

A Development Strategy for Structural Health Monitoring Applications¹

Peter Cawley

Imperial College London,
South Kensington,
London SW7 2BX, UK
e-mail: p.cawley@imperial.ac.uk

Permanently installed structural health monitoring (SHM) systems are now a viable alternative to traditional periodic inspection (nondestructive testing (NDT)). However, their industrial use is limited, and this article reviews the steps required in developing practical SHM systems. The transducers used in SHM are fixed in location, whereas in NDT, they are generally scanned. The aim is to reach similar performance with high temporal frequency, low spatial frequency SHM data to that achievable with conventional high spatial frequency, and low temporal frequency NDT inspections. It is shown that this can be done via change tracking algorithms such as the generalized likelihood ratio (GLR), but this depends on the input data being normally distributed, which can only be achieved if signal changes due to variations in the operating conditions are satisfactorily compensated; there has been much recent progress on this topic, and this is reviewed. Since SHM systems can generate large volumes of data, it is essential to convert the data to actionable information, and this step must be addressed in the SHM system design. It is also essential to validate the performance of installed SHM systems, and a methodology analogous to the model-assisted probability of detection (POD) (MAPOD) scheme used in NDT has been proposed. This uses measurements obtained from the SHM system installed on a typical undamaged structure to capture signal changes due to environmental and other effects and to superpose the signal due to damage growth obtained from finite element predictions. There is a substantial research agenda to support the wider adoption of SHM, and this is discussed in this study. [DOI: 10.1115/1.4051974]

Keywords: continuous and periodic condition assessment, diagnostic decision support, on-line diagnostic approaches, ultrasonics

1 Introduction

The replacement of periodic nondestructive testing (NDT) by permanently installed structural health monitoring (SHM) systems has been discussed for many years, and there is a vast literature on SHM [1]. However, the industrial take-up of SHM technology has been slow, with very few widely deployed applications. This contrasts with rotating machine condition monitoring [2] that is very well established in many sectors after initial issues with false calls were overcome. Most machine condition monitoring applications use passive measurements of vibration or oil debris, whereas monitoring of nonrotating structures usually requires active measurements involving both transmitting and receiving transducers. The two fields are therefore qualitatively different although they are both often termed SHM [3].

The slow take-up of structural health monitoring technology is partly a result of organizational and business case issues [1], but there are also significant technical problems that make deployment and reliable damage detection and quantification difficult and so impact the business case. This article reviews the technical factors hindering wider substitution of NDT by SHM and proposes a system design process, together with a methodology for performance validation.

Most NDT applications involve scanning a transducer or array over the region of structure to be inspected. The scan can be manual or automated, and the results are usually interpreted on

site by a qualified technician; inspection is typically carried out at widely spaced time intervals, often at plant shutdowns. In contrast, scanning is not generally practical in SHM, the transducers being permanently fixed in position. It is not usually economically feasible to deploy an array of transducers at the scan pitch used in NDT, so in SHM, the transducers are usually relatively sparsely deployed over the structure, but frequent measurements can be taken. Therefore, in NDT, measurements are taken at high spatial frequency and low temporal frequency, whereas in SHM, they are taken at low spatial frequency and high temporal frequency.

Unless the probable damage locations are precisely defined and identified in advance, or the degradation will affect a large area so a few sample points are sufficient to give a reliable estimate of its severity, then successful SHM requires an area monitoring capability; it is unlikely to be practical to cover the structure with point sensors, so a method that gives significant area coverage per sensor is needed. This makes the use of ultrasonic guided waves very attractive, and there is a huge literature on their use in SHM [1]. This article mainly focuses on example applications using guided waves, and the application of bulk ultrasonic waves to thick structures is also discussed. However, the methodology is equally applicable to other techniques that may be attractive for some SHM applications such as potential drop [4–8], current deflection [9], and magnetic measurements [10–12].

If the potential of SHM is to be realized widely in industry, a strategic approach to its development is required. This involves:

- (1) Analysis of the business case and selection of applications where the case is strong; development of an SHM system requires considerable design, development, and validation effort. This can be justified by a large volume of applications

¹Paper presented at the ASME 47th Annual Review of Progress in QNDE, Virtual Conference, August 25–26, 2020.

Manuscript received October 19, 2020; final manuscript received December 23, 2020; published online August 27, 2021. Assoc. Editor: Paul Fromme.

such as corrosion monitoring in the oil and gas industry² or very high value cases such as in nuclear plant [13].

- (2) Careful analysis of the design requirements: what defects must be detected, under what conditions, with what constraints; who will interpret the data; what decisions are to be made from it; and what are the costs of incorrect indications.
- (3) Design of the transduction and instrumentation systems to survive the operating environment and have appropriate stability, life (including battery if used), and connectivity.
- (4) Design of the data processing system including compensation for environmental changes and conversion to usable information for the recipient; consideration of the scale of deployment is crucial here as a human operator can deal with reviewing raw data from one or two systems, but automation is increasingly essential when the number of deployments reaches 100 s and 1000 s.
- (5) Validation of performance: many applications with a strong business case will be on safety critical structures, so evidence of the efficacy of the system must be provided to the structure operator and regulators.

This article concentrates on aforementioned points (2), (4), and (5); Sec. 2 discusses SHM system design requirements and Sec. 3 then reviews the possibility of exploiting the high temporal data frequency provided by SHM to overcome the loss of defect sensitivity implied by its low spatial frequency. The use of receiver operating characteristic (ROC) curves as a performance metric and a proposed performance validation methodology are discussed in Sec. 4. Section 5 discusses the remaining research requirements and presents the conclusions of the paper.

2 System Design Requirements

When considering implementation of an SHM system rather than a traditional periodic inspection approach, it is vital to consider the design of the whole system not just, for example, the transduction. Todd and Flynn [14] propose a structured design approach for SHM systems, starting with key questions:

- (1) What are the failure modes the system is being designed to monitor, and, to whatever degree possible, what are their expected probabilities of occurrence? This is vital as not only it is imperative that the system detect all critical forms of damage at all likely locations but also it is not overdesigned and so unnecessarily expensive.
- (2) What specific actions will the SHM system direct in response to the failure mode(s)? This forces consideration of the decision-making process and who will be authorizing actions, e.g., mandating the shutdown of a plant.
- (3) What are the costs associated with the deployment and operation of the SHM system and the actions/decisions that the SHM system directs. This highlights the need to identify the capital and revenue costs of both the SHM system itself and the decision-making process.
- (4) What are the constraints present in the design space? For example, temperature range, intrinsic safety requirements, sensor and instrumentation mass, feasibility of cabling versus wireless operation, and availability of power.

Todd and Flynn [14] then go on to use a Bayesian approach to the design of the system, further details on an optimal sensor placement study being given in Refs. [15,16].

The area of structure that must be monitored has significant implications for the cost of an SHM system and the choice of measurement technique. For example, consider simple 20 mm deep, 32 mm wide, and 150 mm long beams and their corresponding stress distributions in three-point and four-point

loading, as shown in Figs. 1(a) and 1(b). The corresponding probabilities of cracks occurring at different locations in the two cases are shown in Figs. 1(c) and 1(d), respectively, these being obtained using the Weibull weakest link theory [17] as explained in Ref. [18], where the probabilities are quantified; here, the concern is to illustrate the qualitative differences, so the scales are simply low high. Figure 2 shows two candidate monitoring methods, a fixed 25 mm diameter, 45 deg ultrasonic shear wave probe in Fig. 2(a), and a potential drop system in Fig. 2(b). The probabilities of detecting a 6 mm × 6 mm vertical crack at different locations of the beam using the two techniques are shown in Figs. 2(c) and 2(d), respectively; these were estimated using finite element analysis with assumptions about the signal-to-noise ratio taken from experiments, as described in Ref. [18]. Again, the results shown here are qualitative as the intention is to illustrate the issues involved in technique selection. As would be expected, the ultrasonic system gives very high sensitivity at the middle of the beam, but this drops away severely toward the ends of the beam and also at the edges due to the beam being wider (32 mm) than the transducer (25 mm diameter). In contrast, the potential drop system gives relatively uniform coverage over the beam, but the peak sensitivity is lower than with the ultrasonic system. The sensitivity and probability of defect occurrence at different locations maps of Figs. 1 and 2 can be combined to give overall probability of detection (POD) values as presented in Table 1. (The values given are for a very low probability of false alarm (PFA), but the relative performance is replicated at other PFA values.) This shows that the single-point ultrasonic test cannot be used reliably with the four-point bend loading, but it performs little better than the potential drop system for the three-point loading case; this is because of the reduction in performance at the edges of the beam shown in Fig. 2(c). This demonstrates that unless the likely defect location is known very precisely, or it is acceptable to cover the structure with transducers, wide area and less-sensitive monitoring systems are likely to be preferable to the very sensitive, highly localized systems that are more often used in NDT. The difference is that in NDT, the transducer can be scanned so that the high sensitivity is achieved at all scan locations. The question in SHM is whether obtaining frequent readings in time from a low sensitivity, wide area system can reproduce the performance obtained by scanning at infrequent intervals in NDT; this is addressed in the next section.

