A kNN algorithm for locating and quantifying stiffness loss in a bridge from the forced vibration due to a truck crossing at low speed

A kNN algorithm for locating and quantifying stiffness loss in a bridge from the forced vibration due to a truck crossing at low speed

Kun Feng*, Arturo González, Miguel Casero

School of Civil Engineering, University College Dublin, Dublin 4, Ireland

ARTICLE INFO

Article history:
Received 25 August 2020
Received in revised form 25 November 2020
Accepted 29 December 2020
Available online 7 January 2021

Keywords:
Frequency
Short-time Fourier Transform
Structural Health Monitoring
Damage Detection
k-Nearest Neighbours

ABSTRACT

This paper proposes a k-Nearest Neighbours (kNN) algorithm for locating and quantifying bridge damage based on the time-varying forced frequencies due to a moving truck. Eigenvalue analysis of a simplified vehicle-bridge coupled system, consisting of a three-axle rigid truck model and a simply supported finite element beam model, shows how the eigenfrequencies of the coupled system vary with the locations of the vehicle and with the damage represented by a stiffness loss. The computational efficiency of eigenvalue analysis is exploited to generate a vast sample of patterns for training a kNN algorithm. In the field, acceleration due to the crossing of a test vehicle would be measured and analysed using a time–frequency signal processing tool to obtain the instantaneous frequencies. The crossing must take place at a low speed to achieve sufficiently high resolution and to minimise deviations from the eigenvalue solution. Then, the kNN algorithm searches for the patterns of forced eigenfrequencies that are closest to the on-site instantaneous frequencies to determine the location and severity of the damage. For theoretical testing purposes, field measurements are simulated here using coupled equations of motion and dynamic transient analysis.

© 2021 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

1. Introduction

Many bridges built in the middle of the last century are getting close to the end of their design life and giving signs of deterioration. The latter has been aggravated by loads higher than foreseen at the design stage [1]. Therefore, proactive and effective maintenance is key to ensure the safe operation of Europe’s ageing transport infrastructure during its remaining life cycle. Structural Health Monitoring (SHM) methods aim to meet this need via two approaches: (i) response-based methods, and (ii) model-based methods [2].

Response-based methods, also called model-free methods, rely on the measured response from the structure only. Hence, they do not require complicated structural models, which are often difficult to calibrate on site. Amongst these methods, the extraction of mode shapes from measured accelerations is a popular technique for locating damage [3]. However, a requirement for obtaining accurate mode shapes and mode shape curvatures is the installation of a high number of sensors covering a significant proportion of the structure. Accelerometers are commonly used, although there are alternatives such as the recording and processing of video measurements proposed by Yang et al. [4], who extract three high-resolution mode shapes.
from a structure excited by an impact hammer. Sun and Chang [5] report that mode shapes and their derivatives usually exhibit a larger statistical variability than natural frequencies. Meanwhile, damage severity may prove difficult to estimate using only response-based methods due to missing information about the bridge structure.

Compared to response-based methods, model-based methods are generally better suited for both locating and quantifying damage. A detailed numerical model establishes an early statistical baseline of the healthy structure, and the model is subsequently updated to match the responses measured at different stages in time. Monitored responses commonly include natural frequencies, given that their acquisition through accelerometers is easy and energy-efficient, i.e., one sensor may be sufficient to extract the frequency range of interest. Guan et al. [6] propose an algorithm to locate and estimate the severity of the damage from shifts in the natural frequency and changes in the mode shapes. The algorithm is tested on a composite bridge and results show that it is possible to locate and quantify the damage using a minimum amount of data from a post-damaged stage. Koh and Dyke [7] locate damage in a full-scale cable-stayed bridge model by observing the level of correlation between the variations in the measured and analytically synthesized natural frequencies. Lederman et al. [8] use signal processing and machine learning approaches to analyse the forced vibration responses collected from a vehicle-bridge interaction (VBI) model. The magnitudes of the natural frequency spectrum obtained by principal component analysis are successfully employed to diagnose the location and severity of damage in a laboratory bridge model.

The natural frequency of a bridge does not only change with the magnitude of a moving load but also with its location. Cantero and OBrien [9] demonstrate that the variations in natural frequencies greatly depend on the vehicle-to-structure frequency and mass ratios. Chang et al. [10] theoretically and experimentally verify that the frequencies of the bridge alone may differ due to the presence of a vehicle parked on the bridge. Cantero et al. [11] measure the evolution of the frequency shifts and changes in mode shape due to a truck parked in different positions of a steel girder bridge using the Continuous Wavelet Transform (CWT). If the presence of a vehicle can alter the natural frequencies of a bridge, clearly the operational loads can mask the underlying true dynamic features of a bridge structure. Nevertheless, González et al. [12] use eigenvalue analysis to demonstrate that it is possible to distinguish healthy and damage patterns of forced frequencies induced by a vehicle. Therefore, if a similar vehicle were used to obtain the forced frequencies of a bridge in two separate moments in time, their differences could potentially be exploited to assess the damage in the structure. The latter possibility is investigated here with the aim of developing a damage detection algorithm.

But if the frequencies of a bridge alone are characterised best in free vibration, is there any advantage gained from measuring accelerations in forced vibration? For a specific value of frequency measured in free vibration, there are infinite combinations of damage location and severity leading to that same value for the eigenfrequency associated with the first mode. Incorporating the second mode into the analysis will reduce this uncertainty and narrow down the number of potential solutions. Even then, there will still be at least two valid solutions if the structure and modes are symmetric, one to either side of the axis of symmetry. The non-uniqueness of the solution, environmental conditions and experimental uncertainties such as inaccuracies in measurements and the influence of operational loads that can cause shifts in frequency superior to those caused by damage, is the motivation for the new method proposed here.

The use of free vibration measurements is well extended in the SHM field due to undeniable merits, but the location of the vehicle is obviated in such signals. Unlike a free vibration record, a time-history of forced frequencies contains information about its unique relationship with the vehicle position, the damage location, and the damage severity. Moreover, the free vibration data is not lost in a forced vibration record since it dominates the start and end periods of the vehicle passage. In other words, the forced vibration record presents challenges due to VBI, but if understood, it also contains a wealth of information beyond that given by a free vibration record.

kNN algorithms have been successfully applied to the SHM of bridges in the past and are an often-used tool because of its simplicity, ease of implementation, and high efficiency [13]. These algorithms are based on searching the k points in a reference database that are closer to the measured data according to a function representing the distance between the two [14]. Then, the final solution is computed based on the k points that provide the minimum values of the distance. For instance, Diez et al. [15] use a kNN-based approach to evaluate the similarity between joints at different bridge locations and areas, as well as to classify those joints with similar features.

In this paper, the authors exploit the relationship of the frequency response of the structure with the location of a vehicle to propose a novel model-based SHM method that locates and quantifies damage. The method relies on the changes in the forced frequency associated with the first mode of the bridge as a result of a vehicle crossing. For this purpose, eigenvalue analysis is used to generate a database of theoretical forced frequencies for a wide range of damage scenarios that will be compared to the instantaneous frequencies obtained from measurements of the bridge response. The measurements are simulated here using a numerical VBI model to replicate the acceleration signal of a bridge. A kNN algorithm is calibrated to carry out the comparison between the theoretical and simulated forced frequencies. The goal of implementing a kNN algorithm based on random sampling is threefold: (i) to avoid the potential bias due to systematic sampling; (ii) to facilitate a generalized approach to the solution; and (iii) to deal with inaccuracies derived from the level of discretization of the eigenvalue sample, deviations introduced by the speed of the vehicle or the road profile, or deviations between the vehicle-bridge finite element matrices based on numerical models and the real system.

