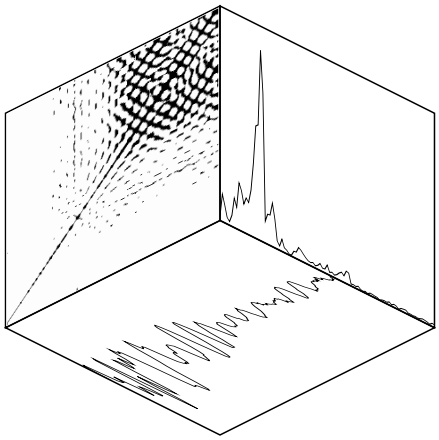
Recurrence Quantification Compared to Fourier Analysis for Ultrasonic Non-Destructive Testing of Carbon Fibre Reinforced Polymers



Dissertation
zur Erlangung des akademischen Grades
Doktor der Ingenieurwissenschaften (Dr.-Ing.)

Carsten Brandt

23.03.2020

vorgelegt am Fachbereich 3 Mathematik/Informatik
Universität Bremen

Dipl.-Ing. Carsten Brandt

Recurrence Quantification Compared to Fourier Analysis for Ultrasonic Non-Destructive Testing of Carbon Fibre Reinforced Polymers

Gutachter:

1. Prof. Dr. Dr. h.c. Peter Maaß (Universität Bremen)
2. Dr. habil. Norbert Marwan (Potsdam-Institut für Klimafolgenforschung)

Kolloquium:
12.06.2020, Bremen

An advice to the inventor

'The inventor should have the very practical consciousness that for many years his greatest contribution will not be a single gadget ... He should ... realize that ... he himself and his work will have to serve rather as a stepping stone to the future than as the end to which science and technique must finally arrive.'

Norbert Wiener [1956: 267]

Publications

The author applied for a patent in Germany and later on in the USA via his employer Airbus Operations GmbH prior to start of this thesis for the basic idea — using recurrence quantification analysis for non-destructive testing:

Brandt, C. (31st July 2014). 'Technik der zerstörungsfreien Prüfung'. German pat. req. DE 10 2014 011 424 A1 2016.02.04. URL: https://depatisnet.dpma.de/DepatisNet/depatisnet?action=pdf&docid=DE102014011424A1&xxxfull=1.

Brandt, C. (28th July 2015). 'Technique for Non-destructive Testing'. U.S. pat. req. US 2016/0034422 A1. URL: https://depatisnet.dpma.de/DepatisNet/depatisnet?action=pdf&docid=US020160034422A1&xxxfull=1.

Some parts of this thesis have already been published as conference manuscripts or as a reviewed contribution to the proceedings of the International Symposium on Recurrence Plots:

Brandt, C. & P. Maaß (2015). 'A State Space Approach for the Non-Destructive Evaluation of CFRP with Ultrasonic Testing'. In: 7th International Symposium on NDT in Aerospace 2015 (Bremen, Germany, 16th–18th Nov. 2015). URL: http://www.ndt.net/article/aero2015/papers/Mo_3_A_1_Brandt_v2.pdf. (pp. 2, 13, 117)

Brandt, C. & P. Maaß (2016). 'Recurrence Quantification Analysis for Non-Destructive Evaluation with an Application in Aeronautic Industry'. In: 19th World Conference on Non-Destructive Testing 2016 (Munich, Germany, 13th–17th June 2016). URL: http://www.ndt.net/article/wcndt2016/papers/we1i4.pdf. (pp. 4, 10, 13)

Brandt, C. (2016). 'Recurrence Quantification Analysis as an Approach for Ultrasonic Testing of Porous Carbon Fibre Reinforced Polymers'. In: *Recurrence Plots and Their Quantifications: Expanding Horizons. Proceedings of the 6th International Symposium on Recurrence Plots, Grenoble, France, 17-19 June 2015*. Ed. by C. L. Webber Jr., N. Marwan & I. Cornel. Vol. 180. Springer Proceedings in Physics. Cham: Springer. Chap. 19, pp. 355–377. DOI: 10.1007/978-3-319-29922-8_19. © Springer International Publishing Switzerland 2016. (pp. 56, 114)

Brandt, C. (2018). *Recurrence Quantification Analysis for Non-Destructive Testing of Porosity in Carbon Fibre Reinforced Polymers*. URL: https://www.researchgate.net/publication/323695223_Recurrence_Quantification_Analysis_for_Non-Destructive_Testing_of_Porosity_in_Carbon_Fibre_Reinforced_Polymers.

The choice of methods in time domain and of Fourier analysis in this work has been accompanied by work in three *Modellierungsseminare* (modelling projects) within the master degree studies in Industrial Mathematics at the University of Bremen, conducted by Matthias Rick & Kai Wah Chan, Matthias Otten & Jim van Kleef and Johannes Leuschner & Marco Hamann. Johannes Leuschner had the idea of performing linear regression on the amplitude spectrum. These results and some work on wavelets (obtained and conducted on a different data set of ultrasonic scans of the CFRP specimens used in this work) can be found in

Brandt, C., M. Hamann & J. Leuschner (2019). *Regression Models for Ultrasonic Testing of Carbon Fiber Reinforced Polymers*. URL: `http://www.math.uni-bremen.de/zetem/cms/media.php/262/BerichtTechnomathematik1901english.pdf`. (pp. 21, 65, 88, 111)

Abstract

Recurrence quantification analysis (RQA) is investigated on an application of ultrasonic testing on carbon fibre reinforced polymer (CFRP) and compared to methods in time domain and Fourier analysis, especially using linear regression. The application — as one example of deployment of RQA in the engineering domain, in which it has been rarely used yet — concerns the detection of porosity in CFRP with ultrasonic normal incidence pulse-echo testing.

The work aims at creating an equivalent to the established criterion of ultrasonic pulse-echo testing for porosity, the height of the back-wall echo (BWE), which is not available in the inspection of certain CFRP structures such as bonded parts or stringers (stiffening elements of an aircraft fuselage or wing). For such structures, a back-wall echo equivalent (BWE-equivalent) shall be generated out of the time series that consist of intermediate echoes from the inside of the CFRP.

Specimens of 6–7 mm thickness from a fabric and a unidirectional CFRP material, both epoxy resin based, were investigated. The latter material was available in two different ply thicknesses. The results depend largely on whether a resonance effect with ply thickness approximately half the main wavelength occurs, leading to rather regular, sinusoidal intermediate echoes.

It is not possible to create a 1:1 correspondence to the BWE, since the spread of the BWE-equivalent is too large despite partially large correlation coefficient $r \approx 0.9$.

A classification into porous or non-porous based on the BWE-equivalent created in this thesis is possible: for the unidirectional CFRP with resonance effect, linear regression on the amplitude spectrum provided excellent cross-validation results with balanced accuracies, mean of true positive (detection) rate and true negative rate (1 minus false alarm rate), greater than 96%. Applying a Hann window on the intermediate echo time series prior to Fourier transform is essential, and linear regression relies heavily on frequencies above the main frequencies, once leakage is reduced through the Hann window. For the fabric material (also with resonance effect), the simple tool quartile coefficient of dispersion (QCD) leads to best results with balanced accuracies $> 91\%$; linear regression on amplitude spectrum with Hann window and RQA, feature determinism DET using Euclidean distance, being second best. For the unidirectional material without resonance effect worse but still promising balanced accuracies from $\approx 81\%$ to $\approx 91\%$ were achieved with RQA, Euclidean distance, feature $RATIO$, as well as with linear regression on amplitude spectrum with Hann window.

An ultrasonic inspection with a frequency leading to a resonance effect plus linear regression on the amplitude spectrum after applying a Hann window is recommended for industrial application to classify into porous and non-porous without access to a back-wall echo.

Zusammenfassung

Recurrence Quantification Analysis (RQA) wird für eine Anwendung der Ultraschallprüfung auf Kohlenstofffaserverstärktem Kunststoff (CFK) untersucht und mit Methoden im Zeitbereich und der Fourier-Analyse verglichen. Die Anwendung – als ein Beispiel des Einsatzes der RQA im Ingenieurwesen, in dem sie bisher selten angewendet wurde - ist die Detektion von Porosität in CFK mit der Ultraschall-Impuls-Echo-Prüfung in Senkrechteinschallung.

Das Ziel ist die Erzeugung eines Äquivalents zum etablierten Kriterium der Ultraschallprüfung in Senkrechteinschallung auf Porosität, der Rückwandechohöhe, die bei der Prüfung von bestimmten CFK-Strukturen, z.B. Klebestrukturen oder Stringer (Versteifungselemente eines Flugzeugrumpfes oder -flügels), nicht genutzt werden kann. Für solche Strukturen soll ein Rückwandechoäquivalent aus den Zeitreihen der Zwischenechos aus dem Inneren des CFK generiert werden.

Bauteile mit 6–7 mm Dicke aus einem Gewebe- und einem unidirektionalen CFK-Material, beide basierend auf Epoxidharz, wurden untersucht. Das zweite Material war in zwei verschiedenen Lagendicken verfügbar, und die Ergebnisse hängen deutlich davon ab, ob ein Resonanzeffekt auftritt. Dieser entsteht, wenn die Lagendicke circa der halben Hauptwellenlänge entspricht, und führt zu eher regelmäßigen, sinusförmigen Signalen.

Ein 1:1 Ersatz des Rückwandechos ist nicht möglich, da die Streuung des Rückwandechoäquivalents trotz z.T. hoher Korrelationskoeffizienten $r \approx 0.9$ zu groß ist.

Eine Klassifikation in porös oder nicht-porös auf Basis des in dieser Arbeit erzeugten Rückwandechoäquivalents ist möglich: Für das unidirektionale CFK-Material mit Resonanzeffekt führt lineare Regression auf dem Amplitudenspektrum zu exzellenten Kreuzvalidierungsergebnissen mit mittleren Korrektklassifikationsraten, Mittelwert zwischen Richtig-positiv-Rate (Fehlerfindungsrate) und Richtig-negativ-Rate (1 minus Falschalarmrate), größer 96%. Die Anwendung eines Hann-Fensters auf den Zwischenechozeitreihen vor der Fouriertransformation zur Reduzierung des Leckeffektes ist von grundlegender Bedeutung, und die lineare Regression betont Frequenzen oberhalb der Hauptfrequenzen. Für das Gewebematerial werden mit dem simplen quartilen Streuungskoeffizienten die besten Ergebnisse mit mittleren Korrektklassifikationsraten $> 91\%$ erreicht. Lineare Regression auf dem Amplitudenspektrum mit Hann-Fenster sowie RQA, Merkmal Determinismus DET, liefern die zweitbesten Ergebnisse. Für das unidirektionale Material ohne Resonanzeffekt werden schlechtere, aber immer noch vielversprechende mittlere Korrektklassifikationsraten von $\approx 81\%$ bis $\approx 91\%$ erreicht, sowohl mit RQA, Merkmal $RATIO$, als auch mit linearer Regression auf dem Amplitudenspektrum mit Hann-Fenster.

Eine Ultraschallprüfung mit einer Frequenz, die zum Resonanzeffekt führt, plus linearer Regression des Amplitudenspektrums nach Anwendung eines Hann-Fensters wird für die industrielle Anwendung zur Klassifikation in porös und nicht-porös ohne Verfügbarkeit eines Rückwandechos empfohlen.

Contents

Acknowledgements

I am very happy and grateful for the opportunity to live out my disposition to scientific-mathematical work through this dissertation, parallel to my job in industrial research for non-destructive testing. There are several people giving me inspiration before my dissertation started, for which I am thankful, for example Professor Günter Ludyk for his lectures on dynamical systems in control engineering more than 20 years ago, and Detlef Scheer for his professional coaching, which was the final impetus to start this thesis.

I am absolutely thankful to my supervisors: Professor Peter Maaß dared to take (after unfinished earlier experiences) another external doctoral student and has been so helpful in his advices and positive feedback — I also thank him for the simple, but efficient idea to use linear regression, never before deployed for porosity detection in CFRP with ultrasonic testing. Dr Norbert Marwan has taken the effort of this supervision, helped so much in his advices for writing and gave me the certainty to be on the right track with my evaluation of recurrence plots and their quantification.

I also especially thank Professor Christian Große and Professor Werner Brannath for participating in the examination board of my dissertation's colloquium. Professor Große furthermore provided helpful comments on my writing style and regarding the theory of ultrasonic testing. Professor Brannath — together with Dr Martin Scharpenberg — also helped a lot with his advices regarding approaches on training and test within data evaluation and machine learning for my thesis.

I am also grateful to Judith Barthel and Dörthe Mindermann, assistants to Professor Maaß at the Center for Industrial Mathematics at the University of Bremen, for their kind support throughout.

I thank my employer, Airbus Operations GmbH, for providing specimens and equipment for ultrasonic measurements and its ideal support; in person foremost my manager Christian Rückert and the manager of non-destructive testing for production, Dr Clemens Bockenheimer, and his predecessor, Rudolf Henrich. I am grateful to many colleagues for useful discussions: Jens Kethler about all kinds of ultrasonic questions, Dr Andreas Kück about statistics, Dr Oliver Bullinger regarding Latex and layout, Dr Jörg Jendrny about his interesting style comments, Wolfgang Bisle for his general encouragement; and Helge Hicken for providing two additional ultrasonic images. I thank my British colleagues Dr Kathryn Atherton, Dr Adam Pitman and David Steer for their immediate replies to my English language questions every now and then. My special thanks go to Stephan Schwaneberg for digging microsections out of an old database and to Dr Rainer Stößel from Airbus Central Research & Technology and his former doctoral student, Dr Denis Kiefel, now with Testia GmbH, for providing and helping with two CT images.

Special thanks go to Professor Jens Rademacher as expert for the theory of dynamical systems for the exchanges and comments in his field of expertise.

My regular exchanges with other doctoral candidates were helpful throughout. I especially want to mention Johannes Leuschner from the Center for Industrial Mathematics for the helpful discussions about my topic and results, and furthermore Lennert Heilmann from DLR, Thies Gerken from the Center for Industrial Mathematics, and particularly my friends Simon Grützner, knowing both worlds of mathematics and engineering, as well as Nicolas Dobmann for the numerous delighting talks during our lunch breaks.

Thanks go especially to the students working within the scope of my dissertation on modelling projects, in addition to Johannes Leuschner this was Matthias Rick, Kai-Wah Chan, Matthias Otten, Jim van Kleef and Marco Hamann.

I am grateful to exchanges with many more people: Professor Robert Smith and also Professor Paul Wilcox as well as Dr Nicolas Dominguez for the exchanges to effects of ultrasound in CFRP, Dr Marcel Rennoch and Dr Tobias Boskamp for general exchanges, Pascal Fernsel on the topic of delay embedding, and, in the area of recurrence quantification analysis, Professor Reik Donner and Professor Charles Webber, Hauke Krämer, Dr Sebastian Oberst and Merten Stender. I am furthermore grateful for having had the opportunity to listen to Professor Sylvie Oliffson Kamphorst in her speech about the compilation of the first paper introducing recurrence plots and the nice conversation with her during the symposium of recurrence plots.

I thank my mother Elli and my late father Heiko — without the conscientiousness they provided me with this work would not have been possible.

I am grateful to my children Elian, Muriel and Maurice for their enduring understanding when their father worked on his dissertation every other weekend.

And the warmest thanks go to my love, my wife Martina: your encouragement and support means everything to me.

Mathematical Notation

For general reference and definition, some mathematical notation is given hereafter, preceding a list of symbols.

Sets and spaces are presented with calligraphic uppercase letters, e.g. $\mathcal{A}$, $\mathcal{X}$. Exceptions are the sets of real and natural numbers $\mathbb{R}$, $\mathbb{N}$. The elements of a specific set are given with braces, e.g. $\mathcal{A} = \{1, 2\}$ for the set consisting of the elements 1 and 2.

A complex number $c = a + jb$ with $j = \sqrt{-1}$ has a complex conjugate $c^* = a - jb$. The absolute values are equal $|c| = |c^*|$.

Intervals with square brackets include the number (in contrast to round brackets), $y \in [2, 3) \Longleftrightarrow 2 \leq y < 3$.

A *function* $f(x)$ is understood in the restricted sense of relating each element x of any set or space to exactly one element of the set of numbers (real or complex). The general case of assigning to exactly one element of *any* set or space is called a *mapping*. (*Map* is exclusively used for the evolution operator of a dynamical system with discrete time, cf. sect. 3.3.2.)

Vectors are given in bold lower case roman letters $\mathbf{x}$ and *matrices* in bold upper case roman letters $\mathbf{A}$ (an exception are vectors of back-wall echo and back-wall echo equivalent, respectively, which are given in their upper case three letter abbreviation in bold face **BWE** and $\mathbf{BWE}_{equiv}$). *Components of vectors and matrices* are given in italic face (as are variables in general) and as greek letters ξ_k; an exception are time series treated as a vector, where the components are x_i. One *norm* of a vector $\|\mathbf{x}\|$ is the *Euclidean norm*

$$\|\mathbf{x}\| = \sqrt{\sum_{i=1}^{n} \xi_i^2}. \tag{1}$$

This norm induces the *Euclidean distance*

$$d(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\| = \sqrt{\sum_{i=1}^{n} (\xi_i - \zeta_i)^2}. \tag{2}$$

A *ball of radius* x is the neighbourhood of a point x, $B_r x = \{y \in \mathcal{R}^n | d(y, x) < r\}$.

The first derivative with respect to time t is displayed with a dot

$$\frac{dx}{dt} = \dot{x}.$$

An *inner product* is a mapping $\langle \cdot, \cdot \rangle$ of two elements x, y of a space $\mathcal{X}$ (the inner product space) to the real line or complex plane such that

$$\begin{aligned} \langle x + y, z \rangle &= \langle x, z \rangle + \langle y, z \rangle, \\ \langle \alpha x, y \rangle &= \alpha \langle x, y \rangle, \\ \langle x, y \rangle = \langle y, x \rangle &\text{ if } \mathcal{X} \text{ is a real space}, \\ \langle x, x \rangle &\geq 0, \\ \langle x, x \rangle = 0 &\iff x = 0. \end{aligned}$$

[Kreyszig 1989: 128f.]

A familiar example in engineering is the dot or scalar product of vectors e.g. in a three-dimensional space.

The inner product allows to generalise the notion of orthogonality as known from vectors in engineering: an element $x \in \mathcal{X}$ is orthogonal to an element $y \in \mathcal{X}$ if

$$\langle x, y \rangle = 0. \tag{3}$$

An orthonormal set $\mathcal{G}$ in an (inner product) space is a set whose elements are pairwise orthogonal and have the norm 1,

$$\langle x, y \rangle = \begin{cases} 0 & \text{if } x \neq y \\ 1 & \text{if } x = y. \end{cases} \tag{4}$$

[Kreyszig 1989: 151f.]

An example from engineering are the three unit vectors in three-dimensional space.

In non-destructive testing, to which the methods of data analysis in this work are applied, signals are often presented (or rather: compared) in the (logarithmic) decibel scale; sound pressure amplitude A_1 equals in comparison to a reference pressure A_0

$$A_{1\mathrm{ref}A_0} = 20 \log \frac{A_1}{A_0}$$

with unit dB (decibel).

Abbreviations or short forms of features, which are obtained from the analysis conducted in this work, are written italic since they are variables, but in upper case letters as for all abbreviations, e.g. *RR* (recurrence rate), *DET* (determinism), *QCD* (quartile coefficient of dispersion), *TPR* (true positive rate) or *BWE* (back-wall echo).

Symbols

A_{bw3}	integral of bandwidth: integral of amplitude spectrum (magnitudes of Fourier transformed time series) from F_{left} to F_{right}
acc_{bal}	balanced accuracy, mean of *TPR* and *TNR*
AUC	area under curve
B_{3dB}	difference $F_{right} - F_{left}$
BWE	(scalar) back-wall echo
BWE_{equiv}	(scalar) back-wall echo equivalent
BWE	(column) vector with back-wall echoes
$\mathbf{BWE}_{equiv}$	(column) vector with back-wall echo equivalents
c	velocity of (ultra)sound
CD	cross distance matrix (cross recurrence plot if plotted)
c_L	velocity of longitudinal (ultra)sonic wave (compression wave)
COR	RQA feature correlation in adaptation of the feature trend: correlation coefficient between distance to line of identity (main diagonal of recurrence plot) and recurrence rate of the single lines, taking values between -1 and 1
c_T	velocity of transverse (ultra)sonic wave (shear wave)
D	distance matrix (distance plot or unthresholded recurrence plot if plotted)
d	dimension of reconstructed state space (via time delay embedding out of time series)
Δ_M	number of diagonal lines furthest away from main diagonal left out for calculaton of *TND* and *COR*, $\Delta_M = M - \widetilde{M}$
DET	determinism (RQA feature)
DIV	divergence (RQA feature), reciprocal of maximum occuring ascending diagonal line length in recurrence plot

$\mathbf{DR}$	difference recurrence matrix (difference recurrence plot if plotted)
d_s	diameter of a spherical scatterer (reflector)
E	Young's modulus (modulus of elasticity)
$\emptyset$	empty set
ENT	entropy (RQA feature)
ϵ	recurrence threshold
ϵ_j	error in linear model when approximating the back-wall echo by the intermediate echo time series (or its Fourier transform)
$\boldsymbol{\epsilon}$	errors in linear model when approximating back-wall echoes by intermediate echo time series (or their Fourier transform) (column vector)
Φ	evolution operator of a dynamical system
f	frequency
$\mathbf{f}$	function of a dynamical system (discrete case)
f_{bwee}	function to generate a back-wall echo equivalent for ultrasonic inspection of CFRP, the overall aim in this thesis, investigated with recurrence quantification analysis, Fourier analysis and several approaches in time domain
F_{centre}	centre frequency: mean of F_{cr} and F_{cl}
F_{cl}	left cut-off: frequency at which (the first) decrease of 3 dB corresponding to a signal increase on $\approx {}^1/\sqrt{2} \approx 0.71$ of F_{max} occurs for smaller F
F_{cr}	right cut-off: frequency at which (the first) decrease of 3 dB corresponding to a signal increase on $\approx {}^1/\sqrt{2} \approx 0.71$ of F_{max} occurs for greater F
f_{equiv}	function to map BWE-equivalent on BWE
F_n	frequency component of discrete Fourier transform
FN	false negatives
FP	false positives
F_{peak}	peak frequency: frequency at which maximum magnitude is achieved
FPR	false positive rate
$\mathbf{f}_{rec}$	reconstruction function from original to reconstructed (by time delay embedding) state space via one-dimensional time series
f_{sa}	sampling frequency
f_{td}	reconstruction function from time series to reconstructed (by time delay embedding) state space

h	measurement (or read-out) function from m-dimensional state space to real line
$H_D(l)$	Histogram of diagonal lines in a recurrence plot
θ	Heaviside function; $\theta(x) = 1$ for $x > 0$, $\theta(x) = 0$ for $x < 0$
JR	joint recurrence matrix (joint recurrence plot if plotted)
L	RQA feature, average length of diagonal lines (of length greater than predefined l_{min}) in recurrence plots
λ	wavelength
L_{nor}	RQA feature, average length of diagonal lines (of length greater than predefined l_{min}) in recurrence plots, normalised by the according l_{min} to get values between 0 and 1
$\mathcal{M}$	state (or phase) space of a dynamical system
m	dimension of original state space
M	length of reconstructed (time delay embedded) time series
$\widetilde{M}$	length of reconstructed time series, equalling number of diagonal lines in recurrence plot, reduced by Δ_M for calculation of TND and COR
μ	Poisson's ratio
$\mathbb{N}$	set of natural numbers
N	length of time series (number of data points)
Ω	set of all non-wandering points of a point $\mathbf{x}_0$ in state space
p_i	sound pressure of an incident wave
p_r	sound pressure of reflected wave
p_t	sound pressure of transmitted wave
Q_1	first quartile of a data set; median of the first half of a data set sorted in ascending order (for an even number of data points; for an odd number, the 'half' ends with the median)

Q_3	third quartile of a data set; median of the second half of a data set sorted in ascending order (for an even number of data points; for an odd number, the 'half' starts with the median)
QCD	quartile coefficient of dispersion
$\mathbb{R}$	set of real numbers
$\mathbf{R}$	recurrence matrix (recurrence plot if plotted)
ρ	material density
r	(Pearson's) correlation coefficient
R	coefficient of reflection, $R = p_r/p_i$
$RATIO$	DET/RR (RQA feature)
RR	recurrence rate (RQA feature)
RR_k	local recurrence rate per diagonal
s	standard deviation of a sample (of a time series)
s^2	variance (standard deviation squared) (of a time series)
$s_x(\upsilon)$	autocovariance (of a time series)
s_{xy}	covariance (of two time series)
T	(time) period
$\mathcal{T}$	time as a set
t	time, $t \in \mathcal{T}$
$thres$	decision threshold to classify into positive and negative from feature
TN	true negatives
TND	RQA feature trend
TNR	true negative rate
t_p	CFRP ply thickness
TP	true positives
TPR	true positive rate
T_t	coefficient of transmission, $T = p_t/p_i$
U	neighbourhood of a point in (state) space
υ	time shift of autocovariance
$\mathbf{v}$	function of a dynamical system (continuous)
w	weight (coefficient) of linear model in simplest form
w_e	excluded window, auxiliary variable to exclude central diagonals of the recurrence plot from calculation of RQA features

$\mathbf{w}_{FFT}$	weights (coefficients) of linear model applied on Fourier transformed time series
w_{th}	Theiler window, auxiliary variable to exclude w_{th} diagonal lines near to the main diagonal of the recurrence plot from the calculation of RQA features
$\mathbf{w}_{time}$	weights (coefficients) of linear model applied on time series (column vector)
x	input of linear model (scalar)
$\mathbf{X}_{FFT}$	matrix with (portions of) amplitude spectra of intermediate echo time series in rows
x_i	(elements of an) (intermediate echo) time series, $i = 1, 2, ..., N$, taken here at constant time intervals
$\mathbf{x}_i$	reconstructed time series as d-dimensional vector, $i = 1, 2, ..., M$
$\mathbf{X}_{ie}$	matrix with intermediate echo time series $\mathbf{x}_i$ in rows
$\mathbf{x}_M$	state in original state space
x_{max}	maximum absolute value (of a time series)
$\bar{x}$	mean of a sample, here generally of a time series x_i
y	ouput of linear model (scalar)
Z	(specific) acoustic impedance, $Z = \rho c$

Acronyms

See the Symbols list above for acronyms that are also variables and written italic, such as *BWE* (back-wall echo) or *RR* (recurrence rate).

BWE-equivalent	back-wall echo equivalent
CFRP	carbon fibre reinforced polymer
CRP	cross recurrence plot
DFT	discrete Fourier transform
DRP	difference recurrence plot
FFT	fast Fourier transform
JRP	joint recurrence plot
LOI	line of identity
NDE	non-destructive evaluation
NDT	non-destructive testing
PoD	probability of detection
ROC	receiver operating characteristic
RP	recurrence plot
RPA	recurrence plot analysis
RQA	recurrence quantification analysis
TCG	time corrected gain

1 Introduction

Signal processing and data analysis are an integral part in many engineering domains and, with the ascend of machine learning, have been further increasing in recent years. Standard tools in time and frequency domain are based on the assumption that the signals stem from a linear system. Non-linear time series analysis on the other hand is rarely used in engineering [cf. Colloquium 2018]. One tool within non-linear time series analysis is recurrence quantification analysis (RQA), which derived from the introduction of recurrence plots (RPs) in 1987 within the scope of research on chaotic dynamical systems. Nowadays, RQA is being used in numerous disciplines from earth sciences to biology and physiology [cf. Marwan et al. 2007: 242]. Usage of RQA has been constantly rising over the last years [cf. Website RP Bibliography 2020]; it plays however yet a minor or no role in signal processing and time series analysis in general [e.g. Alessio 2016; Mitra 2011]. One reason for the low amount of approaches to use this tool in the engineering domain may be the restricted computer power (and memory) just, say, 20 years ago, which made it costly to use recurrence quantification in an industrial context. Another reason may be that current engineering courses in signal processing focus on Fourier analysis [e.g. Alessio 2016; Oppenheim et al. 2014] — i.e. the decomposition of a time signal into its periodic, frequency components — and thus on the linear view of this signal. Its representation in state or phase space, basis for non-linear time series analysis, is unfamiliar to engineers (with some exceptions in control engineering, [cf. Hsu 2014: chap. 7; Ludyk 1995]). The proof of the theorem of delay embedding — main method to reconstruct a multidimensional representation in state space out of a one dimensional signal — involves differential topology, with which again engineers are not acquainted; the method itself though is rather simple [cf. Huke 2006].

This work compares recurrence quantification analysis with tools in time domain and Fourier analysis (especially regression models) for an engineering application. Transient and thus nonstationary signals from ultrasonic testing are evaluated to detect porosity in carbon fibre reinforced polymer.

1.1 The task

Carbon fibre reinforced polymers (CFRP) play an increasing role in aeronautic industry due to their high ratio of mechanical strength to weight. Their potential to save weight and, thus, fuel burn and costs has led to an increasing usage in aircrafts, for example from 25% in the Airbus A380 to 53% in the Airbus A350XWB [Brandt 2010]. 100% non-destructive testing (NDT) of produced CFRP laminates is standard for material used in structural components of Airbus aircrafts, generally performed using

ultrasonic testing in pulse-echo mode [Brandt & Maaß 2015]. Defects such as delaminations (material separations between plies of the layered material) or porosity have to be detected [Schnars & Henrich 2006]. Porosity is evaluated via the height of the so-called back-wall echo (*BWE*), the reflection of the ultrasonic wave at the opposite side of the part. However, there are special situations — complex geometries, bonded components or sandwich structures (consisting of a honeycomb core between two CFRP laminates) — in which this analysis is impossible [cf. Lozak et al. 2014; Brandt & Maaß 2015]: parts have to be designed as if maximum porosity were present. It is therefore beneficial to create an evaluation method without needing the back-wall echo, solely using the echoes of the inside of the inspected part (intermediate echoes) to create a back-wall echo equivalent (BWE-equivalent).

This work aims at investigating whether and how recurrence quantification analysis (RQA) can generate a back-wall echo equivalent for the detection of porosity in carbon fibre reinforced polymers, compared with Fourier analysis and some methods in time domain.

1.2 Structure of the thesis

This thesis is organised as follows: chapter 2 describes the physical background of the data treated in this work, CFRP and ultrasonic testing. Chapter 3 explains the methods for analysis of the data, including recurrence quantification analysis, Fourier analysis and regression models, to obtain a back-wall echo equivalent, the goal of this work. Furthermore, methods of correlation and classification to determine the quality of the back-wall echo equivalent are described. Chapter 4 collects the state of the art especially regarding detection of porosity without a back-wall echo. The measurements and evaluation on different CFRP materials, which form the central part of this thesis, are presented in chapter 5. The thesis ends with the conclusion in chapter 6.

2 Ultrasonic Testing of Carbon Fibre Reinforced Polymers

This chapter describes the process that generates the data analysed in this work. After briefly explaining carbon fibre reinforced polymers in section 2.1, section 2.2 describes ultrasound in general and ultrasonic testing as a method of non-destructively detecting defects in materials, especially the standard procedure for detection of porosity in CFRP.

2.1 Carbon fibre reinforced polymers

Composite materials or composites consist — in the restricted sense — of one continuous (matrix) material and one or more discontinuous (e.g. fibre) materials. The discontinuous material (the reinforcement) is the load-carrying constituent of the composite; the surrounding matrix keeps the discontinuous material in place, transfers stresses between these fibres and protects them against environmental influences. The composite material has as a result properties which are more than the sum of the properties of the two or more constituent materials alone [e.g. Berthelot 1999: 3ff.]. When using fibres as discontinuous material, the composite shows anisotropy, i.e. it exhibits different material properties in different directions. In fibre direction, the tensile strength for example is distinctly higher than perpendicular to fibre direction [Hull & Clyne 1996: 1f.; Berthelot 1999: 5].

Carbon fibre reinforced polymer (CFRP) is a composite consisting of a polymer material as matrix and carbon fibres as discontinuous material [e.g. Mallick 2008: 1ff.]. Epoxy resin is a thermoset polymer (molecules forming a rigid three dimensional network structure, which cannot be melted [Mallick 2008: 62]) frequently used as matrix for CFRP in aerospace industry. It is based on the epoxy (epoxide, oxirane) atom group [Bunsell & Renard 2005: 110f.; Chawla 2013: 81]. Carbon fibres have a graphitic structure with carbon atoms arranged in hexagonal layers; see Hull & Clyne [1996: 9ff.] and Chawla [2013: 24f.] for further details.

One way of producing CFRP parts is to build them up from prepregs: these are layers of fibres (called laminae or plies) with thickness typically in the 0.1 mm range, which are pre-impregnated with resin [Chawla 2013: 145; Hull & Clyne 1996: 40, 43ff., 274]. The fibres can be either orientated in one direction per layer (unidirectional lamina) or can be woven with fibres being perpendicular to each other to form a fabric. The plies are placed upon each other and then cured in an autoclave, applying pressure and heat to the part [Chawla 2013: 145ff.; Hull & Clyne 1996: 275]. For unidirectional material, the direction of the fibres in the different plies (the lay-up or

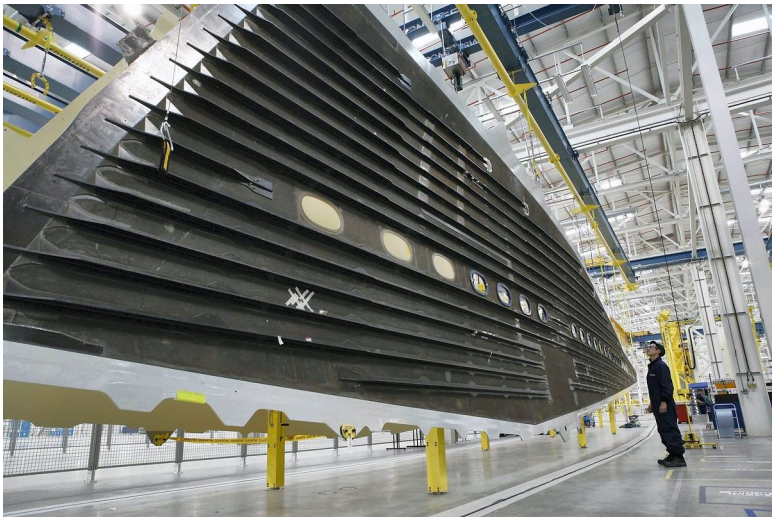


Figure 2.1: CFRP wing cover with stringers (stiffening elements) of an Airbus A350XWB (©Airbus S.A.S. — photo by H. Goussé, master films)

stacking sequence) has to be taken into account. In aircraft industry, plies are placed upon each other typically in a way that the fibre direction differs by 0°, 45° or 90° [Chawla 2013: 401f.]. The final surface of a manufactured CFRP part may be covered with peel-ply [MIL-HDBK 2002: 2-50], a fabric that is removed after curing of the CFRP before an additional bonding step.

All specimens in this work are from carbon fibre reinforced polymers that are used in aircraft industry containing an epoxy matrix.

The most important defects in CFRP for aircraft production are delaminations — areas in which two plies of CFRP have not been connected during curing — and pores (porosity), which may occur if cure parameters are outside the qualified range [MIL-HDBK 2002: 2-58], especially pressure [Rubin & Jerina 1994]. Pores are bubbles of trapped air, gas or vacuum [MIL-HDBK 2002: 1-27]. Porosity decreases some mechanical properties, especially interlaminar shear strength [cf. Brandt & Maaß 2016; Mehdikhani et al. 2019], at least if its amount exceeds approximately 3% in volume [cf. Niu 2000: 506].

A method to join two CFRP parts is the cobonding process. For example an already cured stringer (a stiffening element at the inside of the fuselage or the wing of an aircraft, cf. fig. 2.1) is placed on the not-yet-cured prepreg of the fuselage shell or wing cover with an adhesive film inbetween. This mutual part is then cured in the autoclave; thus the production of one part and joining to the other part is achieved in one step [MIL-HDBK 2002: 2-52].

CFRP is also used as a ‘skin’ of sandwich parts: a core — often resin impregnated paper forming honeycomb cells — is covered on top and bottom with CFRP [Niu 2000: 111ff.; Mallick 2008: 4].

2.2 Ultrasonic testing

This section reviews ultrasound as well as its use for non-destructive testing.

2.2.1 Ultrasound

Sound are pressure changes which propagate through a medium — solid, liquid or gas. A transducer (cf. sect. 2.2.2) as one source of sound oscillates in form of vibrations, i.e. its surface moves back and forth. These movements introduce pressure changes into the adjacent medium, which propagate through this medium. This propagation of pressure changes is one form of a mechanical wave. When used for non-destructive testing, these pressure changes are always elastic deformations of the tested solid part, i.e. no permanent alteration of the material is induced.

Two principal types of waves can be distinguished: a compression or longitudinal wave consists of alternating compression and dilatation, i.e. the particles of the medium move fore- and backwards in the direction of wave propagation (sketched in fig. 2.2). A shear or transverse wave exists if material particles move up and down, perpendicular to the direction of wave propagation. In this thesis, longitudinal waves and propagation perpendicular to the surface (at which the waves enter the part, i.e. normal incidence) are employed. The dimensions of the inspected part do not influence the wave propagation (besides the reflection on the side opposite to that surface at which the wave enters), the part under inspection is significantly larger than the size of the sound beam. These waves are often called bulk waves [Schmerr & Song 2007; Cheeke 2002] to distinguish from guided waves [Rose 2014: 5] (cf. sect. 2.2.2). The smallest time in which one point of the material, in which a wave propagates, returns to the same state of the particle (displacement and velocity) is called (time) period T. Its reciprocal is the frequency f [in Hz = 1/s] of the particle vibration and thus of the wave. The smallest (spatial) distance between two points of the wave of same particle state (displacement and velocity) is the wavelength λ (fig. 2.2). Thus the wave velocity c is

$$c = \lambda f \tag{2.1}$$

(also velocity or speed of sound in the case of mechanical waves).

Ultrasound or ultrasonics is sound above the audible range. Obviously because the latter depends on individuals, the lower limit of ultrasound is given between (15–)16 kHz [Blitz 1967: 1; Šutilov 1988: 1] and 20 kHz [Cheeke 2002: 2; Krautkrämer & Krautkrämer 1990: 1; Halmshaw 1987: 108; Richardson 1962: 1]. Occasionally, an upper bound of 1 GHz is set, beyond which the hypersonic regime may be defined [Cheeke 2002: 2].

The velocity of longitudinal waves c_L as well as the (different) one for shear waves in solids c_T depends on the modulus of elasticity (Young's modulus) E, on Poisson's ratio μ and on the density of the material ρ [see Krautkrämer & Krautkrämer 1990: 13f.; Blitz 1967: 150f.]. The two velocities are linked via Poisson's ratio [Krautkrämer & Krautkrämer 1990: 14]; $c_T \approx c_L/\sqrt{(3)}$ for steel for example. In composite materials, the velocity of sound is composed of the velocities of the individual constituents

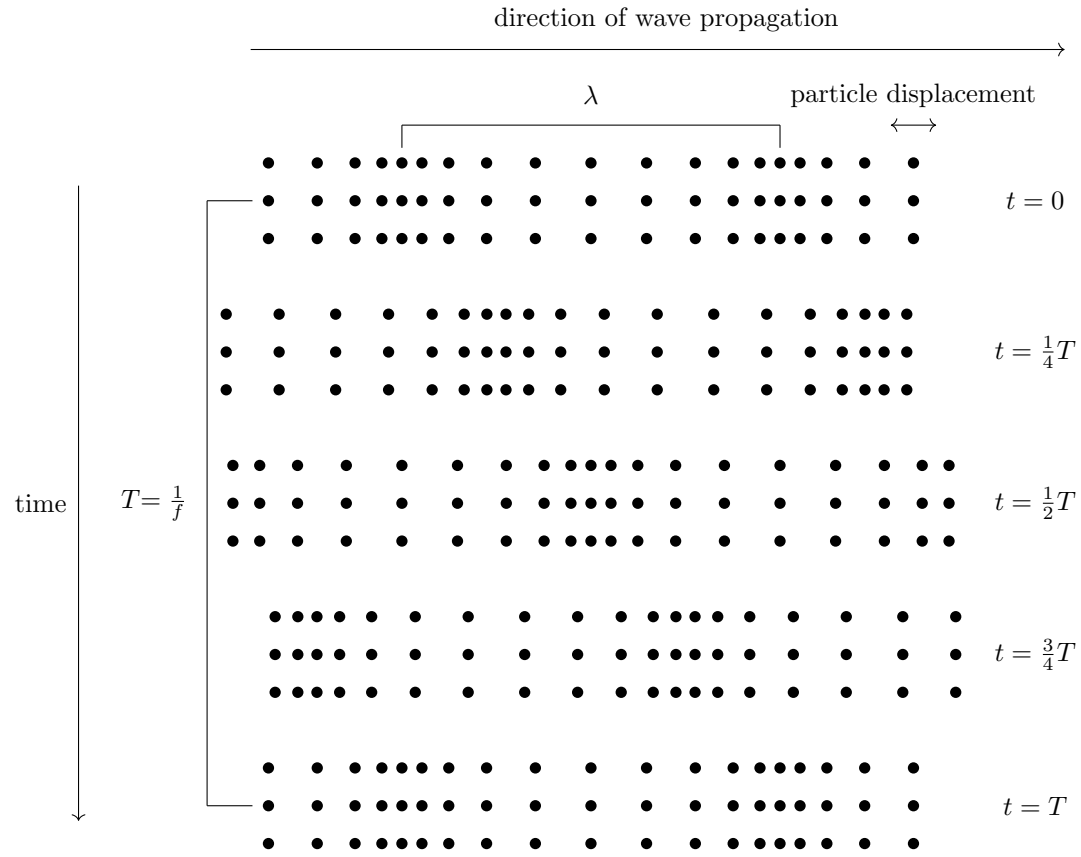


Figure 2.2: Propagation of a longitudinal (compression) wave with period T, frequency f and wavelength λ; particles are only displaced around its original position

and is often determined purely experimentally. For CFRP, the velocity is roughly $c_L \approx 3000\,\mathrm{m/s}$.

The difference in (specific) acoustic impedance Z, the product of density ρ and sound velocity c,

$$Z = \rho c, \tag{2.2}$$

between two materials determines the amount of reflection and transmission of an ultrasonic wave at their interface [Krautkrämer & Krautkrämer 1990: 13; Blitz & Simpson 1996: 16]. The coefficients of reflection and of transmission are the ratios of the sound pressures of the reflected and the transmitted wave, respectively, to the sound pressure of an incident wave hitting an interface perpendicularly (cf. fig. 2.3),

$$R = \frac{p_r}{p_i}, \tag{2.3}$$

$$T_t = \frac{p_t}{p_i}. \tag{2.4}$$

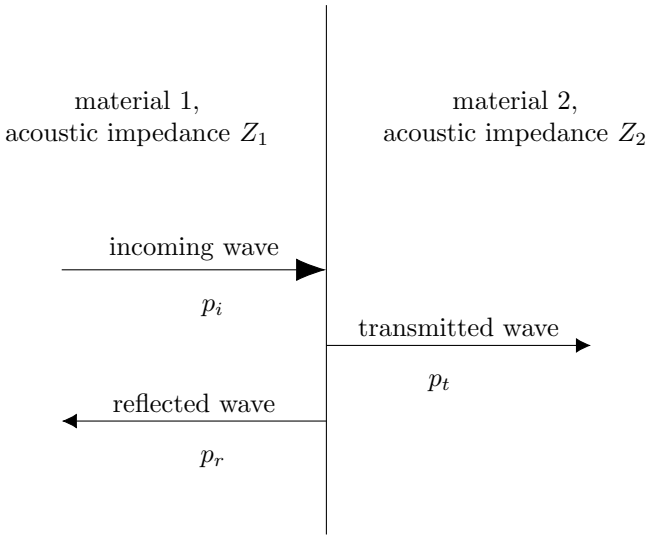


Figure 2.3: Reflection and transmission of ultrasound at perpendicular incidence

They can be determined via

$$R = \frac{Z_2 - Z_1}{Z_2 + Z_1}, \tag{2.5}$$

$$T_t = \frac{2Z_2}{Z_2 + Z_1}, \tag{2.6}$$

where Z_1 and Z_2 are the acoustic impedances of the first and second material. See Rose [1999: sect. 4.2] for a derivation of equation (2.5) and (2.6).

Ultrasonic inspection under oblique incidence, i.e. if a wave hits an interface under an angle different from 90°, is not part of this work. The occurring phenomena of refraction (transmission under an angle different from the angle of the incidence wave) and mode conversion are described e.g. in Blitz & Simpson [1996: sect. 2.5.2] and Rose [1999: chap. 5].

If the wavelength of the ultrasonic wave is small compared to the size of the reflector, specular reflections take place. At the edges of finite reflectors or sources such as ultrasonic transducers the phenomenon of diffraction occurs, due to the fact that reflector or sources can be thought of as composed of numerous single point sources, which emit a spherical wave each, according to Huygens principle. This diffraction leads to constructive and destructive interference in the near field of a transducer with maxima and minima of sound pressure, see e.g. Blitz & Simpson [1996: sect. 2.7] and Cheeke [2002: sect. 6.1.2.2] for details. The maxima and minima are however not that extreme for ultrasonic testing in practice, because transducers send ultrasonic pulses containing numerous frequencies [c.f. Rose 1999: A10].

Scattering can (implicitly) refer to the interaction of a wave with reflectors of any size [e.g. Cheeke 2002: 104; DeSanto 1992: 79f.], but often it refers to the interactions of reflectors of size in the magnitude of wavelength or smaller [Morse & Ingard 1968: 449; Blitz & Simpson 1996: 45f.; Fahr 2014: 186]. For these sizes, a sphere reflects a rather

half-spherical wave [Cheeke 2002: sect. 6.2.2; cf. Fahr 2014: 186]. If the diameter of a spherical scatterer (reflector) d_s is much smaller than the wavelength, Rayleigh scattering occurs [Blitz & Simpson 1996: sect. 3.3.1; Blitz 1967: 155; cf. Adler et al. 1986; Smith 2010: 59ff.], which is proportional to the fourth power of frequency and the cube of diameter, f^4 and d_s^3.

An ultrasonic wave travelling through CFRP is attenuated, and thus, performing ultrasonic testing, an identical reflector in a greater depth is shown with a smaller echo. The attenuation [Fahr 2014: sect. 7.1.4; Rose 1999: 355] in CFRP is due to geometrical beam opening in the far field (beyond the near field); reflections at the CFRP–resin–CFRP interface between two plies; scattering at the carbon fibres within the resin matrix — and at pores — and absorption due to transformation of elastic motion into heat. Furthermore, the ultrasonic energy entering the part can be largely reduced by a rough surface due to scattering [Krautkrämer & Krautkrämer 1990: sect. 15.1; Henneke 1990; Smith 2010: 82].

Potential reflections between the plies of CFRP (as for the materials investigated in this work) occur because a pure resin layer (few 10 µm thin) exists between two CFRP plies with slightly different acoustic impedance compared to the CFRP [cf. Dominguez 2005: 13]. This resin layer acts as a thin reflector and interference occurs due to reflection and transmission [Krautkrämer & Krautkrämer 1990: sect. 2.2] at the CFRP–resin and resin–CFRP interfaces [Smith 2010: sect. 4.2.2.]. Furthermore, a resonance effect occurs: if CFRP plies are of thickness $t_p \approx \lambda/2$, a wave reflected at a resin layer between e.g. ply 7 and 8 adds constructively to its reflection between ply 6 and 7 (fig. 2.4). This effect occurs at least for pulses (wave trains) that are two wavelengths long, cf. fig. 2.4, and leads to rather sinusoidal intermediate echo time series. It is designated resonance effect henceforth [cf. Smith 2010: sect. 4.2.3; Dominguez & Mascaro 2006; Chen et al. 2016], though multiple reflections and standing waves within one ply should not play a role due to the low reflection coefficient: thin resin layers of few 10 µm between CFRP plies have a reflection coefficient of roughly 0.1 [Smith 2010: fig. 4-2]; a reflection that is reflected again twice reduces to $\approx 1\%$ of the original reflection and cannot noteworthy contribute.

2.2.2 Ultrasonic inspection, a non-destructive test

Non-destructive testing (NDT) comprises (besides visual inspection) methods such as eddy current testing, radiography (X-ray testing), penetrant testing, thermography and ultrasonic testing [Halmshaw 1987; Fahr 2014]. It primarily aims at detection of imperfections (discontinuities) of materials without decreasing the integrity and mechanical performance of the tested part. These techniques are used in production as well as in maintenance of a product or its parts and are embedded in a quality control process [cf. Große 2016].

Non-destructive testing, non-destructive evaluation (NDE) and non-destructive inspection (NDI) are often used interchangeably [ISO/TS 2005: 3; Henneke 1990: ix]. In some contexts, NDE has been given other meanings [Ehrhart 2016: 8; Halmshaw 1987: 2]. Within this thesis, NDT is accepted as the most common expression [DIN EN 2015;

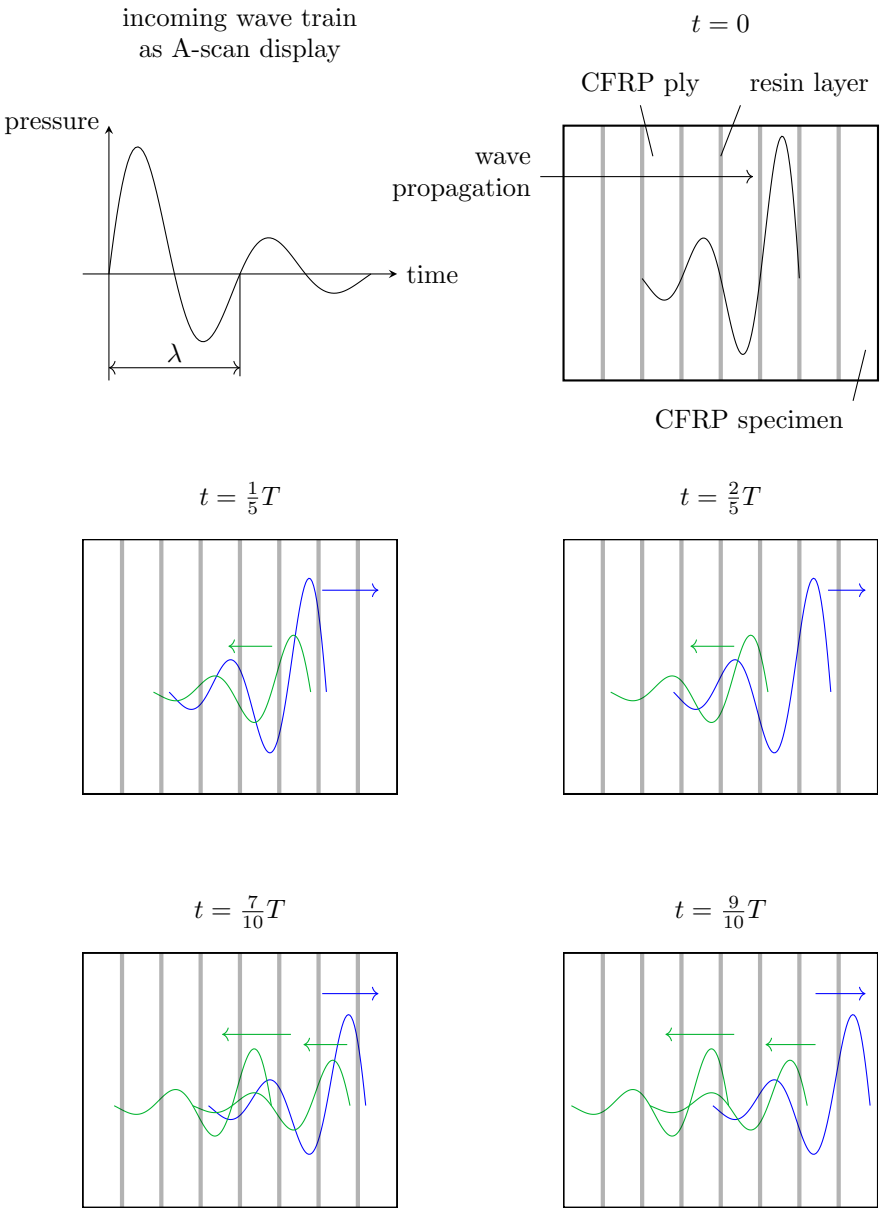


Figure 2.4: Resonance effect of a longitudinal (compression) wave caused by resin layer between each two CFRP plies, CFRP ply thickness $t_p \approx \lambda/2$; transmitted wave in blue, reflected waves in green: reflections at subsequent resin layers (here between CFRP plies 6 and 7, and 7 and 8) superimpose constructively

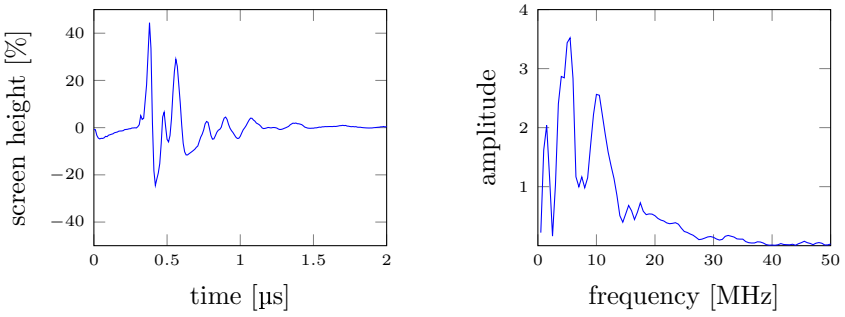


Figure 2.5: Reflection of the probe used in this thesis (sect. 5.1.1), put on the bottom of the immersion tank filled with water: A-scan (left) and its amplitude spectrum (right); peak frequency $F_{peak} \approx 5\,\mathrm{MHz}$

Halmshaw 1987; Henneke 1990; Int. Committee f. NDT 2019; World Conf. of NDT 2019], whereas NDE may emphasise the aspect of the analysis performed on data generated by NDT.

Ultrasonic testing for materials uses the propagation of ultrasonic waves and their interaction with the inspected part. Reflections at material discontinuities; reductions of reflected echoes at boundaries of the part or changes in wave velocity are examples for utilised effects. Different methods exist, partly based on guided waves [Rose 2014] (such as surface waves or plate waves [Cheeke 2002: chap. 8–10 and 15; Schmerr & Song 2007: sect. E.2, E.3]), which are not part of this thesis. Within this work, longitudinal waves (sect. 2.2.1) are used in pulse-echo mode [Blitz & Simpson 1996: sect. 5.3]: this method is generally employed on Airbus aircraft CFRP parts to inspect for inner defects [Brandt & Maaß 2016]; one transducer (also known as probe) sends an ultrasonic pulse (a pulsed wave, fig. 2.5) into the part and receives the reflected echoes. The echo time series generated is called A-scan (fig. 2.6 left). The centre frequency of the amplitude spectrum (sect. 3.2) of the reflection of this pulse at a defined reflector (cf. fig. 2.5) is the operating frequency [DIN EN 2016] or working frequency [Krautkrämer & Krautkrämer 1990], referred to as inspection frequency for the remainder of this work. Transducers are devices that transform one form of energy into another; ultrasonic transducers transform electrical into mechanical energy and vice versa. The piezoelectric effect is mostly used [Cheeke 2002: 210f.; Blitz & Simpson 1996: 52ff.]: a voltage causes a piezoelectric material to compress or expand, depending on the direction of the current (indirect or converse piezoelectric effect). Compression or expansion inducing a voltage in the material is the direct piezoelectric effect.

In this work, an ultrasonic array transducer is used [Drinkwater & Wilcox 2006; Schmerr 2014]. This is a multi-element probe in one housing (fig. 2.7), having for instance in one row 128 elements. Each element can be steered independently with the appropriate electronics. In this work only linear scanning is used: the elements of an aperture (a subset of all elements) are used simultaneously. With one aperture, one A-scan is recorded. The subset of elements changes, for example element 1 to 16 is

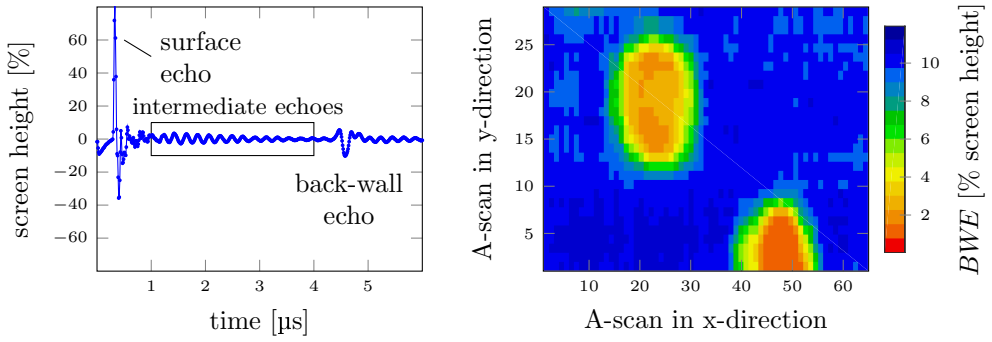


Figure 2.6: A-scan (left, example of measurements of fabric specimens, M15-1x41y12, cf. page 52 for designation; ≈ 4 µs time of flight between surface and back-wall echo correspond — with wave velocity of ≈ 3000 m s^{-1} — to ≈ 7 mm thickness), and C-scan (right) of same specimen, showing the height of the back-wall echo. The A-scan corresponds to one pixel in the C-scan

activated, then 2 to 17, 3 to 18 etc. (cf. fig. 2.7). This way, several A-scans along one line are generated without moving the transducer — an electronic scan. Moving the ultrasonic array transducer in one direction (perpendicular to its length), a grid of A-scans is recorded. The whole set of A-scans in one recording is henceforth designated volume scan.

Phased arrays is a common expression for such multi-element transducers, since they offer the opportunity for beam forming — by delaying the pulsing of the elements in one aperture — to achieve focussing or oblique incidence of the sound beam [Schmerr 2014: 3]. A recent approach is the use of full matrix capture (FMC) and total focussing method (TFM), for which each element is fired once and all elements are receiving, capturing the whole information a phased array can obtain [Grager et al. 2016, 2018]. However none of these methods are applied in this work.

An A-scan (fig. 2.6 left) shows echo heights (corrsponding to the sound pressure) in % screen height. Because these are arbitrary units depending on the amplification, the % is skipped in all further representations (such as embeddings) in this work. A B-scan (not used in this thesis) is a section of a part with several A-scans, showing the echo heights in colours or grey shades [cf. Blitz & Simpson 1996: 97ff.]. A C-scan shows one feature for each of the A-scans of an area of the part (or the whole part, respectively) as a top view of the area (or part) (fig. 2.6 right). This feature can be the amplitude or the time of flight of an echo signal (the latter being the point in time at which it appears); the echo can be for example the one with the maximum amplitude, or the one with the smallest time of flight of all signals greater than a predefined threshold. The signal is taken either from the whole A-scan or from a portion of it, the so-called gate.

For CFRP in aeronautic industry, the inspection is mostly — as is in this thesis — performed in normal incidence, i.e., the ultrasonic wave propagates perpendicularly to the surface at which it enters the part. Due to the large difference in acoustic

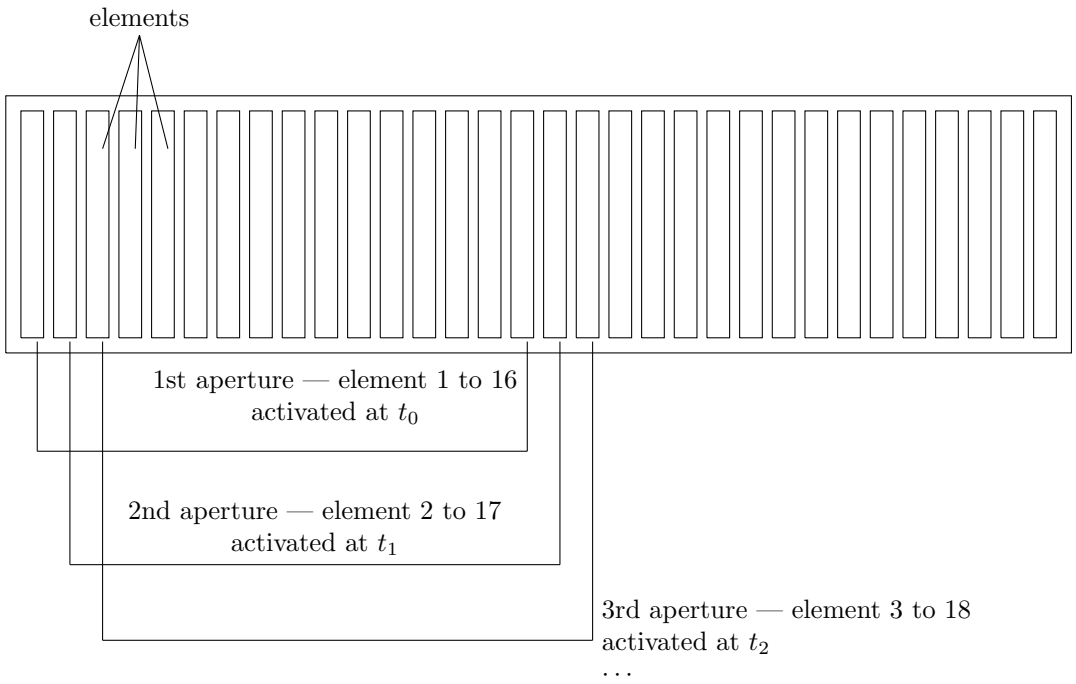


Figure 2.7: Sketch (bottom view) of an ultrasonic array transducer with 32 elements with example of first 3 apertures of a linear scan

impedances of air and solids, most often a liquid is used to couple the transducer to the solid under inspection [Fahr 2014: sect. 7.1.8] — in contrast to air-coupled ultrasound [e.g. Schnars & Henrich 2006; Fahr 2014: sect. 7.5.1], which can be used for e.g. sandwich inspection.

