the Pearson correlation coefficient p, which is calculated with the covariance cov and the standard deviation σ as

$$p = \frac{\mathrm{cov}(u, \tilde{u})}{\sigma_u \cdot \sigma_{\tilde{u}}}. \tag{1}$$

The R^2 score is given by

$$R^2 = p^2. \tag{2}$$

The R^2 takes a value between 1 and 0, which can be interpreted as the waveforms being identical (no change/difference) and completely different (or unrelated), respectively [23].

A metric that describes the similarity between two random variables in the frequency domain is the MSCI. It is according to Gonzales [24] defined as

$$\mathrm{MSCI} = \overline{\mathrm{MSC}_{u,\tilde{u}}(f)} = \frac{|P_{u,\tilde{u}}(f)|^2}{P_{u,u}(f) \cdot P_{\tilde{u},\tilde{u}}(f)} \tag{3}$$

where $P_{u,u}(f)$ and $P_{\tilde{u},\tilde{u}}(f)$ are the power spectral densities and $P_{u,\tilde{u}}(f)$ is the cross power spectral densities, which are calculated using Welch's method [25]. The MSC(f) indicates how similar u is to $\tilde{u}$ at each frequency. To get the one scalar value MSCI, MSC(f) is averaged over a frequency range from 200 to 600 kHz. This range was determined manually on the given datasets. Here the upper and lower frequency is chosen where the signal is dominating compared to noise [26]. In this study, the excitation of the ultrasound transmitter was done with a center frequency of 400 kHz (see Subsection 3.1).

Both the R^2 score and the MSCI are dimensionless and independent of the maximum amplitude of each of the signals, which means you can multiply u or $\tilde{u}(t)$ by a scalar value and still get the same value for the metrics.

The similarity metrics R^2 and MSCI are decreasing with increasing perturbation of the material. In order to have an increasing metric with an increasing damage extent 1-MSCI and 1-R^2 are used as metrics. This is also done in the literature, where, e.g. R^2 is interpreted as the decorrelation coefficient [27].

Further, other simpler metrics such as the maximal amplitude of the signals or the least square norm of two signals were also investigated in this study, but they did not lead to sufficient results for the given datasets.

2.2. Time series-based deep learning for ultrasound full waveform evaluation

In this study, a deep learning model is used, which takes an ultrasound waveform as input and predicts the x-position, y-position, and diameter d of a hole. The problem is formulated as a time series regression task, where continuous outputs are predicted based on a time series input.

Deep learning has demonstrated remarkable achievements across multiple domains, surpassing conventional machine learning models and making it the preferred choice for this application [28]. Moreover, one of its key advantages is the automated extraction of features from raw time series data, streamlining the feature engineering process. In recent times, transformer models have gained significant success in natural language processing tasks. Among these models, ChatGPT has emerged as a highly successful chatbot. Transformer models, introduced by Vaswani et al. [29], leverage an attention mechanism that enables them to focus on important sections of the data. This characteristic is particularly valuable when working with time series data.

A transformer model adapted for time series data was introduced in Zerveasi et al. [30] with the time series transformer (TST). The fastai's [31] implementation of it is used in this study. To determine the appropriate learning rate for each trained model, the fastai learning rate finder was employed. The learning rate was set to $4 \cdot 10^{-5}$ and each model was trained for 40 epochs. To train the models, the fastai fit-one-cycle policy, which incorporates a learning rate scheduler [31], was employed. Throughout all the experiments, the Adam optimiser was utilised for stochastic gradient-based optimisation.

Two datasets were used, comprising simulated and measured data, respectively. The measured dataset is explained in Subsection 3.1 and consists of 1600 signals with varying hole diameter and hole x- and y-positions. For the simulated dataset, deviations based on temperature variation and coupling influences were introduced randomly. The simulation setup is introduced in Subsection 3.2. In total, 2500 signals were simulated with randomly varying hole diameters $d \in [1.0\text{ mm}, 10\text{ mm}]$, x-positions $\in [-30\text{ mm}, 30\text{ mm}]$, and y-positions $\in [-30\text{ mm}, 30\text{ mm}]$.

For the evaluation of the model's performance, a 10-fold cross-validation was conducted. Hereby, the dataset is divided into 10 groups. Each group is used once as a validation dataset to evaluate the performance of the model, which is trained on the remaining 90% of the data. The overall performance is estimated by combining the predictions of the 10 trained models on the respective validation dataset. For more information on cross-validation the reader can refer to James et al. [32]. For the measured data, repeated measurements (see Subsection 3.1) were either in the training or validation dataset to mitigate overfitting on these more similar samples, giving a false impression of the models' performance. The performance of the models was evaluated based on the root mean square error (RMSE), which is calculated as

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(y_i - \hat{y}_i\right)^2}, \tag{4}$$

where y_i is the label of one signal, $\hat{y}_i$ is the respective prediction and n is the number of signals in the dataset. The RMSE was averaged for the three labels hole diameters d and hole x- and y-position.

For training the models, a workstation with the following hardware configuration was used: an NVIDIA GeForce RTX 3090 Ti GPU with 24 GB of GPU-RAM, an AMD Ryzen Threadripper PRO 3955WX with 16 cores, and 64 GB of RAM. This workstation was also used for the ultrasound simulations.

2.3. Probability of detection curves

A POD evaluation is commonly used in NDT to assess the capabilities of a measurement system and its evaluation procedure. In industry, the handbook MIL-HDBK-1823A [33] is widely established. The POD method has been standardised by ASTM [34,35], which conforms with MIL-HDBK-1823A. The POD evaluations of this study are conducted according to this guideline. A POD analysis estimates the probability with which a hole with a defined characteristic (e.g. hole diameter) can be found using the given measurement and evaluation procedure. There are two variations of the POD method: the hit/miss and the $\hat{a}$ vs. a approach. The hit/miss POD is applied when the output of the measurement system is binary, so either hit or miss. This is often the case when the evaluation is done by a human inspector [36]. Since the $\hat{a}$ vs. a POD is used in this study, its working principle will be explained further.

The $\hat{a}$ vs. a POD was originally introduced by Berens [37]. The working principle is sketched in Figure 1. It evaluates the signal response $\hat{a}$ of a defect with a given characteristic a. An $\hat{a}$ vs. a plot is shown at the top of Figure 1. Hereby, three thresholds are set. $\hat{a}_{\text{noise}}$ defines the noise of the measurement system under which no defect could be separated from the noise. The decision threshold $\hat{a}_{\text{dec}}$ defines when a response is seen as a hit or a miss, and is always greater than $\hat{a}_{\text{noise}}$ [33]. The saturation of the measurement system is given by $\hat{a}_{\text{sat}}$. In the case of the comparison metrics, this value is set to 1, since they converge to 1 for large differences between the signals.

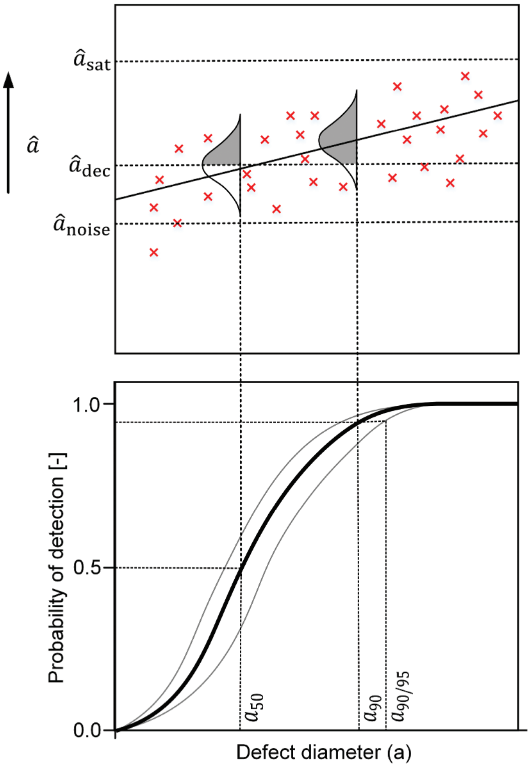


Figure 1. Schematic illustration of $\hat{a}$ vs. a POD analysis.

