

D3.2

Optimised transducer design (first generation)

General information		
Grant agreement number	755500	
Start date of the project	01/09/2017	
Project duration	48 months	
Due date of the deliverable	31/08/2018	
Actual submission date	18/02/2019	
Lead beneficiary	6 – UoB	

Keywords

Array optimisation, probability of detection, classification accuracy, sizing accuracy

Туре	Meaning	
R	Document, report	х
DEM	Demonstrator, pilot, prototype	
DEC	Websites, patent fillings, videos, etc.	
OTHER	Software, technical diagram, etc.	

Dissemination Level		
PU	Public	x
СО	Confidential, only for members of the consortium (including the Commission Services)	



Table of Content

1	SUMMARY	5
2	INTRODUCTION	6
3	THE DEFECT+GRAINS MODELLING APPROACH	7
4	INSPECTION PERFORMANCE EVALUATION	. 10
5	CONCLUSIONS	. 14
6	BIBLIOGRAPHY	. 15



Table of Tables

Table of Figures

FIGURE 1: DETECTION AND CHARACTERISATION OF AN EXAMPLE 1MM CRACK	8
FIGURE 2: D-INDEX (SEE EQ. (8)) OF CRACKS FOR GRAIN SIZES (A) 0.2MM AND (B) 0.3MM.	11
FIGURE 3: PROBABILITY OF DETECTION OF CRACKS WHEN THE GRAIN SIZE IS 0.2MM.	
FIGURE 4: C-INDEX (SEE EQ. (9)) OF CRACKS FOR GRAIN SIZES (A) 0.2MM AND (B) 0.3MM.	12
FIGURE 5: S-INDEX (SEE EQ. (10)) OF CRACKS FOR GRAIN SIZES (A) 0.2MM AND (B) 0.3MM.	



Glossary

Abbreviations/Acronym	Description
ВС	Bhattacharyya Coefficient
FMC	Full Matrix Capture
NDT	Non-Destructive Testing
РСА	Principal Component Analysis



1 Summary

This report aims to understand and evaluate the achievable detection/characterisation performance using ultrasonic arrays, which is determined by the nature of a defect and the measurement configuration adopted in experiments. For this purpose, we propose the "defect+grains modelling approach", in which the statistical distribution of the defect data is obtained from different realisations of the grain structure. This statistical distribution, termed the defect+grains model, is shown to contain information that is needed for detection and characterisation of defects. Hence, the characterisation result can be obtained by constructing a defect+grains model based on multiple realisations of each possible defect and calculating their probability given some measurement. The proposed approach gives insight into the detection/characterisation problem, leading to an evaluation of the fundamental limits of the achievable inspection performance.

The work presented in this report has important implications in transducer array design. In particular, we show that the detectability index, classification index, and sizing index (which can all be calculated using the training data only) are related to a lower bound of the achievable detection/characterisation performance. Depending on the aim of the ultrasonic inspection, one or more of these index values can be used as the cost function in the array optimisation process.



2 Introduction

Ultrasonic arrays [1] are now widely used in non-destructive testing (NDT) of nuclear power plant components thanks to their ability to form high resolution images from FMC datasets. WP3 of the ADVISE project, "Inspection technique optimisation", aims to improve the inspection performance of ultrasonic testing for difficult materials such as austenitic welds and cast austenitic steel which have microstructural features (*i.e.* grain structures) that could severely affect the detectability of defects. To address the performance degradation caused by grain scattering noise, it is important to design ultrasonic arrays that can achieve best possible detection and characterisation of defects. Key array parameters including centre frequency, number of elements, and the element pitch size can be optimised (*e.g.* by using a genetic algorithm [2]) to obtain the optimum array configuration for a given measurement scenario. Typically, the probability of detection is selected as the main inspection objective [2]. As the use of the scattering matrix facilitates improved detection and characterisation of small defects [3], the objective function used in the array optimisation process can also be selected to optimise defect characterisation results.

