

https://dspace.org/

A datadriven hybrid approach to generate synthetic data for unavailable damage scenarios in welded rails for ultrasonic guided wave monitoring

D Ramatlo; D Wilke; Philip Loveday

https://hdl.handle.net/10539/42035
WOS

Downloaded from a DSpace repository. https://dspace.org/

Check for updates

Original Article

Structural Health Monitoring
2024, Vol. 23(3) 1890–1913
© The Author(s) 2023
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/14759217231197265
journals.sagepub.com/home/shm
Sage

A data-driven hybrid approach to generate synthetic data for unavailable damage scenarios in welded rails for ultrasonic guided wave monitoring

Dineo A Ramatlo[1], Daniel N Wilke[1] and Philip W Loveday[2]

Abstract

Developing reliable ultrasonic-guided wave monitoring systems requires a significant amount of inspection data for each application scenario. Experimental investigations are fundamental but require a long period and are costly, especially for real-life testing. This is exacerbated by a lack of experimental data that includes damage. In some guided wave applications, such as pipelines, it is possible to introduce artificial damage and perform lab experiments on the test structure. However, in rail track applications, laboratory experiments are either not possible or meaningful. The generation of synthetic data using modelling capabilities thus becomes increasingly important. This paper presents a variational autoencoder (VAE)-based deep learning approach for generating synthetic ultrasonic inspection data for welded railway tracks. The primary aim is to use a VAE model to generate synthetic data containing damage signatures at specified positions along the length of a rail track. The VAE is trained to encode an input damage-free baseline signal and decode to reconstruct an inspection signal with damage by adding a damage signature on either side of the transducer by specifying the distance to the damage signature as an additional variable in the latent space. The training data was produced from a physics-based model that computes virtual experimental response signals using the semi-analytical finite element and the traditional finite element procedures. The VAE reconstructed response signals containing damage signatures were almost identical to the original target signals simulated using the physics-based model. The VAE was able to capture the complex features in the signals resulting from the interaction of multiple propagating modes in a multi-discontinuous waveguide. The VAE model successfully generated synthetic inspection data by fusing reflections from welds with the reflection from a crack model at specified distances from the transducer on either the right or left side. In some cases, the VAE did not exactly reconstruct the peak amplitude of the reflections. This study demonstrated the potential and highlighted the benefit of using a VAE to generate synthetic data with damage signatures as opposed to using superposition to fuse the damage-free responses containing reflections from welds with a damage signature. The results show that it is possible to generate realistic inspection data for unavailable damage scenarios.

Keywords

Ultrasonic guided waves, deep learning, synthetic data, damage, welded rail track

Introduction

The use of guided wave ultrasound (GWU) for structural health monitoring has become prevalent due to its potential. A key aspect of GWU is the ability to allow for the rapid full-volume interrogation of long lengths of structures from a single transducer location. Propagating waves are multi-modal and can be reflected from different benign features such as structural discontinuities and damage. For pulse-echo measurements, the reflections are received by the transmitting transducer. A second transducer is required to measure reflected signals if a pitch-catch mode is adopted. Typical applications of GWU are in pipelines and railway lines where

[1]Department of Mechanical and Aeronautical Engineering, University of Pretoria, Pretoria, Gauteng, South Afric
[2]School of Mechanical, Industrial and Aeronautical Engineering, University of the Witwatersrand, Johannesburg, South Africa

Corresponding author:
Dineo A Ramatlo, Department of Mechanical and Aeronautical Engineering, University of Pretoria, cnr Lynnwood Road and Roper Street, Hatfield, Pretoria, Gauteng 0002, South Africa.
Email: dineo.ramatlo@up.ac.za

researchers have developed permanent monitoring systems. In pipelines, a permanent monitoring system uses a ring of transducers to excite a torsional mode that reflects strongly from the growth of defects produced by corrosion and erosion.[1] Such a system has been used to investigate guided wave behaviour in simple inspection set-ups for above-ground pipelines and in complex scenarios where sections of a pipe are inaccessible due to being insulated, coated or buried underground.[2] An ultrasonic broken rail detection system developed for railway lines is permanently installed on 840 km of rail between Sishen and Saldanha in South Africa.[3–5] The system uses alternate transmit and receive stations of transducers located approximately 1 km apart. Complete rail breaks are detected using a guided wave mode with energy concentrated in the head of the rail. Unfortunately, a rail break can also occur under the locomotives of a passing train, leading to a part of the train derailing.

Railway operators have highlighted the need to prevent complete rail breaks by detecting initial damage such as cracks, as they are the leading cause of train derailments. Early detection of initial damage will minimise the maintenance period, which could improve train operations. Train operators will also have good planning and will not have to wait long before train operation resumes. An example application where researchers have demonstrated that it is possible to detect small damage using ultrasonic guided waves is in pipelines.[6–8] Liu et al.[8] demonstrated how the performance of an installed monitoring system could be evaluated using receiver operator characteristic curves (ROCs). However, challenges often encountered are, first, ROCs require a lot of monitoring data collected under practical environmental and operational conditions (EOCs). Second, such data should contain damage signatures for different damage evolution stages. Inspection data containing damage evolution is unavailable for rail track applications.

To address the challenge posed by the lack of monitoring data with damage evolution, researchers in reference[1] generated experimental data with damage evolving on a 6 m long pipe subjected to varying temperature conditions. Artificial damage was induced in the test structure by drilling a 1 mm diameter, 4 mm deep hole. The diameter of the hole was gradually increased in 1 mm steps from 1 to 7 mm in successive temperature cycles to simulate damage growth under varying temperature conditions simulated in their laboratory. In railway lines, it is impossible to generate synthetic data using laboratory experiments as the end reflections from a short section of rail will mask reflections from welds and damage. Monitoring data can be collected from an operational rail in the field subjected to real EOCs. However, the data will not contain damage signatures as damaged sections of rail are immediately replaced with new ones.

Alternatively, synthetic data can be generated by modelling damage reflections using the finite element method and adding the signals to experimental data collected over multiple environmental cycles from a structure without damage.[8,9] This approach was shown to be successful on a rudimentary laboratory pipe set-up. The framework proposed by Liu et al.[8] was validated in the study by Heinlein et al.[10] by investigating damage detection performance on synthetic data produced from the superposition of simulated damage from a finite element model (FEM) with experimental data. The performance of this scheme was compared to data produced by inducing artificial damage on a test structure. Very good agreement was obtained between the results produced using the purely experimental data and those obtained from the superposition of FEM and experimental responses. This work thus suggests that it is possible to generate synthetic data for railway lines by fusing damage simulated from a FEM with response signals collected from an operational rail in the field.

Recently, there have been developments in the use of digital twin models to bridge the gap between the simulation and the real world. A digital twin model can be defined as having a digital model and a physical system where the two are exact replicas of each other. The benefit of having digital models that are equivalent to physical systems is that they offer the ability to thoroughly interpret the data to understand better how those systems work. Another significant benefit is that digital models make it possible to predict data that a physical system cannot produce, for example, inspection data with growing damage signatures. Ramatlo et al.[11] developed a finite element modelling framework to simulate guided wave inspections in welded rails. The modelling framework only accounts for direct reflections from welds. The waves are excited by a resonant transducer model validated in reference.[12] The model for calculating the scattering caused by complex discontinuities such as welds is based on a technique presented by Benmeddour et al.[13] and validated by Long and Loveday[14] for aluminothermic welds in rails. This procedure adopts a hybrid model combing a 3D FEM of a weld with two semi-analytical finite element (SAFE) models of the incoming and outgoing rails. The modelling framework is continuously enhanced as the aim is to attain a digital model that can capture the complexities in an operational system. In reference,[15] the digital model was improved to account for multiple reflections, which are most common when working in the web section of the rail. The digital model was validated using a field experiment from an operational railway line. Reflections from

welds were accurately predicted, implying that it is possible to produce virtual experimental responses and to model and simulate realistic responses for unavailable damage scenarios.

When generating synthetic data for the development of monitoring systems, it is essential to consider all possible application scenarios. In addition to including practical EOCs in the data, different types of damage, possible damage growth rates, damage location and other factors should be considered. For pipeline applications, Liu et al.[8] demonstrated a procedure to model damage growth rates for corrosion. Another critical factor to consider is to make the synthetic data as realistic as possible. The simulated damage signatures need to be realistic, and the fusing procedure adopted to combine simulated damage with experimental data should also be realistic. Therefore, it should be impossible to distinguish between a simulated reflection and a true reflection measured from the field experiment. For pipeline applications, synthetic data was developed by following the principle of superposition, where a damage signal simulated using a finite element procedure is added to an experimental signal.[8–10]

In welded railway lines, dispersion behaviour, modal interaction and overlapping reflections from different sources add complexity to the response signals. Furthermore, in addition to direct reflections, multiple discontinuities introduce double reflections that occur when the waves reverberate between discontinuities. This increases the complexity of the response, which renders superposition too simplistic to generate realistic damage responses, as the double reflections resulting from the damage are ignored. The superposition method only works on signals with well-defined characteristics when the reflection caused by the damage is small, does not interact with other discontinuities, and the transmitted signal is not affected. Hence, it cannot deal with complex data variations resulting from inspection set-up complexities. It will thus not be possible to account for significant multiple reverberations that the damage introduced in the waveguide could cause. In addition, the signal transmitted through the damage could be reduced, resulting in decreased reflections from further defects. This becomes even more complicated when mode conversion is present at the damage. To accommodate this additional complexity, machine learning algorithms are of great use as they can recognise the pattern change due to damage.

Recently, many demonstrations of machine learning techniques that address several tasks in the field of GWU have started to emerge in the research community. These techniques show a great advantage as they can model complex behaviour with high efficiency though they require extensive training data. The application of machine learning techniques in structural health monitoring using GWU has been proven successful in studies concerned with damage detection, severity, location and characterisation.[16] The main building block in most of these approaches is a convolutional layer that applies a filter to the training data to extract the underlying features. In the studies by Khurjekar and Harley,[17,18] a deep convolutional neural network-based framework for damage localisation in the presence of uncertainty was proposed and applied to a 1 m square plate. Damage location was modelled as a multi-modal probability distribution, which made it possible to identify multiple damage locations in the plate. The neural network was trained solely with simulated data, and the analysis was extended to experimental data with temperature variations.[18] This approach proved to be robust to uncertainty and showed a competitive performance to traditional localisation methods. Deep convolutional neural networks were also employed by Zhang et al.,[19] Wu[20] and Rautela and Gopalakrishnan[21] in plate-like structures for damage detection, though the techniques were applied to 2D time-frequency wavelet transform image data instead of 1D time series signals. A pipeline waveguide was considered in the study by Pyle et al.,[22] but the model was simplified by approximating a large diameter pipe wall as a 10 mm thick plate.

Machine learning techniques which are mostly used for synthetic data generation, include generative adversarial networks (GANs) and variational autoencoders (VAEs). In general, GANs tend to be more widely used with image data, while VAEs see more use in signal data. The VAE model is based on the principle of dimensionality reduction, where the encoder compresses the input data into a compact form representative of the underlying structure of the data. The low-dimensional data is then transmitted to a decoder which will then generate an output which will be based on the input data used for training. The primary goal of a VAE is to learn the mapping from the input data to low-dimensional output data. Deep convolutional neural layers or other sophisticated layers such as the long short-term memory are often used as the main building blocks of the encoder and decoder networks when dealing with time series data.

The VAE is trained by minimising the reconstruction error between the input and output data when a bottleneck is imposed in the latent space. VAEs have consistently proven to accurately predict the temporal attributes of the original data,[23] though their application to synthetic data generation in GWU is limited. To the authors' knowledge, no attempt has been made to apply deep convolutional neural networks or VAEs to generate synthetic data for long-range GWU inspections of welded railway lines. The authors are only aware of a recent contribution by Mahajan and

Banerjee,[24] where classification and regression-based machine learning algorithms were employed for the task of damage detection in railway lines. They utilised high-frequency GWU signals in a 1.5 m long International Railway Solutions (IRS) 52 rail section to detect damage, characterise it and determine its severity and location zone in the rail head. The guided waves in the rail were generated and received by two single contact-type piezoelectric wafer transducers attached to the rail web in a pitch-catch arrangement. The machine learning framework was trained and validated using time history, frequency spectrum and wavelet transform of the response signals with different types of damage simulated in ABAQUS explicit finite element software. The framework was also tested using laboratory experiment results. Transverse damage was created using a grinder over the surface of the rail head and was located 50 mm from the transducer. It was found that it is possible to detect a minimum defect size of 5% of the head cross-section area with 1 mm thickness. In plate structures, Khurjekar and Harley[25] proposed a VAE framework for damage detection, and the VAE was trained on simulated data.

From a review of past literature, it was found that although machine learning techniques have been applied to GWU monitoring for damage detection, training is mostly performed using purely simulated or synthetic data. This is because obtaining experimental data under practical conditions with many different damaged geometries and locations is generally infeasible. However, the subject of generating realistic synthetic signals continues to be important as the performance of monitoring systems will depend on the practicability of the data used during the development stage of such systems. Realistic inspection data for unavailable damage evolution scenarios can be used to compute ROC curves to predict the performance of monitoring systems.