Section 2 describes the numerical VBI model for healthy and damaged scenarios, which serves to develop and test the proposed method. Section 3 explains the steps involved in determining the damage location and severity by the kNN algorithm. Section 4 tests the algorithm, first in an ideal scenario consisting of a truck travelling over a smooth profile at a very low speed. Secondly, the analysis is extended to highlight the limitations as speed increases and the road profile gets
rougher. Then, the span of the bridge under analysis is varied to assess the generality of the results. Finally, Section 5 provides conclusions and recommendations for further research.

2. Vehicle-bridge structural system

Fig. 1 shows the theoretical bridge and vehicle employed in the simulations. They are modelled as a simply supported beam discretised into finite elements and a 3-axle rigid truck.

2.1. Definition of the vehicle model

The truck model has 5 Degrees of Freedom (DoFs), namely the vertical displacements of the front axle, rear tandem and body masses (labelled $Z_1$, $Z_2$ and $Z_3$, respectively, in Fig. 1) and the rotations of the rear tandem mass ($\theta_1$) and the body mass ($\theta_2$). The mass $[M_v]$, damping $[C_v]$, and stiffness $[K_v]$ matrices of the vehicle are obtained by imposing the equilibrium of all forces and moments acting on the 5 DoFs of the vehicle [16]. These matrices are provided in Appendix A at the end of the article. The parameters of the vehicle are identified in Fig. 1, and defined in Table 1, together with the typical values that are adopted in the simulations based on Kim et al. [16] and Grave [17].

2.2. Definition of the bridge model

The finite element model (FEM) of the bridge consists of 4-DoF elementary beams, namely vertical translation and in-plane rotation of each node. Typical properties for a short span bridge are selected here. The bridge model extends over a 15 m simply supported span, while the cross-section has 15 m width and 0.75 m height ($h$) resulting in a 2nd moment of area $I = 0.5273 \text{ m}^4$. The material properties are based on reinforced concrete, i.e., modulus of elasticity $E = 35 \text{ GPa}$ and density of $2500 \text{ kg/m}^3$ leading to a mass per unit length $m = 28.125 \text{ kg/m}$. Bridge damping is typically low and neglected here. The FEM is discretised into 150 beam elements, based on a length of each element $L_e = 0.1 \text{ m}$. Hence, there are 151 nodes and 302 DoFs in total, and the dimensions of the global matrices of the bridge model, $[M_b]$ and $[K_b]$, are $302 \times 302$ following the assemblage of all the elementary beam matrices $[M_e]$ and $[K_e]$. The detailed formulation of the latter is given in Appendix A.

From amongst the various approaches available for modelling a crack, this paper adopts the stiffness reduction proposed by Sinha et al. [18]. Fig. 2(a) shows the representative parameters of the crack model: the damaged depth, $h_d = \chi h$, where $\chi$ is usually employed to characterise the damage level as a percentage of the total beam height $h$; and the total length affected by the damage, denoted by $2l_d$, which according to Sinha covers a region equal to $1.5h$ at either side of the crack. In this region, the stiffness loss varies linearly until reaching a maximum at the crack location. When analysing a specific damage scenario, the stiffness matrix $[K_e]$ of the beam elements affected by the crack will be modified accordingly. Fig. 2(b) illustrates...
the stiffness distribution along the bridge for crack depths that vary from $k = 0\%$ (healthy) to $k = 26.67\%$ (200 mm) of the cross-sectional height.

2.3. Eigenvalue analysis

An eigenvalue analysis consists of multiplying the inverse of the mass matrix by the stiffness matrix to obtain the eigenfrequencies of the system. It is then possible to obtain the natural frequencies given in Table 2, associated with the modes of vibration of the vehicle and the bridge.

By considering the time-dependent contact forces between the vehicle and the bridge, the $5 \times 5$ stiffness matrix of the vehicle, $[K_v]$, and the $302 \times 302$ stiffness matrix of the bridge, $[K_b]$, are combined into one $307 \times 307$ global stiffness matrix $[K_{gb}]$ of the vehicle-bridge system that varies with the position of the vehicle. The individual mass matrices of the truck and bridge are also combined into one $307 \times 307$ global mass matrix $[M_{gb}]$ that remains constant as long as the same axles are located on the bridge.

As a result of applying eigenvalue analysis to $[M_{gb}]$ and $[K_{gb}]$, Fig. 3 shows the changes in forced frequency associated with the first mode of vibration of the bridge in healthy condition as the vehicle moves along the bridge. The six vertical dashed lines indicate the front, 1st rear and 2nd rear axles entering (labelled as 1E, 2E and 3E) and leaving the bridge (labelled as 1L, 2L and 3L). For instance, the horizontal coordinate of 0 m corresponds to the front axle being located at the left support (1E) and 15 m corresponds to the front axle being located at the right support (1L). For horizontal coordinates greater than 15 m, the front axle has left the bridge, but other axles remain on the bridge. It can be noticed that the maximum difference between the forced and free eigenfrequencies is 0.032 Hz, and it occurs when the front axle is at 12.9 m from the left support.

For the specific vehicle and bridge parameters included in Sections 2.1 and 2.2, the eigenvalue analysis shows an increase of $f_{b1}$ in forced frequency with respect to its value in free vibration. The possibility of an apparent increment in frequency, rather than a reduction, has been reported by other authors in the literature [11,12,19].

2.4. Dynamic transient analysis

The interaction between the vehicle and the bridge is implemented using MATLAB. A dynamic transient simulation is performed to simulate the bridge response in the field. The coupled equations of motion are integrated to obtain the bridge response as described by González [20]. Appendix A summarizes the coupling procedure. The Newmark-beta integration scheme is used here, where the value of $\gamma$ is 0.5. Only one measurement output is necessary for the damage detection algorithm to be tested. Fig. 4 shows the selected output, i.e., bridge acceleration at midspan, due to the truck travelling at 1 m/s.
on a smooth surface for the case of a healthy bridge condition. The scanning frequency, $f_s$, is 1000 Hz, i.e., 1000 values of acceleration per second. Additionally, the front, 1st rear and 2nd rear axles entering and leaving the bridge are highlighted by six vertical dashed lines.

The instantaneous frequencies are obtained from the time series of the acceleration response at midspan. The fast Fourier transform (FFT) converts time-series signals into a frequency domain representation in which the main frequencies of a structure are easily identified, but the FFT cannot depict the frequency changes over the time domain due to a vehicle crossing. To overcome this limitation, the Short-time Fourier transform (STFT) is proposed for providing a representation of the signal in the time–frequency domain. The STFT divides the measured response into shorter segments of the same window length. Usually, these segments overlap with each other over a constant length to improve the resolution in the time domain. The overlap is the result of moving the window along the signal by an amount (hop size) smaller than the window length. Then, the FFT is applied to each segment separately in order to derive its frequency component [21]. Therefore, those components may contain frequency changes over time. The STFT is restricted by the selection of the window length. A large window length will provide good resolution in the frequency domain and poor resolution in the time domain, and vice-versa.

![Diagram](image_url)

**Fig. 2.** Definition of Sinha’s damage model: (a) Model parameters, (b) Stiffness profiles for 6 scenarios with different damage levels for a crack located at 4 m from the left support.

**Table 2**

Frequencies of the vehicle and healthy bridge in free vibration.