The POD is a cumulative normal distribution function Φ and is defined as

$$\mathrm{POD}(a) = \mathrm{Probability}(\hat{a} > \hat{a}_{\mathrm{dec}}). \tag{5}$$

The relation between $\hat{a}$ and $\hat{a}$ is assumed to be linear which can be described as

$$\hat{a} = \beta_0 + \beta_1 \cdot a + \delta \tag{6}$$

with β_0 and β_1 being the parameters of a linear regression and δ is the random error with the standard deviation σ_δ. With this linear relation, the POD can be estimated as

$$\mathrm{POD}(a) = 1 - \Phi[\frac{\hat{a}_{\mathrm{dec}} - (\beta_0 + \beta_1 \cdot a)}{\sigma_\delta}]. \tag{7}$$

If the relationship between $\hat{a}$ and a is nonlinear, it can be linearised by applying a logarithm to $\hat{a}$ or a [33]. The parameters of Equation 7 are fitted to the data using maximum likelihood estimation, which also estimates confidence bounds using the confidence bounds [38]. Here, most often, and also in this study, the 95% confidence bounds are chosen. For an $\hat{a}$ vs. a POD, at least 30 measurements are necessary [36]. For more on the POD calculation process, please refer to [33].

From the POD curves, several metrics can be extracted. Commonly used metrics include a_{50}, a_{90}, and the $a_{90/95}$ value, which is defined as the intersection between the lower 95% confidence bound and the 0.9 probability of detection [38]. The performance of the evaluation approaches is compared based on the a_{50} value in this study, which gives the hole diameter for which a POD of 50% is reached.

3. Methodology

In this Section, the investigated test specimens, the experimental design, and the ultrasound measurement setup are described. Further, the simulation considerations for modelling the random influence factors are explained.

3.1. Laboratory experiment

Experimental measurements were conducted on aluminium plates with dimensions 140 mm × 100 mm × 3 mm. The ultrasound p- and s-wave velocities were determined experimentally to be $v_p = 6382\ \frac{\mathrm{m}}{\mathrm{s}}$ and $v_s = 3191\ \frac{\mathrm{m}}{\mathrm{s}}$ using an OmniScan M×2ultrasonic testing instrument. The measurements were carried out using a 2.25 MHz p-wave transducer (Panametrics V104-RM) and a 5 MHz s-wave transducer (Panametrics V157-RM) in pulse-echo configuration. The density ρ was calculated as $2638\ \frac{\mathrm{kg}}{\mathrm{m}^3}$ by weighing the specimens and using their nominal dimensions. The plates were made of the aluminium alloy AlMg3, and 64 plates were laser-cut to ensure the dimensions are accurate. Each plate featured a single drilled hole. Five different hole diameters, $d = 1.5, 3, 4.5, 6,$ and 7.5 mm were manufactured. For each hole diameter, 16 distinct hole positions, characterised by varying x- and y-coordinates, were drilled. Leveraging the symmetry of the plate by flipping it and interchanging the positions of the transmitter and receiver, 64 hole positions per diameter were tested. This is visualised in Figure 2(a). Notably, some hole positions are common to multiple plates and transmitter-receiver configurations.

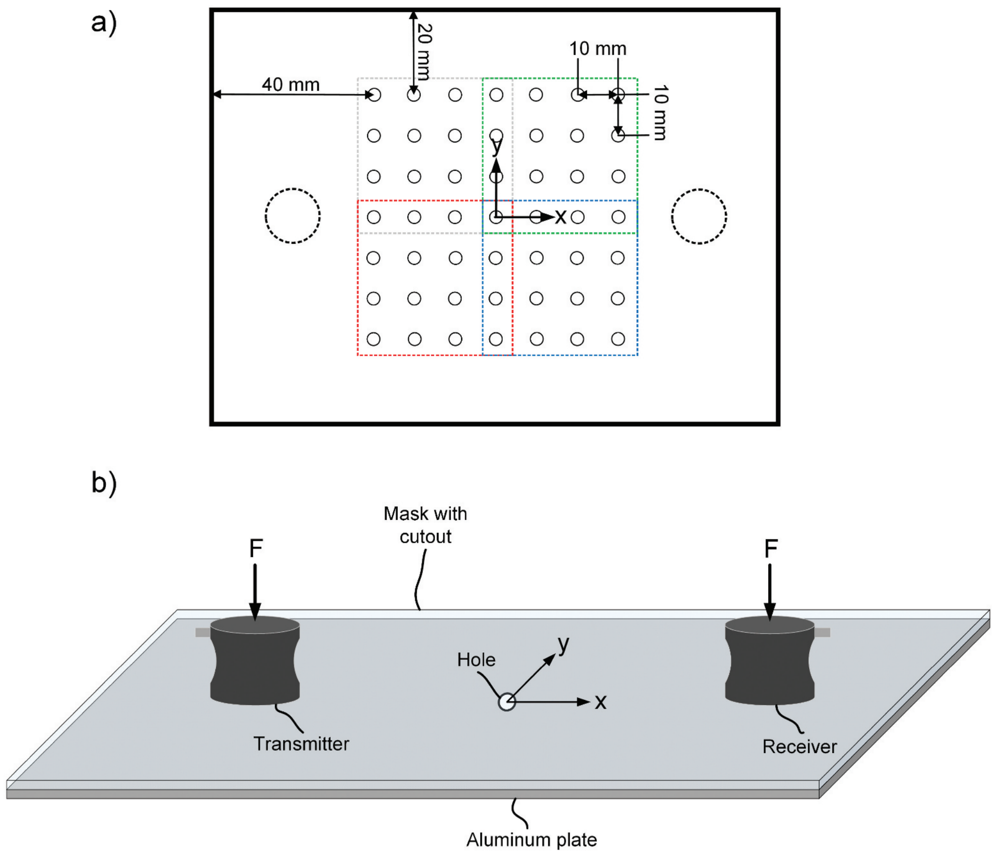


Figure 2. Illustration of a test specimen with all tested hole positions in (a). The colored squares visualize four different plates (flipped) and transducer-receiver (switched) configurations. In (b), the measurement setup with the mask is shown along with the coordinate system.

Each measurement was repeated five times and in between these measurements, the transmitter and receiver were removed from the plate and then remounted. In total, 1600 measurements were conducted. Further, measurements on five plates without a hole, hereafter referred to as reference plates, were carried out.

To achieve accurate positioning of the transmitter and receiver on each plate, a plastic mask with cut-outs for both components was employed, as illustrated in Figure 2(b). Consistent coupling was ensured by applying weights to both the transmitter and receiver and a shear wave couplant was used as a coupling agent. The selected transmitter and receiver were the Panametrics V103 p-wave probes with a resonance frequency of 1 MHz.

Illustration of a test specimen with all tested hole positions in (a). The coloured squares visualise four different plates (flipped) and transducer-receiver (switched) configurations. In (b), the measurement setup with the mask is shown along with the coordinate system.

To excite the transmitting transducer, a waveform generator (TiePie Handyscope HS5), along with a Trek 2100 HF amplifier, was employed. The transmitter was excited at 400 kHz and 150 V with a square pulse. For 400 kHz, the wavelength of the p-wave in the material is 16 mm, while for the s-wave, it is 8 mm. This relatively low frequency was chosen to reduce the attenuation of the ultrasound wave and, therefore, capture longer signals. The signals were recorded with a sampling frequency of 20 MHz and a signal duration of 8 ms, representing 16,000 samples. For each measurement, 100 signals were recorded and averaged, to increase the signal-to-noise ratio (SNR).

3.2. Numerical simulations

The conducted measurements differ not only due to different hole diameters and positions but also due to other random effects. In order to model these effects, a Monte Carlo approach is used in this study. All deviations were assumed to be normal distributed. The applied standard deviations σ for the individual influence factors are given in Table 1.

Three categories of random influence factors are considered and are associated with the test specimen, the transducers, and the hole.

The influence factors related to the test specimen include the error in the plate x- and y-dimensions, which were introduced by the laser cutting process. To estimate the deviation in the dimensions, all plates were measured with a caliper. The standard deviation for both the x- and y-dimensions of the plates was determined to be 0.05 mm, while the plate height exhibited a standard deviation of 0.01 mm. The deviation in the height most likely stems from the extrusion moulding of the raw aluminium material. Furthermore, the standard deviation of temperature in the plate during the experimental measurements was assumed to be $\pm 2\ ^{\circ}$C. A change in temperature affects the ultrasound velocities with which the temperature was changed indirectly. To account for this in the simulation, the estimates of Ginzel and Ginzel [39] were used for aluminium, which state that for a deviation of $\pm 5\ ^{\circ}$C from room temperature, a linear relationship can be approximated. The estimated relationships for the temperature, T and the p- and s-wave velocities, v_p and v_s are, $\tilde{v}_p = 1.47 \cdot \Delta T + v_p$ and $\tilde{v}_s = 1.47 \cdot \Delta T + v_s$, respectively. Since density variations only influence the wave propagation at partially reflective interfaces, which are not present in our simulation scenarios, deviations in the densities were not considered.