Therefore, in this paper, we explore the use of a VAE to generate synthetic data for unavailable damage scenarios in welded rail ultrasonic guided wave inspections. We consider an operational rail with multiple aluminothermic welds as discontinuities. A mode with energy concentrated in the head of the rail is excited by a piezoelectric transducer and propagated to distances up to 400 m in a pulse-echo set-up. The training data used in the VAE is obtained from a physics-based digital model of the field experiment at a specific EOC. This physics-based model employs models of excitation by a piezoelectric transducer, propagation of guided waves and scattering from discontinuities as presented in the study by Ramatlo et al.,[15] and accounts for the complex reverberating reflections during propagation.

Traditionally, VAEs are used to reconstruct the input data through a lower-dimensional latent representation. In this paper, we extend the novelty of reconstructing an output target resembling the input data but enhanced in some respect, reminiscent of denoising VAEs.[26] Unlike the denoising VAEs that reduce the complexity of the input data by removing information from the input signal, we enhance the complexity of the input data by adding information to the input signal in the form of the measured response in the presence of an additional defect. Given a baseline signal with reflections from welds as an input, the reconstructed output will contain an additional reflection from a defect at a user-specified position in the rail. Therefore, the procedure that we propose, illustrated in Figure 1, is an enhancing VAE as it adds more complexity and new features to the input signal. The procedure consists of three phases.

In the first phase, we prepare the training data and initialise the VAE architecture. The training data generated from a physics-based model is a set of simulated baseline signals containing only reflections from welds and corresponding simulated target signals that combine the baseline signals with damage at specific positions. This data did not include aleatoric (inherent randomness) and epistemic (model error) uncertainties, as the focus of the study is on the ability of a VAE to capture the relevant physics such as reflections, as opposed to some superposition of noise. In the second phase, the VAE is trained to generate synthetic data containing a damaged signature at a user-specified distance on either the left or the right side of the transducer by learning the mapping from the baseline signal to the target signals for a user-specified distance. During training, a subset of baseline and simulated target signals is sampled from the training data. The baseline signals are inputted into the encoding network of the VAE, which compresses the data into a lower-dimensional space represented by the mean and variance of a normal distribution. The data is then reparameterised and fed into the decoding network of the VAE. The decoder learns to add damage to each baseline signal based on the specified positions. The output of the VAE is the predicted target signals, which are compared to the sampled simulated target signals by calculating the VAE loss. If the reconstruction ability of the VAE is unsatisfactory, the training phase is repeated with a different subset of training data until an optimal VAE model with an acceptable latent representation and reconstruction ability is achieved.

Finally, in the last phase, the trained VAE model is utilised to generate response signals with damage fused at user-specified positions.

In this study, the primary objective is not to detect damage features but rather to train a VAE model to integrate known damage signatures into a baseline signal that is initially damage-free. To verify the

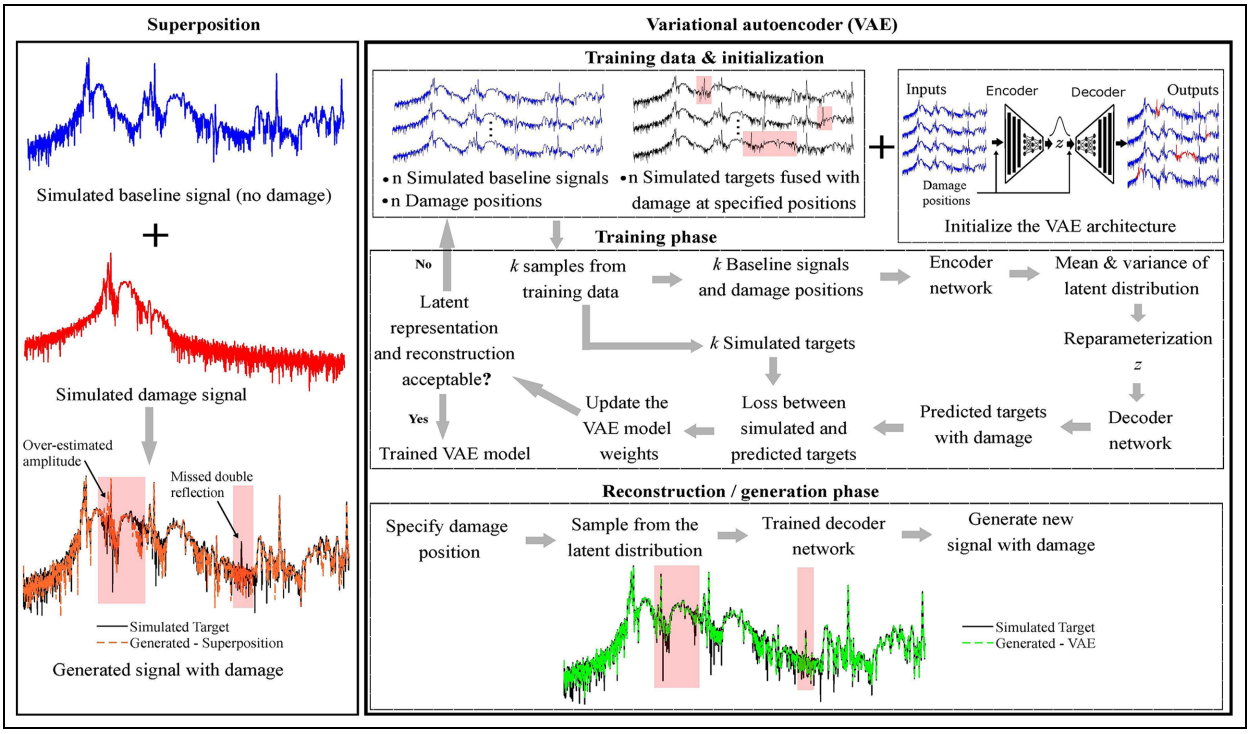


Figure 1. The generation of synthetic data containing damage signatures using the VAE and the superposition principle.

agreement between the damage signatures included in the simulated target signals using the physics-based model and those reconstructed by the VAE, the authors adopted a distance domain analysis approach. Instead of arbitrarily selecting specific damage features to assess, the authors employed fundamental error measures such as mean absolute errors (MAE), auto-correlation and maximum absolute errors to evaluate the signals. This approach allowed for a comprehensive evaluation of the performance and fidelity of the VAE model in fusing the desired damage signal with the baseline signal. We evaluate and compare our proposed method with the results obtained from the traditional superposition method. The performance of the trained VAE model is first validated in a supervised setting by using simulated experimental data that enhances a baseline signal with a simulated damage signature. Second, the ability of the VAE model to generate synthetic data containing damage is validated using experimental field data collected from a damage-free operational railway line. These evaluations aim to demonstrate the effectiveness and potential of the VAE approach in generating synthetic data for unavailable damage scenarios. It is important to note that any further refinements, exploration of alternative architectures and other related investigations are left as future work and require dedicated scientific studies.

The virtual experimental data used to train the VAE is described in the section 'Simulation of guided wave inspection in the head of a rail'. The section 'VAE fusing damage signature with virtual experimental data' explains the VAE architecture, followed by a presentation of results. The final section gives the main conclusions of the study.

Simulation of guided wave inspection in the head of a rail

One important aspect to consider when generating synthetic data by fusing simulated damage signatures with experimental measurements collected from the field is to what extent they are distinct from actual experimentally measured damage. Ideally, a VAE-fused reflection should be indistinguishable from a real reflection. However, this is usually not feasible due to a lack of experimental data. In this paper, we circumvent this by first investigating in a supervised setting the augmentation of finite element simulated responses, seen as simulated experimental data, with an additional simulated damage signature. We then compare this against the finite element simulated response when additional damage is present in the FEM. This allows us to critically investigate the difference between the damage reflection of the VAE-fused response and the response from the finite element simulated damage reflection. Given the virtual experimental response with reflections from known welds as our baseline signal, this

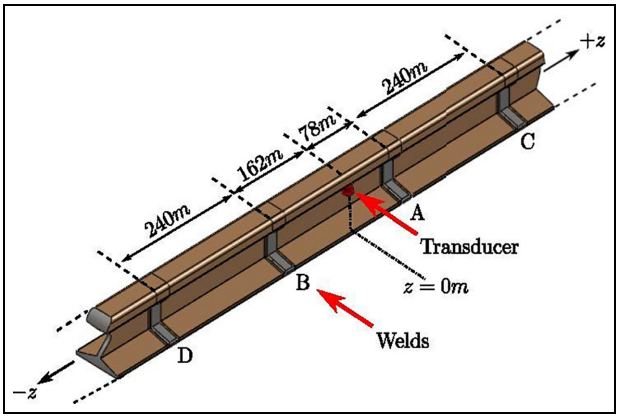


Figure 2. Schematic representation of a section of the field operational rail containing four welds.

paper uses the VAE to fuse a simulated damage reflection at a user-specified distance. The damage model is presented in subsection 'Transverse damage in the head of the rail', and virtual experimental responses containing damage are discussed in subsection 'Virtual experimental response with damage in the head of the rail'. The models explained in this section will be used to simulate response signals for training and testing the VAE. These signals did not include aleatoric and epistemic uncertainties as the focus of the study is on the ability of a VAE to capture the relevant physics such as reflections, as opposed to some superposition of noise.

Simulated virtual experimental baseline signal for a damage-free rail

A field experiment was previously performed on a UIC60 rail in an operational heavy-haul rail track.[3] The rail consisted of 240 m long sections welded together by four aluminothermic welds. The field experiment was performed by exciting the guided waves propagating in the head of a rail with a piezoelectric transducer located at approximately 78 m from the nearest weld. The pulse-echo transducer was driven by a 17.5-cycle Hanning windowed tone burst voltage signal with a centre frequency of 35 kHz. The excited guided waves were transmitted in both directions along the rail. The experimental response was constructed by measuring the reflections from welds using the transducer. A schematic representation of the field layout of the section of rail considered is illustrated in Figure 2 and the measurement details are provided in the study by Loveday et al.[3]

The virtual experimental responses for training the VAE were simulated using a digital model replicating the experimental field scenario in Figure 2. This digital model is a physics-based framework presented and validated in the study by Ramatlo et al.[15] The digital model for the set-up considered in this paper is illustrated in Figure 3(a). For completeness, the model will be briefly explained, but the reader is referred to the studies by Ramatlo et al.[11,15] for additional details.

First, the modelling framework considers the individual elements of the inspection set-up. These elements include excitation using a transducer, propagation of waves in regions of constant cross-section and scattering from discontinuities. For excitation, the piezoelectric transducer was modelled using 3D FEM accounting for the mechanical and electrical properties of the transducer. This resonant transducer was coupled to a 2D cross-section of the waveguide through the computation of the frequency-dependant dynamic stiffness, as explained in the study by Long et al.[12] A voltage signal was applied to the transducer, and equivalent mechanical

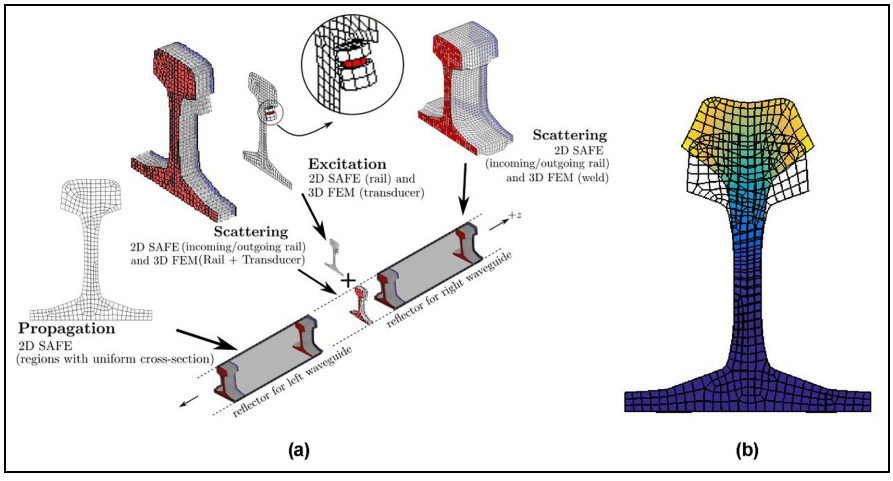


Figure 3. (a) Digital model of the considered section of an operational rail and (b) the target mode shape with energy concentrated in the head of the rail.