An idea of the probability of damage occurrence is important as it will be a key factor in determining an acceptable false call rate. If damage is very unlikely to occur, then a very low false call rate will be required to ensure that there is a reasonable probability that a positive damage call correctly indicates that damage is present; this is a very well-known issue in medical screening and is related to the base rate fallacy [19]. For example, suppose that the probability of false call (PFA) in an inspection is 10% and the POD of a real defect is 90%, and further suppose that the a priori probability of a callable defect being present is 5%. Then, in 1000 inspections, we expect 95 false calls, 45 correct defect calls, and 5 missed defects; therefore of 140 defect calls, ~70% are false. If the a priori probability of a callable defect being present is reduced to 1%, then more than 90% of defect calls will be false. It is essential that the actions to be taken in the event of a defect call are clearly specified so that their cost is known, and so an acceptable false call rate can be defined. If the cost of investigating each call is high, then a very low false call rate will be required, while maintaining the required POD; this puts severe demands on the SHM system design.

Much in-service NDT is performed during plant shutdowns, so the temperature is close to ambient and the structures are not significantly loaded; this means that standard transducers and instrumentation can be used. In contrast, an SHM system is permanently attached to the monitored structure so the transducers and instrumentation must survive the operational environment; it is sometimes decided only to take readings with the plant off load, but

²<http://www.permasense.com>

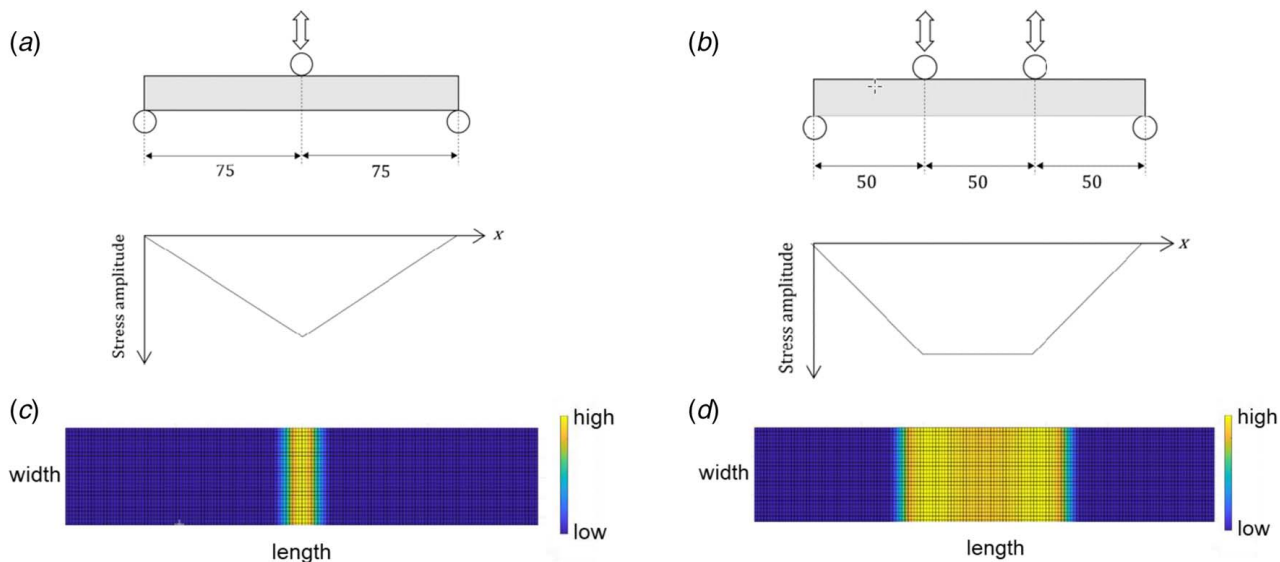


Fig. 1 Schematic (a) three-point and (b) four-point loading and stress distribution on a beam; corresponding probabilities of defect occurrence: (c) three point and (d) four point (adapted from Ref. [18])

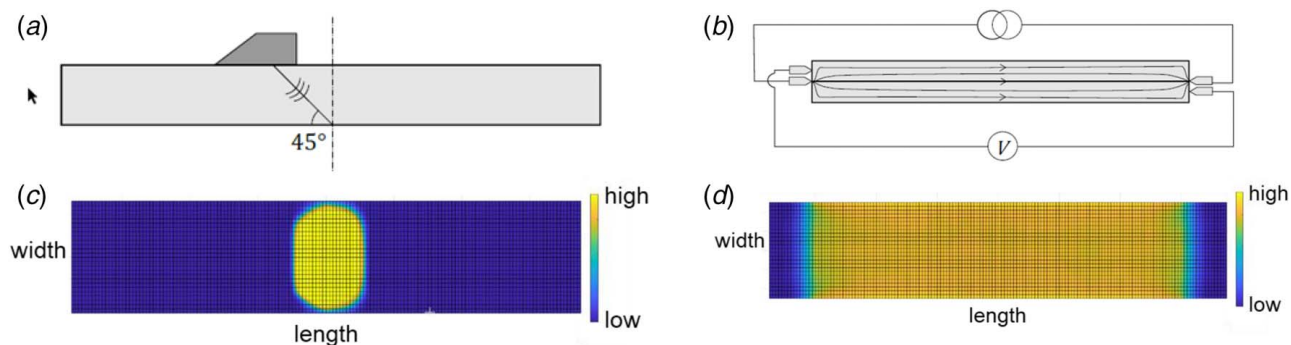


Fig. 2 Schematic (a) ultrasonic corner-echo and (b) potential drop monitoring systems on beams of Fig. 1; probabilities of detecting defect if present at each location: (c) ultrasonic corner-echo and (d) potential drop monitoring (adapted from Ref. [18])

the system must still survive the full range of conditions, even if it only has to operate over a narrower range. Design of an SHM system is therefore often much more demanding than in NDT. For example, Fig. 3 shows a high-temperature ultrasonic thickness monitoring system³ [20] in which the transducers and instrumentation are outside the insulation over the high-temperature structure, the ultrasonic waves being sent and received along stainless steel strip waveguides, the resultant signals being transmitted to the control room via a wireless link. Since the system was designed to operate in zoned areas of oil and gas plant, it also had to meet intrinsic safety requirements that, inter alia, limit the voltage and current that can be used.

3 Exploiting Frequent Data

3.1 Background. As discussed earlier, SHM will become much more attractive if wide coverage can be obtained from each transducer position, but since wide area systems are typically less sensitive than the point inspection systems that are scanned in NDT, they can be employed only if the lost sensitivity can be recovered via processing the frequent, time series data that can be obtained.

³See Note 2.

Table 1 Expected probabilities of detection for 6 mm × 6 mm crack at 10⁻⁶ probability of false alarm in three-point and four-point loading with the monitoring systems of Fig. 2

	Expected POD ultrasonic sensor (%)	Expected POD potential drop sensor (%)
Four-point bend	32.7	90.5
Three-point bend	92.5	90.6

Note: Adapted from Ref. [18].