In manual inspection in contact technique, the couplant is often coupling gel or water under a probe moved by hand, potentially with a plastic delay line, directly placed on the part [Krautkrämer & Krautkrämer 1990: 190f.].

In automatic inspection, probes are often coupled in contact technique [Engl et al. 2003] by constantly providing water in an automated way, or the probes are placed in a water tank (immersion testing), with a certain distance to the part under test [Blitz & Simpson 1996: sect. 7.5].

Geometrical features such as the front surface (where the sound wave enters the part) and the opposite surface (the so-called back-wall), reflect the ultrasonic wave when inspecting a part. The reflection from the back-wall is commonly called back-wall echo (*BWE*). The echoes between surface and back-wall echo may be referred to as intermediate echoes (cf. fig. 2.6), also known as backscatter, cf. section 4.1; the whole of intermediate echoes will be called intermediate echo time series in the remainder of this text. Defects such as delaminations produce additional echoes (fig. 2.8). Furthermore, the *BWE* can be used as a measure for the overall reflections within the part, especially for detecting porosity, cf. sect. 2.2.3. In this usage, the *BWE* is similar to the through-

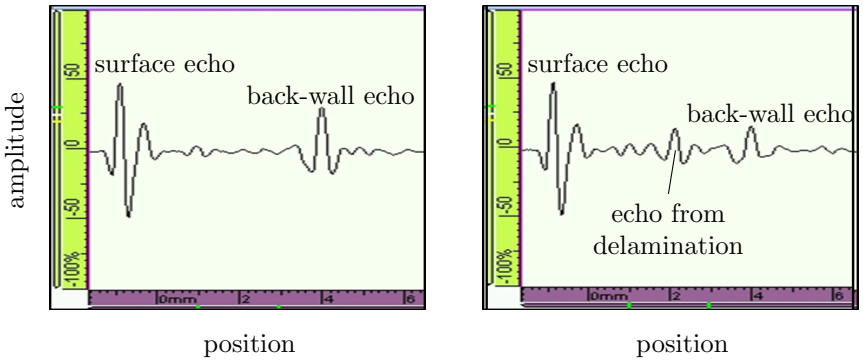


Figure 2.8: A-scan of a part without (left) and with delamination (right; delamination is smaller than the sound beam: back-wall echo reduced, but still existent) [figure from Brandt & Maaß 2016]

transmission inspection (if used under normal incidence), a different type of ultrasonic test: two transducers are used on either side of the part under inspection, one sending and one receiving [Krautkrämer & Krautkrämer 1990: chap. 12]. The received signal gives information about potential defects — similar to the way the *BWE* does.

For pulse-echo inspection often a time-corrected gain is used, which sets amplification dependent on the time of flight (and thus the depth) in a way that echoes of similar reflectors in different depths have similar heights [Blitz & Simpson 1996: 116]. Time corrected gain thus compensates for attenuation (cf. sect. 2.2.1).

2.2.3 Ultrasonic inspection for porosity in CFRP

The evaluation of ultrasonic signals in industrial inspection begins with observing any additional echo in the A-scan, compared to data from ultrasonic inspection of a defect-free part. Such echoes arise from additional interfaces in the part as for example delaminations (fig. 2.8). Porosity however presents numerous small scatterers, which either reflect back to the probe (backscatter [Blitz & Simpson 1996: 46]) — but give only very small echoes that cannot be evaluated directly — or they reflect in other directions than back to the probe. In both cases, the ultrasonic energy reaching the back-wall of the part is reduced. Thus, porosity can be detected via a reduction of the back-wall echo (*BWE*), which is the industrial method used at Airbus in pulse-echo normal incidence ultrasonic testing [Schnars & Henrich 2006]. The existence of delaminations at the inspection location is excluded in a step before, by verifying that no significant additional reflections show up in the A-scan.

The standard approach for evaluation of porosity is to compare the back-wall echo height of a potential indication with a back-wall echo average in a nearby defect-free area of similar build-up and thickness. For Airbus, a reduction of the *BWE* by half, by 3/4 or 7/8 (approximately corresponding to 6 dB, 12 dB or 18 dB in logarithmic scale), depending on thickness of the part, is considered critical [Brandt & Maaß 2015].

Several situations exist in which an ultrasonic inspection of CFRP structures cannot be evaluated on porosity in the way that is described above, such as

- *Co-bonding* (sect. 2.1) of complex parts as for example for skin-stringer structures (cf. fig. 2.1). No back-wall and thus no *BWE* is available in the area of the stringer web.
- *Sandwich structures* (sect. 2.1). The bonding of the skin to the core presents a varying echo.
- *Complex geometries.* In case of non plane-parallel parts, the back-wall echo may not or only partly reach the probe. Especially in case of spherically shaped structures with varying angles between surface and back-wall, the *BWE* varies.

For these examples, either the back-wall echo (or alternative echoes in its place) can vary due to other reasons than porosity, such that reductions of the *BWE* can be mistaken for porosity; or the back-wall echo does not exist.

If porosity cannot be detected, as in these cases, the structure has to be designed as if maximum porosity were present, due to the decreasing effect of porosity on mechanical properties. This thesis aims at creating a representation of the back-wall echo, the back-wall echo equivalent (BWE-equivalent), out of the intermediate echoes of the A-scan.

The dependence between porosity and *BWE* height has been investigated in numerous publications [e.g. Daniel et al. 1992; Jeong & Hsu 1995; Birt & Smith 2004], see Mehdikhani et al. [2019] and Smith [2010: sect. 2.3.2] for overview, also regarding the effect of porosity on ultrasonic velocity.

An overview about research approaches in the literature to find a back-wall equivalent is given in chapter 4, after describing the methods used in this thesis in chapter 3.

3 Analysis and Classification of Time Series Data

This chapter explains the methods of data analysis that are applied in chapter 5. The goal is to enable an inspection of carbon fibre reinforced polymer (CFRP) for porosity with ultrasonic non-destructive evaluation in situations in which it is not performed today. The ultrasonic signals are scalar values that evolve over time, recorded discretely at regular intervals. Such an observation over time, a 'finite... sequence of real numbers' [Broer & Takens 2011: 206] is called a (discrete) time series [Anderson 1994: 1; Prado & West 2010: 1]

$$x_i, \;\; i = 1, 2, \ldots, N, \;\; N\text{: length of time series} \tag{3.1}$$

(not to be confused with an infinite series [e.g. Beals 1987: 14]). Time series generated in ultrasonic testing are called A-scans (cf. sect. 2.2.2); time series used in this work consist of the intermediate echo portion between surface and back-wall echo of the A-scan (fig. 2.6) and are called intermediate echo time series in this text.

In standard inspection, the echo of the opposite side of the part, the back-wall echo (*BWE*), is used for the evaluation of porosity; this is impossible in special inspection situations (cf. sect. 2.2.3). This thesis aims at generating an equivalent to the *BWE* out of the intermediate echo time series:

$$BWE_{equiv} = f_{bwee}\left(x_i\right). \tag{3.2}$$

f_{bwee} in equation (3.2) are in this work functions from recurrence quantification analysis (sect. 3.3.5), Fourier analysis (sect. 3.2) or approaches in time domain (sect. 3.1). The aim is to resemble the back-wall echo (which is available for the specimens used in this thesis, cf. sect. 5.2). The goodness of this resemblance is assessed via correlation (regression) and classification (sect. 3.4).

3.1 Approaches in time domain

The term time series analysis is often used for methods regarding time series as being generated by a random (stochastic) process, which cannot be (fully) deterministically described [Priestley 1981: 10f.; 100ff.; Brockwell & Davis 1991: 8f.; Prado & West 2010: 1]. A prominent example for models describing such a process are the autoregressive and the moving average processes (or the combination of both), where the values of a time series depend partially on a purely random (white noise) process.

Several simpler approaches than the one described above are utilised in this work. It starts with the maximum absolute value of the intermediate echo time series (the maximum echo)

$$x_{max} = \max_{i=1,\dots,N} |x_i|. \tag{3.3}$$

Seeing a time series as a data set, disregarding the order (in time) of the individual data points, a further basic evaluation is the variance (standard deviation squared) of a sample

$$s^2 = \frac{1}{N-1} \sum_{i=1}^{N} (x_i - \bar{x})^2, \tag{3.4}$$

with $\overline{x}$ being the mean of the data series.

A further example is the quartile coefficient of dispersion (QCD) [NIST 2017], also called coefficient of quartile variation (CQV) [Kokoska & Zwillinger 2000: 15]

$$QCD = \frac{Q_3 - Q_1}{Q_3 + Q_1}. \tag{3.5}$$

Q_1 and Q_3 is the first and third quartile, which is the median of the first half and the second half, respectively, of a data set sorted in ascending order (for an even number of data points; for an odd number, each 'half' ends or starts, respectively, with the median of the whole data set).

An important tool for this work is based on linear regression. Linear regression is concerned with the dependence of an output (also explained variable or response) y from an input (also explanatory variable or regressor) x [Spokoiny & Dickhaus 2015: 4]. In its simplest form it is

$$y_j = x_j w + \epsilon_j, \quad (j = 1, 2, \dots, m, \text{ for example}), \tag{3.6}$$

with ϵ_j being errors, accounting for deviations from a perfect correlation (e.g. through measurement uncertainties) between y_j and x_j. Here, the single weight w equals the gradient of the regression line. This simple regression [Freedman 2008: sect. 2.5] is a model for the correlation between the back-wall echo and the back-wall echo equivalent, see section 3.4.1.

In a more general form — multiple regression [Freedman 2008: sect. 2.5; Hastie et al. 2009: sect. 3.2.3] — one output value depends on several input values

$$y_j = \sum_{i=1}^{n} x_{ji} w_i + \epsilon_{ji}, \quad j = 1, 2, \dots, m. \tag{3.7}$$

The basic hypothesis of this work is that the back-wall echo height can be extracted out of the intermediate echo time series. For linear regression this will read for one back-wall echo BWE_j

$$BWE_j = \sum_{i=1}^{n} (x_{ji} w_{time_i} + \epsilon_{ji}), \tag{3.8}$$

in vector form

$$\mathbf{BWE} = \mathbf{X}_{ie}\mathbf{w}_{time} + \boldsymbol{\epsilon}. \tag{3.9}$$

BWE, $\mathbf{w}_{time}$ and $\boldsymbol{\epsilon}$ represent column vectors and $\mathbf{X}_{ie}$ is a matrix with each row representing one time series. Equation (3.9) and equation (3.8), respectively, represent the linear model [Spokoiny & Dickhaus 2015: 4; Priestley 1981: 310f. Hastie et al. 2009: 11f.], also called general linear model [Brockwell & Davis 1991: §2.6].

The first summand on the right hand side of equation (3.8) (or equation (3.9), respectively) equals the back-wall echo equivalent, $BWE_{equiv,j} = \sum_{i=1}^{n} x_{ji} w_{time,i}$, and the error $\epsilon_j = BWE_j - \sum_{i=1}^{n} x_{ji} w_{time_i}$ shall be minimised. Weights $\mathbf{w}_{time}$ are determined on the training data such that the squared errors $\sum_j \epsilon_j^2 = \|\boldsymbol{\epsilon}\|^2$ minimise (the method of least squares [Kreyszig 1970: 287f.; Spokoiny & Dickhaus 2015: sect. 1.2]),

$$\mathbf{w}_{time} = \underset{\mathbf{w}_{time}}{\operatorname{argmin}} \|\mathbf{BWE} - \mathbf{X}_{ie}\mathbf{w}_{time}\|^2 \tag{3.10}$$

$$= \underset{w_j}{\operatorname{argmin}} \sum_{j=1}^{m} \left(BWE_j - \sum_{i=1}^{N} x_{ij} w_j \right)^2, \tag{3.11}$$

where $\|\cdot\|$ is the Euclidean norm or distance, respectively.

Calculation of $\mathbf{w}_{time}$ out of $\mathbf{BWE}_{equiv} = \mathbf{X}_{ie}\mathbf{w}_{time}$ is performed with the Matlab backslash operator [Mathworks 1993: 28]

```
weights=Xie\BWEs,
```

which performs a linear regression analysis if $\mathbf{X}_{ie}$ is not a square matrix.

3.2 Fourier Analysis

The sequences

$$\frac{\cos nt}{\sqrt{\pi}}, \quad n = 0, 1, 2, \ldots \text{ and} \tag{3.12}$$

$$\frac{\sin nt}{\sqrt{\pi}}, \quad n = 1, 2, \ldots \tag{3.13}$$

or

$$e^{jnt} = \cos nt + j \sin nt, \; n = 0, \pm 1, \pm 2, \ldots; \text{ with } j = \sqrt{-1} \tag{3.14}$$

form an orthonormal set (cf. Mathematical Notation and Beals [1987: 125]): the inner product

$$\langle x, y \rangle = \int_0^{2\pi} x(t) y(t) dt \tag{3.15}$$

equals zero for $x = e^{jnt}$ and $y = e^{jmt}$ if $n \neq m$.

Any continuous differentiable periodic function $f(t)$ with periodicity $T = 2\pi$ can be represented by an infinite Fourier series

$$f(t) = \sum_{n=-\infty}^{\infty} c_n e^{jnt}. \tag{3.16}$$

(In fact, piecewise continuity with existence of left- and right-hand derivative suffices [Kreyszig et al. 2011: 480].) Any continuous differentiable function $f(t)$ with periodicity $T = l$ of arbitrary l can be represented by

$$f(t) = \sum_{n=-\infty}^{\infty} c_n e^{2\pi jnt/l}. \tag{3.17}$$

Each summand of equation (3.17) is one frequency component of $f(t)$. Each Fourier coefficient c_n [Stade 2005: 11] can be calculated using now the orthogonality of e^{jnt} in the inner product space:

$$c_n = \langle f(t), e^{2\pi jnt/l} \rangle = \frac{1}{T} \int_0^l f(t) e^{-2\pi jnt/l} dt, \; n \text{ fixed, period } T = l \text{ with arbitrary } l. \tag{3.18}$$

Data obtained for analysis is finite and most often discrete. A time series x_k of length N, data sampled at instances in time $t_k = t(k)$, $k = 0, 1, \ldots, N-1$ is infinitely continued (periodically extended) [cf. Stade 2005: sect. 1.11] and can then be represented through discrete Fourier transform (DFT) as a finite discrete Fourier series

$$f(t_k) = \sum_{n=0}^{N-1} F_n e^{2\pi jnk/N}. \tag{3.19}$$

The Fourier coefficients are now calculated with a sum instead of an integral:

$$F_n = \langle x_n, e^{2\pi jnk/l} \rangle = \frac{1}{N} \sum_{k=0}^{N-1} x_k e^{-2\pi jnk/N}. \tag{3.20}$$

$F_0 = \frac{1}{N} \sum_{k=0}^{N-1} x_k e^0$ is the mean of the time series. Furthermore, for $x_i \in \mathbb{R}$, $F_{N-n} = F_n^*$ for all $n = 1, 2, \ldots, N-1$. In that case, the amplitude spectrum [Alessio 2016: 77] or simply spectrum [Stade 2005: 62] (hence also the name spectral analysis [e.g. Alessio 2016] for Fourier analysis) is $F_0, 2F_n$, $n = 1, 2, \ldots, (N-1)/2$ (for odd N) against frequency [Stade 2005: 60]. Amplitudes of frequency components up to half the sampling frequency ($f_{sa} = 1/\Delta_t$, $\Delta_t = T/(N-1)$) can be represented. Whereas Fourier transform is a linear operation on a time series, creation of the amplitude spectrum is non-linear due to taking magnitudes.

Often the factor $1/N$ is omitted from calculation of equation (3.20) [Kreyszig et al. 2011: 529f.]; in this work equation (3.20) is basis for the presented amplitude spectra though.

The standard algorithm to perform a DFT is the fast Fourier transform (FFT), which allows an extremely faster computation [Stade 2005: sect. 7.6].

If a single-frequency time series is sampled in a way that the number of sample points N is an integer multiple of the signal's period divided by the time between two sampling points, i.e.

$$N\Delta_t/T = p,\ p \in \mathbb{N}, p > 0, \tag{3.21}$$

a discrete Fourier transform delivers exactly the one frequency component of the original signal, since the signal, sampled over a finite time but infinitely periodically extended, is correctly reproduced by sampling.

For a single frequency signal not exactly sampled over a multiple integer of its period, or real world signals containing often numerous different frequency components such that it is impossible to fulfill equation (3.21) for all components, leakage [León et al. 2011: sect. 5.5.6] occurs: the discrete frequency spectrum of e.g. a continuous but not optimally sampled single-frequency signal contains frequencies around the original one. Figure 3.1 shows the effect with a sampling frequency of 100 MHz, $\Delta_t = 10$ ns, and a sampling from 0 to 1 µs: a single frequency signal of 5 MHz (fig. 3.1 top left) already has some frequency components besides the main frequency (fig. 3.1 bottom left) merely because the sampling window is one sampling point too long and a small jump occurs from end to beginning for the signal continuted after 1 µs. A single frequency signal of 2.5 MHz with a step from minimum to maximum signal height from end to beginning of the sampled signal (fig. 3.1 top right) suffers most of leakage: the true main frequency is between two frequency components 2 MHz and 3 MHz due to the finite resolution of the DFT, and the amplitude spectrum is $\neq 0$ for all frequencies (fig. 3.1 bottom right).

Window functions applied on the time signals can reduce leakage. In this work, a Hann window (Hanning window, cosine bell window) [Stade 2005: 427; Alessio 2016: table 5.1]

$$0.5\left(1 - \cos\left(\frac{2\pi n}{N}\right)\right),\ n = 0, 1, \ldots, N \tag{3.22}$$

is applied, reducing all time signals but the centre one, and setting first and last value to 0, such that no step between these values for the periodic extension occurs. A Hann window (fig. 3.2 top left) is multiplied with e.g. a 2.5 MHz single frequency time series (fig. 3.2 top right) as in figure 3.1; the spectrum (fig. 3.2 bottom right) of the resulting time series (fig. 3.2 bottom left) has amplitude < 0.02 from 5 MHz on.

Several features out of the amplitude spectrum are evaluated as back-wall echo equivalent (BWE-equivalent) in this thesis, cf. fig. 3.3:

- The frequency of the frequency component with the largest amplitude, the peak frequency

$$F_{peak} = \max(|F_0|, \max_{n=1,\ldots,N-1} |2F_n|). \tag{3.23}$$

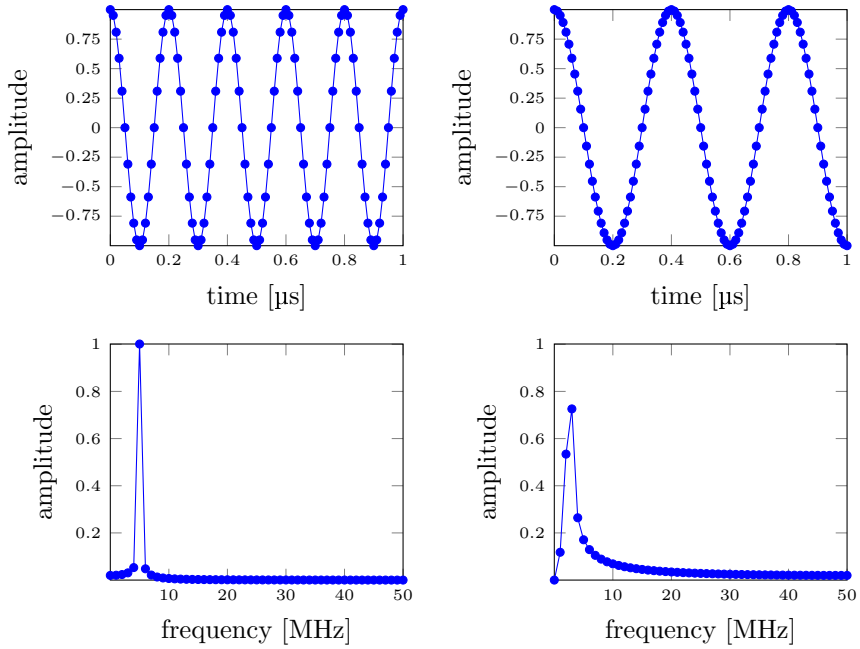


Figure 3.1: Single frequency time series of 5 MHz (top left) and 2.5 MHz (top right), sampled with 100 MHz; their discrete Fourier transform (bottom left and right) showing leakage

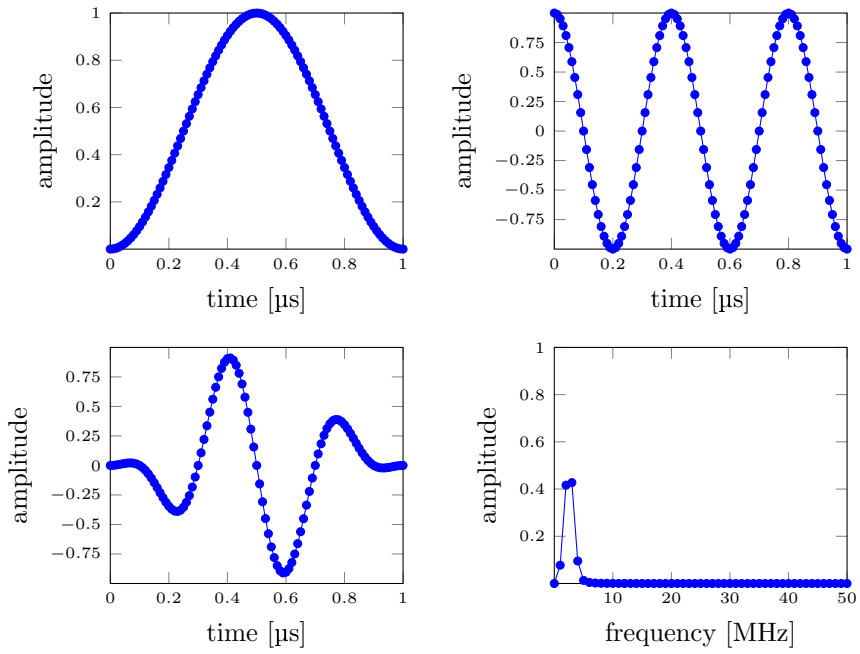


Figure 3.2: Hann window (top left), single frequency time series of 2.5 MHz (top right), Hann window applied on time series (bottom left) and the amplitude spectrum (bottom right) of this time series

- The left and right cut-off frequency F_{cl} and F_{cr}. This is the frequency at which the amplitude spectrum decreases by $-3\,\mathrm{dB}$ relative to the peak frequency left and right to it.

- Bandwidth, this is the difference between right and left cut-off frequency;

$$B_{3dB} = F_{cr} - F_{cl}. \tag{3.24}$$

- Mean frequency (centre frequency), the mean of the left and right cut-off frequency:

$$F_{centre} = \frac{F_{cr} - F_{cl}}{2}. \tag{3.25}$$

  The centre frequency of ultrasonic transducers (cf. sect. 2.2.2) is according to standards [DIN EN 2016; ASTM 2014] defined with cut-off frequencies at $-6\,\mathrm{dB}$; here for the feature out of the amplitude spectrum of intermediate echoes it is defined with $-3\,\mathrm{dB}$ following Blitz & Simpson [1996: 10] and Krautkrämer & Krautkrämer [1990: 212ff.].

- The integral below the amplitude spectrum between left and right cut-off frequency, here written as sum

$$A_{bw3} = \sum_{F_{cl}}^{F_{cr}} 2F_n. \tag{3.26}$$

- Linear regression as for the original time series (sect. 3.1), but here on all connected subsets of the amplitude spectrum $(0;\ 0, 0.33;\ 0, 0.33, 0.67;\ 0, 0.33, 0.67, 1;\ \ldots;\ 0.33;\ 0.33, 0.67;\ 0.33, 0.67, 1;\ \ldots;\ 49.67, 50;\ 50)$ MHz:

$$\mathbf{BWE} = \mathbf{X}_{FFT}\mathbf{w}_{FFT} + \boldsymbol{\epsilon}, \tag{3.27}$$

  cf. equation (3.9). $\mathbf{X}_{FFT}$ is a matrix with each row containing (the window of) an amplitude spectrum of one intermediate echo time series and $\mathbf{w}_{FFT}$ is a column vector with weights. As for linear regression on time series, $\mathbf{BWE}_{equiv} = \mathbf{X}_{FFT}\mathbf{w}_{FFT}$.

Any methods treating a localised frequency (time-frequency) analysis [Stade 2005: chap. 8; Hogan & Lakey 2005] (such as short time Fourier transform [STFT]) or wavelets [e.g. Louis, Maaß & Rieder 1997]) have not been treated in this thesis due to excellent results with Fourier analysis (chap. 5). Some results obtained with wavelets for the generation of a BWE equivalent can be found in Brandt et al. [2019], based on an approach proposed in Kim et al. [2013].

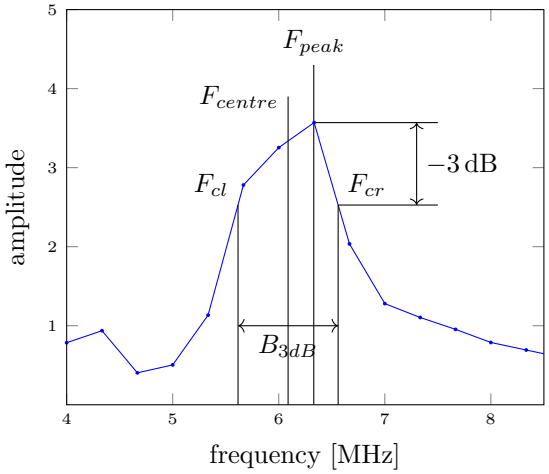


Figure 3.3: Features extracted from the amplitude spectrum of a Fourier transformed time series

3.3 Recurrence plots and recurrence quantification analysis

Recurrence plots (RPs) and recurrence quantification analysis (RQA) are special variants of non-linear time series analysis, which deal with relationships between states rather than with their exact values [cf. Webber et al. 2016: viii]. These methods are described in the next sections after briefly explaining non-linear time series analysis.

3.3.1 Non-linear (deterministic) time series analysis

Non-linear time series analysis may be divided into two areas: the first one stems from methods regarding a time series as being derived from a stochastic process, cf. section 3.1. Alterations of these methods (such as e.g. thresholding) introduce non-linearities into the models. Priestley [1981: sect. 11.5f.] or Diks [1999: chap. 3] give a short, Priestley [1988] a broader overview. The second area of non-linear time series analysis, concerned in this work, has its roots in dynamical systems and chaos theory [Broer & Takens 2011; Just & Schuster 2005; Ott 1993]. The latter marked a paradigm change: before the awareness of deterministic chaotic behaviour, scientists thought that random inputs (or random parameter changes) are the only possible reason for irregular behaviour of systems. Closely connected is the notion that small changes in inputs lead to small changes in outputs, which is valid for linear systems. With chaos research, scientists realised that even simple (low dimensional) non-linear, completely deterministic systems can exhibit irregular behaviour, and the evolution of time series generated by a chaotic system can completely vary after a short time (cf. sect. 3.3.2). Though only short time (quantitative) prediction is possible for these systems, still some information can be gained for dynamical systems with chaotic behaviour, such as Lyapunov exponents or correlation dimension, see Kantz & Schreiber [2005] for a general overview; Kurths & Herzel [1987] and Kurths et al. [1995] are examples for

applications. The targets of non-linear deterministic time series analysis are similar to the ones of its linear counterpart, e.g. prediction (though quantitatively only on a short time scale), noise reduction or understanding the mechanisms of time series generation.

Before explicating recurrence quantification analysis as one method of non-linear deterministic time series analysis, dynamical systems theory and the reconstruction in state space out of a time series as the bases will be explained.

3.3.2 Dynamical systems

Any (biological, economical, engineering...) system that evolves over time can be thought of as dynamical system.

> **Definition 1.** A **dynamical system** consists of a state (or phase) space $\mathcal{M}$ together with an evolution operator $\Phi : \mathcal{M} \times \mathcal{T} \to \mathcal{M}$ involving a time set $\mathcal{T} \subseteq \mathbb{R}$; Φ satisfies the properties:
>
> $$\Phi(\mathbf{x}, 0) = \Phi(\mathbf{x}), \tag{3.28}$$
>
> $$\Phi(\Phi(\mathbf{x}, t_1), t_2) = \Phi(\mathbf{x}, t_1 + t_2) \tag{3.29}$$
>
> [Broer & Takens 2011: chap. 1.2 definition 1.1; Pilyugin 2012: 1].

Remark 1: A continuous dynamical system as in definition 1 or its evolution operator Φ, respectively, is often referred to as a flow [Broer & Takens 2011: 15f.; Ott 1993: 6].

Remark 2: There is a different definition of a dynamical system in control theory, distinguishing between state, input and output variables [e.g. Hinrichsen & Pritchard 2010: 77].

A dynamical system for continuous time is typically not directly given by its evolution operator, but by a set of first-order autonomous (time-invariant) differential equations; this is in vector form

$$\dot{\mathbf{x}}(t) = \mathbf{v}(\mathbf{x}(t)),\ t \in \mathcal{T} \subseteq \mathbb{R},\ x \in \mathcal{M}. \tag{3.30}$$

In the discrete case, $\mathcal{T} = \mathbb{N}$, the system is given as an iterated map

$$\mathbf{x}_{k+1} = \mathbf{f}(\mathbf{x}_k),\ t \in \mathcal{T} \subseteq \mathbb{Z},\ x \in \mathcal{M}; \tag{3.31}$$

$\mathbf{f}$ can e.g. be constructed by — if possible — solving equation (3.30) and sampling the evolution, $\mathbf{x}(t = k\Delta_t) = \mathbf{f}(x(t = 0)), t = k\Delta_t$, with Δ_t as sampling time interval.

Remark 3: Often within dynamical systems theory a system according to equation (3.31) is meant when speaking of a map [Ott 1993: 7; Ott et al. 1994: 2; Kantz & Schreiber 2005: 31].

If the evolution operator Φ is given, definition 1 together with the state at one instant in time, say, $\mathbf{x}_0 = \mathbf{x}(t = 0) \in \mathcal{M}$ determines the evolution of the system for any point in the future; the system is completely deterministic.

The state space $\mathcal{M}$ has the same dimension as the describing equation (3.30) (or 3.31) or a smaller, possibly non-integer, dimension m of a set to which the dynamical

system is attracted after some transient time (a manifold, see sect. 3.3.3). One point in this space describes the state of a dynamical system for a given instant in time completely, e.g. angle and angular velocity for a pendulum [Broer & Takens 2011: 2ff.]. The time evolution of a dynamical system in state space is called trajectory or orbit [Ott 1993: 6; Arrowsmith & Place 1992: 12f.].

The mathematical theory of dynamical systems is concerned with qualitative properties of solutions of equation (3.30) [cf. Arrowsmith & Place 1992; Robinson 1999; Teschl 2012]. If the governing differential equations cannot be solved analytically, these investigations still provide insights into e.g. stability, asymptotic behaviour, fixed points (equilibria) or chaotic behaviour. The latter can occur for non-linear systems and is closely connected to sensitive dependence on initial conditions.

Definition 2. A map $\mathbf{f}$ on a state space $\mathcal{M}$ has **sensitive dependence on initial conditions** if there is an $r > 0$ (independent of the point), so that, given any point $\mathbf{x}_0 \in \mathcal{M}$ and any $\epsilon > 0$, there exists a point $\mathbf{y}_0 \in \mathcal{M}$ and a k such that $\|\mathbf{y}_0 - \mathbf{x}_0\| < \epsilon$ and

$$\|\mathbf{y}_k - \mathbf{x}_k\| \geq r \tag{3.32}$$

[Robinson 1999: 85].

In other words, however tiny the distance between two initial conditions x_0 and y_0 in state space is, a certain distance r between the future states x_k and y_k will always be exceeded if k is large enough.

Sensitive dependence on initial conditions is not a sufficient condition for chaos to appear, but chaotic systems always exhibit sensitive dependence on initial conditions. See Teschl [2012: sect. 11.3] for discussion and mathematical definition of chaotic systems. An overview about chaos theory and dynamical systems can be found in further textbooks such as Broer & Takens [2011], Robinson [1999], Guckenheimer & Holmes [2002] and Argyris et al. [2015].

3.3.3 State space reconstruction — time delay embedding

The only available information about a real world system is often, as in this work, one-dimensional data (time series). To be able to deploy tools of dynamical system theory, a representation in multi-dimensional state space (phase space) has to be generated.

The time series can be thought of as generated through a read-out or measurement function h from the m-dimensional dynamical system

$$h : \mathcal{M} \to \mathbb{R},\ x = h(\mathbf{x}_M). \tag{3.33}$$

(Here $\mathbf{x}_M$ denotes a state in the original state space.) An equivalent system can be reconstructed out of this time series ($\Phi(\Phi(\mathbf{x_M})) = \Phi^2(\mathbf{x_M})$):

Theorem 1. (**Reconstruction Theorem** [Takens 2010: 349]) The mapping

$\mathbf{f}_{rec} : \mathcal{M} \to \mathbb{R}^d$,

$\mathbf{x} = (h(\mathbf{x}_M), h(\Phi(\mathbf{x}_M)), h(\Phi^2(\mathbf{x}_M)), \ldots, h(\Phi^{d-1}(\mathbf{x}_M))),\ \Phi$ is differentiable,

$\mathbf{x} \in \mathbb{R}^d$,

$\mathbf{x_M} \in \mathcal{M},\ \mathcal{M}$ a compact manifold of dimension m,

$h : \mathcal{M} \to \mathbb{R},\ x = h(\mathbf{x}_M),\ h$ is differentiable,

is an embedding for generic pairs of Φ, h and $d > 2m$.

(This is also known as **Takens' Embedding Theorem** [Noakes 1991; Robinson 2010].)

A (differentiable) manifold is a topological, possibly non-linear space that can be locally mapped on a normed linear space so that differentiation is possible [Chillingworth 1976: sect. 3.1]. In theorem 1 it is the region in state space of a dynamical system to which the system is attracted; a simple example is the oval of a forced damped pendulum in its two-dimensional state space of (angle) position and velocity, to which this dynamical system settles asymptotically after the transient period [Broer & Takens 2011: 9f.]. For chaotic systems, the attractors can be complicated [Place & Arrowsmith 1991: 161] potentially with non-integer dimension [Kantz & Schreiber 2005: sect. 11.3]).

An embedding is a diffeomorphic mapping (a diffeomorphism), that is a differentiable mapping that has a differentiable inverse, see Chillingworth [1976: 150, 69] for details. Most important in this context here is that an embedding is one-to-one, i.e. a point in the reconstruction space in $\mathbb{R}^d$ is mapped to a unique point in the original state space (manifold) $\mathcal{M}$. This way the embedding keeps the complete qualitative structure of $\mathcal{M}$, and especially '...all the recurrence properties of the evolution [in original state space] are still present in the corresponding sequence of reconstruction vectors...' $\mathbf{x}_i$ [Broer & Takens 2011: 209].

The proof of theorem 1 was originally given in Takens [1981], also see Aeyels [1981] [cf. Broer & Takens 2011: sect. 6.3.2]; Sauer et al. [1991] extended it, and a comprehensive treatment is Huke [2006].

From theorem 1 follows for a time series x_i:

Definition 3. A **time delay embedding** of a discrete time series $x_i,\ i = 1, 2, \ldots, N$ is

$$f_{td} : \mathbb{R} \to \mathbb{R}^d,$$

$$\mathbf{x}_i = \left(x_i, x_{i+\tau}, x_{i+2\tau}, \ldots, x_{i+(d-1)\tau}\right),\ i = 1, \ldots, M;\ M = N - \tau(d-1) \quad (3.34)$$

with embedding dimension d, time delay τ, N length of the time series and M length of the embedded time series.

Equation (3.34) can also be presented as $\mathbf{x}_i = \sum_{k=0}^{d-1} x_{i+\tau k}\mathbf{e}_{k+1}$ with $i = 1, \ldots, M$, where the reconstruction state space of dimension d is spanned by unit vectors $\mathbf{e}_{k+1}, k = 0, \ldots, d-1$.

The important condition in theorem 1 is that Φ and h are generic, see Takens [2010: 245, 208] for details. Later work altered this condition [Huke 2006: theorem 2; cf. Ott et al. 1994: 46f., 52f.]. The conditions for the evolution operator Φ became more specific [Sauer et al. 1991: theorem 2.5]. One of these conditions is that τ has to be chosen in a way that there are no periodic orbits with period time $T = \tau$ or $T = 2\tau$ [Ott et al. 1994: 51ff.]. However, in this work no such restrictions are taken into account (it is difficult to achieve with varying T in the time series).

It is difficult to verify whether the measured time series sufficiently represents the underlying dynamics. According to Sauer [2006], 'A mathematically generic observation, by definition, monitors all degrees of freedom of the system. The extent to which this is true affects the faithfulness of the reconstruction. If there is only a weak connection from some degrees of freedom to the observation function, the data requirements for a satisfactory reconstruction may be prohibitive in practice.' [Cf. also Kantz & Schreiber 2005: 146f.] In general, researchers do not explicitly check whether the assumption of the measurement function h to be generic is fulfilled. Moreover, the embedding theorem is valid for attractors, i.e. the settled (stationary) state of a dynamical system [cf. Yap & Rozell 2010]; non-linear time series analysis in general and specifically RQA are however explicitly used for nonstationary signals [e.g. Kantz & Schreiber 2005: chap. 13; Digulescu et al. 2016].

Overall, the conditions of theorem 1 are hard to verify in general [Robinson 2010]. However, delay embedding reconstruction has been proven in hundreds of cases to effectively obtain information about the underlying dynamical system for characterisation, prediction, classification etc. In this sense it is also used in this thesis for recurrence quantification analysis.

In practice, embedding dimension d and time delay τ have to be chosen, for which different approaches exist [Kantz & Schreiber 2005: sect. 3.3; also Marwan & Webber 2015]. In this thesis, the analysis — of recurrence plots for detection of porosity in CFRP through ultrasonic testing — starts 'right away' and d and τ leading to best results are chosen [Kantz & Schreiber 2005: 36; cf. Zbilut & Webber 1992].

3.3.4 Recurrence plots

In common English, if something recurs or is recurrent, it 'happens more than once' [Collins Cobuild 1999]. Mathematically, if a dynamical system recurs, it returns infinitely close to a point in state space it visited before [Furstenberg 1981: 8]. Poincaré [1890: 69, theorem 1] set up the first recurrence theorem, now known as Poincaré recurrence theorem. It shall be given here by use of non-wandering points.

> **Definition 4.** A point $\mathbf{x_0}$ is a **non-wandering point** for a map $\mathbf{f}(\mathbf{x}_k)$ on $\mathcal{M}$ if, given any neighbourhood U of $\mathbf{x_0}$, there are arbitrarily large $n > 0$ for which $U \cap \mathbf{f}^n(U) \neq \emptyset \qquad (\mathbf{f}^n(U) = \{\mathbf{f}^n(\mathbf{x});\ \ \mathbf{x} \in U\})$.
>
> [Guckenheimer & Holmes 2002: 236; Chillingworth 1976: 198]

The set of all non-wandering points shall be denoted Ω.

Based on this definition, there is the

> **Theorem 2. Poincaré Recurrence Theorem**: For a (time-invariant) dynamical system $(\mathcal{X}, \Phi, \mathcal{T})$ that preserves volume on an open set $\mathcal{X} \in \mathcal{R}^n$, the set of all non-wandering points of $\mathcal{X}$ equals $\mathcal{X}$: $\Omega(\mathcal{X}) = \mathcal{X}$.
> [Hinrichsen & Pritchard 2010: 208 (including proof)]

The property of preserving volume is nowadays generally given by saying that the measure of $\mathcal{X}$ is constant (and finite) [cf. Hinrichsen & Pritchard 2010; Furstenberg 1981: 8]. The constant (invariant) measure is a 'natural generalization' of the integral invariant — the latter can be seen, in the case of an incompressible fluid as dynamical system, as the fact that the volume of the fluid does not change over time [cf. Nemytskii & Stepanov 1972: 447, 425f.; also Carathéodory 1919: 580f.].

Different expressions of Poincaré's recurrence theorem can be found in Furstenberg [1981: 8] and Barreira [2003].

Eckmann et al. [1987] proposed a tool which they named recurrence plot (RP), utilising recurrences for the evaluation of dynamical systems. It is introduced here via distance matrix and recurrence matrix:

> **Definition 5.** A **distance matrix D** contains the distance between each and every point of a (discretely given) trajectory in state space $\mathcal{M}$ (often obtained via reconstruction out of a time series). The ith row contains, as well as the ith column, the distance between the ith point in state space and every other point

$$D_{i,j} = \|\mathbf{x}_i - \mathbf{x}_j\|, \qquad i, j = 1, \ldots, M;\ M\text{: length of embedded time series} \tag{3.35}$$

> The distance is often the Euclidean metric (used in this thesis) or the maximum norm [Marwan & Webber 2015; Marwan et al. 2007].

In this work, angular distance is used in addition, as recently proposed as alternative distance measure for recurrence plots [Ioana et al. 2014; Birleanu 2012; Carrión et al. 2015b; cf. Carrión et al. 2015a]:

$$\|\mathbf{x}_i - \mathbf{x}_j\|_{ang} = \arccos\left(\frac{\mathbf{x}_i \cdot \mathbf{x}_j}{\|\mathbf{x}_i\|\|\mathbf{x}_j\|}\right). \tag{3.36}$$

A distance plot (or unthresholded recurrence plot) is obtained when the distance matrix is plotted in a two-dimensional figure, colours or grey shades presenting distances [Marwan et al. 2007: 253].

> **Definition 6.** A **recurrence matrix R** is generated out of a distance matrix by thresholding

$$\mathbf{R}(\epsilon) = \theta(\epsilon\mathbf{1} - \mathbf{D}), \quad \mathbf{1} \in \mathbb{R}^{M \times M}\text{ : matrix of 1s} \tag{3.37}$$

$$R_{i,j}(\epsilon) = \theta(\epsilon - D_{i,j}) = \theta\left(\epsilon - \|\mathbf{x}_i - \mathbf{x}_j\|\right), \quad i,j = 1,\ldots,M;\ M\text{: length of embedded time series.} \tag{3.38}$$

The Heaviside function θ maps positive values to 1 (recurrent points) and negative values to 0 (non-recurrent points).

The recurrence threshold ϵ is either constant or it varies in a way that for each state (each point in time) there is a fixed amount of recurrence points.

Today, a constant ϵ is most frequently used [Marwan & Webber 2015]. A fixed amount of recurrence points (fixed amount of nearest neighbours FAN [Marwan & Webber 2015]), as originally proposed in Eckmann et al. [1987] can also be achieved by fixing the local recurrence rate (in the sense of recurrence rate of a row or a column, respectively, of the recurrence matrix).

No systematic approach to the choice of ϵ exists, see Marwan et al. [2007] and Marwan & Webber [2015] for discussion and an overview of the different approaches in literature. In this thesis, ϵ values are varied over almost the whole possible range, and the values leading to optimum results for detection of porosity are used. The determination of ϵ out of the distribution of distances in the distance matrix, achieving thus a fixed global recurrence rate, has been recently proposed by Krämer et al. [2018].

With ϵ = constant, $\mathbf{R}$ is symmetric; with a varying ϵ to achieve the same number of recurrent points for each point it is not symmetric. Independent of determination of ϵ, the main (ascending) diagonal (the line of identity LOI [Marwan & Webber 2015]) naturally — since $\|\mathbf{x}_i - \mathbf{x}_i\| = 0$ — consists of recurrence points only for RP and joint recurrence plots.

The two-dimensional plot of a recurrence matrix (black for recurrent and white for non-recurrent points, or vice versa) is a recurrence plot (RP) [Marwan & Webber 2015].

In the paper introducing the recurrence plot, Eckmann et al. [1987] presented examples for this tool being used on two (simulated) chaotic systems (one with an added term varying linearly with time) and on one experimental series from Eckmann et al. [1986]. A transient, a drift and periodic behaviour was shown.

The following three extensions of recurrence plots enable the comparison of two (or more) trajectories in state space.

Cross recurrence plots

Cross recurrence plots (CRPs) [Marwan et al. 2007: sect. 3.3; Marwan & Webber 2015] and the related matrices compare one trajectory to another:

Definition 7. A **cross distance matrix CD** contains the distance for all possible combinations between points of a (discretely given) trajectory of one dynamical system $\mathbf{x_i}$ with the ones of another dynamical system $\mathbf{y_j}$

$$CD_{i,j} = \|\mathbf{x}_i - \mathbf{y}_j\|,\ i,j = 1,\ldots,M;\ M\text{: length of embedded time series.} \tag{3.39}$$

Cf. definition 5; the cross recurrence matrix and thus the CRP is derived identically to definition 6.

A cross recurrence plot thus contains a recurrent point if the states of the two different systems are close to each other. Two identical but time shifted systems for example lead to a CRP with a recurrent line not on the main diagonal, but shifted away from it. Diagonal lines in general represent, in a similar way as in RPs, trajectories that run parallel for some time. A CRP of two identical systems, but one of them rotated in state space relative to the other, looks completely different than the original RPs [Marwan & Webber 2015]. A CRP of two identical systems, of which one is stretched in time compared to the other one, contains bowed diagonal lines [Marwan et al. 2007: 263].

It can be concluded that CRPs are more appropriate to compare similar systems [Marwan & Webber 2015; Marwan et al. 2007: 263], as for example in this thesis ultrasonic signals from porous and non-porous areas of the same material.

Joint recurrence plots

Joint recurrence plots (JRPs) [Marwan et al. 2007: sect. 3.4; Marwan & Webber 2015] compare two (or more, not considered here) RPs:

Definition 8. A **joint recurrence matrix JR** is generated out of two recurrence matrices

$$JR_{i,j}^{\mathbf{x},\mathbf{y}} = R_{i,j}^{\mathbf{x}} R_{i,j}^{\mathbf{y}} = \theta\left(\epsilon - \|\mathbf{x}_i - \mathbf{x}_j\|\right)\theta\left(\epsilon - \|\mathbf{y}_i - \mathbf{y}_j\|\right), \quad i,j = 1,\ldots,M;\ M\text{: length of embedded time series}, \tag{3.40}$$

deriving from two trajectories $\mathbf{x}$, $\mathbf{y}$ of two different dynamical systems.

A point in a JRP is recurrent only if the corresponding points in both RPs that form the basis of the JRP are recurrent. Thus the main diagonal consists of recurrent points only (in contrast to the CRP). A JRP of two identical systems, but one of them rotated in state space relative to the other, is identical to the RPs of each of the system (which are again identical). A JRP can thus detect that two systems are identical up to a rotation [Marwan & Webber 2015]. On the other hand, a JRP of two identical systems, of which one is stretched in time, can be almost empty.

It can be concluded that JRPs are more appropriate to compare two interacting systems [Marwan & Webber 2015; Marwan et al. 2007: 263].

A CRP as well as a JRP of two identical systems is identical to a standard recurrence plot of this system. This is different for the following recurrence plot.

Difference recurrence plots

The author of this thesis developed the idea of a difference recurrence plot (DRP):

Definition 9. A **difference recurrence matrix DR** subtracts one distance matrix from another and creates a recurrence matrix out of this difference distance matrix.

$$DR_{i,j}^{\mathbf{x},\mathbf{y}} = \theta\left(|D_{i,j}^{\mathbf{x}} - D_{i,j}^{\mathbf{y}}|\right) = \theta\left(\epsilon - |\,\|\mathbf{x}_i - \mathbf{x}_j\| - \|\mathbf{y}_i - \mathbf{y}_j\|\,|\right), \\ i,j = 1,\ldots,M;\ M\text{: length of embedded time series.} \quad (3.41)$$

A difference recurrence plot of two identical systems contains merely recurrent points, since the difference distance matrix contains zeros only. A DRP may be used to compare similar systems A and B and detect local variations in system B compared to A: the distances between states in this variation area and other states for system B are different from these distances of system A. Such distance differences are represented by non-recurrent points. DRPs will be heuristically tested for the first time for one application in this work (sect. 5.6.5).

3.3.5 Recurrence quantification analysis

Only three publications before 1990 are known with direct reference to the paper introducing RPs [Scheinkman & LeBaron 1989; Babloyantz 1989; Mayer-Kress & Hübler 1989]. These publications treat recurrence plots as qualitative, visually inspected tools as Eckmann et al. [1987] introduced them. With beginning of the 1990s, Zbilut [1991] [also Zbilut et al. 1990] and Webber [1991] picked up the topic; the latter introduced the first quantification approach for recurrence plots with percent recurrence (definition 10) (and also an entropy H, but at that time interestingly defined as the entropy of the distribution of distances, i.e. based on the distance matrix, and not the recurrence matrix, cf. definition 13).

Zbilut & Webber [1992] mutually continued to introduce determinism (definition 11) (at that time 'percent line segments') as the next quantification measure and, then, put five qualitative features derived from recurrence plots — percent recurrence, determinism (definition 11, coining the nowadays used name), entropy (this time as definition 13), ratio (definition 12) and trend (definition 17) — into one framework named recurrence plot analysis (RPA) [Webber & Zbilut 1994].

The name recurrence quantification analysis (RQA) for the quantitative analysis of recurrence plots appeared first in Trulla et al. [1996]. It is also the name of a software package [Webber et al. 2009]; the name of the software was changed from recurrence plot analysis (RPA) to RQA (a name Charles L. Webber came up with [personal email communication Aug. 2017]) after detecting a severe error [Webber 2014].

Recurrence quantification analysis (RQA) is a flourishing niche of dynamical system analysis and signal evaluation, starting in the 1990s inspired by the quantification efforts of Webber and Zbilut after the first introduction of recurrence plots by Eckmann et al. [1987] — with an almost constantly growing number of publications per year up to today [cf. Website RP Bibliography 2020].

RQA features

In the following, eight RQA features used in this thesis — two of which (definition 15, definition 18) are newly proposed adaptations of known features — are described. The definitions are based on Marwan et al. [2007] and Marwan & Webber [cf. 2015], but the equations are extended by distinguishing whether or not the main diagonal and potentially a Theiler window shall be taken into account. (The Theiler window shall exclude the tangential motion, i.e. directly in time following or preceding points of the trajectory in state space, [cf. Marwan et al. 2007; Theiler 1986].) To do so, an auxiliary parameter excluded window w_e is introduced

$$w_e = \begin{cases} 1 & \text{LOI excluded from calculation of RQA features} \\ 1 + w_{th} & \text{LOI and Theiler window of length } w_{th} \text{ excluded} \\ 0 & \text{else.} \end{cases} \tag{3.42}$$

Before starting with the first feature, the histogram of diagonal lines parallel to LOI that consist of exactly l recurrence points shall be defined

$$H_D(l) = \sum_{\substack{i,j=1 \\ |i-j|>w_e-1}}^{M-w_e} \left((1 - R_{i-1,j-1})(1 - R_{i+l,j+l}) \prod_{k=0}^{l-1} (R_{i+k,j+k}) \right), \tag{3.43}$$

M: length of embedded time series,

which is necessary for computation of all features based on diagonal lines. Such diagonal lines of length l present periodicity in the dynamical system, since they appear if trajectories in state space are approximately parallel to each other (more exactly, one trajectory is within an ϵ-tube around the other) [cf. Marwan et al. 2007: 244].

Definition 10. The **recurrence rate** RR is the proportion of recurrent points in a recurrence matrix (or the corresponding plot) to the sum of all points in this matrix (here calculated via $H_D(l)$)

$$RR = \frac{\sum_{l=l_{min}}^{M-w_e} l H_D(l)}{M^2 - \theta(w_e - 0.5)\left(M + 2 \sum_{w=1}^{w_e-1} (M - w) \right)}. \tag{3.44}$$

The denominator considers the LOI and Theiler window taken out of computation if applicable; the Heaviside function θ sets the second term to zero if $w_e = 0$.

Definition 11. Determinism DET is the proportion of recurrent points in diagonal lines — parallel to the line of identity, with a length $l \geq l_{min}$ — to all

recurrent points in the recurrence plot

$$DET = \frac{\sum_{l=l_{min}}^{M-w_e} lH_D(l)}{\sum_{l=1}^{M-w_e} lH_D(l)}. \tag{3.45}$$

The next feature $RATIO$ is considered useful to detect dynamic transitions [Marwan & Webber 2015; cf. Webber & Zbilut 1994].

Definition 12. ***RATIO*** is the ratio of DET to recurrence rate (RR)

$$RATIO = \frac{DET}{RR}. \tag{3.46}$$

For the next feature the (probability) distribution of diagonal lines is used

$$p(l) = \frac{H_D(l)\Big|_{l \geq l_{min}}}{\sum_{l=l_{min}}^{M-w_e} H_D(l)}, \tag{3.47}$$

which corresponds for a chosen $l \geq l_{min}$ to the amount of diagonal lines of length l divided by the amount of all diagonal lines with length $l \geq l_{min}$. It thus depends (as well as DET and $RATIO$) on the defined l_{min}, and so do the next two features. The first, ENT, is based on the entropy in information technology [Mitchell 1997: 55ff.; Webber & Zbilut 1994] and gives information about the homogeneity of the distribution of diagonal lines:

Definition 13. Entropy is defined as

$$ENT = -\sum_{l=l_{min}}^{M-w_e} p(l) log_2 p(l). \tag{3.48}$$

Furthermore, the average line length can be derived from the distribution of diagonal lines, being a measure for the mean time two trajectory segments in state space are close to each other:

Definition 14. L is the **average line length** of all (ascending) diagonal lines (parallel to the line of identity) $\geq l_{min}$ in a recurrence plot

$$L = \frac{\sum_{l=l_{min}}^{M-w_e} lH_D(l)}{\sum_{l=l_{min}}^{M-w_e} H_D(l)}. \tag{3.49}$$

The author of this thesis developed the idea of a normalised version in an attempt to use as much as possible RQA features ranging between 0 and 1, achieved via the reciprocal of the average line length:

Definition 15. The **normalised average line length** L_{nor} is the minimum line length divided by the average length of all (ascending) diagonal lines (parallel to the line of identity) $\geq l_{min}$ in a recurrence plot.

$$L_{nor} = \begin{cases} \dfrac{l_{min}}{L} = \dfrac{l_{min} \sum\limits_{l=l_{min}}^{M-w_e} H_D(l)}{\sum\limits_{l=l_{min}}^{M-w_e} l H_D(l)} & \text{for } L > 0 \\ 0 & \text{for } L = 0. \end{cases} \tag{3.50}$$

L_{nor} is defined here to be 0 if $L = 0$, i.e. if the RP contains only diagonal lines of length $< l_{min}$. It takes the value 1 if the recurrence matrix contains only lines of length l_{min}.

DET, $RATIO$, ENT and L_{nor} depend on the minimum line length l_{min}, which is thus an RQA parameter for calculation of these features in addition to embedding dimension d, time delay τ and recurrence threshold ϵ. It is varied in this work over a certain range and the values leading to best results of porosity detection in CFRP are used, as for the three former parameters.

The last three features used in this work do, as recurrence rate RR, not depend on l_{min}. The first of these features measures how fast trajectory segments in state space diverge:

Definition 16. Divergence DIV is the inverse of the length of the greatest diagonal line appearing in the recurrence plot (excluding potentially the line of identity and the Theiler window)

$$DIV = \frac{1}{l_{max}}. \tag{3.51}$$

For the next feature, trend, the RR_k in the recurrence matrix is taken into account [Marwan et al. 2007: 265]

$$RR_k = \frac{1}{M-k} \sum_{j=1}^{N-k} R_{j+k,j}. \tag{3.52*}$$

The index k denotes the distance to the main diagonal (LOI in RPs and JRPs) so that RR_k is the recurrence rate in the diagonal in distance k from the main diagonal (and RR_0 is the recurrence rate of the main diagonal).

Equation (3.52*) is RR_k for a symmetric RP (i.e. RP or JRP with fixed recurrence threshold or fixed global recurrence rate). For non-symmetric RPs, the author of this

thesis proposes to define RR_k as the mean of the recurrence rate of the two diagonals with distance k from the main diagonal

$$RR_k = \frac{1}{2(M-k)} \left(\sum_{j=1}^{N-k} R_{j+k,j} + \sum_{i=1}^{N-k} R_{i,i+k} \right). \tag{3.52}$$

Equation (3.52) is universally valid, also for symmetric RPs.

Definition 17. Trend *TND* calculates the slope of the (least squares) best fit line [Webber & Zbilut 1994] between recurrence rates RR_k of the diagonals against their distance to the main diagonal

$$TND = \frac{\sum_{k=w_e}^{\widetilde{M}} \left(k - \frac{\widetilde{M} - w_e}{2} \right) (RR_k - \overline{RR_k})}{\sum_{k=w_e}^{\widetilde{M}-w_e} \left(k - \frac{\widetilde{M} - w_e}{2} \right)}, \quad \widetilde{M} < M \tag{3.53}$$

[Marwan 2003: 25; Marwan et al. 2007: 268].

Diagonal lines ($\Delta_M = M - \widetilde{M}$ diagonal lines) at the edges of the RP furthest away from the main diagonal are excluded from the calculation of *TND* because of the insufficient number of recurrence points. See section 5.4.1 for the choice of this parameter.

The author of this thesis proposes an adaptation of *TND*:

Definition 18. Correlation *COR* is the correlation coefficient between local recurrence rate RR_k and distance to LOI

$$COR = \frac{\sum_{i=w_e}^{\widetilde{M}} \left(i - \frac{\widetilde{M} - w_e}{2} \right) \left(RR_k - \overline{RR_k} \right)}{\sqrt{\sum_{i=w_e}^{\widetilde{M}} \left(i - \frac{\widetilde{M} - w_e}{2} \right)^2 \sum_{i=w_e}^{\widetilde{M}} \left(RR_k - \overline{RR_k} \right)^2}}. \tag{3.54}$$

This feature takes values between -1 and 1; it does not give any information about the slope of the best fit line, but about how near the points of distribution are to the best fit line (it may theoretically be 1 with a very small slope and all points lying on the best fit line). A stationary process presumably does virtually not have any correlation, and a nonstationary process does; this feature is heuristically tested for the first time in this work.

3.4 Correlation and classification

The last three sections described the methods that will be used in chapter 5. Each method provides a specific function $f_{bwee}(x_i) = BWE_{equiv}$ to create a BWE-equivalent out of intermediate echo time series of ultrasonically tested CFRP parts for evaluation on porosity, called a model in machine learning [Flach 2012: 13].

The aim is to have the back-wall echo equivalent resemble the back-wall echo as well as possible; the function

$$f_{equiv}(BWE_{equiv}) = BWE \tag{3.55}$$

shall be at best linear; f_{equiv} is expected to be individual for each material and ply thickness and for each method to generate the BWE-equivalent.

3.4.1 Correlation and regression

Correlation is the first approach in this thesis to determine how well the BWE-equivalent resembles the BWE. The correlation between two variables can be expressed by Pearson's r (Pearson's correlation coefficient or just correlation coefficient) [Press et al. 2007: 745; Kreyszig 1970: 329]

$$r = \frac{\sum_{i=1}^{n}(x_i - \overline{x})(y_i - \overline{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \overline{x})^2 \sum_{i=1}^{n}(y_i - \overline{y})^2}}, \tag{3.56}$$

which informs how much data points x_i and y_i, $i = 1, \ldots, n$ deviate in the same way from their respective means $\bar{x}$, $\overline{y}$ in units of their standard deviations s ($s_x = \sqrt{1/(n-1)\sum_i (x_i - \overline{x})^2}$). r can take values between -1 and +1; -1 and +1 correspond to perfect correlation; 0 corresponds to no correlation.

It can be shown that $r^2 = R^2$, where $1 - R^2$ is proportional to the variation between predicted and actual values (the residual sum of squares $RSS = \sum_{i=1}^{n}(y_i - \hat{y}_i)^2$ [James et al. 2013: 69ff.]). r^2 (and r itself, respectively) is thus a measure of the goodness of fit.

If the dependence of (in the simplest case) one variable (dependent variable y) on another variable (independent variable x) is emphasized, one often speaks of linear regression (cf. sect. 3.1) [Kreyszig 1970: 285]. The term was coined by [Galton 1886b,a, 1888] through investigating that children of tall parents tend to be shorter and children of short parents tend to be taller than their parents. A regression curve can be drawn through the distribution of variable pairs (x_i, y_i), in the simplest case a regression line (equation (3.6)).

The concept of confidence intervals (sect. 3.4.2) can also be applied to a regression line. In this work, the prediction interval is used, which gives the interval in which to a (for example) 95% probability a future single value y_0 lies. The prediction interval [Mann 2011: sect. 3.8.2] of a certain independent variable x_0 is given, with the value on the regression line $\hat{y}_0$, as

$$y_0 = \hat{y}_0 \pm t_{0.95,\text{2-sided}}(n-2)\sqrt{\frac{1}{n-2}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2\left(1 + \frac{1}{n} + \frac{(x_0 - \overline{x})^2}{\sum_{i=1}^{n}(x_i - \overline{x})}\right)}. \tag{3.57}$$

n is the number of data points; $t_{0.95,2\text{-sided}}$ is taken from the t-distribution table, confidence interval 95%, for $n-2$ degrees of freedom.