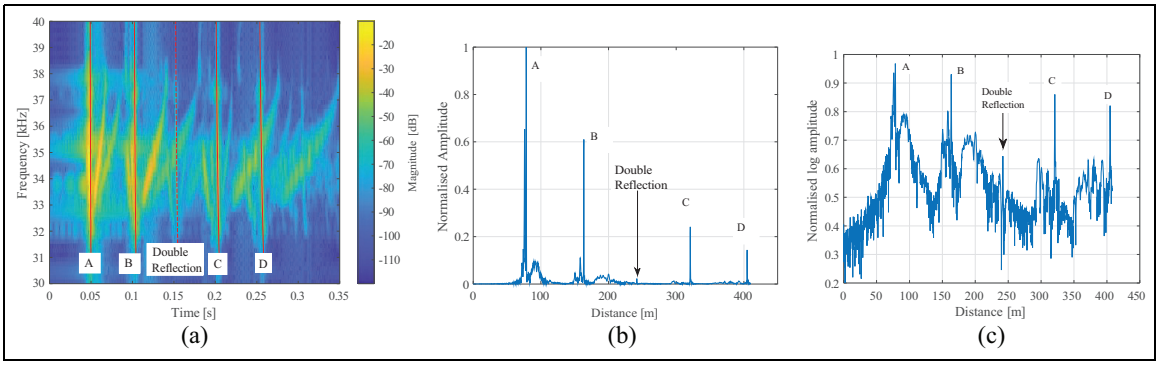


Figure 4. (a) Spectrogram of the simulated baseline signal for the experimental response and (b) & (c) the equivalent distance domain signals.

forces were applied to the 2D cross-sectional model using the dynamic stiffness matrix. This 2D cross-sectional model is based on a SAFE method used to model wave propagation in regions of constant cross-section. The hybrid model combining the 3D model of the transducer and the 2D model of the rail computed the mode shapes and associated modal amplitudes resulting from the transducer excitation. These wave modes can be propagated along the waveguide length by applying analytical variations in the direction of propagation. The mode that is strongly excited by the transducer is shown in Figure 3(b). The scattering of guided waves from discontinuities such as welds is modelled using a second hybrid model, which couples a 3D FEM model of the reflector with two SAFE models to represent the semi-infinite incoming and outgoing rails on either side of the reflector. The hybrid model for scattering from discontinuities solves the modal amplitudes of the reflected and transmitted guided waves by enforcing continuity and equilibrium on the boundaries of the left and right semi-infinite waveguides intersecting the 3D volume of the reflector. The propagation properties calculated from the SAFE models account for dispersion and attenuation in the rail. The results of the individual elements of the inspection set-up for the UIC60 rail considered were published in the study by Ramatlo et al.[11]

Second, the modelling framework combines several local discontinuities as one reflector by introducing a single scattering matrix, according to the study by Baronian et al.[27] The two weld models for welds A and C, located to the right of the transducer, are coupled together to create a single domain of reflectors. This domain is referred to as the right waveguide. Hybrid models for welds B and D are also coupled together to create a domain for the left waveguide. The general scattering matrix for each region is computed using the reflection and transmission matrices for each weld and propagation terms from dispersion properties. The computation procedure accounts for the complex reverberations of guided waves between the welds.

Third, the method of Baronian et al.[27] is further employed according to Ramatlo et al.[15] to create a model for the entire section of rail considered. To account for the computation of reverberating reflections, a scattering model of the transducer is included. This model will allow the waves to be transmitted and reflected through the transducer region. The models of excitation and scattering from the transducer are thus coupled with models of the left and right waveguides, and the wave modes are propagated between these domains. Response signals are then calculated at the transducer location in the frequency domain. They are finally converted to the distance domain by applying an inverse Fourier Transform followed by a dispersion compensation procedure.[28] Details of the coupling procedure and computation of response signals can be found in the studies by Ramatlo et al.[15] and Baronian et al.[27]

Figure 4 shows a spectrogram of the simulated virtual experimental response and the equivalent distance domain signals. The response was validated using the measured experimental response in the study by Ramatlo et al.[11] The spectrogram shows that the simulated experiment contains many propagation modes that are reflected from each weld. The energy of the reflections decays exponentially with time due to attenuation in the rail. The mode with energy concentrated in the head of the rail was strongly excited and appeared to be almost non-dispersive. This mode is identified by a vertical trace in the spectrogram and the highest amplitude in each weld reflection in the distance domain signals. Other modes reflecting from welds are also evident. These modes are very dispersive as their propagation velocity differs as a function of frequency. Some of these modes are coupled in pairs implying that a particular incident mode is reflected as a different mode from the weld. Another important feature of the virtual experimental response is the double reflection caused when the waves reverberate between welds A and B. The VAE model will be trained and tested using

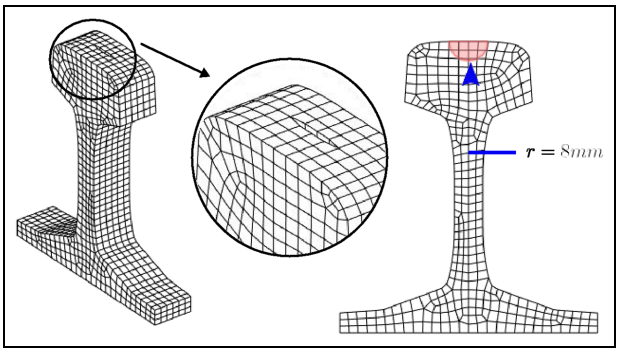


Figure 5. A hybrid model of the defect geometry with an 8 mm radius crack in the head of the rail.

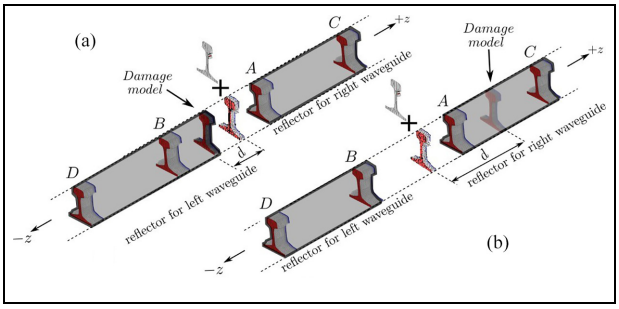


Figure 6. Digital models of the considered section of an operational rail with fake damage at specified distances *d* (a) between the transducer and weld B, and (b) between welds A and C.

the distance domain response signals with the amplitude on a log scale as shown in Figure 4(c). This is because, when comparing Figure 4(b) and (c), the features of the response signal become more apparent when the log of the amplitude is computed.

Transverse damage in the head of the rail

The scattering of guided waves from damage in the rail was modelled using a hybrid modelling technique proposed by Benmeddour et al.[13] The technique, which was also used to model the scattering from welds, calls for both a SAFE and a solid 3D finite element implementation. Two SAFE cross-sectional meshes are attached on either side of the 3D mesh enclosing a transverse damage originating in the head of the rail. The solid 3D finite elements used to model the region with damage are standard displacement-based elements. The correctness of our implementation has been confirmed by Long and Loveday,[29] who looked at several defect geometries of different sizes located at different regions in a UIC60 rail. Figure 5 shows the hybrid model of the defect geometry considered in this paper. The defect is modelled as a circular crack at the top of the head section with an 8 mm radius. The hybrid model was generated using 8-noded quadrilateral SAFE elements and 20-noded solid brick elements. This hybrid model is employed in the next section to simulate inspections with damage at different positions along the length of the rail.

Virtual experimental response with damage in the head of the rail

Let us consider the case of an inspection in the presence of damage in the rail at a user-specified position. As this scenario is virtual, we can simulate it using a physics-based digital model of the rail. The model of the crack in Figure 5 is incorporated into the digital model of the considered section of an operational rail in Figure 3. The virtual experimental response containing damage was simulated using the modelling framework procedure discussed in the section 'Simulated virtual experimental baseline signal for a damage-free rail'. The simulation contained four models of aluminothermic welds located according to the set-up in Figure 2 and one model of a crack placed at an arbitrary distance. Figure 6 shows two scenarios of the digital model for damage at different locations. In Figure 6(a), the crack is located between the transducer and weld B. In Figure 6(b), the crack is located between welds A and C. Response signals were generated by varying the crack position in the considered section of rail with four welds.

Examples of simulated responses with damage at different locations in the rail are plotted in Figure 7. The amplitude of the damage decreases with an increasing distance, implying that attenuation in the rail was well captured. The attenuation model employed was validated in the study by Ramatlo et al.[15]

In the next section, we demonstrate how a VAE can be used to reconstruct virtual experimental signals containing damage by fusing the baseline signal in Figure 4(c) with damage reflections at user-specified distances.

VAE fusing damage signature with virtual experimental data

Dispersion behaviour, modal interaction and overlapping reflections from different sources add complexity to the response signals. Furthermore, in addition to direct reflections, multiple discontinuities introduce double reflections that occur when the waves reverberate between discontinuities. This increases the complexity of the response, which renders superposition too simplistic to generate realistic damage responses, as the double reflections resulting from the damage are

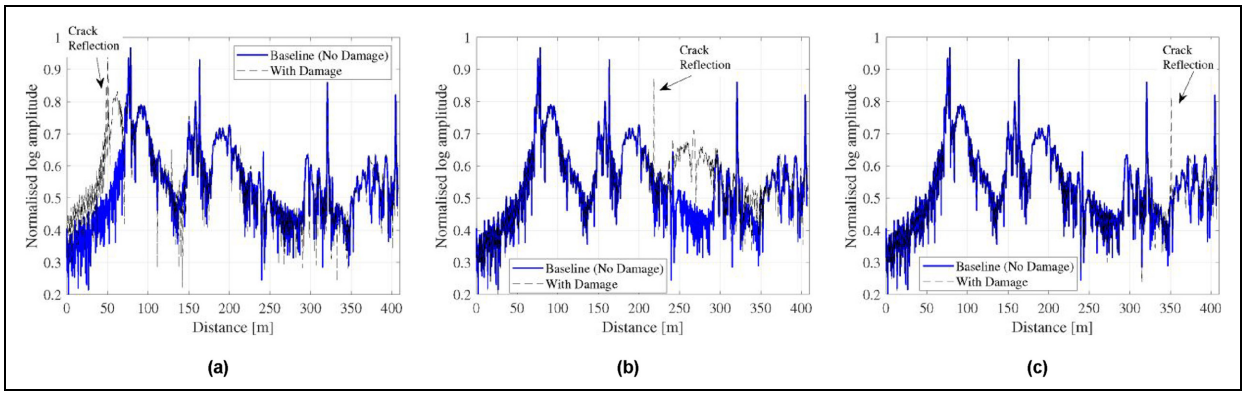


Figure 7. Virtual experimental response signals (y_i) containing damage reflections at user-specified distances of (a) 50 m, (b) 220 m and (c) 350 m.

ignored. To accommodate this additional complexity, machine learning algorithms are of great use as they can recognise the pattern change due to damage. In this section, a VAE has been developed, trained and tested using virtual experimental signals from a physics-based model.

The proposed framework for generating synthetic data with damage signatures is based on the principle of dimensionality reduction. The VAE consists of two independent networks, an encoder and a decoder, connected through a lower-dimensional latent space $\boldsymbol{z}$. The encoder compresses the input data to a lower-dimensional space that maps the data to a continuous latent vector $\boldsymbol{z}$. The decoder then takes the latent variable and maps it to a higher dimensional space to reconstruct an output approximating the target. Traditionally, VAEs are used to reconstruct an output target that approximates the input data. In this paper, we introduce the novelty of reconstructing an output target that is an enhanced input signal. Given a baseline signal with reflections from welds as an input, the reconstructed output will contain a response in the presence of an additional reflection from the crack. Therefore, the procedure that we propose is an enhancing VAE as it adds more complexity and new features to the input signal.

Training and testing data

As explained in the section 'Simulation of guided wave inspection in the head of a rail', the simulation results will be used for training and testing the VAE for synthetic data generation. The VAE requires two sets of virtual experimental data to solve this task. The first set of input data is the baseline signal. This signal is free of damage and contains reflections from four identical weld geometries, as explained in the subsection 'Simulated virtual experimental baseline signal for a damage-free rail' and shown in Figure 4(c). The second set of data is response signals containing damage reflections at arbitrary distances. This damage can either be included on the left or right side of the transducer, between any two discontinuities, as illustrated by examples of the rail set-up in Figure 6. The VAE model will then learn the mapping from the damage-free baseline signal with only reflections from welds to the target signal with reflections from welds and a crack geometry at a given distance.

For training and testing the VAE, the baseline signal, and a dataset of $n = 3173$ samples of target signals with damage were simulated. The VAE was trained using 70% of this data (2220 randomly selected signals), and the remaining 30% (remaining 953 signals) were used for testing the model. This training and testing data did not include noise, as the focus of the study is on the ability of a VAE to capture the relevant physics, such as the scattering from discontinuities. The distance domain response signals obtained from applying dispersion compensation to the time domain responses were first normalised according to the amplitude of the reflection from weld A. The logarithm of the amplitude of the signals was then computed.