<table>
<thead>
<tr>
<th>Vehicle</th>
<th>Frequency (Hz)</th>
<th>Bridge</th>
<th>Frequency (Hz)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mode</td>
<td></td>
<td>Mode</td>
<td></td>
</tr>
<tr>
<td>Body mass bounce, $f_{v1}$</td>
<td>2.076</td>
<td>First, $f_{b1}$</td>
<td>5.655</td>
</tr>
<tr>
<td>Body mass pitch, $f_{v2}$</td>
<td>2.178</td>
<td>Second, $f_{b2}$</td>
<td>22.621</td>
</tr>
<tr>
<td>Rear tandem pitch, $f_{v3}$</td>
<td>10.849</td>
<td>Third, $f_{b3}$</td>
<td>50.898</td>
</tr>
<tr>
<td>Front axle hop, $f_{v4}$</td>
<td>15.540</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rear tandem hop, $f_{v5}$</td>
<td>15.812</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The instantaneous frequencies are obtained from the time series of the acceleration response at midspan. The fast Fourier transform (FFT) converts time-series signals into a frequency domain representation in which the main frequencies of a structure are easily identified, but the FFT cannot depict the frequency changes over the time domain due to a vehicle crossing.
Fig. 5 shows the time–frequency spectrum obtained by applying the STFT to the acceleration in forced-vibration provided in Fig. 4. In the figure, dark and light colours correspond to low and high amplitudes of the FFT, respectively. Then, the frequency associated with the maximum value of the amplitude is identified at each instant in time. In other words, if the contour plot is sliced following a vertical line, a plot similar to the result of applying the FFT to a windowed signal is obtained, and the frequency corresponding to the peak with the maximum amplitude would be extracted. Hence, it is possible to obtain a curve showing the evolution of the instantaneous frequency with respect to the location of the front axle. This curve is marked in Fig. 5 using a dash-dotted dark line. The curve is clearly evolving around a range of frequencies close to the 1st frequency of the bridge ($f_{b1} = 5.655$ Hz).

Fig. 6 zooms into the frequency range of interest to highlight the small changes in the first frequency as the vehicle moves along the bridge. There are various types of window functions that can be utilized in the STFT, being Hamming, Gaussian and Hann amongst the most popular options [23]. Fig. 6 shows the difference between these three window types compared to the eigenvalue analysis from Fig. 3. The Hann window outperforms the other two alternatives in matching the eigenvalue curve. In order to optimise the trade-off between time and frequency resolution, the selected length for the Hann window ($L_{win}$) is $2^{10}$ points, the hop size is $2^3$ points and the total number of FFT points is $2^{20}$. It can be seen that there is a good agreement between the results from the STFT and the eigenvalue analysis, although the former slightly over predicts the change in
frequency and the latter is smoother. Nonetheless, the STFT appears to be a valid tool for capturing the main changes in forced frequency identified by eigenvalue analysis for this particular scenario.

It is obvious from Fig. 6 that the forced frequencies obtained from applying the STFT are polluted with high uncertainty around the instants when axles are entering or leaving the bridge, which are signalled by vertical dashed-dotted lines. These inaccurately predicted frequencies, corresponding to the 1st and 2nd rear axles entering the bridge and the front and 1st rear axles leaving the bridge, are caused by discontinuities in the transient acceleration response. The discontinuities produce unreliable frequency results during the shift from the time domain to the time–frequency domain by the STFT. These polluted areas will be removed prior to the application of the kNN algorithm. The polluted time, \( t_{ps} \), for a single side of each corrupted instant is defined here as: \( L_{win}/2f_s \). Therefore, the total length of the polluted area of each corrupted instant in the spatial domain can be described as: \( w_p = 2nt_{ps} \), where \( \nu \) is the vehicle speed. For example, the values selected for the parameters in this scenario, where \( f_{win} = 2^{10} \) points, \( f_s = 1000 \) Hz and \( \nu = 1 \) m/s, lead to a polluted area that extends approximately 0.512 m to either side of the corrupted instant.

Fig. 5. STFT frequency-time spectrum of the signal in Fig. 4.

Fig. 6. Comparison of the forced frequency variation with vehicle location using eigenvalue analysis and the STFT.
3. Forced frequency-based kNN damage detection algorithm

3.1. Description

In practice, the amount of available measured forced frequencies will be limited, and experimental data for damaged conditions of the bridge will be scarce or non-existent. Therefore, the algorithm proposes to use eigenvalue analysis to obtain theoretical curves of forced frequencies versus vehicle position for a healthy condition and a vast range of theoretical damage scenarios based on Sections 2.2 and 2.3. Eigenvalue analysis is computationally very efficient and allows to quickly generate theoretical curves of forced frequencies versus vehicle position for a healthy condition and a vast range of theoretical damage conditions. The latter will serve for training a simple and easy-to-implement kNN algorithm. Using theoretical eigenvalue results, Eq. (1) is calculated for a wide range of damage locations and severities.

\[
\begin{align*}
\{f_{\text{diff}}^i\} &= \{f_{\text{dmg}}^i\} - \{f_h\}, \quad i = 1, 2, 3, \ldots, n, \\
\end{align*}
\]

where the parameter \(i\) represents the damage scenario number, \(n\) is the number of damage scenarios (i.e., the total number of training samples), \(\{f_h\}\) is a vector containing the forced frequencies for the healthy case (assumed to be the bridge in its original form), \(\{f_{\text{dmg}}^i\}\) is a vector containing the forced frequencies for the damaged case, and \(\{f_{\text{diff}}^i\}\) is the difference between the vectors \(\{f_{\text{dmg}}^i\}\) and \(\{f_h\}\). It must be noted that each vector \(\{f_{\text{diff}}^i\}\) is associated with a known crack location and severity.

The difference between an ‘unknown-health status’ and the healthy bridge can also be computed from on-site acceleration measurements following application of the STFT, as per Eq. (2).

\[
\{f_{\text{diff}}^i\} = \{f_{\text{unk}}\} - \{f_h\}
\]

where \(\{f_h\}\) is a vector containing the forced frequencies for a previous reference period where the bridge is deemed to be healthy, \(\{f_{\text{unk}}\}\) is a vector containing the forced frequencies for an ‘unknown-health status’, and \(\{f_{\text{diff}}^i\}\) is a vector containing the differences between \(\{f_h\}\) and \(\{f_{\text{unk}}\}\). For the vector \(\{f_{\text{diff}}^i\}\), the location and severity of the crack will be typically unknown, except for the validation phase.

Once the vectors \(\{f_{\text{diff}}^i\}\) and \(\{f_{\text{diff}}^j\}\) have been calculated, the output of the kNN algorithm will be based on the average of the \(k\) nearest neighbours. Therefore, the input consists of the different cracks associated with the \(k\) training cases \(\{f_{\text{diff}}^i\}\) closest to \(\{f_{\text{diff}}^i\}\). Here, the Euclidean distance [24] is used to quantify the discordancy between \(\{f_{\text{diff}}^i\}\) and \(\{f_{\text{diff}}^j\}\). For a set of two \(m\)-dimensional vectors \(\{f_{\text{diff}}^i\}\) and \(\{f_{\text{diff}}^j\}\), the Euclidean distance is defined by Eq. (3):

\[
E_{\text{diff}}(\{f_{\text{diff}}^i\}, \{f_{\text{diff}}^j\}) = \sqrt{\sum_{i=1}^{m} (f_{\text{diff}}^i - f_{\text{diff}}^j)^2}, \quad i = 1, 2, 3, \ldots, n; j = 1, 2, 3, \ldots, m.
\]

where \(f_{\text{diff}}^i\) is the component \(j\) of the vector \(\{f_{\text{diff}}^i\}\), regarded as the ‘unknown-health status’ scenario, and \(f_{\text{diff}}^j\) is the component \(j\) of the vector \(\{f_{\text{diff}}^j\}\), regarded as one of the known-damage scenarios. The parameter \(m\) constitutes the size of both vectors. It must be noted that the size of vector \(\{f_{\text{diff}}^i\}\) will be adjusted to match the size of \(\{f_{\text{diff}}^j\}\) when calculating the Euclidean distance.

Two phases can be distinguished in the implementation of the kNN algorithm: calibration, and testing. In the calibration phase, a random validation sample is employed to establish the parameter \(k\) and the size of the training database \(n\). For a given \(k\), the number \(n\) is increased until the error in damage location and severity incurred when assessing the validation sample is below an acceptable threshold. Once \(k\) and \(n\) have been established, a new random sample of damage scenarios is used to test the accuracy and ability of the algorithm to generalise. Fig. 7 gives a schematic representation of the procedure. Ultimately, the kNN algorithm is comparing the differences between the instantaneous forced frequencies versus vehicle position in the healthy and ‘unknown-health status’ conditions to the theoretical curves derived from eigenvalue analysis for locating and quantifying the damage.