3.4.2 Classification

Classification is the process of mapping objects — instances in machine learning terminology [Flach 2012: 49]; here: intermediate echo time series — to two or more pre-defined classes [Duda et al. 2001: 12f.; Hand 1981: 2]. Binary classification is classification into two classes (as in this thesis) [James et al. 2013: 355; Flach 2012: 52]. Discrimination may be defined as the process of finding a mapping rule from already classified objects (classified here into porous and non-porous by back-wall echo height) [Hand 1981: 2]. These tasks are different from clustering or cluster analysis, where no classes are known a priori; in language of machine learning, the instances used for finding the mapping rule (for training) are unlabelled. One speaks of unsupervised learning, in contrast to supervised learning [Hastie et al. 2009: 485ff.; James et al. 2013: sect. 2.1.4; Duda et al. 2001: 16f.]; the latter is performed in this work with intermediate echo time series labelled by the back-wall echo out of the same A-scan. The result of a classification algorithm (or classifier) to sort an instance into one of two classes is called prediction [Flach 2012: sect. 5.2; James et al. 2013: 17ff.]; a somewhat different meaning than prediction of future values of a time series or the prediction interval of a regression (sect. 3.4.1). A feature [Duda et al. 2001: 1ff. Flach 2012: 13] is for example the maximum value of the intermediate echo time series or the RQA feature recurrence rate (generated out of the intermediate echo time series), used as back-wall echo equivalent. The mapping generated by the back-wall echo equivalent to the classes non-porous and porous is a model [Flach 2012: 13]. The work in this thesis is confined to using only one back-wall echo equivalent at once for classification, to show principal relations and physical connections. The combination of different features in a feature vector lying in a feature space [Duda et al. 2001: 4] to improve the classification result is not part of this thesis.

After training of the classification algorithm on the training set, the algorithm is tested (evaluated) on a different, independent test set [James et al. 2013: 30]. (The term evaluation is in this work however meant to be the whole work of analysis and finding a back-wall echo equivalent, including the 'training'.) Quite common in machine learning and also used in this work is the approach of cross-validation: for an x-fold cross-validation, a data set is (randomly) separated into x subsets; all subsets but one are used for training and the remaining data set is used for test. This is performed for all possible x combinations, that is, the test set changes from 1 to x [fig. 3.4, Richert & Coelho 2013: 39; Witten et al. 2011: sect. 5.3; Flach 2012: sect. 12.2].

A thorough differentiation between training and test is of utmost importance in machine learning. Overfitting can lead to much better training results than the algorithm can actually achieve on 'real' (test) data (independent of the training set), which the algorithm has not seen before, identified by test results. Noise leads to overfitting; the algorithm overfits to the noise in the data and thus performs considerably worse on test data [Abu-Mostafa et al. 2012: chap. 4; Mitchell 1997: 67].

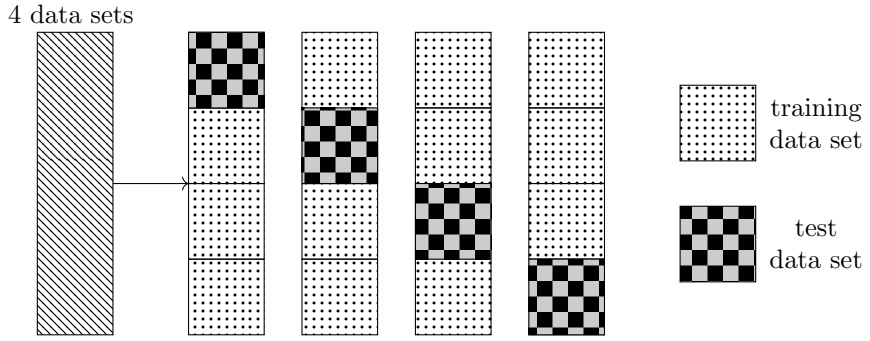


Figure 3.4: Principle of (fourfold) cross-validation

The performance of a classifying algorithm is assessed with different finding rates, which can be visualised with the contingency table (confusion matrix) (table 3.1 and 3.2) [Flach 2012: chap. 2.1; cf. also James et al. 2013: 145f.].

Table 3.1 gives dummy figures of detection etc. for the application in this thesis; table 3.2 shows designations for the performance indicators. The classes are labelled as postitive and negative; positive corresponds to a defect. Taking the result of the algorithm, the prediction, into account, four different areas are spanned. The most familiar in NDT inspection are the true positives (the detected defects) and the false positives. The latter is to be understood as falsely as positive predicted instances (which are indeed negative) — false alarms.

There are different ways to assess the performance of a classifying algorithm. The accuracy (or success rate) is the number of correctly predicted instances, true positives and true negatives, as proportion of overall instances. The error rate is the proportion of incorrectly predicted instances, being one minus accuracy.

Nearer to the field of NDT and adapted in this work is the proportion of true positives to the overall number of positives, the true positive rate (*TPR*) (detection rate, [positive] recall), and the proportion of false positives to all negatives, the false positive rate (*FPR*) (false alarm rate). Both will be in the remainder of this work given in %. True positive rate is in the area of NDT frequently designated probability of detection (PoD; often in connection with a confidence level, see below). True positive rate and false positive rate depend on a decision threshold *thres* if the original outcome of the algorithm is continuous. The value of the decision threshold determines the proportion of predicted positives and negatives and thus the performance. Plotting *TPR* against *FPR* for different values of the decision threshold gives the receiver operating characteristic (ROC) curve [Duda et al. 2001: 49f.] (fig. 3.5). Integrating the ROC curve gives the area under curve (*AUC*) [Witten et al. 2011: 177]

$$AUC = \int_0^1 TPR \, dFPR, \tag{3.58}$$

which has the advantage to be independent of the decision threshold.

Table 3.1: Contingency table (confusion matrix) for binary classification into porosity/non-porosity with artificial example values

		true classes (determined via ultrasonic back-wall echo height)		
		porosity	non-porosity	
prediction of the classifying algorithm (back-wall echo equivalent generated out of ultrasonic intermediate echoes)	porosity	**30**	**6**	number of instances indicated as porosity by algorithm: 36
	non-porosity	**4**	**60**	number of instances indicated as non-porosity by algorithm: 64
		number of true porosity instances: 34	number of true non-porosity instances: 66	correctly identified instances: **90**

For this thesis it is considered advantageous to have a single figure to assess the classification performance. *AUC* is rather abstract, not directly related to finding rates. Accuracy (success rate) does not reflect unbalanced data sets with for example many negative (defect-free) instances and only a few positive instances (defects). To overcome this, balanced accuracy [Brodersen et al. 2010; Sokolova & Lapalme 2009] (also average recall [Flach 2012: 60]) is utilised

$$acc_{bal} = \frac{TPR + TNR}{2}. \tag{3.59}$$

Balanced accuracy is chosen as indicator for the best classification performance. This is done with the additional constraint

$$TPR \approx TNR \tag{3.60}$$

to avoid 'unbalanced balanced accuracy' with for example $TPR = 99\%$ and $TNR = 81\%$. ($TPR = TNR$ may not be achievable due to numerical constraints.)

Standard requirement for NDT in aeronautic industry is a probability of detection (TPR) of 90% (with a confidence level of 95%, see below) [e.g. ISO 2005: 11, 16; Dep. of Defense 1998: 405; Fahr 2014: 14f.]. The goal set for this work is $acc_{bal} = 90\%$ with $TPR \approx TNR$. The corresponding false alarm rate $FPR = 10\%$ may possibly be too high from economic point of view [cf. Schnars & Kück 2009a], but accepted within

Table 3.2: Contingency table (confusion matrix) for binary classification with performance indicators [Flach 2012: sect. 2.1]; number of overall instances: M_{inst}

		true classes		error rate $^{(FP+FN)}/_{M_{inst}}$
		porosity	non-porosity	
prediction of the classifying algorithm	porosity	**TP** (in NDT often: hits; detected defects)	**FP** **(false alarms)**	number of instances indicated as porosity by algorithm: TP + FP
	non-porosity	**FN** (in NDT often: misses, missed defects)	**TN** (correctly identified non-defect areas)	number of instances indicated as non-porosity by algorithm: FN + TN
		true positive rate (detection rate, sensitivity, recall) $TPR = {}^{TP}/_{(TP+FN)}$	true negative rate (specificity, negative recall) $TNR = {}^{TN}/_{(FP+TN)}$; false positive rate (false alarm rate) $FPR = {}^{FP}/_{(FP+TN)}$; $FPR = 1 - TNR$	accuracy (success rate) $acc = {}^{(TP+TN)}/_{M_{inst}}$; balanced accuracy (average recall) $acc_{bal} = {}^{(TPR+TNR)}/_{2}$

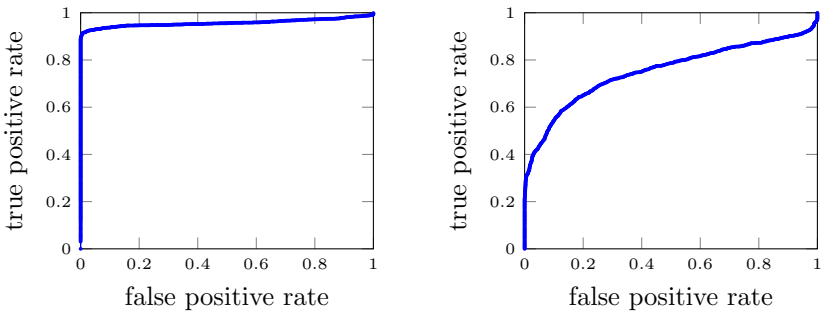


Figure 3.5: Example of ROC curves, *TPR* against *FPR* for different decision thresholds with an area under curve (*AUC*) of 0.96 (very good classification, left) and of 0.75 (worse classification, right)

this academic work, considering that a higher false alarm rate does not pose any safety issue.

The statistical concept of confidence intervals [Kreyszig 1970: chap. 12; Altman et al. 2003] shall be only briefly touched on: the basic concept is making a statement for the population (here perhaps: every CFRP part that will ever be inspected and evaluated with the newly developed method) based on information gained from a sample out of this population (here for example all parts investigated for this thesis). This statement is then done with a certain confidence level: assuming for example the (unknown) true positive rate of a population to be 90%, then, the 95% confidence interval of one sample will contain this true value in 95% of cases — that is, the confidence intervals of 19 out of 20 samples will contain this true value. In NDT, one speaks of probability of detection of 90% at a confidence level of 95%.

In non-destructive testing there are several methods in use for calculating a probability of detection, depending on whether or not a signal height as outcome of the inspection is known and taken into account for a calculation, cf. Schnars & Kück [2009b] (and references therein). If signal height is not taken into account, i.e. only defect/no defect is distinguished, the method of Clopper & Pearson [1934] is used [see also Newcombe 1998; Altman et al. 2003: 54] — termed '29/29 method' in Schnars & Kück [2009b] — or hit–miss analysis [Schnars & Kück 2009b].

Within this work, confidence intervals have not been applied due to following reasons: each scan contains more than 3000 time series. Calculating a confidence interval on such a massive data source gives values of true positive rates with a confidence level of 95% which are only 1–2% below true positive rates calculated without confidence level. This statement would however pretend a result which is not serious — repeating a scan or scanning a different specimen of the same material can have an influence on the results that is distinctly higher than a variation of true positive rate in the range of few %.

4 State of the Art

A literature research regarding earlier attempts to detect porosity in carbon fibre reinforced polymer (CFRP) without a back-wall echo, a comparison of recurrence quantification analysis (RQA) with more standard time series analysis and the use of RQA for non-destructive testing is presented in section 4.1, 4.2 and 4.3.

4.1 Generation of a back-wall echo equivalent in the literature

Backscattering — intermediate echoes from the inner of a material reflected back to the sending transducer [cf. Blitz & Simpson 1996: 46; Goebbels 1980] — has long been used for the evaluation of metals, e.g. for grain size determination or inclusion detection in steel [Goebbels 1980], using oblique incidence; later extended to porosity detection e.g. in aluminium [Nagy et al. 1987] using normal incidence.

First research applications on porosity in composites used oblique incidence (in contrast to the standard method for inspection of CFRP, normal incidence). Roberts [1987] for example performed (following former work e.g. in Kak et al. [1983]) measurements on CFRP specimens with different porosity contents. The power spectrum (amplitude spectrum squared) allowed clear separation between e.g. 0.2% and 1.1% porosity (values which had been determined from ultrasonic measurements in Qu & Achenbach [1987]), but also depended on the angle between the projection of the ultrasonic beam on the surface of the investigated specimen to the fibre direction. Furthermore, dependence on the depth distribution of porosity was reported. The data was obtained by spatial averaging [cf. Goebbels 1980], i.e. averaging of several A-scans from different (neighbouring) locations.

For all following applications on detection of porosity in composites via intermediate echoes (backscattered signals), normal incidence ultrasonic testing was used. The important features of the investigations of each research group are summarised in table 4.1. Given percentages of porosity are relative to the volume of the whole part; they were destructively determined via microscopy with subsequent image analysis, cf. Mehdikhani et al. [2019], if not otherwise stated.

The first application of intermediate echoes from normal incidence ultrasonic testing for porosity in CFRP known to the author is reported in three publications by Grolemund & Tsai [1992, 1995, 1998]. The authors investigated 11 porous PMR-15 polymide resin based fabric CFRP specimens of 6.4 mm thickness (porosity from 0–9%). In Grolemund & Tsai [1995] they determined the influence of phase-sensitivity: in a quasi normal incidence and quasi pulse-echo set-up (emitting and receiving probe

Table 4.1: Research on the characterisation of porosity in carbon fibre reinforced polymer via intermediate (backscatter) echoes from normal incidence ultrasonic testing. No entry in a cell means that this information cannot be deduced from reference

reference	inspection frequency	ply thickness	relevant inter-ply reflections	increasing porosity $\Longrightarrow$ intermediate echoes generally...	spatial averaging of A-scans	features generated out of intermediate echoes
Grolemund & Tsai [1998]	5 MHz	0.04 mm		increase	yes	mean of absolute values, bandwidth of autocovariance
Hillger & Elze [2002]	2–12 MHz	0.35 mm	yes	increase	via histogram	histogram of maximum absolute value of several A-scans
Dominguez & Mascaro [2006]	4 and 5 MHz	0.25 mm	yes, resonance effect	partially increase	no	(local) peak frequency
Kim et al. [2013]	20 MHz	0.25 mm	yes	decrease		maximum, amplitude of peak frequency, maxima of discrete wavelet transform coefficients & (presumably) their amplitude spectrum
Kurz & Rabe [2016]	5 MHz				yes	linear predictive coding (on first portion of intermediate echoes)
Karabutov & Podymova [2013]	0.2–8 MHz		apparently no	increase	apparently no	power spectrum (squared amplitude spectrum)
Chen et al. [2016]	10 MHz	0.25 mm	yes, resonance effect	presumably decrease	comparison via C-scans	centre frequency (mean of frequencies between cut-off frequencies)
Jin et al. [2019]	5 MHz	0.25 mm		increase	no	RQA: recurrence rate; maximum and variance shown but not evaluated
Wang et al. [2019a]	5 MHz	0.125 mm	presumably yes	rather decrease	apparently no	RQA: recurrence rate, determinism, entropy

placed close to each other) with 2.25 MHz inspection frequency, results of a 13.5 mm diameter probe were compared to the sums of maxima of a 1.5 mm diameter probe of measurements from 9 x 9 different locations, covering the same area as the 13.5 mm probe. Doing so, the influence of phase-sensitivity, i.e. potential interferences of reflections of near-by scatterers (pores), should be investigated. The results show an improved correlation between amplitude and porosity for the phase-insensitive measurements, i.e. from the 1.5 mm probe. Means of six positions on each specimen were taken. Overall, in all three publications amplitudes increased with increasing porosity, presumably only small or no inter-ply reflections occurred. Grolemund & Tsai [1998] (work started with Grolemund & Tsai [1992]) investigated the same specimens with a 12.7 mm diameter 5 MHz transducer in normal incidence in a water tank. An intermediate echo time series of 300 data points (3 µs) was used, and filtering to remove 'some of the frequency dependence of the system' as well as depth-amplitude correction (time corrected gain) was applied. The authors used as features the mean of the absolute values of the intermediate echo time, these means spatially averaged over an area of 5 x 5 A-scans (step size 0.5 mm), as well as the 2-D autocovariance over 20 A-scans on a line with 0.5 mm distance, out of which the maximum full width (at) half-maximum (the bandwidth at −6 dB) is determined. (Assuming zero mean of the intermediate echoes, variance — used in this thesis — is the mean of the *squared* values; and the Fourier transform of the autocovariance function of a time series equals the power spectrum [amplitude spectrum squared] of the time series [Mitra 2011: 899; Zbilut & Marwan 2008].) The averaged means show good correlation with porosity content; results are not that clear for the bandwidth. Again, these results were obtained through spatial averaging presumably without large inter-ply resonances.

Hillger & Elze [2002] [also Hillger et al. 2002] showed the detection of localised porosity fields at a 4.1 mm thick fabric specimen of the same series as investigated in this thesis (table 5.2). With ultrasonic equipment and parameters (including time corrected gain) optimised for detection of small intermediate echoes — transducer of 2–6 MHz frequency range and 12 mm diameter for thickness $\geq$ 3 mm; transducer of 6–12 MHz, 6 mm diameter for thickness $<$ 3 mm — the maxima of intermediate echoes of numerous A-scans of an area of the specimen were plotted as histograms, which showed different distributions for areas with and without porosity.

Dominguez & Mascaro [2006] investigated Fourier analysis to localise porosity in the depth of CFRP specimens without relying on spatial averaging. They obtained intermediate (backscatter) echoes with a resonance effect (cf. sect. 2.2.1). Eight ply (thickness 0.25 mm) CFRP with a high porosity of 10% with low porosity plies on top and bottom delivered with an inspection frequency of 5 MHz direct echoes of a lower frequency than the resonance (using short time Fourier transform). With eight plies of 3% porosity at the bottom of 20 low porosity plies and an inspection frequency of 4 MHz the authors observed a drop in the local intermediate echo frequency compared to the resonance.

Kim et al. [2013] investigated discrete wavelet transfrom on intermediate (backscattered) echoes on different CFRP materials (unidirectional and woven epoxy based and woven polymide PMR-15 based) with porosity content between 0% and 5–9%,

determined destructively via acid digestion. They used a 20 MHz transducer, 6.45 mm diameter, for investigating specimens with a correspondingly small thickness (not explicitly given, but according to A-scan ≈ 2 mm). The authors used as features the maxima of the original signal, of its amplitude spectrum, of the approximate and detail (low and high frequency) components of a level 5 Daubechies wavelet transform and apparently of the amplitude spectra of the approximate and detail components. The authors reported significant inter-ply reflections; the maxima decreased with increasing porosity in an exponential relation, for all four specimen types with correlation coefficients from ≈ 0.6 to 0.9. All maxima show a similar performance, the ones of the original signal and of the wavelet components are slightly better than those of their amplitude spectrum.

Karabutov & Podymova [2013] investigated the frequency spectrum of the backscatter signals of porous CFRP. They utilised laser ultrasound [Fahr 2014: sect. 7.5.3] in a contact set-up with a laser to thermoelastically excite an ultrasonic wave in the specimen and a conventional piezoelectric transducer for reception. The laser generated ultrasonic wave has a broad spectrum (compared to the ones generated by standard transducers) and the receiver captured 0.2 MHz to 8 MHz, which the authors claim to be advantageous for detection of pores with varying sizes. The investigated unidirectional CFRP of 12 mm thickness apparently did not contain significant interfaces for ultrasound between the plies. The authors used the ultrasonic velocity, determined via the existing back-wall echo, as reference for porosity content. The investigated feature was the power spectrum (squared amplitude spectrum) of the whole A-scan, including the entrance echo (presumably without the back-wall echo) minus the computed spectrum of the entrance echo to obtain what the authors call structural noise power. An empirical dependence of porosity — calculated via the ultrasonic velocity — on the structural noise power was shown, noise power increasing with increasing porosity.

Kurz & Rabe [2016] reported the usage of linear predictive coding, a tool of language processing, following Lozak et al. [2014]. They investigated two sets of samples of 2 mm to 8.6 mm thickness and porosity from 0% to 16.3%, determined via computed tomography, and used only a first portion of the intermediate echoes, measured with a 5 MHz transducer, to identify porosity. Averages of the derived feature over a whole area or specimen (spatial averaging) correlated with the porosity content determined by microsections; a 1:1 correlation of the feature with porosity was shown not to work.

Chen et al. [2016] observed, similarly to Dominguez & Mascaro [2006], a change in the main frequency of the intermediate echo (backscatter) time series. Using a 10 MHz transducer with specimens of 0.125 mm ply thickness, obtaining a clear resonance effect, they first found a decreasing centre frequency for increasing porosity. The centre frequency is here the mean of all frequencies between the cut-off frequencies. A subsequent investigation of two samples with 3 mm and 4–5 mm thickness, respectively, was done via a comparison of C-scans of back-wall echo and of centre frequency, which showed an at least qualitative agreement of back-wall echo reduction and centre frequency. Large back-wall echo reductions were observed (colour bar going from ≈ −10 dB to ≈ −40 dB to −50 dB). After identification of porous areas, B-scans

with a short time Fourier transform and thus a depth information of amplitudes led to identification of depth of porosity. One of two microsections to verify the results showed a large void of > 0.5 mm (assuming mm and µm were mixed up in the scale).

Two research groups in China picked up the idea of using RQA for detecting porosity in CFRP with ultrasonic testing after first publications within the scope of this thesis: Jin et al. [2019] [cf. He et al. 2018] measured ultrasonically on 61 points of specimens of 3–4 mm thickness with a 5 MHz transducer and subsequently determined the porosity content with micrographs. In parallel, the authors simulated ultrasonic signals with 35 different real morphology void models, representing seven different porosity contents. Though, according to the A-scans, a resonance effect seems to occur, signals increased with increasing porosity. The authors calculated recurrence rate (RR), which decreased with increasing porosity. Maximum and variance of the intermediate echo time series were also presented and in principal increase with porosity, but were not further evaluated. Experiments and simulation agreed in showing varying back-wall echo heights as well as varying recurrence rates for the same porosity content, based on different pore sizes and morphology. The authors showed a smaller variation of RR and thus advantage compared to back-wall echo (BWE) height for a porosity content $> 2\%$.

Wang et al. [2019a] [also Wang et al. 2019b] investigated RQA features RR, DET and ENT. Five specimens of 4 mm thickness, 0.125 mm ply thickness and different porosity contents were investigated with a 5 MHz transducer. Obviously a resonance effect occurs, and RR, DET (determinism) and ENT (entropy) increase with porosity (DET only rather slightly though). Afterwards 12 areas of a specimen of 9 mm thickness and also 0.125 mm ply thickness were inspected. Subsequently performed microsections of all areas confirmed that two areas with relatively high RR and ENT correspond to large voids concentrated in one or two interply regions — the authors thus claim that RQA is appropriate to detect localised porosity.

4.2 Recurrence quantification analysis compared to standard time series analysis

Not much literature compares recurrence quantification analysis with other, more conventional methods in time or frequency domain. The distinctly different approach of embedding a time series into a multi-dimensional space is very likely one reason; another may be the failure of linear methods for signals stemming from chaotic systems, e.g. not being able to differentiate stochastic from chaotic signals with Fourier analysis [Kantz & Schreiber 2005: 21; Loistl & Betz 1993: 50]. Now that RQA — originally created for chaotic dynamical systems [Eckmann et al. 1987] — is also deployed for complex signals from non-chaotic systems, to the author's opinion such comparisons with results of more traditional methods are enlightening.

Rohde et al. [2008] showed that the (squared) cross distance matrix (definition 7) can be related to the covariance (of the two time series) $s_{xy} = {}^1\!/_N \sum_{i=1}^{N} (x_i - \overline{x})(y_i - \overline{y})$. Zbilut & Marwan [2008] derive this for the simpler case of a distance matrix (defini-

tion 5) of one time series and the autocovariance

$$s_x(\upsilon) = \frac{1}{N} \sum_{i=1}^{N-\upsilon} (x_i - \overline{x})(x_{i+\upsilon} - \overline{x}). \tag{4.1}$$

Without embedding ($d = 1$), the mean value of the squared distances $d^2(\upsilon) = {}^1/_N \sum_{i=1}^{N} D^2_{i;j=i+\upsilon}$ (that is the mean of the υth diagonal [the main diagonal being the 0th diagonal] in the distance matrix) is related as

$$d^2(\upsilon) = 2(-s_x(\upsilon) + s^2) \tag{4.2}$$

to the autocovariance s_x and the variance s^2. The derivation involves the assumption of a stationary process with same mean and variance of the shifted series $x_{i+\upsilon}$.

Rohde et al. [2008] furthermore compare 'well-known detectors based on the likelihood ratio' with RQA using simulated data, claiming the latter being 'sub-optimal'. The simulated data is a linear stochastic process, a harmonic process and a Duffing system (non-linear, potentially chaotic) [Argyris et al. 2015: sect. 10.5]; and furthermore the authors investigated detection of (known or unknown) deterministic signals in noise. From the point of view of this thesis' author, their results are interesting, but of limited value: RQA may not be the first choice for signals of linear origin that can be easily represented; and regarding the chaotic Duffing system, Rohde et al. [2008] compare the squared distances with the autocovariance for 1000 realisations, i.e. time series out of the Duffing system, and find indeed only small errors between both.

Marwan & Kurths [2009] comment on Rohde et al. [2008], having concerns about 'such a general statement' [that RQA performs below traditional detectors], and claiming that 'the conclusions of Rohde et al. [2008] may hold for certain linear systems and under certain assumptions'.

4.3 Recurrence quantification analysis in the application for non-destructive testing

Some applications of RQA for non-destructive testing are known. First applications are for structural health monitoring (with permanently installed sensors): Nichols et al. [2006] used simulated results of a steel plate with a cut from nine different (virtual) sensors from a chaotic vibration excitation and cross recurrence plots. Lu & Michaels [2006] deployed recurrence rate from the convolution of the signals from guided waves measurements from transducers in the MHz range on aluminium plates with artificial defects (a drilled hole and a notch) with chaotic signals. Nomura et al. [2013] also investigated recurrence rate (here designated %REC) in an application of civil engineering by applying a chaotic excitation and measuring the response, with which the authors reported a successful damage localisation.

Cacciola et al. [2007] investigated artificial defects (cavities and cracks) using eddy current non-destructive testing (here magnetic field measurements). They created cross correlation plots of (unthresholded) recurrence plots (where the process of this

creation is not described in detail and apparently different to cross recurrence plots) and subsequently quantified higher order statistics presumably of each point of this plot.

Carrión et al. [2014][cf. Carrión et al. 2015a] investigated different techniques of nonlinear time series analysis on the determination of different porosity levels in cement pastes with intermediate echoes of ultrasonic pulse-echo testing: detecting determinism by continuity from Wayland et al. [1993], based on Kaplan & Glass [1992] or by determining smoothness of trajectories in state space [cf. Jeong et al. 2002]. Furthermore, they used determinism (here denoted DRP) and observed higher values for higher porosity for one water-cement ratio. The same group later used the feature determinism DET to determine cracks introduced by compressive load on cement cubes in an ultrasonic through transmission set-up and compared them with measurements of ultrasonic velocity and ultrasonic attenuation, where DET partially performed better than these traditional measures [Carrión Garcia et al. 2017].

Besides several publications picking up the idea of RQA for ultrasonic testing of porosity in CFRP (cf. sect. 4.1), Teng et al. [2019] used RQA for detection of discrete defects. They investigated the detection of three side drilled holes of diameter $\leq 1\,\mathrm{mm}$ in 6 mm depth of a 10 mm thick CFRP part with 0.125 mm ply thickness. Features RR, DET and laminarity (the occurrence of horizontal and vertical lines in a recurrence plot) were used; the authors performed 100 measurements (presumably at exactly the same location) of three defect-free areas and at the three defects with an ultrasonic transducer of 7.5 MHz (no resonance effect). Median, interquartile range (difference between first and third quartile), maximum and minimum of 100 values of each feature were calculated, based on recurrence plots of the second half of the intermediate echo time series. The authors claim that the 0.7 mm and 1 mm side drilled hole can be detected by changes in median and interquartile range, whereas these defects are not visible as direct echoes in the A-scan; they say that the transfer to real defects has however still to be made.

Furthermore, Yang et al. [2019] extended the idea of using ultrasonic backscattered signals for the detection of local defects to metallic materials. The authors investigated 0.8 mm flat bottom holes in steel and reported observation of the defect in the recurrence plot as well as with RQA features such as determinism and entropy.

5 Finding a Back-Wall Echo Equivalent. Evaluation of Ultrasonic Data

This central chapter presents the evaluation to generate a back-wall echo out of the intermediate echo time series of ultrasonic testing using recurrence quantification analysis, standard tools in time domain and Fourier analysis.

In section 5.1 specimens, measurement equipment and process are described, and section 5.2 shows the general approach of evaluation. The evaluation of fabric material with porosity introduced through Freon capsules are presented in section 5.3. Section 5.5 and 5.6 contain evaluation of unidirectional material with porosity introduced by process parameters outside the specified range, with resonance effect and without resonance effect. Prior to this, some parameters are pre-chosen based on separate data sets of unidirectional material, shown in section 5.4.

5.1 Measurements

The following two subsections describe the equipment, carbon fibre reinforced polymer (CFRP) specimens and the process of data acquisition through ultrasonic measurements. The basic approach is to use CFRP specimens of simple plane parallel geometry with constant thickness. The ultrasonic test of these specimens delivers a back-wall echo (*BWE*) as well as intermediate echoes. For this thesis, the *BWE* serves as label, providing information about porosity content (cf. sect. 2.2.3). The intermediate echo time series — out of which the back-wall echo equivalent is generated — stem from the same A-scan as the reference *BWE*; the intermediate echoes are however not influenced by existence or non-existence of a back-wall since they occur earlier in time.

5.1.1 Ultrasonic equipment and parameters

Measurements have been performed with commercially available ultrasonic electronics *Multi2000 32x128* by company M2M, including a laptop for controlling and recording, and an Olympus *5L64* 5 MHz ultrasonic array probe (fig. 5.1). The probe consists of 64 elements, covering an area of 46 mm × 10 mm. It was driven with 5 MHz frequency and a voltage of 40 V with an aperture of eight elements firing simultaneously. The step between two apertures was two elements (so, elements 1 to 8, then 3 to 10, etc. were activated), achieving an electronic scan step of 1.44 mm and 29 A-scans per one electronic scan (in the following referred to as a 'column'). Pulse repetition frequency was 2 kHz. An amplification of 18.9 dB (fabric specimens) and 31 dB (unidirectional

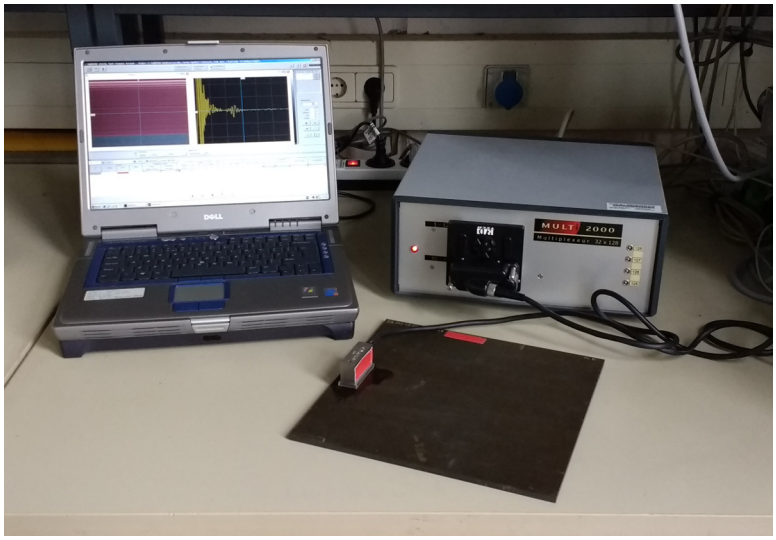


Figure 5.1: Ultrasonic measurement equipment: electronics (top right), one specimen and phased array probe (bottom) and controlling laptop (left)

specimens) was applied, not using any time corrected gain (TCG). No filters on the received signal were applied; the device sampled the signals with a frequency of 100 MHz (acquiring an amplitude value every 10 ns) and digitized with 10 bit, i.e. 1024 steps for ±100% screen height.

5.1.2 CFRP specimens, execution of measurements and data preparation

All specimens are of materials used in aeronautic industry; Airbus Operations GmbH kindly provided them.

Two different types of CFRP material were used: a fabric material 913C-926-35%F of ply thickness 0.35 mm and a unidirectional material B (of two different ply thicknesses, 0.25 mm and 0.19 mm), both epoxy based. The latter is the more relevant state-of-the-art aerospace material; specimens either without porosity or with porosity uniformly distributed over an entire specimen are available, generated by changing the autoclave parameters (temperature, pressure) to be outside the specified range. The former material, 913C-926-35%F, contains artificially introduced porosity fields (using the blowing agent Freon) of different amounts [Damaschke 1996: 12ff.; Gausmann 1995: 29ff.] in the resin areas between the fabric plies (fig. 5.2).

Airbus Central Research & Technology and Testia GmbH kindly provided an X-ray computed tomography (CT) [Fahr 2014: sect. 10.7] image (fig. 5.3) of one porous sample of unidirectional material B used in this work (of an area different to the areas ultrasonically tested in this thesis; the specimen was cut prior to this thesis to have smaller pieces for CT). Porosity is located in the (few 10 µm thick) pure resin layers between the CFRP plies; the pores are rather uniformly distributed over thickness with sizes in the range of below 100 µm to several 100 µm.

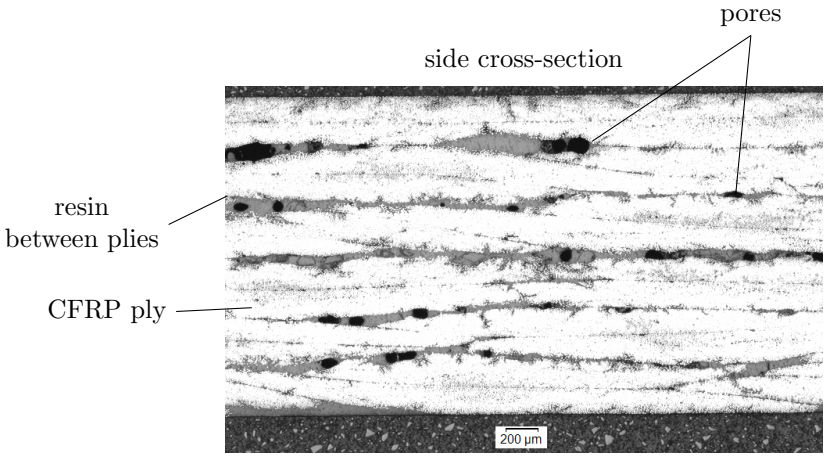


Figure 5.2: Microsection of a 2 mm thick specimen of the same series as M13–15, material 913C-926-35%F, porosity introduced with Freon capsules, ©Airbus Operations GmbH

All specimens are plane parallel of 120 mm × 100 mm (three fabric specimens) or approximately 290 mm × 290 mm (eight unidirectional specimens).

The CFRP materials of the specimens contain resin layers between the plies, such that for the fabric material and for the unidirectional material of thickness 0.25 mm a resonance effect occurs with the chosen inspection (mean) frequency of 5 MHz, corresponding to $\lambda = 0.6$ mm (cf, sect. 2.2.1).

The author created each scan manually in contact technique by moving the transducer in one line horizontally over the whole length of the specimen, partially three lines per specimen were inspected, each row in table 5.1 to 5.3 corresponds to one scan area (line).

Fabric specimens M13–M15 were inspected with spraying a small amount of water on them to ensure the water coupling between probe and specimen. Specimens of material B were placed in a small immersion tank just covering the specimen surface with water to ensure water coupling; the repeatability of scans was too poor with only a small water film (applied by spraying) because of the rough surface of some of the specimens.

The acquisition of data in horizontal direction was time-triggered by the M2M hard- and software (one linear scan taken every 0.1 s). The recorded scans of M13–M15 consist of 101 columns of 29 A-scans, the ones of material B contain 251 columns (corresponding to a speed of ≈ 10 mm/s for both materials), so that for all specimens one column was recorded roughly every millimetre. All A-scans of one scan together are called volume scan.

To avoid areas with lower *BWE*s due to coupling issues — the probe wobbled because of manual scanning — cut-outs of the recorded physical scans are used for evaluation in this work, each cut-out contains 65 × 29 A-scans.

The positions of the cut-outs in the scans can be found in table B.1 and B.2 in

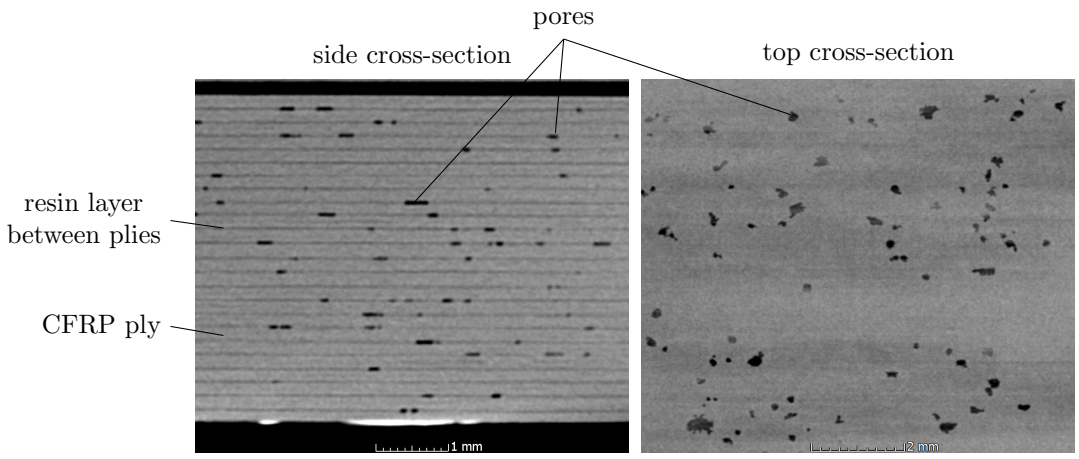


Figure 5.3: X-ray Computed Tomography images of a cross-section perpendicular to the surface (left) and parallel to the surface (right) of material B, specimen 26 with porosity (cf. table 5.3), porosity created with autoclave parameters outside the specified range, ©Airbus Operations GmbH, Airbus Central Research & Technology and Testia GmbH

Table 5.1: CFRP specimens and scans used for preadjustment of parameters; one row in the table corresponds to one (area of one) specimen. The classification of porosity content into medium refers to a *BWE* reduction in the range of or close to 50%. The areas are different to those in table 5.2 and 5.3.

CFRP material	ply thickness (appr.)	porosity content	thickness	specimen-scan	scan date	number of cut out scans
unidirectional material B	0.25 mm	low (reference specimen)	6 mm	50-α1–2	May 2019	2
				50-β1–2	May 2019	2
unidirectional material B	0.25 mm	medium	6 mm	56-α1–2	May 2019	2
				56-β1–2	May 2019	2
unidirectional material B	0.19 mm	low (reference specimen)	6 mm	24-α1–2	May 2019	2
				24-β1–2	May 2019	2
unidirectional material B	0.19 mm	medium	6 mm	38-α1–2	May 2019	2
				38-β1–2	May 2019	2

appendix B. Due to slightly varying speed of manual scanning in connection with time-triggered recording, these cut-outs do not cover one-to-one the same areas of a specimen from scan to scan.

For main investigations, one scan of specimen M13 and M14 and two scans of M15 are used in evaluation (table 5.2; for classification only one scan of M15 is used, cf. sect. 5.3.1). For material B, four or two areas, respectively, of every specimen, with three scans each are evaluated (table 5.2 and 5.3). The three (volume-)scans of one area are referred to as scan set in the following.

In addition, for pre-investigations to adjust few parameters (sect. 5.4) prior to the investigations on unidirectional specimens with porosity created by process parameters outside the specified range, additional areas out of the specimens are used (table 5.1, independent of the ones used for main investigations table 5.2 and 5.3). These consist of two scan sets (each with two scans) for material B both without and with resonance effect, respectively. The measurement data obtained with the M2M equipment was read into the evaluation software *Ultis* [cf. Barut et al. 2012] and stored in data format *.nka*. This data format was read into Matlab. RQA computations are carried out with Python (with the exception of computations used for correlation results, for which Matlab is used); algorithms in time domain and Fourier analysis as well as classification for all methods are performed with Matlab. Self-programmed functions are used throughout this thesis (core functions to calculate features are given in app. D).

The data sets are designated

M15-2,

for the second scan of specimen M15 or

50-A2,

for the second scan of the first measurement area of specimen 50.

If a certain A-scan out of one volume scan shall be referenced, the horizontal position is referred to with x and the vertical position with y; this way, the e.g. A-scan at the 11th horizontal position and the 8th vertical position in first scan of specimen M15 is designated by

M15-1x11y8.

5.2 General approach for evaluation

Section 5.3 to 5.6 show results for fabric material (with resonance effect); for unidirectional material with resonance effect; and for the latter material without resonance effect. Section 5.3, 5.5 and 5.6 each divide into investigation of recurrence quantification analysis (RQA), of standard methods in time domain and of Fourier analysis. Section 5.4 contains preparation work for the evaluation of unidirectional specimens.

After starting with correlation coefficient for fabric material specimens, results of bi-classification (sorting A-scans in 'porous/non-porous' according to back-wall echo height) are used as quality measure for the back-wall echo equivalent — due to large spread of values despite high correlation coefficients, cf. section 5.3. The classification

Table 5.2: Investigated CFRP specimens and scans, ply thickness of 0.35 mm and 0.25 mm (resonance effect with inspection frequency of ≈ 5 MHz); one row in the table corresponds to one (area of one) specimen. The classification of porosity content into medium and large is according to whether the BWE reduction is close to 50% or significantly larger, respectively. The areas are different to those in table 5.1.

CFRP material	ply thickness (appr.)	porosity content	thickness	specimen-scan	scan date	number of cut out scans
fabric 913C-926-35%F	0.35 mm	fields of medium porosity	7 mm	M13	July 2015	1
fabric 913C-926-35%F	0.35 mm	fields of medium porosity	7 mm	M14	July 2015	1
fabric 913C-926-35%F	0.35 mm	fields of large porosity	7 mm	M15-1,2	July 2015	2
unidirectional material B	0.25 mm	low (reference specimen)	6 mm	50-A/B1–3	May 2019	6
				50-C1–3	May 2019	3
				50-D1–3	May 2019	3
unidirectional material B	0.25 mm	medium	6 mm	56-A/B1–3	May 2019	6
				56-C1–3	May 2019	3
				56-D1–3	May 2019	3
unidirectional material B	0.25 mm	low (reference specimen)	4 mm	49-A/B1–3	March 2019	6
unidirectional material B	0.25 mm	medium	4 mm	55-A/B1–3	March 2019	6

Table 5.3: Investigated CFRP specimens and scans, ply thickness of 0.19 mm (no resonance effect with inspection frequency of ≈ 5 MHz); one row in the table corresponds to one (area of one) specimen. The classification of porosity content into medium refers to a *BWE* reduction in the range of or close to 50%. The areas are different to those in table 5.1.

CFRP material	ply thickness (appr.)	porosity content	thickness	specimen-scan	scan date	number of cut out scans
unidirectional material B	0.19 mm	low (reference specimen)	6 mm	24-A/B1–3	May 2019	6
				24-C1–3	May 2019	3
				24-D1–3	May 2019	3
unidirectional material B	0.19 mm	medium	6 mm	38-A/B1–3	May 2019	6
				38-C1–3	May 2019	3
				38-D1–3	May 2019	3
unidirectional material B	0.19 mm	low (reference specimen)	4 mm	26-A/B1–3	May 2019	6
unidirectional material B	0.19 mm	medium	4 mm	36-A/B1–3	May 2019	6

is assessed through balanced accuracy, which is the mean of true positive rate (detection rate) and true negative rate (1 minus false alarm rate, sect. 3.4.2): the greatest balanced accuracy is searched for, with the additional constraint of $TPR \approx TNR$ to avoid 'unbalanced balanced accuracy' of e.g. 99% detection with a false alarm rate of 19%.

Classification results are evaluated in a cross-validation approach (sect. 3.4.2): three scans (or four scan sets, respectively) are split into

- all minus one scan (or set, respectively) used for training,
- the remaining scan (or set, respectively) is used for test and
- this is rotated in a way that all three scans (or four scan sets) are used for test once (cf. fig. 3.4 regarding this principle).

(A random split into training and test sets is not applied to ensure that one specimen or specimen area, respectively, is completely kept out of training.)

'Training' is here to be understood in a broader sense, being either

- only a determination of the decision threshold to determine positives and negatives for methods without any other parameters to adjust, such as taking the variance of the intermediate echo time series, or
- finding the parameters for which optimum classification (balanced accuracy) is achieved for recurrence quantification analysis, or
- direct parameter determination on the training set(s) for linear regression.

Tables with all classification results and the corresponding correlation coefficients are given in appendix C. Following sections 5.3 to 5.6 present the important results.

5.3 Evaluation on fabric specimens

The three specimens M13, M14 and M15 out of fabric material with artificially introduced porosity (from Freon capsules) are chosen to be investigated first, since they contain well defined and localised porosity. In this section and following subsections, features generated with recurrence quantification analysis (sect. 5.3.1), with standard tools in time domain (sect. 5.3.2), and through Fourier analysis (sect. 5.3.3) are used as a back-wall echo equivalent and tested for their capability to replace the established *BWE* criterion. Up to four scans are evaluated: one of M13 and M14 each and two of M15.

5.3.1 Recurrence quantification analysis

For specimens M13–15, *RR* and *DET* are investigated as back-wall echo equivalent.

Correlation coefficient as quality measure

Investigations are presented in Brandt [2016], the evaluation part is reproduced in appendix A; a summary and some amendments follow. For these evaluations, an intermediate echo gate from 0.99 µs to 3.99 µs (included) is used, and LOI has been taken into account for feature calculation (cf. sect. 3.3.5). The time of flight of 3 µs corresponds with a longitudinal wave velocity in CFRP of $\approx 3000\,\mathrm{m\,s^{-1}}$ to approximately 4.5 mm thickness.

Choice of embedding dimension d and time delay τ for time-delay embedding of time series prior to recurrence matrix generation is data-driven throughout this thesis, i.e. the parameters are varied and chosen in a way that the best back-wall echo equivalent (BWE-equivalent) is obtained. A quick test of false nearest neighbour method and of first zero crossing of the autocorrelation (app. A) led to the decision to vary $d \in \{1, 2, \ldots, 10\}$ and $\tau \in \{1, 2, \ldots, 10\}$.

Threshold ϵ for determining recurrence is in general altered within the whole range, achieving recurrence rates from almost zero to almost 1. Euclidean as well as angular distance is used as distance measure. The minimum line length l_{min} (necessary to determine *DET*) is essentially varied between 2, 4 and 8.

For all parameter variations, the correlation between BWE-equivalent (*RR* or *DET*, respectively) and *BWE* is calculated, leading to more than thousand computations for Euclidean and for angular distance. Looking for RQA parameters (d, τ, ϵ and l_{min}) that lead to the greatest positive or smallest negative correlation coefficients, following main results are obtained:

- The optimum correlation coefficients r for Euclidean distance are in general better than for angular distance.
- Two out of four of these optimum r, using Euclidean distance (minimum and maximum both for *RR* and *DET*), go along with either very sparse recurrence plots (corresponding to low recurrence rates), containing merely the line of identity, which gets slightly thicker in case the A-scan has been taken from a porous

area; or with very high recurrence rates (almost 1) with a very low range, i.e. the difference between RR of all RPs in scan M15-1 is low, $R = 0.05$.

- The minimum negative r achieved with determinism DET using Euclidean distance leads to reasonable recurrence plots and RR, but the range of DET is rather low. The result is extremely sensitive to changes of parameters (plus a potential non-optimal embedding with lines perpendicular to the line of identity). However:

- The correlation for DET (minimum negative correlation coefficient using Euclidean distance) is significantly improved with $l_{min} > 2$. A correlation coefficient of $r = -0.87$ is achieved with $\epsilon = 2$ (Euclidean distance), $d = 3$, $\tau = 6$ and $l_{min} = 8$. This goes along with a reasonable range of DET and with robustness against changes of parameters.

- Three out of four optimum correlation coefficients for angular distance (minimum and maximum both for RR and DET) have an absolute value less than 0.7. The fourth value is $r = 0.75$ with DET, the corresponding embedding parameters of $\epsilon = 0.5\pi$, $d = 4$, $\tau = 8$ and $l_{min} = 8$ lead to reasonable recurrence rates (the RPs have however rather thick lines) and a relatively large range of DET.

An amplitude effect causes the good results in terms of r for Euclidean distance using determinism DET as BWE-equivalent: pores reflect the ultrasound in other directions than to the transducer and lead to a reduction of ultrasonic energy and thus of the intermediate echoes. Interestingly, due to this effect determinism DET *increases* for intermediate echo time series obtained from CFRP with porosity.

Using angular distance, recurrence plots with rather thick diagonal lines (due to large ϵ) loose their structure for measurements on areas of porosity, showing the reduced periodicity in the signal: determinism decreases with higher porosity.

The investigations so far reveal a good correspondence between determinism DET as back-wall echo equivalent and the back-wall echo (fig. 5.4). But even the results with best correlation coefficient $r = -0.89$ show an immense spread and fail to provide a good one-to-one correspondence to the BWE: taking the prediction interval into account, a measurement delivering a certain DET will correspond in 95% of all cases to a BWE within a range of $\pm \approx 2.5\%$ (fig. 5.5). This corresponds to ≈ 6 dB with regard to the mean back-wall echo in defect-free area, similar to the threshold separating porous from non-porous. It is thus not feasible to use the BWE-equivalent as a one-to-one representation of the BWE. The results are however presumably well suited for bi-classification, i.e. distinguishing good from porous areas (where the reference is again the back-wall echo height). The quality of the methods will thus hereafter be evaluated with a bi-classification, using balanced accuracy (mean of true positive and true negative rate), aiming at a value of 90%. The search for a maximum balanced accuracy is performed under the constraint true positive rate (TPR) $\approx$ true negative rate (TNR) (cf. sect. 3.4.2).

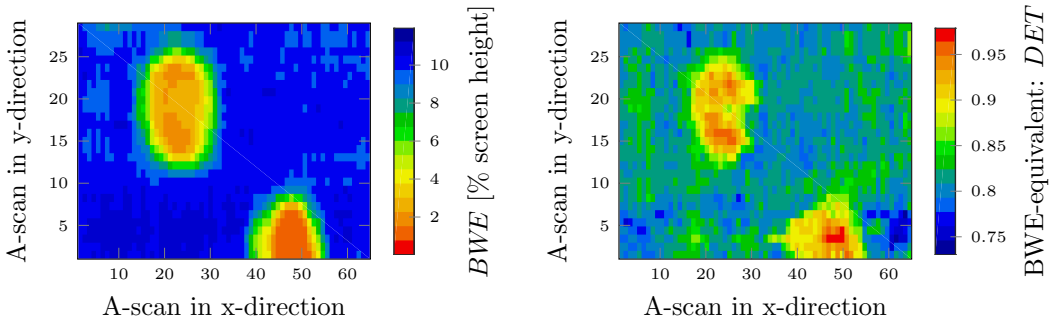


Figure 5.4: C-scans of M15 first scan (M15-1): *BWE* (left) and *DET* as BWE-equivalent with RQA parameters $\epsilon = 2$ (Euclidean distance), $d = 3$, $\tau = 6$ and $l_{min} = 8$ (right) (reproduced after fig. A.4 but with reversed colours)

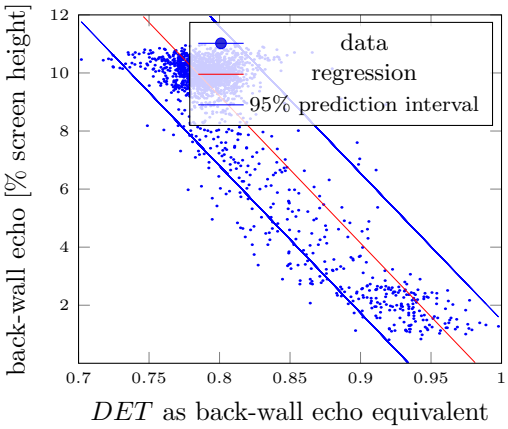


Figure 5.5: Back-wall echo against determinism *DET* as BWE-equivalent, using Euclidean distance, with regression line and 95% prediction interval; scan M15-2, $d = 3$, $\tau = 6$, $\epsilon = 2$, $l_{min} = 8$; correlation coefficient $r = -0.89$

Classification — Euclidean distance

The scans of the fabric specimens are henceforth evaluated in terms of bi-classification. An intermediate echo gate of 1 µs to 4 µs is used.

For the remainder of this thesis, the following applies:

- Parameters recurrence threshold ϵ and minimum line length l_{min} are varied in a broader range and (in most cases) in finer steps (table 5.4, table 5.17).

- Based on the experience in search of a maximum absolute value of correlation coefficient, certain constraints are posed on mean and range (over all A-scans of one volume scan or scan set, respectively) of recurrence rate and determinism (table 5.5).

Table 5.4: RQA parameters when searching for optimum classification of fabric specimens with artificial porosity

RQA parameter	Min	Max	Steps
Time delay τ	1	10	1
Embedding dimension d	1 (2 for ang. distance)	10	1
Recurrence threshold ϵ for Euclidean distance	0.1	10	0.1 for $\epsilon \in [0.1, 2]$; 0.5 for $\epsilon \in [2.5, 10]$
Recurrence threshold ϵ for angular distance	0.025π	0.9π	0.025π for $\epsilon \in [0.025\pi, 0.5\pi]$; 0.05π for $\epsilon \in [0.55\pi, 0.9\pi]$
Minimum line length l_{min}	2	10	1

- A cross validation approach is applied (cf. sect. 3.4.2 and 5.2): RQA parameters leading to optimal balanced accuracy are determined on all but one scans (or scan sets, respectively), and the left out scan (or scan set, respectively) is used for test.

- C-scans contain a colour and grey scale visualising the threshold to distinguish between porosity and non-porosity for ultrasonic back-wall echo (50% of mean *BWE*) and for the back-wall echo equivalent. The latter threshold is the decision threshold to distinguish whether or not a given back-wall echo equivalent predicts porosity. Non-porous areas (or predicted non-porous areas, respectively) are shown in a grey scale, and porous areas (or predicted porous areas, respectively) are shown in colours, see e.g. fig. 5.7. (It is accepted that in general not the whole range of either grey shades or colours, respectively, of the scale is present in the C-scans due to this adaptation.)

- All results are given in tables in appendix C. This section to section 5.6 present the important results.

- The line of identity (LOI) of the recurrence matrix is not taken into account for RQA computations.

No constraint on the minimum recurrence rate is set, taking recommendations in the literature — *RR* values between 0.1% and 2% [Zbilut & Webber 2006] or, given nonstationary data as in this work, values of $\approx 1\%$ [Zbilut et al. 2002] — into account.

In the following, only one scan of each specimen with artificial porosity is used, i.e. scan M13, M4 and M15-1; with two very similar scans of M15, one would always be included into training, leading to results potentially too positive.

Table 5.5: Constraints for search of optimum classification with recurrence rate and determinism as BWE-equivalent

classification based on	maximum of mean of feature over all A-scans	minimum of range of feature over all A-scans
recurrence rate *RR*	0.9	0.1
determinism *DET*	1	0.1
in addition for *DET* in this section, and *DET* and all further RQA features in sect. 5.5.1 and 5.6.1: maximum of mean of *RR* over all scans	0.9	

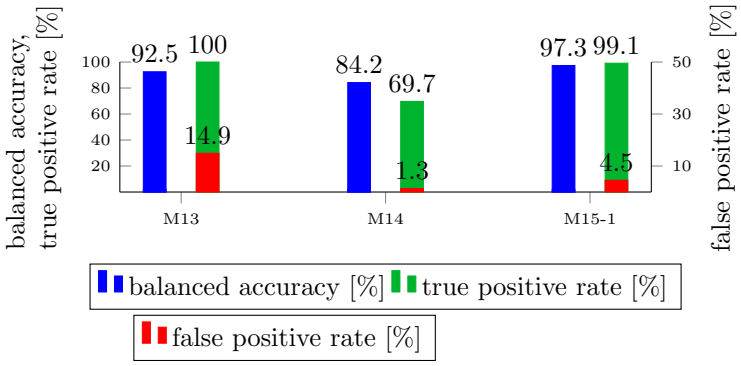


Figure 5.6: Classification results, RQA, with *DET* as BWE-equivalent using Euclidean distance, for fabric CFRP specimens. RQA parameters determined for each scan on two remaining scans, including decision threshold to achieve $TPR \approx TNR$

Recurrence quantification analysis using Euclidean distance delivers with *DET* as BWE-equivalent in two out of three test cases good to excellent balanced accuracies and *TPR*s (fig. 5.6).

C-scans of *BWE* and *DET* for M13 show the good classification results (fig. 5.7, with unbalanced *TPR* and *TNR* through the decision threshold determined on M14 and M15-1 though).

Recurrence plots of M13 (fig. 5.8 bottom) reveal the same effect as identified in search of optimum correlation coefficient (app. A, fig. A.3): porosity leads to a drop in amplitude especially in the second half of the intermediate echo time series (fig. 5.8 top right); most signals fall into the ball of radius ϵ, thus recurrence plots contain more recurrent points. *DET* increases if porosity is present and represents rather a larger recurrence rate. The feature *RR* as BWE-equivalent however only delivers insufficient test results with balanced accuracy below 80% and largely varying RQA parameters, cf. table C.2.

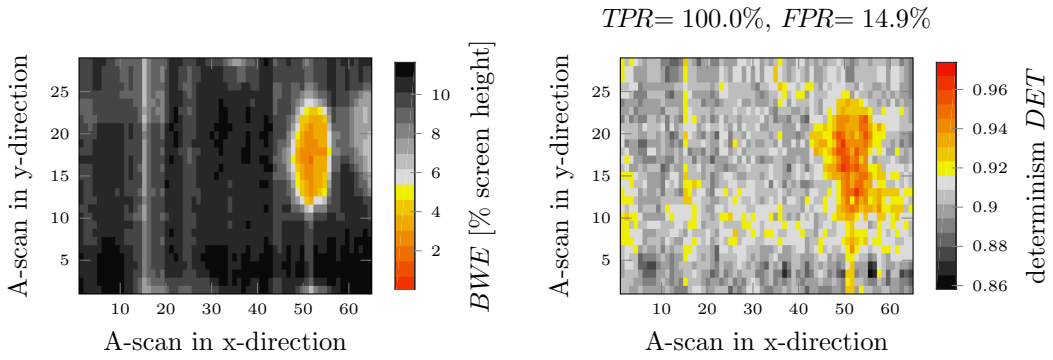


Figure 5.7: C-scans of M13; left: back-wall echo (porosity — *BWE* reduced by more than 50% — shown in colours); right: test results for *DET* as BWE-equivalent, using Euclidean distance, RQA parameters leading to best classification on M14 and M15-1 (embedding dimension $d = 3$, delay $\tau = 5$, recurrence threshold $\epsilon = 1.7$, minimum line length $l_{min} = 4$); decision threshold to distinguish (predict) porosity and non-porosity determined from training on M14 and M15-1 leads to unbalanced *TPR* and *TNR* (predicted porosity shown in colours, predicted defect-free areas in grey shades)

Classification — angular distance

RQA with angular distance as metric to generate recurrence matrices with features *RR* and *DET* as BWE-equivalent does not lead to balanced accuracies as high as for Euclidean distance. Determinism *DET* provides worse classification results than *RR* with balanced accuracies from 71.8% to 89.1% compared to 84.9%–87.2% (table C.5 compared to table C.4). A C-scan of M13 represents the moderate classification results (fig. 5.9) with a balanced accuracy of 86.5%

Recurrence plots show the effect of largely reduced recurrence in the case of porosity (fig. 5.10 right compared to left).

5.3.2 Evaluation in time domain

Time series' maximum absolute value x_{max}, variance s^2 and quartile coefficient of dispersion (QCD) do not have any parameters to adjust (to train). The training for these features simply consists of finding a threshold for predicting porous and non-porous such that $TPR \approx TNR$. QCD is used with an offset for the time series to contain merely nonnegative values (app. A).

Maximum and variance do not deliver reliable results (table C.6). Balanced accuracies are clearly below 90% and the correlation coefficient r is largely varying, i.e. the effect of these simple features is largely changing from specimen to specimen. r even changes its sign in case of variance as BWE-equivalent between M13 on one hand and M14 and M15-1 on the other hand, meaning that for M13 variance goes down for porous areas and for the other two scans it goes up for porous areas.