Table 1. Standard deviations of the individual parameters selected to model the random influences on the simulated ultrasound signals. All influence factors except the transmitter angle were assumed to be normally distributed. The transmitter angle was selected from a uniform distribution between 0 and 25%.

Parameter	σ
Plate x- and y-dimensions	0.05 mm
Height	0.01 mm
Temperature	2°C
Amplitude/Force	0.05
Trans. Rec. x- and y-position	0.07 mm
Transmitter angle of incidence	25%
Hole diameter	0.005 mm
Hole x- and y-position	0.05 mm

The contact area of the transmitter was modelled as a 'Fibonacci sphere', which is displayed in Figure 3(c). This pattern is intended to mimic a disc-shaped transmitter with point sources, which On the Feasibility of Ultrasonic FullWaveform Evaluation are distributed as evenly as possible over a circular area. The excitation was modelled with a timevarying vector force, and a force normal to the surface was applied. The excitation waveform was extracted from a measurement where the transmitter and receiver were pressed directly onto each other and can be seen in Figure 3(a) with its frequency spectrum in (b). The receiver was modelled with the same pattern as (c), and the measured component normal to the surface was averaged over all points of the receiver.

Influence factors stemming from the transmitter and receiver were regarded. Initially, the energy input into the specimen is considered, which is assumed to be linearly correlated to the force with which the transmitter is pressed against the specimen. For that reason, the force in the simulation was scaled such that the maximal amplitude of the simulated signal matched the measured one.

To get an estimate of the variation in the energy, the standard deviation in the maximal amplitude in the measured signals of the reference plates was determined. The standard deviation was further scaled with the same factor as the force in the simulation relative to the measured signals, which resulted in a value of 0.05. With this standard deviation, the force of the vector source in the simulation was scaled by

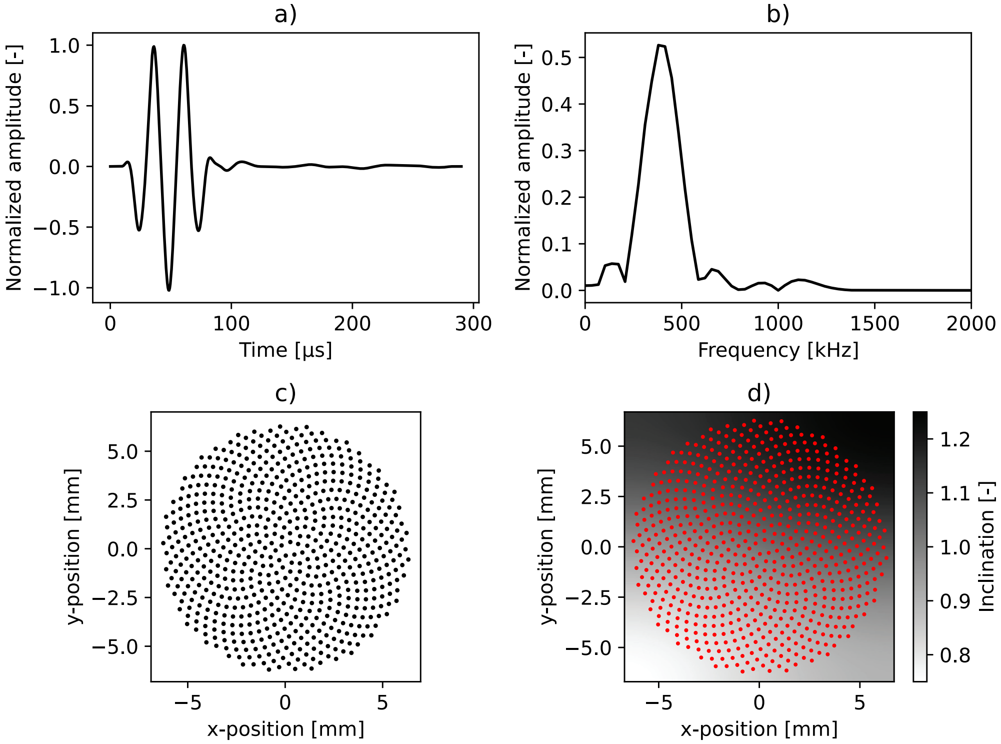


Figure 3. Modeling of the ultrasound transmitter. In (a), the excitation wavelet is displayed with its frequency spectrum in (b). In (c), the ensemble of point sources is visualized, and in (d) the Gaussian random field for modeling the transmitter.

sampling from a Gaussian distribution. The amplitude of the simulated signal is linearly correlated to the applied force in the simulation.

A deviation of the transmitter's angle of incidence from the normal was modelled. It was assumed that the angle of incidence has a linear effect on the travel time in the coupling agent and the force on the surface. For this, a Gaussian random field was estimated, a sample of which is given in Figure 3(d). Regarding the angle of incidence, On the Feasibility of Ultrasonic FullWaveform Evaluation

This pattern is intended to mimic a disc-shaped transmitter with point sources, which are distributed as evenly as possible over a circular area. The excitation was modelled with a time-varying vector force, and a force normal to the surface was applied. The excitation waveform was extracted from a measurement where the transmitter and receiver were pressed directly onto each other and can be seen in Figure 3(a) with its frequency spectrum in (b). The receiver was modelled with the same pattern as (c), and the measured component normal to the surface was averaged over all points of the receiver.

Influence factors stemming from the transmitter and receiver were regarded. Initially, the energy input into the specimen is considered, which is assumed to be linearly correlated to the force with which the transmitter is pressed against the specimen. For that reason, the force in the simulation was scaled such that the maximal amplitude of the simulated signal matched the measured one.

To get an estimate of the variation in the energy, the standard deviation in the maximal amplitude in the measured signals of the reference plates was determined. The standard deviation was further scaled with the same factor as the force in the simulation relative to the measured signals, which resulted in a value of 0.05. With this standard deviation, the force of the vector source in the simulation was scaled by sampling from a Gaussian distribution. The amplitude of the simulated signal is linearly correlated to the applied force in the simulation.

A deviation of the transmitter's angle of incidence from the normal was modelled. It was assumed that the angle of incidence has a linear effect on the travel time in the coupling agent and the force on the surface. For this, a Gaussian random field was estimated, a sample of which is given in Figure 3(d). Regarding the angle of incidence, it was assumed that the transmitter is at its outer radius maximal 25% closer or farther away from the surface of the sample than the centre of the transducer. In contrast to the other influence factor here, a uniform distribution between 0 and 25% was assumed. With the given transmitter radius of 6.35 mm, this corresponds to an angle of incidence of 0.023°, considering a couplant layer thickness of 0.01 mm. This angle, with which the random field was scaled, was determined randomly. In Figure 3, for example, an inclination of 25% was applied. With the respective value for the scaled random field at each source position, the force was multiplied. Further, the starting time of each wavelet was expedited or delayed by the same factor, assuming that the transmitter is on one side closer to the surface of the specimen.

The transmitter and receiver's x- and y-positions were placed randomly and independently. Since they were positioned with a mask oriented along the edges of the plate, and the plate dimensions had a deviation of 0.05 mm, the location error should have at least this magnitude. The mask cut-out also has some dimensional inaccuracies. Therefore, the total standard deviation was set to be 0.07 mm.

Finally, the error in the manufacturing of the hole was considered. Since the drilling process is highly reproducible, the deviation in the hole diameter d was assumed to be low and was set to 0.005 mm. For the x- and y-positions of the hole, the same error as for the plate dimensions was taken, as the positioning of the hole was done based on the edges of the plate. All used standard deviations of the individual parameters are listed in Table 1.

Some effects are not accounted for in this modelling approach, such as the inclination or radial changing sensitivities of the receiver, the frequency response of the receiver, as well as the propagations of the waves into the transmitter and receiver or over other interfaces.