The VAE architecture

Given a sample of n baseline signals containing reflections from four welds, each denoted by vectors $\boldsymbol{x_i} \in \mathrm{R}^{\mathrm{m}}$ where $m = 2400$ is the length of the signal and $i = 1, 2, 3, \ldots, n$; the distances $d_i \in \mathrm{R}$, and a sample of ones and zeros $s_i \in \{0, 1\}$, with one indicating *right* and zero indicating *left*; we want to train the VAE to fuse each baseline signal, with a new reflection from the damage located at a distance of d_i meters from the transducer on either the right or left side. We specify the side where the damage reflection should be included so that the two reflectors between which the user intends to add

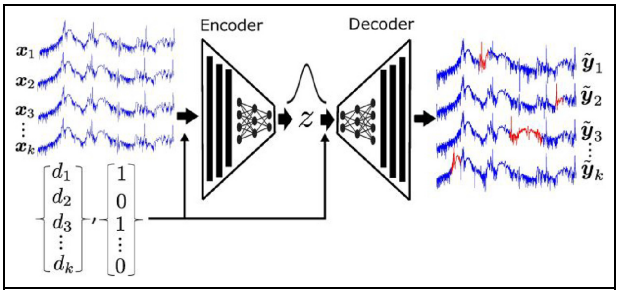


Figure 8. VAE architecture illustrating the construction of $k < n$ random target signals from the training dataset with n samples.
VAE: variational autoencoder.

the damage, as illustrated in Figure 6, become obvious. However, it is possible to train the VAE without providing this information. The distance domain baseline response signals in Figure 4 form the main inputs for our VAE. The second meaningful inputs are the distances where the damage should be included. The third input is a vector of ones and zeros to indicate the side relative to the transducer, where damage should be included. The VAE should be trained to approximate the target signals $\boldsymbol{y}_1$, $\boldsymbol{y}_2$, …, $\boldsymbol{y}_n$ with damage reflections fused according to the distance d_i and the specified side s_i relative to the transducer. The length of the signals, $m = 2400$, was selected such that the reflections from the four welds were all included. The same window length was used for all signals to allow us to focus on the premise of the study. Shorter signals may require the VAE to be combined with other techniques, such as the autoregressive models, and the accuracy of the model might be affected.

This paper aims to investigate whether autoencoding has the potential to fuse damage features to a baseline signal to generate synthetic data for unavailable damage scenarios. For this reason, a standard VAE architecture illustrated in Figure 8 and detailed in Table 1 was employed. The design of this VAE architecture was guided by Khurjekar and Harley,[25] and refinements were made based on the nature of the input data employed in this study. The architecture used in the study by Khurjekar and Harley[25] was for only one input variable and consisted of only three different types of layers; the Conv1D layers for extracting the underlying structure in the input signals, the Dense layers to decrease and increase dimensionality, and to connect the encoder and the decoder through the latent space, and the Conv1DTranspose layers to apply a transposed 1D convolution operation. In this paper, we also include Embedding layers since the VAE needs to enhance the input baseline signals with damage signatures according to the user-specified positions. In reference,[25] the Embedding layers were not used as the VAE architecture was used to reconstruct the input data without adding new features. By adding new features through embedding layers, the VAE can learn to capture the relationship between the original features in the baseline signals and the newly introduced damage features and learn the meaning and interpretation of how the features vary with respect to the distance and the side indicator inputs. This is achieved by placing similar inputs close together in the embedding space. The learning process of the Embedding layers is

Table 1. VAE architecture.

Layer	Layer description	Activation
Encoder		
Embedding ×2	Embeddings for the distance and side indicator	–
Concatenate	To concatenate the baseline signals and the two embeddings	–
Time Distributed Dense	Fully connected; nodes = 1 for each time step	Sigmoid
Conv1D	Filters = 8; kernel size = 3; stride length = 2	eLu
Conv1D	Filters = 16; kernel size = 3; stride length = 2	eLu
Conv1D	Filters = 32; kernel size = 3; stride length = 2	eLu
Dense	Fully connected; nodes = 150	Sigmoid
Dense ×2	Latent dimension = 2	–
Decoder		
Dense	Fully connected; nodes = 150	Sigmoid
Dense	Fully connected; nodes = 150 × 2 × 32	Sigmoid
Conv1DTranspose	Filters = 16; kernel size = 3; stride length = 2	eLu
Conv1DTranspose	Filters = 8; kernel size = 3; stride length = 2	eLu
Conv1DTranspose	Filters = 1; kernel size = 3; stride length = 2	eLu
Concatenate	To concatenate the output from the previous layer and the two embeddings	
Time Distributed Dense	Fully connected; nodes = 1 for each time step	Sigmoid

VAE: variational autoencoder.

dependent on the objective function of the VAE, Equation (2), which is minimised and used to determine the VAE weights that better capture the meaning of the data. We further included Concatenate and Time Distributed Dense layers for dealing with the three input variables, the baseline signal, user-specified distance and the specified side.

For each training step, the encoder receives as input 3D sequences resulting from the concatenation of a batch of $i = 1, 2, \ldots k < n$ randomly sampled baseline signals $\boldsymbol{x}_i$, the distance embeddings d_i and the embeddings s_i that indicates the side where the damage should be added, right or left relative to the transducer. A Time Distributed layer is then used to apply the same instances of a Dense layer to every temporal slice of the 3D sequences. The encoder consists of a stack of three sequential Conv1D layers connected to a fully connected Dense layer with 150 nodes, which is then connected to two dense layers to approximate the mean μ and the variance σ of the 2D latent space as a normal distribution.

The decoder samples from the 2D latent distribution

$$\boldsymbol{z} \sim \boldsymbol{N}(\mu, \boldsymbol{\sigma}), \tag{1}$$

and increases the dimensionality of the data using two stacks of fully connected Dense layers, with 150 nodes and 150×2 nodes, respectively. A stack of three Conv1DTranspose layers is used to increase the dimensionality further. The generated sequences are then concatenated with the two embeddings and passed through a Time Distributed Dense layer to approximate the target signals $\boldsymbol{y}$ containing reflections from the crack.

The VAE is trained by minimizing the reconstruction loss, which is the MAE between the original target $\boldsymbol{y}$ from Figure 7 and the approximation $\tilde{\boldsymbol{y}}$:

$$\text{Reconstruction loss} = \frac{\sum_{i=1}^{k} |\tilde{\boldsymbol{y}}_i - \boldsymbol{y}_i|}{k} \tag{2}$$

A batch size of $k = 128$ samples was used for each training step.

Results

The results of synthetic inspection signals for the rail set-up containing a crack geometry, described in the section 'Simulation of guided wave inspection in the head of a rail', are presented here. We consider five cases of results from the testing set. In the section 'VAE reconstruction', we first compare the simulated target constructed from the physics-based modelling framework and the approximated response signals reconstructed using the VAE architecture. After that, we highlight the benefit of using a VAE to generate the synthetic data with damage as opposed to using superposition to fuse the damage-free response containing reflections from welds with a damage signature in the section 'VAE comparison to superposition for damage fusion'. The VAE model was trained using the simulated virtual experimental data. In section the 'Field Experiment', we attempt to use the trained VAE model to fuse the simulated damage signature to an experimental signal collected from an operational rail.

VAE reconstruction

In Figure 9, response signals reconstructed using a VAE model are compared to the original target responses for several crack locations in the rail. The location of the crack in the rail for each signal is highlighted using a black dot marker and a black vertical dotted line. The area shaded in green indicates the region over which the energy reflected from the damage in the target and reconstructed response signals is significant.

First, we notice that the VAE was able to capture the complex features in the reflected signals. For the rail set-up considered in this paper, the most dominant mode excited by the piezoelectric transducer is the least dispersive mode with energy concentrated in the head of the rail, plotted in Figure 3(b) and referred to as mode 7. This mode was used to compensate for dispersion according to the procedure in the study by Wilcox.[28] Hence the mode is identified as a sharp peak in each reflection group. Other modes with energy in the head of the rail contribute to the response signals, as explained in the section 'Simulation of guided wave inspection in the head of a rail' and in detail in the study by Ramatlo et al.[11] Some of these modes exist individually, while others exist as coupled modes. These modes were not perfectly compensated for dispersion. Hence, their energies are spread out. In Figure 9(a) to (e), it is evident that the reflections from welds in the VAE reconstructed responses are comparable to the finite element simulated virtual experiment target. Furthermore, the VAE was able to reconstruct the double reflection located at approximately 240 m, resulting from the reverberation of mode 7 between welds A and B, as shown in Figure 9.

Secondly, we notice that the VAE model successfully generated synthetic inspection data by fusing reflections from welds with the reflection from a crack model at the specified distance from the transducer on either the right or left side. In all five cases of results considered, damage reflection patterns associated with mode 7 and other contributing modes were well approximated. However, in some cases, the VAE struggled to

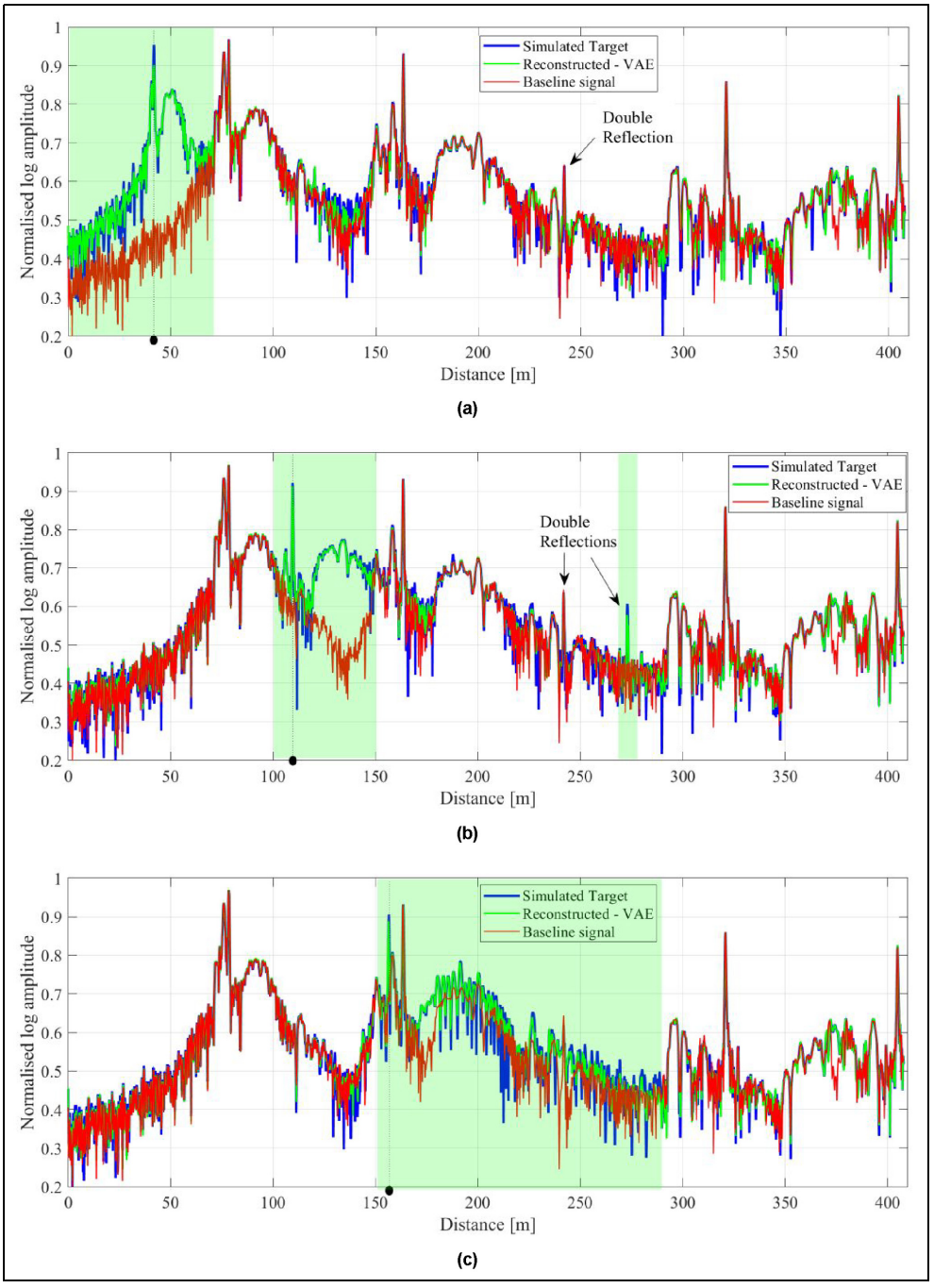


Figure 9. (continued)