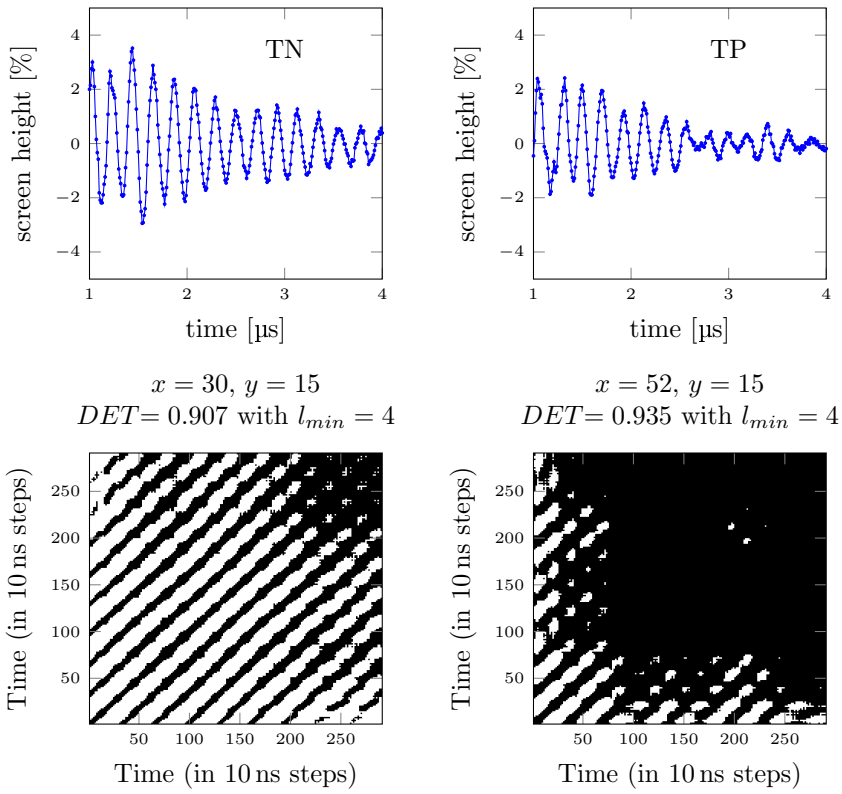


Figure 5.8: Bottom: recurrence plots of M13 with RQA parameters for optimum classification on M14 and M15-1, using DET, Euclidean distance ($d = 3$, $\tau = 5$, $\epsilon = 1.7$); top: corresponding A-scans; example for correctly predicted non-porous area (left) and for correctly predicted porous area (right)

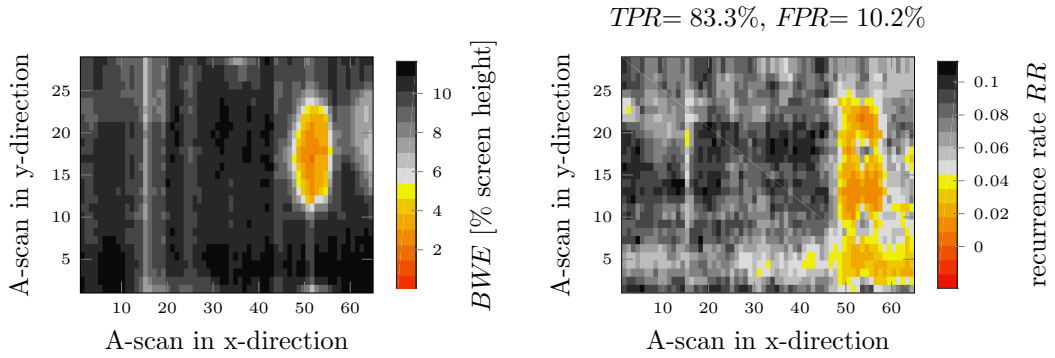


Figure 5.9: C-scans of M13; left: back-wall echo (porosity — back-wall echo reduced by more than 50% — shown in colours); right: test results optimum classification with RR using angular distance, RQA parameters leading to best classification on M14 and M15-1 ($d = 10$, $\tau = 7$, $\epsilon = 0.125\pi$) (predicted porosity in colours, predicted defect-free areas in grey shades)

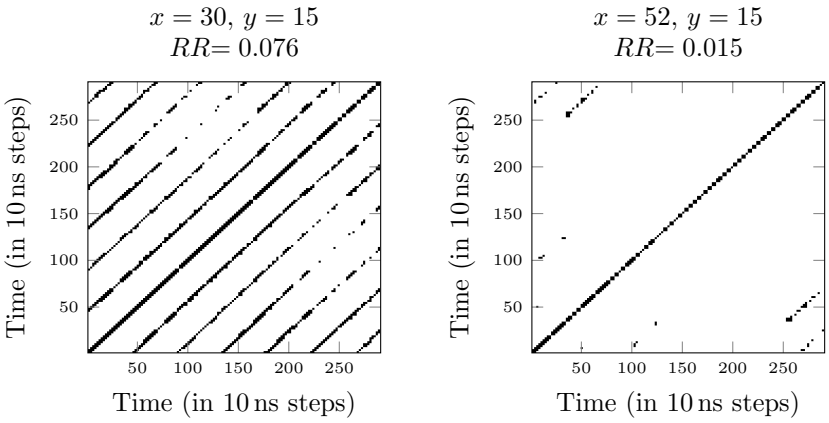


Figure 5.10: Recurrence plots of M13 with RQA parameters for optimum classification on M14 and M15-1, using RR, angular distance ($d = 10$, $\tau = 7$, $\epsilon = 0.125\pi$); example for correctly predicted non-porous area (left) and for correctly predicted porous area (right). Corresponding A-scans in figure 5.8

Quartile coefficient of dispersion delivers good to excellent results (fig. 5.11). Only M14 has slightly unbalanced TPR and TNR with $TPR < 90\%$ when using the threshold determined on M13 and M15-1. This is still considered sufficient, cf. corresponding C-scans (fig. 5.12). A-scan examples of M14 (fig. 5.13) show a somewhat different effect from the one observed in investigations of RQA (fig. 5.8): approximately within the first microsecond, intermediate echoes are *larger* in case of porosity. The porosity that was scattered between the plies as Freon capsules during the production of the specimens may here lead to larger reflections of the CFRP–resin layer–CFRP interfaces between the first plies. In larger thickness, the intermediate echoes are smaller (comparable with the effect seen in fig. 5.8).

Linear regression in time domain

The next investigated feature, linear regression, needs training: based on a training set containing intermediate echo time series and back-wall echoes, weights for each time point are determined in a way that summing the products of these weights and the corresponding time point resembles the back-wall echo as well as possible (cf. sect. 3.1). Test results for linear regression are clearly below 90% balanced accuracy and very varying from scan to scan; furthermore, TPR and TNR are very unbalanced for all scans (table C.7). This is worse than looking on first and third quartile via QCD. Echoes in the first microsecond of the evaluated time series being for porosity partly larger (fig. 5.13 right) and partly smaller (fig. 5.8 top right) may 'confuse' linear regression, whereas these large echoes are for QCD not taken into account due to their small fraction of all echoes.

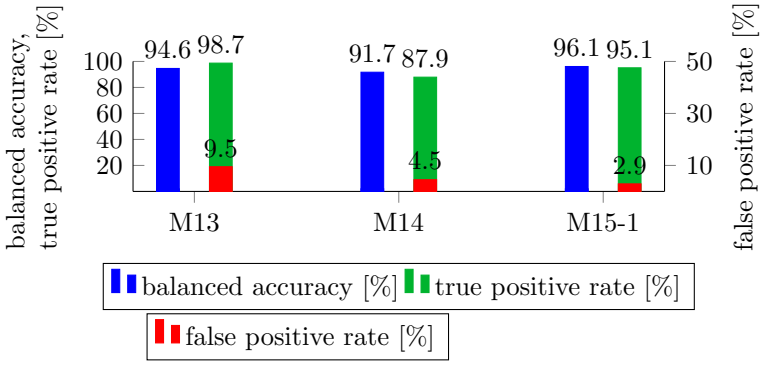


Figure 5.11: Classification and correlation results with quartile coefficient of dispersion as BWE-equivalent for fabric specimens. Decision threshold to achieve $TPR \approx TNR$ determined on two scans, results with this threshold on remaining scan

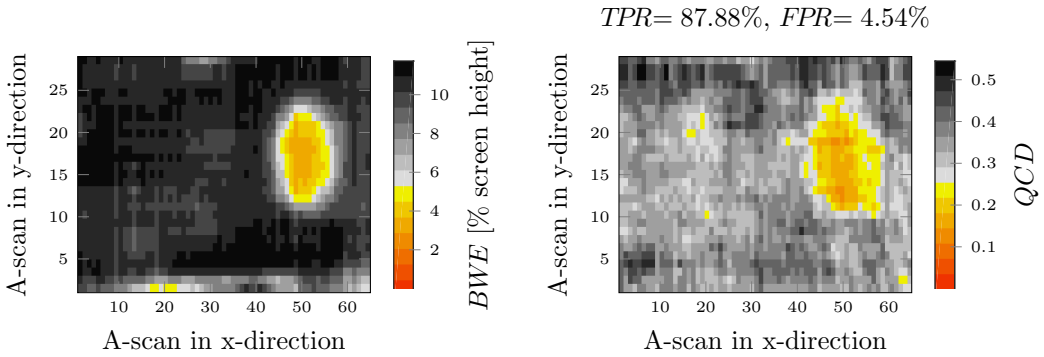


Figure 5.12: C-scans of M14; left: back-wall echo (porosity — BWE reduced by more than 50% — shown in colours); right: quartile coefficient of dispersion (QCD), decision threshold to achieve $TPR \approx TNR$ on M13 and M15-1 (predicted porosity shown in colours, predicted defect-free areas in grey shades)

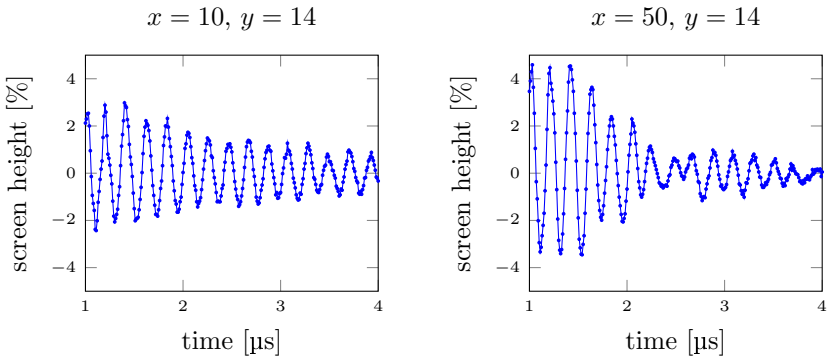


Figure 5.13: A-scan intermediate echo time series of M14, of non-porous area (left) and of porous area (right)

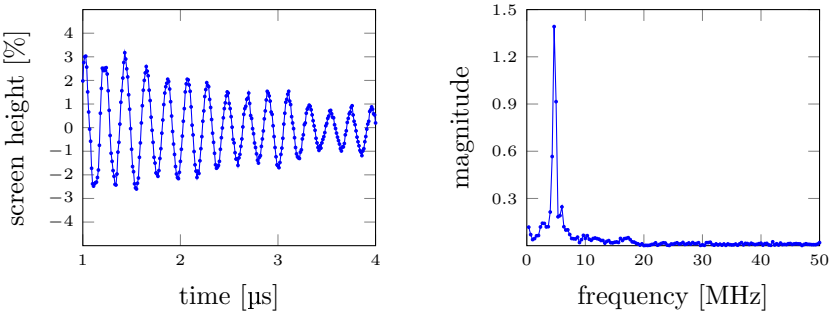


Figure 5.14: A-scan intermediate echo time series (left) and corresponding amplitude spectrum (right) of M15-1x41y12 (non-porous area)

5.3.3 Fourier analysis

The Fourier transform of the intermediate echo time series, or more specifically the amplitude spectrum (fig. 5.14) is considered in the following. $N/2 + 1$ values (with N being the length of the time series) are used, since for time series of real values the magnitudes of the amplitude spectrum are symmetric.

A first approach is to look at the peak frequency, the left and right cut-off frequency and related values (sect. 3.2). These features are, taken as a BWE-equivalent, very varying between specimen M13, M14 and M15-1 with values of balanced accuracies taking often not much more than 50%, thus near a random classifier (table C.8 and C.9).

Linear regression on amplitude spectrum

As proposed and first shown in Brandt, Hamann & Leuschner [2019], linear regression is tested on the amplitude spectrum. For specimens M13–M15, the whole frequency spectrum of 151 values is used (see however discussion in sect. 5.4.2). If a Hann window is applied to avoid the leakage effect (sect. 3.2), best results in frequency domain are achieved with $acc_{bal} = 83.4\%$ to $acc_{bal} = 94.8\%$ (fig. 5.15). Without applying a Hann window to the time series prior to Fourier transform, results are clearly worse with balanced accuracy in two of three scans even below 65% (table C.10). The effect of the Hann window is examined on material B in section 5.4.2.

5.3.4 Summary of evaluation on fabric specimens

The relatively simple feature quartile coefficient of dispersion (QCD) leads to best classification test results on fabric specimens with artificial porosity, achieving balanced accuracies $> 91\%$ for each of the three investigated specimens. Second best are RQA, Euclidean distance, with feature DET (picking up an amplitude effect similar to QCD) and linear regression of the amplitude spectrum, obtained with a Fourier transform after applying a Hann window on the intermediate echo time series. Both features lead to $acc_{bal} > 90\%$ for M13 and M15-1.

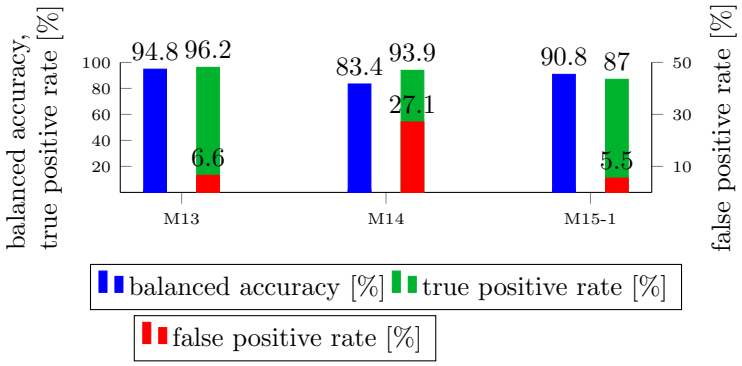


Figure 5.15: Classification and correlation results with linear regression on amplitude spectrum, Hann window applied on time series prior to Fourier transform, for fabric specimens. Training for each scan on two remaining scans, including decision threshold to achieve $TPR \approx TNR$

The other investigated features — other RQA features, including the ones obtained with angular distance; linear regression on time series or on the amplitude spectrum without applying a Hann window; or standard features obtained from the amplitude spectrum — are distinctly worse, not any balanced accuracy reaching 90%.

5.4 Pre-selection of parameters for evaluation on unidirectional specimens

Some parameters for the investigation on specimens of unidirectional material (with porosity created by alteration of production parameters outside the specified range) in section 5.5 and 5.6 are determined in advance, for recurrence quantification (sect. 5.4.1) and Fourier analysis (sect. 5.4.2). Reasons are computational cost and possible overfitting.

Since it is very difficult to create localised porous areas with an alteration of autoclave parameters, only specimens with a rather uniform quality (porosity level) are available. To ease the search for a back-wall echo equivalent, a scan of a good (low porosity) specimen and one of a porous specimen are combined into one 'virtual' volume scan. Such a virtual volume scan consists of 130 columns of data (each column containing 29 A-scans); 1–65 of a good specimen and 66–130 of a porous specimen. For this pre-selection of parameters, smaller data sets of physically different areas of the specimens are used: two scan sets (with two scans each) for specimens with and without resonance effect (for the chosen inspection frequency), respectively (table 5.1). A virtual scan is designated e.g.

5056-α2,

to refer to the second (virtual) volume scan of the first data set, i.e. the first area, of specimens 50 and 56 for this pre-selection of parameters.

A twofold cross-validation is deployed, taking each scan set once for training and once for test.

5.4.1 RQA parameters trend and correlation

For features *TND* and *COR* (definition 12 and 17), an additional parameter is $\widetilde{M} = M - \Delta_M$ (M length of embedded time series, $\Delta_M \geq 0$) to exclude Δ_M outer diagonal lines (furthest away from the line of identity) from computation. Due to small length of these outer lines a reliable determination of RR_k (recurrence rate per diagonal line) is presumably difficult — besides the fact that only a reduced amount of discrete values of RR_k is possible with such short lines, e.g. 0%, 50% and 100% for a line of length 2.

Four different choices of Δ_M are calculated for the whole RQA parameter range (cf. table 5.17) on one scan set, and the RQA parameters with the best *training* results (the highest balanced accuracy with $TPR \approx TNR$) are used for *test* on the two scans of the other scan set. Test results are compared for all four choices of Δ_M.

Choices for Δ_M are (table 5.6)

- 50, following the recommendation to use a value in the order of magnitude of five times the first zero crossing of the autocorrelation [Marwan & Webber 2015]; this zero crossing (or the first minimum of the autocorrelation) lies for the materials in this work (determined at a few A-scans) between 4 and 7;
- 10, following the recommendation this being sufficient for noise [Marwan & Webber 2015];

- 25 as intermediate step between the two former values;
- 2, considered as the lowest sensible choice (shortest considered diagonal line of RP has thus length 3, i.e. four different values of RR_k are possible).

Table 5.6: Values for $\Delta_M = M - \widetilde{M}$ used in pre-selection of this parameter for RQA features *TND* and COR

$\Delta_M =$	2	10	25	50

Unidirectional specimens, resonant case

Balanced accuracy for test results on scan set 5056-α and -β, respectively (with parameters leading to optimum training results on the other set, respectively) decreases for most scans with a greater Δ_M both for Euclidean (fig. 5.16) and angular distance (fig. 5.17).

Unidirectional specimens, non-resonant case

Classification results of *TND* and COR are in general worse on specimens 2438 (fig. 5.18 and 5.19), with 2438-β1 having especially bad results. Balanced accuracies depend less on Δ_M than for 5056, the general tendency is however similar.

Overall, in this application and the rather small time series, it is generally negative for classification purposes to have greater Δ_M, i.e. to exclude a larger number of outer diagonals of the RPs from calculation of *TND* and the newly introduced feature COR. For main investigations (sect. 5.5.1 and sect. 5.6.1) $\Delta_m = M - \widetilde{M} = 2$ is used.

5.4.2 Window size for linear regression on amplitude spectrum

For the application of linear regression on the amplitude spectrum on unidirectional specimens of material B (with porosity introduced by alteration of the production parameters outside the specified range) the following is considered: with sampling frequency of 100 MHz frequency components up to 50 MHz can be measured. Ultrasound from a probe with inspection frequency of 5–6 MHz has hardly any frequency components other than noise above presumably 20 MHz. Thus, higher frequencies may contribute to an overfitting (cf. sect. 3.4.2) of linear regression. Such effects can only be detected with test (and not training) results; this shall be done prior to the cross-validation results in the following sections. Furthermore, it is computational more efficient if a smaller number of amplitudes suffices to achieve satisfactory results.

Unidirectional specimens, resonance case

Firstly portions of the amplitude spectrum starting at 0 MHz and ending between 0 MHz and 50 MHz are considered. Linear regression weights are computed on both

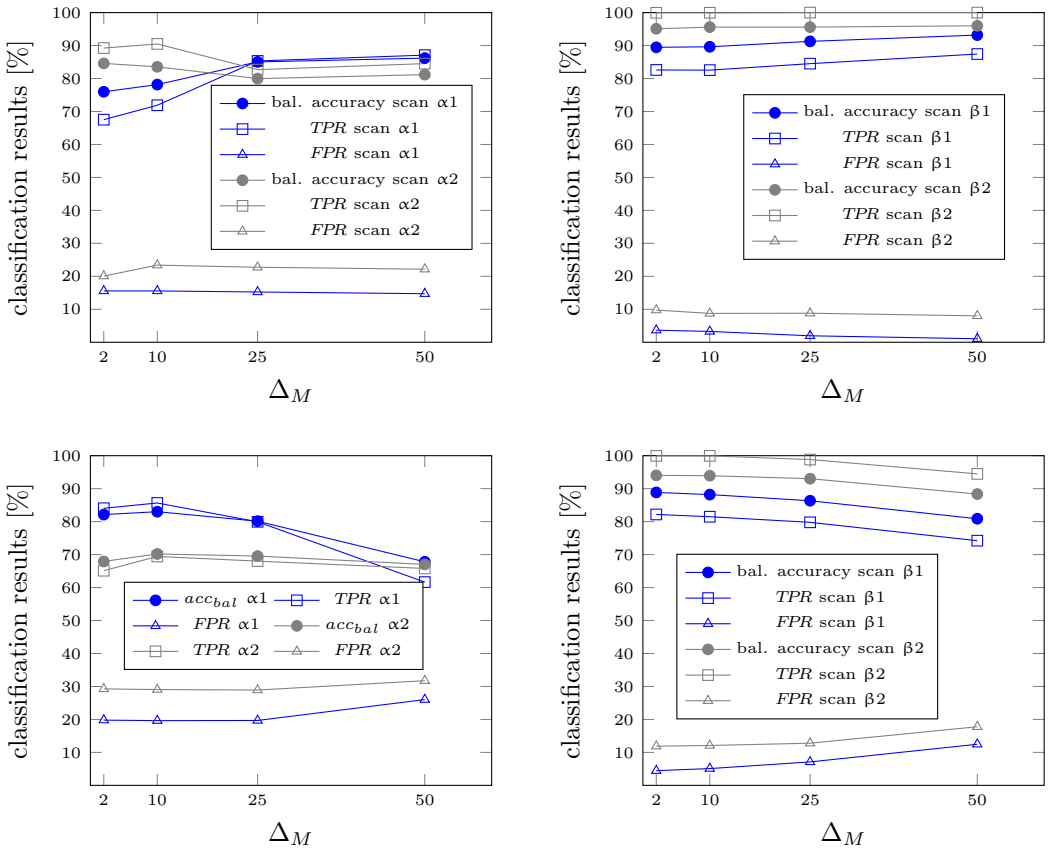


Figure 5.16: Classification test results for 5056-α/β for RQA-feature *TND* (top) and *COR* (bottom), Euclidean distance, in dependence of Δ_M; training for α on β (left) and vice versa (right)

scans of the training scan set, and the BWE-equivalent is calculated with these weights on the same scan set — training — as well as on the two scans of the other scan set (test).

Maximum balanced accuracies for training are achieved for linear regression on windows of the amplitude spectrum of 0 MHz to roughly 10 MHz and greater, taking values of almost 100% (fig. 5.20). Test results do for one scan (5056-β2) even exceed training results, but for the three other scans balanced accuracies are lower. Moreover, balanced accuracies do decrease (presumably by overfitting) if the window ends above ≈ 8 MHz for these three scans (with the exception of 5056-α1, where it increases again at 18 MHz).

To complete the picture, linear regression on all possible combinations of connected windows is calculated, i.e. on (0; 0, 0.33; 0, 0.33, 0.67; 0, 0.33, 0.67, 1; ...; 0.33; 0.33, 0.67; 0.33, 0.67, 1; ...; 49.67, 50; 50) MHz. Again, the weights — and decision threshold to reach *TPR* ≈ *TNR* — are calculated on both scans of the training scan set and are applied (on the same window) for each of the two scans of the test scan

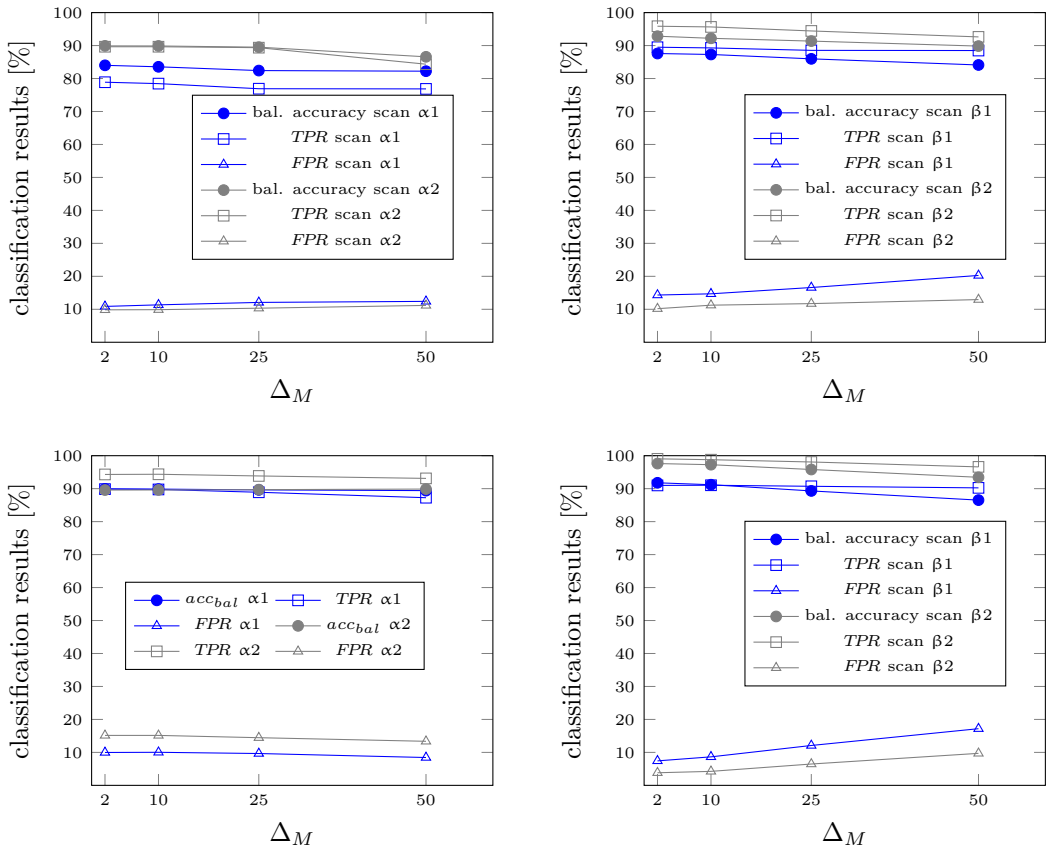


Figure 5.17: Classification test results for 5056-α/β for RQA-feature *TND* (top) and *COR* (bottom), angular distance, in dependence of Δ_M, training for α on β (left) and vice versa (right)

set. The window with optimum classification results is determined for each of the four test scans. Results are excellent with balanced accuracies $\geq 95\%$ (table 5.7) and the optimum windows vary between the four scans.

The goal is however to have a common window of the amplitude spectrum for specimens 5056. Therefore the optimum window for all scans together (5056-α/β1/2) are calculated (using again test results, i.e. weights and decision thresholds from 5056-β for 5056-α1/2 and vice versa). The test results of each scan with this one optimum window (table 5.8) are almost as good as for the individual optimum windows, balanced accuracies maximal 2.1% lower.

When applying a Hann window on the intermediate echo time series prior to Fourier transform, the dependence of test results on the length of the window of the amplitude spectrum (fig. 5.21) is different to FFT without applying a Hann window: test results are in general better, and the progression of test results resemble those of training results — a drop in test results for larger window sizes, taking larger frequency values into account, does not appear. The larger frequencies of the amplitude spectrum

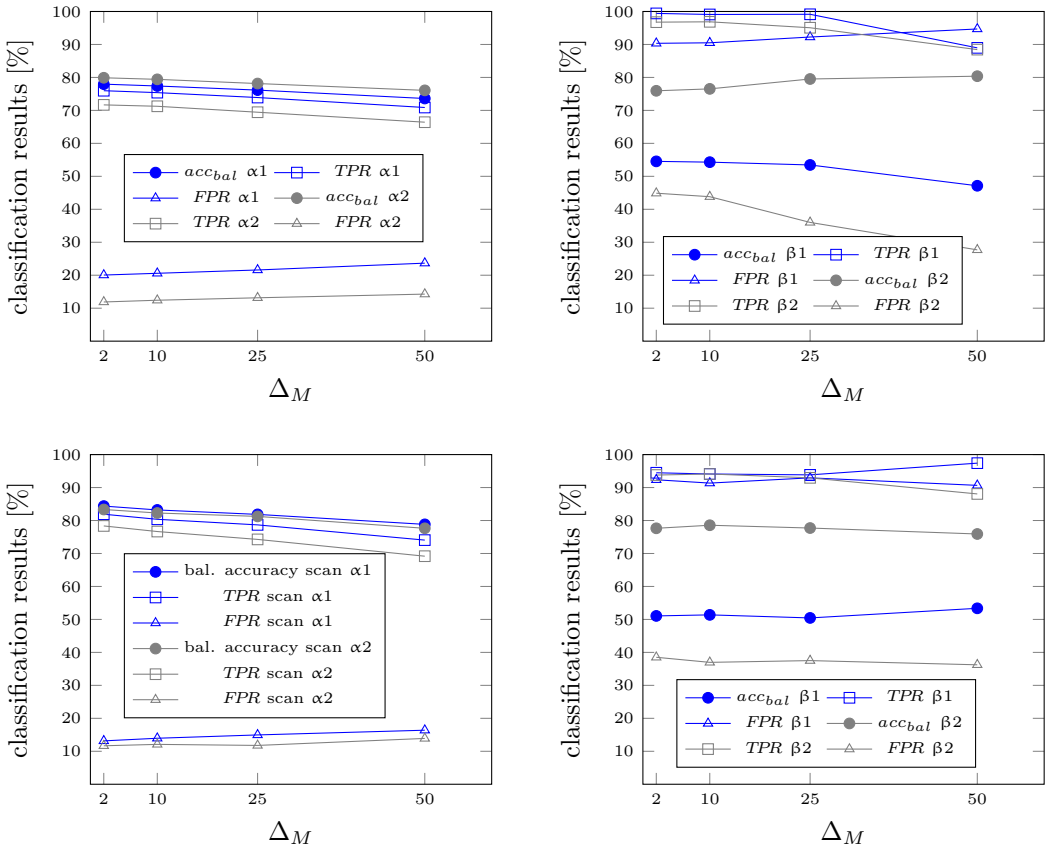


Figure 5.18: Classification test results for 2438-α/β for RQA-feature *TND* (top) and *COR* (bottom), Euclidean distance, in dependence of Δ_M; training for α on β (left) and vice versa (right)

Table 5.7: Windows of amplitude spectrum leading to best balanced accuracy acc_{bal} for test on 5056-α/β

test scan	start of window	end of window	acc_{bal}	*TPR*	*FPR*
5056-α1	2.33 MHz	7 MHz	96.7%	98.6%	5.2%
5056-α2	0 MHz	7.33 MHz	95.0%	94.3%	4.4%
5056-β1	2.33 MHz	7 MHz	96.4%	92.8%	0%
5056-β2	1.33 MHz	5.67 MHz	100%	100%	0%

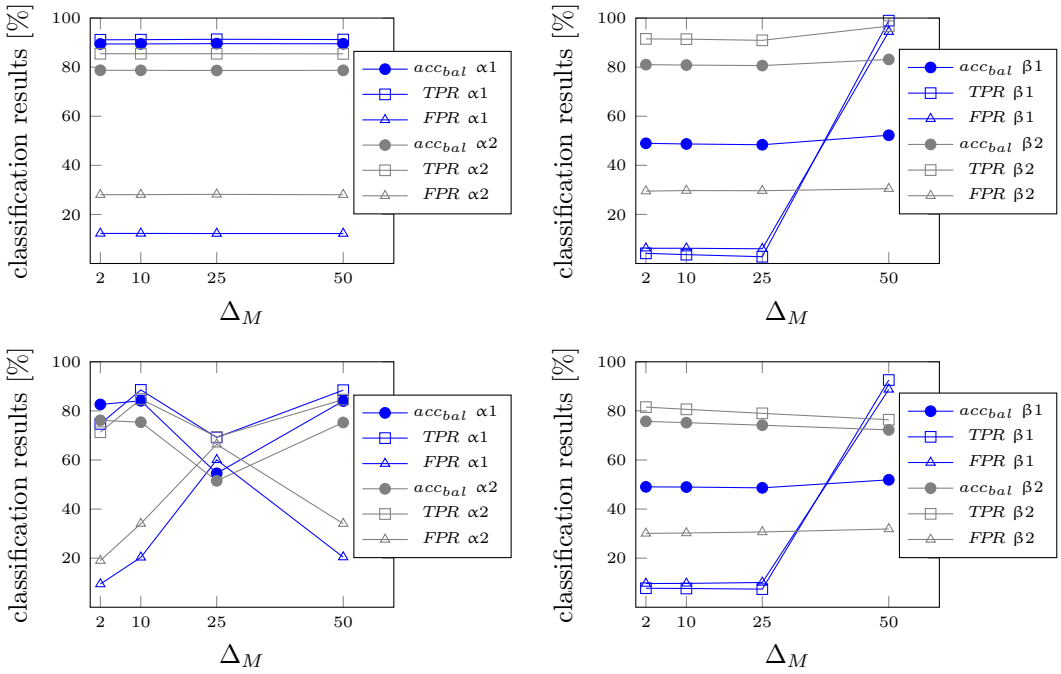


Figure 5.19: Classification test results for 2438α/β for RQA-feature *TND* (top) and *COR* (bottom), angular distance, in dependence of Δ_M; training for α on β (left) and vice versa (right)

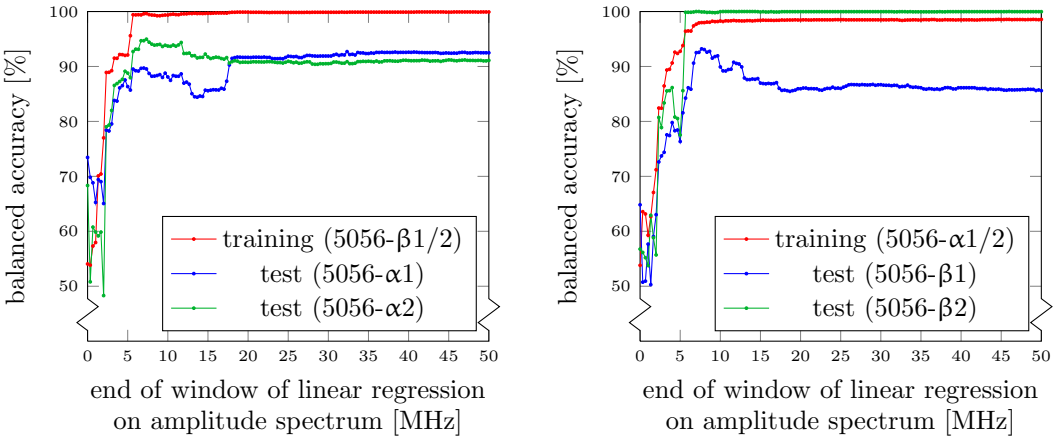


Figure 5.20: Balanced accuracy for 5056 in dependence of end of window of amplitudes on which linear regression has been performed; start is at 0 MHz. Training on two scans combined; test results on each single scan

Table 5.8: Classification test results for each scan of 5056-α/β with linear regression on the fixed window of the amplitude spectrum that leads to best classification results with all scans combined

window	test scan	acc_{bal}	*TPR*	*FPR*
0.66 MHz–7 MHz	5056-α1	96%	94.5%	2.6%
	5056-α2	94.9%	95.9%	6%
	5056-β1	94.3%	88.6%	0%
	5056-β2	99.97%	100%	0.05%

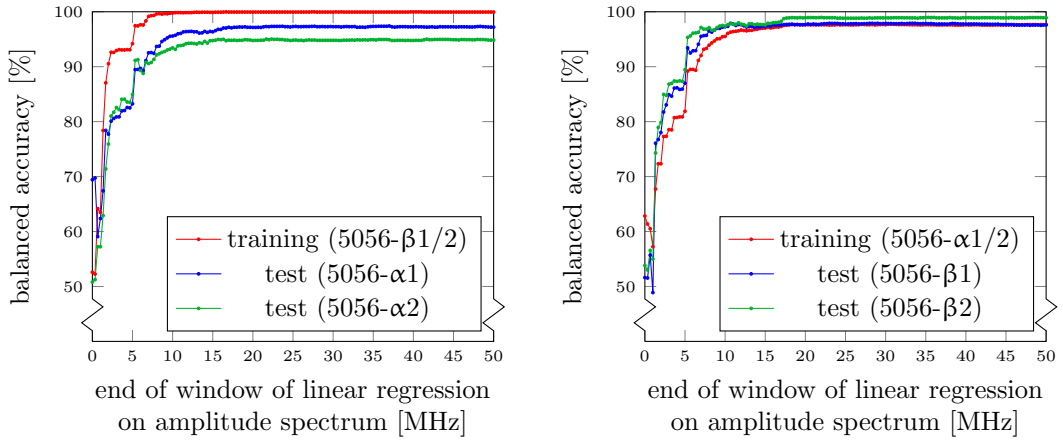


Figure 5.21: Balanced accuracy for 5056 in dependence of end of window of amplitudes on which linear regression has been performed; start is at 0 MHz. Training on two scans combined; test results on each single scan. Fourier transform after applying a Hann window on the intermediate echo time series

without using a Hann window may be 'contaminated' through the leakage effect such that overfitting occurs, which the Hann window prevents. The window on the amplitude spectrum leading to the best common classification results over all test scans (analogous to table 5.8 for data without using a Hann window) starts at 0 MHz and is larger. It end at 23 MHz and produces generally slightly better results (table 5.9) than FFT without applying a Hann window on the time series. The individual windows on each scan to achieve best classification results differ in their sizes, but for three of four scans balanced accuracy is better only in tenth percent range (table 5.10). The Hann window reduces amplitudes of the time series greatly, as can be noticed especially regarding the large amplitudes in the first approximately 1 µs of the intermediate echo time series (fig. 5.22 top, right compared to left). The corresponding amplitude spectra are accordingly reduced (fig. 5.22 bottom), and the peak frequency shifts from 6 MHz to 5.33 MHz. The weights determined by linear regression on an amplitude spectrum after applying a Hann window (fig. 5.23 left) greatly reduce frequencies between approximately 1 MHz and 7–8 MHz. Multiplying weights with amplitudes of

Table 5.9: Classification test results for each scan of 5056-α/β with linear regression on the fixed window of the amplitude spectrum that leads to best classification results with all scans combined; Fourier transform after applying a Hann window on the intermediate echo time series

window	test scan	acc_{bal}	*TPR*	*FPR*
0 MHz–23 MHz	5056-α1	97.3%	94.9%	0.3%
	5056-α2	95%	92%	2%
	5056-β1	97.6%	96.8%	1.5%
	5056-β2	99.1%	99.6%	1.4%

Table 5.10: Windows of amplitude spectrum leading to best balanced accuracy acc_{bal} for test on 5056-α/β; Fourier transform after applying a Hann window on the intermediate echo time series

test scan	start of window	end of window	acc_{bal}	*TPR*	*FPR*
5056-α1	0 MHz	37 MHz	97.4%	95.1%	0.3%
5056-α2	2 MHz	22.33 MHz	96.1%	96%	3.8%
5056-β1	0 MHz	23 MHz	97.9%	98%	2.2%
5056-β2	1 MHz	39.67 MHz	99.2%	99.95%	2.5%

a measurement at non-porosity (fig. 5.23 right), the importance of the frequencies are approximately equal over the whole window — with two exceptions, below 1 MHz and at 5.33 MHz (the peak of the amplitude spectrum after applying a Hann window). Presumably frequencies above the main ones (> 8 MHz) contain the effect of porosity in a good way once leakage is reduced via the Hann window.

The amplitude spectra *without* applying a Hann window of a correctly predicted non-porosity (fig. 5.24 top left) and of a porous area falsely predicted as non-porous (fig. 5.24 top right) are qualitatively the same. That is probably due to the large influence of leakage, which worsens results for windows ending beyond 7 MHz (fig. 5.20) if no Hann window is applied. It is difficult to judge from amplitude spectra *with* applying a Hann window on the intermediate echo time series (fig. 5.24 bottom) what the difference between non-porosity (bottom left) and porous area (bottom mid and right) exactly is. However, there are evidently differences sufficient for correct classification when applying linear regression.

Unidirectional specimens, non-resonance case

Considering the dependence of classification results for specimens 2438 — non-resonant case with inspection frequency 5 MHz (fig. 5.25) — bad test results are striking, whereas training balanced accuracies are still above 90%. For these specimens without using a Hann window apparently an overfitting takes place.

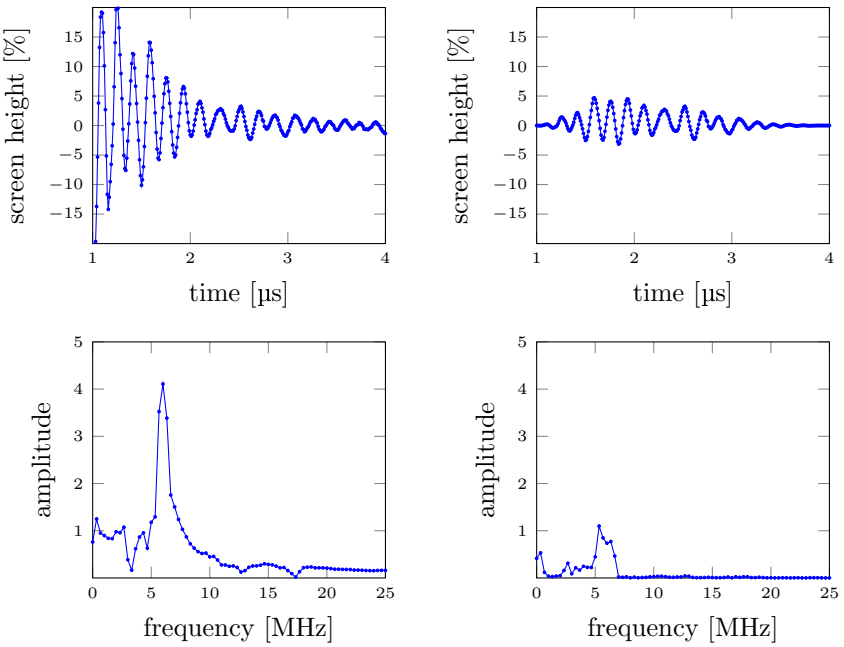


Figure 5.22: Intermediate echo time series of non-porous area, specimen 50 (5056-β1x21y20), without (top left) and with applied Hann window (top right), and corresponding amplitude spectra

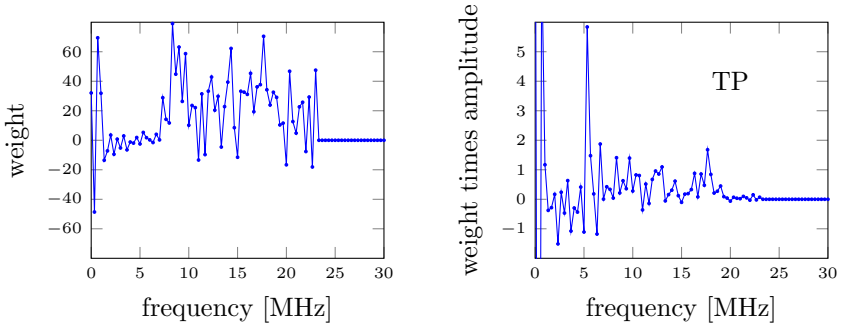


Figure 5.23: Weights for linear regression on amplitude spectrum, Fourier after applying a Hann window on intermediate echo time series, determined on 5056-α (left), weights times amplitude spectrum of correctly predicted non-porous area (5056-β1x21y20, right)

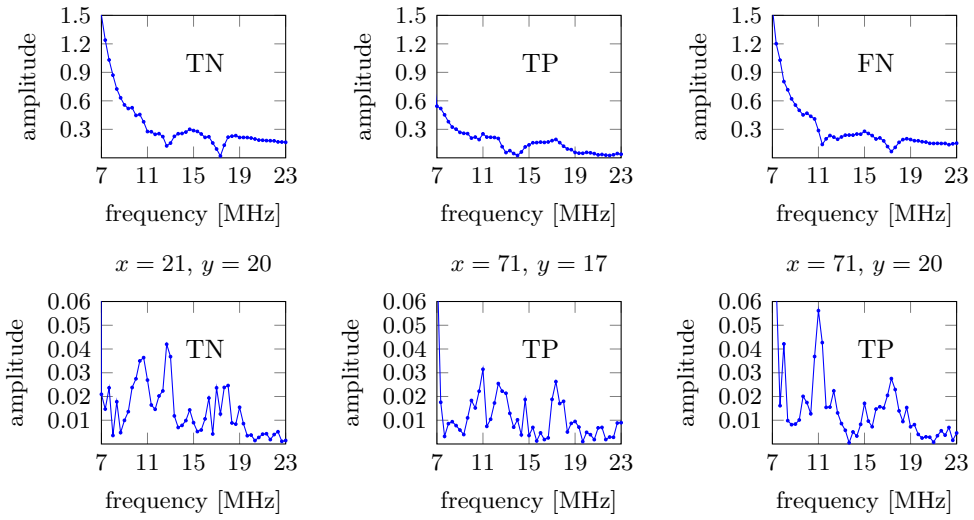


Figure 5.24: Amplitude spectra from 7 MHz to 23 MHz without (top) and with (bottom) applying a Hann window on intermediate echo time series prior to Fourier transform of 5056-β1; using linear regression on amplitude spectrum: correctly predicted non-porous area (left); correctly predicted porous area (middle); falsely as non-porous predicted area without using a Hann window (top right), but correctly as porosity predicted with using a Hann window (bottom right)

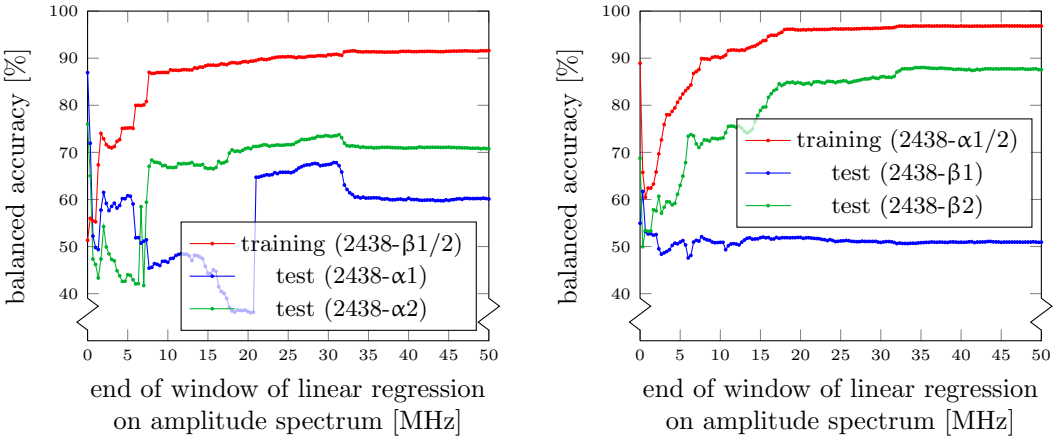


Figure 5.25: Balanced accuracy for 2438 in dependence of end of window of amplitudes on which linear regression has been performed; start is at 0 MHz. Training on two scans combined; test results on each single scan

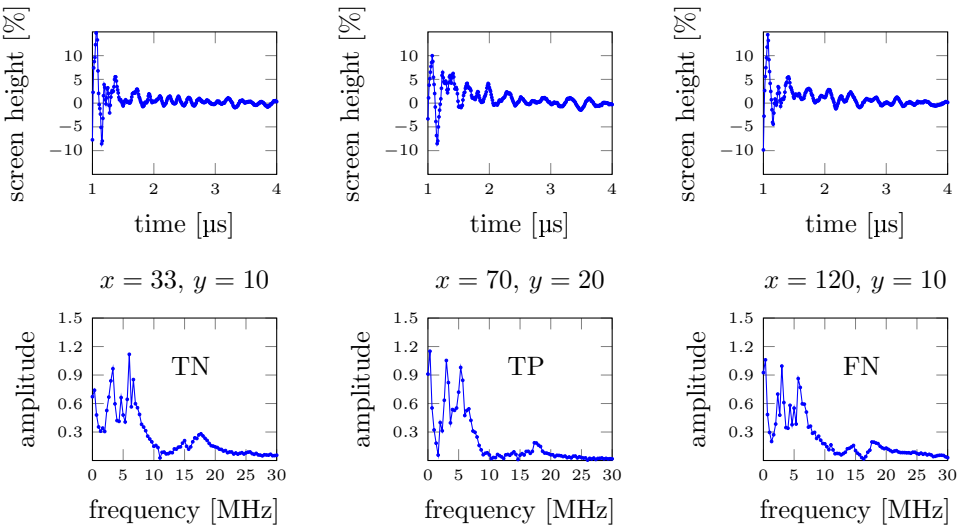


Figure 5.26: Intermediate echo time series (top) and the according amplitude spectra (bottom) without applying a Hann window prior to Fourier transform of 2438-α2; with linear regression on amplitude spectrum: correctly predicted non-porous area (left); correctly predicted porous area (middle); falsely as non-porous predicted area (right)

The intermediate echoes from the A-scans (fig. 5.26 top) are more irregular than their counterparts from measurements with resonance effect. The corresponding amplitude spectra (fig. 5.26 bottom) are accordingly different from the ones for intermediate echo time series obtained with resonance effect, they exhibit rather two peak areas around 3 MHz and 5–6 MHz.

Linear regression for measurements without resonance ignores or largely reduces the influence of frequencies between ≈ 2 MHz and ≈ 7 MHz (fig. 5.27). The emphasis of the area around 8 MHz (next to the peak frequency 7 MHz) and of ≈ 18 MHz evidently leads to bad classification results.

When determining the window of all possible connected values of the amplitude spectrum that leads to optimum test results for each scan separately, balanced accuracies are between 73.6% to 88% (table 5.11). The window sizes largely vary; consequently, with a fixed window determined on all scan sets, balanced accuracies only from 57.7% to 83.4% are achieved (table 5.12). Linear regression on amplitude values of Fourier transform without a Hann window will thus not be considered for the non-resonance case in the main part (sect. 5.6.3).

Applying a Hann window on the intermediate echo time series before Fourier transform distinctly improves classification results (fig. 5.28), only 2438-β1 appears to be an outlier.

The difference between weights for amplitudes of frequencies from 6 MHz–18 MHz and those for ≈ 1–6 MHz is larger if a Hann window is applied (fig. 5.30 left compared to fig. 5.27 left, taking different scales into account). Once leakage is reduced, amplitudes

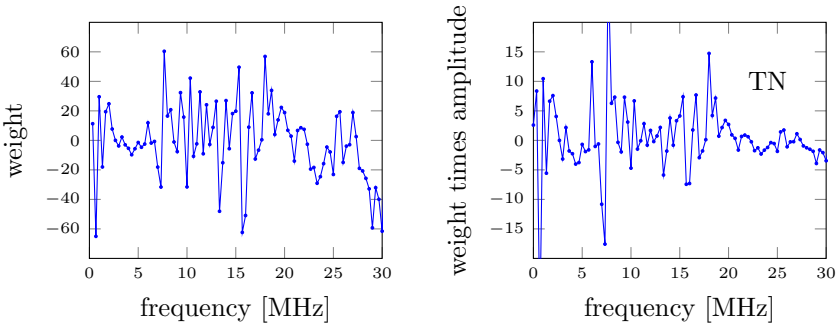


Figure 5.27: Weights for linear regression on amplitude spectrum determined on 2438-β (left), weights times amplitudes of correctly identified non-porous area (2438-α2x33y10, right)

Table 5.11: Windows of amplitude spectrum leading to best balanced accuracy acc_{bal} for test on 2438-α/β

test scan	start of window	end of window	acc_{bal}	*TPR*	*FPR*
2438-α1	0 MHz	0 MHz	86.9%	97.2%	23.3%
2438-α2	3.67 MHz	31.33 MHz	79.9%	78.3%	18.4%
2438-β1	6 MHz	6.33 MHz	73.6%	85.5%	38.3%
2438-β2	0 MHz	35.33 MHz	88%	93.7%	17.6%

of frequencies greater 6 MHz apparently transport information about porosity in a better way than amplitudes of smaller frequencies. As for data without resonance effect, it is however difficult to assess from the amplitude spectra (fig. 5.29) which the direct effect is.

The individually — per scan — windows on the amplitude spectrum for linear regression with Hann window that lead to best classification differ between 2438-β1 and the other three scans (table 5.13). The window of the amplitude spectrum leading to greatest balanced accuracies on all scans of all scan sets together for linear regression after applying a Hann window goes from 1 MHz to 50 MHz (table 5.14). For the final choice of window it is considered that:

- It is still doubted from a physical point of view that high frequencies contribute to the classification; the reduced data set may not reflect reality here. Furthermore, it is computationally advantageous to use a smaller window size.
- The individual window sizes (table 5.13) start for all but 2438-β1 at 0 MHz.

Two different window sizes are thus chosen for the main evaluation in section 5.6.3, from 0 MHz to 30 MHz and from 1 MHz to 30 MHz. The window 1 MHz–30 MHz gives only slightly worse results (compared to regression over 1–50 MHz, table 5.14) for 2438-α1/2 and -β1 and even better results for 2438-β2 (table 5.15 bottom). The

Table 5.12: Classification test results for each scan of 2438-α/β with linear regression on the fixed window of the amplitude spectrum that leads to best classification results with all scans combined

window	test scan	acc_{bal}	*TPR*	*FPR*
3.67 MHz–27.33 MHz	2438-α1	83.4%	83.4%	16.55%
	2438-α2	79.8%	78.5%	18.8%
	2438-β1	57.7%	41.4%	26.1%
	2438-β2	74%	91.7%	43.8%

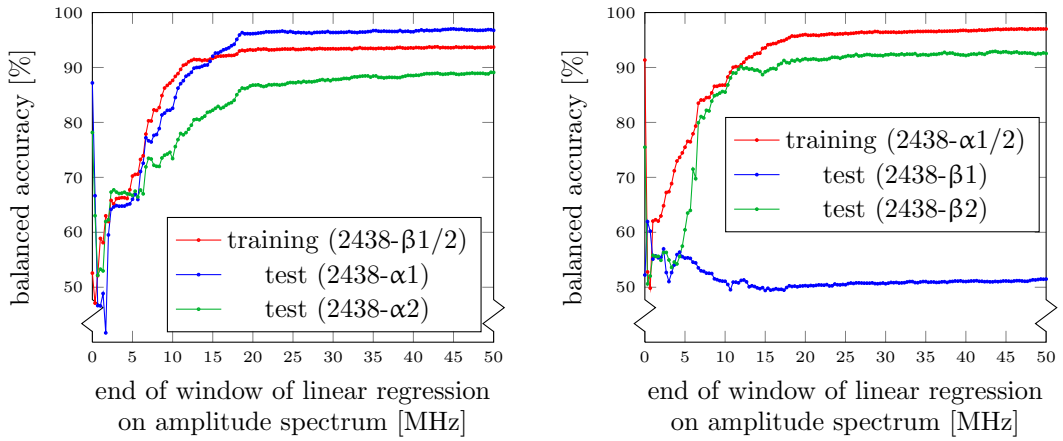


Figure 5.28
Balanced accuracy for 2438 in dependence of end of window of amplitudes on which linear regression has been performed; start is at 0 MHz. Training on two scans combined; test results on each single scan. Fourier transform after applying a Hann window on the intermediate echo time series

Table 5.13: Windows of amplitude spectrum leading to best balanced accuracy acc_{bal} for test on 2438-α/β; Fourier transform after applying a Hann window on the intermediate echo time series

test scan	start of window	end of window	acc_{bal}	*TPR*	*FPR*
2438-α1	0 MHz	45 MHz	97%	98.4%	4.3%
2438-α2	0 MHz	49.67 MHz	89.1%	91.5%	13.2%
2438-β1	3 MHz	35 MHz	78.9%	74.6%	16.8%
2438-β2	0 MHz	43.67 MHz	92.9%	94.6%	8.8%

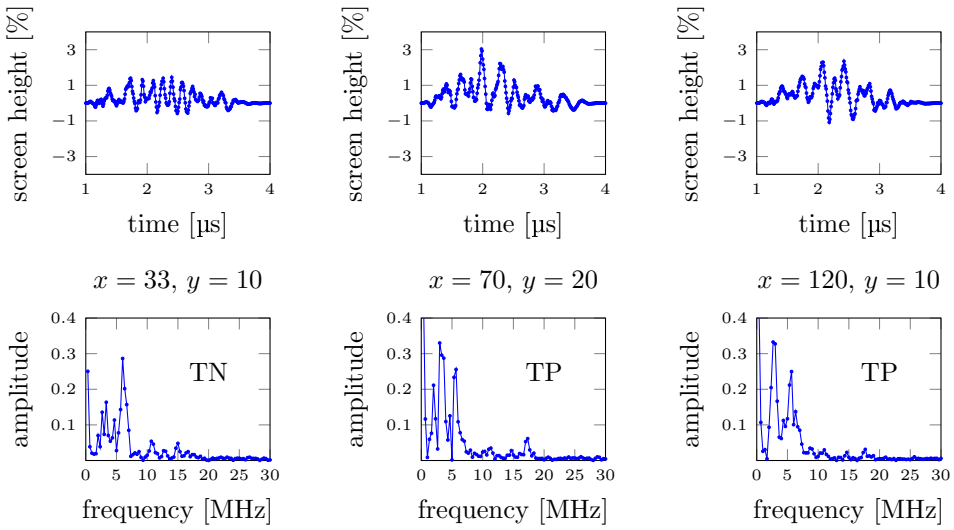


Figure 5.29: Intermediate echo time series (top) and the according amplitude spectra (bottom), applying a Hann window prior to Fourier transform of 2438-α2; with linear regression on amplitude spectrum: correctly predicted non-porous area (left); correctly predicted porous area (middle); correctly predicted porous area (right)

window 0 MHz–30 MHz improves results for 2438-α1/2 and -β2, but reduces 2438-β1 to extremely small 50.6% (table 5.15 top).

The final decision for choice of the windows of the amplitude spectrum for linear regression for material B both without and with resonance effect is summarised in table 5.16.

Table 5.14: Classification test results for each scan of 2438-α/β with linear regression on the fixed window of the amplitude spectrum that leads to best classification results with all scans combined; Fourier transform after applying a Hann window on the intermediate echo time series

window	test scan	acc_{bal}	*TPR*	*FPR*
1 MHz–50 MHz	2438-α1	96.4%	96.4%	3.6%
	2438-α2	87.9%	87.6%	11.9%
	2438-β1	77.6%	74.6%	19.5%
	2438-β2	86.6%	95.5%	22.3%

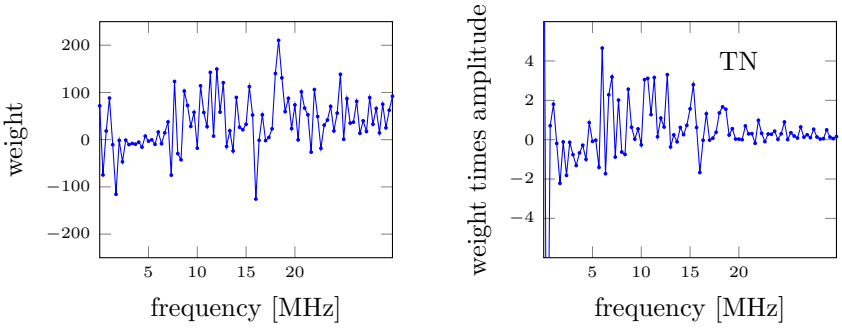


Figure 5.30: Weights for linear regression on amplitude spectrum determined on 2438-β (left) with appyling a Hann window prior to Fourier transform, weights times amplitudes of correctly predicted non-porous area (2438-α2x33y10, right)

Table 5.15: Classification test results for each scan of 2438-α/β with linear regression with two choices of fixed window of the amplitude spectrum

window	test scan	acc_{bal}	*TPR*	*FPR*
0 MHz–30 MHz	2438-α1	96.4%	97.9%	5.2%
	2438-α2	87.8%	90.2%	14.6%
	2438-β1	50.6%	3.8%	2.5%
	2438-β2	92.3%	94.1%	9.5%
1 MHz–30 MHz	2438-α1	95.8%	95.9%	4.2%
	2438-α2	86.2%	85.4%	13.1%
	2438-β1	74.6%	59.5%	10.4%
	2438-β2	89%	89.9%	11.9%

Table 5.16: Windows of amplitude spectrum used for linear regression in section 5.5.3 and 5.6.3

material	preparation of intermediate echo time series	FFT window
material B, resonance case	none	0.66 MHz–7 MHz
	Hann window	0 MHz–23 MHz
material B, non-resonance case	Hann window	0 MHz–30 MHz
		1 MHz–30 MHz

5.5 Evaluation on unidirectional specimens, resonance case

This section investigates the classification of porosity in CFRP out of the intermediate echo time series of ultrasonic testing in unidirectional material B — more relevant to current aerospace industry than the fabric specimens investigated in section 5.3 — with a ply thickness of 0.25 mm, which leads to a resonance effect with the inspection frequency of 5 MHz (sect. 2.2.3). Recurrence quantification analysis (sect. 5.5.1), features generated in time domain (sect. 5.5.2) and Fourier analysis (sect. 5.5.3) are studied for two specimens of 6 mm thickness, 50 and 56 (table 5.2). The best method is then applied on reduced intermediate echo time series on these specimens to eventually test it on specimen 49 and 55 of 4 mm thickness (sect. 5.5.5).

Four or two, respectively, different areas of each specimen have been measured three times each (cf. table 5.2); the three volume scans of one area are referred to as scan set.

The ('virtual') volume scans combining one scan from specimen 50 and one from specimen 56 are designated e.g.

5056-C2,

to refer to the second (virtual) volume scan of the third data set, i.e. the third area of specimens 50 and 56.

The process of training and test in section 5.5.1 to 5.5.3 is similar to the one for classification for the fabric specimens, cf. section 5.2 and figure 3.4:

- Three out of four scan sets are used for training, the remaining one is used for test.
- This is repeated four times, so that each scan set is test set once.

Here, for material B, with three volume scans per scan set

- all (volume) scans of three scan sets (that is nine volume scans) are used commonly to find the optimum parameters (training);
- results on the test scan set are shown for its three volume scans separately.

The A-scans of specimens of material B have a different appearance than those for the fabric specimens (cf. fig. 5.31 with fig. 2.6). Amplification has been adjusted to be approximately four times as high as for the fabric material (+12.1 dB). The back-wall echo is however just twice as high; specimens 50, 56, 49 and 55 have a rougher surface (after pulling of peel-ply) that increase reflections and thus reduces the transmission of ultrasound into the specimen. This amplification was chosen to obtain intermediate echoes high enough (considering digitization) especially in the last part of the A-scan, cf. fig. 5.31; a surface echo signal cut off with signal $\geq 100\%$ was accepted, since the surface echo is not used for the investigations.