With the assumed deviations, simulations were conducted, and the signals of two randomly selected parameterised simulations on a plate without a hole are presented in Figure 4 (bottom Figure). In the top Figure, two sample measured signals from different reference plates without a hole are shown in red and black. Here signals with visually large differences were chosen. In the simulation, the attenuation was modelled using Q-values, which were manually adapted, such that the decrease in the amplitude over time of the simulated signals matches the measured signals. The positions of the different wave packets vary between the measured and simulated signals. This discrepancy may arise from not modelling the interfaces and assuming perfectly reflecting boundaries, which is not the case in reality. In the experiment, the plate was lying on a table, where continuously some of the energy is coupled out over the plate/wood table interface. For signals exhibiting numerous reflections and mode conversions, minor inaccuracies in the

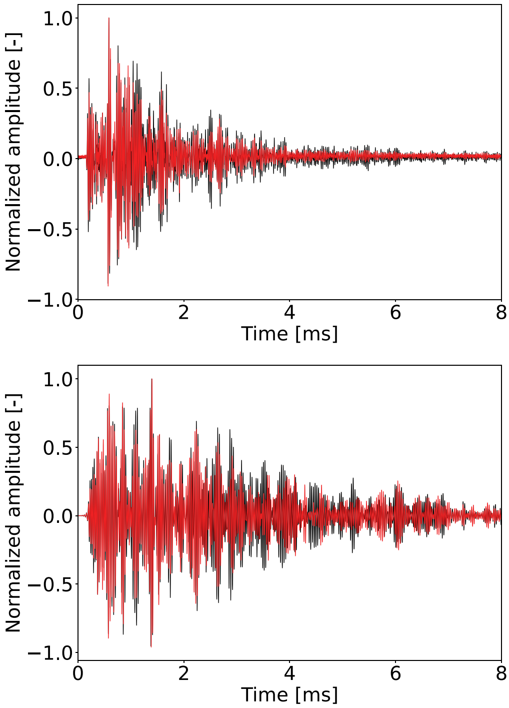


Figure 4. Comparison of sample measured (top) and simulated (bottom) signals for plates without holes. The respective red and black signal represent two samples from different measured or simulated plates.

model can result in significant differences in the waveforms. Nevertheless, both the simulated and measured signals displayed significant visual differences. Given the study's objective to explore the extent to which such data can be utilised for the detection of the holes, the simulated data remains suitable despite the modelling inaccuracies.

All simulated signals were downsampled to a sampling rate of 20 MHz (same as measured signals) using Lanczos resampling. For all simulations, boundary-conforming meshes were used, which were generated within the software Salvus from Mondaic. Salvus is based on the spectral element method, which enables an efficient solution, allowing simulations of high-frequency wave propagations, of the wave equation with little restrictions on the geometry. The spectral element method is a variant of the finite element method with a diagonal mass matrix [40]. For the smallest simulated hole diameter (1 mm), 16 hexahedral elements circumscribe the hole. Together with the use of a tensor order of four and 4th-order Lagrange polynomials. The error in the shape representation is negligible.

4. Results and discussion

In this section, the results are presented and discussed. First, the optimal signal length is determined, and the influence of the hole position on the comparison metrics is estimated using the simulated data. Next, the performance of the two comparison metrics for defect detection is evaluated. Subsequently, the results of the deep learning model are presented. Finally, the outcomes of the comparison metrics and the deep learning model are compared, based on both simulated and measured data.

4.1. Influence of signal length and hole position

At first, the effect of the position of the hole and the length of the investigated signal on the comparison metric is investigated. Hereby, simulated data was used without introducing random effects to show the idealistic relations. To investigate the influence of the signal length on the MSCI and R^2 scores, signal durations of 0.25 to 8 ms, which represent 500 to 16,000 samples, respectively, were used to calculate the comparison metrics. It can be expected that later portions of the signal are more sensitive to other influences, like changes in the temperature [6]. Nevertheless, they are also more sensitive to holes being present in the material. Hereby the hole was positioned at the center of the plate, and hole diameters ranging from 1.0 mm to 10.5 mm were simulated. The results are given in Figure 5.

It can be observed that for the longer signal lengths, both the MSCI and R^2 scores show the largest slope over increasing hole diameters. For that reason, the entire signals, which have a length of 8 ms, were used for all evaluations with the comparison metrics. For small hole diameters, the R^2 score shows at first a decreasing trend. This can be due to a change in the dominant interaction from diffraction to reflection-based for holes smaller than half of the wavelength. Further, the R^2 score also shows for larger defect diameters short decreasing trends, which is not the case for the MSCI.

Further, the spatial sensitivity of the comparison metric on the hole positions was evaluated. For this variations of the hole position were introduced within the simulation,

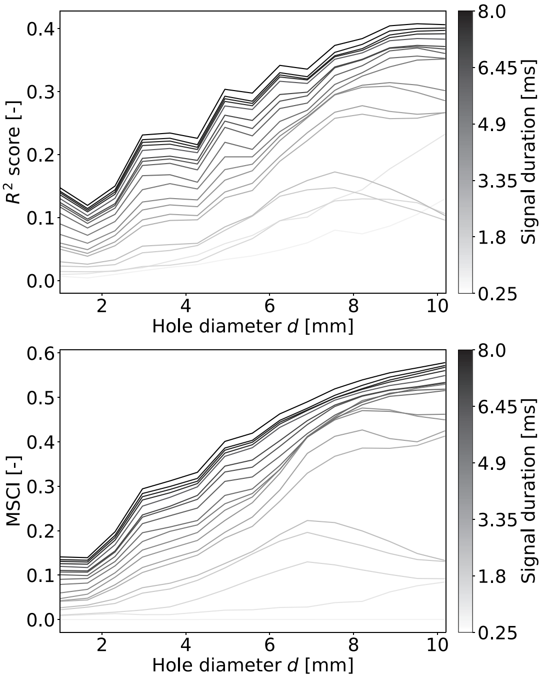


Figure 5. Evaluation of how the R^2 score (top) and MSCI (bottom) change with varying signal durations. The colorbar represents the duration or length of the signals used for the calculation of the comparison metrics.

ranging from −30 to 30 mm for both the x- and y-positions. This represents the same range for the hole positions as for the laboratory experiments (see Figure 2). A hole diameter $d = 5$ mm was selected. The simulated signals were then compared to a reference signal of a plate without a hole using the comparison metrics MSCI and R^2 score. The result is shown in Figure 6.

It can be seen that both comparison metrics show a similar relation, and the metrics are symmetric towards the centre axis in the x- and y-direction of the plate. High metric values are reached at the symmetry axis of the plate. This can be due to focusing effects, which lead to more energy of the waves being present at these positions. Another explanation for these phenomena is that the reflected waves pass through the centre axis of the plate most often, and therefore, a hole affects the wave propagation there the most in that specific region. The symmetry in the y-direction is expected since the simulation and measurement setup are symmetric. The symmetry in the x-direction, however, is surprising. An explanation could be that the position of the transducer and receiver is not that important when long signals are evaluated. In general, the metrics deviate strongly over the hole positions. The maximal values of the metrics are reached at the centre of the plate.

4.2. Evaluation of the comparison metrics

At first, the variation in the R^2 score and MSCI for different hole diameters is explored, using both simulated and measured data. Centrally positioned holes (x = 0 mm and y = 0

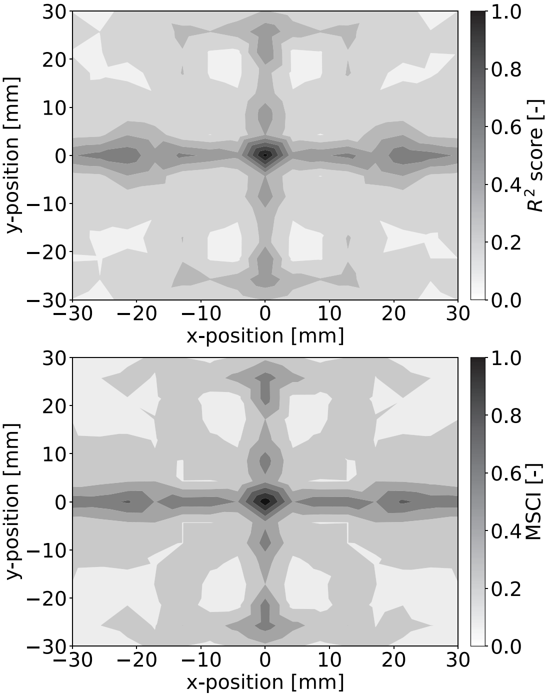


Figure 6. Effect of hole position on the comparison metrics for the R^2 score (top) and MSCI (bottom).

mm) are considered only. The simulated dataset involves simulations with 15 distinct hole diameters ranging from 1.0 mm to 10.5 mm. For each hole diameter, 30 simulations with randomly initialised influence factors are conducted. The reference for R^2 score and MSCI was estimated through a simulation without a hole and with no random influences. Additionally, a POD evaluation is conducted. For this, the decision threshold $\hat{a}_{\mathrm{dec}}$ was determined by estimating the mean comparison metrics of 50 simulations without a hole but including random influence factors and adding two time the standard deviation to this mean value. This represents 95% of the data assuming a Gaussian distribution. In a The calculated values of the comparison metrics for each simulation are depicted in Figure 7(a) for the R^2 score and in (b) the MSCI in red. The mean values, together with an error band denoting the standard deviation, are also shown. In Figure 7(c), the signal-to-noise ratio, calculated as the mean divided by the standard deviation, is given for each hole diameter. The MSCI is shown in blue and the R^2 score is in black. The same colours were used in the result of the POD evaluation in Figure 7(d). Further, the a_{50} values are given for the R^2 score and the MSCI, which are 5.5 and 2.9 mm, respectively.