This report is Deliverable 3.2 of WP3 and we introduce three index values which are related to different aspects of defect characterisation (*i.e.* detection, classification of the defect type, and sizing). In *Deliverable 3.1 (Inspection sensitivity analysis),* it was shown that the statistical distribution of grain noise can be modelled based on multiple random realisations of the grain structure and used for evaluation of the characterisation uncertainty. In this report, we extend the "noise modelling approach" used in *Deliverable 3.1* by combining the defect manifold (representing scattering matrices of all possible defects of a given type) [3] and the grain noise model, which enables us to better explain the defect characterisation process and evaluate the achievable detection, classification and sizing performance without explicitly preparing a large number of test data through experiments or time-consuming forward simulations. In particular, each point on the defect manifold (representing noise-free defect scattering matrix) combined with the defect-dependent grain noise distribution can be considered as a statistical "defect+grains" model, which describes the variability of the measurement given the grain size and frequency. Then the evaluation of the characterisation uncertainty can be performed within the same defect characterisation framework as in [3].

A main finding of this report is that characterisation uncertainty arises when there is overlapping between two or more defect+grains models and the achievable characterisation accuracy can be evaluated by quantifying this overlap. Besides optimisation of transducer arrays, this quantitative information can be used in different aspects of ultrasnoic NDT within the scope of the ADVISE project, including inversion framework optimisation (*e.g.* selecting optimal resolution/interval of defect parameters), and selection of the measurement channels and/or wave modes in ultrasonic data fusion [4] (*i.e.* only include data which are expected to improve the detection/characterisation performance).



3 The defect+grains modelling approach

In this section, we introduce the defect+grains modelling approach by considering the characterisation problem of an example 1mm crack in a Type 304 stainless steel sample where the mean grain size is 0.2mm. The same forward modelling approach as in *Deliverable 3.1* is used to simulate the array data of the defect for a 64 element array (element pitch: 0.5mm). Figure 1(a) shows the TFM image of the 1mm crack, obtained at 2.5 MHz. It can be seen that the defect is barely distinguishable from the grain indications, and thus, detection of the small crack from the image is challenging.

Alternatively, we can extract the scattering matrix from the defect location (indicated with an arrow in Fig. 1(a)) and use it for characterisation. In particular, we extend the noise modelling approach proposed in *Deliverable 3.1* by considering the distribution of the defect data (*i.e.* defect+grains) as opposed to only considering the noise distribution. Defect characterisation requires to determine the conditional probability of the defect parameter p (representing size and/or type of a defect which can be quantified using, for example, the aspect ratio of the defect shape [3]) based on the measurement of the noisy scattering matrix S_n . Using Bayes theorem, we have [3]

$$P(\boldsymbol{p}|\boldsymbol{S}_n) = CP(\boldsymbol{S}_n|\boldsymbol{p}). \tag{1}$$

In Eq. (1), $P(S_n|p)$ describes the probability of measuring the noisy scattering matrix S_n from a defect with parameter p, the normalisation constant C is given by $C = (\int P(S_n|p)dp)^{-1}$, and we have assumed that p and S_n are uniformly distributed for simplicity. If there is prior knowledge about the occurrence of different defects or scattering matrices, such information can be used to describe the relationship between the conditional probabilities in Eq. (1) more precisely (*i.e.* different normalisation constants can be derived for different defects) [3]. The statistical distribution $P(S_n|p)$ is termed the defect+grains model in this report, and based on Eq. (1), the defect characterisation problem can be formulated by constructing a defect+grains model for each target defect and calculating the conditional probability of them given the measurement.

Considering the high dimensionality of a scattering matrix, S_n is converted into the defect principal component (PC) space by the use of principal component analysis (PCA) [5] to obtain S_n^{pc} . The role of PCA is to represent the original dataset in a different coordinate system so that the variance of the data can be more efficiently described using a small number of the coordinate axes. This helps to reduce the dimensionalty of the scattering matrix and as a result, actual defect+grains modelling is performed for $S_n^{\rm pc}$ instead of S_n (note here that the defect PC space is different from the noise PC space used in *Deliverable* 3.1). The PCA process effectively constructs a defect manifold [3] for modelled types of defects, and a large number of noise-free scattering matrices are needed for accurate reconstruction of the defect manifold (i.e. smooth surface for defects defined on 2D parameter spaces). In this report, we consider characterisation of small cracks and holes of sizes 1mm, 2mm, and 3mm, in the frequency range between 1 MHz and 3 MHz. The noise-free dataset used for PCA is obtained by sampling in both the defect size and aspect ratio (defined as the ratio between the width and length of a defect [6]) axes of the parameter space. Defect size is sampled in 0.05 λ intervals between 0.05 λ and 2 λ , and aspect ratio is sampled in 0.1 intervals between 0 (cracks) and 1 (holes). The considered defect sizes fall within the modelled size range of $[0.05\lambda,$ 2λ] when the frequency is between 1 MHz and 3 MHz. Although defect+grains modelling can be performed for each point on the defect manifold in principle, this is done only for the six cracks and holes in the current work because of the considerable computational time needed for forward modelling. The variation