Farrar and Worden [3] identify two broad methodologies for processing the data: the inverse problem and pattern recognition approaches. In this context, the inverse problem approach usually adopts a physics-based model of the structure and tries to relate changes in the measurements to changes in the model; this is strongly related to the more recently termed digital twin concept [21,22] that in the civil engineering field has been named structural identification [23,24]. It is very challenging to invert a large model to match measured signal changes to the structural changes (in this case damage growth) that produce them although there has been significant progress in the full wave inversion of sonic and

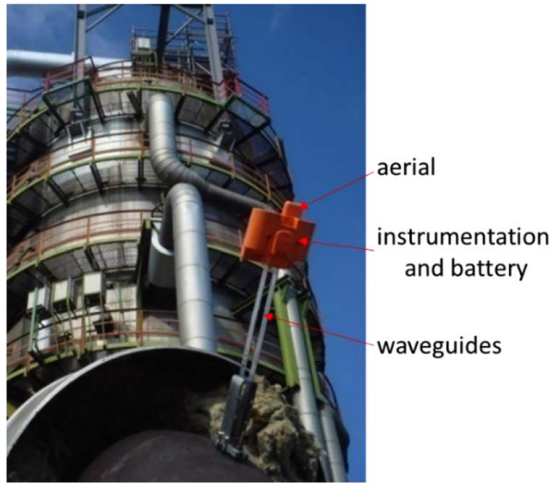


Fig. 3 High-temperature ultrasonic thickness monitoring system [20]

ultrasonic signals, starting in the geophysics field [25]. The pattern recognition approach is therefore more commonly followed in machine condition monitoring and SHM although once change has been identified, located, and its magnitude estimated, it is then possible to investigate the size of the structural change that has produced it using damage models; this is discussed further in Sec. 4.

In this context, pattern recognition becomes a change-point detection [26,27] or trend identification problem [28], and these have been widely studied in other fields such as machine condition monitoring [3,2], medical monitoring [29], and finance [30]. Many rotating machines operate in protected environments and under fairly constant conditions that ease the problem [3]. In SHM, it is very common for measurements to be taken under varying environmental conditions, and this must be accounted for. This was recognized in early attempts at bridge monitoring using natural frequency measurements where it was initially found that the measured natural frequency increased with the introduction of damage, whereas it would be expected that damage would decrease the stiffness and so decrease the natural frequencies; this anomaly turned out to be caused by temperature changes between the predamage and post-damage measurements [3].

3.2 Temperature Compensation. This article concentrates on the use of ultrasonic-guided wave measurements for SHM, and there is a vast literature on signal compensation for environmental changes, particularly temperature. Two widely used methods for the temperature compensation of guided wave signals are baseline signal stretch (BSS) and optimal baseline selection (OBS). BSS requires only one baseline signal that is used as a reference to which any current measurement is compressed or stretched in time to minimize the residuals [31,32], so compensating for velocity and dimension changes with temperature. In OBS, a set of baseline signals is stored so that the one deemed most similar to the specific current measurement is used for amplitude subtraction, often after also applying BSS on the selected baseline [31,33]. Both these techniques are in principle applicable to bulk wave and guided wave signals.

Unfortunately, the signal phase often changes with temperature with both piezoelectric and electromagnetic acoustic transducer (EMAT) transducer systems [34–37], and Mariani et al. have extended the BSS method to compensate for phase changes [38]. There have also been several attempts to deal with the difficulty of temperature compensation on multimode guided wave signals [39,40]. All these methods seek to minimize the difference between the current and baseline signals, and the function that is

minimized usually is the rms residual when the two signals are subtracted; therefore, they process the whole signal and the residual tends to be dominated by changes in any large reflections that are present. A blind trial of a guided wave pipe monitoring system [41] on the pipe shown schematically in Fig. 4 reveals that it was possible to reliably detect defects about five times smaller than could be detected by interpreting the signal obtained in a single test.⁴

The residual signal amplitude as a function of time at a location where no defect was introduced when processed with the phase compensation (PSC) method of Ref. [38] is shown in Fig. 5(a), along with the temperature of the pipe (this result is slightly superior to that obtained with BSS compensation). It can be seen that the residual varies significantly and that this variation correlates with temperature. To limit false calls, the defect “call” level has to be set above the level of variations in the absence of damage, so these variations limit the improvement in sensitivity that can be obtained. Therefore, it was clear that improved temperature compensation was the key to improving the defect sensitivity obtained in monitoring.

Mariani et al. hypothesized that the remnant temperature dependence was due to interference between different guided wave modes present in the signal, the interference being a function of temperature; this is due to the different modes having different velocity coefficients of temperature and also due to the relative amplitudes of the excited modes being a function of temperature. They developed a new temperature compensation method that denoted location-specific temperature compensation (LSTC) [42]; unlike the BSS, OBS, and other previous methods that consider the whole signal in a single process, LSTC treats each point on the signal, corresponding to a particular location on the test structure separately. The method comprises a calibration phase and a monitoring operation phase. In calibration, a set of baseline signals is acquired across the temperature range of interest, which is used to construct a calibration curve for each signal sample, i.e., each point on the captured waveform. Each curve shows how the expected signal amplitude at each location in the absence of damage varies with temperature. In the monitoring operation phase, when a new measurement is acquired, at each point, the expected value at the relevant temperature obtained from the calibration curves is subtracted from the measurement itself. Thus, in the absence of damage, the expected value of the residual signal is zero. Further details are given in the study by Mariani et al. [42].

Figure 5(b) shows the residual signals corresponding to those of Fig. 5(a) when LSTC compensation is applied following initial PSC compensation. The temperature dependence of the residuals has been removed, and Fig. 5(c) shows that the residual signal is now normally distributed (the data comfortably passed the Lilliefors normality test [43] at a 5% significance level [42]). Figures 6(a) and 6(b) show the results corresponding to those of Figs. 5(a) and 5(b) at the location of defect 2 shown in Fig. 4. The solid curve labeled defect 2 shows the history of defect introduction that started after about 330 measurements and grew to a cross section loss of about 1.8% (note that the amplitude scales on Figs 5 and 6 are different). The residual signal of Fig. 6(a) again correlates with temperature, making reliable defect calling difficult until the defect has grown to above about 1% cross section loss, whereas it would be possible to call the defect reliably much earlier from the residual signal of Fig. 6(b) that remains low with zero mean before the defect introduction, and then tracks its growth; the final residual signal level expressed as a percentage of the end reflection before damage was grown is slightly above the percentage cross section loss since the reflected signal is a function of the defect shape and the section loss [44].

The defect signal amplitude, and hence defect size, that can be reliably detected at a given false call rate is determined by the variability in the signal obtained from an undamaged structure. If the signal from an undamaged structure is normally distributed, as

⁴<http://www.guided-ultrasonics.com>

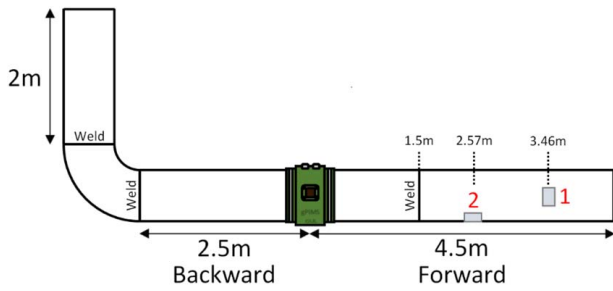


Fig. 4 Geometry of the 8 in. schedule 40 test pipe used in blind trial [41]. The 7 and 2 m pipe sections are connected by a 90 deg elbow with the bend radius being 1.5 times the outer diameter of the pipe. The defects introduced in the “forward” direction are highlighted and numbered. The distances given of the defects and weld in the forward direction are measured from the middle of the sensor produced by Guided Ultrasonics Ltd.

shown in Fig. 5(b), then the probabilities of detection and false alarm are a function of the ratio of the defect signal amplitude to the standard deviation of the signal from an undamaged structure; this is discussed in more detail in Ref. [45].