As for classification into porous and non-porous for specimens in section 5.3, a standard intermediate echo gate from 1 µs to 4 µs applies in section 5.5.1 to 5.5.3.

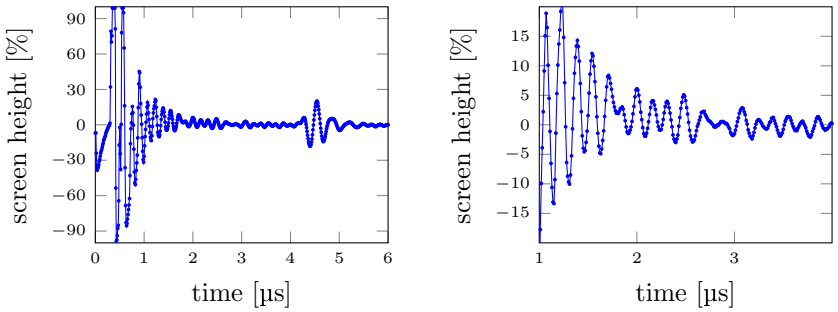


Figure 5.31: A-scan time series of material B specimen 50 (non-porous), 50-A1x5y5 (left); magnified intermediate echo time series (of the same A-scan) (right)

Table 5.17: RQA parameters for optimum classification of unidirectional specimens with natural porosity

RQA parameter	Min	Max	Steps
Time delay τ	1	10	1
Embedding dimension d	1 (2 for ang. distance)	10	1
Recurrence threshold ϵ for Euclidean distance	0.1	20	0.1 for $\epsilon \in [0.1, 2]$; 0.5 for $\epsilon \in [2.5, 20]$
Recurrence threshold ϵ for angular distance	0.025π	0.9π	0.025π for $\epsilon \in [0.025\pi, 0.05\pi]$; 0.5π for $\epsilon \in [0.55\pi, 0.9\pi]$
Minimum line length l_{min}	2	10	1

5.5.1 Recurrence quantification analysis

The optimum RQA parameter set (including decision threshold) is again determined via greatest balanced accuracy, trying to achieve a value of at least 90% with parity of *TPR* and *TNR*. For the remainder of this text the evaluation is expanded to all RQA features described in section 3.3.5. Constraints in terms of the maximum of *RR* and minimum range of *RR* and *DET* of table 5.5 in section 5.3.1 are applied as well. For all further RQA features no constraints in terms of minimum range are used in calculation; their feasibility as BWE-equivalent including a sufficiently large spread of their values will be considered during the following evaluations if relevant.

Range and steps of parameter changes are the same as for classification computations in section 5.3.1, only threshold ϵ for Euclidean distance has been enhanced to 20 (adapted to the larger signals), table 5.17.

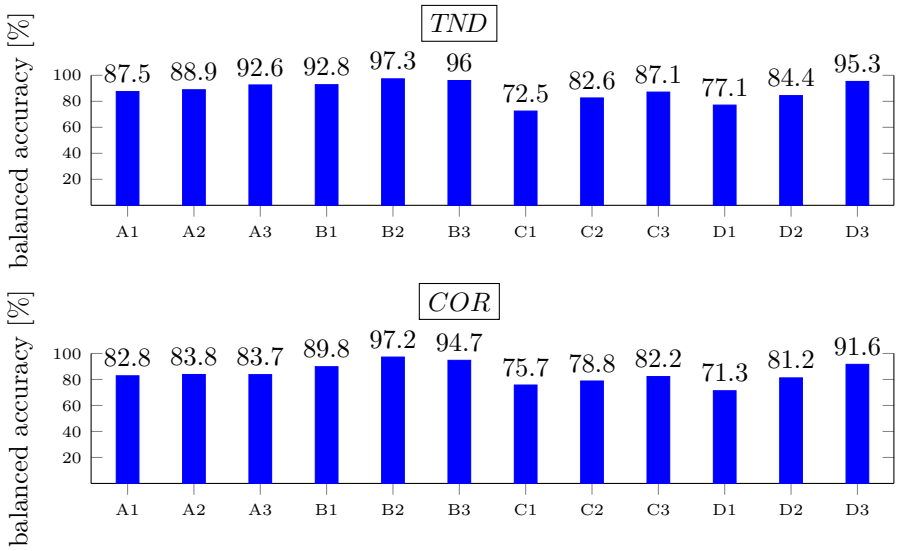


Figure 5.32: Classification test results (balanced accuracy acc_{bal}) for recurrence quantification features *TND* and *COR* as BWE-equivalent, using Euclidean distance, on scan sets 5056-A–D; RQA parameters (and decision threshold) leading to best classification results on three remaining scan sets, i.e. for A on B–D and so on. For *TND* and *COR*, $\tilde{M} = M - 2$

Euclidean distance

Using features from RQA as BWE-equivalent with Euclidean distance delivers unsatisfying classification test results. Features *TND* and *COR* lead to balanced accuracies below 80% for two or three, respectively, out of all 12 scans (of four scan sets) (fig. 5.32). *RR* and *DIV* are even worse with eight and six, respectively, balanced accuracies $< 80\%$, out of which one equals 65.2% and 58.3%, respectively (table C.28 and C.29).

Classification results for L_{nor} and *ENT*, Euclidean distance, are similarly bad than the ones for *TND* and *COR* (fig. 5.33). *RATIO* is worse, comparable to *RR* (but with worse results for scan set 5056-B, table C.33).

DET is the only feature with merely one out of 12 scans having balanced accuracies below 80% (fig. 5.33). Values of $> 90\%$, the goal that has been set in this work, are achieved for five scans. High values of ϵ lead to these positive results (cf. table C.32) via an amplitude effect: RPs show complete recurrence for time points beyond 100 ($> 2\,\mu s$ in original A-scan), whereas in approximately 1–2 µs signals are uniformly damped in non-porous areas compared to more complex behaviour with overall smaller signals for measurements of porous regions (fig. 5.34 left compared to right). *DET* increases here for porosity. Interestingly, the optimum results are achieved with $\tau = 7 \triangleq 70\,\text{ns} \approx \lambda/2$, creating diagonal lines perpendicular to the LOI.

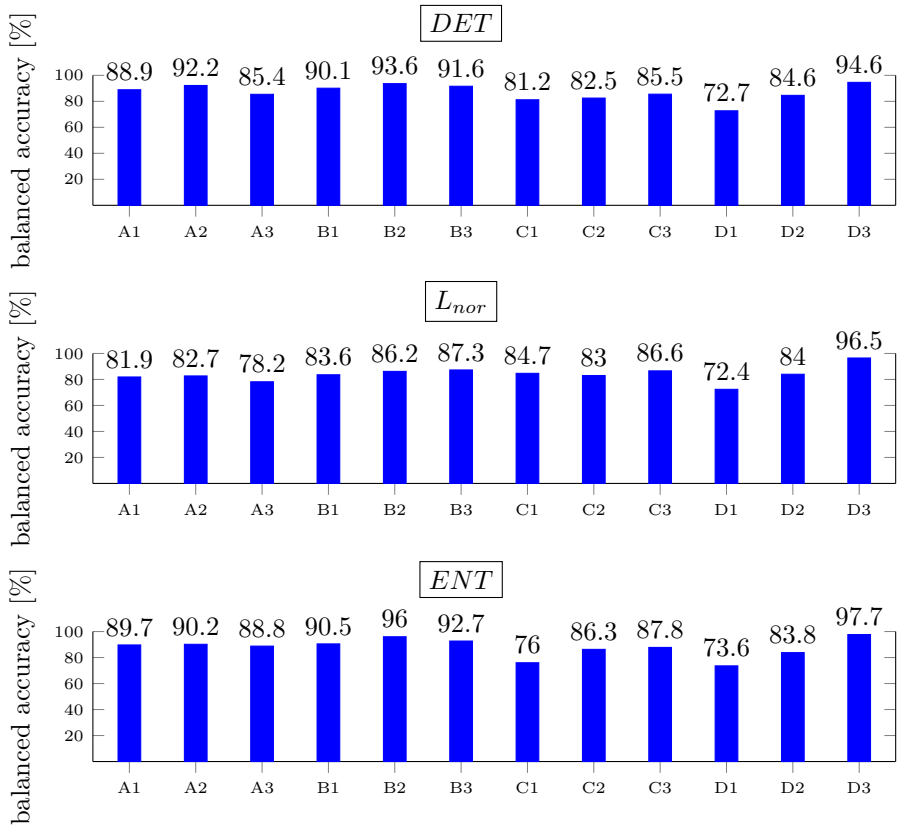


Figure 5.33: Classification test results (balanced accuracy acc_{bal}) for recurrence quantification features DET, L_{nor} and ENT as BWE-equivalent, using Euclidean distance, on scan sets 5056-A–D; RQA parameters (and decision threshold) leading to best classification results on three remaining scan sets, i.e. for A on B–D and so on

Angular distance

RQA feature DIV as BWE-equivalent, generated using angular distance, leads to bad classification results with balanced accuracies below 80% for most of the scans (table C.37). RR is better, but still contains one scan (C1) with $acc_{bal} = 70.5\%$. TND and COR deliver similar results (fig. 5.35) with only one of 12 scans below 80% balanced accuracy; the newly in this work proposed feature COR may be considered slightly better: the worst results, for C1, and D2 and D3, are with COR better than with TND.

Recurrence plots for COR, using RQA parameters leading to optimum classification results, are rather sparse (fig. 5.36). Signals of the intermediate echo time series, measured at non-porous specimen 50, in the first approximately 1 µs are periodic, leading to diagonal lines in the RP, whereas recurrence and thus periodicity breaks down almost completely for a measurement at porous specimen 56. This effect is picked up by COR as a smaller correlation coefficient between the recurrence rate per

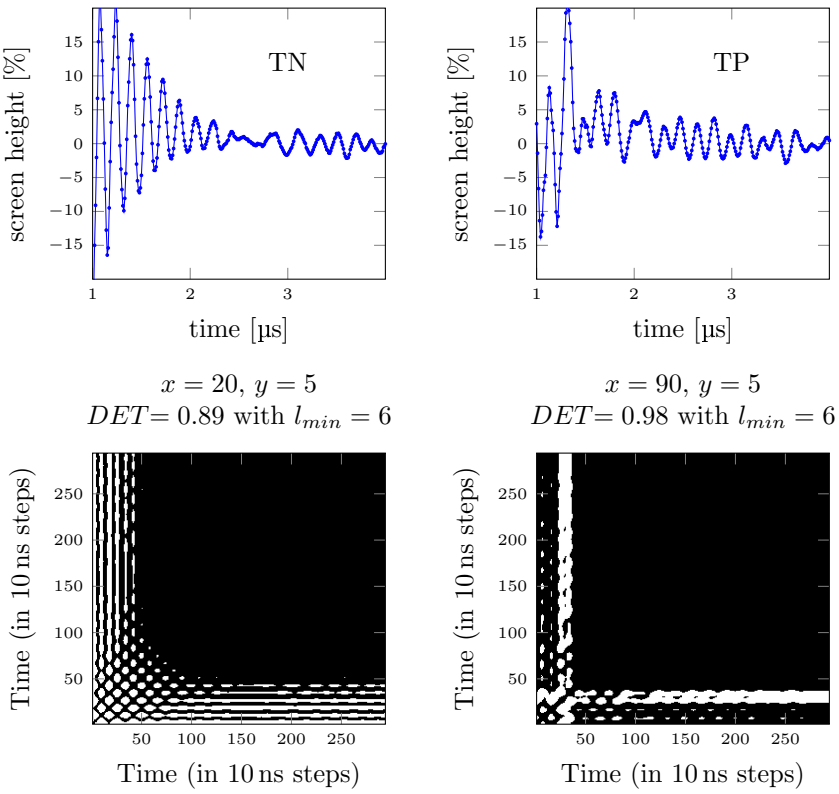


Figure 5.34: Bottom: recurrence plots of 5056-C3 with RQA parameters for optimum classification on 5056-A/B/D, using DET, Euclidean distance ($d = 2$, $\tau = 7$, $\epsilon = 16.5$); top: corresponding A-scans; example for correctly predicted non-porous area (left) and for correctly predicted porous area (right)

diagonal line and its distance to the main diagonal.

$RATIO$, using angular distance, delivers the best classification results of features depending additionally on the parameter l_{min} and slightly better than those for COR. Values for DET are for most of the scans up to ca. 1% smaller (fig. 5.37); balanced accuracies of L_{nor} and ENT are several percent lower.

RQA parameters for optimum classification with $RATIO$ are similar to those for COR (cf. table C.41 with table C.39) and thus RPs look similar (fig. 5.38). For measurements on the porous part, recurrence rate reduces; interestingly, determinism increases at the same time, because the largely reduced number of recurrent points almost merely occur around the LOI, and thus $RATIO{=}DET/RR$ increase with porosity.

5.5.2 Evaluation in time domain

The simple features x_{max}, s^2 and QCD are unreliable: they largely differ between different scans of one scan set and especially between 5056-B on one hand and 5056-A,

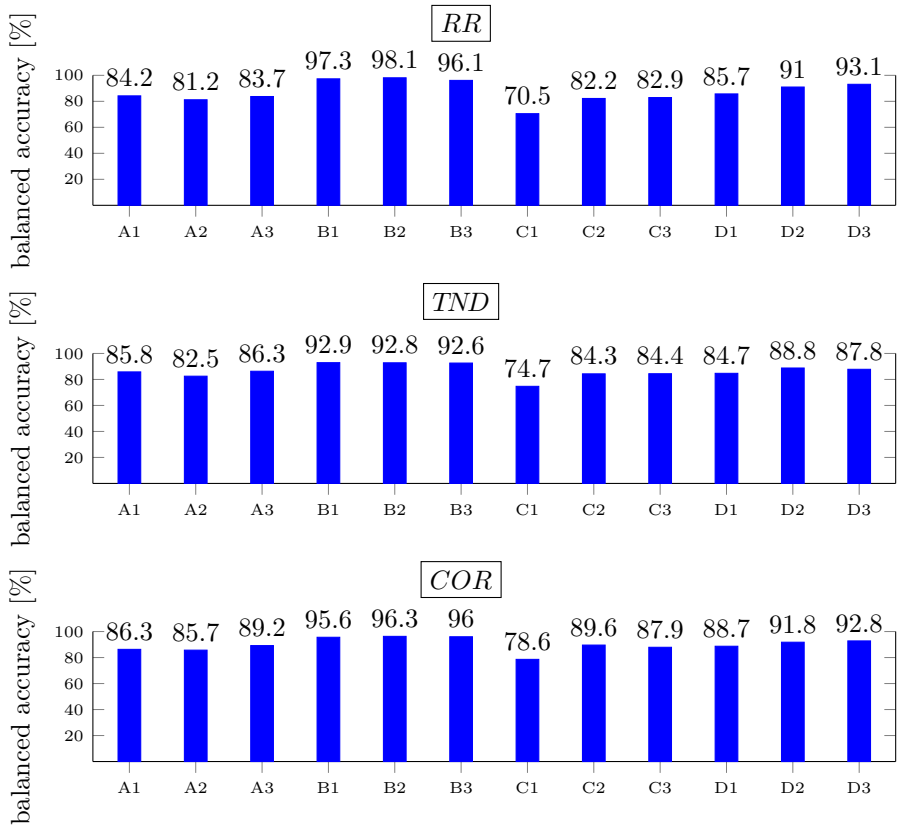


Figure 5.35: Classification test results (balanced accuracy acc_{bal}) for recurrence quantification features *RR*, *TND* and *COR* as BWE-equivalent, using angular distance, on scan sets 5056-A–D; RQA parameters (and decision threshold) leading to best classification results on three remaining scan sets, i.e. for A on B–D and so on

-C and -D on the other hand (table C.44 to table C.46); for 5056-A/C/D, balanced accuracies are between 47.5% and 84.5% for the three features.

Linear regression in time domain — trying to resemble the true back-wall echo as well as possible by weighting each point of the intermediate echo time series and summing the products of each time point and its weight — leads to excellent results with balanced accuracy and *TPR*s of $> 95\%$ for ten of twelve scans (fig. 5.39). The first two scans of scan set 5056-D have massively lower values of 55.8% and 79.4%. Small values of *TPR* cause these low accuracies, cf. figure 5.40.

Comparing, based on linear regression in time domain, a correct prediction of non-porous (true negative) with an incorrect prediction of non-porous (in reality porosity, false negative) (scan 5056-D1, fig. 5.41 top left with bottom left), the intermediate echoes look indeed qualitatively similar. The intermediate echo time series of the (roughly, taking into account manual scanning) same location (porous area) in another scan (5056-D3) has distinctly lower echoes (fig. 5.41 bottom right compared to bottom left). An A-scan of correctly predicted porosity in the outlier scan 5056-D1 consists

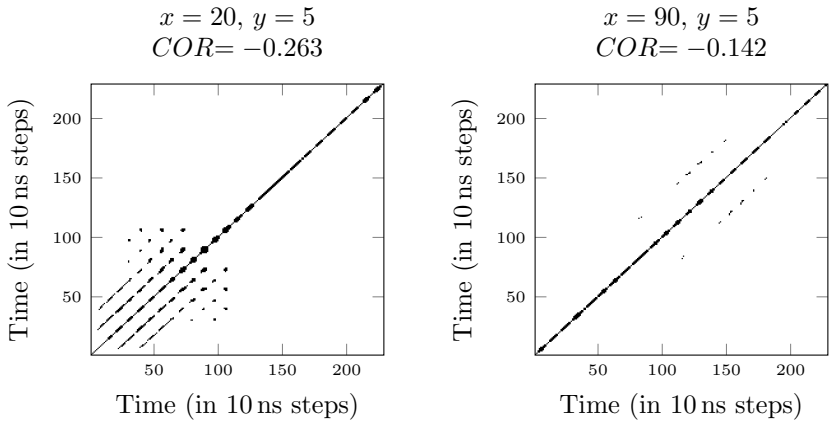


Figure 5.36: Recurrence plots of 5056-C3 with RQA parameters for optimum classification on 5056-A/B/D, using COR, angular distance ($d = 9$, $\tau = 9$, $\epsilon = 0.1\pi$); example for correctly predicted non-porous area (left) and for correctly predicted porous area (right). Corresponding A-scans in figure 5.34

of echoes in general *larger* than for the non-porous area (fig. 5.41 top right compared to top left). It is unclear why such an apparently different behaviour occurs between two repeated scans of roughly the same area — both recorded within five minutes. This different behaviour may be the reason for the maximum value and variance as BWE-equivalent delivering excellent *TPR*s for 5056-D3 and very bad values for 5056-D1 (table C.44, table C.45). These bad values also occur for example for all three scans of 5056-A. Eventually, doubts about the feasibility of the investigated tools in time domain cannot be ruled out.

Linear regression on time data has another important drawback: this method depends on a stable intermediate echo gate, i.e. the same cut-out has to be taken out of each A-scan time series for the weights to be correctly applied. Brandt et al. [2019] contains some investigations on the influence of a shift on results for linear regression on time data as well as on their Fourier transform.

5.5.3 Fourier Analysis

Standard features out of the amplitude spectrum (peak frequency, bandwidth etc.) are poor in a similar way as for the fabric specimens, with most balanced accuracies $< 70\%$ (table C.48 and C.49).

Linear regression on the amplitude spectrum from 0.66 MHz to 7 MHz (the window for optimum test results determined in sect. 5.4.2) delivers better results than linear regression applied directly on the time data, but still with outliers: smallest balanced accuracy is 83.2%, lowest $TPR = 72.4\%$ (fig. 5.42).

The same false negative as for linear regression on time series (fig. 5.41 bottom left) is also false negative for linear regression on the amplitude spectrum (fig. 5.43 bottom left): the algorithm cannot pick up the effect.

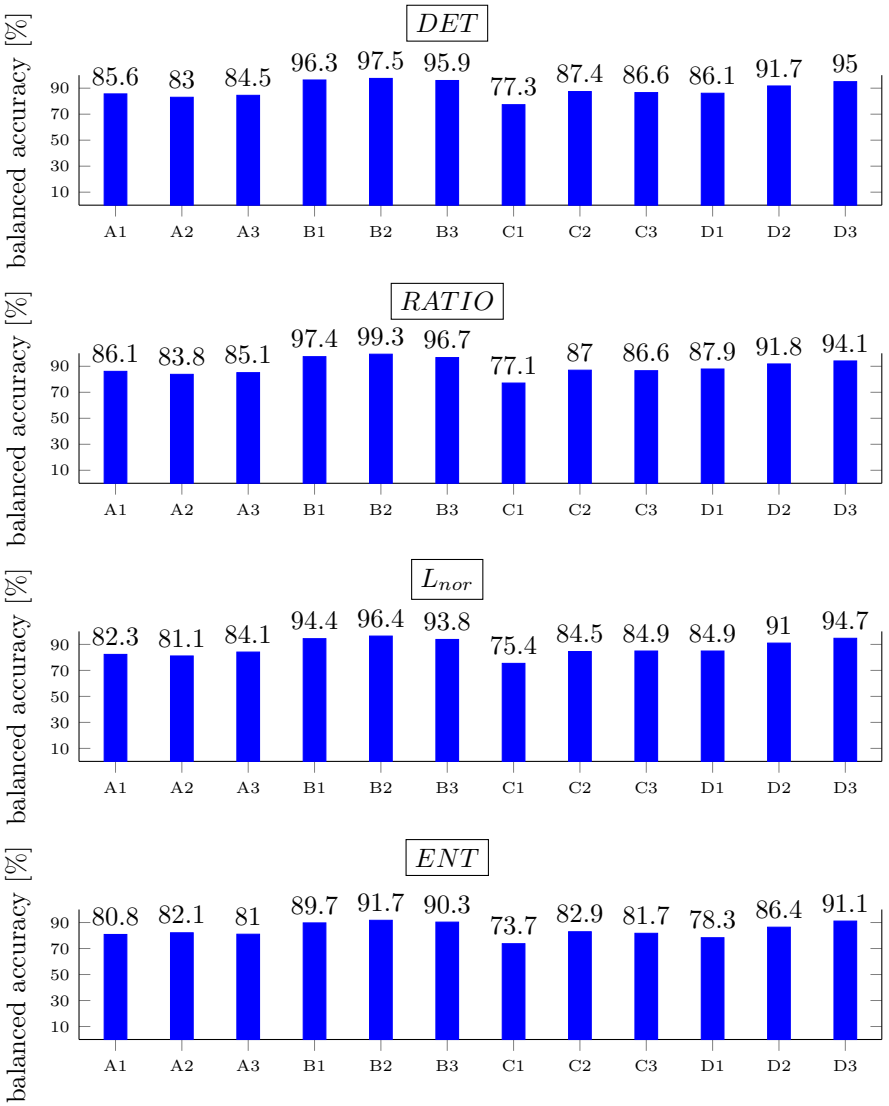


Figure 5.37: Classification test results (balanced accuracy acc_{bal}) for recurrence quantification features *DET*, *RATIO*, L_{nor} and *ENT* as BWE-equivalent, using angular distance, on scan sets 5056-A–D; RQA parameters (and decision threshold) leading to best classification results on three remaining scan sets, i.e. for A on B–D and so on

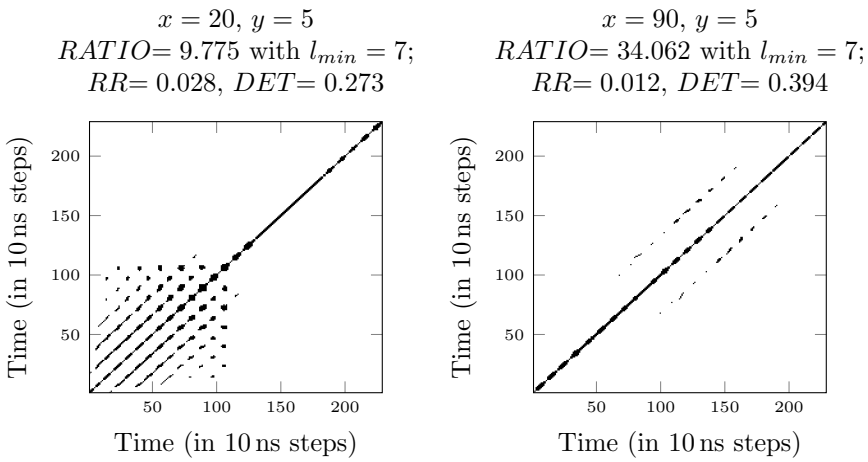


Figure 5.38: Recurrence plots of 5056-C3 with RQA parameters for optimum classification on 5056-A/B/D, using $RATIO$, angular distance ($d = 9$, $\tau = 9$, $\epsilon = 0.125\pi$); example for correctly predicted non-porous area (left) and for correctly predicted porous area (right). Corresponding A-scans see fig. 5.34

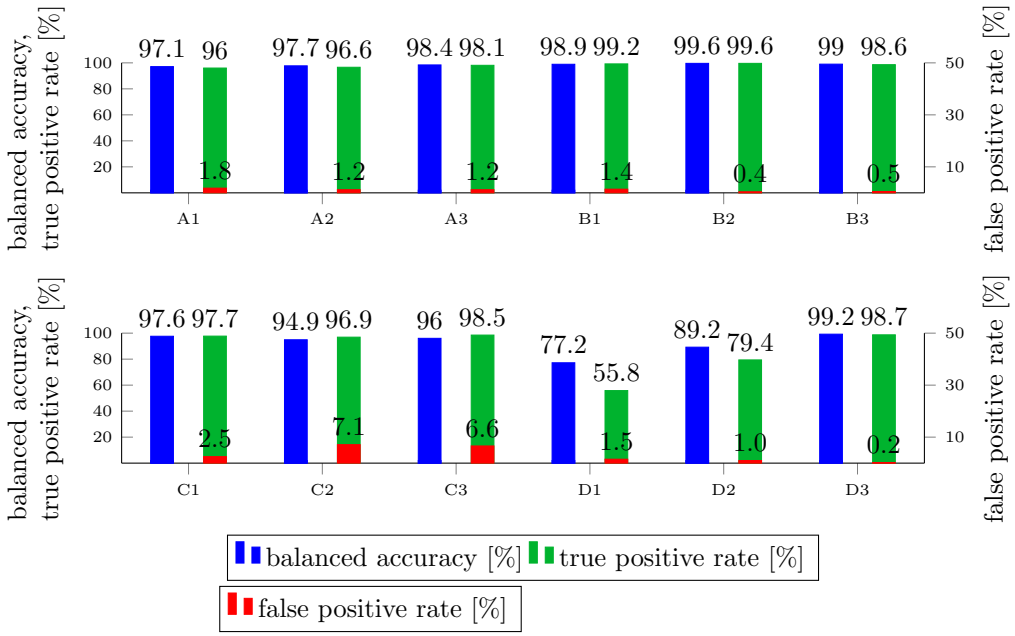


Figure 5.39: Classification results with linear regression in time domain for 5056A–D. Training (including decision threshold to achieve $TPR \approx TNR$) performed on 3 areas (scan sets), test results on scans of the remaining scan set, i.e. for A on B–D and so on

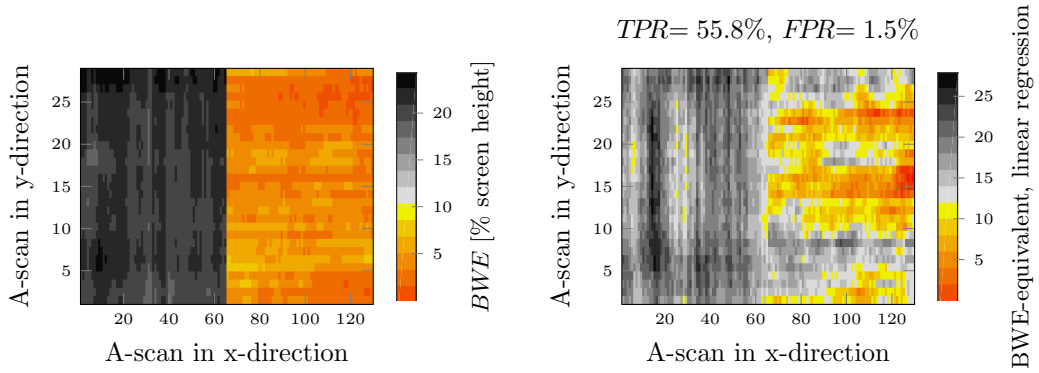


Figure 5.40: C-scans of 5056-D1; left: back-wall echo (porosity — *BWE* reduced by more than 50% — shown in colours); right: BWE-equivalent, linear regression on time data, test results (predicted porosity shown in colours, predicted defect-free areas in grey shades)

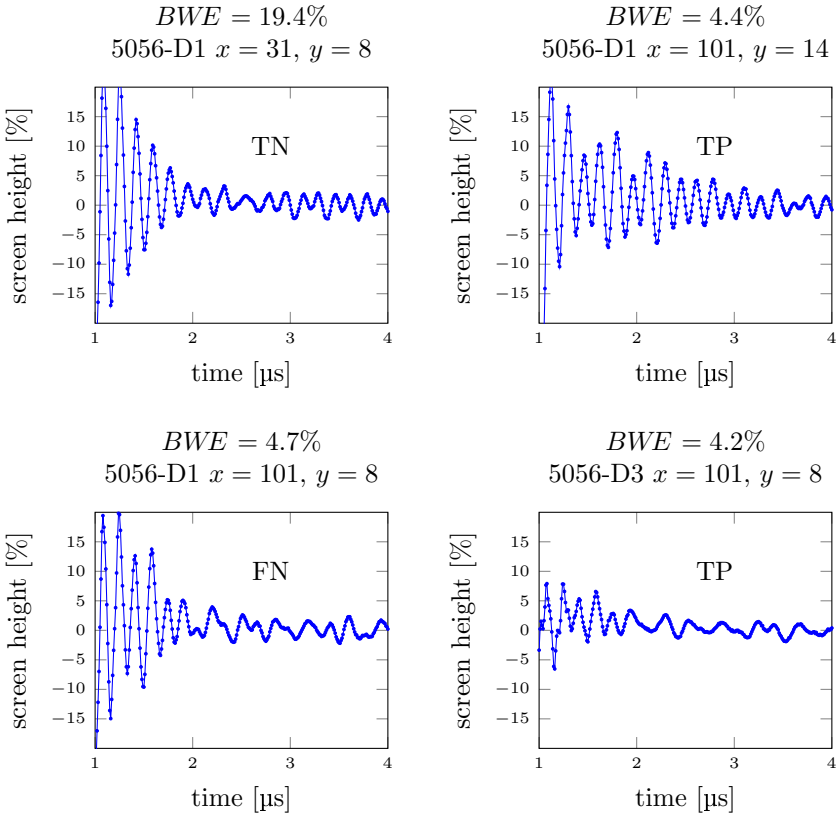


Figure 5.41: A-scan intermediate echo time series of scan 5056-D; with linear regression on intermediate echo time series: top left: correctly predicted non-porous area of 5056-D1; top right: correctly predicted porous area of 5056-D1; bottom left: falsely as non-porous predicted area of 5056-D1; bottom right: same location in 5056-D3: correctly predicted porous area)

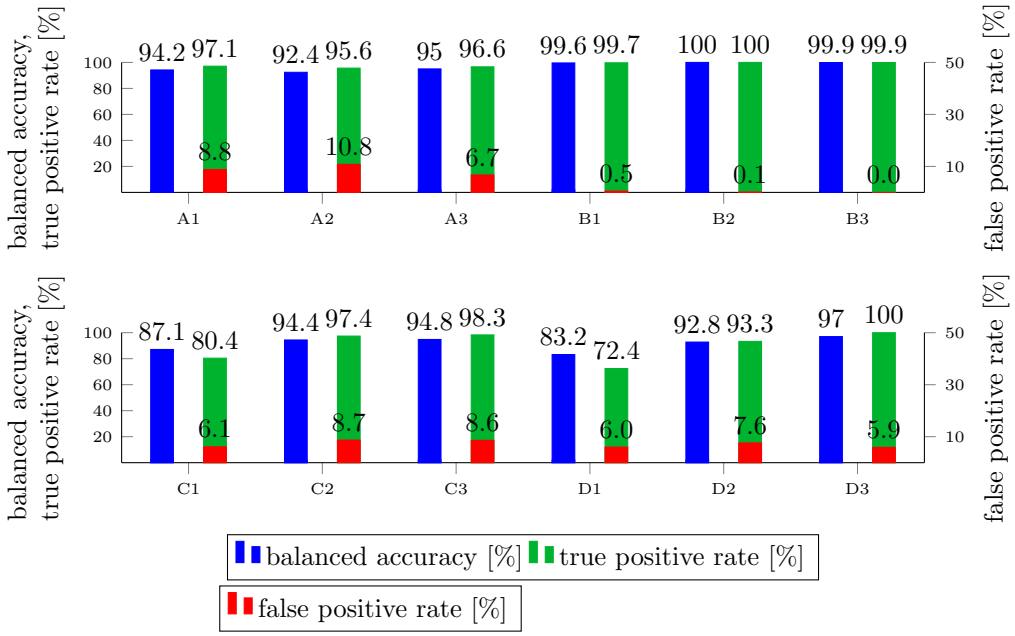


Figure 5.42: Classification results with linear regression on amplitude spectrum 0.66 MHz to 7 MHz for 5056A–D. Training (including decision threshold to achieve $TPR \approx TNR$) performed on 3 areas (scan sets), test results on scans of the remaining scan set, i.e. for A on B–D and so on

Applying a Hann window on the intermediate echo time series prior to Fourier transform distinctly increases classification results for scans 5056-C1, 5056-D1 and 5056-D2. All balanced accuracies take excellent values $> 96\%$ with TPRs $\geq 94.7\%$ (fig. 5.44, 5056-D1 visualised in fig. 5.45).

Weights determined on 5056-A–C and weights times amplitudes for a true negative (correctly predicted non-porous area) (fig. 5.46) are similar as for the pre-selection data (fig. 5.23). The weights largely reduce the influence of the dominating frequencies between approximately 1 MHz and 7 MHz (fig. 5.23 left) and thus make the influence of frequencies more uniform, largely pronouncing $\approx$ 0–1 MHz and 5 MHz and moderately pronouncing 7–17 MHz (fig. 5.23 right). The larger frequencies seem to contribute greatly to classification of porosity once leakage is reduced through the Hann window.

5.5.4 Summary of evaluation on unidirectional specimens, resonance case

RQA delivers good classification test results when using angular distance (superior to Euclidean distance in this case) with balanced accuracies for 11 out of 12 scans between 85.7% and 96.3% for the new feature COR, closely followed by DET and $RATIO$, for which results differ by $\approx 1\%$ compared to COR. Linear regression clearly outperforms RQA though, especially when used on the amplitude spectrum obtained through Fourier transform after applying a Hann window: balanced accuracies greater than 96% with $TPR \geq 94.7\%$ are achieved. Without applying a Hann window, linear

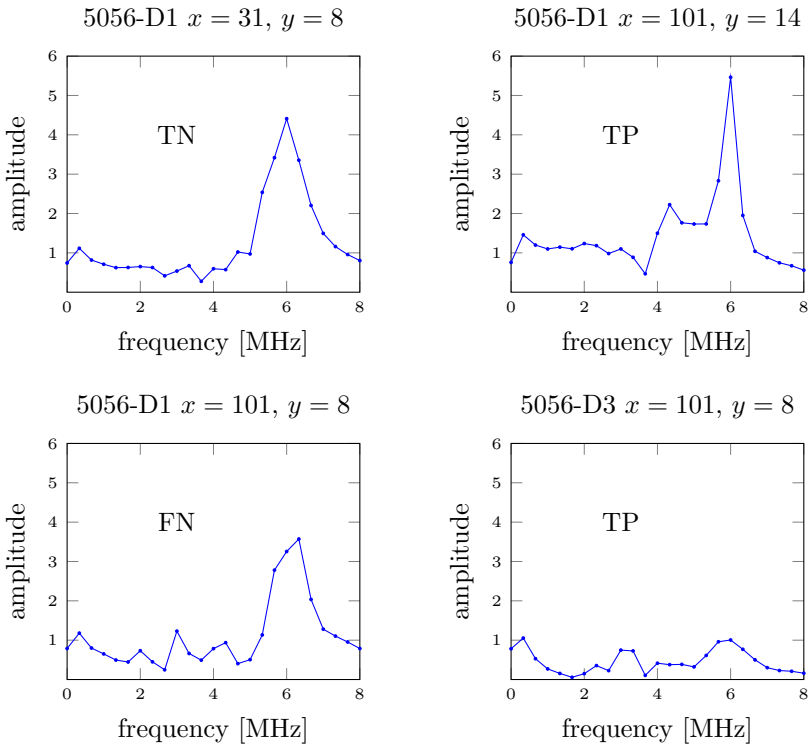


Figure 5.43: Amplitude spectrum of intermediate echo time series of scan 5056-D (of A-scans shown in fig. 5.41); classification with linear regression on amplitude spectrum: top left: correctly predicted non-porous area of 5056-D1; top right: correctly predicted porous area of 5056-D1; bottom left: falsely as non-porous predicted area of 5056-D1; bottom right: same location in 5056-D3: correctly predicted porous area)

regression on the amplitude spectrum and particularly on the original time series show a few outliers going down to balanced accuracies of $\approx 72\%$ or $\approx 56\%$, respectively.

Standard features in time and frequency domain (such as variance or peak frequency) provide poor results (including, in large contrast to the investigated fabric material, QCD).

The excellent results obtained with linear regression on the amplitude spectrum after applying a Hann window are achieved through weights that largely reduce the influence of main frequencies and rely on frequencies above 7 MHz. An additional test shall thus be conducted to provide more certainty in this method.

5.5.5 Test results on thinner specimens

Merely one pair of good and moderately porous 6 mm thick specimens of ≈ 0.25 mm ply thickness (specimen 50 and 56) was available for the investigations in previous sections. Therefore two 4 mm specimens (49 and 55) of same ply thickness shall be used

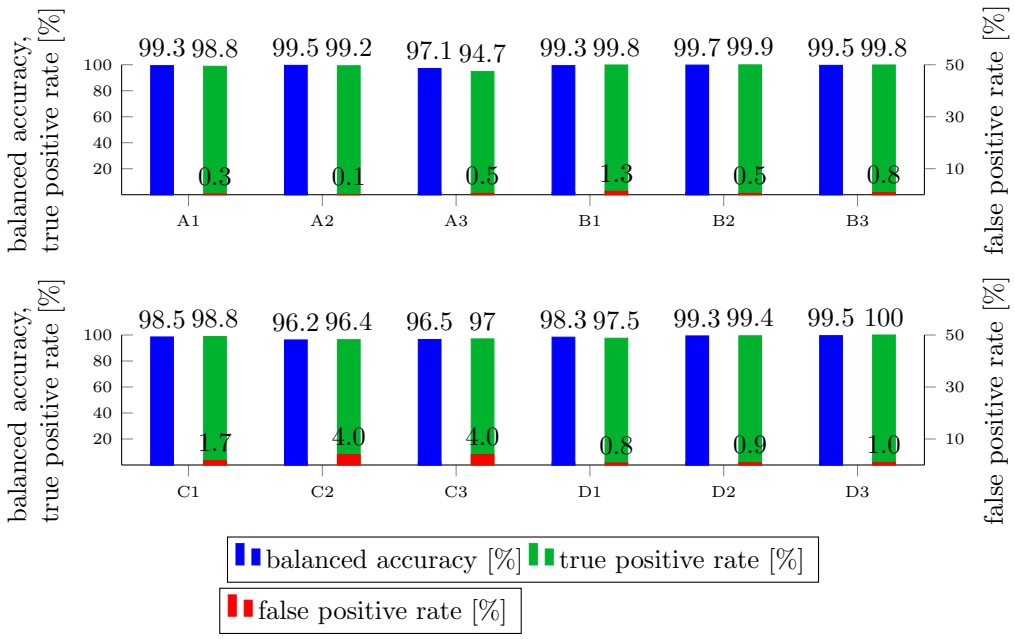


Figure 5.44: Classification results with linear regression on amplitude spectrum 0 MHz to 23 MHz for 5056A–D. Hann window applied on intermediate echo time series prior to Fourier transform. Training (including decision threshold to achieve *TPR* ≈ *TNR*) performed on 3 areas (scan sets), test results on scans of the remaining scan set, i.e. for A on B–D and so on

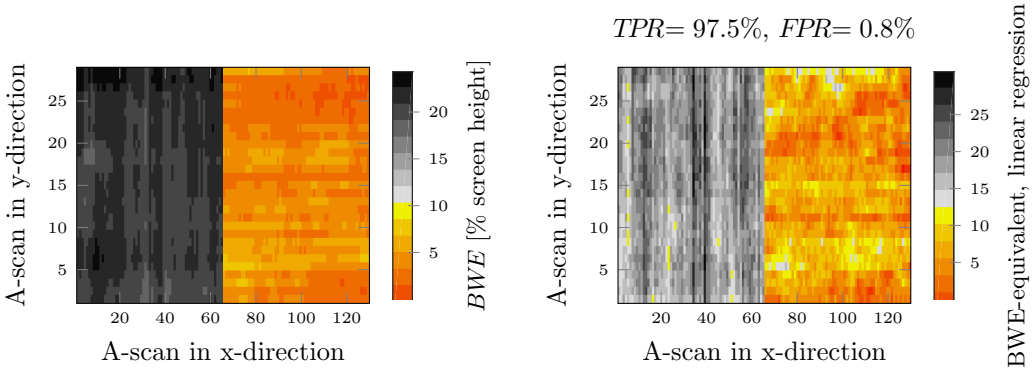


Figure 5.45: C-scans of 5056-D1; left: back-wall echo (porosity — *BWE* reduced by more than 50% — shown in colours); right: BWE-equivalent, linear regression on amplitude spectrum, Fourier transform after applying a Hann window on time data, test results (predicted porosity shown in colours, predicted defect-free areas in grey shades)

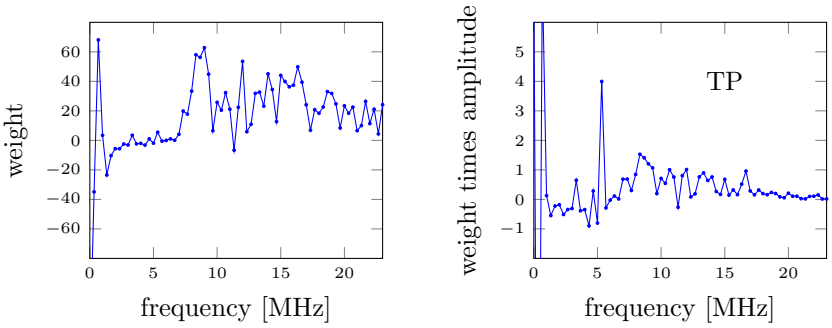


Figure 5.46: Weights for linear regression on amplitude spectrum, Fourier transform after applying a Hann window on intermediate echo time series, determined on 5056-A–C (left), weights times amplitudes of correctly predicted non-porous area (5056-D1x21y20, right)

for additional test of the best algorithm, linear regression on the amplitude spectrum with applying a Hann window prior to Fourier transform.

Training is performed on the first half of the intermediate echo gate of specimens 50 and 56, from 1 µs to 2.5 µs, to determine whether the algorithm and its parameters can be transferred from one specimen set to another.

Results on 6 mm thick specimens, reduced intermediate echo gate

Cross-validation test results on specimens 50 and 56 with a reduced intermediate echo gate from 1 µs to 2.5 µs still lead to balanced accuracies and *TPR*s greater than 91% with one exception of $TPR = 86.6\%$ for 5056-C1 (fig. 5.47). Results are thus still good to excellent, though the intermediate echoes out of ≈ 2 mm thickness of the specimens are compared to the back-wall echo, which collects information over the whole thickness. This suggests a quite uniform distribution of porosity over the whole thickness (cf. fig. 5.3).

Test on 4 mm thin specimens

The weights of linear regression and the decision threshold determined for the reduced intermediate echo gate on scan sets 5056-A/B/D are tested finally on two scan sets, three scans each, of 4 mm thin specimens 49 and 55. Classification results are not as good as on specimens 50 and 56, but balanced accuracy is for four scans $> 91\%$, and $TPR \geq 92\%$ in five of overall six scans (fig. 5.48).

The amplitude spectrum has its largest values up to ≈ 8 MHz (fig. 5.49 top right), and the weights (determined on 5056-A/B/D) are largest for frequencies $\gtrsim 8$ MHz (fig. 5.49 bottom left). Thus weights times amplitude spectrum are more uniform over the whole range of frequencies used for linear regression (fig. 5.49 bottom right) — similar as for the intermediate echo gate from 1 µs to 4 µs for 6 mm thickness.

This test is considered successful with mostly good classification results (e.g. fig. 5.50), especially taking into account that this is achieved with a model completely trained on other, thicker specimens.

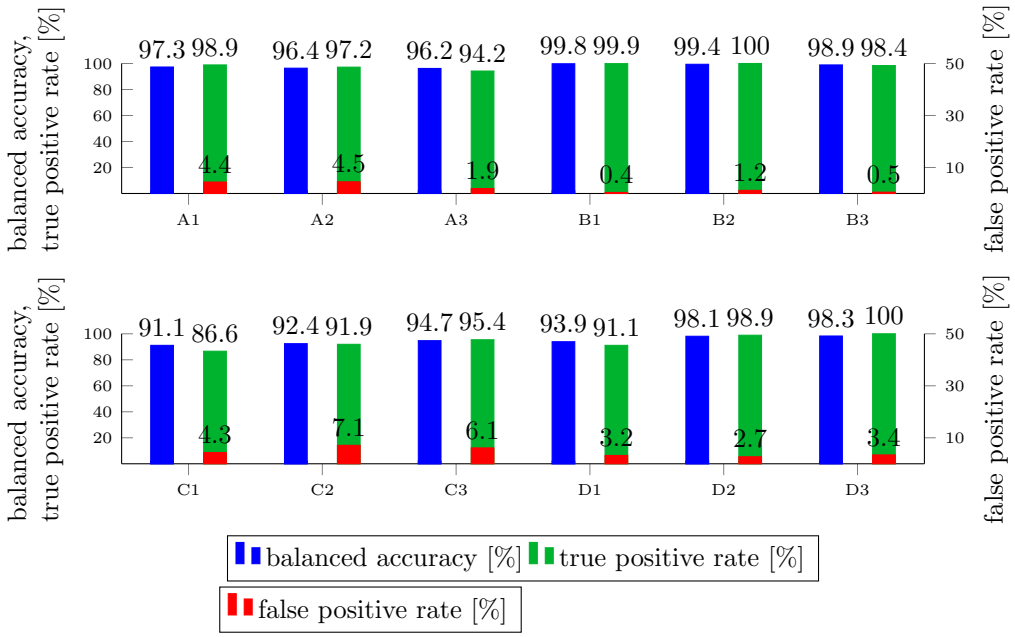


Figure 5.47: Classification results with linear regression on amplitude spectrum from 0 MHz to 22.67 MHz for 5056-A–D; reduced IE gate from 1 µs to 2.5 µs. Hann window applied on intermediate echo time series prior to Fourier transform. Training (including decision threshold to achieve $TPR \approx TNR$) performed on 3 areas (scan sets), test results on scans of the remaining scan set, i.e. for A on B–D and so on

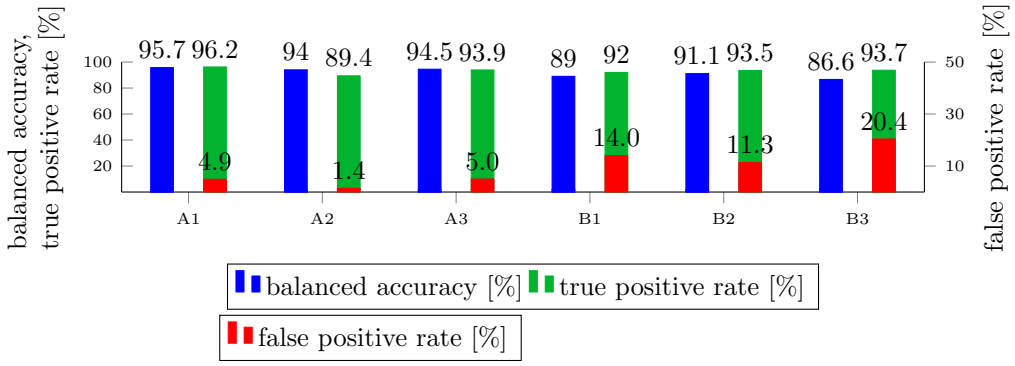


Figure 5.48: Classification results with linear regression on amplitude spectrum 0 MHz to 22.67 MHz for 4955-A/B; IE gate from 1 µs to 2.5 µs. Hann window applied on intermediate echo time series prior to Fourier transform. Training (including decision threshold to achieve $TPR \approx TNR$) performed on scan sets 5056-A/B/D, IE gate 1 µs–2.5 µs

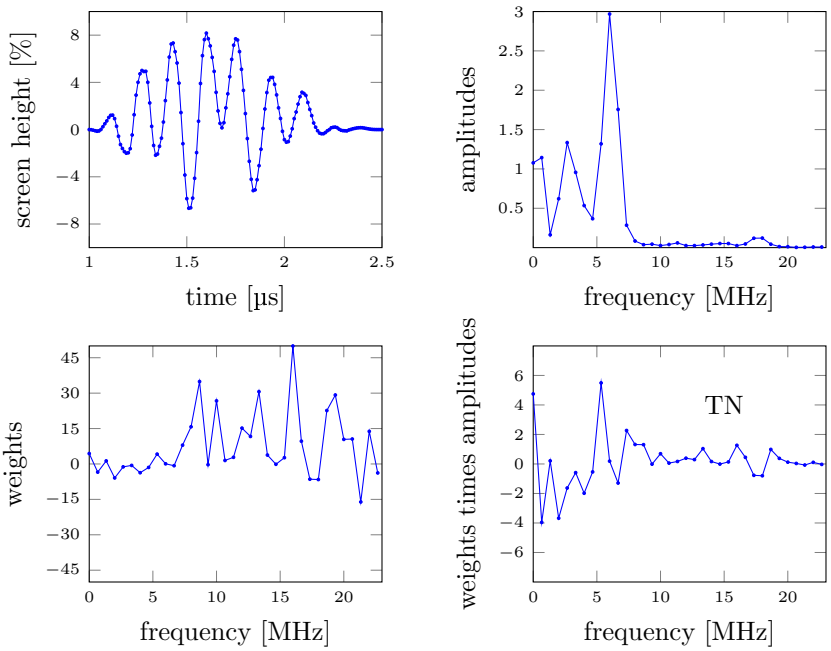


Figure 5.49: Intermediate echo time series of non-porous area after applying a Hann window, specimen 49 (4955-B2x100y8) (top left), its amplitude spectrum (top right), weights from 5056-A/B/D used for test linear regression on amplitude spectrum (bottom left) and weights times amplitude spectrum (bottom right)

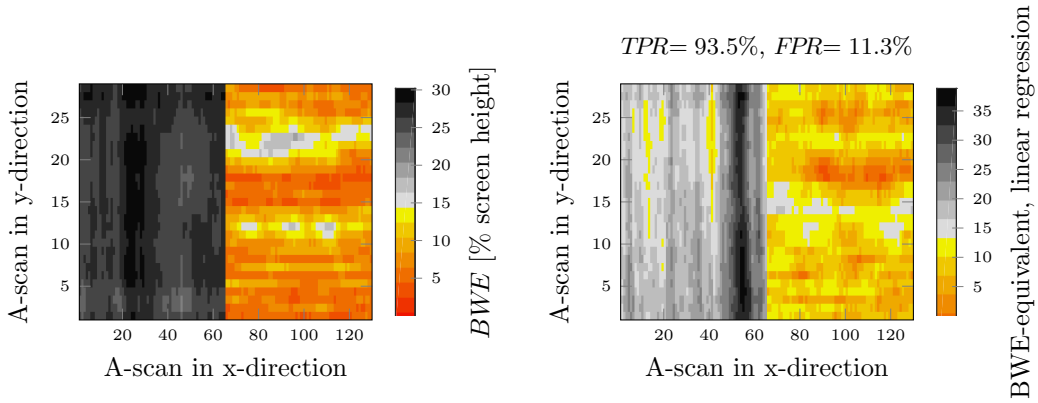


Figure 5.50: C-scans of 4955-B2; left: back-wall echo (porosity — *BWE* reduced by more than 50% — shown in colours); right: BWE-equivalent, linear regression on amplitude spectrum from 0 MHz to 22.67 MHz, Fourier transform after applying a Hann window on time data, intermediate echo gate from 1 µs to 2.5 µs, test results (predicted porosity shown in colours, predicted defect-free areas in grey shades)

5.6 Evaluation on unidirectional specimens, non-resonance case

The results for classification into porous or non-porous shown so far have all been obtained on specimens with a ply thickness t_p approximately equal half the main wavelength λ of the ultrasonic wave so that a resonance effect occurs, leading to rather sinusoidal (damped) signals (cf. fig. A.1, fig. 5.31). The inspection frequency of $\approx 5\,\mathrm{MHz}$ leads to this effect with the ply thickness of 0.25 mm–0.35 mm.

For smaller ply thickness it would be possible to use a smaller frequency (leading to $t_p \approx \lambda$), this has however the disadvantage of decreased detectability of small defects such as porosity. To the other end, ultrasonic pulsed waves with a peak frequency much larger than 5 MHz (to achieve $t_p \approx \lambda/2$ for smaller t_p) do not reach the back-wall of parts of larger thickness, because high frequencies are damped by the CFRP material.

Furthermore and important, the resonance effect is not favoured in a general industrial inspection setting, because direct echoes of larger defects (such as delaminations) are more difficult to be detected.

Classification results from time series analysis on ultrasonic intermediate echo signals obtained at specimens with smaller ply thickness, keeping the testing frequency of 5 MHz, are thus very interesting. Sections 5.6.1 to 5.6.3 show results of RQA, evaluation in time domain and with Fourier analysis on two 6 mm thick specimens 24 and 38. The best methods are tested on 4 mm thin specimens 26 and 36 in section 5.6.6.

These specimens have a smooth surface (without removed peel-ply) similar to specimens M13–15. Back-wall echoes have thus roughly similar heights, taking the 12 dB higher amplification into account (fig. 5.51 compared to fig. 2.6).

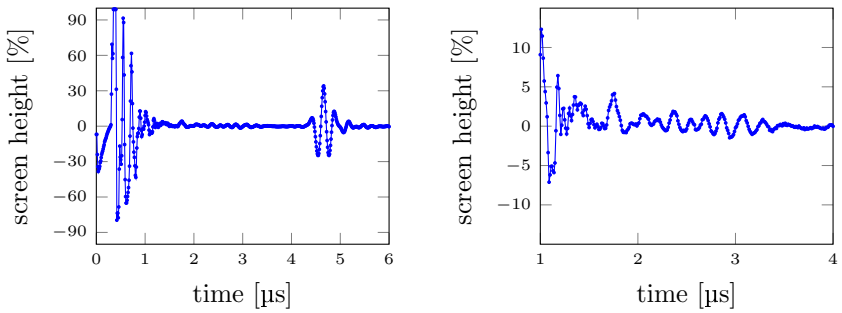


Figure 5.51: A-scan time series of material B specimen 24 (non-porous), 24-A1x5y5 (left); magnified intermediate echo time series (of the same A-scan) (right)

The intermediate echo time series of ultrasonic inspection without the resonance effect have distinctly lower amplitudes (fig. 5.51), even more so when compared to the back-wall echo, which is greater than for specimen 50 (fig. 5.51) and 56 due to the smoother surface. Furthermore, though a sinusoidal signal can be recognised, the intermediate echoes from non-resonant inspection are clearly more complex.

5.6.1 Recurrence quantification analysis

Test results are determined as in section 5.5.1: RQA parameters leading to maximum balanced accuracy with parity of *TPR* and *TNR* on three scan sets are used to determine classification results on three scans of the remaining fourth scan set. This is repeated three times to have each scan set as test set once.

As throughout this work, the standard Euclidean distance as well as recently proposed angular distance are used to determine recurrence in recurrence matrices before subsequent computation of RQA features.

Euclidean distance

Classification test results (balanced accuracies) for features not depending on parameter minimum line length l_{min} (RR, DIV, TND and COR) are not feasible for this specific application. DIV delivers worst results, with $acc_{bal} < 70\%$ for all scans (table C.55), far below the goal in this work to achieve 90%. RR is better, but still has five scans with $acc_{bal} < 80\%$ (three of them even below 70%, cf. table C.54). The newly proposed COR (table C.57) surpasses the classical TND (table C.56), but still has $acc_{bal} < 80\%$ for seven out of twelve scans. RQA features depending on l_{min} DET, L_{nor} and ENT are not feasible for this specific application either. Balanced accuracies are below 80% for most scans (table C.58, C.60), for ENT even below 71% for 11 of 12 scans (table C.61).

The feature $RATIO$, using Euclidean distance, leads to balanced accuracies greater than 80% for all scans, for one scan exceeding the goal of $acc_{bal} = 90\%$. True positive rates are more varying (from 67.7% to 98.2%, fig. 5.52): in average, these are considered to be moderate results with sufficient difference of $RATIO$ between TNR and TPR (fig. 5.53).

Recurrence plots show almost no recurrence in the first third of the embedded time series. For the last two thirds, recurrence reduces for porosity (fig. 5.54 bottom right compared to left). A smaller RR together with slightly greater DET leads to the best classification results for specimen 2438 from all features of RQA with Euclidean distance. The delay is again $\tau \approx \lambda/2$, leading to diagonal lines perpendicular to the LOI.

Angular distance

Classification results of RQA using angular distance are worse than those for Euclidean distance, in contrast to results with resonance effect (sect. 5.5.1), where angular distance improves RQA results. Balanced accuracies for features not depending on l_{min} RR (table C.62), DIV (table C.63) and TND (table C.64) are below 80% or even below 70% for almost all scans and thus far below the goal $acc_{bal} = 90\%$. The new feature COR even leads to $acc_{bal} < 60\%$ for four out of twelve scans (table C.65).

Features depending on minimum line length DET (table C.66), $RATIO$ (table C.67) and L_{nor} (table C.68) lead to balanced accuracies around 80% for scan set 2438-C and

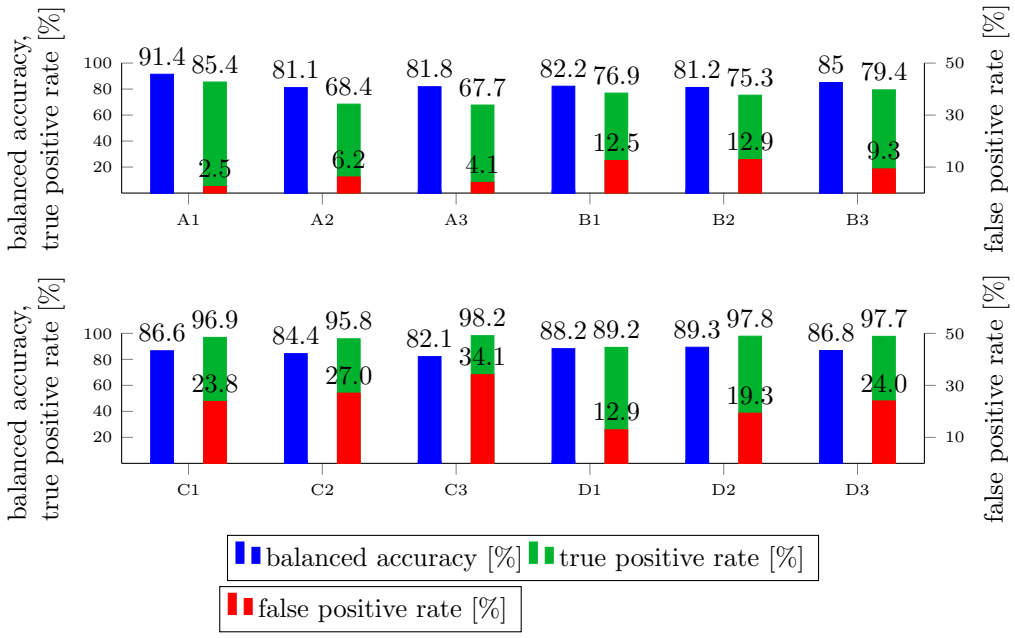


Figure 5.52: Classification test results (balanced accuracy acc_{bal}) for recurrence quantification feature *RATIO* as BWE-equivalent, using Euclidean distance, on scan sets 2438-A–D; RQA parameters (and decision threshold) leading to best classification results on three remaining scan sets, i.e. for A on B–D and so on

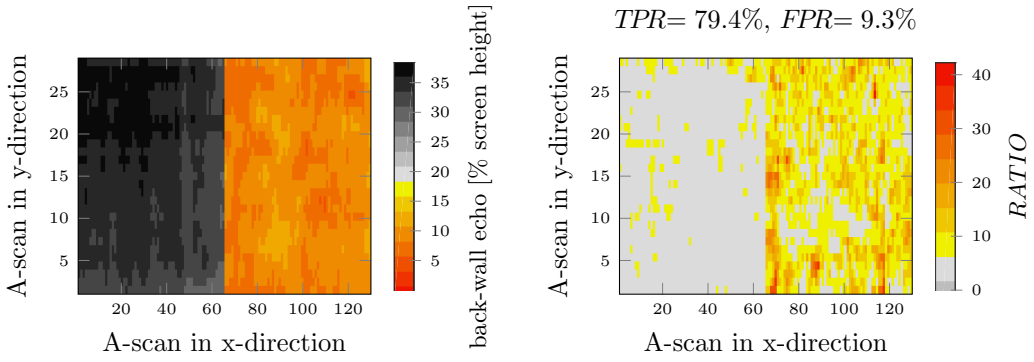


Figure 5.53: C-scans of 2438-B3; left: back-wall echo (porosity — *BWE* reduced by more than 50% — shown in colours); right: *RATIO*, Euclidean distance. RQA parameters (and decision threshold) leading to best classification results on 2438-A/C/D ($d = 5$, $\tau = 9$, $\epsilon = 1.3$, $l_{min} = 4$; predicted porosity shown in colours, predicted defect-free areas in grey shades)

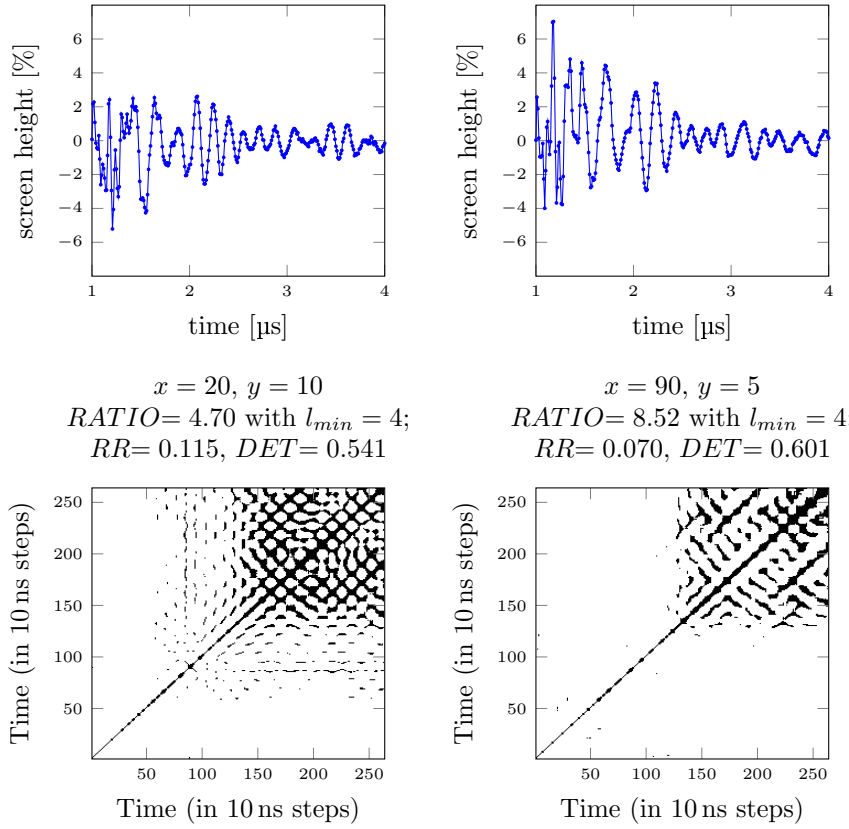


Figure 5.54: Bottom: recurrence plots of 2438-B3 with RQA parameters for optimum classification on 2438-A/C/D, using $RATIO$, Euclidean distance ($d = 5$, $\tau = 9$, $\epsilon = 1.3$); top: corresponding A-scans; example for correctly identified non-porous area (left) and for correctly identified porous area (right)

2438-D and partly 2438-A, but test results on scan set 2438-B are even below 60%; ENT is even slightly worse for scan set 2438-C (table C.69).