explained by first 10 PC directions is given in Table 1, which indicates that we can use a small number of PCs to represent a defect scatteing matrix without much information loss.



Figure 1: Detection and characterisation of an example 1mm crack: (a) TFM image, (b) defect+grains model of a 1mm crack, (c) defect+grains model of a 2mm crack, (d) grains model, (e) all considered defect+grains models and the grains model, and (f) characterisation result obtained by using 7 PCs. The red circle in (b)-(e) represents the scattering matrix of the 1mm crack in defect PC-space, and the contour lines in (e) show the normalised probabilities 0.15 to 0.95 (in 0.2 intervals).

PC number	Variance
1	48.07
2	11.94
3	1.23
4	0.37
5	0.14
6	0.07
7	0.06
8	0.03
9	0.02
10	0.01

Table 1: Variance explained by (i.e. importance of) the first 10 principal component directions of the defect database.

In this report, defect+grains model $P(S_n^{pc}|p)$ is assumed to follow a multivariate Gaussian distribution, which is given as

$$P(\boldsymbol{S}_{n}^{\mathrm{pc}}|\boldsymbol{p}) = \frac{1}{(2\pi)^{N_{p}/2}|\boldsymbol{\Sigma}|^{1/2}} \exp\left[-\frac{1}{2}(\boldsymbol{S}_{n}^{\mathrm{pc}}-\boldsymbol{\mu})^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}(\boldsymbol{S}_{n}^{\mathrm{pc}}-\boldsymbol{\mu})\right],$$
(2)

where N_p denotes the number of PCs considered for characterisation, and is set to 7 in this report since the other PCs are less than 1% (relative to the first PC) in terms of the significance (*i.e.* variation of the database explained, see Table 1). The mean scattering matrix μ and the covariance matrix Σ can be estimated from the available training data by



$$\boldsymbol{\mu} = \sum_{i=1}^{M} \boldsymbol{S}_{n,i}^{\mathrm{pc}} / M, \qquad (3)$$

$$\boldsymbol{\Sigma} = \sum_{i=1}^{M} (\boldsymbol{S}_{n,i}^{\mathrm{pc}} - \boldsymbol{\mu}) (\boldsymbol{S}_{n,i}^{\mathrm{pc}} - \boldsymbol{\mu})^{\mathrm{T}} / (M-1), \qquad (4)$$

where M = 50 is the number of the noisy scattering matrices simulated for each target defect, and $S_{n,i}^{pc}$ denotes the *i*-th training data.

The contour plot in Fig. 1(b) represents the defect+grains model of a 1 mm crack, and the red circle denotes the extracted scattering matrix of the 1mm crack for which detection was shown to be difficult from the image (Fig. 1(a)). Given the defect+grains model and the measurement point in defect PC-space, the conditional probability $P(S_n^{pc}|C_1)$ can be calculated as 0.0065, where we used the notation C_1 to denote the defect+grains model of the 1mm crack. Similarly, we can obtain $P(S_n^{pc}|C_2) = 4.4 \times 10^{-5}$ for the conditional probability with respect to a 2mm crack shown in Fig. 1(c). As can be seen, the conditional probability of observing the measurement is very low if the measurement is tested against a "wrong" defect+grains model. In Fig. 1(d), the contour plot represents the "grains model", which shows the distribution of the scattering matrices measured from defect-free samples. The grains model is seen to be close to the measurement, and it gives conditional probability of the measurement that is higher than the individual defect+grains models of the cracks (*i.e.* $P(S_n^{pc}|N) = 0.0084$ where N denotes the grains model). This result can also explain why detection is inherently difficult from the image for this crack. The considered defect+grains models are plotted in Fig. 1(e) alongside the grains model, and it can be seen that the measurement lies within the overlapping region of the grains model and defect+grains models of the 1mm crack and holes of all sizes. Note that the defect+grains models are shown in 2D PC-space for purposes of visualisation and the probability of detection for the considered crack is only 63.3% in this case. The characterisation result can be improved by using more PCs than 2, and an excellent characterisation result is obtained by using 7 PCs as shown in Fig. 1(f). However, it should be noted that the example case shown here is taken from the training set, and actual characterisation performance can potentially drop. In the next section, the defect+grains models are used to evaluate the achievable characterisation performance which can be expected from real experimental measurements.