3.3 Data Analysis. In most conventional inspections, the signals are analyzed on site by the inspection technician who interprets the data and will often carry out any necessary follow-up inspection to confirm a defect call or to size the indication more accurately. Therefore, those responsible for the integrity management of the structure are only alerted when there is an abnormality. In contrast, SHM data are generated automatically, and it is typically much more frequent than NDT inspections (e.g., daily rather than annual) and is transmitted direct to the structure operator. As the number of monitoring locations on a structure or plant increases, this data stream can become unmanageable unless some automatic preprocessing is applied. When the thickness monitoring system⁵ [21] was first deployed at 100 s of locations on a plant, operators described the experience of frequent, multipoint data as being like “drinking from a hosepipe.” Goulet and Smith [46] note that with increasing availability of communication systems and decreasing cost of sensors, more and more structures are being measured, but our capacity to analyze large amounts of data is only marginally increasing.

The Cambridge Centre for Smart Infrastructure and Construction (CSIC) stresses the importance of data handling and interpretation and present a pyramid model [47], a simplified version of which is shown in Fig. 7. Here, data are processed to a digestible form before being passed to the decision-making level, different degrees of automation being possible at each level.

While the number of sensors on a plant remains modest, it may be sufficient simply to compress the information displayed to the operator in time and space. For example, the display used in the thickness monitoring system⁶ [21] was initially a single-point A-scan as shown in Fig. 8(a) and then graduated to a thickness versus time plot as shown in the uncompensated graph of Fig. 8(b). This showed sudden jumps in thickness that caused confusion for operators until it was understood that they were due to temperature excursions changing the speed of sound; this led to a thermocouple being incorporated in the transducer of Fig. 3, which enabled the temperature compensated thickness versus time plot of Fig. 8(b) to be produced. In turn, this enabled reliable corrosion rates to be computed, and the different rates shown in Fig. 8(b) proved to correlate with feedstock changes in the plant. Figure 8(c) shows the next level of data compression with corrosion rates being displayed at all locations in the plant, enabling the operator to see at a glance

⁵See Note 2.

⁶See Note 2.

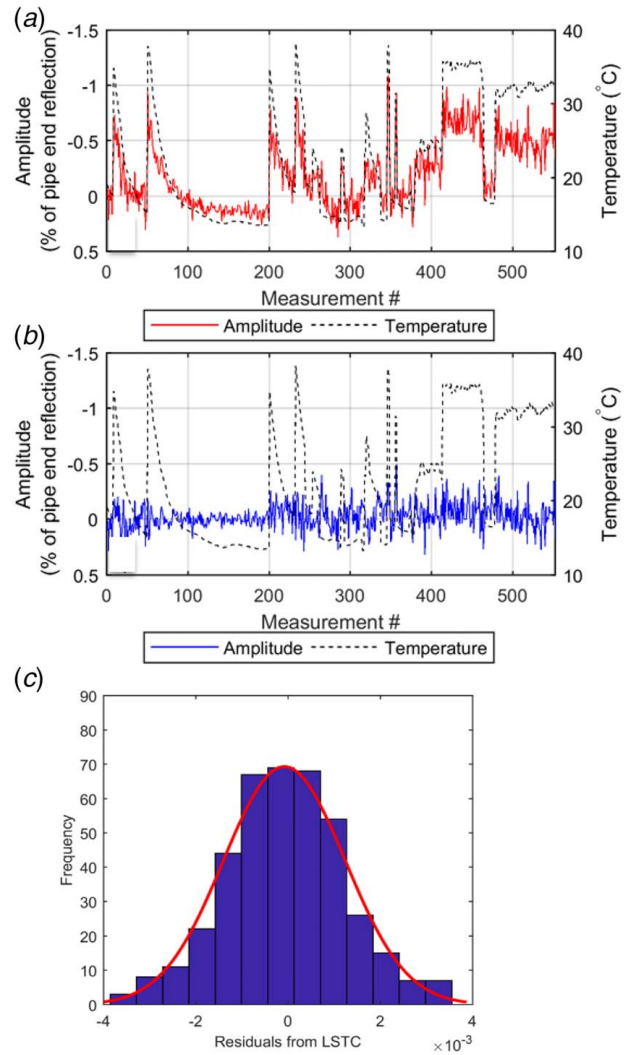


Fig. 5 Residual signal and temperature as a function of measurement number (time) at 1.45 m location on pipe of Fig. 4, where no defect was grown when residuals obtained using method of (a) Ref. [38] and (b) Ref. [38] followed by LSTC method [42]. (c) Distribution of residual signals of (b) with normal curve fit shown.

where rates are highest and also to see commonality between different locations that is also valuable information not readily seen from individual plots.

3.4 Automated Analysis and Change Detection. As data volumes increase, it is essential to increase the automation of processing, a reasonable target to keep the absolute volume of data requiring manual processing constant with time. Therefore, the fraction of data requiring manual intervention decreases in proportion to the number of sensors installed. Gradually increasing automation in this way ensures that the development team doing the manual processing sees a large number of cases with their associated difficulties, so informing the development of automated processing. This also gives a significant database of cases that can be used for supervised learning if machine learning algorithms are to be considered; this is discussed further below.

Since the a priori probability of the structure being defective is usually low, the most effective and valuable way to reduce the volume of data to be processed manually is to automatically identify those cases where there is high confidence that no significant defect is present. This usually corresponds to the cases with the cleanest signals, so this is relatively simple to automate. In the thickness

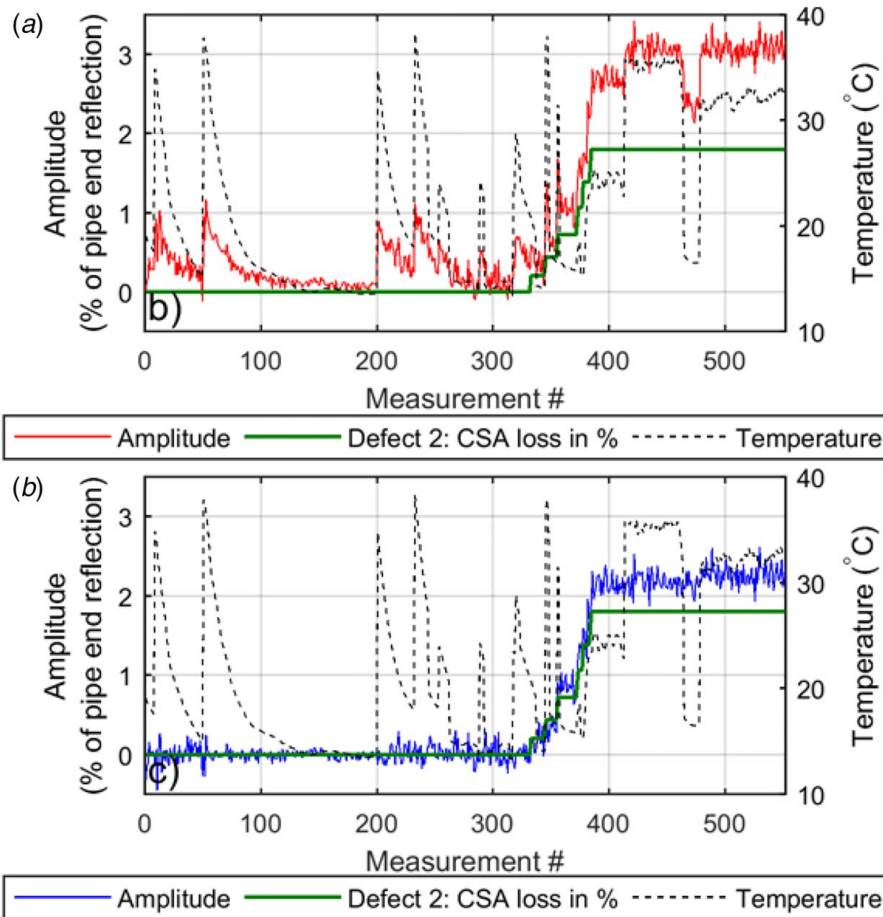


Fig. 6 Residual signal and temperature as a function of measurement number (time) at location of defect 2 on pipe of Fig. 4 when residuals obtained using method of (a) Ref. [38] and (b) Ref. [38] followed by the LSTC method [42]

monitoring example of Fig. 8, this might involve identifying those cases where there is high confidence that the corrosion rate is less than, for example, 0.1 mm/year.