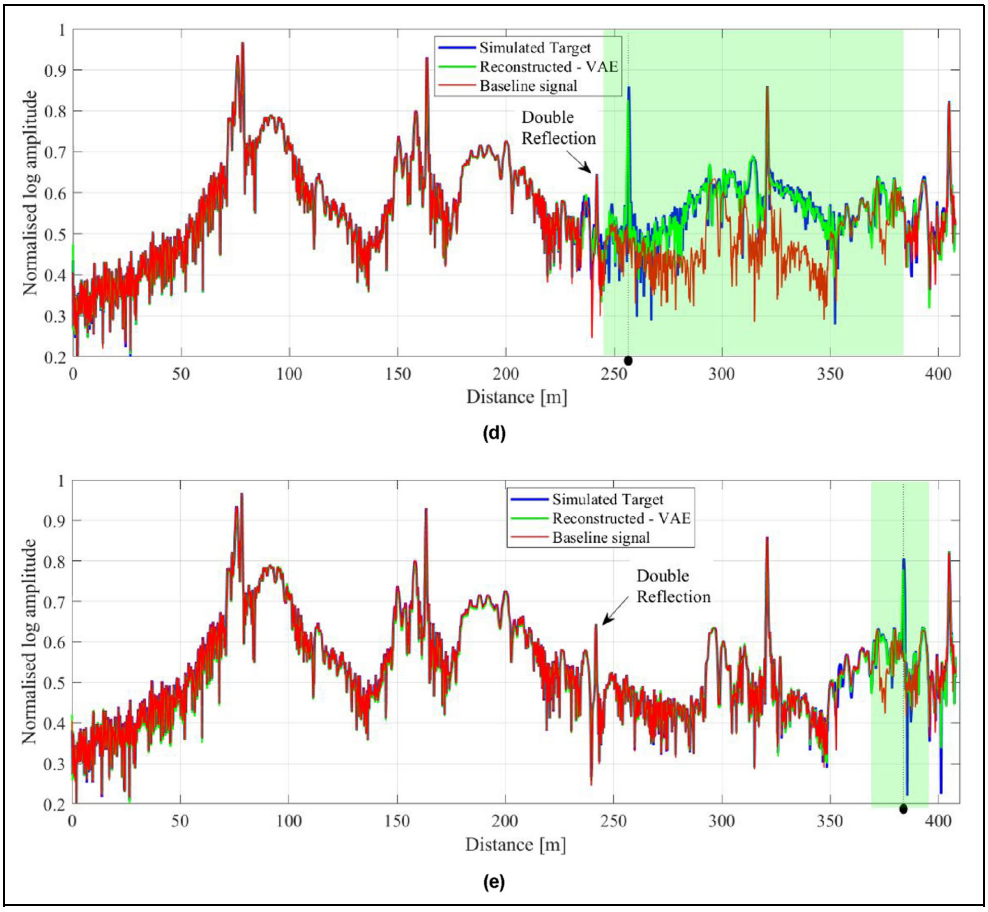


Figure 9. Response signals reconstructed using a VAE compared to the original target response for a crack located at (•···) (a) 41.375 m from the transducer on the left side, (b) 108.5 m from the transducer on the right side, (c) 154.8750 m from the transducer on the right side, and (d) 253 m from the transducer on the left side and (e) 379.25 m from the transducer on the right side. The area shaded in green indicates the region over which the energy reflected from the damage in the target and reconstructed response signals is significant.
VAE: variational autoencoder.

match the signal energy for mode 7. This observation is clear in Figure 9(a), where the reconstructed peak was approximately 5% smaller than the target peak and slightly visible in Figure 9(d) to (e).

Thirdly, in scenarios where the crack was located closer to a weld or the two were located on the two opposite sides but at a distance closer to each other, the overlap between the reflections from the damage and the weld was captured correctly. This is highlighted in Figure 9(c) to (d). In Figure 9(c), the crack was located at 154.75 m to the right of the transducer, while weld A is also to the right of the transducer at 162 m. From Figure 9(c), the reflections from these two discontinuities caused an overlap. The VAE model can approximate the overlap of two reflections since the prediction is comparable to the simulated target. In Figure 9(d), the overlap was caused by discontinuities located on two opposite sides of the transducer. For this set-up, the crack was located at 253 m from the transducer on the left side, and the second discontinuity involved in the said overlap is weld C. This weld is located approximately 320 m to the right of the transducer. Although the locations of the two discontinuities are distinct, their reflections overlapped with each other as the reflection from the crack spread out over a large distance, and the VAE result had a reconstruction loss of 0.0093. In the section 'VAE comparison to superposition for damage fusion', we show that when superposition is employed to generate synthetic inspection signals, it fails to approximate parts of these overlapping reflections.

The last feature worth noting in the response signals that the VAE could predict is the double reflection that resulted from the inclusion of damage in the rail set-up. In most cases, this double reflection is not evident as it is masked by other reflections or coherent noise.

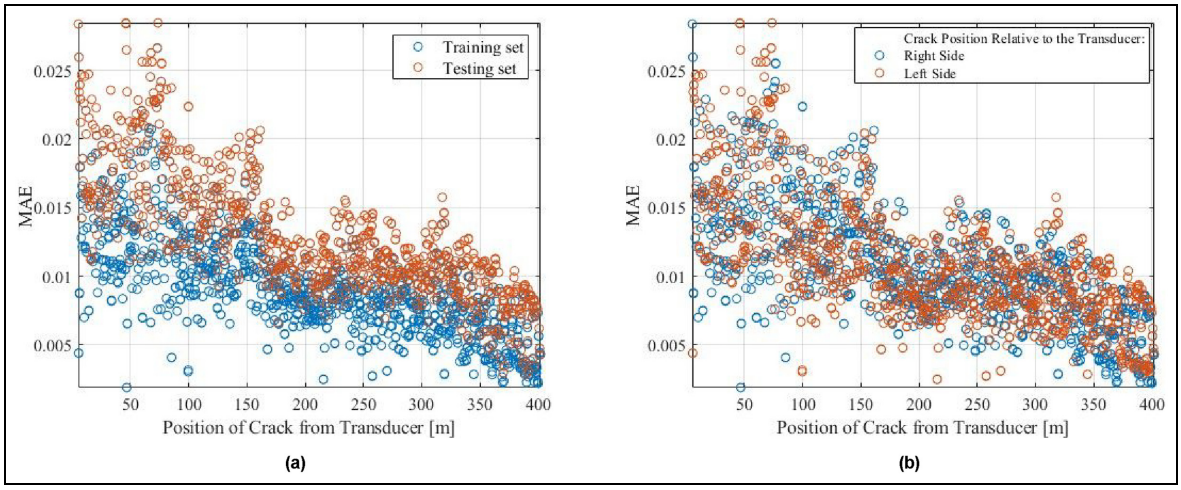


Figure 10. The VAE MAE for reconstructed response signals in (a) the training and testing sets and (b) samples located to the right and left of the transducer.
VAE: variational autoencoder; MAE: mean absolute error.

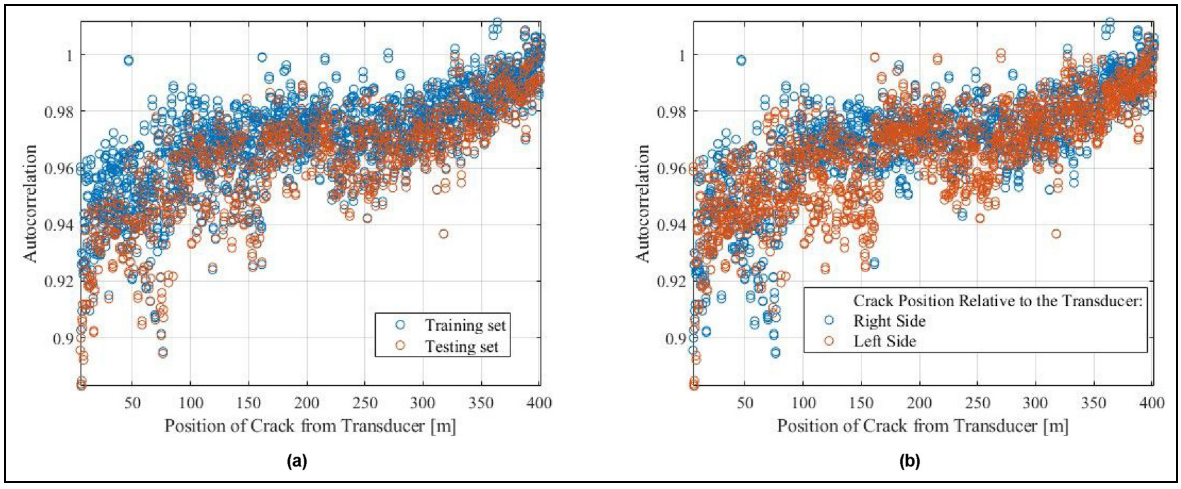


Figure 11. The autocorrelation coefficients for reconstructed response signals in (a) the training and testing sets and (b) samples located to the right and left of the transducer.

For example, when damage is located at 41.375 m to the left of the transducer, it is expected that a double reflection between this damage and weld A would be approximately 119 m. However, the double reflection is not visible in Figure 9(a) due to the noise caused by multiple interacting modes in this region. On the other hand, in addition to the crack reflection at 108.5 m in Figure 9(b), we further notice another reflection reconstructed by the VAE at approximately 273 m. This double reflection is the result of reverberation between the crack and weld B, which the VAE can predict. This result highlights the importance of accounting for multiple reflections by using the physics-based modelling procedure in the study by Ramatlo et al.[15] to simulate the training and testing data.

The MAE, autocorrelation coefficients and maximum absolute errors for the reconstructed response signals in the training and testing sets are plotted in Figures 10 to 12, respectively, and were calculated using the following equations:

$$\text{MAE} = \frac{\sum|\tilde{\boldsymbol{y}} - \boldsymbol{y}|}{m} \tag{3}$$

$$\text{Autocorrelation} = \frac{\sum(\boldsymbol{y} - \mu_y)(\tilde{\boldsymbol{y}} - \mu_y)}{\sum(\boldsymbol{y} - \mu_y)^2} \tag{4}$$

$$\text{Maximum absolute error} = \max\{|\tilde{\boldsymbol{y}} - \boldsymbol{y}|\} \tag{5}$$

where μ_y is the mean of the simulated target y.

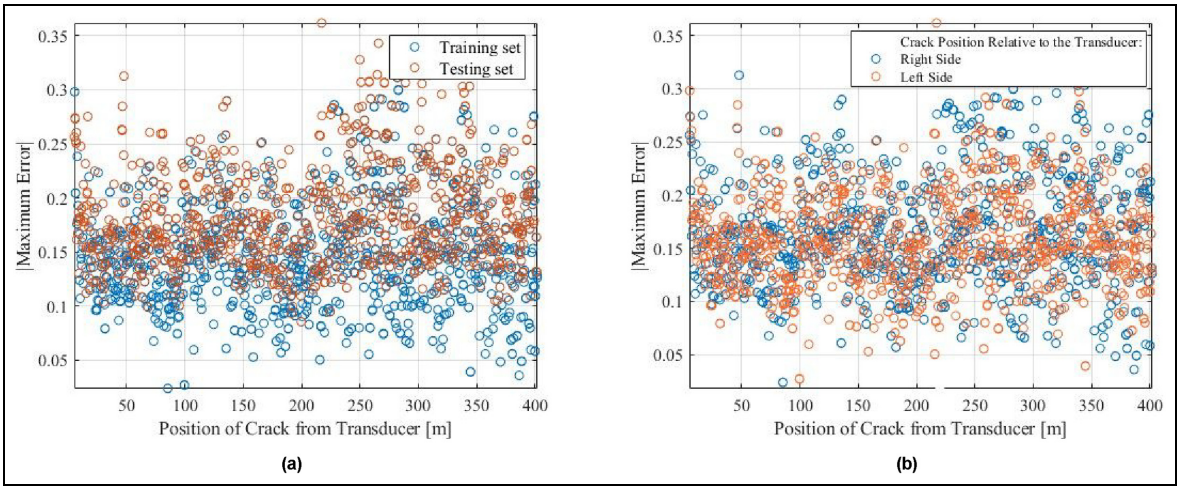


Figure 12. The VAE maximum absolute error for reconstructed response signals in (a) the training and testing sets and (b) samples located to the right and left of the transducer.
VAE: variational autoencoder.

Figure 10(a) shows the highest MAE of 0.028 when damage is located very close to the transducer, within 100 m. This highest error was obtained for samples in the testing set, as highlighted in Figure 10(a). The errors indicate a notable decreasing trend for an increase in the damage location as the reflections are more complex for short-range propagation. The same behaviour is also observed in Figure 11(a), where a lower correlation is obtained when damage is located very close to the transducer, within 100 m, and where the autocorrelation coefficients improve with an increase in the damage location. When propagating over long distances, the energy of the reflections from the damage decays due to attenuation and dispersion. This leads to highly dispersive modes contributing less to the response as their contribution gets masked by the noise in the signal. In Figure 12(a), the highest error of 0.36 was obtained in the maximum error for a testing sample where the damage was approximately 217 m from the transducer. The maximum absolute error indicates that the training set performed better than the testing set, though it does not show a distinct trend associated with the damage location. The error values are up to 36%. However, it is noted that these maximum errors could have been located within insignificant regions within the signals, outside the damage zone and where noise is dominant.

In the section 'VAE comparison to superposition for damage fusion', the maximum absolute errors for the response signals in Figure 9(a) to (e) will be computed within the highlighted regions where damage reflections are significant. Overall, Figures 10 to 12 highlight that the training set performed slightly better than the testing set. This is because the VAE had seen the training set several times during training, while the testing set was fed to the VAE only after the model had been trained. The MAE, autocorrelation coefficients and maximum absolute errors for the reconstructed response signals in the training and testing sets are also compared in Figure 13(a) to (c) using bar plots. A total count of 2220 signals was used in the training set, and 953 signals were used for testing the VAE model. It is observed that the distributions of the training and testing sets have similar trends for both three cases, indicating that the model is generalizing well to unseen data.

In Figures 10 to 12(b), the errors are plotted again, but now to investigate the influence of inserting damage on either the left or right side of the transducer. The results indicate that the crack position relative to the transducer does not influence the performance of the VAE.

VAE comparison to superposition for damage fusion

Next, we compare the VAE results with the results obtained by superposing[8] the virtual experiment damage-free baseline signal with the damage signatures to highlight further the benefit of using machine learning-based algorithms for synthetic data generation.