4 Inspection performance evaluation

In *Deliverable 3.1* it was shown that the achievable classification accuracy and sizing accuracy of a defect are not necessarily the same. This can also be found from Fig. 1(e) --- the defect+grains model of the 3mm crack is well separated from those of the holes which can result in excellent classification of 3mm cracks, while there is still uncertainty in sizing due to the overlap between defect+grains models of 2mm and 3mm cracks. Hence, it is necessary to evaluate and quantify the inspection performance of different aspects of defect characterisation, and as in *Deliverable 3.1*, we select probability of detection, classification accuracy, and sizing accuracy as main inspection objectives.

The defect+grains model introduced in this report describes the variability of the measurement due to noise, and hence, contains information about the achievable characterisation performance. As explained in Section 3, if the measurement is within the overlapping region of several defect+grains models, its characterisation result will have a probability distribution over the corresponding defect parameters, *i.e.*, there is characterisation uncertainty. In order to quantify the amount of overlap between two defect+grains models, we can use the Bhattacharyya coefficient (BC) which is defined as [7]

$$BC(f,g) = \int \sqrt{f(x)g(x)} dx,$$
(5)

where f(x) and g(x) are probability density functions of variable x and satisfy $\int f(x)dx = \int g(x)dx = 1$. It follows from Cauchy-Schwarz inequality [8] that $0 \leq BC(f,g) \leq \int f(x)dx \cdot \int g(x)dx = 1$. The Bhattacharyya coefficient is 0 if f and g do not overlap and is 1 if f is equal to g. In other cases, the Bhattacharyya coefficient is within the range (0,1), and higher values indicate more severe overlapping of two distributions.

For detection problem, the physical meaning of BC defined in Eq. (5) is that it gives an estimation of the achievable probability of detection using the defect+grains modelling approach proposed in Section 3. Let C_i and N denote the defect+grains model of a crack with size i (unit: mm) and the grains model, respectively. Because they are modelled as multivariate Gaussian distributions (see Eq. (2)), their BC value can be calculated as [9]

$$BC(\boldsymbol{C}_i, \boldsymbol{N}) = \exp(-D_B), \tag{6}$$

where the Bhattacharyya distance D_B has the expression

$$D_B = \frac{1}{8} (\boldsymbol{\mu}_2 - \boldsymbol{\mu}_1)^T \boldsymbol{\Sigma}^{-1} (\boldsymbol{\mu}_2 - \boldsymbol{\mu}_1) + \frac{1}{2} \ln \frac{\left| \frac{\boldsymbol{\Sigma}_1 + \boldsymbol{\Sigma}_2}{2} \right|}{\sqrt{|\boldsymbol{\Sigma}_1||\boldsymbol{\Sigma}_2|}}.$$
 (7)

In Eq. (7), μ_1 , μ_2 are the mean scattering matrices of the conisered defect+grains model and grains model, resepectively, and Σ_1 , Σ_2 are their covariance matrices (Σ is the average covariance matrix). Then it can be shown that the achievable probability of detection using the proposed approach has a lower bound that is given by $1 - 0.5 \times BC(C_i, N)$ for a binary classification problem (*i.e.* defect versus noise) [10]. Here we have assumed that the occurrence of the defect and non-defect cases are equally probable, and as a result, the expected probability of detection is 0.5 when there is no useful characterisation information available in the worst case scenario. However, in practice, large defects tend to have lower BC values which could result in higher probability of detection than small defects. This also indicates that for a given defect, best detection performance can be achieved by minimising the BC between the defect+grains model of the defect and the grains model.