In the absence of field measurements to extract \(\{f_{\text{diff}}^i\}\), the latter is simulated in this paper by solving the equation of motion and applying dynamic transient analysis (Section 2.4). More specifically, the acceleration response of a beam model (Section 2.2) traversed by a 3-axle truck (Section 2.1) is simulated. The Hann-based STFT is employed to obtain a time-history of instantaneous forced frequencies, although alternative time–frequency domain techniques such as the CWT and the Hilbert–Huang transform could also be applied. The same vehicle is assumed to be employed in all scenarios. When implementing the algorithm in the field, the properties of the test vehicle driven over the bridge must remain unaltered. Typical values have been adopted for these properties based on available truck data, that lead to realistic axle weights.
and frequencies of body and axle vibrations. Certainly, it would be theoretically possible to find properties, more sensitive than typical values, that would amplify the changes in the frequencies of the bridge to a larger extent. Although this situation would be optimal from a theoretical point of view, it will also be more difficult and costly to implement in practice. For instance, the latter would imply adjusting the test vehicle and its properties for each bridge under investigation.

Two moments in time are considered: (i) an initial time from the past in which the condition of the structure is known, and (ii) another time when the structure has an 'unknown-health status'. First, the models of the bridge and test vehicle are validated based on initial tests conducted on the vehicle and the bridge at a time (i). Moreover, the goal of this first test is also
to obtain and validate \(\{f_i\}_h\). Then, the kNN algorithm is calibrated based on the numerical models to allow for changes in damage severity and location. A second test is carried out on the bridge at a time (ii) using the same vehicle. The forced frequencies for the ‘unknown-health status’ \(\{f_i^{\text{unk}}\}_h\) are obtained and fed into the algorithm, which will assess whether there is damage or not. Using eigenvalue analysis, Section 3.2 highlights the differences in forced frequencies \(\{f_i^\text{diff}\}_h\) as a result of a change in damage location. Section 3.3 demonstrates how the Euclidean distance is used to group those distributions of bridge stiffness yielding minimum values of the distance and to calibrate the kNN algorithm.

**Fig. 8.** Difference between forced eigenfrequencies for a 100 mm crack located at: (a) 0 m, (b) 2 m, (c) 4 m, (d) 6 m, (e) 8 m, (f) 10 m, (g) 12 m, (h) 14 m from the left support.
3.2. Changes in eigenfrequencies as a result of damage

Before proceeding to the calibration of the algorithm, it is worth illustrating the differences in forced frequencies versus vehicle position between a damage scenario and the healthy scenario. Fig. 8 plots \(f_{\text{diff}}\) for eight different cracks with a 10 cm depth (13.33\% of the cross-sectional height) that vary location between 0 m and 14 m, measured from the left support. The vertical dotted lines in each subfigure signal the real damage location. There appear to be two troughs that move in the same direction than the location of damage, and that increase in magnitude as the damage gets closer to midspan. The trough closer to the left support matches the dotted line for the subfigures from (d) to (h), i.e., within 6–14 m from the left support. The values of \(f_{\text{diff}}\) near the supports are much smaller than near midspan, denoting a lack of sensitivity to damage, especially for 0–2 m and 12–14 m. A maximum relative change of around 3.2\% can be observed in subfigure (e), while relative changes in subfigure (a) are only in the range of 0.014\%.

For illustration purposes, Fig. 9 shows the first eigenfrequency in free vibration for locations and severities of the crack varied between 0 and 15000 mm and 0–200 mm, respectively. As expected, the contour plot is symmetric with respect to midspan. For a specific location of the crack, the value of the eigenfrequency decreases as the damage severity increases. As in forced vibration, the extent of this decrease becomes more significant as the location of the crack gets closer to midspan. On the other hand, multiple combinations of crack location and depth yield the same value of the first frequency in free vibration. Fig. 9 in free vibration shares with Fig. 8 in forced vibration the low sensitivity to damage in locations close to the nodes of the first mode, i.e., the supports. However, a time-history of forced frequencies such as the one shown in Fig. 8 mitigates the impact of the symmetry and the large uncertainty around the damage that is associated with a particular crack when using one frequency value in free vibration.

3.3. Calibration

For the training sample, an \(n\) number of \(f_{\text{diff}}\) vectors is generated using eigenvalue analysis. For the validation sample, 1600 \(f_{\text{diff}}\) vectors are derived from dynamic transient runs (Section 2.4). For both training and validation samples, the damage associated with each vector is randomly varied using Monte Carlo simulation based on uniform distributions within the range [0, 15,000 mm] for location and [0, 200 mm] for severity and with a millimetric resolution, i.e., the sampling is made from a population of \(15,000 \times 200 = 3 \times 10^6\) possible combinations. Firstly, for each validation vector \(f_{\text{diff}}\), the \(n\) Euclidean distances, \(E_{\text{dis}}(\{f_{\text{diff}}\})\), are calculated according to Eq. (3). Each Euclidean distance has its own unique severity and location associated with it.

The Euclidean distance will not be exactly 0 due to discrepancies between eigenvalue analysis and dynamic transient analysis, i.e., as a result of edge effects and the limited duration of the signal in the transient analysis, unexplored space in the sampling (the same damage severity and location present in the transient analysis may have not been covered by eigenvalue analysis), etc. However, it is possible that considering the average of the first two nearest neighbours \((k = 2)\), three nearest neighbours \((k = 3)\) or more, yielded a better solution than simply taking the highest ranked combination of crack

![Fig. 9. Contour plot of eigenfrequencies (Hz) in free vibration vs damage severity and damage location.](image)
depth and severity \((k = 1)\). Therefore, the performance of the kNN algorithm greatly relies on the selection of an appropriate \(k\) value. Hence, a for-loop as in Fig. 7 is used to compute damage localization and quantification errors \((e_k^l \) and \(e_k^s\)) for \(k = 1, \ldots, n\), as defined by Eqs. (4) and (5).

\[
e_k^l = Location_{\text{est}}^k - Location_{\text{real}}, \quad k = 1, 2, 3, \ldots, n. \tag{4}
\]

\[
e_k^s = Severity_{\text{est}}^k - Severity_{\text{real}}, \quad k = 1, 2, 3, \ldots, n. \tag{5}
\]

where Location_{\text{est}}^k and Severity_{\text{est}}^k represent the location and severity of damage estimated by the kNN algorithm for a particular \(k\) value, and Location_{\text{real}} and Severity_{\text{real}} represent the true location and severity associated with the dynamic transient run.

A training sample of size \(n = 10^5\) based on eigenvalue analysis and a validation sample of 1600 dynamic transient runs are used to establish that \(k = 1\) leads to the best solution for the largest proportion of the problem at hand. Hence, the unique attributes of the objective transient vector will be estimated directly using the values of the closest training vector in the sections that follow.

4. Theoretical testing

4.1. Vehicle crossing a 15 m bridge with a smooth profile at 1 m/s

A testing sample composed of 1600 transient vectors \(\{\mathbf{f}_{\text{diff}}\}\), distinguished by damage severity and/or location, is generated randomly in a similar fashion, but independently from the validation sample. The conditions of this test are rather ideal, with a very low vehicle speed of 1 m/s and a smooth road surface. In order to analyse the accuracy of the algorithm in predicting damage location and severity, the transient vectors have been classified into 75 groups according to 5 damage severity ranges \((0–4, 4–8, 8–12, 12–16, \text{and} \ 16–20 \text{ cm})\) and 15 location ranges \((0–1, 1–2, \ldots, 13–14, \text{and} \ 14–15 \text{ m})\), i.e., there is an average of 21 scenarios per group. For each group, the mean and standard deviation (STD) of the errors in localization, \(e_k^l\) – Eq. (4) –, and quantification, \(e_k^s\) – Eq. (5) –, are calculated.