Firstly, we highlight the similarities and differences between the two methods. Both the VAE and superposition methods employ a baseline signal containing reflections from the four welds. We notice that the VAE was able to capture the complex features in the reflected signals. This baseline signal is a virtual experimental signal simulated using the physics-based model explained in section 'Simulated virtual experimental baseline signal for a damage-free rail' and optimised to reflect the field experiment closely. The baseline signal, therefore, captures all the significant complex behaviour of guided wave propagation in a damage-free

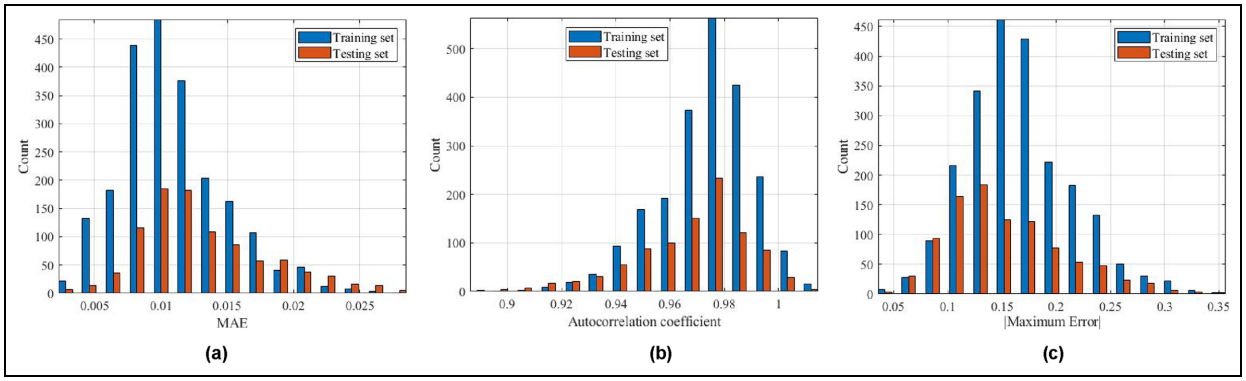


Figure 13. Bar plots comparing the (a) MAE, (b) autocorrelation coefficients and (c) maximum absolute errors for the training and testing data.
MAE: mean absolute error.

rail. The same baseline signal was used in both the VAE model and the superposition method.

Further to employing a virtual experimental baseline signal, the VAE model also employed a virtual experimental target signal, simulated using the same physics-based model validated in reference.[15] In addition to capturing the complex interaction of reflections from multiple discontinuities, the ability of this model to account for complex reverberations makes it even more attractive. Though for guided wave inspection in the head of the rail, these reverberations are not as significant as when the inspection is in the web of a rail. This is because the mode with energy in the head of the rail (Figure 3) reflects less strongly from welds, and these reverberating reflections are thus mostly masked by noise and other dominant reflections. This study demonstrated a case of inspection in the head of a rail where a double reflection can be considered significant, Figure 9(b). All the data (baseline and target signals) used to train the VAE model was obtained using the physics-based model in the study by Ramatlo et al.[15] The target signals containing damage signatures are thus believed to be more representative of a hypothetical scenario of an operational railway line with a head defect.

The second set of input data required for the superposition method was defect signatures for different distance positions. These defect signals were modelled using the physics-based modelling framework validated in the study by Ramatlo et al.[11] This modelling framework computes only direct reflections from discontinuities. The model in the study by Ramatlo et al.[15] was an extension of the modelling framework in the study by Ramatlo et al.[11] to account for reverberating reflections. The simulated damage signals obtained from the procedure in the study by Ramatlo et al.[11] are also believed to be a good approximation of a hypothetical scenario of an operational railway line with a head defect. However, the only drawback is that double reflections were not accounted for. Damage signatures simulated using the procedure in the study by Ramatlo et al.[11] for a crack located at different distances are plotted in Figure 14 and Figure A1 in Appendix A. These response signals contain direct reflections from only the damage geometry, as no welds were included in the model. The superposition results were obtained by adding the baseline signal and the damage signatures in the time domain. The resultants were then post-processed by applying dispersion compensation and scaling the amplitudes accordingly.

The VAE results are compared to those obtained from superposition in Figure 15. The location of the crack in the rail for each signal is highlighted using a dot marker and a vertical dotted line. The area shaded in green indicates the region over which the energy reflected from the damage in the target and VAE reconstructed response signals is significant and was determined according to Figure 9. The area shaded in orange indicates the region over which the energy reflected from the damage in the superposition reconstructed response signals was over-estimated.

Firstly, we notice that the superposition results captured the complex patterns in the response signals as expected. This is because the two physics-based models used to generate the data accounted for almost all features of the virtual experimental setup. However, the second expectation observed from the superposition results is that the multiple reflections caused by the crack defect were not reconstructed since the physics-based model in the study by Ramatlo et al.[11] does not account for reverberating reflections. This observation is evident in Figure 15(b). Last, in regions around the damage reflections, superposition resulted in an over-estimate of some parts of the resultant reflections.

The absolute errors in the reconstructed signals were calculated between the VAE results and the target signals and between the superposition results and the

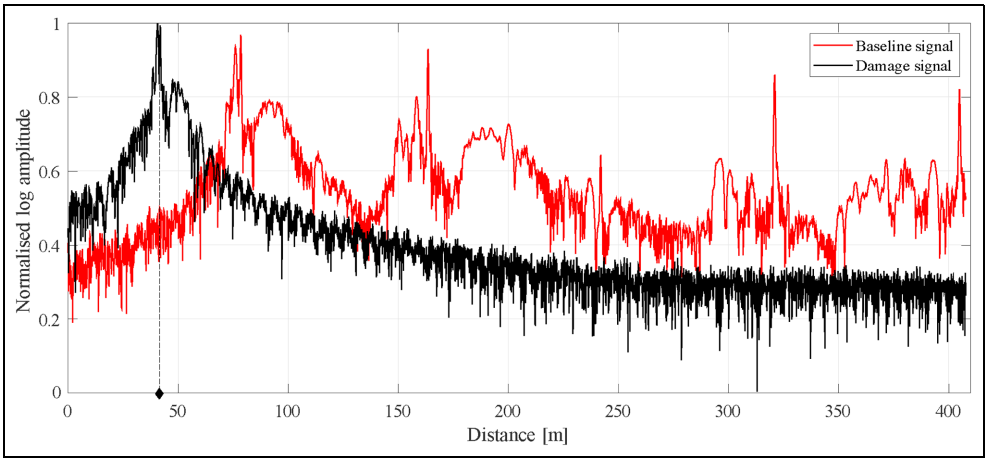


Figure 14. Damage signatures simulated using the procedure in the study by Ramatlo et al.[11] for a crack located at (•···) 41.375 m from the transducer on the left side.

target signals. The results for the five cases considered in Figure 15(a) to (e) are plotted in Figure 16(a) to (e). The absolute errors in the signals demonstrate that VAE is almost consistently better than superposition regarding the worst error.

The MAE, autocorrelation coefficients and maximum absolute errors for the results in Figure 15 are presented in Tables 2 to 4, respectively.

In Table 2, the MAE was computed using all the points in each signal, and the reconstruction performance of the VAE compared to superposition was evaluated according to $\frac{\text{Superposition MAE} - \text{VAE MAE}}{\text{Superposition MAE}} \times 100\%$. As observed from the plot in Figure 10, Table 2 also highlights that the VAE MAE tends to decrease with an increasing crack position from the transducer, though the error when the crack was fused at 154.875 m is slightly higher than when it was located at 108.5 m. The same is observed in the superposition MAE results. The performance evaluation for the five cases considered shows that the VAE performed far better than superposition when the crack was fused at 108.5 m. This is because superposition failed to reconstruct the double reflection highlighted in Figure 15(b). For this case, the VAE performed 63.07% better than superposition. On average, the VAE reconstructed results were 53.41% better than the superposition reconstructed results.

The autocorrelation coefficients in Table 3 indicate that both the VAE and superposition reconstructed results are very close to the target signals. It is noted that similar to what the MAE suggested, the VAE model performed better than superposition in all five cases, though the improvement in terms of autocorrelation was below 4%.

In Table 4, the maximum absolute errors were first calculated based on regions where the damage reflections are significant, as highlighted in green and orange in Figure 15. The distances where the maximum error for each response signal was read off are also presented in the table. The results show that the maximum absolute errors for the VAE and superposition results were read off at different distances in the signals within the highlighted regions. The results will thus not give a reliable performance indication. However, it is observed that superposition yielded larger errors except for the last case when damage was fused at 379.25 m.

The maximum absolute errors were also calculated at the peaks of the damage reflections, and the results are presented in Table 5. When it comes to estimating the peaks of the damage reflection, superposition performed better than the VAE in three out of the five cases evaluated.

Overall, the results suggest that although the VAE slightly underpredicts the peak amplitudes of the damage reflections, it outperforms the superposition method when capturing the more complex features of the response signals. The authors would like to highlight that each time the VAE model is run, it generates a novel reconstructed output due to aleatoric uncertainty in the model. The randomness introduced in the reconstructed output results from variability in the underlying variables and is representative of a real-life scenario where the measured response signals from an operational railway line will contain aleatoric uncertainty due to unknowns that differ each time we run the same experiment. On the other hand, superposition will yield only one solution as it does not account for aleatoric uncertainty in the data for a fixed input signal.

Field experiment

In this section, we use the trained VAE model to fuse the simulated damage signature to an experimental signal collected from an operational rail. The VAE model was

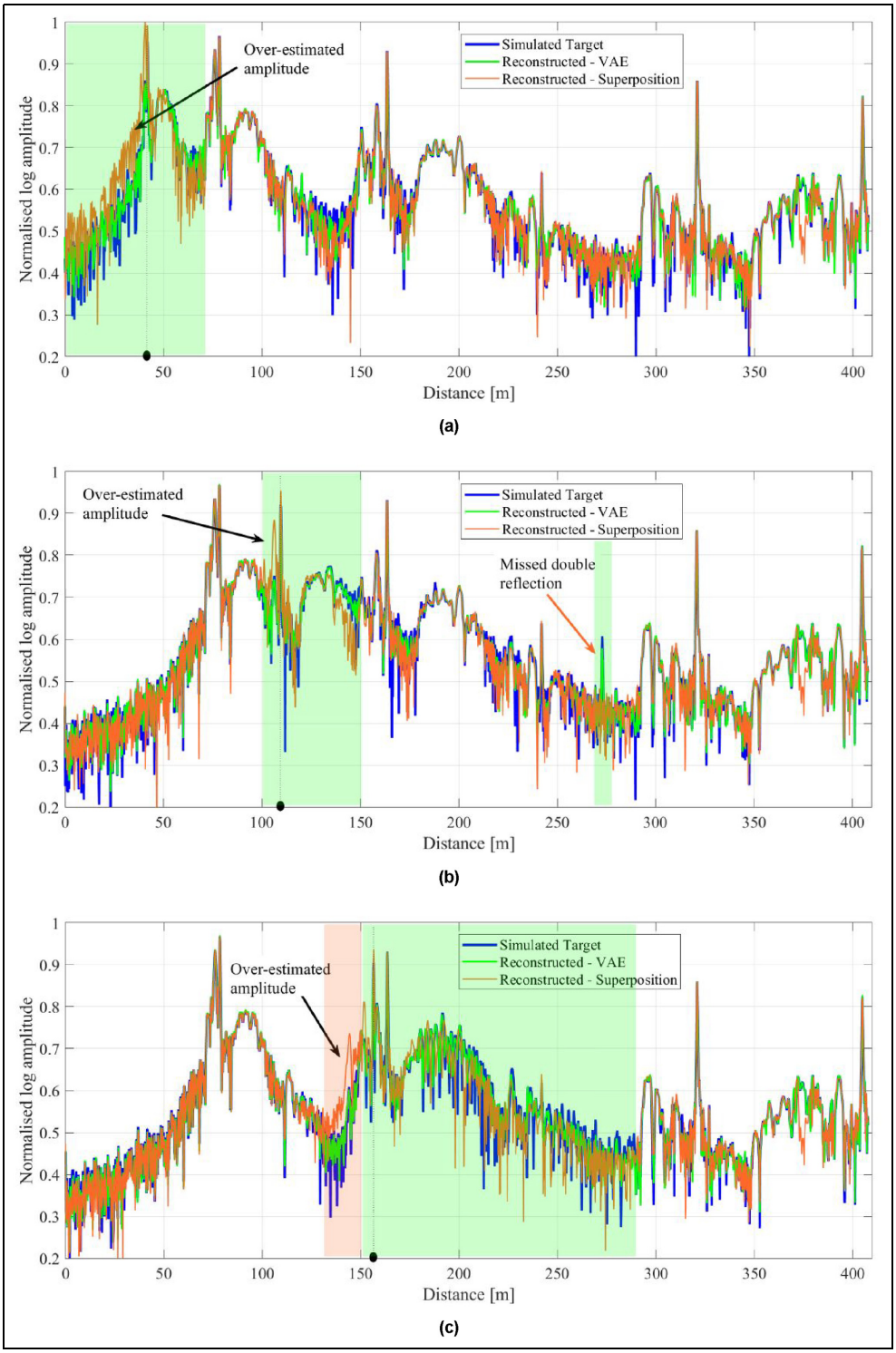


Figure 15. (continued)