The distribution of the comparison metric values for a given diameter is skewed, with a few instances exhibiting relatively high values. For the simulated dataset, the MSCI performs overall better than the R^2 score. The R^2 score shows outliers for individual samples, exhibiting a much larger spread with a higher standard deviation, as also evident in the signal-to-noise ratio. In the SNR, there is a noticeable jump at approximately 8 mm, corresponding to half of the wavelength for the p-wave ($\lambda/2 = 8$ mm). This change may be attributed to a shift in the dominant wave interaction from reflection to scattering

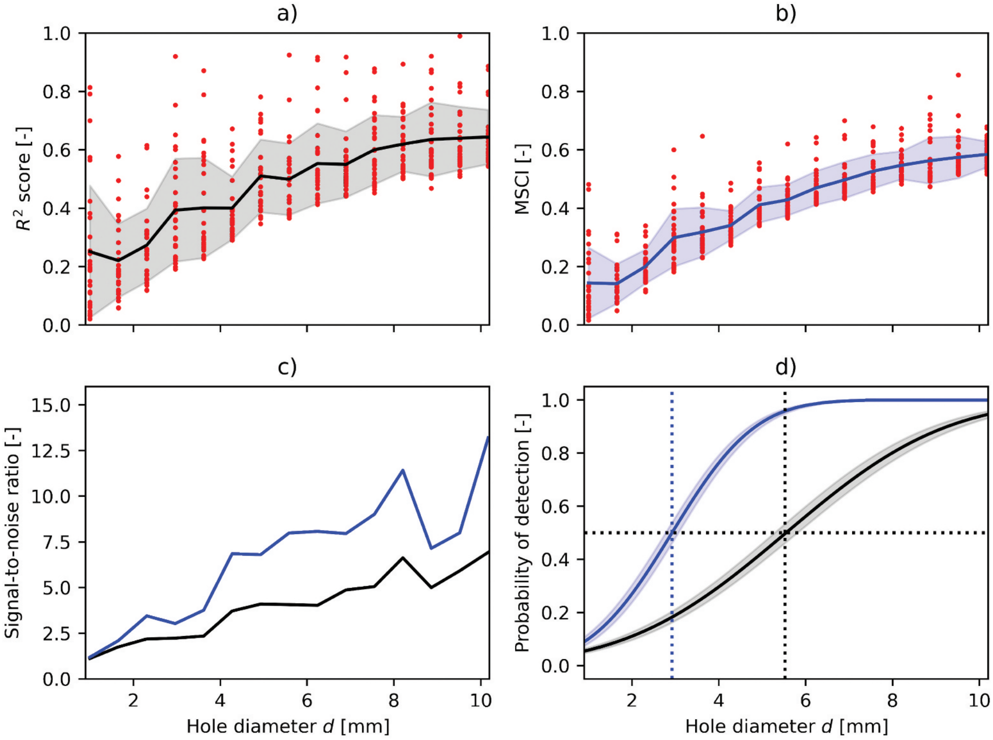


Figure 7. Evaluation of the response of the comparison metrics for different hole diameters with simulated data. The holes are located in the center of the plate. The R^2 score is plotted in black, and the MSCI in blue. The shaded error bars display, for the R^2 and MSCI, the standard deviation, and for the POD curve, the 95% confidence bound. In (a) and (b) the response of the R^2 score and MSCI is plotted. In (c) the signal-to-noise ratio is given and in (d) the POD evaluation is shown. The horizontal and vertical lines in (d) give the POD 50% and the a_{50} values. The decision thresholds $\hat{a}_{\mathrm{dec}}$ are 0.70 for the R^2 score and 0.39 for the MSCI.

[41]. In the POD evaluation, the probability of detection is generally higher for the MSCI than for the R^2 score.

The same evaluation was conducted on the experimentally measured dataset using the holes in the centre, which included 100 measurements with 5 repetitions for each plate configuration.

The evaluations are given in Figure 8. For the reference signal, the signals of half of the measured plates without a hole were averaged. For the POD evaluation (see Figure 8(d)), the decision threshold $\hat{a}_{\mathrm{dec}}$ was established using the mean of the comparison metrics from measurements of a plate without a hole, which was not utilised in computing the reference signal. In this process, half of the measurements were averaged to estimate a reference, while the other half was employed to determine $\hat{a}_{\mathrm{dec}}$.

For the measured data, the MSCI consistently outperforms the R^2 score, and the a_{50} value, where the POD reaches 50%, is 1.9 mm for the MSCI, which is significantly smaller than its simulated counterpart of 2.9 mm. The R^2 score does not exhibit a monotonically increasing trend over the hole diameter and is failing to reach a POD of 50% within the

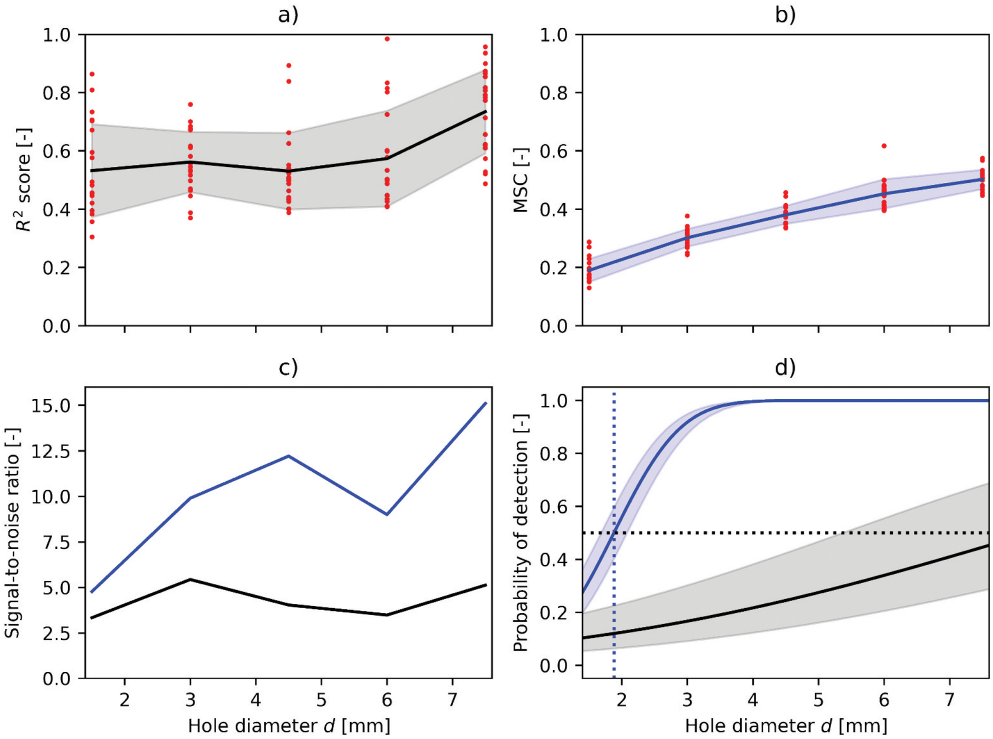


Figure 8. Evaluation of the response of the comparison metrics for different hole diameters on the measured dataset. The holes are located in the center of the plate. The R^2 score is plotted in black, and the MSCI in blue. The shaded error bars display, for the R^2 and MSCI, the standard deviation, and for the POD curve, the 95% confidence bound. In (a) and (b) the response of the R^2 score and MSCI is plotted. In (c) the signal-to-noise ratio is given and in (d) the POD evaluation is shown. The horizontal and vertical lines in (d) give the POD 50% and the a_{50} values. The decision thresholds $\hat{a}_{\text{dec}}$ are 0.85 for the R^2 score and 0.24 for the MSCI.

investigated hole diameter range. One possible explanation for the worse performance of the R^2 score on measured data is its inability to handle noise in the reference signals, a factor absent in the simulation. Further, since the MSCI evaluates the changes in the frequency domain, it is independent of time shifts in the signals (e.g. different arrival times), which makes it more robust. Given the superior performance of the MSCI, it is utilised in all subsequent evaluations.