Fig. 10 shows the statistics for \(e_k^l\) as a function of the true location and severity of the damage. It should be noted that the magnitude of the errors found in the validation and testing samples is similar. Negative and positive values of the error in Fig. 10(a) imply that the estimated location is to the left or the right of the real damage location, respectively. As a result, all errors are negative in the range closest to the right support and positive in the range closest to the left support. For those scenarios within the ranges \([4–20 \text{ cm}]\) for severity and \([2–13 \text{ m}]\) for location, the mean error in locating damage is smaller or equal to 0.335 m. Even further, the error tends to decrease as the damage increases in severity, and its location approaches midspan. For instance, when the damage is located between 5 and 6 m, the mean errors are \(-0.272, -0.228, -0.172, \text{and} \ -0.182\); 118 m for severity ranges \([4–8, 8–12, 12–16, \text{and} \ 16–20 \text{ cm}]\), respectively. As expected, the algorithm has more difficulty in estimating the location as the severity becomes smaller, leading to significantly larger errors within the lowest damaged range \([0–4 \text{ cm}]\) \((<5.3\% \text{ of the cross-sectional height})\), than within higher ranges \([4–20 \text{ cm}]\). Additionally, there is a loss of accuracy near the supports due to the reasons exposed in Section 3.2. A similar trend can be found for the STD values in Fig. 10(b); the degree of uncertainty surrounding damage locations within the ranges \([4–20 \text{ cm}]\) for severity and \([2–13 \text{ m}]\) for location is far smaller than for the lowest severity range or locations near the supports. In particular, the largest STD when estimating a crack located between 2 and 13 m with a severity between 4 and 20 cm is 0.735 m. Furthermore, 90% of the errors when locating damages between 2 and 13 m for severities between 4 and 20 cm are smaller than 1.459 m. This 90% percentile decreases to 0.467 m for cracks located at midspan \([7–8 \text{ m}]\) with depths above 4 cm. In summary, the kNN algorithm can conciliate eigenvalue analysis with transient results at 1 m/s to locate damage with high accuracy, except for very low severities or locations near the supports.

The mean value and STD of \(e_k^s\) – the error in estimating the crack severity (or depth) in mm – are given in Fig. 11. Positive and negative values represent an overestimation and underestimation of the true depth, respectively. Similar conclusions to those found for the estimation of the location of the crack are drawn here, i.e., the mean and STD values of the error decrease as the damage location gets closer to midspan. Nonetheless, it is worth highlighting that this time there is no significant loss of accuracy for the lowest severity range \([0–4 \text{ cm}]\). For crack locations between 2 and 13 m, the maximum mean error is 2.5 cm regardless of the severity range, and for crack locations between 5 and 10 m, that value comes down to 0.4 cm with most of the groups yielding mean errors of 0 and 0.1 cm. In contrast, mean errors within 2 m from the supports reach values in the order of a few centimetres. Regarding the STD, there are higher values, signalling a larger degree of uncertainty, in the outer 4 m, while the STD for the groups covering the inner 11 m of the bridge remains below 5 cm. Whereas cracks located within the range \([5–10 \text{ m}]\) lead to values of \(e_k^s\) that fall within the interval 0.1–1.0 cm in 90% of the cases, cracks located within the ranges \([2–5 \text{ m}]\) and \([10–13 \text{ m}]\) lead to a 90% probability of the errors being in the interval 0.4–11.2 cm. Finally, the errors in estimating the severity of cracks located near the supports – range \([0–2, 13–15 \text{ m}]\) – have a 90% probability of being over 2.4 cm and up to 18.7 cm.
Two cases, characterised by the location and depth of cracks given in Table 3, are extracted from the testing sample for further investigation. Case ‘A’ is a scenario near the midspan where the damage has been predicted with high accuracy, and case ‘B’ is a scenario near the supports with poor accuracy.

Fig. 12(a) and (b) shows the top 20 ranked solutions for cases ‘A’ and ‘B’, respectively. They are ordered from left to right starting by the lowest value of the Euclidean distance. For each bar group, the first, second, and third bars provide the value of the Euclidean distance, and the crack location and severity associated with each solution, respectively. The values of the Euclidean distance for case ‘A’ are much larger than those from case ‘B’. Hence, a smaller value of the Euclidean distance does not necessarily mean a more accurate solution. On the other hand, the difference in the Euclidean distance between the 1st and 20th ranked solutions is in the order of $300 \times 10^{-5}$ and $2 \times 10^{-5}$ for cases ‘A’ and ‘B’. Thus, there is a relatively large degree of uncertainty in the selection of the optimal solution for case ‘B’.

The scatter of the top 20 ranked solutions in terms of damage location and severity for these two cases can be visualised in Fig. 13. Fig. 13(a) shows that there are a few solutions within the close neighbourhood of the real solution for case ‘A’, regarding both damage location and severity. It should be noted that the number of potential solutions near the true solution will be limited by the size of the training sample. For case ‘B’, the top 20 solutions fall at the opposite end of the bridge where the true solution lies. Moreover, the scatter of the solutions in terms of damage location and severity is significantly larger than for case ‘A’, i.e., the solutions span over a range of 4 mm in severity and 1611 mm in location for case ‘A’ as opposed to 163 mm and 5022 mm in case ‘B’.

Furthermore, Fig. 14 compares the true forced frequency versus front axle location curve from eigenvalue analysis to the STFT curve and the top 3 solutions ranked by the kNN algorithm. Although the top ranked solutions resemble very closely the STFT curve for both cases ‘A’ and ‘B’, the differences between the true eigenvalue and STFT curves are greater for case ‘B’. Hence, the poor accuracy for case ‘B’ is motivated by the negligible change in forced frequency near the supports – Fig. 8.

Fig. 10. Errors in locating the crack: (a) Mean value, (b) STD.
of the signal, which has been further shortened by removing edge effects prior to the analysis. Overall, Figs. 11 and 12 have shown that, except for locations near the supports, the proposed technique has the potential to reach a level 3 damage detection category with somewhat better accuracy in quantifying damage than in locating it. Nevertheless, the STFT introduces discrepancies with respect to the eigenvalue solution that become more significant the smaller the damaged-induced changes in forced frequency and the shorter the signal. This aspect will be worsened with the use of higher vehicle speeds, which are assessed together with the impact of a road profile in the next section.

Fig. 11. Errors in quantifying the crack depth: (a) Mean value, (b) STD.

<table>
<thead>
<tr>
<th>Scenario</th>
<th>Damage location (mm)</th>
<th>Damage severity (mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Real</td>
<td>Estimated</td>
</tr>
<tr>
<td>Case 'A'</td>
<td>7177</td>
<td>7193</td>
</tr>
<tr>
<td>Case 'B'</td>
<td>14,339</td>
<td>2891</td>
</tr>
</tbody>
</table>

(a) and (h) –, the compromise between time and frequency resolution by the STFT and the duration of the signal, which has been further shortened by removing edge effects prior to the analysis. Overall, Figs. 11 and 12 have shown that, except for locations near the supports, the proposed technique has the potential to reach a level 3 damage detection category with somewhat better accuracy in quantifying damage than in locating it. Nevertheless, the STFT introduces discrepancies with respect to the eigenvalue solution that become more significant the smaller the damaged-induced changes in forced frequency and the shorter the signal. This aspect will be worsened with the use of higher vehicle speeds, which are assessed together with the impact of a road profile in the next section.
4.2. Vehicle crossing a 15 m bridge with a rough surface at higher speeds