The presence of uncompensated environmental effects such as the excursions with temperature shown in the uncompensated thickness plot of Fig. 8(b) or the residual guided wave signals of

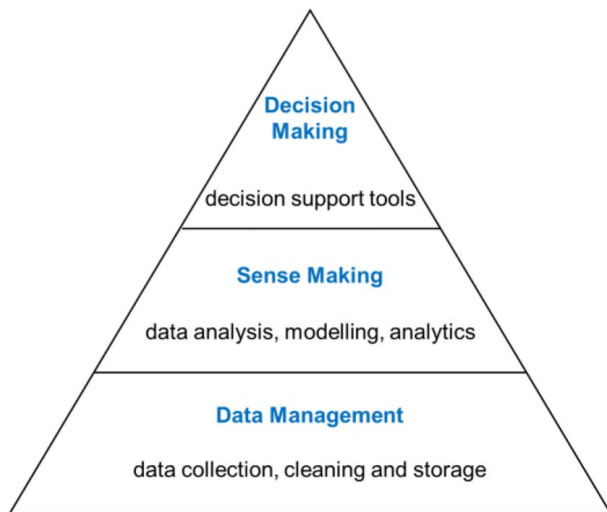


Fig. 7 Pyramid model of decreasing data volume and increasing data value (adapted from Ref. [47])

Figs. 5(a) and 6(a) makes the automated reliable calling of small defects very difficult. However, if the environmental effects are compensated satisfactorily, then automated processing is relatively straightforward as shown by the corrosion rate calculations of Fig. 8(b). Likewise, the normal distribution of the residual guided wave signals after LSTC compensation of Figs. 5(b) and 5(c) makes it possible to apply standard change tracking algorithms that have been developed in the statistical process control (SPC) field.

Over the past century, researchers in SPC have produced a number of methods for the quality control of manufacturing processes that involve the analysis of monitored parameters (e.g., temperature, pressure, humidity) that are typically assumed to follow normal distributions [48,49]. Increasing computer power has resulted in increased attention from SPC researchers to the generalized likelihood ratio (GLR) approach, which was first proposed by Lorden in 1971 [50]. Although the GLR is rather computationally intensive, it is particularly attractive as it does not require the extent or time of change to be specified in advance. Mariani and Cawley [45] tested it on guided wave pipe monitoring data of the type shown in Figs. 5 and 6 and showed that it can enable the reliable detection of defects equivalent to as little as 0.1% cross section loss in cases similar to those of Fig. 6 without a significant number of false calls. This assumes that the sensor is stable with time and that the residual signal in a given structure state remains normally distributed. Nonstationarity of the data coming from an unchanged system is a recognized problem in change-point detection that is still the subject of current research [27]. Therefore, as signal processing is used to drive down the minimum detectable defect size, the demands on sensor stability with time increase so it is necessary

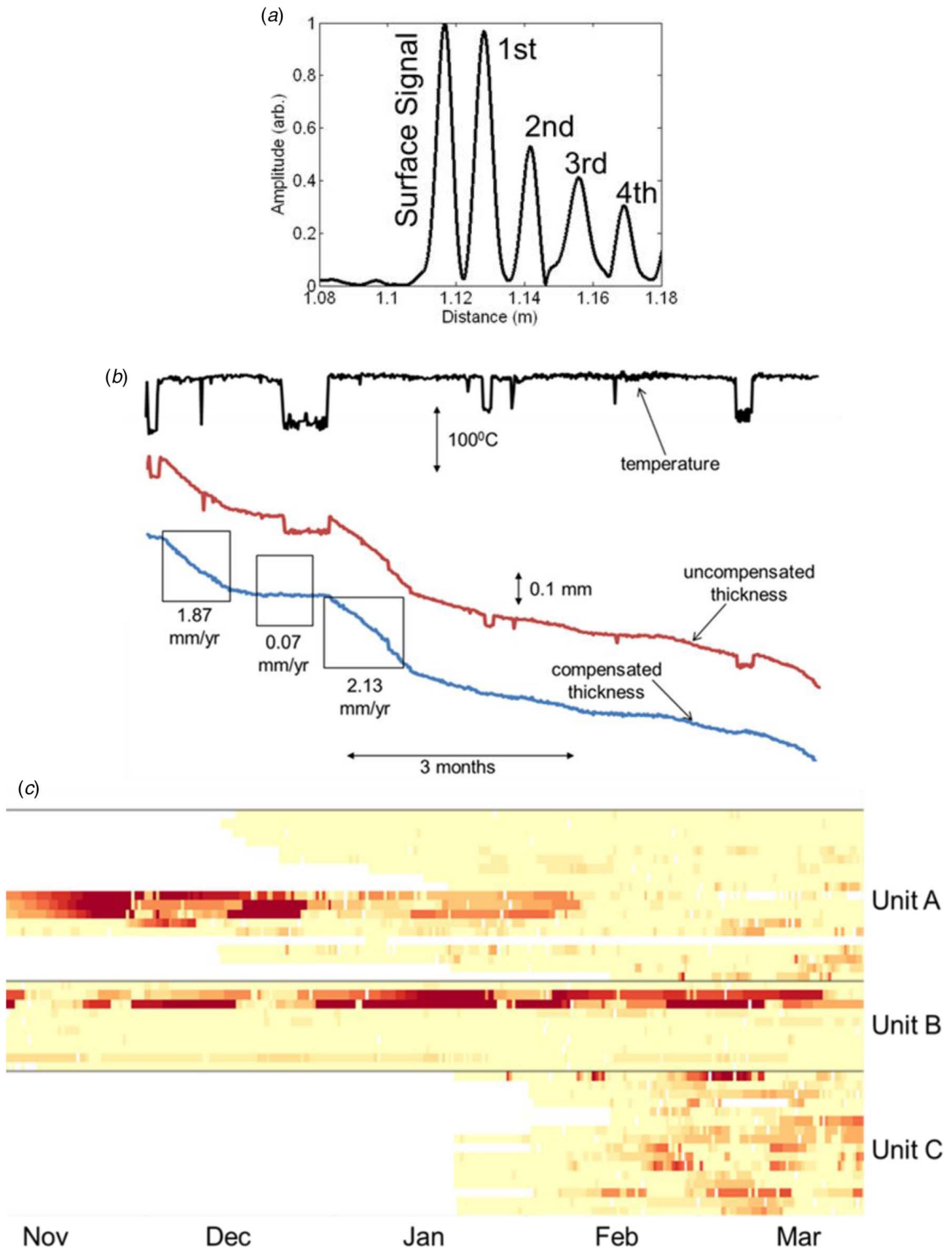


Fig. 8 Increasing compression of data presentation: (a) single-point A-scan, thickness being inferred (automatically) from time between successive peaks and speed of sound, (b) thickness versus time plots uncompensated and compensated for temperature variations, and (c) corrosion rate at all sensor locations on three units on plant (darker color—higher wall loss rate). Sensors on unit C and some on unit A installed later than on unit B.

to pay increased attention to the design of transducers and their attachment to the structure; it is also essential to check the statistics of the signals as part of the automated analysis process.

In some systems, it is possible to employ signal processing methods that eliminate the effect of drift, one example being a bulk wave ultrasound test in which the amplitude of the echo from the back face of the test piece is monitored. In this case, Mariani et al. [51] used a method first proposed by Achenbach et al. [52] in which instead of using the amplitude of the back wall echo as the measure of structural health, the ratio between successive back wall echoes is employed; the ratio is much less sensitive to drift. However, this technique is only applicable in a limited number of cases, so more research is needed on the problem of sensor drift.