4.3. *Effect of the environmental and coupling influences*

As described in Section 3.2, eight different random influence factors were modelled. Three categories of influence factors were regarded, which are related to the plate, the transmitter and receiver, and the hole within the plate. To investigate the effect of each influence factor on the MSCI, they were sampled from a Gaussian distribution with a variation of $\pm 3 \cdot \sigma$, where σ represents the given range in Table 1. An exception here is the transducers' angle of incidence, which was varied between 0 and 100%. A hole with a diameter of 5 mm was introduced in the middle of the plate. The resulting MSCI values are given in Figure 9. For the variation of the plate x-dimensions, the left and right sides of the plate were varied positive and negative x-direction. The results of all four combinations (left pos./right pos., left neg./right neg., left neg./right pos., left pos./right neg) were averaged. The same was done for the plate y-dimensions and the transmitter and receiver positions.

It can be seen that the influence factors on the whole plate, plate dimensions, plate height, and temperature have the largest effect on the MSCI (see Figure 9(c)). It shall be noted that the y-axis is scaled differently for these factors since larger MSCI values are reached. Especially, changing the height has a high impact. The MSCI value of the 5 mm hole without the influence factors is 0.38, which is the value at $\sigma = 0$. For some parameter variations, e.g. temperature, plate dimensions, or hole positions, MSCI values less than 0.38 are reached for certain values of that influence factor. This can lead to an underestimation of the defect size. The MSCI appears to be constant over the range of contact forces considered (see Figure 9(a)). The reason for this is that the changes in the force change the amplitude in the signal linearly and in the calculation of the MSCI, the amplitudes are normalised. For real measurements, the force can have an effect on the MSCI for very large and very low forces due to signal clipping

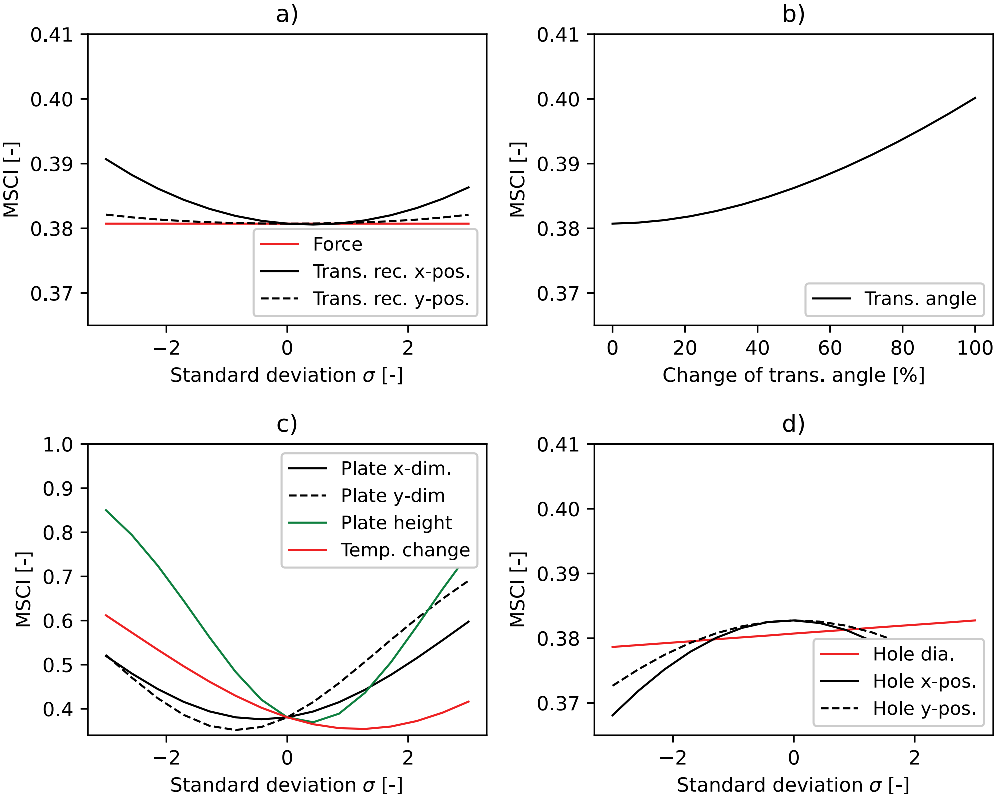


Figure 9. Effect of individual influence factors on the MSCI. In (a) and (b) the influence factors related to the transmitter and receiver are given, while (c) addresses those related to the whole plate. In (d) the factors affecting the hole are shown.

and the presence of noise. Further, changing the transmitter and receiver x-position has a larger effect than the y-position; the same is true for the hole position. A reason for this could be that by changing the x-position, the direct path of the wave from the ultrasound transmitter to the receiver is affected, whereas changing the y-position has more influence on the reflections at the side of the plate. The angle of incidence has a relatively small effect for the chosen deviation (25%) but increases non-linearly for larger changes (see Figure 9(b)). Most of the influence factors change the MSCI non-linearly over their parameter range, which makes it more difficult to calculate the introduced error due to the respective influence. A parameter with a linear relationship is the change in the hole diameter (see Figure 9(d)).

4.4. Estimation of the hole position and diameter with deep learning and comparison metrics

Based on the previous findings, it can be concluded that variations of the defect diameter *d* (see Figures 7 and 8) and the hole x- and y-positions (see Figure 6) significantly affect the MSCI. These two factors are likely difficult to distinguish based on one comparison metric, thereby compromising the ability to detect defects. This is evident in Figure 10, where the MSCI was assessed for both the simulated dataset (top) comprising 2500 randomly

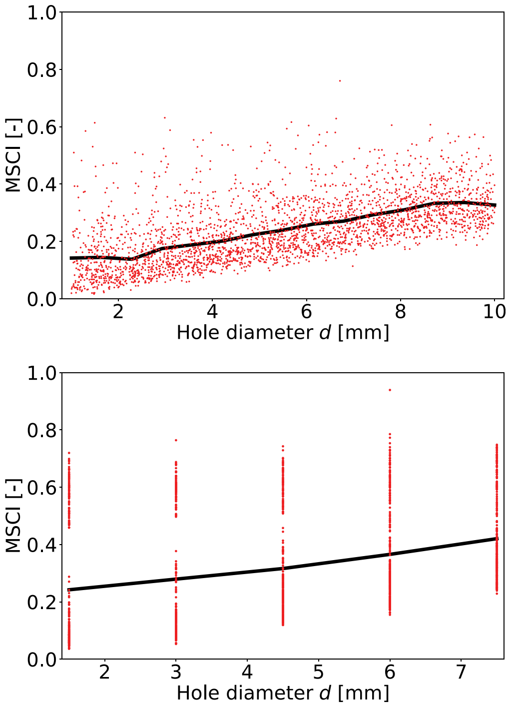


Figure 10. Evaluation of MSCI over the hole diameter *d* with random influence factors and varied hole positions. The simulated dataset is presented at the top, while the measured dataset is shown at the bottom. The black line indicates the MSCI for each respective hole diameter.

initialised simulations and the measured dataset (bottom) consisting of 1600 measurements. Both datasets contain various hole positions with x and y-positions ∈ [−30 mm, 30 mm], along with different defect diameters.

Since there was no visible relation when plotting the MSCI over the x-position or the y-position for both datasets, these plots are not shown here. It can be seen that the relation is now significantly weaker in comparison to the evaluations of the hole position in the centre of the plate (see Figures 7 and 8), and the noise level in the data is higher. Nevertheless, the mean of the MSCI is still increasing over the defect diameter. For the simulated dataset, the mean MSCI was calculated for evenly spaced hole diameter ranges.

Moreover, a TST model was trained on both simulated and measured data. Since the model has three outputs, the hole diameter *d*, and the x- and y-position, it has the capability of distinguishing these three influence factors in the signals. The models were trained with a 10-fold cross-validation on the respective dataset. Only the validated dataset of each fold is used for further evaluation. Notably, it was observed that when labeling x and y-positions within ∈ [−30 mm, 30 mm] with a positive and negative range, the model failed to learn anything and simply predicted the mean of these labels, which is zero. This can be due to symmetries in the measurement setup, which lead to similar signals at different hole positions. To address this, absolute values for the x- and y-positions were utilised as labels, resulting in a range from 0 to 30 mm. The test results are presented in Figure 11.