The road surface profile is assumed to be a zero-mean stationary Gaussian random process and theoretically generated from power spectral density functions through an inverse Fourier transformation. A very good road profile with a geometric spatial mean of $0.1 \times 10^{-6}$ m$^2$/cycle, road class ‘A’ according to ISO 8608 [25], is adopted for all testing simulations in this section. Four new testing samples, each consisting of 1600 dynamic transient runs, are generated. Each testing sample is associated with a different vehicle speed (1 m/s, 2 m/s, 4 m/s, and 8 m/s) and covers the same damage location and severity combinations as the testing sample in Section 4.1. For visualisation purposes, the 1600 quantification errors from each of the four testing samples are classified into 15 groups according to 3 damage severity ranges (0–4, 4–12, and 12–20 cm) and 5 location ranges (0–3, 3–6, 6–9, 9–12, and 12–15 m). The algorithm has limited success analysing those scenarios with lower damage severity, especially for crack depths below 4 cm (<5.3% of the section height). Moreover, the presence of the class ‘A’ road profile and the increase in vehicle speed have a negative effect in the localization capabilities of the algorithm. In the most favourable scenarios, i.e., damage severity range [12–20 cm] and location range [6–9 m] with the vehicle moving at

![Fig. 12. Rank of the 20 minimum values of Euclidean distances: (a) Case ‘A’, (b) Case ‘B’.](image-url)
1 m/s and 2 m/s, the percentage of errors less than 2 m for locating damage are 67% and 52%, respectively. Hence, the results presented in this section will focus on the quantification of damage.

Fig. 15(a) shows the percentage distribution of the damage quantification errors across 3 categories, $e_k^q < 0.025h$, $e_k^q < 0.05h$, $e_k^q < 0.1h$, for those scenarios within the damage severity range [12–20 cm] (i.e., crack depth between 16% $h$ and 27% $h$) and the vehicle travelling at 1 m/s over a class 'A' road. The percentage of predictions within each category decreases as the location of damage moves towards the supports, which is consistent with the results of Fig. 11 based on a smooth profile. This trend is somewhat expected since Fig. 8(a) and (h) suggest that changes in frequency near the supports will be close to negligible and far more sensitive to small deviations from eigenvalue analysis. On the other hand, Fig. 8(d) and (e), corresponding to damage near midspan, exhibit the largest changes in frequency, and as a result, these locations are less sensitive to inaccuracies. Although there is an underlying trend with minimum errors near midspan, the figure is not exactly symmetric since neither is the force vibration record. Hence the edge effects impact differently the assessment of the left and right parts of the bridge. For those scenarios with crack locations between 6 and 9 m, 65%, 89% and 100% of the damage quantification errors are $< 2.5\% h$ (=1.875 cm), $< 5\% h$ (=3.75 cm) and $< 10\% h$ (=7.5 cm), respectively.

Fig. 13. Predictions by the top 20 ranked solutions: (a) Case 'A', (b) Case 'B'.
Fig. 15(b) and 15(c) compare the percentage of damage quantification errors that are smaller than 5% $h$ and 10% $h$, respectively, as a function of the vehicle speed. Fig. 15(b) shows how the proposed algorithm can quantify the damage for cracks in the severity range $[12–20 \text{ cm}]$ with an error below 5% $h$ for a meaningful population, although there is a deterioration of the results as speed increases. If a larger margin for error were accepted, i.e., up to 10% $h$, Fig. 15(c) indicates that 99% of the cracks can be identified regardless of the tested speed for locations between 6 and 9 m. Within this less restrictive threshold of 10% $h$, the proposed kNN algorithm is able to detect the damage with a high probability for all speeds (over 91%) for cracks located in the range $[3–12 \text{ m}]$. As it is well known, the detection and quantification of damage near the supports $[0–3$ and $3–15 \text{ m}]$ become problematic due to a lower sensitivity (Section 3.2). In summary, the kNN algorithm performs best for lower speeds (1 and 2 m/s). Each time the speed is doubled, the duration of the signal is cut by half, with the subsequent loss of accuracy for determining instantaneous forced frequencies in the time domain, as well as an increase in sensitivity to edge effects and road irregularities. Therefore, the algorithm is recommended for controlled situations where the vehicle can be moved at very low speeds on a bridge with a smooth surface, or even parked at different locations on a bridge while absorbing ambient vibrations.

Fig. 16 compares the mean STFT and eigenvalue solutions for damage cases in the severity range $[12–20 \text{ cm}]$ and two different speeds, i.e., 2 m/s and 4 m/s. Fig. 16(a) corresponds to 109 transient simulations for which the cracks are located...
Fig. 15. Percentage distribution of damage quantification errors versus damage location in 15 m span bridge: (a) 1 m/s scenario, (b) 1, 2, 4, and 8 m/s scenarios with errors below 5% of the cross-sectional height, (c) 1, 2, 4, and 8 m/s scenarios with errors below 10% of the cross-sectional height.
in the range [3–6 m], and Fig. 16(b) to 99 transient simulations with cracks in the range [6–9 m]. The difference in the num-
ber of simulations arises from the randomness of the process. Therefore, there are 109 and 99 frequency-time spectra (Fig. 5)
behind the mean STFT shown in Fig. 16(a) and (b), respectively. The same scenarios are used to calculate the mean value of

**Fig. 16.** Comparison of the mean STFT and eigenvalue solutions for damage cases in the high severity range and two different speeds: (a) damage location 3–6 m (all predictions), (b) damage location 6–9 m (all predictions), (c) damage location 6–9 m (predictions with errors ≤5% h).
the corresponding exact eigenvalue solutions. As expected, the values of the forced frequencies associated with damage in the range that includes midspan [6–9 m] are lower than in the range [3–6 m]. The larger the separation between the mean STFT and the mean of the exact eigenvalue solutions, the more inaccurate the overall prediction of damage severity will become. It must be noted that edge effects are removed prior to the application of the algorithm and they are not included in the comparison (Section 2.4). Moreover, the higher the speed, the longer the edge effect in relation to the total duration of the signal. For the predictions of damage located in the range [3–6 m] (Fig. 16(a)), the difference between the mean STFT at 2 m/s and 4 m/s is not significant, which is consistent with Fig. 15(c). The same happens for the runs considered in Fig. 16(b) covering the prediction of damage located in the range [6–9 m].

On the other hand, Fig. 16(c) is a subset of Fig. 16(b), containing only those predictions with an error ≤5% h, i.e., 72 runs in the case of 2 m/s, and 33 runs in the case of 4 m/s. Fig. 16(c) shows that the runs captured within the 5% h threshold produce lower forced frequencies than the total 99 runs shown in Fig. 16(b). This is to be expected since lower frequencies result in greater changes with respect to the original frequency baseline. Even further, Fig. 16(c) shows that the mean forced frequencies captured within the 5% h threshold are noticeably lower at 4 m/s than at 2 m/s. In other words, runs at 4 m/s require larger changes in forced frequency to achieve an established degree of accuracy, to deal with the larger edge effects and shorter signal available. It is worth noting that Fig. 15(b) indicates that the percentage of estimations at 4 m/s with errors ≤5% h improves when the location of the crack moves from the midspan range [6–9 m] towards [3–6 m] or [9–12 m]. These differences in accuracy are attributed to deviations caused by road irregularities in critical locations. The latter does not appear to be problematic for very low speeds (1 or 2 m/s) when the dynamic forces of the vehicle are relatively small and the duration of the signal allows completing one or more cycles around the sought mean value, but it becomes more significant as the vehicle speed increases. Nevertheless, if a higher error threshold of 10% h was deemed to be acceptable, then the aforementioned deviations will be captured within the error and the percentage of successful runs will increase dramatically, even for 4 m/s and 8 m/s, as evidenced by Fig. 15(c).