An alternative to the aforementioned traditional approaches is to employ supervised machine learning. There has been substantial research interest in applying supervised machine learning, usually in the form of multilayer perceptron (MLP) networks, to NDT signals, see, e.g., [53]. Due to the limited hardware capabilities in early work, shallow MLP networks consisting of a limited number of hidden layers were the preferred choice, and the signals were typically preprocessed to reduce their dimensionality by extracting features deemed to be sensitive to defect reflections, see, e.g., Ref. [54]. However, this work has not found widespread application in NDT, partly due to the need for skillful operators to manually determine defect-sensitive features effective for each specific application.

It would be attractive to feed full time-domain waveforms to the machine learning algorithm, hence using the network itself to learn the defect detection features, rather than relying on operator input; this requires more complex architectures able to model the complexities of ultrasonic wave propagation, especially when dealing with multimodal propagation of guided waves.

In the last decade, with the advent of GPU-based machines, there has been a proliferation of deep learning architectures involving the use of multiple hidden layers. Mariani et al. [55] have reviewed the field and shown that in two guided wave monitoring examples, the WaveNet convolutional neural network originally developed for audio signals [56] was able to learn features and/or patterns related to the presence of waves scattered from damage, thus eliminating the need for any feature input from human operators. The network outperformed the optimal baseline selection and baseline signal stretch compensation methods discussed earlier, and it was especially encouraging that the improvements over the conventional approach were particularly marked when the “current” signals were taken at temperatures well outside the temperature range available in the set of baseline signals. This suggests that this class of network can complement or replace existing methods, especially when testing occurs under new environmental and operational conditions. This is also a possible approach for dealing with the effects of sensor drift, but this requires much more investigation.

The need for very large training data sets makes the application of supervised machine learning in SHM difficult although it is possible to supplement experimental data on damaged structures with data from undamaged structures on which predicted damage signatures are superposed; this is discussed further in the next section. A further concern is whether the training data set covers a wide enough range of cases, and it is also more difficult to qualify a supervised machine learning system for use in safety critical applications than one based on predefined signal processing operations.

4 Performance Validation

Receiver operating characteristic curves [57] are routinely used in the evaluation of diagnostic tests in medicine and engineering; they plot the sensitivity (POD in NDT/SHM terminology) for the test against its specificity (PFA in SHM) for a given change in condition (defect size and type in SHM), both axes being on a (0,1) scale. The ideal operating point is at the top left-hand corner of the plot,

corresponding to unity probability of detection and zero probability of false alarm, but of course this is not practically attainable. The ROC curve enables the following questions to be answered:

- What is the POD at an allowable false call rate of 1% for a given damage size, type, and location and under a specified range of operating conditions?
- What is the smallest damage size that can be detected at 95% POD and 1% false call rate given a specified measurement frequency and range of operating conditions?

For example, Fig. 9 shows ROC curves for a guided wave pipe monitoring system using location-specific temperature compensation (LSTC) for three different levels of cross section loss. They are plotted on a semi-log scale, so that low levels of probability of false alarm can be seen, recognizing that the PFA must be very low for the system to be acceptable, as discussed earlier. The performance improves dramatically as the defect size (cross section loss) increases, being almost perfect at unit POD and <0.01% PFA for 0.36% cross section loss. The dashed line corresponds to the performance expected from a random guess strategy; this would be a straight, 45 deg line on a linear plot

Unfortunately, it is typically even more costly to define ROC curves experimentally in SHM applications than it is to conduct POD trials in NDT. This is because the SHM system must be installed on many structures of the required design that are cycled through the full range of operational conditions, damage being introduced in some of the structures at a representative range of locations. In most cases, this is totally impractical and its cost would preclude the application of the SHM system. In NDT, there is an increasing move toward model-assisted POD (MAPOD) evaluation [58], and Liu et al. [59] have proposed a related approach for ROC curve generation in SHM.

Modern computational resources mean that the signature produced by damage in bulk wave ultrasound, guided wave ultrasound, and other techniques can be reliably predicted, even when it has complex shape [44]. In contrast, the reliable prediction of signal changes due to environmental and other variability is not possible so it is not feasible to predict the signal changes seen in the absence of damage. However, it is straightforward to obtain

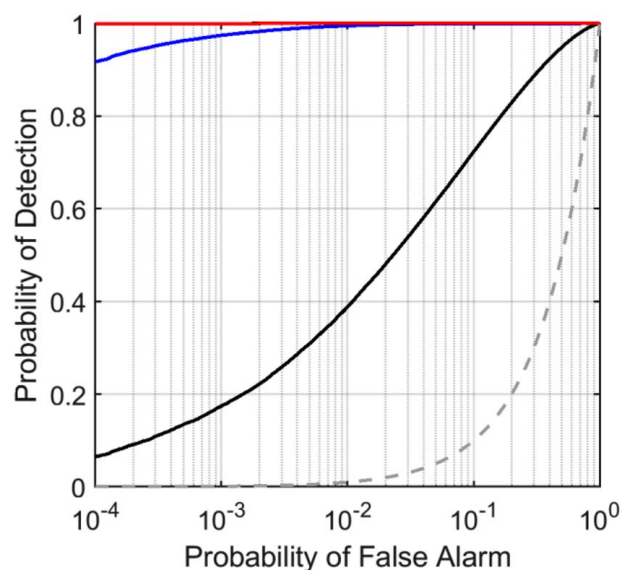


Fig. 9 Example ROC curves plotted on semi-log scale. Case of guided wave pipe monitoring system using LSTC temperature compensation with defects corresponding to 0.12% cross section loss (black – lowest solid curve), 0.24% loss (blue – middle solid curve), and 0.36% loss (red – top solid curve). Data are used to compute Fig. 13(a) of Ref. [45]. Dashed line is random guess that is 45 deg line on a linear plot.

experimental data with environmental variation from an SHM system installed on a typical undamaged structure. Therefore, measured data can be obtained over multiple environmental cycles on an undamaged structure, and predicted damage signatures can be added at different locations with different growth patterns. The effect of different degrees of environmental variation on the ROC curves is then straightforward to simulate by appropriate selection of signal sets, and the effect of varying damage severity, multiple damage sites, frequency of readings, and so on is easy to assess. The approach has been validated on data obtained in the blind trial of a guided wave pipe monitoring system [41] by Heinlein et al. [60].

Crucially, this approach can also be used to validate performance of an installed SHM system. Suppose that the regulator requires assurance that the SHM system installed on a nuclear power plant is correct in indicating that no defect larger than a specified size has grown at a particular location. Signals from defect growth at the specified location can be predicted and added to the sequence of raw signals measured from the installed SHM system; the resulting synthetic signal sequence can then be processed by the algorithms used to determine the health of the system, so verifying whether the required defect would be detected if present. The statistics of the measured signals can also be checked to ensure that any assumptions about, for example, normality of residuals as discussed earlier are valid.

5 Conclusions and Future Research Needs

Permanently installed SHM systems are now a viable alternative to traditional periodic inspection. The SHM system must have high reliability over the envelope of operating conditions, a particular concern being the avoidance of false calls due to, for example, temperature or load changes. Since scanning a transducer over the structure, as is routinely done in NDT, is not possible with the fixed transducers used in SHM, the data obtained in SHM typically have a much lower spatial frequency than that in NDT, and it is desirable to choose methods that enable high volume coverage from each transducer. On the other hand, it is possible to acquire data much more frequently in SHM than in NDT; hence, the aim is to achieve similar performance with low spatial frequency, high temporal frequency SHM data as is obtained in NDT with high spatial frequency, and low temporal frequency information. This can be done via change tracking algorithms such as the GLR, but this depends on the input data being normally distributed, which can be achieved only if signal changes due to variation in the operating conditions are satisfactorily compensated. There has been much progress in this area, and the recently developed LSTC method [42] has shown excellent results on both ultrasonic-guided wave and bulk wave data. There is also increasing interest in the application of supervised deep learning methods to SHM signals.

