World Conference on Timber Engineering Oslo 2023

ULTRASONIC-BASED DEFECT IDENTIFICATION IN TIMBER USING UNIFORM MANIFOLD APPROXIMATION AND PROJECTION (UMAP)

Prashanth Gunasekaran[1], Mehrisadat Makki Alamdari[2], Hamid Valipour Goudarzi[3]

ABSTRACT: Naturally occurring strength-impacting defects in timber, such as grain deviation and knots, are common in commercial timber. Additionally, defects caused by the environment such as fungal decay and termite damage are also attributed to strength reduction which can lead to the failure of structural timber. This work provides a demonstration, of applying Uniform Manifold Approximation and Projection (UMAP), a novel feature extraction tool, for the purpose of identifying such defects in structural timber, with the use of an ultrasonic-based examination of the test specimens. Two case studies are presented, the first demonstrating UMAP's capability in identifying grain deviation and knots in timber, and the second, showing its ability to distinguish healthy sections from that with known internal damage. In both case studies, 54kHz pressure waves were transmitted transversely through pre-defined sections of known damage state and were analysed with UMAP – producing highly favourable clustering and data separation between defect states, compared to wavelet packet decomposition, a state-of-the-art signal analysis framework. The key contributions of the study are – UMAP is successfully applied for the first time for damage identification in commercial bare timber, and it is demonstrated to be able to identify signals sent through sections of different defect types. This demonstration in this area opens the door to future works, which may then scrutinise in-depth material-specific factors for the aim of defect characterisation in timber and timber composites using UMAP.

KEYWORDS: Ultrasonic testing, Pressure wave, Wavelet decomposition, Feature extraction, Damage identification

1 INTRODUCTION

Strength-reducing defects are highly common in commercial timber and are either naturally occurring at growth, or induced due to the service environment or loading. Some examples of these defects include grain deviations and knots, associated with the former, and that related to termite damage, fungal decay and splitting damage associated with the latter. Additionally, local splitting failure caused by the rupturing of in-service structural timber due to higher loads also leads to the creation of weak zones. These weak zones incited by these defects have been shown to often be the initiation point of further damage [1], and hence their detection while in-service can be used to inform rehabilitation decisions.

Non-destructive testing (NDT) methods used for the purposes of assessing defects in timber have evolved over the last few decades. Current ultrasonic-based NDT techniques adopted on timber [2] to investigate these defects, face challenges in the signal processing stage. Wavelet packet decomposition (WPD) is one of the leading current signal processing approaches adopted for these purposes in the literature [3]. In [4], the structural integrity of timber utility poles was assessed using WPD, and their embedment depth has also been investigated using similar techniques [5]. In these studies, it was found that its chief challenge was parameter calibration. In WPD, different hyperparameters (levels of decomposition, choice of mother wavelet and filter functions) need to be trialled in order to attain separation of data between healthy sections and defective sections – a time-consuming and multi-variate process.

In this paper, a novel dimensionality reduction and feature extraction tool, Uniform Manifold Approximation and Projection (UMAP) [6], is applied for the first time, for the purposes of timber defect identification. The study is presented in two case studies. In the first case study, a timber specimen exhibiting multiple states including healthy sections, a section with visible grain deviation, and one with a visible knot are tested ultrasonically. The choice to adopt a test specimen with visible defects is justified in the study via the following consideration – the received signal is a function of the medium it was propagated through, regardless of the external visibility of the defect. Hence, the same outputs can be drawn from a test specimen with identical defects that were not externally visible. Accordingly, through visual inspection, the ground truth of the defect state of the specimen can be ascertained and thereafter compared with the output of the adopted signal analysis techniques, serving as an evidence-based mechanism in this study to demonstrate the efficacy of the algorithm.

[1] Prashanth Gunasekaran, University of New South Wales p.gunasekaran@unsw.edu.au
[2] Mehrisadat Makki Alamdari, University of New South Wales m.makkialamdari@unsw.edu.au
[3] Hamid Valipour Goudarzi, University of New South Wales h.valipour@unsw.edu.au

https://doi.org/10.52202/069179-0109

54 kHz ultrasonic pressure waves (P-waves) are transmitted between the sections transversely, and the signals are analysed first using WPD. Thereafter, the signals are comparatively analysed using UMAP, with the differences between the two methods examined and the superiority of the latter being demonstrated. In the second case study, a test specimen exhibiting healthy sections is adopted, with a void fabricated internally in the middle, creating loss of material – serving as a proxy for a piece of timber that has experienced interior termite damage, fungal decay or internal splitting damage. 54 kHz P-waves were transmitted transversely through pre-defined discretised locations in the timber. Similar to the first case study, the signals were analysed with WPD first, and thereafter compared with a UMAP analysis, their results further discussed.

2 THEORETICAL BACKGROUND

The theoretical background behind the two signal analysis approaches compared in this study is briefly recalled in the succeeding sections.

2.1 WPD METHODOLOGY

WPD is a technique used to fundamentally capture the frequency characteristics of the assessed signal while time-domain information is still retained [3]. A two-stage process is used in the decomposition: (1) The signal is processed with a wavelet transform, wherein features of the signal similar to the adopted mother wavelet are extracted at a particular scale and time. The mother wavelet is scaled and translated over the input signal with the aid of discrete scaling and translation parameters in the transform function. This wavelet transform is recalled in Equation 1. Thereafter; (2) A pair of low-pass filters (LPF) and high-pass filters (HPF) are used to conduct sub-band coding on the signal to delineate time-frequency features, in which the approximate and detail wavelet coefficients are obtained, respectively. The LPF and HPF formulas are recalled in Equation 2 and 3.

$$G_{k,j} = \int_{-\infty}^{\infty} f(t)\psi_{k,j}(t)\,dt \tag{1}$$

where G = wavelet coefficients, ψ = mother wavelet function, k = scaling parameter and j = translation parameter.

$$y_{low}[v] = \sum_{t} f(t) \cdot c(2v - t) \tag{2}$$

$$y_{high}[v] = \sum_{t} f(t) \cdot u(2v - t) \tag{3}$$

where f = input signal, c = low-pass filter impulse response factor, u = high-pass filter impulse response factor, t = translation parameter and v = subsampling parameter.