In both the simulated and measured datasets, the TST model demonstrates more accurate predictions for the hole diameter, *d* than for the hole positions. Hole positions near the centre axis of the plates (x = 0 mm or y = 0 mm) or at the boundary (x = 30 mm or y = 30 mm) exhibit higher prediction errors for both

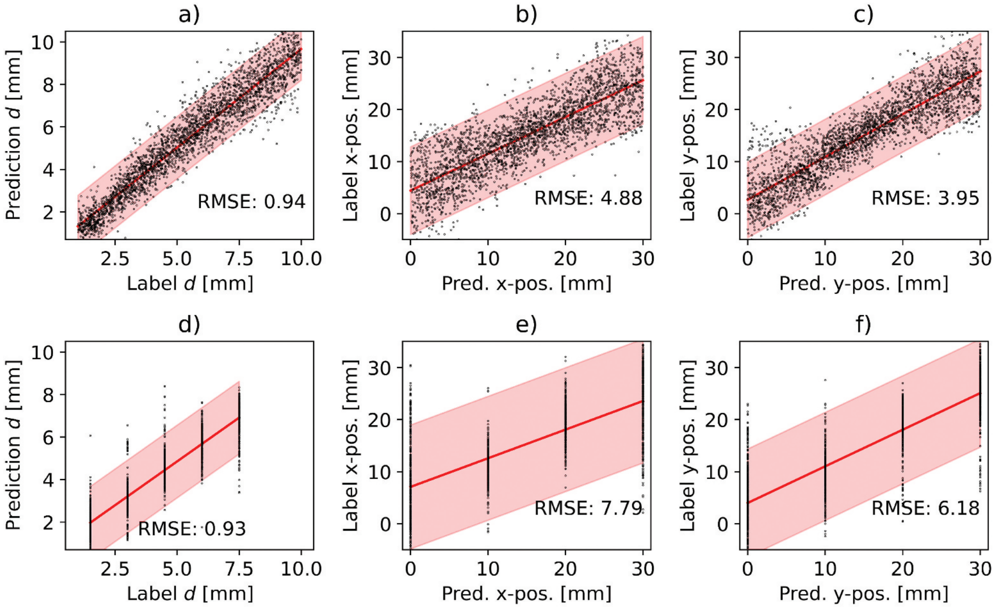


Figure 11. Predictions generated by the TST model for the simulated dataset in (a)-(c), and the measured dataset in (d)-(f). A line fit with its 95% prediction bounds is given in red.

datasets. Moreover, larger diameters are predicted less accurately. Overall, the models exhibit poorer performance on the measured dataset in (d)-(f) compared to the simulated one in (a)-(c). An exception is observed in the case of hole diameter, where the RMSE is smaller. It shall be noted that the simulated dataset contains larger hole diameters than the measured dataset, which leads to a higher RMSE.

4.5. Evaluation of the probability of detection for the comparison metrics deep learning model

To quantitatively compare the MSCI and the deep learning models, a POD evaluation was performed. The estimation of $\hat{a}_{\mathrm{dec}}$ on a plate without a hole is not directly applicable to the deep learning model. Therefore, the mean response plus one standard deviation of the model to hole diameters $\leq 1,5$ mm was used as $\hat{a}_{\mathrm{dec}}$ for both the simulated and measured datasets. This represents 68% of the data for that hole diameter range. To evaluate the TST model and the MSCI comparable this was also done for the MSCI. The POD curves are illustrated in Figure 12, with the simulated data displayed at the top and

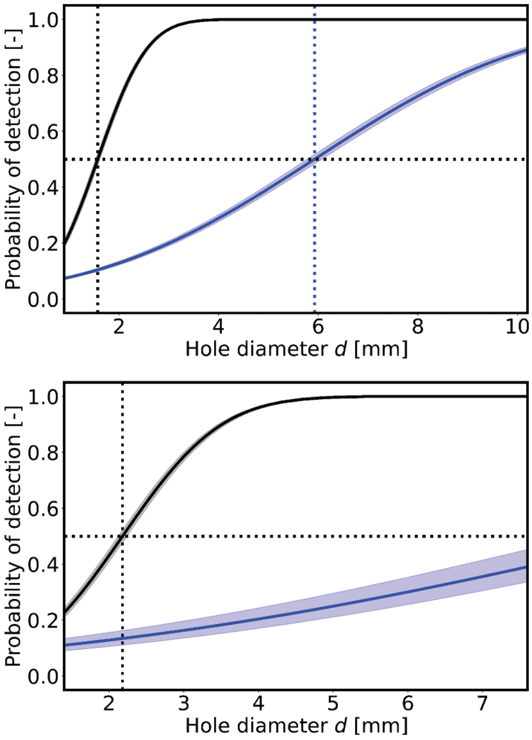


Figure 12. POD curves for the simulated (top) and measured (bottom) datasets. The TST response is depicted in black, and the MSCI response in blue is. The datasets include simulations and measurements with varying hole diameters and positions. The decision thresholds $\hat{a}_{\mathrm{dec}}$ for the simulated data are 1.85 for the TST and 0.25 for the MSCI. For the measured data they are 2.53 and 0.47, respectively.

the measured data at the bottom. The MSCI is represented in blue, and the precision of the deep learning model is depicted in black.

The POD for the simulated dataset is higher than that of the measured dataset, and the deep learning model outperforms the MSCI. In the simulated dataset, the MSCI achieves a 50% POD at 5.93 mm, while the TST model achieves this at 1.57 mm. Notably, the a_{50} value of the TST model is even smaller than that of the MSCI with the defect positioned at the center and the simulated data (see Subsection 4.4). For the measured data, the TST model reaches a 50% POD at 2.18 mm, and the a_{50} value for the MSCI is beyond the investigated range of hole diameters. The relatively narrow confidence bounds, especially in the context of the high variance in the data, can be attributed to the large dataset sizes of 2500 and 1600 samples, respectively.

5. Conclusions and outlook

This study demonstrates the application and limitations of full waveform evaluation for quality control of manufactured parts using a simple ultrasonic transmitter-receiver setup. Plates with holes of varying sizes at different positions were examined, and the measured setup was simulated. Hereby, random influence factors, which represent changes in coupling and environmental conditions, were regarded. Hereby, it was found that influence factors related to the properties of the whole plate, such as temperature or the plate dimensions, have the largest effect on the measured and simulated ultrasonic signals.

The comparison metrics, R^2 score and magnitude squared coherence MSCI, were evaluated for the capability to detect the holes in the plate. For the MSCI hole diameter and position were investigated separately and combined. Additionally, a deep learning model, i.e. a time series transformer (TST), was trained to predict defect positions and diameters. All approaches were compared using probability of detection curves. The key findings of the study can be summarised as follows:

- The MSCI was found to be robust to many random influence factors, related to the coupling or transducer position or the deviation of the position of the hole given at a specific location.
- Both the hole diameter and the hole position have a large effect on the ultrasound signals and the detectability of the defects but are with the given approach impossible to distinguish with the MSCI.
- The best approach for detecting defects was found to be deep learning models, which are robust to different influences and can separate several effects that change the ultrasonic signal.

For the practical implementation of this kind of methodology, initial precision in manufacturing is crucial, requiring minimal variations in part dimensions. Consistent measurement conditions, including temperature, are equally relevant. The accurate positioning of the transmitter and receiver is also of importance, and in such cases, industrial robots could help. The force with which the transmitter is pressed onto the specimen does not influence the evaluation with MSCI under the range studied.

Additionally, when examining the same part, it is important to investigate the signals for various defect positions. Simulation can streamline and optimise this experimentation process. In large-scale manufacturing facilities, establishing a database of ultrasonic measures becomes feasible. With such a database, which contains defects with known characteristics, a deep learning model could be trained, enabling relatively robust defect detection.

In the future, we aim to explore whether utilising both the MSCI and ultrasonic recordings as inputs enhances defect detection. Furthermore, we plan to explore whether pretraining the model on simulated data followed by fine-tuning it on measured data can lead to improved performance on the measured dataset. Further, we want to investigate similar approaches as presented in Lu and Michaels [5] for the temperature valuation in which temperature effects are compensated by using measurements of a baseline dataset, conducted at different temperatures. This approach can be extended to other influencing factors. A further next step could be to test deep learning models on a setup with more than one hole. Here it could be estimated if holes are there (detection), where they are (localisation), and what diameter they have (characterisation). Additionally, we want to test the presented approaches on scattering materials like carbon fibre-reinforced plastic. Finally, measurements using several receivers in a pulse-echo and pitch-catch setup could improve the localisation capability of the TST model significantly. This advancement may bridge the gap needed to significantly boost performance to use such a method in practical applications.