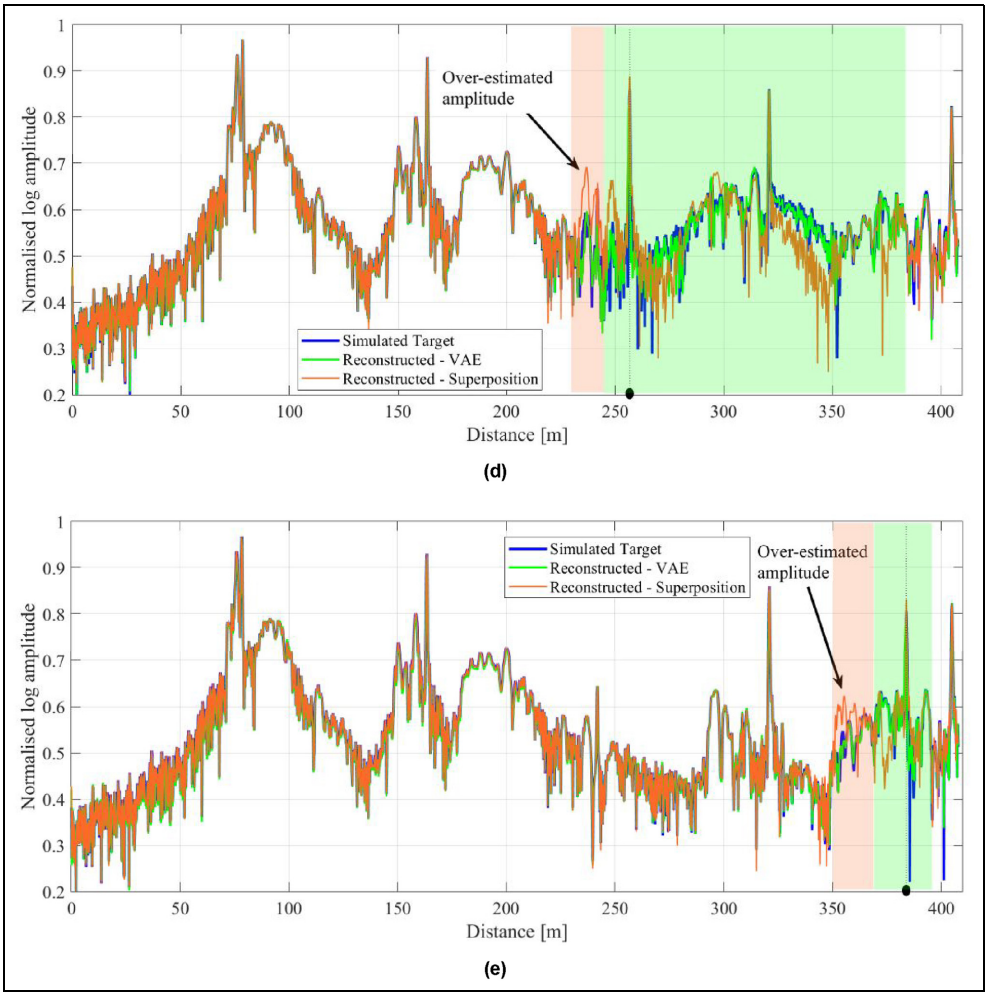


Figure 15. Response signals reconstructed from a VAE compared to superposition and the original target response for a crack located at (•···) (a) 41.375 m from the transducer on the left side, (b) 108.5 m from the transducer on the right side, (c) 154.8750 m from the transducer on the right side, and (d) 253 m from the transducer on the left side and (e) 379.25 m from the transducer on the right side. The area shaded in green and orange indicates the region over which the energy reflected from the damage in the superposition reconstructed response signals is significant.
VAE: variational autoencoder.

first trained with virtual experimental data. The inspection signal generated when an experimental field signal was given to the trained VAE model as a baseline input is plotted in Figure 17. The results are compared to those obtained from superposition. The damage reflection is fused at 93.625 m from the transducer on the right side. The results show that the reconstructed inspection resembles those obtained when the simulated virtual experimental signals were used as baselines instead of experimental baselines from the field. This is because the field experiment contains noise. In regions where the signal is insignificant, this noise dominates the response. However, when there is a reflected signal, the reconstruction is much better, and these are the important regions for monitoring purposes. If the regions without the reflection peaks are of interest, the VAE results could be improved by enhancing the training data with aleatoric uncertainty in the form of representative noise in the measured response. The superposition result is more representative of the baseline signal. However, superposition cannot reconstruct the two double reflections around 250 m. In the experimental signals collected from the field, these double reflections would not be visible due to noise. The VAE can uncover and highlight features that might not be apparent in the field measurements.

Conclusions

This paper attempts to develop a machine learning-based method to generate synthetic inspection data for

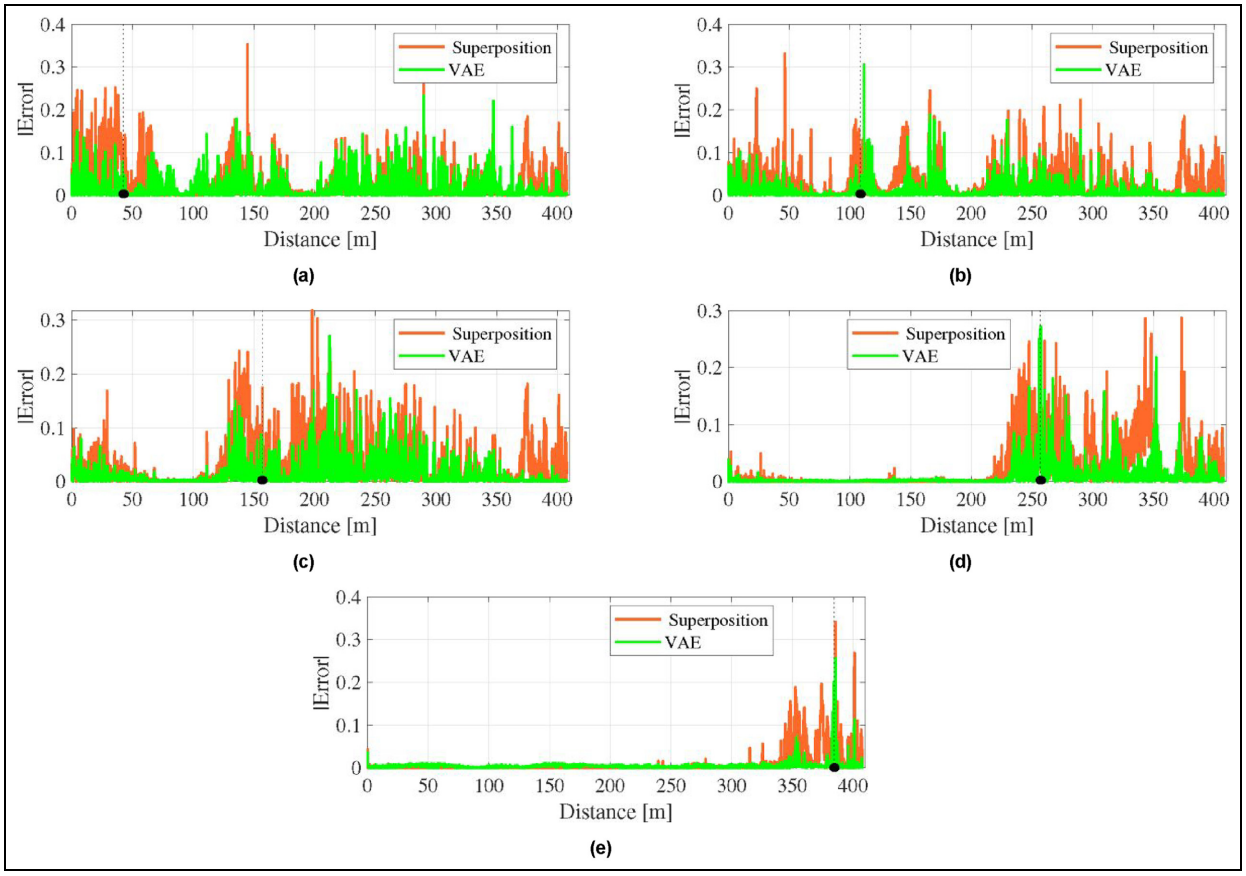


Figure 16. Absolute errors obtained from the VAE and superposition results for a crack located at (•···) (a) 41.375 m from the transducer on the left side, (b) 108.5 m from the transducer on the right side, (c) 154.8750 m from the transducer on the right side, (d) 253 m from the transducer on the left side and (e) 379.25 m from the transducer on the right side.
VAE: variational autoencoder.

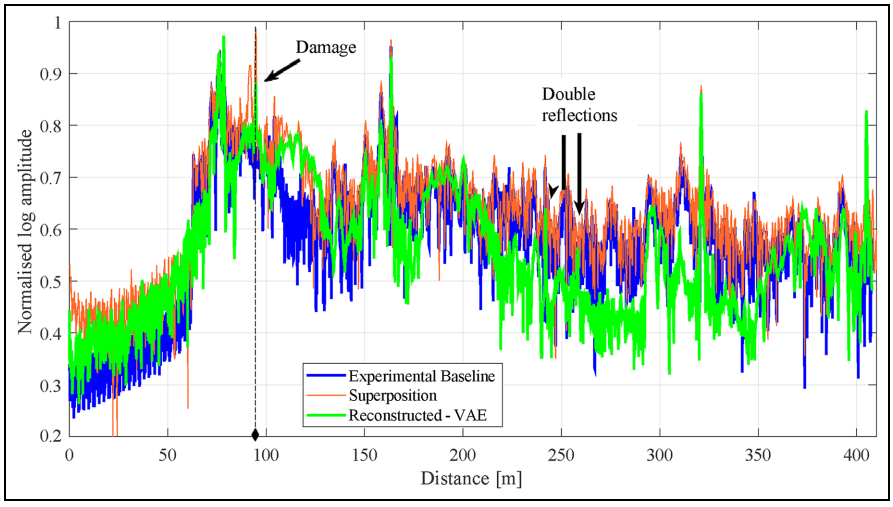


Figure 17. Inspection signal generated from a field experiment and a VAE model trained with virtual experimental signals.
VAE: variational autoencoder.

Table 2. The MAE and VAE performance compared to superposition.

Position of the crack from the transducer (m)	MAE		VAE performance evaluation compared to superposition (%) $= \frac{\text{Superposition MAE} - \text{VAE MAE}}{\text{Superposition MAE}} \times 100\%$
	VAE	Superposition	
41.375	0.0168	0.0324	48.21
108.5	0.0104	0.0282	63.07
154.875	0.0130	0.0317	58.82
253	0.0093	0.0232	59.85
379.25	0.0052	0.0083	37.08

VAE: variational autoencoder; MAE: mean absolute error.

Table 3. The autocorrelation coefficient and VAE performance compared to superposition.

Position of the crack from the transducer (m)	Autocorrelation		VAE performance evaluation compared to superposition (%) $= \frac{\text{VAE corr} - \text{Superposition corr}}{\text{VAE corr}} \times 100\%$
	VAE	Superposition	
41.375	0.9315	0.9160	1.6594
108.5	0.9709	0.9648	0.6322
154.875	0.9780	0.9463	3.2390
253	0.9742	0.9392	3.5841
379.25	0.9869	0.9740	1.3071

VAE: variational autoencoder.

Table 4. The maximum absolute errors are evaluated within the highlighted damage regions in Figure 15.

Position of the crack from the transducer (m)	Maximum error (within the damage region)			
	VAE		Superposition	
	Error	Distance (m)	Error	Distance (m)
41.375	0.0990	65.7895	0.1940	58.8195
108.5	0.0690	105.0591	0.2108	273.0178
154.875	0.0919	286.7876	0.3183	198.0484
253	0.2716	257.0379	0.2870	373.1469
379.25	0.2006	384.1968	0.1966	373.9969

VAE: variational autoencoder.

Table 5. The maximum absolute errors ($|E_{max}|$) are evaluated at the damage reflection peaks.

Position of the crack from the transducer (m)	Maximum error (at the peak)		VAE performance evaluation compared to superposition (%) $= \frac{\text{Superposition } \lvert E_{max}\rvert - \text{VAE } \lvert E_{max}\rvert}{\text{Superposition } \lvert E_{max}\rvert} \times 100\%$
	VAE	Superposition	
41.375	0.0530	0.0396	−33.84 (worse)
108.5	0.0080	0.0325	75.39
154.875	0.0168	0.0318	47.17
253	0.0330	0.0291	−13.40 (worse)
379.25	0.0275	0.0254	−8.27 (worse)

GWU in welded railway lines. Synthetic data generation continues to become important because experimental databases for different scenarios of damage growth are still lacking in applications such as the rail industry, and laboratory experiments are not possible. Large experimental databases are required in the development of reliable monitoring systems. To cater to this challenge, a VAE model for generating synthetic data containing damage signatures at arbitrary positions along the length of a rail track was developed. The VAE model was given a damage-free baseline signal and trained to reconstruct an inspection signal with damage by adding a damage signature on either side of the transducer at the specified distance. The training data was produced from a physics-based model that computes virtual experimental response signals using the SAFE and finite element procedures.