Based on the above consideration, we define the detectability index (d-index) of a crack (with size *i*) as



(8)

$$d(i) = 1 - BC(C_i, N).$$

Using the forward modelling approach described in Deliverable 3.1, we have simulated array data for cracks and holes of the sizes 1mm, 2mm, and 3mm, for two different mean grain sizes 0.2mm and 0.3mm, and at centre frequencies 1-3 MHz. For each defect and grain size, 50 random grain structures are simulated to obtain the training data (i.e. used for constructing the defect+grains models). In addition, array data of defect-free samples are simulated and used for constructing the grains model. Figures 2(a)-2(b) show the d-index results of the cracks for grain sizes 0.2mm and 0.3mm, respectively. As expected, the d-index (i.e. detectability) of the cracks drops as the frequency increases. When the grain size is 0.2mm, the d-index of the cracks is 1 for frequencies 1 MHz and 1.5 MHz, meaning that the cracks can be detected with high confidence at these low frequencies. When the frequency is 2.5 MHz, 2mm and 3mm cracks maintain excellent detectability but 1mm cracks can potentially become undetectable for certain grain structures. All these findings obtained from the d-index results are in good agreement with image-based inspection results of the cracks (see Table 1 of Deliverable 3.1). When the grain size is 0.3mm, the d-index values of the cracks drop more quickly as the frequency increases, meaning that the effect of grain noise is more severe. For example, we find that the d-index of 3mm cracks for the grain size 0.3mm is smaller than that of 1mm cracks for the grain size 0.2mm when the frequency is 3 MHz. This explains why detection can easily become fundamentally challenging even for large defects when the grain size is also large.



Figure 2: d-index (see Eq. (8)) of cracks for grain sizes (a) 0.2mm and (b) 0.3mm.

For crack sizes 1mm, 2mm, and 3mm, 10 new grain structures (which were not included for defect+grains modelling) are used to simulate noisy scattering matrices for the grain size 0.2mm, and the defect characterisation approach described in Section 3 is used to characterise them. Figure 3 shows the average probability of detection of the 10 noisy measurements for each crack size. Good agreement is observed between the d-index results in Fig. 2(a) and actual probability of detection in Fig. 3, indicating that the d-index values can indeed provide a realsitic estimation of the achievable probability of detection, and is thus a useful criterion for transducer array optimisation.

To evaluate the classification accuracy, we define the classification index (c-index) of a crack as

$$c(i) = 1 - \max\{BC(C_i, H_j)\}, \quad j = 1, 2, 3.$$
 (9)

By optimising the c-index, we seek to minimise the overlap between the defect+grains model of a crack and that of the "most undesirable" hole for the considered crack (*i.e.* the hole that is most difficult to

distinguish from the crack). The c-index results of the cracks for different grain sizes are shown in Fig. 4. When the grain size is 0.2mm, we find that the cracks have the maximum c-index values at 2 MHz, meaning that the small defect size (relative to the ultrasonic wavelength) at 1 MHz and the high noise level at 3 MHz are both undesirable for distinguishing cracks from holes. When the grain size is 0.3mm, however, the c-index values of 2mm and 3mm cracks are shown to decrease monotonically as the frequency increases, and the classification performance is now dominated by the grain noise.



Figure 3: Probability of detection of cracks when the grain size is 0.2mm.



Figure 4: c-index (see Eq. (9)) of cracks for grain sizes (a) 0.2mm and (b) 0.3mm.

Based on a similar idea, we define the sizing index (s-index) of a crack as

$$s(i) = 1 - \max\{BC(\boldsymbol{C}_i, \boldsymbol{H}_j), BC(\boldsymbol{C}_i, \boldsymbol{C}_j)\}, \quad j \neq i.$$
(10)

The s-index results of the cracks are shown in Fig. 5, which can be used to compare and evaluate the sizing accuracy of a crack in different measurement scenarios. When the grain size is 0.2mm, different frequencies are shown to be optimum for different crack sizes (*e.g.* 1.5 MHz for 3mm cracks and 2 MHz for 2mm cracks). Interstingly, the s-index value of 1mm cracks is reasonably high at 1 MHz (*i.e.* s = 0.65 and is higher than s-index values of 2mm and 3mm cracks at the same frequency), which means that sizing of 1mm cracks can still be expected to be accurate regardless of their low classification accuracy. When the grain size increases



to 0.3mm, once again, the s-index values decrease quickly as the frequency increases. The s-index values shown in Fig. 5(b) are low (*i.e.* below 0.6) at all frequencies, suggesting that sizing of the cracks has become fundamentally challenging due to the increased grain size.