Overall, $RATIO$, using Euclidean distance, is the most appropriate RQA feature for detection of porosity out of intermediate echo time series of ultrasonic testing to detect porosity for unidirectional CFRP material B in non-resonance case.

5.6.2 Evaluation in time domain

The maximum and variance of the intermediate echo time series are a bad BWE-equivalent similar as for the investigations on porosity in unidirectional specimens with resonance effect. Also quartile coefficient of dispersion leads — in contrast to fabric specimens (porosity introduced with Freon capsules) — to bad results; balanced accuracy is for most scans even below 60% (table C.70 to C.72).

Linear regression on the intermediate echo time series delivers balanced accuracies — besides one poor value of 48.4% — between 72.5% and 88.2%. This is distinctly

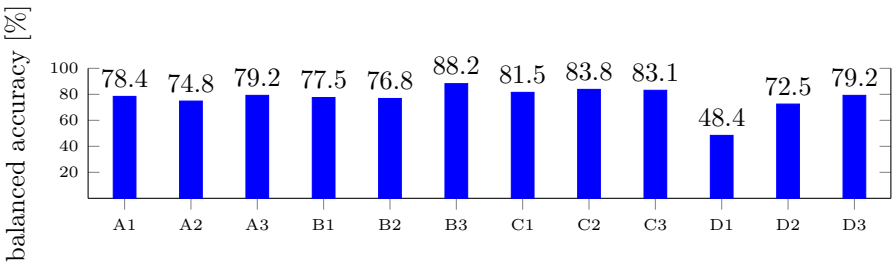


Figure 5.55: Classification test results (balanced accuracy acc_{bal}) for linear regression on intermediate echo time series on scan sets 2438-A–D; training for each scan set on the remaining scan sets, i.e. for A on B–D and so on

better than with maximum, variance or QCD, but also worse than results for RQA feature $RATIO$, using Euclidean distance (sect. 5.6.1) and clearly worse than linear regression on intermediate echo time series for specimens 50 and 56 with resonance effect (sect. 5.5.2).

5.6.3 Fourier Analysis

Some of the standard features determined out of the amplitude spectrum (table C.74 and C.75) deliver for 2438-C and 2438-D balanced accuracies around 80%, but in all other cases $< 68\%$, sometimes around 50% — near a random classifier.

Linear regression on the amplitude spectrum is for material B without resonance effect only performed after applying a Hann window; results without Hann window have already been identified being too poor in pre-investigations (sect. 5.4.2). Classification test results are not as excellent as for material B with resonance effect. They reach however for a window of 1 MHz–30 MHz balanced accuracies between 81.7% and 88.9%, smallest TPR 76.1% (fig. 5.56); results with the second selected window of 0 MHz–30 MHz are from 79.4% to 89.1%, smallest TPR 65.4% (table C.76).

The weights (fig. 5.57 bottom left) reduce the influence of frequencies between $\approx$ 2 MHz and $\approx$ 7 MHz. The resulting weights times amplitudes for a correctly predicted non-porous area are thus more uniform over the frequency range up to $\approx$ 18 MHz (fig. 5.57 bottom right) than the original spectrum (fig. 5.57 top right) out of the intermediate echo time series after applying a Hann window (fig. 5.57 top left). This is a similar behaviour as observed on the pre-selection data sets 2438-α/β (fig. 5.27). Again, the larger frequencies seem to contribute greatly to classification of porosity once leakage is reduced through the Hann window.

5.6.4 Summary of evaluation on unidirectional specimens, non-resonance case

Intermediate echoes from inspections of CFRP without the resonance effect (ply thickness $t_p \neq \lambda/2$) allow a clearly less good classification compared to inspections with

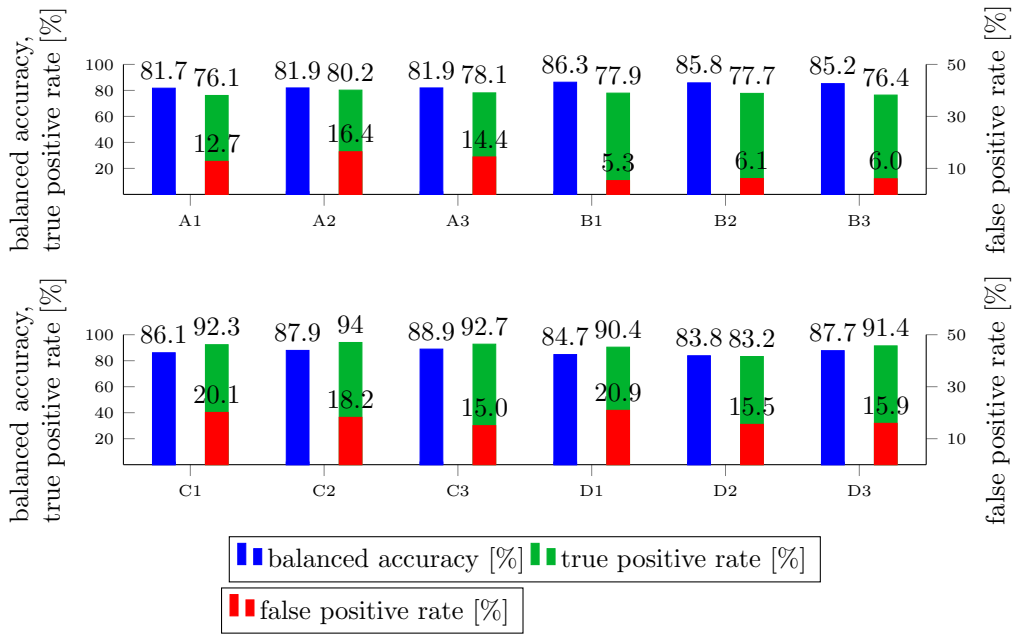


Figure 5.56: Classification results with linear regression on amplitude spectrum 1 MHz to 30 MHz for 2438-A–D. Hann window applied on intermediate echo time series prior to Fourier transform. Training (including decision threshold to achieve $TPR \approx TNR$) performed on 3 areas (scan sets), test results on scans of the remaining scan set, i.e. for A on B–D and so on

resonance effect (sect. 5.5). Best results are achieved with linear regression on the amplitude spectrum with application of a Hann window prior to Fourier analysis (fig. 5.58) and with the RQA feature $RATIO$ using Euclidean distance. Test results of balanced accuracies are between 81.1% and 91.4%. Regarding Fourier analysis, the amplitudes of frequencies above the main frequencies apparently carry information about porosity best, once influence of leakage is reduced.

COR, DET and L_{nor}, RQA using Euclidean distance, follow with balanced accuracies around 80%. All other investigated features from RQA or from tools in time domain and Fourier analysis provide for most scans $acc_{bal} < 80\%$, partly distinctly smaller.

5.6.5 Extended recurrence quantification analysis

Based on best results with RQA for unidirectional material, non-resonance case — $RATIO$, Euclidean distance — four different extensions of RQA beyond the standard recurrence plot (with fixed recurrence threshold ϵ) are investigated:

- Adapted ϵ to achieve a fixed recurrence rate. Based on the recurrence rates obtained for above mentioned best results, recurrence rate (RR) is varied between 1% and 30% (1% steps between 1% and 10%, 2% steps between 12% and

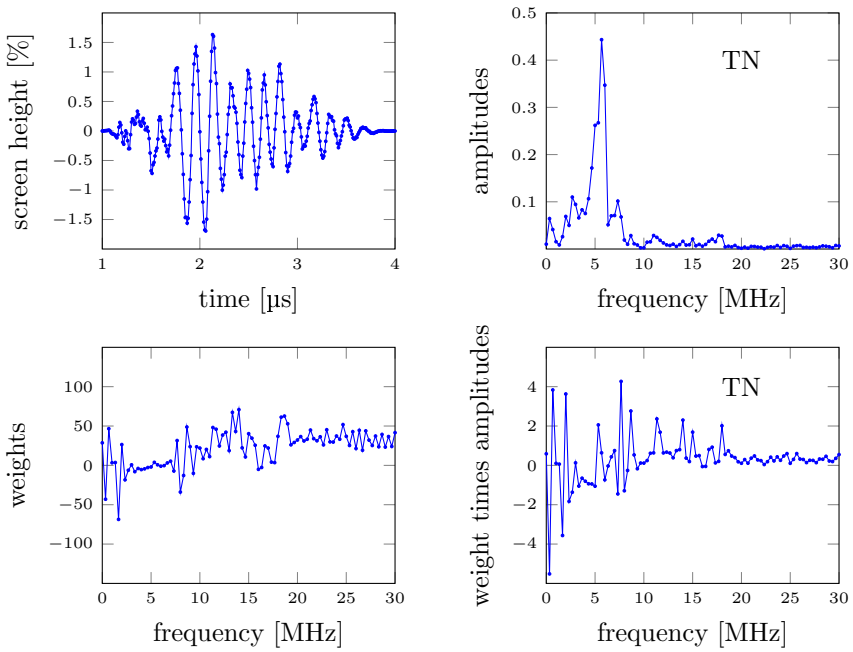


Figure 5.57: Intermediate echo time series of non-porous area after applying a Hann window, specimen 24 (2438-B3x20y10) (top left), its amplitude spectrum (top right), weights from 2438-A/C/D used for test linear regression on amplitude spectrum (bottom left) and weights times amplitudes (bottom right)

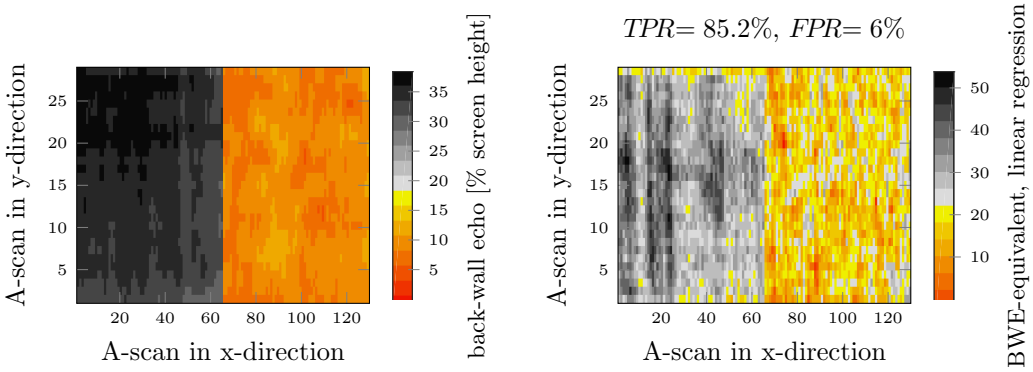


Figure 5.58: C-scans of 2438-B3; left: back-wall echo (porosity — *BWE* reduced by more than 50% — shown in colours); right: BWE-equivalent, linear regression on amplitude spectrum from 1 MHz to 30 MHz, Fourier transform after applying a Hann window on time data, weights and decision threshold determined on 2438-A/C/D (predicted porosity shown in colours, predicted defect-free areas in grey shades)

30%). *DET* is evaluated, since for a fixed recurrence rate this is proportional to *RATIO*. This approach would have the benefit of being less dependent on a correct calibration of the gain of the inspection equipment, because ϵ is 'automatically' determined.

- *RATIO* based on cross recurrence plots (CRPs), Euclidean distance. Reference intermediate echo time series is the mean out of one subset of one volume scan of the data sets for pre-selection of parameters (sect. 5.4) from non-porous sample 24, 24-α1, of the area $x = 28$–30, $y = 20$–22. Recurrence threshold ϵ is varied between 0.1 and 2 in 0.1 steps, based on the recurrence threshold ϵ with values of 1 and 1.3 for best values of *RATIO* (table C.59).

- *RATIO* based on joint recurrence plots (JRPs), Euclidean distance, reference intermediate echo time series as for CRPs; ϵ varied between 0.1 and 2 in 0.1 steps.

- *RATIO* based on newly proposed difference recurrence plots (DRPs), Euclidean distance, reference intermediate echo time series as for CRPs; ϵ varied between 0.1 and 2 in 0.1 steps.

None of the methods provide any benefit for the application investigated in this thesis (fig. 5.59) when compared to standard RP results (fig. 5.52):

- A fixed recurrence rate is evidently not appropriate here, because the feature $RATIO = DET/RR$ with fixed recurrence threshold ϵ takes advantage of the changing recurrence rate. Test results for *DET* with fixed recurrence rate are similar to those for *DET* with fixed ϵ (table C.58).

- Results for cross, joint and difference recurrence plots are similar to the ones for standard RPs — but slightly worse in the range of few percent. Best is CRP; a performance superior to the one of JRPs can be expected from the literature, cf. section 3.3.4 (though differences are small). Because no advantage can be obtained in this specific application based on best results of RPs, cross, joint and difference recurrence plots are not further pursued in this thesis.

5.6.6 Test results on thinner specimens

The two best algorithms, linear regression on the amplitude spectrum after applying a Hann window and the RQA feature *RATIO*, Euclidean distance, are first applied on the first half of the intermediate echoes of ultrasonic measurements of specimens 2438 (1 µs–2.5 µs), table C.82 and C.84. Balanced accuracies are reduced by up to $\approx 10\%$, results for *RATIO* vary more, but are also in average better than for linear regression. The frequency information in the first half of the intermediate echo time

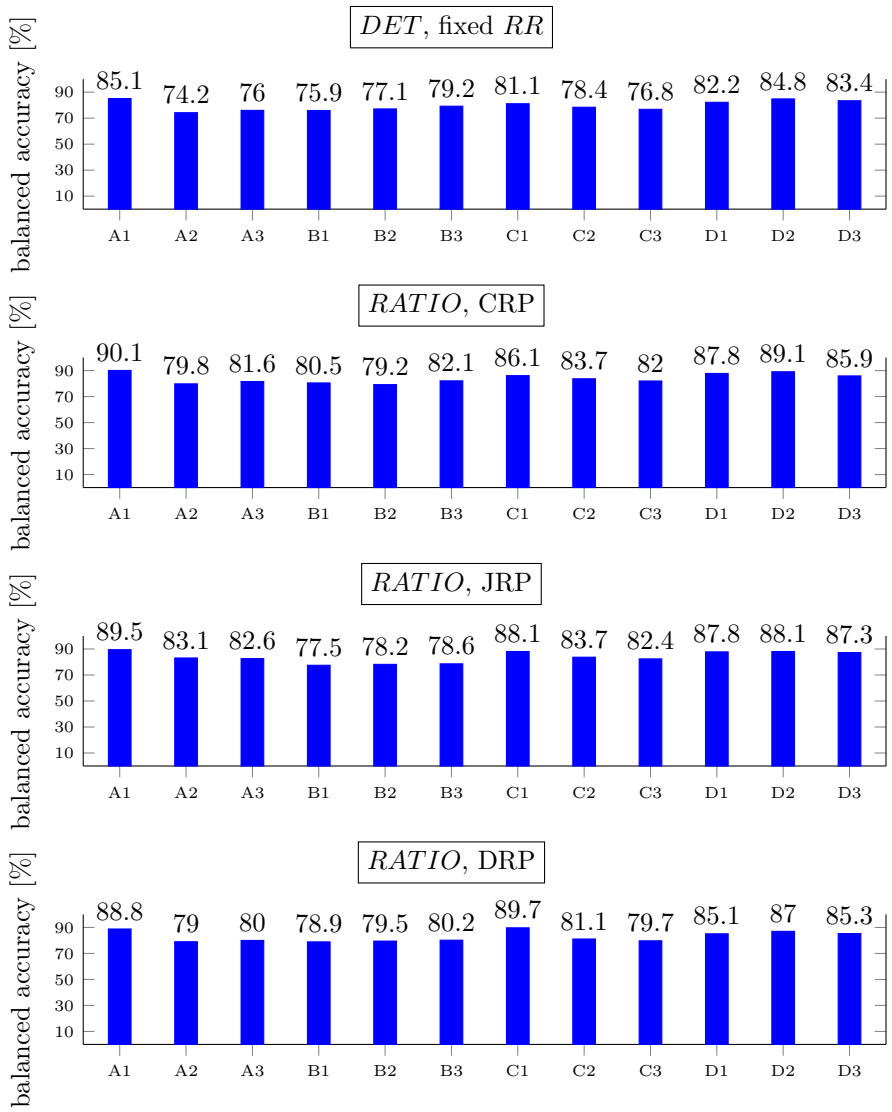


Figure 5.59: Classification test results (balanced accuracy acc_{bal}) for recurrence quantification features: *DET* using a fixed recurrence rate (instead of fixed ϵ; *DET* is proportional to *RATIO* for *RR* = const.; top); *RATIO* based on cross recurrence plot (second from top); *RATIO* based on joint recurrence plot (third from top); *RATIO* based on difference recurrence plots (bottom), all *RATIO* using Euclidean distance, on scan sets 2438-A–D; RQA parameters (and decision threshold) leading to best classification results on three remaining scan sets, i.e. for A on B–D and so on. Reference for cross, joint and difference recurrence plot is the mean of intermediate echo time series out of 24-α1 of the area $x = 28$–30, $y = 20$–22

series obviously does not suffice for classification. *RATIO* relies mainly on the last two thirds of the original full intermediate echo time series (cf. fig. 5.54) and thus performs worse applied on the first half of the time series only. Consequently, test on thinner specimens 2636 are even worse with all but one balanced accuracies $< 70\%$ (table C.83 and C.85).

This test is thus not considered successful; a test of the algorithms has to be performed on further porous specimens of the same 6 mm thickness in the future.

5.7 Comparison and summary of results

Recurrence quantification analysis (RQA), tools in time domain and Fourier analysis, especially using linear regression, have been investigated on fabric CFRP specimens with locally introduced porosity and on unidirectional state-of-the-art aerospace CFRP material. The latter material contains porosity created by altering process parameters outside the specified range. The unidirectional material is available with two different thicknesses of the plies. With the smaller thickness no resonance effect occurred (as compared to the larger ply thickness and to the fabric material, where half the wavelength approximately equals the ply thickness) with the chosen inspection frequency.

It is not possible to find a 1:1 equivalent of the back-wall echo (*BWE*). Though partly large correlation coefficients $r \approx 0.9$ are achieved, the spread is too large for a 1:1 substitution of the *BWE*. Different sizes and shape of pores may influence the *BWE* [Hsu 2013] in a different way than the intermediate echoes [Jin et al. 2019]. The ability of the determined back-wall echo equivalents (BWE-equivalents) to enable a classification into good and porous was investigated then. The reference (labels) was determined from ultrasonic data via setting a threshold on the *BWE* (as in standard industrial inspection). Balanced accuracy — the mean of true positive rate (*TPR*) (detection rate of porosity) and true negative rate (*TNR*) (detection rate of non-porosity) — was used to determine the quality of classification. All values were determined in a cross-validation approach, tested on a specimen (or area of a specimen) different from those on which training was performed. Training is merely the determination of the decision threshold to distinguish prediction of positives and negatives for simple features; determination of parameters of RQA that lead to greatest balanced accuracy; or determination of weights for linear regression.

Table 5.18 summarises the best performing algorithms, further information is given in the following.

For fabric material with artificially introduced porosity fields, the relatively simple feature *QCD* delivers the best results. Second best are the RQA feature determinism *DET*, using Euclidean distance, and linear regression of the amplitude spectrum with Fourier transform after applying a Hann window on the intermediate echo time series. The effect for RQA with Euclidean distance is a reduction of signals in case of porosity.

For unidirectional material, porosity was introduced via alteration of production parameters outside the specified range and thus occurs throughout an entire specimen. Therefore virtual scans, combining a scan of a good and a porous specimen, are used for the evaluation. Excellent results are obtained on specimens with a CFRP ply thickness of 0.25 mm, which leads to a resonance effect similar to the fabric material.

Recurrence quantification analysis performs on these specimens better with angular distance, a metric recently proposed for RQA. The features *DET* and *RATIO* lead to balanced accuracies between $\approx$ 77% and $\approx$ 99%; the feature *COR* (correlation, newly proposed in this work) as a variant of trend *TND* provides balanced accuracies between 78.6% and $\approx$ 96%. The features pick up a reduction of recurrence rate with

Table 5.18: Balanced accuracies (mean of true positive rate and true negative rate) of best performing algorithms for classification into porous and non-porous out of the ultrasonic intermediate echo (backscattered) time series for the different investigated CFRP materials. Colours refer to mean balanced accuracies, cf. table C.1

algorithm	material & ply thickness		
	fabric material 913C-926-35%F	unidirectional material B	
	0.35 mm ply thickness	0.25 mm ply thickness	0.19 mm ply thickness
	porosity locally introduced by Freon capsules	porosity introduced by alteration of production parameters outside the specified range	
quartile coefficient of dispersion (QCD)	91.7%–96.1% mean 94.1%	47.5%–87.3% mean 66.3%	50.7%–68.0% mean 57.7%
linear regression on amplitude spectrum, Hann window applied before Fourier transform	83.4%–94.8% mean 89.7%	96.2%–99.5% mean 98.6%	81.7%–88.9% mean 85.2%
recurrence quantification analysis	84.2%–97.3% mean 91.3% feature DET, Euclidean distance	77.3%–97.5% mean 88.9% feature DET, angular distance	81.1%–91.4% mean 85% feature $RATIO$, Euclidean distance

a simultaneous increase of determinism, which are best reflected with DET as well as $RATIO$ and also shown in a changed distribution of recurrence via the diagonal lines for COR and TND.

Linear regression on the intermediate echo time series and on the amplitude spectrum, without using a Hann window, leads to excellent classification results in ten or nine scans, respectively, out of twelve scans, with balanced accuracies $\gtrsim 97\%$. However, severe outliers occur with these methods. Applying a Hann window prior to Fourier transform to reduce leakage leads to excellent results without outliers (table 5.18). As already identified in the pre-selection of the window of the amplitude spectrum, linear regression picks up a limited number of the large amplitude values plus amplitude values of frequencies from 7 MHz to 17 MHz, which are rather small in the original spectra. Linear regression of the amplitudes after applying a Hann window can furthermore be transferred to application on the first half of the intermediate echo gate with balanced accuracies still above 91%, and finally to another, thinner specimen

set with balanced accuracies $\geq 86.6\%$, in 4 out of 6 scans above 91%. This is stilll considered good taking into account that the reduced IE gate on the thick specimens is not the optimum training.

Results are clearly worse for unidirectional material with a lower ply thickness, for which no resonance effect occurs and intermediate echoes are more irregular. Best results are achieved with two features: RQA using Euclidean distance with the feature $RATIO$ and linear regression on the amplitude spectrum, Fourier transform after applying a Hann window (table 5.18). Linear regression again largely reduces the influence of frequencies with large amplitudes, here between 2 MHz and 7 MHz, and pronounces very low amplitudes of frequencies up to ≈ 18 MHz. The results for $RATIO$ are achieved with an embedding delay of approximately half the wavelength — a value that shall generally be avoided for a proper embedding. Recurrence reduces in case of porosity; DET slightly increases because more of the reduced number of recurrence points appears in diagonal lines, thus the effect on $RATIO = DET/RR$.

RQA parameters have been varied in a broad range, and those parameters, for which best classification results have been obtained (on the training set), have been chosen. Constraints on recurrence rate (at least upper value) and on range of the computed feature may be posed during such a data driven approach. Two new features, COR and L_{nor} have been proposed to cover a range of -1 or 0 to 1 to ease setting of such a range constraint, which finally however was not used for these features. The applied cross-validation approach leads to rather robust results in most cases, however partially quite sparse recurrence plots (low recurrence rate) or rather high recurrence rates resulted from the optimum parameters. A final check of recurrence plots for the determined parameters is recommended.

Several variants of RQA provide in this application on unidirectional material without resonance case no benefit; neither a comparison of two (embedded) time series in a cross, joint or the newly proposed difference recurrence plots, nor a recurrence threshold adapted in such a way that the recurrence rate is fixed.

Comparison with former research

Six publications — out of nine known publications for the characterisation of porosity in carbon fibre reinforced polymer (CFRP) with normal incidence ultrasonic testing without a back-wall echo — are comparable with the work in this thesis in terms of not relying on spatial averaging (averaging over several A-scans from different neighbouring positions), cf. section 4.1 and table 4.1.

In this thesis, none of the simple features (except quartile coefficient of dispersion (QCD)) from time or frequency domain can properly distinguish porosity from non-porous areas for the investigated materials. This is in contrast to the materials researched by Kim et al. [2013]. They achieved already with the maximum signal out of the intermediate echoes correlation coefficients > 0.83 for three out of four material sets (with a maximum improvement of 0.05 for the maximum approximate component of the discrete wavelet transform with Daubechies wavelets, level 5). The results are based on six data points each from another sample of different porosity content,

assuming an exponential dependence. Largely varying correlation coefficients occur in contrast in this thesis, e.g. for the unidirectional material, resonance case, from -0.07 to 0.94. These results are obtained for different areas of one porous and one non-porous sample, based on 3770 A-scans each, assuming a linear dependence between maximum signal and *BWE*. Kim et al. [2013] however used a higher frequency of 20 MHz for 2 mm thick specimens, unusual for industrial inspection of aerospace structures and not suitable to inspect larger thickness parts. Linear regression on the wavelet coefficients has been tested in Brandt et al. [2019] with worse results than for linear regression on the amplitude spectrum.

Karabutov & Podymova [2013] observed an increase of signals (they investigated the sum of the squared amplitude spectrum) with increasing porosity. This is not comparable to the materials in this thesis with clear inter-ply reflections due to the pure resin layers between the CFRP plies, for which rather a decrease of intermediate echoes takes place if porosity is present. Karabutov & Podymova [2013], Dominguez & Mascaro [2006] and Chen et al. [2016] all relied on the (partially local) amplitude spectrum. They agree in this respect with this thesis, in which however no simple features such as (shift of) the peak frequency [Dominguez & Mascaro 2006; Chen et al. 2016] lead to reasonable results. It is an open point whether this is due to the different investigated materials.

Jin et al. [2019] and Wang et al. [2019a] picked up the idea of using RQA from first publications within the scope of this thesis. Jin et al. [2019] used recurrence rate as feature; they emphasised the influence of different pore sizes and morphology on this feature as well as on the back-wall echo and claimed that recurrence rate shows less variations than *BWE* for porosity $\gtrsim 2\%$. Maximum and variance also seem to (positively) correlate with porosity but were not further investigated. Wang et al. [2019a] used recurrence rate, determinism and entropy and identified porosity in two of twelve unknown areas of a specimen using these features. No comparison with standard time domain or amplitude spectrum features is attempted.

Overall, the CFRP material seems to have a large influence on the effects of porosity on the intermediate echoes not only in respect of whether or not relevant inter-ply reflections (and thus potentially a resonance effect) occur.

5.8 Limitations and outlook

This work provides a contribution to the application of recurrence quantification analysis and Fourier analysis on complex signals from ultrasonic testing of two different carbon fibre reinforced polymer materials. The main results on unidirectional material have been obtained at one pair of 6 mm thick specimens good and porous each for two different ply thicknesses. The small number of specimens due to the challenge of producing porous specimens [cf. Grolemund & Tsai 1998] is a limitation, though the successful transfer on thinner specimens — at least for the ply thickness leading to a resonance effect — is a good indication regarding the robustness of the algorithm. As future work, the algorithm needs to be tested on further specimens, preferably

moderately porous. This work should include specimens of larger thickness; good results with Fourier analysis rely on frequencies $\gg 5\,\mathrm{MHz}$, which may be damped in specimens of larger thickness.

Though the investigated specimens contain moderate porosity, the back-wall echo heights that serve as the reference in this work are mostly clearly below the applied threshold of -50%. Future work regarding the robustness of the algorithm should also contain specimens with porosity leading to back-wall echoes around the threshold. This should be combined with automated inspection to ensure most reproducible measurements regarding coupling of the probe.

For the presented investigations, one phased array ultrasonic transducer with basic parameter setting was applied. Tests of different parameters and different ultrasonic transducers (including single-element probes) should be included into future work. This might include oblique incidence, deviating from the standard inspection of CFRP.

Results for inspections without a resonance effect — the more relevant case for industrial inspection — are not as good as the excellent results for the ones with resonance effect. Furthermore, this work contains investigations to use merely one feature at a time to classify into porous or non-porous. Different ideas on usage of multiple created features in a machine learning approach shall be pursued, e.g. linear regression on the amplitude spectrum (after applying a Hann window) plus RQA feature RATIO (using Euclidean distance), potentially including the original intermediate echo time series, using for example decision trees or neural networks.

Recurrence quantification analysis provides comparable results to Fourier analysis in the non-resonance case of inspection. Further methods based on delay embedding into a multidimensional state space may be tested on the current data: Carrión et al. [2014] proposed detecting determinism by continuity from Wayland et al. [1993] or by determining smoothness of trajectories in state space [cf. Jeong et al. 2002], cf. section 4.3. Non-linear prediction is another approach [Kantz & Schreiber 2005: sect. 4.2; Kennel & Isabelle 1992], where predictions are done based on similar states in state space; this might be performed for the application of finding a back-wall echo equivalent based on reference trajectories from non-porous areas because of non-stationarity.

Investigations on best features out of intermediate echo time series from ultrasonic testing have to be repeated for each different material — algorithms perform differently on different materials according to the literature research and partially confirmed for the two materials examined in this thesis. The most obvious effect is whether or not relevant inter-ply reflections occur and, if so, they lead to a resonance effect. The influence of the formation of pores (size, shape and location) might be clarified; a comparison with X-ray computed tomography may reveal different influences on *BWE* and features out of intermediate echoes. These can be a reason for large variations of these features for the same *BWE* height observed in this thesis.

6 Conclusion

This work is an approach to use recurrence quantification analysis (RQA) on complex signals of an engineering application in ultrasonic testing of carbon fibre reinforced polymers (CFRP), compared to applying Fourier analysis and several tools in time domain, especially using linear regression.

The established criterion of back-wall echo (BWE) reduction out of the ultrasonic pulse-echo A-scan for finding porosity shall be substituted to provide porosity detection in special situations, in which the BWE fails due to geometrical or structural reasons. Examples are stringers (stiffening elements of an aeroplane fuselage or wing) or bonded parts. A back-wall echo equivalent (BWE-equivalent) shall be generated out of the intermediate echoes of the A-scan (also known as backscatter), which are reflected from the inside of the CFRP part. The presented work exclusively compares BWE-equivalent with BWE for each A-scan, thus does not rely on spatial averaging over different A-scans (i.e. different measurement points) on a specimen.

Starting the work with correlation as quality measure, aiming at a 1:1 substitution of the BWE, it was identified that — despite a correlation coefficient partly $\gtrsim 0.9$ — the spread of results is too large. Reason might be a different effect of different pore sizes and shapes on BWE on the one hand and on intermediate echoes on the other hand. The work was subsequently directed into classification into porous and non-porous. To do so, cross-validation was applied throughout, i.e. training on two or three scans (each containing 3770 ultrasonic A-scans) or scan sets and testing on a different scan or scan set not seen in training, and rotating this to use every scan or scan set in test once. Balanced accuracy is used as quality measure, being the mean of true positive rate and true negative rate (rates of correctly predicted porosity and of correctly predicted non-porosity).

This thesis shows that the generation of a BWE-equivalent to classify into porous and non-porous areas of CFRP is in principle possible. Results differ between materials and largely depend on the ratio of wavelength λ to ply thickness t_p: for $t_p \approx \lambda/2$ a resonance effect occurs (from the reflections of the interfaces CFRP ply–pure resin layer–CFRP ply). Without resonance effect, results are promising but largely below the goal to achieve 90% balanced accuracy. RQA with the feature $RATIO$ and linear regression on the amplitude spectrum with applying a Hann window prior to Fourier transform are the two best performing features.

For industrial application to classify into porous and non-porous without access to a back-wall echo, an ultrasonic inspection with a frequency leading to a resonance effect is recommended. Linear regression on the amplitude spectrum provides excellent results with balanced accuracies $> 96\%$. The application of a Hann window on the intermediate echo time series prior to Fourier transform is essential.

Appendix A

Results on Fabric Specimens, Correlation

This appendix is a reproduction of the hypotheses of section 3.3 and of section 5 (evaluation results) of Brandt [2016], with adaptation of numbering and correction of some minor errors.

The hypotheses followed are:

1. Pores will lead to a perturbation of the otherwise almost periodic signal (which has decaying amplitude due to the attenuation of the CFRP, fig. A.1 right). The intermediate echoes obtained from the investigated samples (specimens) are almost periodic due to the layered structure of the material. Porosity leads to a perturbation of these signals, bringing higher randomness into the system by randomly distributed scatterers as assumed in Carrión et al. [2014], thus decreasing determinism.

2. Reflections at the pores will also lead to changes of signal amplitudes, hence comparing the time-localised energy of the signal between porous and non-porous structures will reveal a change of energy. This may be a reduction due to reflections not going back to the ultrasonic transducer or an increase due to additional reflections measured by the transducer.

The choice of optimal parameters for the recurrence quantification analysis (RQA) presented in this work is very much driven by the RQA itself (a similar approach to the one proposed in Zbilut & Webber [1992]). ϵ is varied between its extremes for both Euclidean and angular distance (Table 1).

To have a starting point for time delay τ, a quick check using the autocorrelation [Press et al. 2007: 602] is performed, which leads to values between 5 and 6 (time steps) for the different time series; quite similar to the 'good first guess' [Kantz & Schreiber 2005: 39] to take a quarter of the dominating wave length (if present). For the embedding dimension d, the method of false nearest neighbours [Webber & Marwan 2015: 5; Kantz & Schreiber 2005: 37f.] is merely checked on one representative time series using the CRP-toolbox for MATLAB [Marwan et al. 2007; Marwan 2015]. This check delivered a value of 4. Based on these values, τ and d are both varied in a range from 1 to 10.

The minimum line length l_{min} for the RQA feature determinism is initially set to 2 for Euclidean distance. For angular distance, $l_{min} = 2, 4, 8$ is applied.

Table A.1: Applied range of parameters for recurrence quantification analysis

RQA parameter	Min	Max	Steps/values used
Time delay τ	1	10	1
Embedding dimension d	1	10	1
Recurrence threshold ϵ for Euclidean distance	0.1	10	0.1, 0.2, 0.3, 0.4, 0.5, 0.8, 1, 1.2, 1.5, 2, 4, 6, 8, 10
Recurrence threshold ϵ for angular distance	0.01π	0.75π	0.01π, 0.025π, 0.05π, 0.075π, 0.1π, 0.25π, 0.5π, 0.75π
Minimum line length l_{min}	2	8	Euclidean distance: 2, in one case: ≤ 15 Angular distance: $2, 4, 8$

For 29 times 65 A-scans, thus 1885 time series, of one volume scan of sample M15, the RQA features RR and DET are generated for all parameter values shown in table A.1. The linear correlation coefficient r (or Pearson's r, [Press et al. 2007: 745]) is determined between RR and back-wall echo (BWE) (r_{RR}) and between DET and BWE (r_{DET}).

A.1 Euclidean distance

The time series in these investigations have a dominating frequency of approximately 5 MHz (fig. A.1). Theoretical considerations for Euclidean distance based on this fact lead to an assumption of optimum Recurrence Rate of approximately 0.055 (see Annex).

Extrema of correlation coefficients for all variations of RQA parameters and in the optimum Recurrence Rate range are determined (table A.2).

Three sets of parameters (row 3, 4 and 6) lead to Recurrence Rates extremely low or high, respectively. The minimum r_{RR} occurs for $RR = 0.004$. The according recurrence plots merely consist of the line of identity for non-porous areas. In the case of time series measured from an area with porosity, this line simply gets thicker on some sections. It is obvious that no serious RQA is possible with such spurious results.

The maximum correlation coefficients go along with a Recurrence Rate of almost 1 and with a range of Recurrence Rate R_{RR} (difference between minimum and maximum) of only 0.2% (the linear correlation coefficient makes no statement about the magnitude of the slope of the regression line [Press et al. 2007: 745]). The recurrence plots are almost black, recurrence and determinism decrease by a very small amount in the case of porosity. These results are not further considered either.

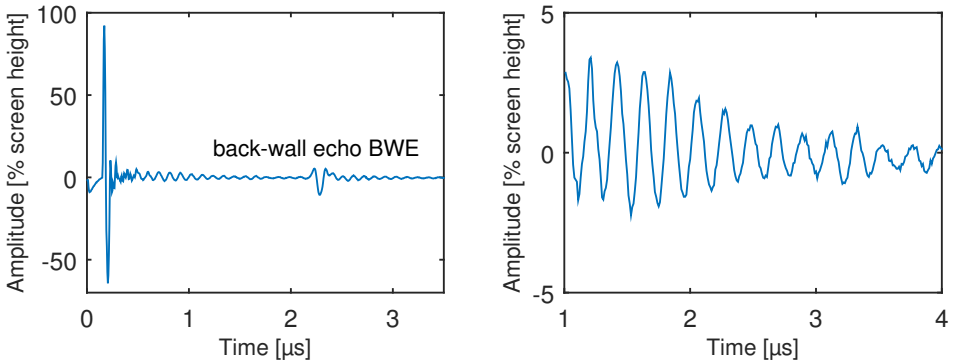


Figure A.1: Example of A-scan out of measurements of sample M15 (left); intermediate echoes from 0.99 µs to 3.99 µs magnified (right): the time series of 301 points used for RQA

Table A.2: Optimum coefficients of linear correlations between BWE-equivalent and *BWE* in dependence of RQA parameters for $l_{min} = 2$ for Euclidean distance

Row nr		correlation coefficient r	mean RR/R_{RR}	mean DET \ R_{DET}	ϵ	d	τ
1	RR driven	$r_{RR} = -0.60$	0.055/0.218	0.931\0.101	1	6	3
2	RR driven	$r_{DET} = -0.75$	0.055\0.147	0.729\0.263	0.3	2	1
3	min r_{RR}	$r_{RR} = -0.65$	0.004\0.005	0.975\0.267	0.4	10	6
4	max r_{RR}	$r_{RR} = 0.84$	0.996\0.048	0.9999\0.023	10	4	1
5	min r_{DET}	$r_{DET} = -0.81$	0.142\0.277	0.904\0.112	0.8	3	1
6	max r_{DET}	$r_{DET} = 0.78$	0.995\0.073	0.9998\0.0019	6	2	5

The minimum correlation coefficient for DET, reaching a large absolute value of -0.81, goes along with a reasonable recurrence rate of 0.142. The range of DET is still rather low (0.112). The recurrence plots for the situation without porosity (fig. A.2 left) as well as with porosity (fig. A.2 right) show lines perpendicular to the main diagonal, which may be caused by non-optimal embedding [Marwan 2011].

This choice of embedding parameters does not prove to be very robust: a reduction of the embedding dimension from 3 to 2 causes the correlation coefficient to change from -0.81 to -0.35, an increase of the time delay from 1 to 2 goes along with a change to $r_{DET} = -0.50$.

The optimum correlation coefficients taking optimum RR into account (row 1 and 2 in table A.2) do not lead to satisfying results either. The best r_{RR} takes a value of merely -0.60. The best $r_{DET} = -0.75$ is achieved via RPs again with a checkerboard-

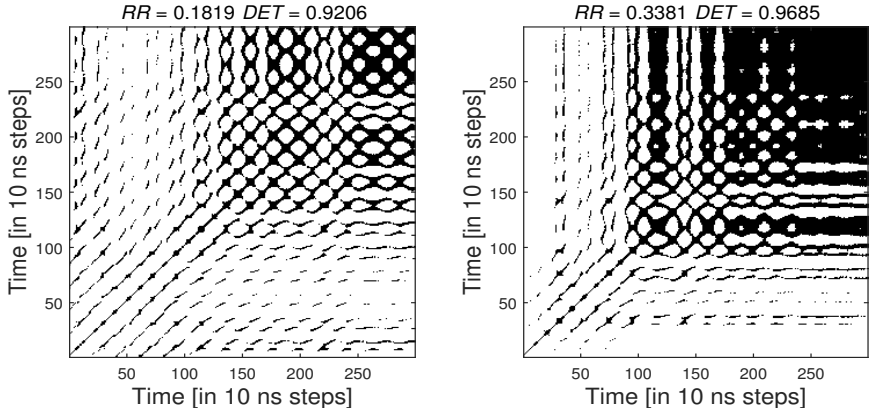


Figure A.2: Recurrence Plots for $\epsilon = 0.8$ (Euclidean distance), $d = 3$, $\tau = 1$ from first scan of M15, non-porous (left, A-scan with coordinates 2,20;) and porous area (right, A-scan with coordinates 2,50; M15-1-x2y50)

like structure as in figure A.2 and the results are again sensitive against change of embedding parameters as for the minimum r_{DET}.

Earlier work by the author with exemplary evaluations revealed that increasing l_{min} can lead to significant improvements. In Brandt & Maaß [2015], results for $\epsilon = 2$, $d = 3$, $\tau = 6$ and $l_{min} = 10$ are presented. The recurrence plots for this choice of recurrence threshold and embedding parameters show for the situation without as well as with porosity partly diagonal thick lines and, especially on the upper right, almost black regions (fig. A.3). The decaying amplitude of the original time series apparently leads to a trajectory which falls for later instants in time completely into the ball with radius ϵ in state space.

Porosity has a decreasing effect on the amplitudes of the intermediate echoes in the current scans. The pores reflect the ultrasonic wave in other directions than to the ultrasonic transducer and thus reduce the measured echoes coming from the CFRP material. This loss in ultrasonic energy is reflected in an increase of recurrence, especially in the upper right of the recurrence plots (fig. A.3 right). This increase is only severe for the area of highest porosity, from which this RP is taken: the correlation between recurrence rate and *BWE* over all the A-scans is merely $r_{RR} = -0.12$. The feature determinism however leads to $r_{DET} = -0.75$ with $l_{min} = 2$. A significantly better value -0.87 is achieved for a minimum line length from 7 to 15 (with the highest in absolute value, precision of 4 decimals, for $l_{min} = 8$). Without the non-recurrent gaps in the recurrence plots due to the reduced amplitude for the situation with porosity, considerable more diagonal lines are detected: the feature *DET* increases with higher porosity (fig. A.4). This effect may also be detected with the RQA features laminarity or trapping time, not investigated here, which are related to the occurrence of vertical lines in the recurrence plot [Webber & Marwan 2015: 16f.].

The range of *DET* equals 0.242 for $l_{min} = 8$, which is considered as sufficient. The choice of embedding parameters is insensitive to variations in comparison to results

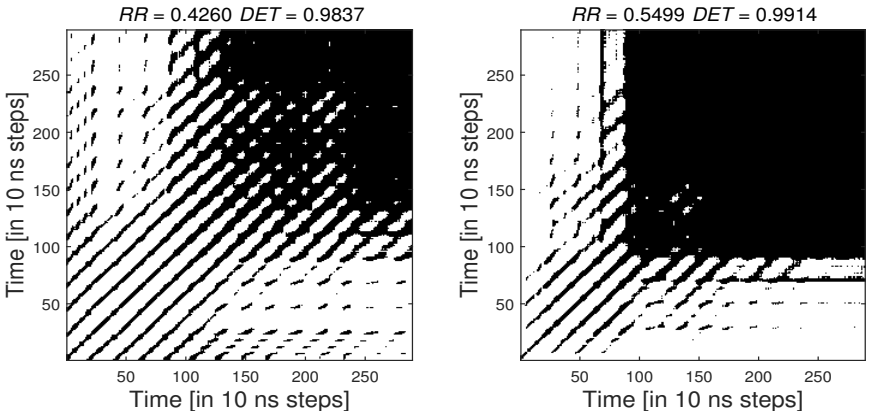


Figure A.3: Recurrence Plots for $\epsilon = 2$ (Euclidean distance), $d = 3$, $\tau = 6$ from first scan of M15, non-porous (left, A-scan with coordinates 2,20; M15-1x2y20) and porous area (right, A-scan with coordinates 2,50; M15-1x2y50)

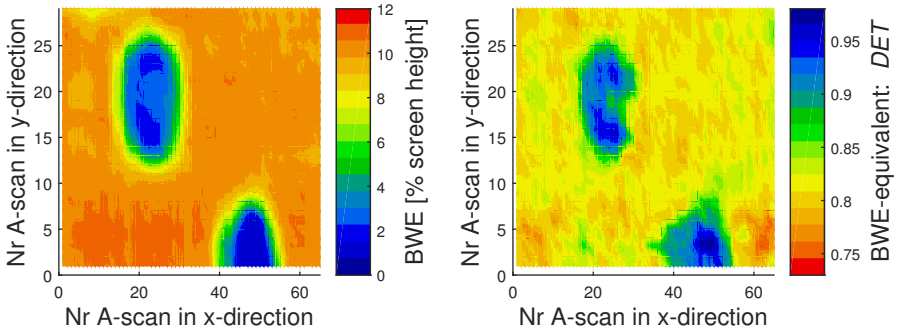


Figure A.4: C-scans of M15 first scan (M15-1): *BWE* (left hand side) and *DET* as BWE-equivalent with RQA parameters $\epsilon = 2$ (Euclidean distance), $d = 3$, $\tau = 6$ and $l_{min} = 8$ (right)

presented above: the correlation coefficient decreases to a minimum in absolute value of -0.70 for a variation of either d or τ by ± 1.

The optimum embedding parameters determined by autocorrelation and false nearest neighbours method ($d = 4$ and $\tau = 5$ to $\tau = 6$) are similar to the ones for which the above described good correlation on sample M15 was achieved.

This correlation on M15 with RQA parameters $d = 3$ and $\tau = 6$, recurrence threshold $\epsilon = 2$ (Euclidean distance) and $l_{min} = 8$ is checked on a second scan of M15 ($r_{DET} = -0.89$), on a scan of M13 ($r_{DET} = -0.69$) as well as of M14 ($r_{DET} = -0.76$). The poorest correlation coefficient of M13 may be related to the fact that the amount of porosity is not as high as in M15 (cf. C-scans fig. A.5; the range of the colour scale is identical to fig. A.4). The poorer scan quality may also negatively affect the correlation. However, even the vertical line obviously showing a slip-stick effect of scanning at approximately $x = 15$ is represented in the *DET*-C-scan.

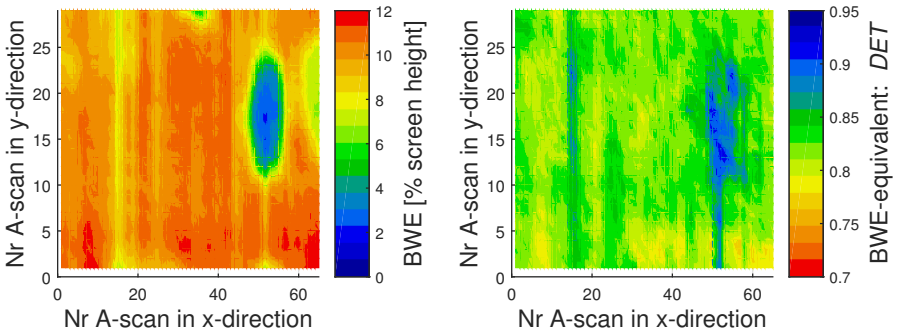


Figure A.5: C-scans of M13: *BWE* (left hand side) and *DET* as BWE-equivalent with RQA parameters $\epsilon = 2$ (Euclidean distance), $d = 3$, $\tau = 6$ and $l_{min} = 8$ (right)

The amplitude effect observed with recurrence quantification analysis turned out to be detectable with the quartile coefficient of dispersion (QCD) [Bonett 2006], too. For its determination, the time series is sorted in ascending order of amplitude values. It is then in the present case calculated as

$$QCD = \frac{x_{sort,226} - x_{sort,76}}{x_(sort, 226) - x_(sort, 76) - 2x_{min}}. \tag{A.1}$$

An offset was added to the time series to make all values nonnegative, the lowest value equalling zero. This is reflected by subtracting x_{min} twice from the denominator in equation (A.1).

Using QCD as back-wall echo equivalent, a correlation coefficient $r_{QCD} = 0.81$ for the first scan of M15 is achieved, slightly lower in absolute value than with recurrence quantification analysis. Refer to table A.3 in sect. A.2 for further results of r_{QCD} for the other scan of the same part and scans of the other investigated parts.

Whether this amplitude effect exists for other CFRP materials and 'natural' porosity, whether it can be detected by RQA and whether it can be determined using the quartile coefficient of dispersion or other simple means will be investigated in further steps of this ongoing research.

To summarise results for Euclidean distance, the RQA feature determinism has been found to provide a good back-wall echo equivalent for the investigated sample set. This is achieved with embedding parameters near to the ones determined with autocorrelation and false nearest neighbours method, leading to higher determinism for higher porosity. The cause is an amplitude effect, corresponding to hypothesis (2) (see app. A) with a loss of energy when porosity is present. This effect can also be observed with the simpler means of quartile coefficient of dispersion, with a correlation almost as good as with RQA.

Next, angular distance is investigated, a metric independent of amplitude effects by definition.

Table A.3: Optimum coefficients of linear correlations between BWE-equivalent and *BWE* in dependence of RQA parameters for $l_{min} = 2, 4, 8$ for angular distance

Extrema r	correlation coefficient r	Mean $RR\backslash R_{DET}$	mean DET $\backslash R_{DET}$	ϵ	d	τ	l_{min}
min r_{RR}	$r_{RR} = -0.64$	0.754\0.111	0.947\0.005	0.75π	8	1	-
max r_{RR}	$r_{RR} = 0.64$	0.101\0.063	0.551\0.102	0.1π	6	1	-
min r_{DET}	$r_{DET} = -0.55$	0.012\0.028	0.549\0.705	0.075π	10	7	8
max r_{DET}	$r_{DET} = 0.75$	0.504\0.201	0.910\0.338	0.5π	4	8	8

A.2 Angular distance

The varied RQA parameters for angular distance (table A.1) are checked on extrema of the coefficients of correlation of RR or DET, respectively, with the back-wall echo. With the experience from Euclidean distance, $l_{min} = 4$ and $l_{min} = 8$ are included into the investigated RQA parameters. Three of the four extrema of the coefficients are less than 0.7 in absolute value (table A.3) and are thus not further considered.

The maximum coefficient is $r_{DET} = 0.75$ for the correlation between DET and BWE. The range of determinism $R_{DET} = 0.338$, larger than any range of optimum correlation coefficients for Euclidean distance. The according recurrence plot derived from an A-scan from a non-porous area exhibits as expected no special dependence on time and consists almost exclusively of thick diagonals (Figure 9 left). For porosity, the recurrence plots loses partly its structure, which shows the reduced periodicity of the signal. The recurrence rate is almost the same as for the non-porous situation, but the determinism is significantly decreasing (Figure 9 right).

The according C-scan for determinism DET displays the correlation (fig. A.7), which is not as good as for the (in absolute value) greater minimum r_{DET} for Euclidean distance.

The optimum embedding dimension here equals the one determined by false nearest neighbours method, whereas the time delay is 2 to 3 time steps higher than the one determined with autocorrelation.

A variation of embedding parameters by ±1 decreases r_{DET} to a minimum of 0.65, which is considered as rather robust. The minimum line length has an influence of the same magnitude: $r_{DET} = 0.72$ and $r_{DET} = 0.69$ for $l_{min} = 4$ and $l_{min} = 2$, respectively.

For the optimum RQA parameters, which led to the maximum $r_{DET} = 0.75$ for M15, the correlation coefficient is determined at a second scan of M15 and at one scan

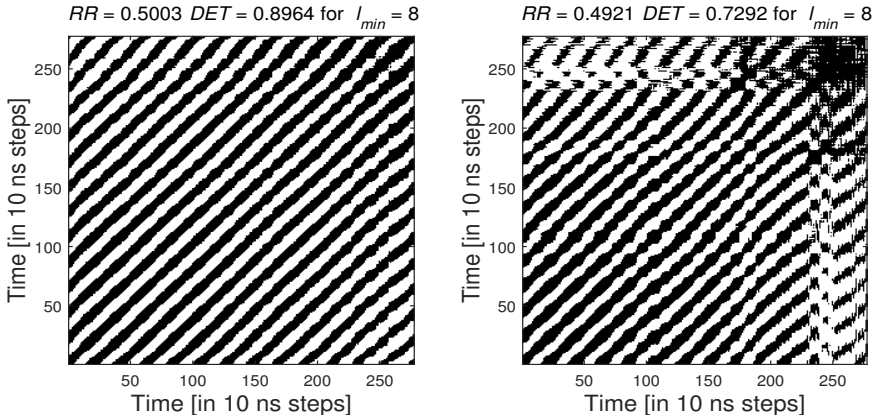


Figure A.6: Recurrence Plots for $\epsilon = 0.5\pi$ (angular distance), $d = 4$, $\tau = 8$ from first scan of M15, non-porous (left, A-scan with coordinates 2,20; M15-1x2y20) and porous area (right, A-scan with coordinates 2,50; M15-1x2y50)

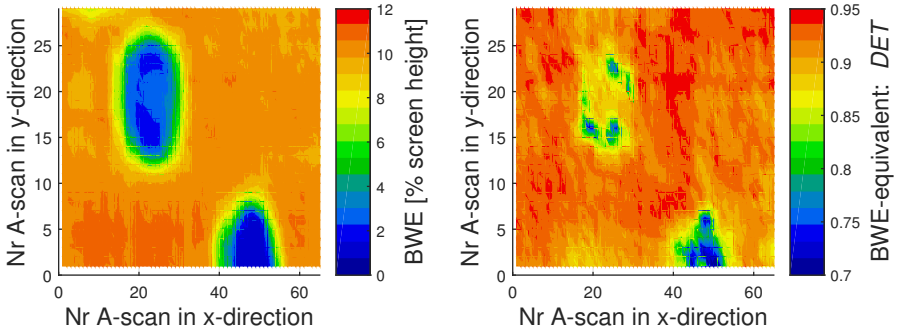


Figure A.7: C-scans of M15 first scan (M15-1): *BWE* (left hand side) and *DET* as BWE-equivalent with RQA parameters $\epsilon = 0.5\pi$ (angular distance), $d = 4$, $\tau = 8$ and $l_{min} = 8$ (right)

each of M13 and M14. The correlation coefficients are $r_{DET} = 0.73$, $r_{DET} = 0.57$ (C-scans in fig. A.7) and $r_{DET} = 0.69$, respectively.

Summarising for angular distance, the RQA feature determinism has been found as an appropriate back-wall echo equivalent. In contrast to the results with Euclidean distance, the feature decreases with increasing porosity. Amplitude effects are not taken into account with angular distance by definition, and the decreasing determinism shows a lower periodicity in the ultrasonic wave propagation, caused by the pores, corresponding to hypothesis (1).

The results with best parameters for angular distance in this section, originating from the reduction of determinism, as well as the results for best parameters with Euclidean distance and the hence derived quartile coefficient of dispersion (sect. A.1), both originating from amplitude effects, are summarised in table A.4.

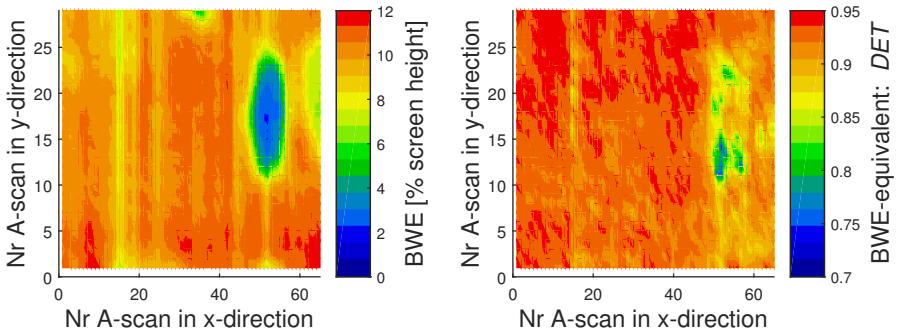


Figure A.8: C-scans of M13: *BWE* (left hand side) and *DET* as BWE-equivalent with RQA parameters $\epsilon = 0.5\pi$ (angular distance), $d = 4$, $\tau = 8$ and $l_{min} = 8$ (right)

Table A.4: Correlation coefficients r with the optimum RQA parameters for Euclidean distance and angular distance and quartile coefficient of dispersion

	M15-1	M15-2	M13	M14
$r_{DET,Euc}$ with $\epsilon = 2$, $d = 3$, $\tau = 6$, $l_{min} = 8$	−0.87	−0.89	−0.69	−0.76
r_{QCD} with time series offset	0.81	0.81	0.67	0.76
$r_{DET,ang}$ with $\epsilon = 0.5\pi$, $d = 4$, $\tau = 8$, $l_{min} = 8$	0.75	0.73	0.57	0.69

Annex: Considerations about optimum recurrence rate

Based on the dominating 5 MHz frequency in the time series (fig. A.1) the recurrence rate for an embedding of a sine wave with a wavelength of 20 time steps (corresponding to 5 MHz at a sampling rate of 100 MHz) is considered. With any recurrence threshold $\epsilon > 0$ (because of the ideal motion exactly equalling itself after one wave length) and less than a value depending on the amplitude of the sine wave and the embedding dimension (e.g. 0.3167 or 0.6258 times the amplitude for $d = 3$ or 8, respectively) this leads to a recurrence rate of 0.05, i.e. 5%. The according recurrence plot consists of diagonal lines [Marwan et al. 2007: figure 1A] occurring every 20 time steps. Thus it can be assumed that recurrence rates around 0.05 are appropriate for recurrence quantification analysis in the present case and may reveal changes in determinism caused by porosity.

The recurrence rate differs significantly with the embedding dimension, as the example with the ideal sine wave above shows (this is due to the larger distances between points in state space with a larger embedding dimension for one and the same time series).

Consequently, the recurrence thresholds, for which the mean recurrence rate (i.e.