The level of decomposition is an inherent hyperparameter to be chosen when conducting the analysis. Representations of both the two-level and three-level sub-band coding schema are presented in Figure 1.

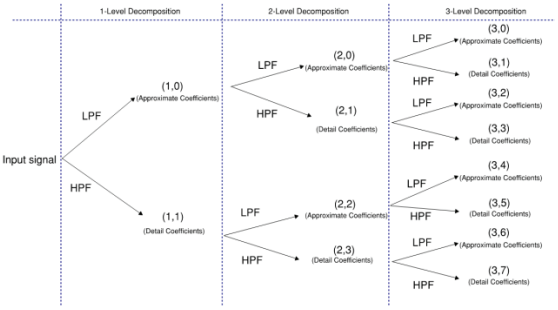

Figure 1: *Illustration of the sub-band coding schema for a 2 and 3-level wavelet packet decomposition.*

Apart from the level of decomposition, another key hyperparameter to select would be the mother wavelet, with a large variety of popular wavelet types adopted in the literature [7-10]. In this study both the Daubechies' 'db2' and 'db4' wavelets were investigated, and both 2 and 3-level decompositions were conducted, their results assessed in the succeeding sections.

2.2 UMAP METHODOLOGY

UMAP as an algorithm exists in the dimensionality reduction space, wherein the signal is represented as a point in a multi-dimensional space whose coordinates correspond to its amplitudes. UMAP then reduces the dimensions of the space to an easily comprehensible output graph.

The aim of the algorithm is to cluster signals with similar characteristics together and accurately and distinctively separate those with differing characteristics. It does so in two main stages. The first stage involves constructing the data's high dimensional topology as simplicial complexes. The Nerve theorem is applied, such that each signal is established as a 0-simplex and connections are made between each closely related signal – thereby approximating the topology of the high dimensional distribution of signals. These connections are made through the following process: (1) Spheres of varying radii are created around each signal based on density, through a fuzzy cover, wherein connections are formed with signals of intersecting radii. Mathematically, this is accomplished by first constructing the high dimensional signals on a Riemannian Manifold and establishing the spheres around each signal with a unit size. The required variable radii property is achieved through the homeomorphisms of these unit sized spheres on the manifold and their counterparts in Euclidian space, the latter allowing for such a property via varying notions of

https://doi.org/10.52202/069179-0109

distance. To maintain the same number of neighbouring signals allowed around each signal, fed into the algorithm as a hyperparameter, smaller notions of distance are used in Euclidian space for regions in the topology with higher density, and vice versa. (2) For efficient processing of the connections for each signal, concepts of Random Projection Trees [11] and Nearest Neighbour Descent [6] are utilised. (3) Connections between signals are mathematically represented through connection probabilities, considered via a 'neighbouring' matrix which holds each individual connection probability between signals in the original high dimension. In some regions of the topology, signals could have multiple connection probabilities associated with neighbouring signals in different metric spaces, due to the overlap of multiple spheres, and hence, distances of notion. Probability theory is utilised to condense these multiple connections to a single ideal connection between pairs of signals, by considering the combined probability of all connections between the signal pair.

In the second stage of the algorithm, the higher dimensional simplicial complex is projected to an easily comprehendible lower dimensional space. This done through a cross-entropy process, applied on the complexes, to cluster signals of similar characteristics together, and separate those with dissimilar features, projecting the topology into a two-dimensional graph [6]. The former, considered an attractive component between signals in the operation, is a commonly found trait of modern dimensionality reduction tools. UMAP's strength lies in the latter component of the cross-entropy process, considered a 'repulsive' operation, wherein the requisite distances between dissimilar signals are established. This results in the final output graph attained at the end of the algorithm.

The key hyperparameters to be calibrated for the algorithm are: (1) 'k' nearest neighbours – the number of nearby signals allowed for a given signal in the local metric space; (2) Minimum distance – referring to that between nearby signals in the final output. This hyperparameter is known to assist with the visualisation and cluster depiction; (3) Distance model – chosen for the mathematical calculation of distance between each signal.

3 ULTRASONIC TESTS ON TIMBER SPECIMENS

3.1 CASE STUDY 1

To investigate the versatility of the algorithm in structural defect detection and localisation in plain timber, the following test was conducted. The test specimen features an externally visible knot and grain deviation.

3.1.1 Specimen description

A 300mm long piece of machine-graded Australian *radiata pine* (MGP10), 90mm in width and 45mm in thickness was utilised, shown in Figure 2. Figure 2 also shows a schematic of the visible features in the test specimen. The knot in the specimen can be considered 'dead' due to a visible loss of fibre continuity between itself and the surrounding pieces of wood. The specimen was discretised into 4 sections. The centre of the knot is aligned with Section 4, and the grain direction change due to the presence of the knot is also within Section 3. Sections 1 and 2 did not show any visible signs of defects in the timber and featured longitudinally aligned grains - hence being classified as healthy sections.

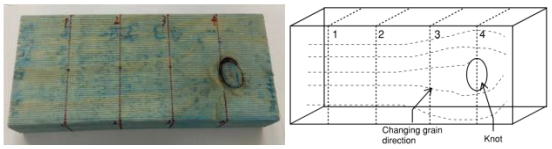

Figure 2: *Machine Graded Pine (MGP10) specimen (left) and schematic expressing defects (right).*

3.1.2 Time signals

P-waves of 54 kHz were transmitted transversely through the specimen, parallel to its 45mm thickness, in the perpendicular to grain direction. At each section of the specimen, 10 signals were transmitted with the transmitter on one end and the receiver on the other, thereafter having their positions swapped and measuring another 10 signals - amounting to a total of 20 signals at each measured section, summarised in Table 1.