Figure 5: s-index (see Eq. (10)) of cracks for grain sizes (a) 0.2mm and (b) 0.3mm.



5 Conclusions

In order to address the defect characterisation problem, we have proposed the defect+grains modelling approach in this report, in which the defect+grains model of a specific defect is constructed using scattering matrices obtained from multiple random grain realisations. Given any measurement, the probability that the scattering matrix is measured from a specific defect+grains model can be calculated, and it is shown that this probability is proportional to the probability of the considered defect+grains model given the measurement. Hence, the defect characterisation problem can be formulated by constructing a defect+grains model for each target defect and calculating the conditional probability of them given the measurement.

Based on the defect+grains modelling approach, we have introduced three index values which can potentially be used as the objective function for transducer array optimisation. The detectability index, classification index, and sizing index are defined based on the Bhattacharyya coefficient and as their names suggest, are related to different levels of defect characterisation. By calculating these index values, it was shown that the achievable characterisation accuracy is different for different defects. Using this approach, Figures 2, 4, and 5 were constructed which allow the optimal selection of transducer frequency to achieve a desired level of detection, classification or sizing. For example, 3mm cracks can be detected, classified and sized with excellent accuracy when the grain size is 0.2mm and the frequency is between 1 MHz and 2.5 MHz. However, the classification accuracy of 1mm cracks is consistently poor at all frequencies, while accurate sizing of them can still be expected when the frequency is 1 MHz. When the grain size is 0.3mm, detection, classification, and sizing of the cracks are all severely affected by grain noise, in particular, at high frequencies.

The defect+grains modelling approach relies on extracting the scattering matrix from the measured array data. Grain noise is an inherent feature of the defect scattering matrix and a spatial filtering approach was used in the current work when extracting the scattering matrix. It is noted that defect characterisation can become more reliable if the noise level in the defect scattering matrix decreases, and future work will aim to explore the possibility of improving the accuracy of the scattering matrix extraction.



6 Bibliography

[1] B. W. Drinkwater and P. D. Wilcox, "Ultrasonic arrays for non-destructive evaluation: A review," *NDT E Int.*, vol. 39, no. 7, pp. 525–541, 2006.

[2] Y. Humeida, P. D. Wilcox, M. D. Todd, and B. W. Drinkwater, "A probabilistic approach for the optimisation of ultrasonic array inspection techniques," *NDT E Int.*, vol. 68, pp. 43–52, Dec. 2014.

[3] A. Velichko, L. Bai, and B. W. Drinkwater, "Ultrasonic defect characterisation using parametricmanifold mapping," *Proc. R. Soc. A*, vol. 473, no. 2202, p. 20170056, 2017.

[4] N. Brierley, T. Tippetts, and P. Cawley, "Data fusion for automated non-destructive inspection," *Proc. R. Soc. A*, vol. 470, no. 2167, p. 20140167, 2014.

[5] I. T. Jolliffe, *Principal component analysis,* 2nd ed. Springer-Verlag, 2002.

[6] L. Bai, A. Velichko, and B. W. Drinkwater, "Characterization of defects using ultrasonic arrays: A dynamic classifier approach," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control*, vol. 62, no. 12, pp. 2146–2160, 2015.

[7] T. Kailath, "The divergence and Bhattacharyya distance measures in signal selection," *IEEE Trans. Commun. Technol.*, vol. 15, no. 1, pp. 52–60, 1967.

[8] M. Abramowitz and I. A. Stegun, *Handbook of mathematical functions with formulas, graphs, and mathematical tables*. Dover, 1972.

[9] K. Fukunaga, *Introduction to statistical pattern recognition*, 2nd ed. Academic Press, 1990.

[10] P. A. Devijver and J. Kittler, *Pattern recognition: A statistical approach*. Prentice hall, 1982.