Acknowledgments

The authors acknowledge the support of Dr. Lion Krischer and Dr. Christian Böhm from Mondaic AG for the implementations in Salvus and Clemens Edelmann and Zhixin Wang for conducting the measurements.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

The work was supported by the Technical University of Munich and the Portland State University.

Data availability statement

The raw data are available from the corresponding author upon reasonable request.

References

[1] Cantero-Chinchilla S, Wilcox P, Croxford A. Deep learning in automated ultrasonic nde – developments, axioms and opportunities. NDT E Int. 2022;131:102703. doi: 10.1016/j.ndteint.2022.102703

[2] Krautkramer J, Krautkramer H. Ultrasonic testing of materials. Berlin, Germany: Springer; 1990.

[3] Marin-Cortes A. Ultrasonic coda wave comparison for quality control of manufactured parts: Proof of feasibility. Civil Environ Eng Master's Project Reports. 2020;51. doi: 10.15760/CCEMP.50

[4] Hafiz A, Schumacher T. Monitoring of stresses in concrete using ultrasonic coda wave comparison technique. J Nondestruct Eval. 2018;37(73). doi: 10.1007/s10921-018-0527-8

[5] Lu Y, Michaels J. A methodology for structural health monitoring with diffuse ultrasonic waves in the presence of temperature variations. Ultrasonics. 2005;43(9):717–731. doi: 10.1016/j.ultras.2005.05.001

[6] Diewald F, Epple N, Kraenkel T, et al. Impact of external mechanical loads on coda waves in concrete. Materials. 2022;15(16):5482. doi:10.3390/ma15165482

[7] Clauß F, Epple N, Ahrens M, et al. Correlation of load- bearing behavior of reinforced concrete members and velocity changes of coda waves. Materials. 2022;15(3):738. doi: 10.3390/ma15030738

[8] Jiang H, Asce M, Zhan H, et al. Detecting stress changes and damage in full-size concrete t-beam and slab with ultrasonic coda waves. J Struct Eng. 2021;147(9). doi: 10.1061/(ASCE)ST.1943-541X.0003090

[9] Grabke S, Clauß F, Bletzinger K, et al. Damage detection at a reinforced concrete specimen with coda wave interferometry. Materials. 2021;14(17):5013. doi: 10.3390/ma14175013

[10] Wang Q, Schumacher T, Hafiz A. Damage detection in glass fiber-reinforced plastics using ultrasonic full-waveform comparison. 2019;218:155–163. doi: 10.1007/978-3-030-12111-214

[11] Chen G, Abraham O, Pageot D, et al. Evaluation of nonlinear interface areas in a multiple scattering medium by nonlinear coda wave interferometry (NCWI): Experimental studies. NDT E Int. 2024;141:102992.

[12] Shokouhi P, Riviere J, Lake C, et al. Dynamic acousto-elastic testing of concrete with a coda-wave probe: comparison with standard linear and nonlinear ultrasonic techniques. Ultrasonics. 2017;81:59–65. doi: 10.1016/j.ultras.2017.05.010

[13] Schurr D, Kim J, Sabra K, et al. Damage detection in concrete using coda wave interferometry. NDT E Int. 2011;44(8):728–735. doi:10.1016/j.ndteint.2011.07.009

[14] Snieder R, Grˆet A, Douma H, et al. Coda wave interferometry for estimating nonlinear behavior in seismic. Science. 2002;295(5563):2253–2255. doi:10.1126/science.1070015

[15] Chen A, Schumacher T. Characterization of flaws in structural steel members using diffuse wave fields. AIP Conf Proc. 2014;1581:761–768.

[16] Ebrahimkhanlou A, Salamone S. Single- sensor acoustic emission source localization in plate-like structures using deep learning. Aerospace. 2018;5(2):50. doi: 10.3390/aerospace5020050

[17] Zhao J, Kaiyue Y, Du X. Automated quantification of small defects in ultrasonic phased array imaging using awga-gcforest algorithm. Case Stud Nondestr Test Eval. 2023;1–22. doi:10.1080/10589759.2023.2274001

[18] Tufoi M, Gillich G, Mituletu I, et al. Location of the corrosion damage in rectangular plates. J Phys Conf Ser. 2015;628:012005. doi:10.1088/1742-6596/628/1/012005

[19] Le T, Ho D, Nguyen C, et al. 2022. Structural damage localization in plates using global and local modal strain energy method. Adv Civil Eng, 4456439:1–16. doi: 10.1155/2022/4456439

[20] Cheng L, Yam L, Wong W. Identification of damage locations for plate-like structures using damage sensitive indices: strain modal approach. Comput Struct. 2001;80(25):1881–1894. doi: 10.1016/S0045-7949(02)00209-2

[21] Zuljan D. Effect of ultrasonic coupling media and surface roughness on contact transfer loss. Cogent Eng. 2022;9(1). doi: 10.1080/23311916.2021.2009092

[22] Payan C, Garnier V, Moysan J, et al. Determination of third order elastic constants in a complex solid applying coda wave interferometry. Appl Phys Lett. 2009;94(1):011904. doi: 10.1063/1.3064129

[23] Stanton G, Slinker B. Primer of applied regression and analysis of variance. New York, USA: McGraw- Hill; 1990.

[24] González A, Rodríguez J, Sagartzazu X, et al. Multiple coherence method in time domain for the analysis of the transmission paths of noise and vibrations with non-stationary signals. In: Proceedings of the 2010 International Conference of Noise and Vibration Engineering; Leuven, Belgium. 2010.

[25] Welch P. The use of the fast Fourier transform for the estimation of power spectra: A method based on time averaging over short, modified periodograms. IEEE Trans Audio Electroacoust. 1967;15(2):70–73. doi:10.1109/TAU.1967.1161901

[26] Schumacher T, Linzer L, Grosse C. Signal- based ae analysis. In: Grosse C, Ohtsu M, Aggelis D Shiotani T, editors. Acoustic emission testing. Berlin, Germany: Springer; 2022. doi:10.1007/978-3-030-67936-15

[27] Larose E, Planes T, Rossetto V, et al. Locating a small change in a multiple scattering environment. Appl Phys Lett. 2010;96(20):204101. doi: 10.1063/1.3431269

[28] LeCun Y, Bengio Y, Hinton G. Deep learning. Nature. 2015;521(7553):436–444. doi: 10.1038/nature14539

[29] Vaswani A, Shazeer N, Parmar N, et al. Attention is all you need. NeurIP. 2017. doi: 10.48550/arXiv.1706.03762

[30] Zerveas G, Jayaraman S, Patel D, et al. A transformer- based framework for multivariate time series representation learning. In: Conference on Knowledge Discovery and Data Mining, p. 2114–2124 (2021). 10.1145/3447548.3467401

[31] Howard J, Gugger S. Fastai: a layered api for deep learning. Information. 2020;11(2):108. doi: 10.3390/info11020108

[32] James G, Witten D, Hastie T, et al. Resampling methods introduction to statistical learning. New York, USA: Springer; 2013. doi: 10.1007/978-1-4614-7138-7

[33] Annis C. MIL-HDBK-1823A: Nondestructive evaluation system reliability assessment. Department Def Handb. 2009;7(15).

[34] ASTM. Standard Practice for Probability of Detection Analysis for Hit/Miss Data, Vol. E2862-12. 2023. ASTM International.

[35] ASTM. Standard Practice for Probability of Detection Analysis for â Versus a Data, ASTM-E3023. 2023. ASTM International.

[36] Virkkunen I, Koskinen T, Papula S, et al. Comparison of â versus a and hit/miss pod-estimation methods: A European viewpoint. J Nondestruct Eval. 2019;38(89). doi: 10.1007/s10921-019-0628-z

[37] Berens A. NDE reliability data analysis. ASM Handb. 1989:17:689–701.

[38] Kanzler D. Reliability analysis of digital radiography systems in the testing of real material holes (published in German). [PhD thesis]. University of Rostock. 2016. https://opus4.kobv.de/opus4-bam/frontdoor/index/index/year/2017/docId/38872

[39] Ginzel E, Ginzel R. Approximate dv/dt values for some materials. e-J Nondestr Test. 2017. https://www.ndt.net/?id=21545

[40] Afanasiev M, Boehm C, van Driel M, et al. Modular and flexible spectral-element waveform modelling in two and three dimensions. Geophys J Int. 2019;216(3):1675–1692. doi:10.1093/gji/ggy469

[41] Seeber A, Vrana J, Mosshofer H, et al. Correct sizing of reflectors smaller than one wavelength. In: 12th European Conference on Non-Destructive Testing (ECNDT). 2018. https://www.ndt.net/article/ecndt2018/papers/ecndt-0177-2018.pdf