Table 1: *Case Study 1 - Section information summary*

Section	Defect Status	Number of signals
1	Healthy	20
2	Healthy	20
3	Grain Deviation	20
4	Knot	20
Total number of signals		80

Figure 3 illustrates a sample of the ultrasonic signals sent through the timber specimen. The vertical axis represents signal energy, and for clearer visualisation, they are translated vertically with respect to each other. From visual inspection of the signals, the presence of the knot in Section 4 causes the most disruption to the received signal relative to the grain deviation in Section 3, when benchmarked against the signals from the healthy sections (1 and 2). The signal from Section 4 exhibits significantly lower wave energy than that of healthy sections, while the signal from Section 3 has similar wave amplitudes.

Prior studies have been conducted using stress waves in timber, wherein recorded phase velocities of signals transmitted transversely (perpendicular to grain) are smaller to that transmitted longitudinally (parallel to grain) - showcasing the significant effect that grain direction has on phase velocities [12,13].

https://doi.org/10.52202/069179-0109

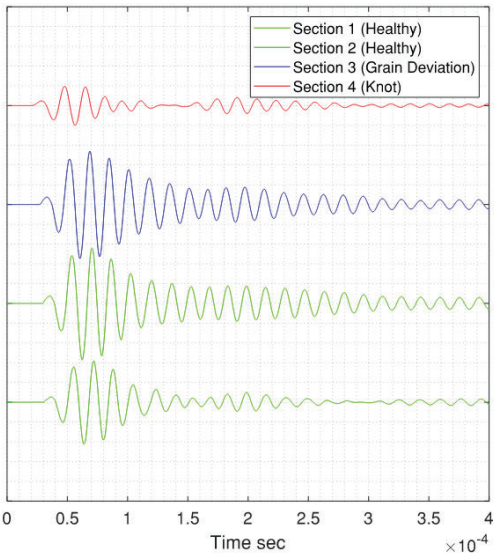

Figure 3: *Sample of time signals measured for each section in Case Study 1. The vertical axis represents signal energy, although each signal has been translated vertically for clarity. The signals are presented in order, with Section 1 featured as the lowermost and Section 4 as the uppermost signal.*

Where phase velocities are concerned in this test, Sections 1 and 2 exhibited similar P-wave speeds, with an average of 1660 m/s. However, in Section 3 with the presence of grain deviation, the average recorded phase velocity was 1838 m/s. In Section 4, wherein the largest disruption of grain direction was visible with the presence of the knot, the recorded average phase velocity of 2340 m/s was the fastest of all sections. This feature is in-line with findings from the aforementioned literature [12,13] demonstrating that in Sections 3 and 4, the grains were deviating in such a way that they were becoming more aligned with the transmitted signals, amounting to a parallel to grain transmission compared to the perpendicular to grain transmission conducted in Sections 1 and 2. This is also complementary with visual inspection of the test specimen shown in Figure 2.

The signal window used for the analysis with UMAP needs to be calibrated prior to use. P-waves exhibit higher phase velocities than surface waves, the latter being a component of ultrasonic signals that are often found to succeed the former [14]. As evidenced in Figure 3, it can be seen that the P-wave component of the signal cannot be easily identified due to significant overlap with the surface wave components. As such, beginning from the largest recorded arrival time from the 80 signals, the upper bound of the signal window used for the analysis needs to be calibrated in order to fine tune the results obtained, while the lower bound of the signal window is set at the start of the signal – although it should be noted that the sensitivity of the output of UMAP to this calibration is low.

3.1.3 Wavelet analysis results and discussion

Firstly, all 80 signals measured in Case Study 1 were analysed using WPD. The two key hyperparameters of WPD to be trialled are levels of decomposition, and mother wavelet type. For this study, both 2 and 3 levels of decomposition are conducted, and both the Daubechies' 'db2' and 'db4' mother wavelets have been trialled in this study.

2-level decomposition

Once both the detail and approximate wavelet coefficients are extracted, following Equation 1 to 3, their wave energies are then calculated. To encapsulate the outputs of the analysis in a single graph, a matrix scatter plot can be used – allowing the relationships of the wave energies of the decomposed signals between frequency band pairs to be expressed. This is represented in the off-diagonal plots presented in Figure 4, making any dependencies between the frequency band pairs easily discernible. The distribution of the wave energies of the decomposed signals within the associated frequency band is plotted on the diagonal plots in Figure 4, allowing single variable dependence to be ascertained.

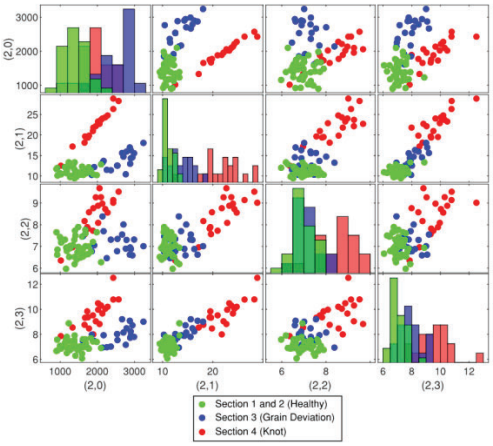

Figure 4: *Case Study 1 - Scatter plot of the signal energies in 4 frequency bands - decomposition level 2 and 'db2' mother wavelet.*

Some of the scatter plots shown in Figure 4 demonstrate some separation between data from Section 4 (knot) and the other sections. In particular, plots comparing frequency bands (2,1 – 2,0) show a promising separation of the majority of the wave energies of Section 4 from the rest, although notably there is still some overlap of the wave energies of some decomposed signals of Section 4 and healthy sections (1 and 2). The same can be said about the overlap of data from Section 3 and that of healthy sections (1 and 2). Investigating the other off-diagonal scatter plots, it can be seen that no distinct separation between data from the sections of each defect type and that of healthy sections can be discerned. The distribution for the wave energies of the individual studied frequency bands computed in each diagonal plot also illustrates

https://doi.org/10.52202/069179-0109

some separation between the section categories, with frequency band (2,1) showcasing the most promise - albeit significant overlap between the studied section categories is still present, with no high-confidence separation between the data.