4.3. Vehicle crossing a 30 m bridge with a rough surface

This section tests the performance of the algorithm on a 30 m long simply supported bridge traversed by the same test vehicle over the same road profile with roughness class ‘A’ used in Section 4.2. The new bridge model is discretised into 150 beam elements, each 0.2 m long. The properties of the model are referred to Li [26], where $m = 21068$ kg/m, $E = 35$ GPa, $I = 2.2643$ m$^4$, and the first three fundamental frequencies of the bridge are 3.385 Hz, 13.540 Hz and 30.465 Hz. Firstly, $10^5$ damage scenarios are characterized using eigenvalue analysis to populate the solution space of the kNN algorithm as per Section 3. The maximum change in forced frequency due exclusively to the presence of the vehicle, i.e., in the absence of damage, for the 30 m bridge is found to be 2.2% in contrast to 0.6% for the 15 m bridge (Fig. 3). Similarly to the 15 m bridge, 1600 dynamic transient simulations are carried out using a time step of 0.002 s (scanning frequency $f_s = 500$ Hz) for three speeds of the vehicle, namely 2 m/s, 4 m/s and 8 m/s. Each of these testing samples corresponds to random combinations of damage location (from 0 to 30 m on the bridge) and severity (from 0 to 40 cm crack depth). It must be noted that the healthy cross-section of the 30 m bridge has a height $h = 150$ cm. As in Section 4.2, the 1600 quantification errors from each of the three testing samples are divided into fifteen groups according to five damage location ranges (0–6, 6–12, 12–18, 18–24 and 24–30 m) and three damage severity ranges (0–8, 8–24 and 24–40 cm).

Fig. 17(a) presents the percentage distribution of the damage quantification errors divided into the same three categories used in Fig. 15, for those scenarios within the damage severity range [24–40 cm] (i.e., crack depth between 16% $h$ and 27% $h$) and the vehicle moving at 2 m/s over a class ‘A’ road profile. It can be seen that 76%, 94% and 100% of the damage quantification errors are ≤2.5% $h$ (=3.75 cm), ≤5% $h$ (=7.5 cm) and ≤10% $h$ (=15 cm), respectively, for those scenarios with damage located in the range around midspan [12–18 m]. Furthermore, for those scenarios with damage located between [6–12 m] and [18–24 m], over 56% and 95% of the damage quantification errors are smaller than 5% $h$ and 10% $h$, respectively. These relatively high percentages of success decay with an increase in vehicle speed. Therefore, Fig. 17(b) and (c) compare the percentage of damage quantification errors ≤5% $h$ and ≤10% $h$ respectively, for 2, 4 and 8 m/s. Unlike the 15 m bridge (Fig. 15(b)), the percentages of errors within 5% $h$ for 4 m/s and 8 m/s are larger in the 30 m bridge (Fig. 17(b)). This is likely due to the deviations introduced by the vehicular forces excited by the road roughness being relatively smaller compared to the inertial forces of the bridge. For damages located within the range [6–24 m], the kNN algorithm is able to detect damage in 94% of the highest speed scenarios with an accuracy ≤10% $h$ (Fig. 17(c)). Nevertheless, damage close to the supports, i.e., between 0 and 6 m and 24–30 m, leads to highly uncertain predictions for the reasons exposed in Section 3.2.

Fig. 18 compares the mean true eigenvalue solution corresponding to the testing sample covering the damage severity range [24–40 cm] to the mean instantaneous forced frequency obtained from applying the STFT to the dynamic transient simulations within the same range. Mean values shown in Fig. 18(a) and (b) are obtained from 126 and 106 dynamic transient runs for damage scenarios located in the ranges [6–12 m] and [12–18 m], respectively. As expected, edge effects at 4 m/s are larger than edge effects at 2 m/s, and the number of oscillations at 4 m/s is smaller than at 2 m/s. Hence, the average of the oscillations at 4 m/s will typically deviate more from the true eigenvalue solution and it will be less likely that the oscillations compensate for each other than the average of the oscillations at 2 m/s. As a result, the accuracy of the results shown in Fig. 17 is better for 2 m/s than for 4 m/s. This difference in accuracy due to vehicle speed is felt more strongly by the damage locations more sensitive to damage, i.e., near midspan (Fig. 17(b)).
Fig. 17. Percentage distribution of damage quantification errors versus damage location in 30 m span bridge: (a): 2 m/s scenario, (b): 2, 4, and 8 m/s scenarios with errors below 5% of the cross-sectional height, (c) 2, 4, and 8 m/s scenarios with errors below 10% of the cross-sectional height.
Eigenvalue analysis demonstrates the different levels of sensitivity to damage of the first forced frequency of the bridge depending on the location of the damage. For both the 15 m and 30 m bridges, the results follow a similar pattern that exhibits a higher percentage of success near midspan, and a lower percentage of success as the location of damage moves towards the supports. Accuracy tends to decrease as edge effects, vehicle speed or road roughness increase. Similar levels of accuracy are found for both bridges in relative terms (i.e., expressed as a % of the total cross-sectional height), except for cracks located near midspan. In the latter, the 30 m bridge outperforms the 15 m bridge, as a result of more significant changes on the forced frequency together with a longer duration of the simulated or measured signal.

5. Conclusions

This paper has proposed a novel damage detection method based on a kNN algorithm that uses instantaneous forced frequencies to attempt the localization and quantification of damage. For a given damage location and severity combination, there is a unique curve of forced eigenfrequency versus vehicle position. The computational efficiency of eigenvalue analysis has been exploited to train a kNN algorithm with $10^5$ eigenfrequency curves that vary in damage severity and location. The points of each curve define the forced frequency corresponding to a different position of a 3-axle rigid truck on a 15 m simply supported bridge.
supported bridge. For testing purposes, instantaneous frequencies have been obtained from the simulated acceleration measurements of a VBI system with a Hann-based STFT. The algorithm has then compared the differences between the forced frequency curves in the healthy and unknown status conditions from dynamic transient test runs to the theoretical curves derived from eigenvalue analysis for estimating damage. Results have shown that damage can be detected, and in optimal cases, located and quantified, with some noticeable unfavourable locations near the supports. However, overall accuracy has been compromised with an increase in speed and road roughness, which broadens the discrepancies between eigenvalue analysis and dynamic transient analysis.

CRediT authorship contribution statement

**Kun Feng:** Conceptualization, Methodology, Software, Formal analysis, Investigation, Data curation, Writing - original draft, Visualization. **Arturo González:** Conceptualization, Methodology, Resources, Writing - review & editing, Supervision, Project administration, Funding acquisition. **Miguel Casero:** Conceptualization, Methodology, Validation, Resources, Writing - review & editing, Visualization.

Declaration of Competing Interest

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

Acknowledgements

This research has received funding from Science Foundation Ireland (SFI)'s US-Ireland R&D partnership programme under the proposal id. 16/US/I3277 titled MARS-Fly.