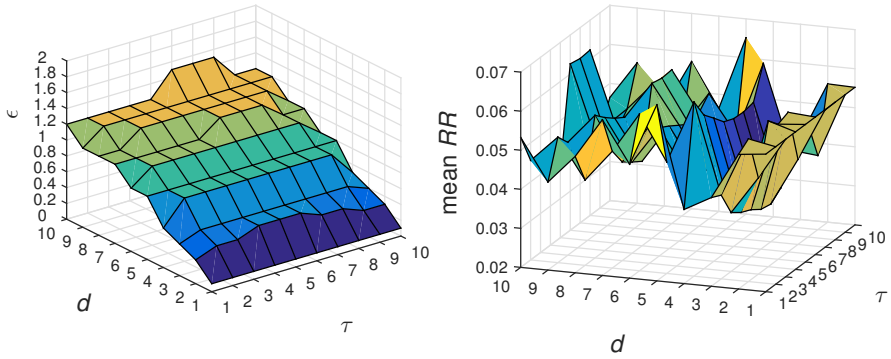


Figure A.9: Left: Recurrence thresholds for which mean values of the recurrence rate are nearest to 0.05 (Euclidean distance); right: mean values of the related recurrence rate

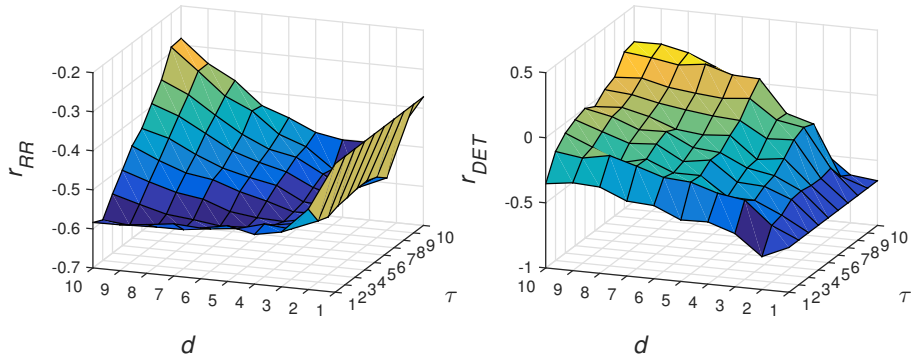


Figure A.10: Coefficients of linear correlation between RR and BWE (left) and between DET and BWE (right) for ϵ as in fig. A.9, i.e. for RR near to 0.05 (Euclidean distance)

the average of the recurrence rates of all measured time series per scan) is nearest to a value of 0.05 (fig. A.9 right), are determined (fig. A.9 left).

For the values of ϵ shown in fig. A.9, the coefficients of linear correlation between recurrence rate RR and back-wall echo BWE are shown in figure A.10 left hand side. The coefficients for the according correlation between determinism DET and BWE are displayed in figure A.10 right hand side.

Appendix B

Scan Data

Only portions of physical scans have been used for this thesis; they are given in table B.1 and B.2.

Table B.1: Cut-outs of physical scans of specimens from unidirectional specimens and fabric specimens, resonance case, used for evaluation in this thesis. Cf. table 5.1 and 5.2. One row in the table corresponds to one physical scan area covering the whole length of the specimen

physical scans	cut-out areas used for evaluation in scan (x) direction
M13	2–66
M14	16–80
M15-1/2	2–66
50-A/B1–3	6–70/126–190
56-A/B1–3	6–70/126–190
50-C1–3/α1–2	6–70/110–179
56-C1–3/α1–2	6–70/110–179
50-D1–3/β1–2	6–70/110–179
56-D1–3/β1–2	6–70/110–179
49-A/B1–3	6–70/126–190
55-A/B1–3	6–70/126–190

Table B.2: Cut-outs of physical scans of specimens from fabric specimens, non-resonance case, used for evaluation in this thesis. Cf. table 5.1 and 5.3. One row in the table corresponds to one physical scan area covering the whole length of the specimen

physical scans	cut-out areas used for evaluation in scan (x) direction
24-A/B1–3	6–70/126–190
38-A/B1–3	6–70/126–190
24-C/α1–3	6–70/130–199
38-C1–3	6–70
24-D1–3	6–70
38-D1–3	6–70
26-A/B1–3	6–70/126–190
36-A/B1–3	6–70/126–190

The areas of the defect-free part of the C-scans, on which the mean back-wall echo has been determined to assess porosity, are listed in table B.3.

Table B.3: Reference mean back-wall echo areas. Mean back-wall echoes minus 50% are the threshold below which back-wall echoes are labelled as representing a porous area.

specimens	reference area to determine mean back-wall echo
M13, M14, M15	x 28–37 y 3–10
50, 24, 49, 36	x 21–30 y 15–24

Appendix C

Result Tables

All classification results are given in this appendix in tabular form (plus the corresponding correlation coefficient for each optimum classification). Balanced accuracies are highlighted by colours (table C.1).

Table C.1: Colour coding for marking of balanced accuracies in tables

$acc_{bal} \geq 95\%$
$90\% \leq acc_{bal} < 95\%$
$85\% \leq acc_{bal} < 90\%$
$80\% \leq acc_{bal} < 85\%$
$acc_{bal} < 80\%$

C.1 Fabric specimens

C.1.1 Recurrence quantification analysis

Euclidean distance

Table C.2: Classification and correlation results, RQA, with recurrence rate (RR) as back-wall echo equivalent (BWE-equivalent), using Euclidean distance, for artificial porosity specimens M13–15. Optimum RQA parameters and decision threshold determined on 2 scans, results with these parameters on the remaining scan.

results of specimen/ scan	test	
	parameters determined on M14 & M15-1	results
M13	$d = 3$ $\tau = 3$ $\epsilon = 4.5$ $thres$= 0.8642	$r = -0.08$ $AUC = 0.525$ $acc_{bal} = 49.7\%$ $TPR = 84.6\%$ $FPR = 85.2\%$
	parameters determined on M13 & M15-1	results
M14	$d = 6$ $\tau = 2$ $\epsilon = 0.6$ $thres$= 0.0203	$r = -0.33$ $AUC = 0.763$ $acc_{bal} = 63.7\%$ $TPR = 45.5\%$ $FPR = 18.1\%$
	parameters determined on M13 & M14	results
M15-1	$d = 6$ $\tau = 1$ $\epsilon = 0.7$ $thres$= 0.0369	$r = -0.58$ $AUC = 0.879$ $acc_{bal} = 79.5\%$ $TPR = 79.8\%$ $FPR = 20.8\%$

Table C.3: Classification and correlation results, RQA, with DET as BWE-equivalent, using Euclidean distance, for artificial porosity specimens M13–15. Optimum RQA parameters and decision threshold determined on 2 scans, results with these parameters on the remaining scan.

results of specimen/ scan	test	
	parameters determined on M14 & M15-1	results
M13	$d = 3$ $\tau = 5$ $\epsilon = 1.7$ $l_{min} = 4$ $thres$= 0.9159	$r = -0.62$ AUC= 0.987 $acc_{bal} = 92.5\%$ TPR= 100.0% $FPR = 14.9\%$
	parameters determined on M13 & M15-1	results
M14	$d = 3$ $\tau = 5$ $\epsilon = 1.7$ $l_{min} = 6$ $thres$= 0.8662	$r = -0.61$ AUC= 0.953 $acc_{bal} = 84.2\%$ TPR= 69.7% $FPR = 1.3\%$
	parameters determined on M13 & M14	results
M15-1	$d = 3$ $\tau = 5$ $\epsilon = 1.7$ $l_{min} = 4$ $thres$= 0.9172	$r = -0.87$ AUC= 0.994 $acc_{bal} = 97.3\%$ TPR= 99.1% $FPR = 4.5\%$

Angular distance

Table C.4: Classification and correlation results, RQA, with *RR* as BWE-equivalent, using angular distance, for artificial porosity specimens M13–15. Optimum RQA parameters and decision threshold determined on 2 scans, results with these parameters on the remaining scan.

results of specimen/ scan	test	
	parameters determined on M14 & M15-1	results
M13	$d = 10$ $\tau = 7$ $\epsilon = 0.125\pi$ $l_{min} = 2$ $thres = 0.0435$	$r = 0.39$ $AUC = 0.932$ $acc_{bal} = 86.5\%$ $TPR = 83.3\%$ $FPR = 10.2\%$
	parameters determined on M13 & M15-1	results
M14	$d = 10$ $\tau = 6$ $\epsilon = 0.125\pi$ $l_{min} = 2$ $thres = 0.0433$	$r = 0.52$ $AUC = 0.913$ $acc_{bal} = 87.2\%$ $TPR = 80.8\%$ $FPR = 6.3\%$
	parameters determined on M13 & M14	results
M15-1	$d = 10$ $\tau = 9$ $\epsilon = 0.100\pi$ $l_{min} = 2$ $thres = 0.0249$	$r = 0.52$ $AUC = 0.927$ $acc_{bal} = 84.9\%$ $TPR = 86.5\%$ $FPR = 16.8\%$

Table C.5: Classification and correlation results, RQA, with DET as BWE-equivalent, using angular distance, for artificial porosity specimens M13–15. Optimum RQA parameters and decision threshold determined on 2 scans, results with these parameters on the remaining scan.

results of specimen/ scan	test	
	parameters determined on M14 & M15-1	results
M13	$d = 4$ $\tau = 8$ $\epsilon = 0.600\pi$ $l_{min} = 8$ $thres = 0.9163$	$r = 0.56$ $AUC = 0.955$ $acc_{bal} = 89.1\%$ $TPR = 88.5\%$ $FPR = 10.2\%$
	parameters determined on M13 & M15-1	results
M14	$d = 9$ $\tau = 8$ $\epsilon = 0.500\pi$ $l_{min} = 10$ $thres = 0.9112$	$r = 0.57$ $AUC = 0.914$ $acc_{bal} = 71.8\%$ $TPR = 44.4\%$ $FPR = 0.8\%$
	parameters determined on M13 & M14	results
M15-1	$d = 2$ $\tau = 5$ $\epsilon = 0.650\pi$ $l_{min} = 9$ $thres = 0.9232$	$r = 0.58$ $AUC = 0.901$ $acc_{bal} = 82.3\%$ $TPR = 84.3\%$ $FPR = 19.7\%$

C.1.2 Evaluation in time domain

Table C.6: Classification and correlation results with maximum value, variance and quartile coefficient of dispersion as BWE-equivalent for artificial porosity specimens M13–15. Decision threshold to achieve $TPR \approx TNR$ determined on 2 scans, results with this threshold on remaining scan.

results of specimen/ scan	x_{max}	s^2	QCD
M13	threshold to achieve $TPR \approx TNR$ on M14 & M15-1	threshold to achieve $TPR \approx TNR$ on M14 & M15-1	threshold to achieve $TPR \approx TNR$ on M14 & M15-1
	$thres = 3.8574$	$thres = 1.7071$	$thres = 0.2639$
	$r = -0.16$ $AUC = 0.770$ $acc_{bal} = 67.4\%$ $TPR = 43.6\%$ $FPR = 8.7\%$	$r = 0.16$ $AUC = 0.619$ $acc_{bal} = 52.9\%$ $TPR = 89.7\%$ $FPR = 83.9\%$	$r = 0.66$ $AUC = 0.982$ $acc_{bal} = 94.6\%$ $TPR = 98.7\%$ $FPR = 9.5\%$
M14	threshold to achieve $TPR \approx TNR$ on M13 & M15-1	threshold to achieve $TPR \approx TNR$ on M13 & M15-1	threshold to achieve $TPR \approx TNR$ on M13 & M15-1
	$thres = 3.7109$	$thres = 1.5662$	$thres = 0.2531$
	$r = -0.44$ $AUC = 0.896$ $acc_{bal} = 85.4\%$ $TPR = 82.8\%$ $FPR = 12.1\%$	$r = -0.21$ $AUC = 0.828$ $acc_{bal} = 75.3\%$ $TPR = 85.9\%$ $FPR = 35.3\%$	$r = 0.65$ $AUC = 0.963$ $acc_{bal} = 91.7\%$ $TPR = 87.9\%$ $FPR = 4.5\%$
M15-1	threshold to achieve $TPR \approx TNR$ on M13 & M14	threshold to achieve $TPR \approx TNR$ on M13 & M14	threshold to achieve $TPR \approx TNR$ on M13 & M14
	$thres = 3.4424$	$thres = 1.4766$	$thres = 0.2570$
	$r = -0.72$ $AUC = 0.961$ $acc_{bal} = 82.3\%$ $TPR = 99.1\%$ $FPR = 34.5\%$	$r = -0.57$ $AUC = 0.851$ $acc_{bal} = 70.0\%$ $TPR = 88.8\%$ $FPR = 48.7\%$	$r = 0.81$ $AUC = 0.990$ $acc_{bal} = 96.1\%$ $TPR = 95.1\%$ $FPR = 2.9\%$

Table C.7: Classification and correlation results with linear regression for artificial porosity specimens M13–15. Training (including decision threshold to achieve $TPR \approx TNR$) for each scan on two remaining scans.

results of specimen/ scan	linear regression
M13	training (including threshold to achieve $TPR \approx TNR$) on M14 & M15-1
	$thres = 7.4842$
	$r = 0.61$ $AUC = 0.943$ $acc_{bal} = 80.3\%$ $TPR = 100.0\%$ $FPR = 39.4\%$
M14	training (including threshold to achieve $TPR \approx TNR$) on M13 & M15-1
	$thres = 6.7144$
	$r = -0.01$ $AUC = 0.483$ $acc_{bal} = 47.4\%$ $TPR = 19.2\%$ $FPR = 24.4\%$
M15-1	training (including threshold to achieve $TPR \approx TNR$) on M13 & M14
	$thres = 7.8590$
	$r = 0.58$ $AUC = 0.833$ $acc_{bal} = 72.8\%$ $TPR = 49.3\%$ $FPR = 3.7\%$

C.1.3 Fourier analysis

Table C.8: Classification and correlation results with maximum frequency, left and right cutoff frequency as BWE-equivalent for artificial porosity specimens M13–15. Decision threshold to achieve $TPR \approx TNR$ determined on 2 scans, results with this threshold on remaining scan.

results of specimen/ scan	F_{peak}	F_{cl}	F_{cr}
M13	threshold to achieve $TPR \approx TNR$ on M14 & M15-1	threshold to achieve $TPR \approx TNR$ on M14 & M15-1	threshold to achieve $TPR \approx TNR$ on M14 & M15-1
	$thres = 5.0000$	$thres = 4.8469$	$thres = 5.3823$
	$r = 0.02$ $AUC = 0.501$ $acc_{bal} = 50.0\%$ $TPR = 0.0\%$ $FPR = 0.0\%$	$r = 0.03$ $AUC = 0.653$ $acc_{bal} = 56.8\%$ $TPR = 30.8\%$ $FPR = 17.2\%$	$r = 0.13$ $AUC = 0.557$ $acc_{bal} = 51.9\%$ $TPR = 98.7\%$ $FPR = 94.9\%$
M14	threshold to achieve $TPR \approx TNR$ on M13 & M15-1	threshold to achieve $TPR \approx TNR$ on M13 & M15-1	threshold to achieve $TPR \approx TNR$ on M13 & M15-1
	$thres = 5.3333$	$thres = 4.8549$	$thres = 5.2449$
	$r = -0.10$ $AUC = 0.527$ $acc_{bal} = 52.7\%$ $TPR = 70.7\%$ $FPR = 65.4\%$	$r = 0.18$ $AUC = 0.797$ $acc_{bal} = 66.3\%$ $TPR = 77.8\%$ $FPR = 45.2\%$	$r = -0.34$ $AUC = 0.937$ $acc_{bal} = 50.0\%$ $TPR = 100.0\%$ $FPR = 100.0\%$
M15-1	threshold to achieve $TPR \approx TNR$ on M13 & M14	threshold to achieve $TPR \approx TNR$ on M13 & M14	threshold to achieve $TPR \approx TNR$ on M13 & M14
	$thres = 5.3333$	$thres = 4.8535$	$thres = 5.4186$
	$r = -0.05$ $AUC = 0.502$ $acc_{bal} = 50.2\%$ $TPR = 0.4\%$ $FPR = 0.0\%$	$r = 0.74$ $AUC = 0.966$ $acc_{bal} = 89.7\%$ $TPR = 87.4\%$ $FPR = 8.1\%$	$r = -0.64$ $AUC = 0.935$ $acc_{bal} = 58.1\%$ $TPR = 16.6\%$ $FPR = 0.4\%$

Table C.9: Classification and correlation results with mean frequency, bandwidth and integral of bandwidth as BWE-equivalent for artificial porosity specimens M13–15. Decision threshold to achieve $TPR \approx TNR$ determined on 2 scans, results with this threshold on remaining scan.

results of specimen/ scan	F_{centre}	B_{3dB}	A_{bw3}
M13	threshold to achieve $TPR \approx TNR$ on M14 & M15-1	threshold to achieve $TPR \approx TNR$ on M14 & M15-1	threshold to achieve $TPR \approx TNR$ on M14 & M15-1
	$thres = 5.1005$	$thres = 0.5222$	$thres = 84.9820$
	$r = 0.15$ $AUC = 0.638$ $acc_{bal} = 53.0\%$ $TPR = 97.4\%$ $FPR = 91.4\%$	$r = 0.11$ $AUC = 0.485$ $acc_{bal} = 53.1\%$ $TPR = 98.7\%$ $FPR = 92.5\%$	$r = 0.27$ $AUC = 0.698$ $acc_{bal} = 50.2\%$ $TPR = 100.0\%$ $FPR = 99.5\%$
M14	threshold to achieve $TPR \approx TNR$ on M13 & M15-1	threshold to achieve $TPR \approx TNR$ on M13 & M15-1	threshold to achieve $TPR \approx TNR$ on M13 & M15-1
	$thres = 5.0501$	$thres = 0.3865$	$thres = 65.2753$
	$r = 0.04$ $AUC = 0.618$ $acc_{bal} = 49.8\%$ $TPR = 0.0\%$ $FPR = 0.4\%$	$r = -0.32$ $AUC = 0.906$ $acc_{bal} = 53.9\%$ $TPR = 100.0\%$ $FPR = 92.3\%$	$r = -0.33$ $AUC = 0.926$ $acc_{bal} = 55.7\%$ $TPR = 100.0\%$ $FPR = 88.7\%$
M15-1	threshold to achieve $TPR \approx TNR$ on M13 & M14	threshold to achieve $TPR \approx TNR$ on M13 & M14	threshold to achieve $TPR \approx TNR$ on M13 & M14
	$thres = 5.1041$	$thres = 0.4805$	$thres = 72.7403$
	$r = -0.55$ $AUC = 0.888$ $acc_{bal} = 62.6\%$ $TPR = 28.3\%$ $FPR = 3.1\%$	$r = -0.69$ $AUC = 0.954$ $acc_{bal} = 83.8\%$ $TPR = 71.3\%$ $FPR = 3.7\%$	$r = -0.71$ $AUC = 0.928$ $acc_{bal} = 86.1\%$ $TPR = 83.4\%$ $FPR = 11.3\%$

Table C.10: Classification and correlation results with linear regression on amplitude spectrum for artificial porosity specimens M13–15. Training (including decision threshold to achieve $TPR \approx TNR$) for each scan on two remaining scans.

results of specimen/ scan	linear regression on amplitude spectrum
M13	training (including threshold to achieve $TPR \approx TNR$) on M14 & M15-1
	$thres = 7.7095$
	$r = 0.44$ $AUC = 0.937$ $acc_{bal} = 86.4\%$ $TPR = 83.3\%$ $FPR = 10.5\%$
M14	training (including threshold to achieve $TPR \approx TNR$) on M13 & M15-1
	$thres = 7.2465$
	$r = 0.30$ $AUC = 0.746$ $acc_{bal} = 54.4\%$ $TPR = 94.9\%$ $FPR = 86.1\%$
M15-1	training (including threshold to achieve $TPR \approx TNR$) on M13 & M14
	$thres = 8.3488$
	$r = 0.39$ $AUC = 0.710$ $acc_{bal} = 64.9\%$ $TPR = 47.1\%$ $FPR = 17.3\%$

Table C.11: Classification and correlation results with linear regression on amplitude spectrum for artificial porosity specimens M13–15. Fourier transform performed after applying a Hann window on time series. Training (including decision threshold to achieve $TPR \approx TNR$) for each scan on two remaining scans.

results of specimen/ scan	linear regression on amplitude spectrum after applying a Hann window on time series
M13	training (including threshold to achieve $TPR \approx TNR$) on M14 & M15-1
	$thres = 7.5104$
	$r = 0.68$ $AUC = 0.987$ $acc_{bal} = 94.8\%$ $TPR = 96.2\%$ $FPR = 6.6\%$
M14	training (including threshold to achieve $TPR \approx TNR$) on M13 & M15-1
	$thres = 7.4677$
	$r = 0.61$ $AUC = 0.932$ $acc_{bal} = 83.4\%$ $TPR = 93.9\%$ $FPR = 27.1\%$
M15-1	training (including threshold to achieve $TPR \approx TNR$) on M13 & M14
	$thres = 7.9313$
	$r = 0.77$ $AUC = 0.973$ $acc_{bal} = 90.8\%$ $TPR = 87.0\%$ $FPR = 5.5\%$

C.2 Pre-selection of parameters for evaluation on unidirectional specimens

C.2.1 RQA parameters trend and correlation

Specimens from unidirectional material with natural porosity, resonant case

Table C.12: Classification and correlation training and test results, RQA feature *TND* as BWE-equivalent, using Euclidean distance, in dependence of different δ_M for natural porosity areas 5056α/β. Optimum RQA parameters and decision threshold determined on 5056β1/2 together, test results with these parameters on 5056α1 and 5056α2.

	training on 5056-β		test on 5056-α	
δ_M	parameters determined on 5056-β1/2	results on 5056-β1/2	results on 5056-α1	results on 5056-α2
2	$d = 7$ $\tau = 10$ $\epsilon = 14.0$ $thres = -3.8631$	$r = -0.68$ $AUC = 0.983$ $acc_{bal} = 94.0\%$ $TPR = 94.0\%$ $FPR = 6.0\%$	$r = -0.49$ $AUC = 0.794$ $acc_{bal} = 76.0\%$ $TPR = 67.5\%$ $FPR = 15.6\%$	$r = -0.67$ $AUC = 0.901$ $acc_{bal} = 84.6\%$ $TPR = 89.2\%$ $FPR = 20.1\%$
10	$d = 5$ $\tau = 10$ $\epsilon = 13.0$ $thres = -3.5047$	$r = -0.73$ $AUC = 0.983$ $acc_{bal} = 94.0\%$ $TPR = 94.0\%$ $FPR = 6.0\%$	$r = -0.51$ $AUC = 0.817$ $acc_{bal} = 78.2\%$ $TPR = 71.9\%$ $FPR = 15.6\%$	$r = -0.66$ $AUC = 0.891$ $acc_{bal} = 83.5\%$ $TPR = 90.5\%$ $FPR = 23.4\%$
25	$d = 5$ $\tau = 2$ $\epsilon = 15.5$ $thres = -2.5581$	$r = -0.79$ $AUC = 0.981$ $acc_{bal} = 94.0\%$ $TPR = 94.0\%$ $FPR = 6.0\%$	$r = -0.73$ $AUC = 0.916$ $acc_{bal} = 85.0\%$ $TPR = 85.3\%$ $FPR = 15.3\%$	$r = -0.64$ $AUC = 0.861$ $acc_{bal} = 80.0\%$ $TPR = 82.7\%$ $FPR = 22.8\%$
50	$d = 3$ $\tau = 3$ $\epsilon = 11.5$ $thres = -2.1758$	$r = -0.82$ $AUC = 0.984$ $acc_{bal} = 94.6\%$ $TPR = 94.6\%$ $FPR = 5.4\%$	$r = -0.74$ $AUC = 0.918$ $acc_{bal} = 86.2\%$ $TPR = 87.1\%$ $FPR = 14.7\%$	$r = -0.68$ $AUC = 0.876$ $acc_{bal} = 81.2\%$ $TPR = 84.5\%$ $FPR = 22.2\%$

Table C.13: Classification and correlation training and test results, RQA feature *TND* as BWE-equivalent, using Euclidean distance, in dependence of different δ_M for natural porosity areas 5056α/β. Optimum RQA parameters and decision threshold determined on 5056α1/2 together, test results with these parameters on 5056β1 and 5056β2.

	training on 5056-α		test on 5056-β	
δ_M	parameters determined on 5056-α1/2	results on 5056-α1/2	results on 5056-β1	results on 5056-β2
2	$d = 3$ $\tau = 5$ $\epsilon = 10.5$ $thres = -3.2170$	$r = -0.71$ $AUC = 0.910$ $acc_{bal} = 86.0\%$ $TPR = 86.0\%$ $FPR = 13.9\%$	$r = -0.68$ $AUC = 0.954$ $acc_{bal} = 89.5\%$ $TPR = 82.6\%$ $FPR = 3.7\%$	$r = -0.93$ $AUC = 1.000$ $acc_{bal} = 95.1\%$ $TPR = 99.9\%$ $FPR = 9.8\%$
10	$d = 5$ $\tau = 3$ $\epsilon = 13.0$ $thres = -3.2587$	$r = -0.71$ $AUC = 0.903$ $acc_{bal} = 86.2\%$ $TPR = 86.2\%$ $FPR = 13.8\%$	$r = -0.68$ $AUC = 0.938$ $acc_{bal} = 89.6\%$ $TPR = 82.5\%$ $FPR = 3.3\%$	$r = -0.92$ $AUC = 1.000$ $acc_{bal} = 95.6\%$ $TPR = 99.9\%$ $FPR = 8.8\%$
25	$d = 3$ $\tau = 5$ $\epsilon = 10.0$ $thres = -3.1263$	$r = -0.73$ $AUC = 0.906$ $acc_{bal} = 85.8\%$ $TPR = 85.8\%$ $FPR = 14.2\%$	$r = -0.74$ $AUC = 0.961$ $acc_{bal} = 91.3\%$ $TPR = 84.5\%$ $FPR = 2.0\%$	$r = -0.93$ $AUC = 1.000$ $acc_{bal} = 95.6\%$ $TPR = 100.0\%$ $FPR = 8.8\%$
50	$d = 4$ $\tau = 2$ $\epsilon = 13.5$ $thres = -2.0398$	$r = -0.71$ $AUC = 0.900$ $acc_{bal} = 83.8\%$ $TPR = 83.8\%$ $FPR = 16.2\%$	$r = -0.83$ $AUC = 0.982$ $acc_{bal} = 93.2\%$ $TPR = 87.4\%$ $FPR = 1.1\%$	$r = -0.94$ $AUC = 1.000$ $acc_{bal} = 96.0\%$ $TPR = 100.0\%$ $FPR = 8.0\%$

Table C.14: Classification and correlation training and test results, RQA feature *COR* as BWE-equivalent, using Euclidean distance, in dependence of different δ_M for natural porosity areas 5056α/β. Optimum RQA parameters and decision threshold determined on 5056β1/2 together, test results with these parameters on 5056α1 and 5056α2.

	training on 5056-β		test on 5056-α	
δ_M	parameters determined on 5056-β1/2	results on 5056-β1/2	results on 5056-α1	results on 5056-α2
2	$d = 4$ $\tau = 9$ $\epsilon = 15.0$ $thres = -0.8632$	$r = -0.61$ $AUC = 0.972$ $acc_{bal} = 92.2\%$ $TPR = 92.2\%$ $FPR = 7.8\%$	$r = -0.67$ $AUC = 0.908$ $acc_{bal} = 82.1\%$ $TPR = 84.1\%$ $FPR = 19.8\%$	$r = -0.46$ $AUC = 0.746$ $acc_{bal} = 67.9\%$ $TPR = 65.1\%$ $FPR = 29.3\%$
10	$d = 4$ $\tau = 9$ $\epsilon = 15.0$ $thres = -0.8579$	$r = -0.60$ $AUC = 0.971$ $acc_{bal} = 91.9\%$ $TPR = 91.9\%$ $FPR = 8.1\%$	$r = -0.67$ $AUC = 0.915$ $acc_{bal} = 83.0\%$ $TPR = 85.7\%$ $FPR = 19.6\%$	$r = -0.48$ $AUC = 0.761$ $acc_{bal} = 70.2\%$ $TPR = 69.4\%$ $FPR = 29.1\%$
25	$d = 4$ $\tau = 9$ $\epsilon = 14.0$ $thres = -0.8605$	$r = -0.59$ $AUC = 0.968$ $acc_{bal} = 91.1\%$ $TPR = 91.1\%$ $FPR = 8.9\%$	$r = -0.59$ $AUC = 0.880$ $acc_{bal} = 80.1\%$ $TPR = 79.9\%$ $FPR = 19.7\%$	$r = -0.46$ $AUC = 0.754$ $acc_{bal} = 69.6\%$ $TPR = 68.0\%$ $FPR = 28.9\%$
50	$d = 6$ $\tau = 5$ $\epsilon = 12.5$ $thres = -0.8937$	$r = -0.50$ $AUC = 0.938$ $acc_{bal} = 87.5\%$ $TPR = 87.5\%$ $FPR = 12.5\%$	$r = -0.36$ $AUC = 0.716$ $acc_{bal} = 67.8\%$ $TPR = 61.7\%$ $FPR = 26.0\%$	$r = -0.41$ $AUC = 0.709$ $acc_{bal} = 67.0\%$ $TPR = 65.8\%$ $FPR = 31.7\%$

Table C.15: Classification and correlation training and test results, RQA feature *COR* as BWE-equivalent, using Euclidean distance, in dependence of different δ_M for natural porosity areas 5056α/β. Optimum RQA parameters and decision threshold determined on 5056α1/2 together, test results with these parameters on 5056β1 and 5056β2.

	training on 5056-α		test on 5056-β	
δ_M	parameters determined on 5056-α1/2	results on	results on 5056-β1	results on 5056-β2
2	$d = 3$ $\tau = 3$ $\epsilon = 11.0$ $thres = -0.9119$	$r = -0.60$ $AUC = 0.859$ $acc_{bal} = 80.3\%$ $TPR = 80.3\%$ $FPR = 19.7\%$	$r = -0.69$ $AUC = 0.961$ $acc_{bal} = 88.9\%$ $TPR = 82.2\%$ $FPR = 4.5\%$	$r = -0.76$ $AUC = 0.999$ $acc_{bal} = 94.0\%$ $TPR = 99.9\%$ $FPR = 11.9\%$
10	$d = 3$ $\tau = 3$ $\epsilon = 11.0$ $thres = -0.9054$	$r = -0.59$ $AUC = 0.853$ $acc_{bal} = 79.7\%$ $TPR = 79.7\%$ $FPR = 20.3\%$	$r = -0.71$ $AUC = 0.959$ $acc_{bal} = 88.2\%$ $TPR = 81.5\%$ $FPR = 5.1\%$	$r = -0.71$ $AUC = 0.997$ $acc_{bal} = 93.9\%$ $TPR = 99.9\%$ $FPR = 12.1\%$
25	$d = 3$ $\tau = 3$ $\epsilon = 10.5$ $thres = -0.9022$	$r = -0.55$ $AUC = 0.839$ $acc_{bal} = 78.3\%$ $TPR = 78.3\%$ $FPR = 21.7\%$	$r = -0.68$ $AUC = 0.946$ $acc_{bal} = 86.3\%$ $TPR = 79.8\%$ $FPR = 7.1\%$	$r = -0.59$ $AUC = 0.991$ $acc_{bal} = 93.0\%$ $TPR = 98.8\%$ $FPR = 12.8\%$
50	$d = 3$ $\tau = 3$ $\epsilon = 9.5$ $thres = -0.9007$	$r = -0.41$ $AUC = 0.760$ $acc_{bal} = 69.7\%$ $TPR = 69.7\%$ $FPR = 30.3\%$	$r = -0.55$ $AUC = 0.895$ $acc_{bal} = 80.9\%$ $TPR = 74.2\%$ $FPR = 12.5\%$	$r = -0.54$ $AUC = 0.962$ $acc_{bal} = 88.4\%$ $TPR = 94.5\%$ $FPR = 17.8\%$

Table C.16: Classification and correlation training and test results, RQA feature *TND* as BWE-equivalent, using angular distance, in dependence of different δ_M for natural porosity areas 5056α/β. Optimum RQA parameters and decision threshold determined on 5056β1/2 together, test results with these parameters on 5056α1 and 5056α2.

	training on 5056-β		test on 5056-α	
δ_M	parameters determined on 5056-β1/2	results on 5056-β1/2	results on 5056-α1	results on 5056-α2
2	$d = 5$ $\tau = 9$ $\epsilon = 0.075\pi$ $thres = -0.0962$	$r = -0.71$ $AUC = 0.958$ $acc_{bal} = 90.6\%$ $TPR = 90.6\%$ $FPR = 9.4\%$	$r = -0.66$ $AUC = 0.910$ $acc_{bal} = 84.0\%$ $TPR = 78.9\%$ $FPR = 10.9\%$	$r = -0.72$ $AUC = 0.950$ $acc_{bal} = 89.9\%$ $TPR = 89.6\%$ $FPR = 9.8\%$
10	$d = 5$ $\tau = 9$ $\epsilon = 0.075\pi$ $thres = -0.1007$	$r = -0.70$ $AUC = 0.953$ $acc_{bal} = 90.2\%$ $TPR = 90.2\%$ $FPR = 9.8\%$	$r = -0.65$ $AUC = 0.905$ $acc_{bal} = 83.6\%$ $TPR = 78.5\%$ $FPR = 11.3\%$	$r = -0.72$ $AUC = 0.948$ $acc_{bal} = 89.9\%$ $TPR = 89.6\%$ $FPR = 9.9\%$
25	$d = 5$ $\tau = 9$ $\epsilon = 0.075\pi$ $thres = -0.1100$	$r = -0.68$ $AUC = 0.941$ $acc_{bal} = 88.9\%$ $TPR = 88.9\%$ $FPR = 11.1\%$	$r = -0.64$ $AUC = 0.895$ $acc_{bal} = 82.4\%$ $TPR = 76.9\%$ $FPR = 12.1\%$	$r = -0.72$ $AUC = 0.944$ $acc_{bal} = 89.5\%$ $TPR = 89.4\%$ $FPR = 10.3\%$
50	$d = 4$ $\tau = 9$ $\epsilon = 0.050\pi$ $thres = -0.0800$	$r = -0.68$ $AUC = 0.936$ $acc_{bal} = 87.4\%$ $TPR = 87.4\%$ $FPR = 12.6\%$	$r = -0.66$ $AUC = 0.891$ $acc_{bal} = 82.2\%$ $TPR = 76.9\%$ $FPR = 12.4\%$	$r = -0.69$ $AUC = 0.927$ $acc_{bal} = 86.6\%$ $TPR = 84.3\%$ $FPR = 11.2\%$

Table C.17: Classification and correlation training and test results, RQA feature *TND* as BWE-equivalent, using angular distance, in dependence of different δ_M for natural porosity areas 5056α/β. Optimum RQA parameters and decision threshold determined on 5056α1/2 together, test results with these parameters on 5056β1 and 5056β2.

	training on 5056-α		test on 5056-β	
δ_M	parameters determined on 5056-α1/2	results on 5056-α1/2	results on 5056-β1	results on 5056-β2
2	$d = 5$ $\tau = 9$ $\epsilon = 0.075\pi$ $thres = -0.1002$	$r = -0.69$ $AUC = 0.930$ $acc_{bal} = 86.8\%$ $TPR = 86.8\%$ $FPR = 13.3\%$	$r = -0.69$ $AUC = 0.939$ $acc_{bal} = 87.6\%$ $TPR = 89.5\%$ $FPR = 14.3\%$	$r = -0.74$ $AUC = 0.974$ $acc_{bal} = 92.8\%$ $TPR = 95.9\%$ $FPR = 10.2\%$
10	$d = 5$ $\tau = 9$ $\epsilon = 0.075\pi$ $thres = -0.1048$	$r = -0.68$ $AUC = 0.926$ $acc_{bal} = 86.3\%$ $TPR = 86.3\%$ $FPR = 13.7\%$	$r = -0.68$ $AUC = 0.935$ $acc_{bal} = 87.3\%$ $TPR = 89.3\%$ $FPR = 14.7\%$	$r = -0.73$ $AUC = 0.969$ $acc_{bal} = 92.2\%$ $TPR = 95.6\%$ $FPR = 11.2\%$
25	$d = 5$ $\tau = 9$ $\epsilon = 0.075\pi$ $thres = -0.1145$	$r = -0.68$ $AUC = 0.920$ $acc_{bal} = 85.7\%$ $TPR = 85.7\%$ $FPR = 14.3\%$	$r = -0.66$ $AUC = 0.922$ $acc_{bal} = 86.0\%$ $TPR = 88.5\%$ $FPR = 16.6\%$	$r = -0.72$ $AUC = 0.958$ $acc_{bal} = 91.4\%$ $TPR = 94.4\%$ $FPR = 11.7\%$
50	$d = 5$ $\tau = 9$ $\epsilon = 0.075\pi$ $thres = -0.1354$	$r = -0.68$ $AUC = 0.915$ $acc_{bal} = 85.4\%$ $TPR = 85.4\%$ $FPR = 14.6\%$	$r = -0.63$ $AUC = 0.902$ $acc_{bal} = 84.1\%$ $TPR = 88.5\%$ $FPR = 20.3\%$	$r = -0.69$ $AUC = 0.943$ $acc_{bal} = 89.8\%$ $TPR = 92.6\%$ $FPR = 12.9\%$

Table C.18: Classification and correlation training and test results, RQA feature *COR* as BWE-equivalent, using angular distance, in dependence of different δ_M for natural porosity areas 5056α/β. Optimum RQA parameters and decision threshold determined on 5056β1/2 together, test results with these parameters on 5056α1 and 5056α2.

	training on 5056-β		test on 5056-α	
δ_M	parameters determined on 5056-β1/2	results on 5056-β1/2	results on 5056-α1	results on 5056-α2
2	$d = 9$ $\tau = 9$ $\epsilon = 0.075\pi$ $thres = -0.1541$	$r = -0.81$ $AUC = 0.984$ $acc_{bal} = 95.4\%$ $TPR = 95.4\%$ $FPR = 4.6\%$	$r = -0.74$ $AUC = 0.952$ $acc_{bal} = 89.9\%$ $TPR = 89.9\%$ $FPR = 10.0\%$	$r = -0.75$ $AUC = 0.960$ $acc_{bal} = 89.6\%$ $TPR = 94.3\%$ $FPR = 15.2\%$
10	$d = 9$ $\tau = 9$ $\epsilon = 0.075\pi$ $thres = -0.1560$	$r = -0.80$ $AUC = 0.983$ $acc_{bal} = 95.1\%$ $TPR = 95.1\%$ $FPR = 4.9\%$	$r = -0.74$ $AUC = 0.952$ $acc_{bal} = 89.9\%$ $TPR = 89.8\%$ $FPR = 10.0\%$	$r = -0.75$ $AUC = 0.960$ $acc_{bal} = 89.6\%$ $TPR = 94.4\%$ $FPR = 15.2\%$
25	$d = 9$ $\tau = 9$ $\epsilon = 0.075\pi$ $thres = -0.1589$	$r = -0.79$ $AUC = 0.976$ $acc_{bal} = 94.5\%$ $TPR = 94.5\%$ $FPR = 5.5\%$	$r = -0.74$ $AUC = 0.951$ $acc_{bal} = 89.6\%$ $TPR = 88.9\%$ $FPR = 9.7\%$	$r = -0.75$ $AUC = 0.957$ $acc_{bal} = 89.7\%$ $TPR = 93.9\%$ $FPR = 14.5\%$
50	$d = 9$ $\tau = 9$ $\epsilon = 0.075\pi$ $thres = -0.1649$	$r = -0.76$ $AUC = 0.958$ $acc_{bal} = 92.6\%$ $TPR = 92.6\%$ $FPR = 7.4\%$	$r = -0.74$ $AUC = 0.946$ $acc_{bal} = 89.4\%$ $TPR = 87.3\%$ $FPR = 8.4\%$	$r = -0.75$ $AUC = 0.952$ $acc_{bal} = 89.9\%$ $TPR = 93.1\%$ $FPR = 13.4\%$

Table C.19: Classification and correlation training and test results, RQA feature *COR* as BWE-equivalent, using angular distance, in dependence of different δ_M for natural porosity areas 5056α/β. Optimum RQA parameters and decision threshold determined on 5056α1/2 together, test results with these parameters on 5056β1 and 5056β2.

	training on 5056-α		test on 5056-β	
δ_M	parameters determined on 5056-α1/2	results on	results on 5056-β1	results on 5056-β2
2	$d = 9$ $\tau = 9$ $\epsilon = 0.100\pi$ $thres = -0.1860$	$r = -0.78$ $AUC = 0.952$ $acc_{bal} = 90.8\%$ $TPR = 90.8\%$ $FPR = 9.2\%$	$r = -0.80$ $AUC = 0.961$ $acc_{bal} = 91.8\%$ $TPR = 91.0\%$ $FPR = 7.4\%$	$r = -0.87$ $AUC = 0.987$ $acc_{bal} = 97.6\%$ $TPR = 99.0\%$ $FPR = 3.8\%$
10	$d = 9$ $\tau = 9$ $\epsilon = 0.100\pi$ $thres = -0.1875$	$r = -0.78$ $AUC = 0.951$ $acc_{bal} = 90.7\%$ $TPR = 90.7\%$ $FPR = 9.3\%$	$r = -0.79$ $AUC = 0.954$ $acc_{bal} = 91.2\%$ $TPR = 91.0\%$ $FPR = 8.6\%$	$r = -0.86$ $AUC = 0.983$ $acc_{bal} = 97.3\%$ $TPR = 98.8\%$ $FPR = 4.2\%$
25	$d = 9$ $\tau = 9$ $\epsilon = 0.100\pi$ $thres = -0.1896$	$r = -0.78$ $AUC = 0.947$ $acc_{bal} = 90.5\%$ $TPR = 90.4\%$ $FPR = 9.5\%$	$r = -0.74$ $AUC = 0.932$ $acc_{bal} = 89.3\%$ $TPR = 90.7\%$ $FPR = 12.1\%$	$r = -0.83$ $AUC = 0.972$ $acc_{bal} = 95.8\%$ $TPR = 98.1\%$ $FPR = 6.5\%$
50	$d = 9$ $\tau = 9$ $\epsilon = 0.100\pi$ $thres = -0.1946$	$r = -0.78$ $AUC = 0.941$ $acc_{bal} = 90.0\%$ $TPR = 90.0\%$ $FPR = 10.0\%$	$r = -0.67$ $AUC = 0.893$ $acc_{bal} = 86.5\%$ $TPR = 90.2\%$ $FPR = 17.2\%$	$r = -0.79$ $AUC = 0.948$ $acc_{bal} = 93.4\%$ $TPR = 96.6\%$ $FPR = 9.7\%$

Specimens from unidirectional material with natural porosity, nonresonant case

Table C.20: Classification and correlation training and test results, RQA feature *TND* as BWE-equivalent, using Euclidean distance, in dependence of different δ_M for natural porosity areas 2438α/β. Optimum RQA parameters and decision threshold determined on 2438β1/2 together, test results with these parameters on 2438α1 and 2438α2.

	training on 2438-β		test on 2438-α	
δ_M	parameters determined on 2438-β1/2	results on 2438-β1/2	results on 2438-α1	results on 2438-α2
2	$d = 4$ $\tau = 9$ $\epsilon = 1.2$ $thres = -0.9216$	$r = -0.55$ $AUC = 0.860$ $acc_{bal} = 78.2\%$ $TPR = 78.2\%$ $FPR = 21.8\%$	$r = -0.61$ $AUC = 0.872$ $acc_{bal} = 78.0\%$ $TPR = 76.0\%$ $FPR = 20.1\%$	$r = -0.62$ $AUC = 0.893$ $acc_{bal} = 79.9\%$ $TPR = 71.7\%$ $FPR = 11.9\%$
10	$d = 4$ $\tau = 9$ $\epsilon = 1.1$ $thres = -0.7965$	$r = -0.54$ $AUC = 0.858$ $acc_{bal} = 77.9\%$ $TPR = 77.9\%$ $FPR = 22.1\%$	$r = -0.60$ $AUC = 0.866$ $acc_{bal} = 77.4\%$ $TPR = 75.4\%$ $FPR = 20.6\%$	$r = -0.60$ $AUC = 0.887$ $acc_{bal} = 79.4\%$ $TPR = 71.2\%$ $FPR = 12.4\%$
25	$d = 4$ $\tau = 9$ $\epsilon = 1.1$ $thres = -0.8621$	$r = -0.53$ $AUC = 0.853$ $acc_{bal} = 77.6\%$ $TPR = 77.6\%$ $FPR = 22.4\%$	$r = -0.58$ $AUC = 0.853$ $acc_{bal} = 76.2\%$ $TPR = 73.9\%$ $FPR = 21.6\%$	$r = -0.59$ $AUC = 0.877$ $acc_{bal} = 78.1\%$ $TPR = 69.4\%$ $FPR = 13.2\%$
50	$d = 4$ $\tau = 9$ $\epsilon = 1.0$ $thres = -0.8057$	$r = -0.50$ $AUC = 0.841$ $acc_{bal} = 76.1\%$ $TPR = 76.1\%$ $FPR = 23.9\%$	$r = -0.53$ $AUC = 0.825$ $acc_{bal} = 73.6\%$ $TPR = 70.9\%$ $FPR = 23.7\%$	$r = -0.54$ $AUC = 0.854$ $acc_{bal} = 76.1\%$ $TPR = 66.4\%$ $FPR = 14.3\%$

Table C.21: Classification and correlation training and test results, RQA feature *TND* as BWE-equivalent, using Euclidean distance, in dependence of different δ_M for natural porosity areas 2438α/β. Optimum RQA parameters and decision threshold determined on 2438α1/2 together, test results with these parameters on 2438β1 and 2438β2.

	training on 2438-α		test on 2438-β	
δ_M	parameters determined on 2438-α1/2	results on	results on 2438-β1	results on 2438-β2
2	$d = 5$ $\tau = 9$ $\epsilon = 6.0$ $thres = -3.4035$	$r = 0.72$ $AUC = 0.939$ $acc_{bal} = 89.7\%$ $TPR = 89.7\%$ $FPR = 10.3\%$	$r = -0.14$ $AUC = 0.666$ $acc_{bal} = 54.5\%$ $TPR = 99.4\%$ $FPR = 90.3\%$	$r = 0.63$ $AUC = 0.854$ $acc_{bal} = 75.9\%$ $TPR = 96.8\%$ $FPR = 44.9\%$
10	$d = 5$ $\tau = 9$ $\epsilon = 6.0$ $thres = -3.2787$	$r = 0.76$ $AUC = 0.944$ $acc_{bal} = 90.2\%$ $TPR = 90.2\%$ $FPR = 9.8\%$	$r = -0.14$ $AUC = 0.649$ $acc_{bal} = 54.3\%$ $TPR = 99.1\%$ $FPR = 90.5\%$	$r = 0.66$ $AUC = 0.854$ $acc_{bal} = 76.5\%$ $TPR = 96.9\%$ $FPR = 43.8\%$
25	$d = 5$ $\tau = 10$ $\epsilon = 5.5$ $thres = -3.2468$	$r = 0.82$ $AUC = 0.957$ $acc_{bal} = 91.1\%$ $TPR = 91.1\%$ $FPR = 8.9\%$	$r = -0.08$ $AUC = 0.578$ $acc_{bal} = 53.4\%$ $TPR = 99.2\%$ $FPR = 92.3\%$	$r = 0.71$ $AUC = 0.862$ $acc_{bal} = 79.5\%$ $TPR = 95.1\%$ $FPR = 36.0\%$
50	$d = 9$ $\tau = 5$ $\epsilon = 7.0$ $thres = -2.7162$	$r = 0.82$ $AUC = 0.960$ $acc_{bal} = 92.1\%$ $TPR = 92.1\%$ $FPR = 7.9\%$	$r = -0.07$ $AUC = 0.586$ $acc_{bal} = 47.1\%$ $TPR = 89.0\%$ $FPR = 94.7\%$	$r = 0.65$ $AUC = 0.848$ $acc_{bal} = 80.4\%$ $TPR = 88.4\%$ $FPR = 27.7\%$

Table C.22: Classification and correlation training and test results, RQA feature *COR* as BWE-equivalent, using Euclidean distance, in dependence of different δ_M for natural porosity areas 2438α/β. Optimum RQA parameters and decision threshold determined on 2438β1/2 together, test results with these parameters on 2438α1 and 2438α2.

	training on 2438-β		test on 2438-α	
δ_M	parameters determined on 2438-β1/2	results on 2438-β1/2	results on 2438-α1	results on 2438-α2
2	$d = 3$ $\tau = 9$ $\epsilon = 0.8$ $thres = -0.6309$	$r = -0.65$ $AUC = 0.891$ $acc_{bal} = 81.2\%$ $TPR = 81.2\%$ $FPR = 18.8\%$	$r = -0.72$ $AUC = 0.930$ $acc_{bal} = 84.4\%$ $TPR = 81.9\%$ $FPR = 13.2\%$	$r = -0.70$ $AUC = 0.921$ $acc_{bal} = 83.3\%$ $TPR = 78.4\%$ $FPR = 11.7\%$
10	$d = 3$ $\tau = 9$ $\epsilon = 0.8$ $thres = -0.6312$	$r = -0.65$ $AUC = 0.889$ $acc_{bal} = 81.1\%$ $TPR = 81.1\%$ $FPR = 18.9\%$	$r = -0.71$ $AUC = 0.923$ $acc_{bal} = 83.2\%$ $TPR = 80.4\%$ $FPR = 14.0\%$	$r = -0.69$ $AUC = 0.914$ $acc_{bal} = 82.3\%$ $TPR = 76.7\%$ $FPR = 12.1\%$
25	$d = 4$ $\tau = 9$ $\epsilon = 1.0$ $thres = -0.5871$	$r = -0.64$ $AUC = 0.884$ $acc_{bal} = 80.4\%$ $TPR = 80.4\%$ $FPR = 19.6\%$	$r = -0.69$ $AUC = 0.907$ $acc_{bal} = 81.9\%$ $TPR = 78.7\%$ $FPR = 15.0\%$	$r = -0.67$ $AUC = 0.903$ $acc_{bal} = 81.2\%$ $TPR = 74.3\%$ $FPR = 11.8\%$
50	$d = 3$ $\tau = 9$ $\epsilon = 0.6$ $thres = -0.4927$	$r = -0.60$ $AUC = 0.871$ $acc_{bal} = 79.2\%$ $TPR = 79.2\%$ $FPR = 20.8\%$	$r = -0.63$ $AUC = 0.873$ $acc_{bal} = 78.8\%$ $TPR = 74.1\%$ $FPR = 16.4\%$	$r = -0.60$ $AUC = 0.866$ $acc_{bal} = 77.6\%$ $TPR = 69.2\%$ $FPR = 13.9\%$

Table C.23: Classification and correlation training and test results, RQA feature *COR* as BWE-equivalent, using Euclidean distance, in dependence of different δ_M for natural porosity areas 2438α/β. Optimum RQA parameters and decision threshold determined on 2438α1/2 together, test results with these parameters on 2438β1 and 2438β2.

	training on 2438-α		test on 2438-β	
δ_M	parameters determined on 2438-α1/2	results on	results on 2438-β1	results on 2438-β2
2	$d = 5$ $\tau = 5$ $\epsilon = 5.5$ $thres = -0.9164$	$r = 0.45$ $AUC = 0.941$ $acc_{bal} = 90.8\%$ $TPR = 90.8\%$ $FPR = 9.2\%$	$r = -0.07$ $AUC = 0.690$ $acc_{bal} = 51.1\%$ $TPR = 94.5\%$ $FPR = 92.4\%$	$r = 0.44$ $AUC = 0.807$ $acc_{bal} = 77.6\%$ $TPR = 93.8\%$ $FPR = 38.5\%$
10	$d = 3$ $\tau = 10$ $\epsilon = 4.5$ $thres = -0.9057$	$r = 0.52$ $AUC = 0.950$ $acc_{bal} = 91.1\%$ $TPR = 91.1\%$ $FPR = 8.9\%$	$r = -0.06$ $AUC = 0.633$ $acc_{bal} = 51.4\%$ $TPR = 94.1\%$ $FPR = 91.4\%$	$r = 0.48$ $AUC = 0.801$ $acc_{bal} = 78.6\%$ $TPR = 94.1\%$ $FPR = 37.0\%$
25	$d = 5$ $\tau = 5$ $\epsilon = 5.5$ $thres = -0.8951$	$r = 0.48$ $AUC = 0.946$ $acc_{bal} = 91.3\%$ $TPR = 91.3\%$ $FPR = 8.7\%$	$r = -0.13$ $AUC = 0.694$ $acc_{bal} = 50.5\%$ $TPR = 93.8\%$ $FPR = 92.9\%$	$r = 0.47$ $AUC = 0.811$ $acc_{bal} = 77.7\%$ $TPR = 92.9\%$ $FPR = 37.5\%$
50	$d = 4$ $\tau = 10$ $\epsilon = 4.0$ $thres = -0.9128$	$r = 0.56$ $AUC = 0.935$ $acc_{bal} = 89.1\%$ $TPR = 89.1\%$ $FPR = 10.9\%$	$r = -0.27$ $AUC = 0.721$ $acc_{bal} = 53.4\%$ $TPR = 97.4\%$ $FPR = 90.7\%$	$r = 0.52$ $AUC = 0.811$ $acc_{bal} = 75.9\%$ $TPR = 88.1\%$ $FPR = 36.2\%$

Table C.24: Classification and correlation training and test results, RQA feature *TND* as BWE-equivalent, using angular distance, in dependence of different δ_M for natural porosity areas 2438α/β. Optimum RQA parameters and decision threshold determined on 2438β1/2 together, test results with these parameters on 2438α1 and 2438α2.

	training on 2438-β		test on 2438-α	
δ_M	parameters determined on 2438-β1/2	results on 2438-β1/2	results on 2438-α1	results on 2438-α2
2	$d = 7$ $\tau = 10$ $\epsilon = 0.050\pi$ $thres = -0.0188$	$r = 0.49$ $AUC = 0.769$ $acc_{bal} = 70.2\%$ $TPR = 70.2\%$ $FPR = 29.8\%$	$r = 0.75$ $AUC = 0.962$ $acc_{bal} = 89.4\%$ $TPR = 91.1\%$ $FPR = 12.3\%$	$r = 0.62$ $AUC = 0.868$ $acc_{bal} = 78.7\%$ $TPR = 85.5\%$ $FPR = 28.0\%$
10	$d = 7$ $\tau = 10$ $\epsilon = 0.050\pi$ $thres = -0.0201$	$r = 0.49$ $AUC = 0.769$ $acc_{bal} = 70.1\%$ $TPR = 70.1\%$ $FPR = 29.9\%$	$r = 0.75$ $AUC = 0.962$ $acc_{bal} = 89.5\%$ $TPR = 91.2\%$ $FPR = 12.3\%$	$r = 0.62$ $AUC = 0.868$ $acc_{bal} = 78.7\%$ $TPR = 85.5\%$ $FPR = 28.1\%$
25	$d = 7$ $\tau = 10$ $\epsilon = 0.050\pi$ $thres = -0.0229$	$r = 0.49$ $AUC = 0.769$ $acc_{bal} = 70.2\%$ $TPR = 70.2\%$ $FPR = 29.8\%$	$r = 0.76$ $AUC = 0.962$ $acc_{bal} = 89.6\%$ $TPR = 91.4\%$ $FPR = 12.2\%$	$r = 0.62$ $AUC = 0.868$ $acc_{bal} = 78.6\%$ $TPR = 85.5\%$ $FPR = 28.2\%$
50	$d = 7$ $\tau = 10$ $\epsilon = 0.050\pi$ $thres = -0.0291$	$r = 0.49$ $AUC = 0.769$ $acc_{bal} = 70.2\%$ $TPR = 70.2\%$ $FPR = 29.8\%$	$r = 0.76$ $AUC = 0.963$ $acc_{bal} = 89.5\%$ $TPR = 91.2\%$ $FPR = 12.2\%$	$r = 0.62$ $AUC = 0.869$ $acc_{bal} = 78.7\%$ $TPR = 85.4\%$ $FPR = 28.0\%$

Table C.25: Classification and correlation training and test results, RQA feature *TND* as BWE-equivalent, using angular distance, in dependence of different δ_M for natural porosity areas 2438α/β. Optimum RQA parameters and decision threshold determined on 2438α1/2 together, test results with these parameters on 2438β1 and 2438β2.

	training on 2438-α		test on 2438-β	
δ_M	parameters determined on 2438-α1/2	results on 2438-α1/2	results on 2438-β1	results on 2438-β2
2	$d = 6$ $\tau = 1$ $\epsilon = 0.025\pi$ $thres = -0.0222$	$r = 0.72$ $AUC = 0.960$ $acc_{bal} = 91.0\%$ $TPR = 91.0\%$ $FPR = 9.0\%$	$r = 0.13$ $AUC = 0.604$ $acc_{bal} = 49.0\%$ $TPR = 4.1\%$ $FPR = 6.2\%$	$r = 0.65$ $AUC = 0.894$ $acc_{bal} = 81.0\%$ $TPR = 91.5\%$ $FPR = 29.5\%$
10	$d = 6$ $\tau = 1$ $\epsilon = 0.025\pi$ $thres = -0.0227$	$r = 0.72$ $AUC = 0.960$ $acc_{bal} = 91.0\%$ $TPR = 91.0\%$ $FPR = 9.0\%$	$r = 0.12$ $AUC = 0.598$ $acc_{bal} = 48.7\%$ $TPR = 3.6\%$ $FPR = 6.2\%$	$r = 0.65$ $AUC = 0.892$ $acc_{bal} = 80.8\%$ $TPR = 91.4\%$ $FPR = 29.7\%$
25	$d = 6$ $\tau = 1$ $\epsilon = 0.025\pi$ $thres = -0.0238$	$r = 0.72$ $AUC = 0.960$ $acc_{bal} = 90.7\%$ $TPR = 90.7\%$ $FPR = 9.3\%$	$r = 0.10$ $AUC = 0.583$ $acc_{bal} = 48.4\%$ $TPR = 2.8\%$ $FPR = 6.0\%$	$r = 0.64$ $AUC = 0.888$ $acc_{bal} = 80.6\%$ $TPR = 90.9\%$ $FPR = 29.7\%$
50	$d = 9$ $\tau = 3$ $\epsilon = 0.075\pi$ $thres = -0.0468$	$r = 0.73$ $AUC = 0.960$ $acc_{bal} = 90.7\%$ $TPR = 90.7\%$ $FPR = 9.3\%$	$r = -0.02$ $AUC = 0.482$ $acc_{bal} = 52.3\%$ $TPR = 98.9\%$ $FPR = 94.4\%$	$r = 0.68$ $AUC = 0.917$ $acc_{bal} = 83.1\%$ $TPR = 96.7\%$ $FPR = 30.5\%$

Table C.26: Classification and correlation training and test results, RQA feature *COR* as BWE-equivalent, using angular distance, in dependence of different δ_M for natural porosity areas 2438α/β. Optimum RQA parameters and decision threshold determined on 2438β1/2 together, test results with these parameters on 2438α1 and 2438α2.

	training on 2438-β		test on 2438-α	
δ_M	parameters determined on 2438-β1/2	results on 2438-β1/2	results on 2438-α1	results on 2438-α2
2	$d = 9$ $\tau = 2$ $\epsilon = 0.025\pi$ $thres = -0.0004$	$r = 0.35$ $AUC = 0.649$ $acc_{bal} = 64.6\%$ $TPR = 48.9\%$ $FPR = 19.7\%$	$r = 0.67$ $AUC = 0.844$ $acc_{bal} = 82.6\%$ $TPR = 74.7\%$ $FPR = 9.4\%$	$r = 0.55$ $AUC = 0.781$ $acc_{bal} = 76.2\%$ $TPR = 71.3\%$ $FPR = 19.0\%$
10	$d = 8$ $\tau = 2$ $\epsilon = 0.025\pi$ $thres = -0.0307$	$r = 0.37$ $AUC = 0.671$ $acc_{bal} = 64.5\%$ $TPR = 64.5\%$ $FPR = 35.5\%$	$r = 0.72$ $AUC = 0.904$ $acc_{bal} = 84.1\%$ $TPR = 88.5\%$ $FPR = 20.3\%$	$r = 0.58$ $AUC = 0.819$ $acc_{bal} = 75.4\%$ $TPR = 84.8\%$ $FPR = 34.1\%$
25	$d = 4$ $\tau = 8$ $\epsilon = 0.700\pi$ $thres = -0.2326$	$r = -0.22$ $AUC = 0.670$ $acc_{bal} = 64.7\%$ $TPR = 64.7\%$ $FPR = 35.3\%$	$r = 0.12$ $AUC = 0.568$ $acc_{bal} = 54.6\%$ $TPR = 69.2\%$ $FPR = 60.1\%$	$r = 0.07$ $AUC = 0.519$ $acc_{bal} = 51.5\%$ $TPR = 69.4\%$ $FPR = 66.4\%$
50	$d = 8$ $\tau = 2$ $\epsilon = 0.025\pi$ $thres = -0.0177$	$r = 0.34$ $AUC = 0.658$ $acc_{bal} = 64.5\%$ $TPR = 64.5\%$ $FPR = 35.5\%$	$r = 0.71$ $AUC = 0.893$ $acc_{bal} = 84.0\%$ $TPR = 88.4\%$ $FPR = 20.4\%$	$r = 0.57$ $AUC = 0.811$ $acc_{bal} = 75.3\%$ $TPR = 84.6\%$ $FPR = 34.1\%$

Table C.27: Classification and correlation training and test results, RQA feature *COR* as BWE-equivalent, using angular distance, in dependence of different δ_M for natural porosity areas 2438α/β. Optimum RQA parameters and decision threshold determined on 2438α1/2 together, test results with these parameters on 2438β1 and 2438β2.