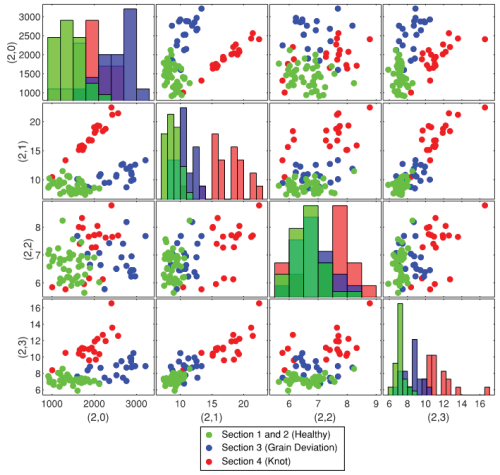

Figure 5: *Case Study 1 - Scatter plot of the signal energies in 4 frequency bands - decomposition level 2 and 'db4' mother wavelet.*

A similar scatter plot to the 'db2' decomposition can be plotted using the 'db4' mother wavelet, as in Figure 5 – the outcome of this analysis also being similar - showing a low sensitivity of the final output to changes in the Daubechies' mother wavelet type. Although some separation in signal energies corresponding to frequency band pairs (2,1 – 2,0) can be seen, none of the off-diagonal scatter plots depict a high-confidence separation of wave energies related to the desired section categories. The distributions of each frequency band plotted on the diagonal also feature large portions of overlap between signal energies of different section categories, also amounting to an undesirable outcome. Ultimately, the analysis using wavelet packet decomposition at 2 levels of decomposition with both the 'db2' and 'db4' mother wavelet does not achieve the desired data separation and clustering required to be able to tell the studied sections apart based on defect status.

3-level decomposition

Similar to the 2-level decomposition, a 3-level decomposition can be conducted on the transmitted signals. As expressed in Figure 1, at the third level of the sub-band coding schema, 8 frequency bands can be analysed, identified from (3,0) to (3,7).

When comparing the wave energies of the 'db2' wavelet decompositions of all 80 measured signals amongst the 8 frequency band pairs in the scatter plot matrix shown in Figure 6, it is clear that the majority of the frequency band pair plots do not show adequate separation between the section categories - a similar result being obtained when a decomposition with the 'db4' wavelet is done, shown in Figure 7. Notably, in Figure 6, some separation of the wave energies of Section 4 is seen, away from that of Sections 1, 2 and 3 in frequency band pairs (3,1 - 3,0) and (3,2 - 3,0).

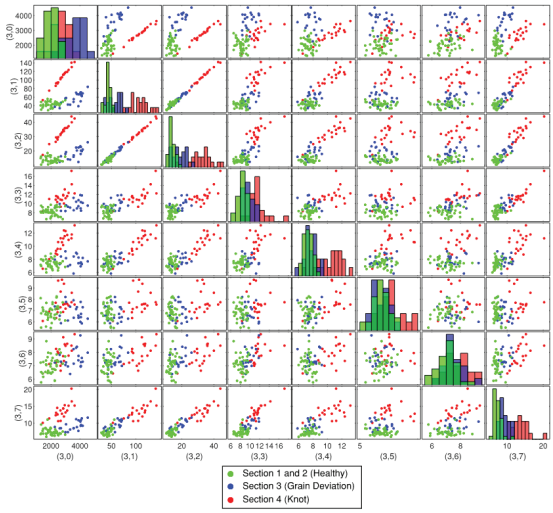

Figure 6: *Case Study 1 - A sample set of decomposed signals over 8 frequency bands with mother wavelet 'db2'.*

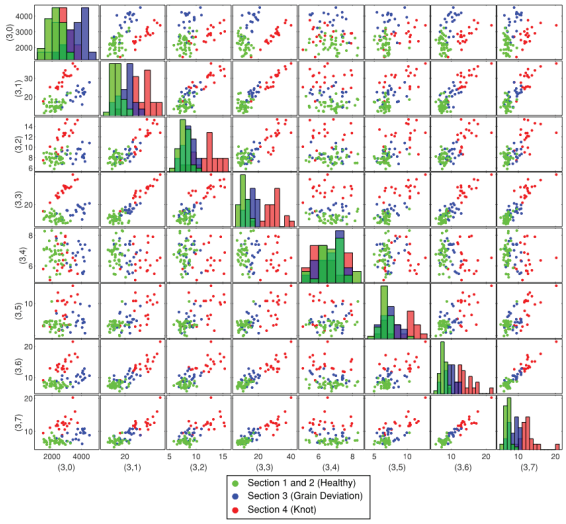

Figure 7: *Case Study 1 - A sample set of decomposed signals over 8 frequency bands with mother wavelet 'db4'.*

However, the distinction of wave energies in Section 3 from Sections 1 and 2 is not sufficient in these frequency band pairs. This feature is missing when studying the scatter plot output of the 'db4' decomposition (Figure 7), revealing the increased sensitivity of the output to the Daubechies' mother wavelet choice at this level of decomposition.

In summary, the preceding investigation could not yield a high-confidence separation between the signals transmitted through healthy sections and defective

sections. Moreover, no separation between signals sent through sections of varying defect types was discernible.

3.1.4 UMAP analysis results and discussion

The UMAP algorithm was also used to analyse the signals. It should be noted that for the UMAP analysis, the same signal window as that used in WPD was adopted. The key hyperparameters needed to be set were: (1) target embedment dimension, (2) k-nearest neighbours, (3) minimum distance, and (4) distance model. (1) refers to the dimension of the final output graph – a value of 2 dimensions being selected for simplicity of comprehension of the final results, additionally allowing for an easy comparison of the UMAP output with that of WPD demonstrated in the preceding sections. (2) and (3) work to control the level of clustering of the signals, changing how condensed or sparsely spaced the data points are, wherein the former sets the number of neighbouring signals which can be associated with a given signal, and the latter controls the minimum separation between neighbouring signals in the output graph. For this investigation, (2) was set at a value of 0.07 and (3) was set to 15 as they provided the most distinct clustering – although it should be noted that the sensitivity of the final output on these parameters is very low. (4) refers to the model used to express the distance between each signal, with many options available in the literature [15-18]. In this study, the Chebyshev distance model [18] was adopted, due to the highly favourable results obtained with this choice.