The transducers and instrumentation used in SHM must survive the operating environment and remain stable over extended periods; this is much more demanding than in NDT where inspection is generally carried out offline. Since SHM systems can generate large volumes of data and often transmit it direct to the engineers responsible for the plant, it is essential to convert the data to actionable information, and this step must be addressed in the SHM system design. It is also essential to validate the performance of installed SHM systems, and a methodology analogous to the scheme used in NDT has been proposed. This involves using measurements obtained from the SHM system installed on a good structure to capture signal changes due to environmental and other effects and to superpose the signal due to damage growth obtained from finite element or other predictions.

The routine industrial use of SHM is in its infancy [1], with high value oil and gas industry applications to the fore. There are many sensors deployed on civil engineering structures, but data processing to provide useful information to operators is far from standardized [46]. There has been a great deal of interest in SHM in the

aerospace sector, and many prototype systems have been developed and flown on a trial basis; however, few systems are in routine use apart from a few local hotspot monitoring systems, particularly on military aircraft.

There is a substantial research agenda to support the wider adoption of SHM, key topics being:

- Transducers and instrumentation to withstand harsh environments and remain stable
- Calibration methods to account for any drift in sensitivity, phase, and so on with time
- Methods to increase area coverage per sensor at the required sensitivity and low probability of false call
- Further techniques for performance validation including both probability of defect detection and false call rate
- Development of use cases including both the technology and business case
- Efficient data handling to give operators information on which decisions should be taken, rather than raw data
- Fusion of data from multiple sources to provide better prognostic information

Acknowledgment

The author thanks Dr. Stefano Mariani for multiple helpful discussions, for his review of the paper prior to submission and for generating the plot of Fig. 9. The author also thanks Dr. Joseph Corcoran for reviewing the paper and for several helpful suggestions.

Conflict of Interest

There are no conflicts of interest.

References

- [1] Cawley, P., 2018, "Structural Health Monitoring: Closing the Gap Between Research and Industrial Deployment," *Struct. Health. Monit.*, **17**(5), pp. 1225–1244.
- [2] Randall, R. B., 2011, *Vibration-Based Condition Monitoring*, Wiley, Chichester, UK.
- [3] Farrar, C. R., and Worden, K., 2013, *Structural Health Monitoring: A Machine Learning Perspective*, Wiley, UK.
- [4] Farrell, D. M., Robbins, B. J., Sikka, P., and Seaman, M., 2004, "On-Line Monitoring and Control of Furnace Wall Corrosion in PF-Fired Boilers," *Power Plant: Oper. Maint. Mater. Issues*, **3**, pp. 16–21.
- [5] Farrell, D. M., Robbins, B. J., Stallings, J., Cardoso, S., and Bakker, W., 2008, "Crack Growth Monitoring on Industrial Plant Using Established Electrical Resistance 'Scanner' Technology," *Insight*, **50**(12), pp. 690–694.
- [6] Strommen, R. D., Horn, H., and Wold, K. R., 1993, "New Technique Monitors Pipeline Corrosion, Cracking," *Oil Gas J.*, **91**, pp. 88–93.
- [7] Corbin, D., and Willson, E., 2007, "New Technology for Real-Time Corrosion Detection," *Tri-Service Corrosion Conference*, Denver, CO.
- [8] Kuroiwa, T., Homada, Y., Ogawa, Y., Iida, H., Kuroki, M., and Hayakawa, M., 2005, "Application of Potential Drop Technique to the Inspection of Welded Boiler High Temperature and Pressure Parts," *11th International Conference on Fracture*, Turin, Italy.
- [9] Jarvis, R., Cawley, P., and Nagy, P. B., 2018, "Permanently Installed Corrosion Monitoring Using Magnetic Measurement of Current Deflection," *Struct. Health. Monit.*, **17**(2), pp. 227–239.
- [10] Draper Inc., 2017, "A New Tool for Detecting Oil and Gas Pipeline Corrosion," <https://www.draper.com/news-releases/new-tool-detecting-oil-and-gas-pipeline-corrosion>, Accessed January 15, 2021.
- [11] Yokogawa Inc., "Pervasive Corrosion Monitoring on Piping," https://www.yokogawa.com/uk/about/research-development/inv_center/pcm/, Accessed January 15, 2021.
- [12] Li, Z., Jarvis, R., Nagy, P. B., Dixon, S., and Cawley, P., 2017, "Experimental and Simulation Methods to Study the Magnetic Tomography Method (MTM) for Pipe Defect Detection," *NDT&E Int.*, **92**, pp. 59–66.
- [13] Plant Integrity Ltd, "Materials Testing and Inspection Services," <http://www.plantintegrity.com/case-studies/teletest-guided-wave-technology-supports-uk-nuclear-industry/>, Accessed January 15, 2021.
- [14] Todd, M. D., and Flynn, E. B., 2011, "A Bayesian Experimental Design Approach for Structural Health Monitoring," *Proc XIV International Symposium on Dynamic Problems of Mechanics (DINAME 2011)*, Sao Sebastiao, Brazil.