Appendix A

Eqs. (A1), (A2) and (A3) represent the vehicle matrices, where the first to last rows refer to the vertical translation of the front axle unsprung mass, vertical translation of the rear tandem unsprung mass, rotation of the rear tandem unsprung mass, vertical translation of the sprung mass, and rotation of the sprung mass, respectively.

\[
[M_v] = \begin{bmatrix}
    m_1 & 0 & 0 & 0 & 0 \\
    0 & m_2 & 0 & 0 & 0 \\
    0 & 0 & I_1 & 0 & 0 \\
    0 & 0 & 0 & m_3 & 0 \\
    0 & 0 & 0 & 0 & I_2 \\
\end{bmatrix}
\] (A1)

\[
[C_v] = \begin{bmatrix}
    C_{11} + C_{12} & 0 & 0 & -C_{11} & C_{11}L_{v3} \\
    0 & C_{21} + C_{22} + C_{32} & \frac{(-C_{22} + C_{33})L_{v3}}{2} & -C_{21} & -C_{21}L_{v4} \\
    0 & \frac{(-C_{22} + C_{33})L_{v3}^2}{4} & \frac{(C_{22} + C_{33})C_{33}L_{v3}}{2} & 0 & 0 \\
    -C_{11} & -C_{21} & 0 & C_{11} + C_{21} & -C_{11}L_{v3} + C_{21}L_{v4} \\
    C_{11}L_{v3} & -C_{21}L_{v4} & 0 & -C_{11}L_{v3} + C_{21}L_{v4} & C_{11}L_{v3}^2 + C_{21}L_{v4}^2 \\
\end{bmatrix}
\] (A2)

\[
[K_v] = \begin{bmatrix}
    K_{11} + K_{12} & 0 & 0 & -K_{11} & K_{11}L_{v3} \\
    0 & K_{21} + K_{22} + K_{32} & \frac{(-K_{22} + K_{33})L_{v3}}{2} & -K_{21} & -K_{21}L_{v4} \\
    0 & \frac{(-K_{22} + K_{33})L_{v3}^2}{4} & \frac{(K_{22} + K_{33})K_{33}L_{v3}}{2} & 0 & 0 \\
    -K_{11} & -K_{21} & 0 & K_{11} + K_{21} & -K_{11}L_{v3} + K_{21}L_{v4} \\
    K_{11}L_{v3} & -K_{21}L_{v4} & 0 & -K_{11}L_{v3} + K_{21}L_{v4} & K_{11}L_{v3}^2 + K_{21}L_{v4}^2 \\
\end{bmatrix}
\] (A3)

Eqs. (A4) and (A5) represent the elementary beam matrices of the bridge, where the 1st and 3rd rows are related to the vertical translation of the left and right end nodes of each elementary beam, whereas the 2nd and 4th rows correspond to the respective rotations. The global matrices of the bridge model, \([M_b]\) and \([K_b]\), are obtained following the assemblage of all the elementary beam matrices \([M_e]\) and \([K_e]\).

\[
[M_e] = \frac{mL_e}{420} \begin{bmatrix}
    156 & 22L_e & 54 & -13L_e^2 \\
    22L_e & 4L_e^2 & 13L_e & -3L_e^3 \\
    54 & 13L_e & 156 & -22L_e \\
    -13L_e & -3L_e^2 & -22L_e & 4L_e^3 \\
\end{bmatrix}
\] (A4)
The coupled equation of motion is formed as:

\[
[M_G] \{\ddot{u}_g\} + [C_g] \{\dot{u}_g\} + [K_g] \{u_g\} = \{F_g\} 
\]  

(A6)

where \([M_G]\), \([C_g]\), and \([K_g]\) represent the global mass matrix, global damping matrix, and global stiffness matrix, respectively; \(\{F_g\}\) is the global force vector; and \(\{u_g\} = \begin{bmatrix} u_v \\ u_b \end{bmatrix}\) is the unknown displacement vector with \(u_v\) corresponding to the DoFs of the vehicle and \(u_b\) to the DoFs of the bridge. In Eq. (A6), the superscripts \(\dot{u}_g\) and \(\ddot{u}_g\) indicate the first and second derivatives with respect to time, respectively. The global mass \([M_G]\), global damping \([C_g]\) and global stiffness \([K_g]\) matrices are given by:

\[
[M_G] = \begin{bmatrix} M_v & 0 \\ 0 & M_b \end{bmatrix}
\]

(A7)

\[
[C_g] = \begin{bmatrix} C_v & C_{vb} \\ C_{bv} & C_b + C_{bb} \end{bmatrix}
\]

(A8)

\[
[K_g] = \begin{bmatrix} K_v & K_{vb} \\ K_{bv} & K_b + K_{bb} \end{bmatrix}
\]

(A9)

The time-varying matrices \(C_{bv}, C_{bb}, C_{vb}, K_{bv}, K_{bb}\) and \(K_{vb}\) are related to the interaction forces between the vehicle and the bridge. The details of these matrices are given below:

\[
C_{bv} = -N_b \begin{bmatrix} C_{12} & 0 & 0 & 0 & 0 \\ 0 & C_{22} & \frac{I_{yv}C_{22}}{I_{yv}} & 0 & 0 \\ 0 & C_{32} & -\frac{I_{yv}C_{32}}{I_{yv}} & 0 & 0 \end{bmatrix}
\]

(A10)

\[
K_{bv} = -N_b \begin{bmatrix} K_{12} & 0 & 0 & 0 & 0 \\ 0 & K_{22} & \frac{I_{yv}K_{22}}{I_{yv}} & 0 & 0 \\ 0 & K_{32} & -\frac{I_{yv}K_{32}}{I_{yv}} & 0 & 0 \end{bmatrix}
\]

(A13)

\[
C_{bb} = N_b \begin{bmatrix} C_{12} & 0 & 0 \\ 0 & C_{22} & 0 \\ 0 & 0 & C_{32} \end{bmatrix}^T
\]

(A11)

\[
C_{vb} = C_{bb}^T
\]

(A12)

\[
K_{bb} = N_b \begin{bmatrix} K_{12} & 0 & 0 \\ 0 & K_{22} & 0 \\ 0 & 0 & K_{32} \end{bmatrix}^T
\]

(A14)

\[
K_{vb} = K_{bb}^T
\]

(A15)

where \(N_b\) is the location matrix at a given instant in time, represented in Eq. (A16).

\[
N_b = \begin{bmatrix} 0 & 0 & 0 \\ N_1 & : & 0 \\ : & N_2 & : \\ 0 & 0 & N_3 \\ 0 & 0 & 0 \end{bmatrix}
\]

(A16)
The size of this matrix is $N_2 \times N_2$, corresponding to the total number of DoFs in the beam elements, $N_2$, and the number of vehicle axles, $n_t$, respectively. The location matrix $N_2$ varies in time given that it affects the nodal displacements and rotations of the beam elements in contact with the vehicle. It is important to impose zero vertical translation to the entries corresponding to the boundary conditions. The Hermitian shape function $N_i$ for the $i$th interaction force located on an element $y$ can be written in global coordinates as below:

$$
N_i = \begin{cases}
1 - 3 \left( \frac{y - (y - 1)L_y}{L_y} \right)^2 + 2 \left( \frac{y - (y - 1)L_y}{L_y} \right)^3 \\
\left( x_i - (y - 1)L_y \right) - 2 \left( \frac{x_i - (y - 1)L_y}{L_y} \right)^2 \\
\left( x_i - (y - 1)L_y \right)^3 \\
- \frac{y - (y - 1)L_y}{L_y} \\
\left( x_i - (y - 1)L_y \right)^2 \\
\left( x_i - (y - 1)L_y \right)^3
\end{cases}
\quad \text{for} \quad (y - 1)L_y < x_i < yL_y
\quad (A17)
$$

where $i = 1, 2, 3$, represents the 1st 2nd and 3rd interaction force, respectively, and $y = 1, 2, \ldots, n_t$ refers to the total number of beam elements, $n_t$. Additionally, $x_i$ corresponds to the location reference for the $i$th vehicle axle measured from the left support. Outside of the range defined in Eq. (A17), the value of $N_i$ is zero. The location matrix is also used to calculate the global force vector $F_g$, which is given by:

$$
F_g = \begin{bmatrix}
K_{12}f_1 + C_{12}f_1 \\
K_{22}f_2 + C_{22}f_2 \\
K_{32}f_3 + C_{32}f_3 \\
0 \\
0 \\
\{ P_1 - (K_{12}f_1 + C_{12}f_1) \} \\
\{ P_2 - (K_{22}f_2 + C_{22}f_2) \} \\
\{ P_3 - (K_{32}f_3 + C_{32}f_3) \}
\end{bmatrix}
\quad (A18)
$$

where $f_1, f_2, f_3$ are the road profile displacements under the wheels 1, 2 and 3, respectively, and $P_1, P_2$ and $P_3$ represent the static axle weights for the front, 1st rear and 2nd rear axles, respectively.

References