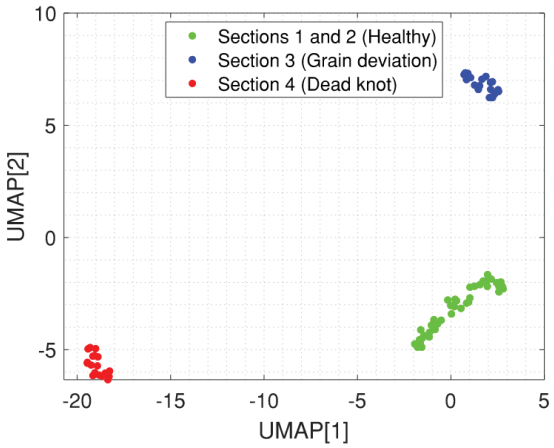

Figure 8: *Case Study 1 - UMAP two-dimensional (2D) reduction output.*

The UMAP results are shown in Figures 8 and 9. There is strong clustering of points with respect to each section. Clusters from both Sections 3 and 4 are separated from that of 1 and 2. Notably, wave behaviour in Section 3 is not clustered together with Section 4, showing the separation between a defect location characterised by grain deviation within a part of the specimen which is still intact with fibre continuity, and a defect location characterised by the centre of a dead knot.

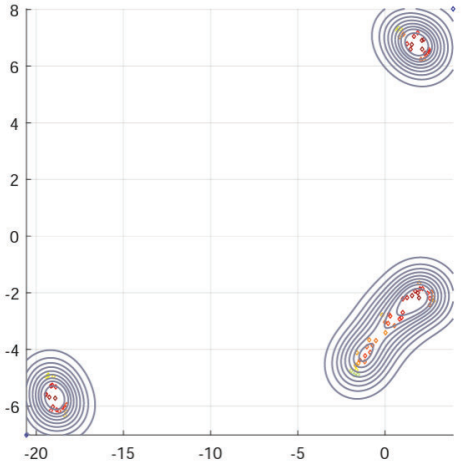

Figure 9: *Case Study 1 - UMAP two-dimensional (2D) cluster topology.*

A schematic of the relevant visual features of the specimen is shown in Figure 2. The outcome is highly desirable, especially since the results are not as sensitive to hyperparameter calibration, compared to the wavelet approach. The results are congruent with visible findings of the timber specimen, clearly distinguishing regions with material defects from those without any. The results demonstrate UMAP's capability of ultimately distinguishing between sections of sound wood and those with strength-reducing characteristics. Additionally, its ability to separate sections experiencing different types of strength-reducing characteristics is highly useful in structural health monitoring applications.

3.2 CASE STUDY 2

To examine the behaviour of both WPD and UMAP for defects caused by fungal decay, termite damage or splitting damage – ultimately eventuating in fibre discontinuity or voids in the centre of the cross-section, Case Study 2 was developed.

3.2.1 Specimen description

A piece of machine-graded Australian *radiata pine* (MGP10), similar to that used in Case Study 1 was employed. The specimen features a 190mm width and 45mm thickness. The specimen was carefully selected, such that no grain deviation or knots were present, with only longitudinally aligned grains visible – signifying a standard healthy section of timber.

3 sections were devised for the case study, a picture of the test specimen and its schematic shown in Figures 10a and 10b, respectively. Sections 1 and 2 are located within the healthy zone of the timber. To simulate the loss of material and fibre discontinuity due to the aforementioned defects, a hole was drilled at the centre of the cross-section, with a 6mm drill bit, as shown in Figures 10a and 10b. This method of simulating either splitting damage or that due to decay or termite attack is justified with two considerations: (1) Termite attack and fungal decay have

https://doi.org/10.52202/069179-0109

been shown to manifest as a material loss with the creation of voids within structural timber [19]; (2) Splitting damage manifests as longitudinally aligned cracks along the length of the timber element, with similar creation of voids used to artificially recreate such damage, as utilised in previous studies [20].

Figure 10a: *Case Study 2 – Test specimen*

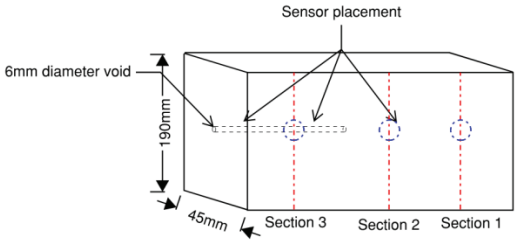

Figure 10b: *Case Study 2 – Specimen schematic*

A point directly aligned over the position of the created void was marked as Section 3, representing the fabricated defective section.

3.2.2 Time signals

54 kHz P-waves were transmitted over each of the sections, similar to the process adopted in Case Study 1. 10 ultrasonic signals were transmitted with a given transmitter and receiver position, and another 10 were recorded with their positions swapped, amounting to 20 signals measured at each location, and 60 signals measured in total for the analysis. A summary of the recorded signals is reflected in Table 2.

Table 2: *Case Study 2 - Section information summary*

Section	Defect Status	Number of signals
1	Healthy	20
2	Healthy	20
3	Drilled	20
Total number of signals		60

The measured time signals reflected a similar phase velocity across all sections, with the recorded averages being 2152 m/s, 2157 m/s and 2159 m/s for Sections 1, 2 and 3 respectively. It can be seen that they have negligible differences between each other, as evidenced by a largely similar wave pattern comparing Section 3 with that of Section 1 and 2, observed in Figure 11, which illustrates a sample of the signals, recorded at each section. This could possibly be due to the small size of the void.