The VAE reconstructed response signals containing damage signatures were nearly identical to the original target signals simulated using the physics-based model. The VAE was able to capture the complex features in the signals resulting from the interaction of multiple propagating modes in a multi-discontinuous waveguide. These complex features included reverberating reflections that resulted from introducing a crack defect in the set-up of the rail model. In railway lines, multiple reflections become even more important when the inspection is in the web of a rail as the employed mode shape reflects strongly from welds. In such a case, the ability to capture all significant reverberations and distinguish the ones associated with growing damage from those not, would be beneficial when designing monitoring systems. For inspection in the head of a rail, reverberating reflections are mostly insignificant. However, this paper demonstrated that when damage occurs in certain locations in the rail, noticeable double reflections may be present in the response signals. In scenarios where the crack was closer to a weld or the two were located on opposite sides but at a distance closer to each other, the VAE gave a good approximation of the overlap between the reflections from the damage and the weld. Overall, the VAE model successfully generated synthetic inspection data by fusing reflections from welds with the reflection from a crack model at an arbitrary distance from the transducer on either the right or left side. In some cases, the VAE struggled to predict a good approximation of the amplitude for the head mode. The VAE results were further compared to results obtained from the superposition method. The superposition results could not capture the double reflection caused by the crack. In regions around the damage reflections, some parts of the resultant reflections were over-estimated by superposition.

In conclusion, this study highlighted the benefit of using a VAE to generate synthetic data with damage signatures as opposed to using superposition to fuse the damage-free response containing reflections from welds with a damage signature. However, when the measurement collected from the field was employed as a baseline signal, the reconstruction ability was not good due to the influence of the random noise. This is because the training data used in the VAE did not include aleatoric and epistemic uncertainties, as the focus of the study was on the ability of the VAE to capture the relevant physics, such as reflections, as opposed to some superposition of noise. If the researcher shifts their focus from capturing just the appropriate physics to capturing uncertainties in the data, noise could be included in the training data, or the VAE could be trained to add noise and EOC variations representative of uncertainties in practical applications. The inclusion of noise will be dealt with as part of future work when the focus will be on obtaining a better agreement for sections of rail where there are no defects. Overall, the results show that it is possible to use a VAE to generate realistic inspection data for unavailable damage scenarios. In future, the predicted response data containing damage features will be employed to develop and investigate damage identification methods. In that case, a thorough analysis of the extracted damage features will be required. The generated data can then be used to compute ROC curves to predict the performance of a monitoring system for damage detection. Future work will also explore alternative VAE architectures and further investigate how different parameters affect its prediction ability based on various representations of the target signals.

Acknowledgement

The authors gratefully acknowledge the contribution of Dr Craig Long who helped with the physics-based modelling procedure.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship and/or publication of this article.

ORCID iDs

Dineo A Ramatlo https://orcid.org/0000-0003-4874-9494
Philip W Loveday https://orcid.org/0000-0002-0515-3204

References

1. Cawley P, Cegla F and Galvagni A. Guided waves for NDT and permanently-installed monitoring. *Insight Non-Destructive Test Cond Monit* 2012; 54(11): 594–601.

2. Leinov E, Lowe MJ and Cawley P. Investigation of guided wave propagation and attenuation in pipe buried in sand. *J Sound Vib* 2015; 347: 96–114.
3. Loveday PW, Long CS and Ramatlo DA. Ultrasonic guided wave monitoring of an operational rail track. *Struct Heal Monit* 2020; 19(6): 1666–1684.
4. Burger F and Loveday P. Ultrasonic broken rail detector and rail condition monitor technology. In: *Proceedings of the 11th International Heavy Haul Association Conference (IHHA 2017)*, Cape Town, South Africa, 2–6 September 2017, pp. 275–280. Red Hook, NY: Curran Associates, Inc.
5. Burger FA, Loveday PW and Long CS. Large scale implementation of guided wave based broken rail monitoring. In: Proceedings of the 41st Annual Review of Progress in Quantitative Nondestructive Evaluation, 2014, Boise, Idaho, 20–25 July 2014, vol. 34, pp. 771–776. Melville, NY: AIP Publishing.
6. Dobson J and Cawley P. Independent component analysis for improved defect detection in guided wave monitoring. *Proc IEEE* 2016; 104(8): 1620–1631.
7. Liu C, Dobson J and Cawley P. Practical ultrasonic damage monitoring on pipelines using component analysis methods. In *19th World Conference on Non-Destructive Testing*, 2016, Munich, Germany, 13–17 June 2016, NY 12571 USA, vol. 5, pp. 4093–4100. Red Hook, NY: Curran Associates, Inc.
8. Liu C, Dobson J and Cawley P. Efficient generation of receiver operating characteristics for the evaluation of damage detection in practical structural health monitoring applications. *Proc R Soc A Math Phys Eng Sci* 2017; 473(2199): 1–26.
9. Mariani S, Heinlein S and Cawley P. Compensation for temperature-dependent phase and velocity of guided wave signals in baseline subtraction for structural health monitoring. *Struct Heal Monit* 2020; 19(1): 26–47.
10. Heinlein S, Cawley P and Vogt T. Validation of a procedure for the evaluation of the performance of an installed structural health monitoring system. *Struct Heal Monit* 2019; 18(5–6): 1557–1568.
11. Ramatlo DA, Long CS, Loveday PW, et al. A modelling framework for simulation of ultrasonic guided wave-based inspection of welded rail tracks. *Ultrasonics* 2020; 108(106215): 1–18.
12. Long CS, Loveday PW, Ramatlo DA, et al. Numerical verification of an efficient coupled SAFE-3D FE analysis for guided wave ultrasound excitation. *Finite Elem Anal Des* 2018; 149: 45–56.
13. Benmeddour F, Treysse'de F and Laguerre L. Numerical modeling of guided wave interaction with non-axisymmetric cracks in elastic cylinders. *Int J Solids Struct* 2011; 48(5): 764–774.
14. Long CS and Loveday PW. Validation of hybrid SAFE-FE guided wave scattering predictions in rail. *AIP Conf Proc* 2015; 1650: 703–712.
15. Ramatlo DA, Long CS, Loveday PW, et al. Physics-based modelling and simulation of reverberating reflections in ultrasonic guided wave inspections applied to welded rail tracks. *J Sound Vib* 2022; 530(116914): 1–20.
16. Cantero-Chinchilla S, Wilcox PD and Croxford AJ. Deep learning in automated ultrasonic NDE – developments, axioms and opportunities. *NDT E Int* 2022; 131(102703): 1–19.
17. Khurjekar ID and Harley JB. Uncertainty aware deep neural network for multistatic localization with application to ultrasonic structural health monitoring. *arXiv Prepr arXiv200706814* 2020; : 1–10.
18. Khurjekar ID and Harley JB. Sim-to-real localization: environment resilient deep ensemble learning for guided wave damage localization. *J Acoust Soc Am* 2022; 151(2): 1325–1336.
19. Zhang B, Hong X and Liu Y. Deep convolutional neural network probability imaging for plate structural health monitoring using guided waves. *IEEE Trans Instrum Meas* 2021; 70: 2510610.
20. Wu J, Xu X, Liu C, et al. Lamb wave-based damage detection of composite structures using deep convolutional neural network and continuous wavelet transform. *Compos Struct* 2021; 276(114590): 1–11.
21. Rautela M and Gopalakrishnan S. Ultrasonic guided wave based structural damage detection and localization using model assisted convolutional and recurrent neural networks. *Expert Syst Appl* 2021; 167(114189): 1–14.
22. Pyle RJ, Hughes RR, Ali AAS, et al. Uncertainty quantification for deep learning in ultrasonic crack characterization. *IEEE Trans Ultrason Ferroelectr Freq Control* 2022; 69(7): 2339–2351.
23. Jin XB, Gong WT, Kong JL, et al. PFVAE: a planar flow-based variational auto-encoder prediction model for time series data. *Mathematics* 2022; 10(4): 610.
24. Mahajan H and Banerjee S. A machine learning framework for guided wave-based damage detection of rail head using surface-bonded piezo-electric wafer transducers. *Mach Learn with Appl* 2022; 7(1–14): 100216.
25. Khurjekar ID and Harley JB. Closing the sim-to-real gap in guided wave damage detection with adversarial training of variational auto-encoders. *ICASSP, IEEE Int Conf Acoust Speech Signal Process – Proc* 2022; 2022: 3823–3827.
26. Diakoloukas V, Lygerakis F, Lagoudakis MG, et al. Variational denoising autoencoders and least-squares policy iteration for statistical dialogue managers. *IEEE Signal Process Lett* 2020; 27: 960–964.
27. Baronian V, Lhémery A and Jezzine K. Hybrid SAFE/FE simulation of inspections of elastic waveguides containing several local discontinuities or defects. *AIP Conf Proc* 2011; 1335(June): 183–190.
28. Wilcox PD. A rapid signal processing technique to remove the effect of dispersion from guided wave signals. *IEEE Trans Ultrason Ferroelectr Freq Control* 2003; 50(4): 419–427.
29. Long CS and Loveday PW. Prediction of guided wave scattering by defects in rails using numerical modelling. In: Proceedings of the 40th Annual Review of Progress in Quantitative Nondestructive Evaluation, 2013, Baltimore, Maryland, USA, 21–26 July 2013, vol. 1581, pp. 240–247. Melville, NY: AIP Publishing.

Appendix A

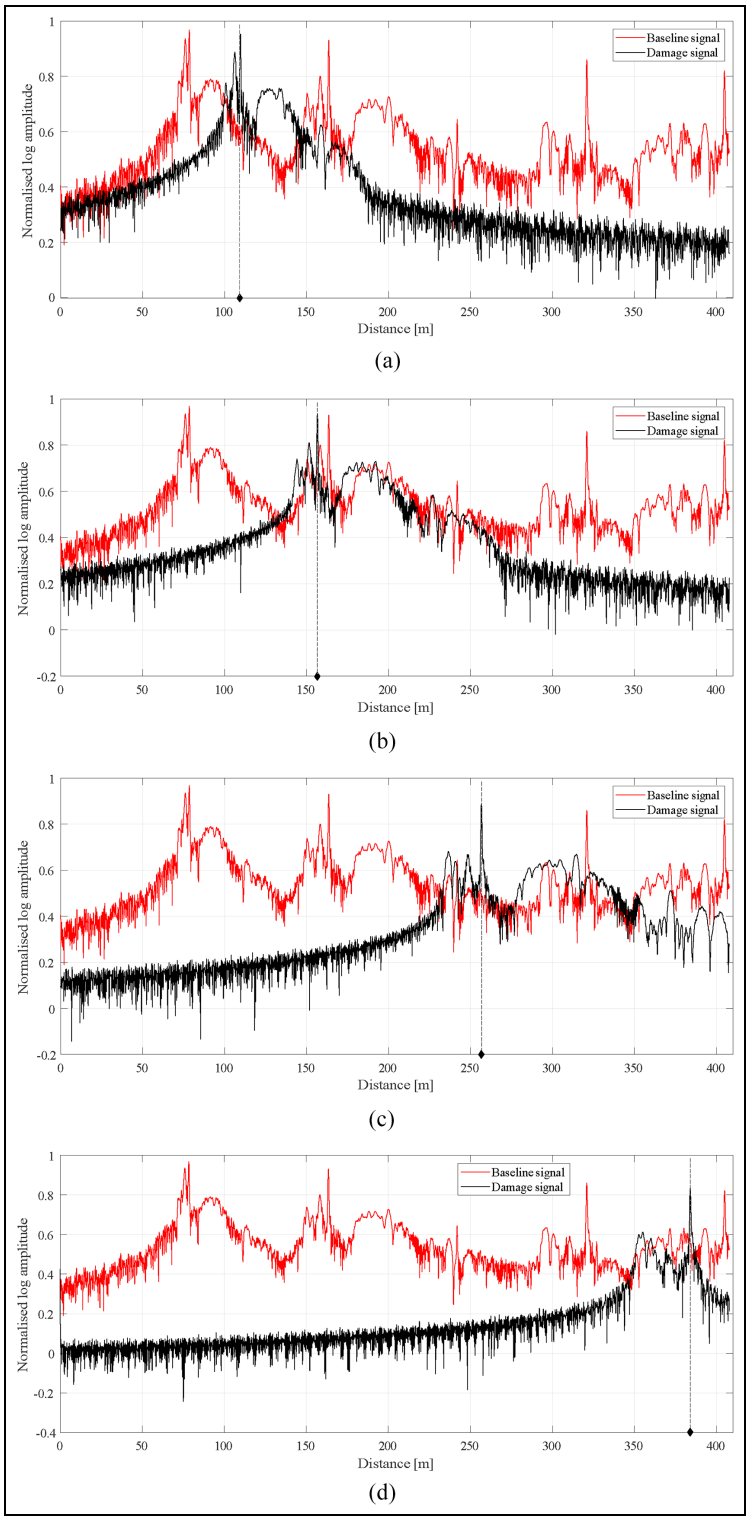


Figure A1. Damage signatures simulated using the procedure in Ramatlo et al.[11] for a crack located at (•···) (a) 108.5 m from the transducer on the right side, (b) 154.8750 m from the transducer on the right side, (c) 253 m from the transducer on the left side and (d) 379.25 m from the transducer on the right side.