- [15] Todd, M. D., and Flynn, E. B., 2010, "A Bayesian Approach to Optimal Sensor Placement for Structural Health Monitoring With Application to Active Sensing," *Mech. Syst. Signal Process.*, **24**(4), pp. 891–903.
- [16] Flynn, E. B., and Todd, M. D., 2010, "Optimal Placement of Piezoelectric Actuators and Sensors for Detecting Damage in Plate Structures," *J. Intell. Mater. Syst. Struct.*, **21**(3), pp. 265–274.
- [17] Weibull, W., 1939, *A Statistical Theory of the Strength of Materials*, Generalstabens Litografiska Anstalts Förlag, Stockholm.
- [18] Leung, M. S., 2020, "Probabilistic Approach to Structural Integrity Assessment of Fatigue Damage Using Permanently Installed Monitoring Systems," Ph.D. thesis, Imperial College, London.
- [19] Bar-Hillel, M., 1980, "The Base-Rate Fallacy in Probability Judgements," *Acta Psychol.*, **44**(3), pp. 211–233.
- [20] Cegla, F. B., Cawley, P., Allin, J., and Davies, J., 2011, "High-Temperature (> 500 °C) Wall Thickness Monitoring Using Dry-Coupled Ultrasonic Waveguide Transducers," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control*, **58**(1), pp. 156–167.
- [21] Tuegel, E. J., Ingrassia, A. R., Eason, T. G., and Spottswood, S. M., 2011, "Reengineering Aircraft Structural Life Prediction Using a Digital Twin," *Int. J. Aerosp. Eng.*, **2011**, pp. 1–4.
- [22] Tao, F., Cheng, J., Qi, Q., Zhang, M., Zhang, H., and Sui, F., 2018, "Digital Twin-Driven Product Design, Manufacturing and Service With big Data," *Int. J. Adv. Manuf. Technol.*, **94**(9–12), pp. 3563–3576.
- [23] Catbas, F. N., and Kijewski-Correa, T., 2013, "Structural Identification of Constructed Systems: Collective Effort Toward an Integrated Approach That Reduces Barriers to Adoption," *ASCE J. Struct. Eng.*, **139**(10), pp. 1648–1652.
- [24] Moaveni, B., Hurlbauss, S., and Moon, F., 2013, "Special Issue on Real-World Applications of Structural Identification and Health Monitoring Methodologies," *ASCE J. Struct. Eng.*, **139**(10), pp. 1637–1638.
- [25] Warner, M., Ratcliffe, A., Nangoo, T., Morgan, J., Umpleby, A., Shah, N., Vinje, V., Stekl, I., Guasch, L., Win, C., Conroy, G., and Bertrand, A., 2013, "Anisotropic 3D Full-Waveform Inversion," *Geophysics*, **78**(2), pp. R59–R80.
- [26] Niu, Y. S., Hao, N., and Zhang, H., 2016, "Multiple Change-Point Detection: A Selective Overview," *Stat. Sci.*, **31**(4), pp. 611–623.
- [27] Aminikhanghahi, S., and Cook, D. J., 2016, "A Survey of Methods for Time Series Change: Point Detection," *Knowl. Inf. Syst.*, **51**, pp. 339–367.
- [28] Maurya, M. R., Paritosh, P. K., Rengaswamy, R., and Venkatasubramanian, V., 2010, "A Framework for On-Line Trend Extraction and Fault Diagnosis," *Eng. Appl. Artif. Intell.*, **23**(6), pp. 950–960.
- [29] Staudacher, M., Telsler, S., Amann, A., Hinterhuber, H., and Ritsch-Marte, M., 2005, "A New Method for Change-Point Detection Developed for On-Line Analysis of the Heart Beat Variability During Sleep," *Phys. A*, **349**(3–4), pp. 582–596.
- [30] Albert, A. A., Lopez, L. F., and Blas, N. G., 2019, "Multilinear Weighted Regression (MWE) With Neural Networks for," *Appl. Soft Comput.*, **82**, p. 105555.
- [31] Lu, Y., and Michaels, J., 2005, "A Methodology for Structural Health Monitoring With Diffuse Ultrasonic Waves," *Ultrason.*, **43**(9), pp. 717–731.
- [32] Harley, J. B., and Moura, J. M., 2012, "Scale Transform Signal Processing for Optimal Ultrasonic Temperature Compensation," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control*, **59**(10), pp. 2226–2236.
- [33] Konstantinidis, G., Wilcox, P. D., and Drinkwater, B. W., 2007, "An Investigation Into the Temperature Stability of a Guided Wave Structural Health Monitoring System Using Permanently Attached Sensors," *IEEE Sens. J.*, **7**(5), pp. 905–912.
- [34] Clarke, T., Simonetti, F., and Cawley, P., 2010, "Guided Wave Health Monitoring of Complex Structures by Sparse Array Systems: Influence of Temperature Changes on Performance," *J. Sound Vib.*, **329**(12), pp. 2306–2322.
- [35] Ha, S., Lonkar, K., Mittal, A., and Chang, F.-K., 2010, "Adhesive Layer Effects on PZT-Induced Lamb Waves at Elevated Temperatures," *Struct. Health Monit.*, **9**(3), pp. 247–256.
- [36] Herdovics, B., and Cegla, F. B., 2018, "Structural Health Monitoring Using Torsional Guided Wave Electromagnetic Acoustic Transducers," *Struct. Health Monit.*, **17**(1), pp. 24–38.
- [37] Herdovics, B., and Cegla, F. B., 2018, "Compensation of Phase Response Changes in Ultrasonic Transducers Caused by Temperature Variations," *Struct. Health Monit.*, **18**(2), pp. 508–523.
- [38] Mariani, S., Heinlein, S., and Cawley, P., 2020, "Compensation for Temperature-Dependent Phase and Velocity of Guided Wave Signals in Baseline Subtraction for Structural Health Monitoring," *Struct. Health Monit.*, **19**(1), pp. 26–47.
- [39] Fendzi, C., Rébillat, M., Mechbal, M., Guskov, M., and Coffigna, G., 2016, "A Data-Driven Temperature Compensation Approach for Structural Health Monitoring Using Lamb Waves," *Struct. Health Monit.*, **15**(5), pp. 525–540.
- [40] Zoubi, A. B., Kim, S., Adams, D. O., and Mathews, V. J., 2019, "Lamb Wave Mode Decomposition Based on Cross-Wigner-Ville Distribution and Its Application to Anomaly Imaging for Structural Health Monitoring," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control*, **66**(5), pp. 984–997.
- [41] Heinlein, S., Cawley, P., Vogt, T., and Burch, S., 2018, "Blind Trial Validation of a Guided Wave Structural Health Monitoring System for Pipework," *Mater. Eval.*, **76**, pp. 1118–1126.
- [42] Mariani, S., Heinlein, S., and Cawley, P., 2020, "Location Specific Temperature Compensation of Guided Wave Signals in Structural Health Monitoring," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control*, **67**(1), pp. 146–157.
- [43] Lilliefors, H. W., 1967, "On the Kolmogorov–Smirnov Test for Normality With Mean and Variance Unknown," *J. Am. Stat. Assoc.*, **62**(318), pp. 399–402.
- [44] Carandente, R., and Cawley, P., 2012, "The Effect of Complex Defect Profiles on the Reflection of the Fundamental Torsional Mode in Pipes," *NDT&E Int.*, **46**, pp. 41–47.
- [45] Mariani, S., and Cawley, P., 2020, "Change Detection Using the Generalized Likelihood Ratio Method to Improve the Sensitivity of Guided Wave Structural Health Monitoring Systems," *Struct. Health Monit.*, p. 147592172098183 (in press).
- [46] Goulet, J. A., and Smith, I. F., 2013, "Predicting the Usefulness of Monitoring for Identifying the Behavior of Structures," *ASCE J. Struct. Eng.*, **139**(3), pp. 479–479.
- [47] CSIC, "The Smart Infrastructure Paper," <http://www-smartinfracture.eng.cam.ac.uk/files/the-smart-infrastructure-paper>, Accessed January 15, 2021.
- [48] Montgomery, D. C., 2012, *Introduction to Statistical Quality Control*, 7th ed., Wiley, New York.
- [49] Basseville, M., and Nikiforov, I. V., 1993, *Detection of Abrupt Changes: Theory and Application*, Vol. 104, Prentice Hall, Englewood Cliffs.
- [50] Lorden, G., 1971, "Procedures for Reacting to a Change in Distribution," *Ann. Math. Stat.*, **42**(6), pp. 1897–1908.
- [51] Mariani, S., Liu, Y., and Cawley, P., 2020, "Improving Sensitivity and Coverage of SHM Using Bulk Ultrasonic Waves," *Struct. Health Monit.*, p. 147592172096512. (in press).
- [52] Achenbach, J., Komsky, I., Lee, Y. C., and Angel, Y. C., 1992, "Self-Calibrating Ultrasonic Technique for Crack Depth Measurement," *J. Nondestruct. Eval.*, **11**(2), pp. 103–108.
- [53] Legendre, S., Massicotte, D., Goyette, J., and Bose, T. K., 2001, "Neural Classification of Lamb Wave Ultrasonic Weld Testing Signals Using Wavelet Coefficients," *IEEE Trans. Instrum. Meas.*, **50**(3), pp. 672–678.
- [54] Sbarufatti, C., Manson, G., and Worden, K., 2014, "A Numerically-Enhanced Machine Learning Approach to Damage Diagnosis Using a Lamb Wave Sensing Network," *J. Sound Vib.*, **333**(19), pp. 4499–4525.
- [55] Mariani, S., Rendu, Q., Urbani, M., and Sbarufatti, C., 2020, "Causal Dilated Convolutional Neural Networks for Automatic Inspection of Ultrasonic Signals in Non-Destructive Evaluation and Structural Health Monitoring," *Mech. Syst. Signal Process.*
- [56] van den Oord, A., Dieleman, S., Zen, H., Simonyan, K., Vinyals, O., Graves, A., Kalchbrenner, N., Senior, A., and Kavukcuoglu, K., 2016, "WaveNet: A Generative Model for Raw Audio," <http://arxiv.org/abs/1609.03499>, Accessed October 14, 2020.
- [57] Krzanowski, W. J., and Hands, D. J., 2009, *ROC Curves for Continuous Data*, Chapman and Hall/CRC, New York.
- [58] Rentala, V. K., Mylavarapu, P., and Gautam, J. P., 2018, "Issues in Estimating Probability of Detection of NDT Techniques—A Model Assisted Approach," *Ultrason.*, **87**, pp. 59–70.
- [59] Liu, C., Dobson, J., and Cawley, P., 2017, "Efficient Generation of Receiver Operating Characteristics for the Evaluation of Damage Detection in Practical Structural Health Monitoring Applications," *Proc. R. Soc. A*, **473**(2199), p. 20160736.
- [60] Heinlein, S., Cawley, P., and Vogt, T., 2019, "Validation of a Procedure for the Evaluation of the Performance of an Installed Structural Health Monitoring System," *Struct. Health Monit.*, **18**(5–6), pp. 1557–1568.