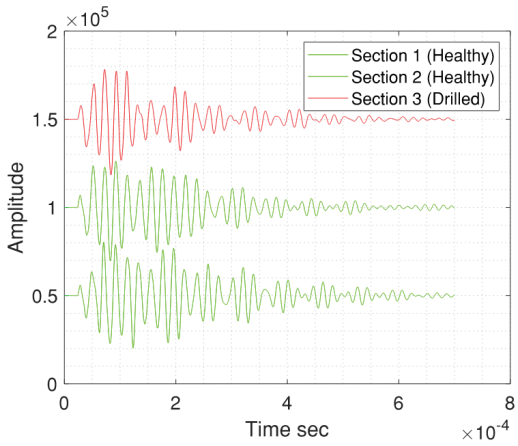

Figure 11: *Sample of time signals measured for each section in Case Study 2. The vertical axis represents signal energy, although each signal has been translated vertically for clarity. The signals are presented in order, with Section 1 featured as the lowermost and Section 3 as the uppermost signal.*

Similar to the observation made in Case Study 1, the distinction between the P-wave component of the signals and the surface waves is not evident from visual inspection and hence, calibration of the signal window for the output of UMAP is required, just as was done in Case Study 1.

3.2.3 Wavelet analysis results and discussion

All 60 recorded signals were analysed using WPD, with both 2 and 3 levels of decomposition trialled, and the 'db2' and 'db4' mother wavelet applied – the identical analysis approach adopted in Case Study 1.

2-level decomposition

Figure 12 provides a depiction of a concise plot comparing the wave energies of all 60 decomposed signals of all 3 sections within pairs of frequency bands to visualise their dependencies, similar to that depicted in Case Study 1. The outcome of the analysis shows significant overlap between the signal energies of the defective section, shown in red, and the healthy sections shown in green, in the off-diagonal scatter plots. The plots comparing wave energies in frequency band pair (2,2 – 2,1) show the strongest positive correlation, with no discernible separation between signals of healthy sections and that of the defective section. The plots comparing frequency band pairs (2,1 – 2,0), and (2,2 – 2,0) show some promising separation, but notably, it can be observed that there is still some overlap between the signals – an unfavourable outcome. This outcome is akin to that demonstrated in Case Study 1.

The same analysis was conducted with the 'db4' mother wavelet, the final scatter plot is shown in Figure 13. Of note, the strong positive correlation of frequency band pairs (2,2 – 2,1) which was evident with the 'db2' mother

wavelet analysis is missing with the 'db4' mother wavelet.

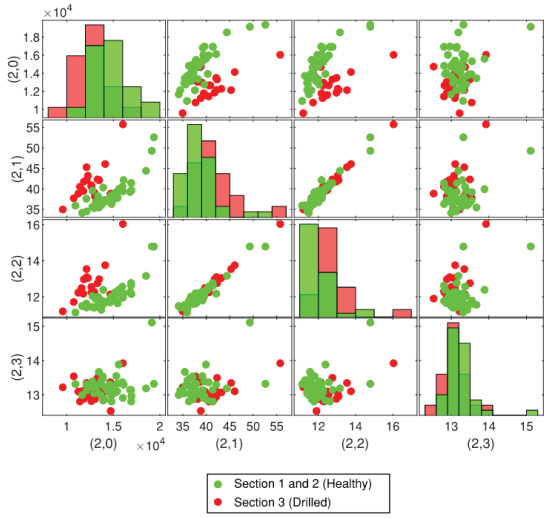

Figure 12: *Case Study 2 - Scatter plot of the signal energies in 4 frequency bands - decomposition level 2 and 'db2' mother wavelet.*

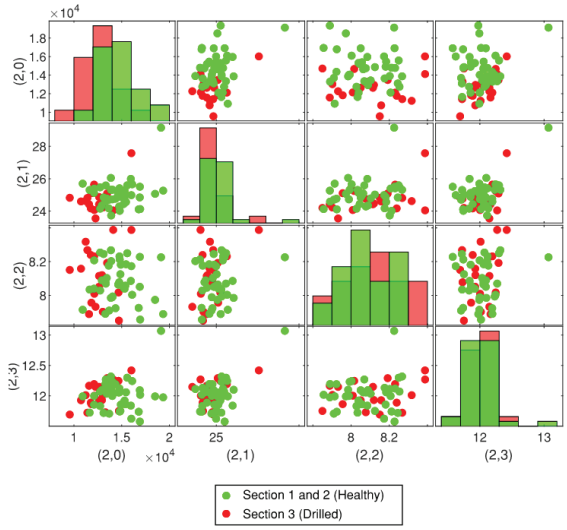

Figure 13: *Case Study 2 - Scatter plot of the signal energies in 4 frequency bands - decomposition level 2 and 'db4' mother wavelet.*

The wave energies are more scattered in the plots compared to that shown in Figure 12, signifying a reduced correlation between the frequency band pairs when 'db4' is adopted, relative to 'db2'. The outcome of the analysis, however, is still not favourable, with no high-confidence separation in the data seen in the bivariate relationship comparisons in the off-diagonal plots, or the univariate analyses depicted on the diagonal plots.

3-level decomposition

Just as was conducted in Case Study 1, 3-level decompositions can also be carried out with the 60 measured signals. As depicted in Figure 1, the 3-level decomposition yields 8 frequency bands ranging from (3,0) to (3,7) – amounting to 56 possible frequency band pair comparisons – shown on the off-diagonal plots, and 8 plots visualising the distribution of wave energies within a given frequency band – shown on the diagonal plots. Figure 14 depicts this analysis with the use of the 'db2' mother wavelet and Figure 15 features that conducted with the 'db4' mother wavelet. In the former, several of the frequency band pairs – (3,2 – 3,1), (3,4 – 3,1), (3,4 – 3,2) – show a strong positive correlation, with high levels of overlap between signals of healthy sections and that of the defective section. This feature is missing from the 'db4' analysis, an artefact also found in the 2-level decomposition described in the preceding section.

Similar to that shown in Case Study 1, both the 'db2' and 'db4' analyses yield an ultimately unfavourable result, with no high-confidence separation of the wave energies between healthy and defective sections, as intended.

In summary, all of the trialled hyperparameter variations in the study – levels of decomposition and mother wavelet type – were unable to yield a desirable outcome. Although some separation of signals was observed in certain cases, as established in the preceding discussion, the lack of complete, high-confidence separation of signal clusters points to an inability of the adopted WPD methodology and trialled hyperparameters to accurately serve the purpose of the detection of a defect of this nature in timber.

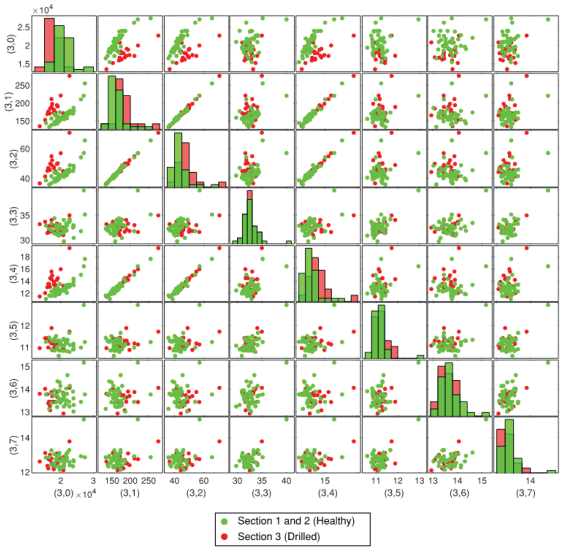

Figure 14: *Case Study 2 - Scatter plot of the signal energies in 4 frequency bands - decomposition level 3 and 'db2' mother wavelet.*

https://doi.org/10.52202/069179-0109

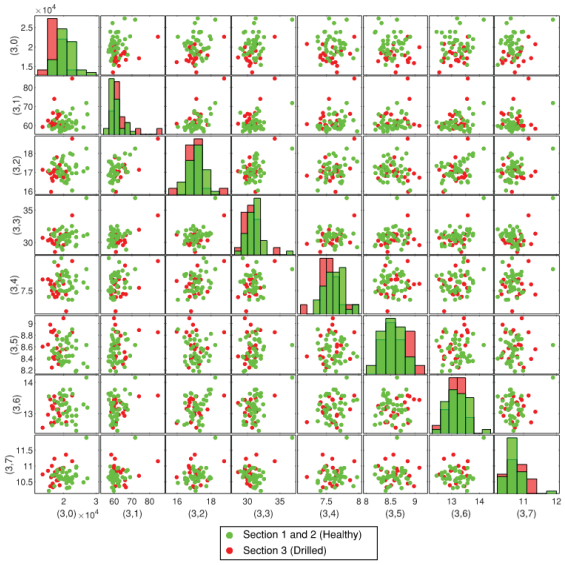

Figure 15: *Case Study 2 - Scatter plot of the signal energies in 4 frequency bands - decomposition level 3 and 'db4' mother wavelet.*

3.2.4 UMAP analysis and results

All 60 signals were thereafter analysed with UMAP, similar to Case Study 1. The hyperparameters used in the analysis were similar to that used in Case Study 1, with minimum distance being set to a value of 0.07 and k-nearest neighbours being set to 15. It should be noted that just as in Case Study 1, the values of both minimum distance and k-nearest neighbours had minimal impact on the output graph, primarily altering the density of the clustering of the signals and influencing the cluster separation. The Chebyshev distance model was adopted for the analysis. The 2-dimensional output graph is shown in Figure 16, with the associated data topology shown in Figure 17.

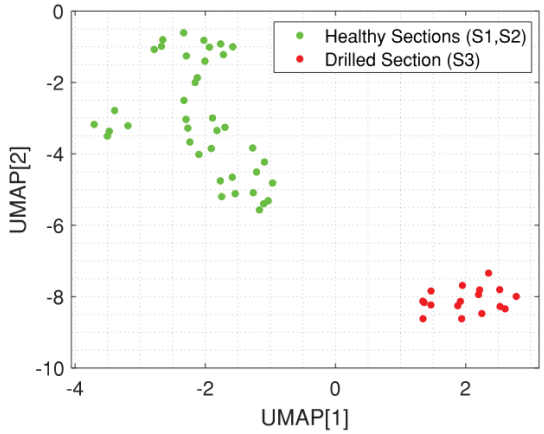

Figure 16: *Case Study 2 - UMAP two-dimensional (2D) reduction output.*

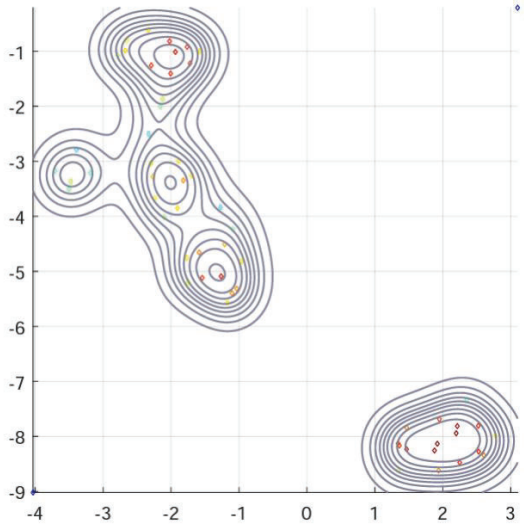

Figure 17: *Case Study 2 - UMAP two-dimensional (2D) cluster topology.*

As depicted, the output of the UMAP analysis is highly favourable, with complete and distinct separation between signals sent through the defective Section 3, and that sent through healthy Sections 1 and 2. Signals from Sections 1 and 2 are treated as being part of the same cluster, as shown in the cluster topology in Figure 17. This positive outcome is congruent to that found in Case Study 1.

4 CONCLUSIONS

This study attempts for the first time to investigate the effectiveness of a novel feature extraction algorithm for the purpose of identifying defects within the timber. UMAP proved to be very effective in separating signals from damaged and healthy sections in timber, requiring little calibration, and providing an easily visualisable output graph. UMAP was versatile in performing successfully under differing defect types in timber. In the first case study, UMAP was able to successfully distinguish between healthy sections and those with naturally occurring defects such as grain deviation and knots. A key positive attribute of the algorithm is its ability to separate the signals of sections featuring differing defect types as opposed to merely clustering them together – providing further insight into the different material characteristics inherent in sections of starkly different grain orientation and fibre make-up. In the second case study, the algorithm proved to be able to successfully distinguish signals transmitted through healthy sections and a defective section fabricated with a void in the centre of the cross-section, meant to simulate material loss and fibre discontinuity caused by defects such as fungal decay, termite damage or splitting damage. The study also juxtaposes the highly favourable outputs of UMAP against competing wavelet-based frameworks, proving the former's robustness and applicability for the intended purpose proposed in the study, compared to the latter. The low sensitivity of the final output of UMAP

due to changes in its hyperparameters, and the versatility of the algorithm in being able to identify various defect types in timber demonstrates the capability of the algorithm for the application of defect identification in timber.

This study acts as an initial proof-of-concept for the viability of UMAP as a signal analysis tool for defect identification in timber. The positive demonstration of the algorithm can hence open the door to future works investigating in-depth the efficacy of the algorithm in more complex, built-up sections of engineered timber, wherein the occurrences of defects are not visible from the exterior of the element. These investigations are currently in progress by the authors.

ACKNOWLEDGEMENT

This work was supported by an Australian Government Research Training Program Scholarship. The assistance provided by Mr Chan Chun-Yen, Mr Jonathan Lui and Ms Benjaporn Samleerangkul is greatly appreciated.

REFERENCES

[1] L. Kia, "Development of Expressive Timber-Steel Hybrid Exoskeletal Systems for Tall Timber Structures," thesis, 2022.
[2] R. H. White and R. J. Ross, Wood and Timber Condition Assessment Manuel. US Forest Service.
[3] W. Ting, Y. Guo-zheng, Y. Bang-hua, and S. Hong, "EEG feature extraction based on wavelet packet decomposition for brain computer interface," Measurement, vol. 41, no. 6, pp. 618–625, 2008.
[4] J. Li, M. Subhani, and B. Samali, "Determination of embedment depth of timber poles and piles using wavelet transform," Advances in Structural Engineering, vol. 15, no. 5, pp. 759–770, 2012.
[5] M. Subhani, J. Li, B. Samali, and N. Yan, "Determination of the embedded lengths of electricity timber poles utilising flexural wave generated from impacts," Australian Journal of Structural Engineering, vol. 14, no. 1, 2013.
[6] L. McInnes, J. Healy, N. Saul, and L. Großberger, "UMAP: Uniform manifold approximation and projection," Journal of Open Source Software, vol. 3, no. 29, p. 861, 2018.
[7] R. S. Stanković and B. J. Falkowski, "The Haar wavelet transform: Its status and achievements," Computers & Electrical Engineering, vol. 29, no. 1, pp. 25–44, 2003.
[8] C. É. Vonesch, T. Blu, and M. Unser, "Generalized daubechies wavelet families," IEEE Transactions on Signal Processing, vol. 55, no. 9, pp. 4415–4429, 2007.
[9] W. Dahmen and C. A. Micchelli, "Biorthogonal wavelet expansions," Constructive Approximation, vol. 13, no. 3, pp. 293–328, 1997.
[10] A. K. Yadav, R. Roy, A. P. Kumar, C. S. Kumar, and S. K. Dhakad, "De-noising of ultrasound image using discrete wavelet transform by symlet wavelet and filters," 2015 International Conference on Advances in Computing, Communications and Informatics (ICACCI), 2015.
[11] S. Dasgupta and Y. Freund, "Random projection trees and low dimensional manifolds," Proceedings of the fortieth annual ACM symposium on Theory of computing, 2008.
[12] U. Dackermann, K. Crews, B. Kasal, J. Li, M. Riggio, F. Rinn, and T. Tannert, "In situ assessment of structural timber using stress-wave measurements," Materials and Structures, vol. 47, no. 5, pp. 787–803, 2013.
[13] X. Wang, F. Divos, C. Pilon, B. K. Brashaw, R. J. Ross, and R. F. Pellerin, "Assessment of decay in standing timber using Stress Wave Timing Nondestructive Evaluation Tools : A guide for use and interpretation," 2004.
[14] H. Zhao, A. Gachagan, G. Dobie, and T. Lardner, "Wavelet analysis of poorly-focused ultrasonic signal of pressure tube inspection in nuclear industry," AIP Conference Proceedings, 2018.
[15] Z. Li, Q. Ding, and W. Zhang, "A comparative study of different distances for similarity estimation," Communications in Computer and Information Science, pp. 483–488, 2011.
[16] S. K. Prabhakar and H. Rajaguru, "Code converters with city block distance measures for classifying epilepsy from EEG signals," Procedia Computer Science, vol. 87, pp. 5–11, 2016.
[17] G. J. McLachlan, "Mahalanobis distance," Resonance, vol. 4, no. 6, pp. 20–26, 1999.
[18] R. Coghetto, "Chebyshev distance," Formalized Mathematics, vol. 24, no. 2, pp. 121–141, 2016.
[19] M. Riggio, R. W. Anthony, F. Augelli, B. Kasal, T. Lechner, W. Muller, and T. Tannert, "In situ assessment of structural timber using non-destructive techniques," Materials and Structures, vol. 47, no. 5, pp. 749–766, 2013.
[20] Y. J. Cai, W. Zhang, and W. P. Zhang, "Experimental study on eccentric compression behavior of timber columns with longitudinal cracks strengthened by CFRP sheets," Advanced Materials Research, vol. 446-449, pp. 3132–3136, 2012.
[21] H. Mitsui, K. Mizutani, and N. Wakatsuki, "Defect detection in square billet using time-of-flight of longitudinal waves," Japanese Journal of Applied Physics, vol. 49, no. 7S, 2010.

https://doi.org/10.52202/069179-0109