	training on 2438-α		test on 2438-β	
δ_M	parameters determined on 2438-α1/2	results on	results on 2438-β1	results on 2438-β2
2	$d = 9$ $\tau = 1$ $\epsilon = 0.025\pi$ $thres = -0.1539$	$r = 0.75$ $AUC = 0.929$ $acc_{bal} = 86.7\%$ $TPR = 86.7\%$ $FPR = 13.3\%$	$r = 0.07$ $AUC = 0.541$ $acc_{bal} = 49.0\%$ $TPR = 7.7\%$ $FPR = 9.7\%$	$r = 0.61$ $AUC = 0.831$ $acc_{bal} = 75.7\%$ $TPR = 81.5\%$ $FPR = 30.1\%$
10	$d = 9$ $\tau = 1$ $\epsilon = 0.025\pi$ $thres = -0.1532$	$r = 0.75$ $AUC = 0.927$ $acc_{bal} = 86.4\%$ $TPR = 86.4\%$ $FPR = 13.6\%$	$r = 0.06$ $AUC = 0.535$ $acc_{bal} = 49.0\%$ $TPR = 7.6\%$ $FPR = 9.7\%$	$r = 0.60$ $AUC = 0.826$ $acc_{bal} = 75.2\%$ $TPR = 80.6\%$ $FPR = 30.2\%$
25	$d = 9$ $\tau = 1$ $\epsilon = 0.025\pi$ $thres = -0.1515$	$r = 0.74$ $AUC = 0.923$ $acc_{bal} = 85.9\%$ $TPR = 85.9\%$ $FPR = 14.1\%$	$r = 0.04$ $AUC = 0.522$ $acc_{bal} = 48.6\%$ $TPR = 7.4\%$ $FPR = 10.1\%$	$r = 0.58$ $AUC = 0.816$ $acc_{bal} = 74.2\%$ $TPR = 79.0\%$ $FPR = 30.7\%$
50	$d = 9$ $\tau = 1$ $\epsilon = 0.025\pi$ $thres = -0.1465$	$r = 0.72$ $AUC = 0.911$ $acc_{bal} = 84.8\%$ $TPR = 84.8\%$ $FPR = 15.2\%$	$r = -0.00$ $AUC = 0.501$ $acc_{bal} = 51.9\%$ $TPR = 92.6\%$ $FPR = 88.8\%$	$r = 0.55$ $AUC = 0.797$ $acc_{bal} = 72.3\%$ $TPR = 76.4\%$ $FPR = 31.9\%$

C.3 Unidirectional specimens; resonance case

C.3.1 Recurrence quantification analysis

Euclidean distance

Table C.28: Classification and correlation results, RQA, with RR as BWE-equivalent, using Euclidean distance, for natural porosity area 5056A–D. Optimum RQA parameters and decision threshold determined on 3 areas (scan sets), test results with these parameters on scans of the remaining scan set.

results of specimen/ scan	parameters	test results		
	determined on 5056-B/C/D	scan A1	scan A2	scan A3
5056-A	$d = 5$ $\tau = 2$ $\epsilon = 16.5$ $thres = 0.7448$	$r = -0.55$ $AUC = 0.827$ $acc_{bal} = 76.0\%$ $TPR = 75.8\%$ $FPR = 23.7\%$	$r = -0.60$ $AUC = 0.847$ $acc_{bal} = 75.6\%$ $TPR = 77.0\%$ $FPR = 25.8\%$	$r = -0.39$ $AUC = 0.753$ $acc_{bal} = 72.8\%$ $TPR = 61.0\%$ $FPR = 15.4\%$
	determined on 5056-A/C/D	scan B1	scan B2	scan B3
5056-B	$d = 5$ $\tau = 2$ $\epsilon = 16.0$ $thres = 0.7320$	$r = -0.83$ $AUC = 0.979$ $acc_{bal} = 92.4\%$ $TPR = 95.2\%$ $FPR = 10.5\%$	$r = -0.93$ $AUC = 0.999$ $acc_{bal} = 98.2\%$ $TPR = 99.5\%$ $FPR = 3.1\%$	$r = -0.81$ $AUC = 0.957$ $acc_{bal} = 95.4\%$ $TPR = 91.5\%$ $FPR = 0.8\%$
	determined on 5056-A/B/D	scan C1	scan C2	scan C3
5056-C	$d = 5$ $\tau = 2$ $\epsilon = 16.0$ $thres = 0.7329$	$r = -0.60$ $AUC = 0.821$ $acc_{bal} = 73.8\%$ $TPR = 73.3\%$ $FPR = 25.7\%$	$r = -0.49$ $AUC = 0.773$ $acc_{bal} = 71.8\%$ $TPR = 63.8\%$ $FPR = 20.1\%$	$r = -0.56$ $AUC = 0.823$ $acc_{bal} = 73.0\%$ $TPR = 76.0\%$ $FPR = 30.0\%$
	determined on 5056-A/B/C	scan D1	scan D2	scan D3
5056-D	$d = 5$ $\tau = 2$ $\epsilon = 16.0$ $thres = 0.7301$	$r = -0.33$ $AUC = 0.695$ $acc_{bal} = 65.2\%$ $TPR = 55.1\%$ $FPR = 24.7\%$	$r = -0.67$ $AUC = 0.866$ $acc_{bal} = 75.1\%$ $TPR = 84.6\%$ $FPR = 34.3\%$	$r = -0.93$ $AUC = 0.999$ $acc_{bal} = 87.1\%$ $TPR = 100.0\%$ $FPR = 25.8\%$

Table C.29: Classification and correlation results, RQA, with *DIV* as BWE-equivalent, using Euclidean distance, for natural porosity area 5056A–D. Optimum RQA parameters and decision threshold determined on 3 areas (scan sets), test results with these parameters on scans of the remaining scan set.

results of specimen/ scan	parameters	test results		
	determined on 5056-B/C/D	scan A1	scan A2	scan A3
5056-A	$d = 10$	$r = 0.60$	$r = 0.68$	$r = 0.43$
	$\tau = 4$	$AUC = 0.830$	$AUC = 0.883$	$AUC = 0.755$
	$\epsilon = 11.0$	$acc_{bal} = 77.0\%$	$acc_{bal} = 80.0\%$	$acc_{bal} = 77.3\%$
		$TPR = 80.8\%$	$TPR = 84.3\%$	$TPR = 69.6\%$
	$thres = 0.0041$	$FPR = 26.8\%$	$FPR = 24.3\%$	$FPR = 15.0\%$
	determined on 5056-A/C/D	scan B1	scan B2	scan B3
5056-B	$d = 2$	$r = 0.82$	$r = 0.88$	$r = 0.80$
	$\tau = 5$	$AUC = 0.972$	$AUC = 0.998$	$AUC = 0.967$
	$\epsilon = 4.5$	$acc_{bal} = 92.0\%$	$acc_{bal} = 97.3\%$	$acc_{bal} = 94.8\%$
		$TPR = 96.6\%$	$TPR = 99.3\%$	$TPR = 92.3\%$
	$thres = 0.0039$	$FPR = 12.6\%$	$FPR = 4.8\%$	$FPR = 2.7\%$
	determined on 5056-A/B/D	scan C1	scan C2	scan C3
5056-C	$d = 2$	$r = 0.52$	$r = 0.39$	$r = 0.49$
	$\tau = 5$	$AUC = 0.790$	$AUC = 0.752$	$AUC = 0.800$
	$\epsilon = 4.5$	$acc_{bal} = 75.2\%$	$acc_{bal} = 79.5\%$	$acc_{bal} = 81.7\%$
		$TPR = 74.3\%$	$TPR = 78.3\%$	$TPR = 88.2\%$
	$thres = 0.0039$	$FPR = 23.8\%$	$FPR = 19.4\%$	$FPR = 24.7\%$
	determined on 5056-A/B/C	scan D1	scan D2	scan D3
5056-D	$d = 6$	$r = 0.40$	$r = 0.67$	$r = 0.93$
	$\tau = 4$	$AUC = 0.713$	$AUC = 0.851$	$AUC = 0.999$
	$\epsilon = 8.0$	$acc_{bal} = 58.3\%$	$acc_{bal} = 77.5\%$	$acc_{bal} = 98.0\%$
		$TPR = 21.2\%$	$TPR = 65.5\%$	$TPR = 99.9\%$
	$thres = 0.0039$	$FPR = 4.6\%$	$FPR = 10.6\%$	$FPR = 3.8\%$

Table C.30: Classification and correlation results, RQA, with *TND* as BWE-equivalent, using Euclidean distance, for natural porosity area 5056A–D. Optimum RQA parameters and decision threshold determined on 3 areas (scan sets), test results with these parameters on scans of the remaining scan set.

results of specimen/ scan	parameters	test results		
	determined on 5056-B/C/D	scan A1	scan A2	scan A3
5056-A	$d = 3$ $\tau = 4$ $\epsilon = 10.5$ $thres = -3.1223$	$r = -0.71$ $AUC = 0.923$ $acc_{bal} = 87.5\%$ $TPR = 86.1\%$ $FPR = 11.1\%$	$r = -0.73$ $AUC = 0.926$ $acc_{bal} = 88.9\%$ $TPR = 86.7\%$ $FPR = 8.9\%$	$r = -0.70$ $AUC = 0.969$ $acc_{bal} = 92.6\%$ $TPR = 91.9\%$ $FPR = 6.7\%$
	determined on 5056-A/C/D	scan B1	scan B2	scan B3
5056-B	$d = 3$ $\tau = 4$ $\epsilon = 11.0$ $thres = -3.0940$	$r = -0.67$ $AUC = 0.984$ $acc_{bal} = 92.8\%$ $TPR = 98.4\%$ $FPR = 12.9\%$	$r = -0.82$ $AUC = 0.999$ $acc_{bal} = 97.3\%$ $TPR = 99.9\%$ $FPR = 5.4\%$	$r = -0.79$ $AUC = 0.993$ $acc_{bal} = 96.0\%$ $TPR = 96.2\%$ $FPR = 4.1\%$
	determined on 5056-A/B/D	scan C1	scan C2	scan C3
5056-C	$d = 3$ $\tau = 4$ $\epsilon = 11.0$ $thres = -3.0796$	$r = -0.50$ $AUC = 0.814$ $acc_{bal} = 72.5\%$ $TPR = 60.7\%$ $FPR = 15.7\%$	$r = -0.62$ $AUC = 0.881$ $acc_{bal} = 82.6\%$ $TPR = 82.8\%$ $FPR = 17.7\%$	$r = -0.68$ $AUC = 0.926$ $acc_{bal} = 87.1\%$ $TPR = 90.9\%$ $FPR = 16.8\%$
	determined on 5056-A/B/C	scan D1	scan D2	scan D3
5056-D	$d = 10$ $\tau = 1$ $\epsilon = 20.0$ $thres = -3.0860$	$r = -0.55$ $AUC = 0.845$ $acc_{bal} = 77.1\%$ $TPR = 66.4\%$ $FPR = 12.3\%$	$r = -0.64$ $AUC = 0.926$ $acc_{bal} = 84.4\%$ $TPR = 88.4\%$ $FPR = 19.7\%$	$r = -0.90$ $AUC = 0.999$ $acc_{bal} = 95.3\%$ $TPR = 100.0\%$ $FPR = 9.3\%$

Table C.31: Classification and correlation results, RQA, with COR as BWE-equivalent, using Euclidean distance, for natural porosity area 5056A–D. Optimum RQA parameters and decision threshold determined on 3 areas (scan sets), test results with these parameters on scans of the remaining scan set.

results of specimen/ scan	parameters	test results		
	determined on 5056-B/C/D	scan A1	scan A2	scan A3
5056-A	$d = 4$	$r = -0.63$	$r = -0.66$	$r = -0.59$
	$\tau = 2$	$AUC = 0.886$	$AUC = 0.892$	$AUC = 0.872$
	$\epsilon = 13.5$	$acc_{bal} = 82.8\%$	$acc_{bal} = 83.8\%$	$acc_{bal} = 83.7\%$
		$TPR = 81.9\%$	$TPR = 82.7\%$	$TPR = 75.6\%$
	$thres = -0.9080$	$FPR = 16.2\%$	$FPR = 15.0\%$	$FPR = 8.2\%$
	determined on 5056-A/C/D	scan B1	scan B2	scan B3
5056-B	$d = 4$	$r = -0.53$	$r = -0.64$	$r = -0.60$
	$\tau = 2$	$AUC = 0.961$	$AUC = 0.998$	$AUC = 0.970$
	$\epsilon = 13.5$	$acc_{bal} = 89.8\%$	$acc_{bal} = 97.2\%$	$acc_{bal} = 94.7\%$
		$TPR = 94.0\%$	$TPR = 99.5\%$	$TPR = 93.8\%$
	$thres = -0.9098$	$FPR = 14.4\%$	$FPR = 5.0\%$	$FPR = 4.4\%$
	determined on 5056-A/B/D	scan C1	scan C2	scan C3
5056-C	$d = 2$	$r = -0.51$	$r = -0.57$	$r = -0.63$
	$\tau = 4$	$AUC = 0.826$	$AUC = 0.830$	$AUC = 0.887$
	$\epsilon = 10.0$	$acc_{bal} = 75.7\%$	$acc_{bal} = 78.8\%$	$acc_{bal} = 82.2\%$
		$TPR = 70.8\%$	$TPR = 76.4\%$	$TPR = 86.0\%$
	$thres = -0.9023$	$FPR = 19.4\%$	$FPR = 18.7\%$	$FPR = 21.6\%$
	determined on 5056-A/B/C	scan D1	scan D2	scan D3
5056-D	$d = 8$	$r = -0.47$	$r = -0.52$	$r = -0.75$
	$\tau = 1$	$AUC = 0.781$	$AUC = 0.901$	$AUC = 0.998$
	$\epsilon = 19.5$	$acc_{bal} = 71.3\%$	$acc_{bal} = 81.2\%$	$acc_{bal} = 91.6\%$
		$TPR = 59.2\%$	$TPR = 87.4\%$	$TPR = 100.0\%$
	$thres = -0.9045$	$FPR = 16.6\%$	$FPR = 25.0\%$	$FPR = 16.9\%$

Table C.32: Classification and correlation results, RQA, with *DET* as BWE-equivalent, using Euclidean distance, for natural porosity area 5056A–D. Optimum RQA parameters and decision threshold determined on 3 areas (scan sets), test results with these parameters on scans of the remaining scan set.

results of specimen/ scan	parameters	test results		
	determined on 5056-B/C/D	scan A1	scan A2	scan A3
5056-A	$d = 3$	$r = -0.80$	$r = -0.82$	$r = -0.64$
	$\tau = 7$	$AUC = 0.964$	$AUC = 0.978$	$AUC = 0.904$
	$\epsilon = 18.0$	$acc_{bal} = 88.9\%$	$acc_{bal} = 92.2\%$	$acc_{bal} = 85.4\%$
	$l_{min} = 5$	$TPR = 93.9\%$	$TPR = 93.2\%$	$TPR = 79.8\%$
	$thres = 0.9467$	$FPR = 16.1\%$	$FPR = 8.8\%$	$FPR = 9.1\%$
	determined on 5056-A/C/D	scan B1	scan B2	scan B3
5056-B	$d = 2$	$r = -0.85$	$r = -0.92$	$r = -0.87$
	$\tau = 6$	$AUC = 0.979$	$AUC = 0.998$	$AUC = 0.975$
	$\epsilon = 17.5$	$acc_{bal} = 90.1\%$	$acc_{bal} = 93.6\%$	$acc_{bal} = 91.6\%$
	$l_{min} = 8$	$TPR = 96.1\%$	$TPR = 99.6\%$	$TPR = 95.2\%$
	$thres = 0.9340$	$FPR = 15.8\%$	$FPR = 12.5\%$	$FPR = 12.0\%$
	determined on 5056-A/B/D	scan C1	scan C2	scan C3
5056-C	$d = 2$	$r = -0.66$	$r = -0.68$	$r = -0.71$
	$\tau = 7$	$AUC = 0.889$	$AUC = 0.899$	$AUC = 0.931$
	$\epsilon = 16.5$	$acc_{bal} = 81.2\%$	$acc_{bal} = 82.5\%$	$acc_{bal} = 85.5\%$
	$l_{min} = 6$	$TPR = 80.9\%$	$TPR = 80.5\%$	$TPR = 87.4\%$
	$thres = 0.9288$	$FPR = 18.5\%$	$FPR = 15.4\%$	$FPR = 16.5\%$
	determined on 5056-A/B/C	scan D1	scan D2	scan D3
5056-D	$d = 2$	$r = -0.52$	$r = -0.74$	$r = -0.95$
	$\tau = 7$	$AUC = 0.805$	$AUC = 0.920$	$AUC = 1.000$
	$\epsilon = 17.5$	$acc_{bal} = 72.7\%$	$acc_{bal} = 84.6\%$	$acc_{bal} = 94.6\%$
	$l_{min} = 7$	$TPR = 58.5\%$	$TPR = 87.0\%$	$TPR = 100.0\%$
	$thres = 0.9218$	$FPR = 13.1\%$	$FPR = 17.7\%$	$FPR = 10.9\%$

Table C.33: Classification and correlation results, RQA, with *RATIO* as BWE-equivalent, using Euclidean distance, for natural porosity area 5056A–D. Optimum RQA parameters and decision threshold determined on 3 areas (scan sets), test results with these parameters on scans of the remaining scan set.

results of specimen/ scan	parameters	test results		
	determined on 5056-B/C/D	scan A1	scan A2	scan A3
5056-A	$d = 5$	$r = 0.60$	$r = 0.66$	$r = 0.33$
	$\tau = 10$	$AUC = 0.888$	$AUC = 0.912$	$AUC = 0.779$
	$\epsilon = 14.5$	$acc_{bal} = 81.4\%$	$acc_{bal} = 82.9\%$	$acc_{bal} = 78.0\%$
	$l_{min} = 10$	$TPR = 85.0\%$	$TPR = 88.3\%$	$TPR = 70.2\%$
	$thres = 1.2258$	$FPR = 22.2\%$	$FPR = 22.4\%$	$FPR = 14.3\%$
	determined on 5056-A/C/D	scan B1	scan B2	scan B3
5056-B	$d = 4$	$r = -0.48$	$r = -0.45$	$r = -0.50$
	$\tau = 9$	$AUC = 0.824$	$AUC = 0.781$	$AUC = 0.868$
	$\epsilon = 3.0$	$acc_{bal} = 75.5\%$	$acc_{bal} = 73.1\%$	$acc_{bal} = 78.8\%$
	$l_{min} = 10$	$TPR = 79.6\%$	$TPR = 66.0\%$	$TPR = 78.8\%$
	$thres = 2.4862$	$FPR = 28.5\%$	$FPR = 19.9\%$	$FPR = 21.2\%$
	determined on 5056-A/B/D	scan C1	scan C2	scan C3
5056-C	$d = 5$	$r = 0.56$	$r = 0.52$	$r = 0.60$
	$\tau = 10$	$AUC = 0.815$	$AUC = 0.808$	$AUC = 0.866$
	$\epsilon = 14.5$	$acc_{bal} = 73.7\%$	$acc_{bal} = 75.9\%$	$acc_{bal} = 78.3\%$
	$l_{min} = 10$	$TPR = 66.6\%$	$TPR = 66.6\%$	$TPR = 78.3\%$
	$thres = 1.2207$	$FPR = 19.1\%$	$FPR = 14.8\%$	$FPR = 21.7\%$
	determined on 5056-A/B/C	scan D1	scan D2	scan D3
5056-D	$d = 5$	$r = 0.45$	$r = 0.71$	$r = 0.89$
	$\tau = 10$	$AUC = 0.777$	$AUC = 0.911$	$AUC = 1.000$
	$\epsilon = 14.5$	$acc_{bal} = 69.4\%$	$acc_{bal} = 80.7\%$	$acc_{bal} = 89.2\%$
	$l_{min} = 10$	$TPR = 59.8\%$	$TPR = 90.1\%$	$TPR = 100.0\%$
	$thres = 1.2283$	$FPR = 21.1\%$	$FPR = 28.6\%$	$FPR = 21.6\%$

Table C.34: Classification and correlation results, RQA, with L_{nor} as BWE-equivalent, using Euclidean distance, for natural porosity area 5056A–D. Optimum RQA parameters and decision threshold determined on 3 areas (scan sets), test results with these parameters on scans of the remaining scan set.

results of specimen/ scan	parameters	test results		
	determined on 5056-B/C/D	scan A1	scan A2	scan A3
5056-A	$d = 2$	$r = 0.68$	$r = 0.72$	$r = 0.51$
	$\tau = 1$	$AUC = 0.904$	$AUC = 0.922$	$AUC = 0.817$
	$\epsilon = 15.0$	$acc_{bal} = 81.9\%$	$acc_{bal} = 82.7\%$	$acc_{bal} = 78.2\%$
	$l_{min} = 2$	$TPR = 84.5\%$	$TPR = 85.8\%$	$TPR = 69.2\%$
	$thres = 0.0731$	$FPR = 20.6\%$	$FPR = 20.4\%$	$FPR = 12.7\%$
	determined on 5056-A/C/D	scan B1	scan B2	scan B3
5056-B	$d = 7$	$r = 0.72$	$r = 0.76$	$r = 0.77$
	$\tau = 9$	$AUC = 0.913$	$AUC = 0.939$	$AUC = 0.944$
	$\epsilon = 2.5$	$acc_{bal} = 83.6\%$	$acc_{bal} = 86.2\%$	$acc_{bal} = 87.3\%$
	$l_{min} = 2$	$TPR = 86.8\%$	$TPR = 90.6\%$	$TPR = 90.7\%$
	$thres = 0.3888$	$FPR = 19.6\%$	$FPR = 18.2\%$	$FPR = 16.1\%$
	determined on 5056-A/B/D	scan C1	scan C2	scan C3
5056-C	$d = 2$	$r = 0.70$	$r = 0.65$	$r = 0.71$
	$\tau = 7$	$AUC = 0.909$	$AUC = 0.890$	$AUC = 0.943$
	$\epsilon = 17.0$	$acc_{bal} = 84.7\%$	$acc_{bal} = 83.0\%$	$acc_{bal} = 86.6\%$
	$l_{min} = 3$	$TPR = 84.0\%$	$TPR = 78.3\%$	$TPR = 89.3\%$
	$thres = 0.0785$	$FPR = 14.5\%$	$FPR = 12.3\%$	$FPR = 16.0\%$
	determined on 5056-A/B/C	scan D1	scan D2	scan D3
5056-D	$d = 2$	$r = 0.54$	$r = 0.73$	$r = 0.95$
	$\tau = 6$	$AUC = 0.829$	$AUC = 0.915$	$AUC = 1.000$
	$\epsilon = 17.5$	$acc_{bal} = 72.4\%$	$acc_{bal} = 84.0\%$	$acc_{bal} = 96.5\%$
	$l_{min} = 2$	$TPR = 54.6\%$	$TPR = 83.6\%$	$TPR = 100.0\%$
	$thres = 0.0475$	$FPR = 9.8\%$	$FPR = 15.5\%$	$FPR = 6.9\%$

Table C.35: Classification and correlation results, RQA, with ENT as BWE-equivalent, using Euclidean distance, for natural porosity area 5056A–D. Optimum RQA parameters and decision threshold determined on 3 areas (scan sets), test results with these parameters on scans of the remaining scan set.

results of specimen/ scan	parameters	test results		
	determined on 5056-B/C/D	scan A1	scan A2	scan A3
5056-A	$d = 2$	$r = -0.79$	$r = -0.80$	$r = -0.77$
	$\tau = 8$	$AUC = 0.967$	$AUC = 0.965$	$AUC = 0.933$
	$\epsilon = 9.5$	$acc_{bal} = 89.7\%$	$acc_{bal} = 90.2\%$	$acc_{bal} = 88.8\%$
	$l_{min} = 2$	$TPR = 92.0\%$	$TPR = 91.7\%$	$TPR = 84.7\%$
	$thres = 2.4126$	$FPR = 12.6\%$	$FPR = 11.4\%$	$FPR = 7.1\%$
	determined on 5056-A/C/D	scan B1	scan B2	scan B3
5056-B	$d = 3$	$r = -0.78$	$r = -0.93$	$r = -0.82$
	$\tau = 7$	$AUC = 0.989$	$AUC = 0.999$	$AUC = 0.977$
	$\epsilon = 18.5$	$acc_{bal} = 90.5\%$	$acc_{bal} = 96.0\%$	$acc_{bal} = 92.7\%$
	$l_{min} = 2$	$TPR = 98.8\%$	$TPR = 99.9\%$	$TPR = 94.9\%$
	$thres = 2.9623$	$FPR = 17.8\%$	$FPR = 7.9\%$	$FPR = 9.4\%$
	determined on 5056-A/B/D	scan C1	scan C2	scan C3
5056-C	$d = 2$	$r = -0.62$	$r = -0.78$	$r = -0.82$
	$\tau = 8$	$AUC = 0.851$	$AUC = 0.940$	$AUC = 0.956$
	$\epsilon = 10.5$	$acc_{bal} = 76.0\%$	$acc_{bal} = 86.3\%$	$acc_{bal} = 87.8\%$
	$l_{min} = 2$	$TPR = 64.4\%$	$TPR = 85.9\%$	$TPR = 91.5\%$
	$thres = 2.4974$	$FPR = 12.4\%$	$FPR = 13.4\%$	$FPR = 16.0\%$
	determined on 5056-A/B/C	scan D1	scan D2	scan D3
5056-D	$d = 3$	$r = -0.53$	$r = -0.68$	$r = -0.96$
	$\tau = 7$	$AUC = 0.810$	$AUC = 0.888$	$AUC = 1.000$
	$\epsilon = 18.5$	$acc_{bal} = 73.6\%$	$acc_{bal} = 83.8\%$	$acc_{bal} = 97.7\%$
	$l_{min} = 2$	$TPR = 54.5\%$	$TPR = 81.6\%$	$TPR = 100.0\%$
	$thres = 2.9999$	$FPR = 7.4\%$	$FPR = 14.0\%$	$FPR = 4.5\%$

Angular distance

Table C.36: Classification and correlation results, RQA, with *RR* as BWE-equivalent, using angular distance, for natural porosity area 5056A–D. Optimum RQA parameters and decision threshold determined on 3 areas (scan sets), test results with these parameters on scans of the remaining scan set.

results of specimen/ scan	parameters	test results		
	determined on 5056-B/C/D	scan A1	scan A2	scan A3
5056-A	$d = 6$	$r = 0.62$	$r = 0.62$	$r = 0.63$
	$\tau = 9$	$AUC = 0.935$	$AUC = 0.930$	$AUC = 0.938$
	$\epsilon = 0.125\pi$	$acc_{bal} = 84.2\%$	$acc_{bal} = 81.2\%$	$acc_{bal} = 83.7\%$
		$TPR = 89.7\%$	$TPR = 91.0\%$	$TPR = 92.1\%$
	$thres = 0.0265$	$FPR = 21.2\%$	$FPR = 28.6\%$	$FPR = 24.7\%$
	determined on 5056-A/C/D	scan B1	scan B2	scan B3
5056-B	$d = 6$	$r = 0.82$	$r = 0.83$	$r = 0.81$
	$\tau = 9$	$AUC = 0.991$	$AUC = 0.998$	$AUC = 0.989$
	$\epsilon = 0.125\pi$	$acc_{bal} = 97.3\%$	$acc_{bal} = 98.1\%$	$acc_{bal} = 96.1\%$
		$TPR = 95.4\%$	$TPR = 97.2\%$	$TPR = 93.7\%$
	$thres = 0.0256$	$FPR = 0.8\%$	$FPR = 1.1\%$	$FPR = 1.5\%$
	determined on 5056-A/B/D	scan C1	scan C2	scan C3
5056-C	$d = 6$	$r = 0.46$	$r = 0.60$	$r = 0.63$
	$\tau = 9$	$AUC = 0.777$	$AUC = 0.904$	$AUC = 0.908$
	$\epsilon = 0.125\pi$	$acc_{bal} = 70.5\%$	$acc_{bal} = 82.2\%$	$acc_{bal} = 82.9\%$
		$TPR = 57.5\%$	$TPR = 82.1\%$	$TPR = 80.4\%$
	$thres = 0.0252$	$FPR = 16.5\%$	$FPR = 17.8\%$	$FPR = 14.7\%$
	determined on 5056-A/B/C	scan D1	scan D2	scan D3
5056-D	$d = 6$	$r = 0.63$	$r = 0.74$	$r = 0.77$
	$\tau = 9$	$AUC = 0.916$	$AUC = 0.971$	$AUC = 0.982$
	$\epsilon = 0.125\pi$	$acc_{bal} = 85.7\%$	$acc_{bal} = 91.0\%$	$acc_{bal} = 93.1\%$
		$TPR = 79.5\%$	$TPR = 93.0\%$	$TPR = 94.7\%$
	$thres = 0.0257$	$FPR = 8.2\%$	$FPR = 10.9\%$	$FPR = 8.6\%$

Table C.37: Classification and correlation results, RQA, with *DIV* as BWE-equivalent, using angular distance, for natural porosity area 5056A–D. Optimum RQA parameters and decision threshold determined on 3 areas (scan sets), test results with these parameters on scans of the remaining scan set.

results of specimen/ scan	parameters	test results		
	determined on 5056-B/C/D	scan A1	scan A2	scan A3
5056-A	$d = 10$	$r = 0.46$	$r = 0.46$	$r = 0.44$
	$\tau = 9$	$AUC = 0.831$	$AUC = 0.852$	$AUC = 0.849$
	$\epsilon = 0.175\pi$	$acc_{bal} = 75.1\%$	$acc_{bal} = 77.7\%$	$acc_{bal} = 79.8\%$
		$TPR = 68.2\%$	$TPR = 71.2\%$	$TPR = 76.6\%$
	$thres = 0.0104$	$FPR = 18.0\%$	$FPR = 15.8\%$	$FPR = 17.0\%$
	determined on 5056-A/C/D	scan B1	scan B2	scan B3
5056-B	$d = 8$	$r = 0.54$	$r = 0.54$	$r = 0.51$
	$\tau = 9$	$AUC = 0.870$	$AUC = 0.856$	$AUC = 0.853$
	$\epsilon = 0.275\pi$	$acc_{bal} = 84.5\%$	$acc_{bal} = 83.2\%$	$acc_{bal} = 82.8\%$
		$TPR = 86.8\%$	$TPR = 88.6\%$	$TPR = 86.5\%$
	$thres = 0.0044$	$FPR = 17.8\%$	$FPR = 22.1\%$	$FPR = 20.9\%$
	determined on 5056-A/B/D	scan C1	scan C2	scan C3
5056-C	$d = 10$	$r = 0.43$	$r = 0.44$	$r = 0.46$
	$\tau = 9$	$AUC = 0.815$	$AUC = 0.864$	$AUC = 0.854$
	$\epsilon = 0.175\pi$	$acc_{bal} = 70.9\%$	$acc_{bal} = 77.7\%$	$acc_{bal} = 76.2\%$
		$TPR = 60.4\%$	$TPR = 76.6\%$	$TPR = 74.6\%$
	$thres = 0.0104$	$FPR = 18.7\%$	$FPR = 21.1\%$	$FPR = 22.2\%$
	determined on 5056-A/B/C	scan D1	scan D2	scan D3
5056-D	$d = 10$	$r = 0.48$	$r = 0.52$	$r = 0.63$
	$\tau = 9$	$AUC = 0.867$	$AUC = 0.931$	$AUC = 0.988$
	$\epsilon = 0.175\pi$	$acc_{bal} = 78.5\%$	$acc_{bal} = 83.8\%$	$acc_{bal} = 90.3\%$
		$TPR = 74.7\%$	$TPR = 92.0\%$	$TPR = 99.4\%$
	$thres = 0.0108$	$FPR = 17.6\%$	$FPR = 24.3\%$	$FPR = 18.8\%$

Table C.38: Classification and correlation results, RQA, with *TND* as BWE-equivalent, using angular distance, for natural porosity area 5056A–D. Optimum RQA parameters and decision threshold determined on 3 areas (scan sets), test results with these parameters on scans of the remaining scan set.

results of specimen/ scan	parameters	test results		
	determined on 5056-B/C/D	scan A1	scan A2	scan A3
5056-A	$d = 5$ $\tau = 9$ $\epsilon = 0.075\pi$ $thres = -0.0944$	$r = -0.64$ $AUC = 0.941$ $acc_{bal} = 85.8\%$ $TPR = 93.5\%$ $FPR = 22.0\%$	$r = -0.62$ $AUC = 0.927$ $acc_{bal} = 82.5\%$ $TPR = 94.3\%$ $FPR = 29.2\%$	$r = -0.68$ $AUC = 0.949$ $acc_{bal} = 86.3\%$ $TPR = 95.4\%$ $FPR = 22.8\%$
	determined on 5056-A/C/D	scan B1	scan B2	scan B3
5056-B	$d = 5$ $\tau = 9$ $\epsilon = 0.075\pi$ $thres = -0.0911$	$r = -0.75$ $AUC = 0.969$ $acc_{bal} = 92.9\%$ $TPR = 91.4\%$ $FPR = 5.5\%$	$r = -0.77$ $AUC = 0.969$ $acc_{bal} = 92.8\%$ $TPR = 91.5\%$ $FPR = 5.9\%$	$r = -0.76$ $AUC = 0.968$ $acc_{bal} = 92.6\%$ $TPR = 91.1\%$ $FPR = 5.8\%$
	determined on 5056-A/B/D	scan C1	scan C2	scan C3
5056-C	$d = 5$ $\tau = 9$ $\epsilon = 0.075\pi$ $thres = -0.0902$	$r = -0.53$ $AUC = 0.817$ $acc_{bal} = 74.7\%$ $TPR = 65.0\%$ $FPR = 15.6\%$	$r = -0.62$ $AUC = 0.900$ $acc_{bal} = 84.3\%$ $TPR = 83.2\%$ $FPR = 14.7\%$	$r = -0.63$ $AUC = 0.906$ $acc_{bal} = 84.4\%$ $TPR = 81.5\%$ $FPR = 12.6\%$
	determined on 5056-A/B/C	scan D1	scan D2	scan D3
5056-D	$d = 5$ $\tau = 9$ $\epsilon = 0.075\pi$ $thres = -0.0905$	$r = -0.64$ $AUC = 0.919$ $acc_{bal} = 84.7\%$ $TPR = 78.7\%$ $FPR = 9.3\%$	$r = -0.69$ $AUC = 0.957$ $acc_{bal} = 88.8\%$ $TPR = 87.7\%$ $FPR = 10.1\%$	$r = -0.70$ $AUC = 0.955$ $acc_{bal} = 87.8\%$ $TPR = 84.9\%$ $FPR = 9.3\%$

Table C.39: Classification and correlation results, RQA, with *COR* as BWE-equivalent, using angular distance, for natural porosity area 5056A–D. Optimum RQA parameters and decision threshold determined on 3 areas (scan sets), test results with these parameters on scans of the remaining scan set.

results of specimen/ scan	parameters	test results		
	determined on 5056-B/C/D	scan A1	scan A2	scan A3
5056-A	$d = 10$	$r = -0.67$	$r = -0.65$	$r = -0.71$
	$\tau = 9$	$AUC = 0.938$	$AUC = 0.946$	$AUC = 0.960$
	$\epsilon = 0.100\pi$	$acc_{bal} = 86.3\%$	$acc_{bal} = 85.7\%$	$acc_{bal} = 89.2\%$
		$TPR = 93.1\%$	$TPR = 95.5\%$	$TPR = 93.9\%$
	$thres = -0.1738$	$FPR = 20.5\%$	$FPR = 24.2\%$	$FPR = 15.5\%$
	determined on 5056-A/C/D	scan B1	scan B2	scan B3
5056-B	$d = 10$	$r = -0.86$	$r = -0.85$	$r = -0.87$
	$\tau = 9$	$AUC = 0.983$	$AUC = 0.985$	$AUC = 0.989$
	$\epsilon = 0.100\pi$	$acc_{bal} = 95.6\%$	$acc_{bal} = 96.3\%$	$acc_{bal} = 96.0\%$
		$TPR = 96.3\%$	$TPR = 96.2\%$	$TPR = 94.6\%$
	$thres = -0.1698$	$FPR = 5.0\%$	$FPR = 3.7\%$	$FPR = 2.6\%$
	determined on 5056-A/B/D	scan C1	scan C2	scan C3
5056-C	$d = 9$	$r = -0.63$	$r = -0.73$	$r = -0.72$
	$\tau = 9$	$AUC = 0.869$	$AUC = 0.941$	$AUC = 0.935$
	$\epsilon = 0.100\pi$	$acc_{bal} = 78.6\%$	$acc_{bal} = 89.6\%$	$acc_{bal} = 87.9\%$
		$TPR = 63.1\%$	$TPR = 84.5\%$	$TPR = 83.4\%$
	$thres = -0.1720$	$FPR = 5.8\%$	$FPR = 5.3\%$	$FPR = 7.5\%$
	determined on 5056-A/B/C	scan D1	scan D2	scan D3
5056-D	$d = 10$	$r = -0.73$	$r = -0.73$	$r = -0.75$
	$\tau = 9$	$AUC = 0.953$	$AUC = 0.972$	$AUC = 0.973$
	$\epsilon = 0.100\pi$	$acc_{bal} = 88.7\%$	$acc_{bal} = 91.8\%$	$acc_{bal} = 92.8\%$
		$TPR = 85.9\%$	$TPR = 93.3\%$	$TPR = 94.7\%$
	$thres = -0.1701$	$FPR = 8.5\%$	$FPR = 9.8\%$	$FPR = 9.1\%$

Table C.40: Classification and correlation results, RQA, with *DET* as BWE-equivalent, using angular distance, for natural porosity area 5056A–D. Optimum RQA parameters and decision threshold determined on 3 areas (scan sets), test results with these parameters on scans of the remaining scan set.

results of specimen/ scan	parameters	test results		
	determined on 5056-B/C/D	scan A1	scan A2	scan A3
5056-A	$d = 10$	$r = -0.72$	$r = -0.73$	$r = -0.75$
	$\tau = 9$	$AUC = 0.926$	$AUC = 0.930$	$AUC = 0.936$
	$\epsilon = 0.125\pi$	$acc_{bal} = 85.6\%$	$acc_{bal} = 83.0\%$	$acc_{bal} = 84.5\%$
	$l_{min} = 6$	$TPR = 86.6\%$	$TPR = 88.9\%$	$TPR = 88.1\%$
	$thres = 0.4428$	$FPR = 15.3\%$	$FPR = 22.9\%$	$FPR = 19.2\%$
	determined on 5056-A/C/D	scan B1	scan B2	scan B3
5056-B	$d = 10$	$r = -0.90$	$r = -0.91$	$r = -0.89$
	$\tau = 9$	$AUC = 0.994$	$AUC = 0.998$	$AUC = 0.989$
	$\epsilon = 0.125\pi$	$acc_{bal} = 96.3\%$	$acc_{bal} = 97.5\%$	$acc_{bal} = 95.9\%$
	$l_{min} = 6$	$TPR = 96.5\%$	$TPR = 98.6\%$	$TPR = 95.0\%$
	$thres = 0.4483$	$FPR = 4.0\%$	$FPR = 3.6\%$	$FPR = 3.1\%$
	determined on 5056-A/B/D	scan C1	scan C2	scan C3
5056-C	$d = 10$	$r = -0.63$	$r = -0.78$	$r = -0.77$
	$\tau = 9$	$AUC = 0.857$	$AUC = 0.947$	$AUC = 0.940$
	$\epsilon = 0.125\pi$	$acc_{bal} = 77.3\%$	$acc_{bal} = 87.4\%$	$acc_{bal} = 86.6\%$
	$l_{min} = 6$	$TPR = 66.2\%$	$TPR = 86.4\%$	$TPR = 83.6\%$
	$thres = 0.4572$	$FPR = 11.7\%$	$FPR = 11.6\%$	$FPR = 10.3\%$
	determined on 5056-A/B/C	scan D1	scan D2	scan D3
5056-D	$d = 10$	$r = -0.76$	$r = -0.83$	$r = -0.88$
	$\tau = 9$	$AUC = 0.942$	$AUC = 0.978$	$AUC = 0.991$
	$\epsilon = 0.125\pi$	$acc_{bal} = 86.1\%$	$acc_{bal} = 91.7\%$	$acc_{bal} = 95.0\%$
	$l_{min} = 6$	$TPR = 81.2\%$	$TPR = 95.7\%$	$TPR = 98.0\%$
	$thres = 0.4474$	$FPR = 9.1\%$	$FPR = 12.3\%$	$FPR = 8.1\%$

Table C.41: Classification and correlation results, RQA, with *RATIO* as BWE-equivalent, using angular distance, for natural porosity area 5056A–D. Optimum RQA parameters and decision threshold determined on 3 areas (scan sets), test results with these parameters on scans of the remaining scan set.

results of specimen/ scan	parameters	test results		
	determined on 5056-B/C/D	scan A1	scan A2	scan A3
5056-A	$d = 10$	$r = -0.72$	$r = -0.73$	$r = -0.74$
	$\tau = 9$	$AUC = 0.943$	$AUC = 0.940$	$AUC = 0.945$
	$\epsilon = 0.150\pi$	$acc_{bal} = 86.1\%$	$acc_{bal} = 83.8\%$	$acc_{bal} = 85.1\%$
	$l_{min} = 8$	$TPR = 88.6\%$	$TPR = 91.1\%$	$TPR = 91.5\%$
	$thres = 8.9332$	$FPR = 16.3\%$	$FPR = 23.5\%$	$FPR = 21.3\%$
	determined on 5056-A/C/D	scan B1	scan B2	scan B3
5056-B	$d = 10$	$r = -0.87$	$r = -0.90$	$r = -0.87$
	$\tau = 9$	$AUC = 0.996$	$AUC = 0.999$	$AUC = 0.993$
	$\epsilon = 0.150\pi$	$acc_{bal} = 97.4\%$	$acc_{bal} = 99.3\%$	$acc_{bal} = 96.7\%$
	$l_{min} = 8$	$TPR = 96.7\%$	$TPR = 99.3\%$	$TPR = 94.1\%$
	$thres = 9.5903$	$FPR = 2.0\%$	$FPR = 0.7\%$	$FPR = 0.6\%$
	determined on 5056-A/B/D	scan C1	scan C2	scan C3
5056-C	$d = 9$	$r = -0.63$	$r = -0.76$	$r = -0.76$
	$\tau = 9$	$AUC = 0.856$	$AUC = 0.937$	$AUC = 0.936$
	$\epsilon = 0.125\pi$	$acc_{bal} = 77.1\%$	$acc_{bal} = 87.0\%$	$acc_{bal} = 86.6\%$
	$l_{min} = 7$	$TPR = 65.1\%$	$TPR = 85.6\%$	$TPR = 82.4\%$
	$thres = 15.5792$	$FPR = 10.9\%$	$FPR = 11.7\%$	$FPR = 9.1\%$
	determined on 5056-A/B/C	scan D1	scan D2	scan D3
5056-D	$d = 9$	$r = -0.75$	$r = -0.83$	$r = -0.86$
	$\tau = 9$	$AUC = 0.948$	$AUC = 0.982$	$AUC = 0.988$
	$\epsilon = 0.125\pi$	$acc_{bal} = 87.9\%$	$acc_{bal} = 91.8\%$	$acc_{bal} = 94.1\%$
	$l_{min} = 6$	$TPR = 83.6\%$	$TPR = 95.5\%$	$TPR = 96.6\%$
	$thres = 23.5143$	$FPR = 7.8\%$	$FPR = 11.9\%$	$FPR = 8.4\%$

Table C.42: Classification and correlation results, RQA, with L_{nor} as BWE-equivalent, using angular distance, for natural porosity area 5056A–D. Optimum RQA parameters and decision threshold determined on 3 areas (scan sets), test results with these parameters on scans of the remaining scan set.

results of specimen/ scan	parameters	test results		
	determined on 5056-B/C/D	scan A1	scan A2	scan A3
5056-A	$d = 10$	$r = 0.69$	$r = 0.69$	$r = 0.73$
	$\tau = 9$	$AUC = 0.907$	$AUC = 0.910$	$AUC = 0.928$
	$\epsilon = 0.150\pi$	$acc_{bal} = 82.3\%$	$acc_{bal} = 81.1\%$	$acc_{bal} = 84.1\%$
	$l_{min} = 4$	$TPR = 81.6\%$	$TPR = 84.8\%$	$TPR = 86.8\%$
	$thres = 0.5492$	$FPR = 17.0\%$	$FPR = 22.5\%$	$FPR = 18.7\%$
	determined on 5056-A/C/D	scan B1	scan B2	scan B3
5056-B	$d = 10$	$r = 0.88$	$r = 0.90$	$r = 0.87$
	$\tau = 9$	$AUC = 0.989$	$AUC = 0.996$	$AUC = 0.982$
	$\epsilon = 0.150\pi$	$acc_{bal} = 94.4\%$	$acc_{bal} = 96.4\%$	$acc_{bal} = 93.8\%$
	$l_{min} = 4$	$TPR = 94.9\%$	$TPR = 98.7\%$	$TPR = 93.5\%$
	$thres = 0.5480$	$FPR = 6.1\%$	$FPR = 5.8\%$	$FPR = 5.9\%$
	determined on 5056-A/B/D	scan C1	scan C2	scan C3
5056-C	$d = 10$	$r = 0.60$	$r = 0.75$	$r = 0.75$
	$\tau = 9$	$AUC = 0.832$	$AUC = 0.929$	$AUC = 0.928$
	$\epsilon = 0.125\pi$	$acc_{bal} = 75.4\%$	$acc_{bal} = 84.5\%$	$acc_{bal} = 84.9\%$
	$l_{min} = 2$	$TPR = 64.1\%$	$TPR = 83.9\%$	$TPR = 82.5\%$
	$thres = 0.4460$	$FPR = 13.3\%$	$FPR = 14.8\%$	$FPR = 12.6\%$
	determined on 5056-A/B/C	scan D1	scan D2	scan D3
5056-D	$d = 10$	$r = 0.73$	$r = 0.82$	$r = 0.90$
	$\tau = 9$	$AUC = 0.929$	$AUC = 0.973$	$AUC = 0.995$
	$\epsilon = 0.150\pi$	$acc_{bal} = 84.9\%$	$acc_{bal} = 91.0\%$	$acc_{bal} = 94.7\%$
	$l_{min} = 4$	$TPR = 79.4\%$	$TPR = 95.5\%$	$TPR = 99.2\%$
	$thres = 0.5492$	$FPR = 9.5\%$	$FPR = 13.5\%$	$FPR = 9.7\%$

Table C.43: Classification and correlation results, RQA, with *ENT* as BWE-equivalent, using angular distance, for natural porosity area 5056A–D. Optimum RQA parameters and decision threshold determined on 3 areas (scan sets), test results with these parameters on scans of the remaining scan set.

results of specimen/ scan	parameters	test results		
	determined on 5056-B/C/D	scan A1	scan A2	scan A3
5056-A	$d = 10$	$r = -0.65$	$r = -0.68$	$r = -0.68$
	$\tau = 9$	$AUC = 0.886$	$AUC = 0.893$	$AUC = 0.896$
	$\epsilon = 0.650\pi$	$acc_{bal} = 80.8\%$	$acc_{bal} = 82.1\%$	$acc_{bal} = 81.0\%$
	$l_{min} = 5$	$TPR = 79.9\%$	$TPR = 78.0\%$	$TPR = 78.8\%$
	$thres = 3.2862$	$FPR = 18.3\%$	$FPR = 13.8\%$	$FPR = 16.8\%$
	determined on 5056-A/C/D	scan B1	scan B2	scan B3
5056-B	$d = 10$	$r = -0.80$	$r = -0.84$	$r = -0.80$
	$\tau = 9$	$AUC = 0.952$	$AUC = 0.975$	$AUC = 0.964$
	$\epsilon = 0.650\pi$	$acc_{bal} = 89.7\%$	$acc_{bal} = 91.7\%$	$acc_{bal} = 90.3\%$
	$l_{min} = 6$	$TPR = 88.5\%$	$TPR = 94.9\%$	$TPR = 91.0\%$
	$thres = 3.4610$	$FPR = 9.1\%$	$FPR = 11.5\%$	$FPR = 10.5\%$
	determined on 5056-A/B/D	scan C1	scan C2	scan C3
5056-C	$d = 4$	$r = -0.59$	$r = -0.69$	$r = -0.67$
	$\tau = 9$	$AUC = 0.825$	$AUC = 0.897$	$AUC = 0.884$
	$\epsilon = 0.750\pi$	$acc_{bal} = 73.7\%$	$acc_{bal} = 82.9\%$	$acc_{bal} = 81.7\%$
	$l_{min} = 2$	$TPR = 58.6\%$	$TPR = 75.3\%$	$TPR = 75.7\%$
	$thres = 3.5285$	$FPR = 11.3\%$	$FPR = 9.5\%$	$FPR = 12.3\%$
	determined on 5056-A/B/C	scan D1	scan D2	scan D3
5056-D	$d = 10$	$r = -0.64$	$r = -0.77$	$r = -0.89$
	$\tau = 9$	$AUC = 0.871$	$AUC = 0.941$	$AUC = 0.995$
	$\epsilon = 0.650\pi$	$acc_{bal} = 78.3\%$	$acc_{bal} = 86.4\%$	$acc_{bal} = 91.1\%$
	$l_{min} = 6$	$TPR = 72.7\%$	$TPR = 90.7\%$	$TPR = 99.7\%$
	$thres = 3.4529$	$FPR = 16.1\%$	$FPR = 17.8\%$	$FPR = 17.6\%$

C.3.2 Evaluation in time domain

Table C.44: Classification and correlation results with x_{max} as BWE-equivalent for natural porosity area 5056A–D. Optimum decision threshold determined on 3 areas (scan sets), test results with these parameters on scans of the remaining scan set.

results of specimen/ scan	decision threshold	test results		
	determined on 5056-B/C/D	scan A1	scan A2	scan A3
5056-A	*thres*= 27.3438	$r = 0.06$ $AUC = 0.521$ $acc_{bal} = 55.7\%$ $TPR = 51.6\%$ $FPR = 40.2\%$	$r = 0.25$ $AUC = 0.616$ $acc_{bal} = 59.5\%$ $TPR = 53.7\%$ $FPR = 34.6\%$	$r = 0.33$ $AUC = 0.698$ $acc_{bal} = 64.9\%$ $TPR = 56.0\%$ $FPR = 26.3\%$
	determined on 5056-A/C/D	scan B1	scan B2	scan B3
5056-B	*thres*= 27.7466	$r = 0.81$ $AUC = 0.958$ $acc_{bal} = 84.1\%$ $TPR = 94.6\%$ $FPR = 26.5\%$	$r = 0.94$ $AUC = 0.999$ $acc_{bal} = 94.5\%$ $TPR = 100.0\%$ $FPR = 11.1\%$	$r = 0.70$ $AUC = 0.900$ $acc_{bal} = 86.3\%$ $TPR = 85.1\%$ $FPR = 12.5\%$
	determined on 5056-A/B/D	scan C1	scan C2	scan C3
5056-C	*thres*= 27.4170	$r = 0.20$ $AUC = 0.615$ $acc_{bal} = 57.0\%$ $TPR = 48.9\%$ $FPR = 35.0\%$	$r = 0.37$ $AUC = 0.709$ $acc_{bal} = 62.5\%$ $TPR = 51.0\%$ $FPR = 25.9\%$	$r = 0.52$ $AUC = 0.798$ $acc_{bal} = 73.1\%$ $TPR = 78.7\%$ $FPR = 32.5\%$
	determined on 5056-A/B/C	scan D1	scan D2	scan D3
5056-D	*thres*= 27.6978	$r = -0.07$ $AUC = 0.505$ $acc_{bal} = 48.2\%$ $TPR = 63.6\%$ $FPR = 67.2\%$	$r = 0.52$ $AUC = 0.742$ $acc_{bal} = 63.2\%$ $TPR = 74.7\%$ $FPR = 48.3\%$	$r = 0.93$ $AUC = 0.999$ $acc_{bal} = 84.5\%$ $TPR = 99.9\%$ $FPR = 31.0\%$

Table C.45: Classification and correlation results with s^2 as BWE-equivalent for natural porosity area 5056A–D. Optimum decision threshold determined on 3 areas (scan sets), test results with these parameters on scans of the remaining scan set.

results of specimen/ scan	decision threshold	test results		
	determined on 5056-B/C/D	scan A1	scan A2	scan A3
5056-A	*thres*= 28.3502	$r = 0.02$ $AUC = 0.515$ $acc_{bal} = 52.0\%$ $TPR = 43.3\%$ $FPR = 39.3\%$	$r = 0.19$ $AUC = 0.598$ $acc_{bal} = 57.9\%$ $TPR = 50.9\%$ $FPR = 35.1\%$	$r = 0.08$ $AUC = 0.588$ $acc_{bal} = 60.1\%$ $TPR = 47.9\%$ $FPR = 27.6\%$
	determined on 5056-A/C/D	scan B1	scan B2	scan B3
5056-B	*thres*= 29.0350	$r = 0.80$ $AUC = 0.954$ $acc_{bal} = 84.7\%$ $TPR = 93.4\%$ $FPR = 24.0\%$	$r = 0.94$ $AUC = 0.999$ $acc_{bal} = 95.9\%$ $TPR = 100.0\%$ $FPR = 8.1\%$	$r = 0.75$ $AUC = 0.931$ $acc_{bal} = 90.2\%$ $TPR = 88.3\%$ $FPR = 7.8\%$
	determined on 5056-A/B/D	scan C1	scan C2	scan C3
5056-C	*thres*= 28.6138	$r = 0.21$ $AUC = 0.576$ $acc_{bal} = 53.8\%$ $TPR = 51.6\%$ $FPR = 44.1\%$	$r = 0.17$ $AUC = 0.575$ $acc_{bal} = 57.5\%$ $TPR = 46.0\%$ $FPR = 31.0\%$	$r = 0.30$ $AUC = 0.663$ $acc_{bal} = 59.8\%$ $TPR = 58.9\%$ $FPR = 39.3\%$
	determined on 5056-A/B/C	scan D1	scan D2	scan D3
5056-D	*thres*= 29.2189	$r = -0.05$ $AUC = 0.508$ $acc_{bal} = 47.8\%$ $TPR = 56.9\%$ $FPR = 61.4\%$	$r = 0.48$ $AUC = 0.749$ $acc_{bal} = 61.5\%$ $TPR = 75.5\%$ $FPR = 52.6\%$	$r = 0.92$ $AUC = 0.998$ $acc_{bal} = 81.5\%$ $TPR = 99.9\%$ $FPR = 37.0\%$

Table C.46: Classification and correlation results with QCD as BWE-equivalent for natural porosity area 5056A–D. Optimum decision threshold determined on 3 areas (scan sets), test results with these parameters on scans of the remaining scan set.

results of specimen/ scan	decision threshold	test results		
	determined on 5056-B/C/D	scan A1	scan A2	scan A3
5056-A	*thres*= 0.0538	$r = 0.03$ $AUC = 0.537$ $acc_{bal} = 51.7\%$ $TPR = 63.4\%$ $FPR = 60.0\%$	$r = 0.01$ $AUC = 0.492$ $acc_{bal} = 47.5\%$ $TPR = 52.7\%$ $FPR = 57.8\%$	$r = -0.32$ $AUC = 0.689$ $acc_{bal} = 64.3\%$ $TPR = 60.8\%$ $FPR = 32.3\%$
	determined on 5056-A/C/D	scan B1	scan B2	scan B3
5056-B	*thres*= 0.0514	$r = -0.55$ $AUC = 0.855$ $acc_{bal} = 72.6\%$ $TPR = 90.1\%$ $FPR = 44.9\%$	$r = -0.74$ $AUC = 0.991$ $acc_{bal} = 87.3\%$ $TPR = 99.3\%$ $FPR = 24.8\%$	$r = -0.69$ $AUC = 0.968$ $acc_{bal} = 84.4\%$ $TPR = 96.8\%$ $FPR = 28.0\%$
	determined on 5056-A/B/D	scan C1	scan C2	scan C3
5056-C	*thres*= 0.0543	$r = 0.06$ $AUC = 0.468$ $acc_{bal} = 49.0\%$ $TPR = 70.5\%$ $FPR = 72.6\%$	$r = -0.36$ $AUC = 0.740$ $acc_{bal} = 68.1\%$ $TPR = 60.3\%$ $FPR = 24.2\%$	$r = -0.40$ $AUC = 0.763$ $acc_{bal} = 69.7\%$ $TPR = 65.8\%$ $FPR = 26.4\%$
	determined on 5056-A/B/C	scan D1	scan D2	scan D3
5056-D	*thres*= 0.0528	$r = -0.09$ $AUC = 0.549$ $acc_{bal} = 51.6\%$ $TPR = 37.9\%$ $FPR = 34.7\%$	$r = -0.47$ $AUC = 0.749$ $acc_{bal} = 66.9\%$ $TPR = 68.6\%$ $FPR = 34.7\%$	$r = -0.72$ $AUC = 0.977$ $acc_{bal} = 82.0\%$ $TPR = 98.9\%$ $FPR = 35.0\%$

Table C.47: Classification and correlation results with linear regression in time domain for natural porosity area 5056A–D. Training (including decision threshold to achieve $TPR \approx TNR$) determined on 3 areas (scan sets), test results on scans of the remaining scan set.

results of specimen/ scan	decision threshold	test results		
	determined on 5056-B/C/D	scan A1	scan A2	scan A3
5056-A	*thres*= 12.7860	$r = 0.91$ $AUC = 0.996$ $acc_{bal} = 97.1\%$ $TPR = 96.0\%$ $FPR = 1.8\%$	$r = 0.92$ $AUC = 0.998$ $acc_{bal} = 97.7\%$ $TPR = 96.6\%$ $FPR = 1.2\%$	$r = 0.92$ $AUC = 0.999$ $acc_{bal} = 98.4\%$ $TPR = 98.1\%$ $FPR = 1.2\%$
	determined on 5056-A/C/D	scan B1	scan B2	scan B3
5056-B	*thres*= 12.3142	$r = 0.91$ $AUC = 1.000$ $acc_{bal} = 98.9\%$ $TPR = 99.2\%$ $FPR = 1.4\%$	$r = 0.93$ $AUC = 1.000$ $acc_{bal} = 99.6\%$ $TPR = 99.6\%$ $FPR = 0.4\%$	$r = 0.90$ $AUC = 0.998$ $acc_{bal} = 99.0\%$ $TPR = 98.6\%$ $FPR = 0.5\%$
	determined on 5056-A/B/D	scan C1	scan C2	scan C3
5056-C	*thres*= 13.5463	$r = 0.88$ $AUC = 0.990$ $acc_{bal} = 97.6\%$ $TPR = 97.7\%$ $FPR = 2.5\%$	$r = 0.89$ $AUC = 0.979$ $acc_{bal} = 94.9\%$ $TPR = 96.9\%$ $FPR = 7.1\%$	$r = 0.89$ $AUC = 0.982$ $acc_{bal} = 96.0\%$ $TPR = 98.5\%$ $FPR = 6.6\%$
	determined on 5056-A/B/C	scan D1	scan D2	scan D3
5056-D	*thres*= 11.8597	$r = 0.73$ $AUC = 0.922$ $acc_{bal} = 77.2\%$ $TPR = 55.8\%$ $FPR = 1.5\%$	$r = 0.85$ $AUC = 0.981$ $acc_{bal} = 89.2\%$ $TPR = 79.4\%$ $FPR = 1.0\%$	$r = 0.93$ $AUC = 1.000$ $acc_{bal} = 99.2\%$ $TPR = 98.7\%$ $FPR = 0.2\%$

C.3.3 Fourier analysis

Table C.48: Classification and correlation results with maximum frequency, left and right cutoff frequency as BWE-equivalent on scan sets 5056A–D. Decision threshold to achieve $TPR \approx TNR$ determined on 3 sets, results with this threshold on remaining set; results of all three scans of a set combined.

results of specimen/ scan	F_{peak}	F_{cl}	F_{cr}
	threshold to achieve $TPR \approx TNR$ on 5056-B/C/D		
	$thres = 6.3333$	$thres = 5.9524$	$thres = 6.7650$
5056-A	$r = -0.10$ $AUC = 0.602$ $acc_{bal} = 53.1\%$ $TPR = 97.4\%$ $FPR = 91.1\%$	$r = -0.13$ $AUC = 0.649$ $acc_{bal} = 69.3\%$ $TPR = 90.5\%$ $FPR = 51.8\%$	$r = 0.17$ $AUC = 0.566$ $acc_{bal} = 47.5\%$ $TPR = 28.6\%$ $FPR = 33.6\%$
	threshold to achieve $TPR \approx TNR$ on 5056-A/C/D		
	$thres = 6.3333$	$thres = 6.0085$	$thres = 6.8150$
5056-B	$r = 0.46$ $AUC = 0.692$ $acc_{bal} = 70.1\%$ $TPR = 41.4\%$ $FPR = 1.3\%$	$r = 0.45$ $AUC = 0.569$ $acc_{bal} = 42.5\%$ $TPR = 66.6\%$ $FPR = 81.6\%$	$r = 0.47$ $AUC = 0.816$ $acc_{bal} = 62.4\%$ $TPR = 87.6\%$ $FPR = 62.8\%$
	threshold to achieve $TPR \approx TNR$ on 5056-A/B/D		
	$thres = 6.3333$	$thres = 5.9644$	$thres = 6.7860$
5056-C	$r = 0.01$ $AUC = 0.476$ $acc_{bal} = 45.7\%$ $TPR = 0.8\%$ $FPR = 9.4\%$	$r = -0.12$ $AUC = 0.704$ $acc_{bal} = 68.5\%$ $TPR = 79.4\%$ $FPR = 42.4\%$	$r = 0.21$ $AUC = 0.643$ $acc_{bal} = 55.7\%$ $TPR = 52.2\%$ $FPR = 40.9\%$
	threshold to achieve $TPR \approx TNR$ on 5056-A/B/C		
	$thres = 6.3333$	$thres = 5.9900$	$thres = 6.7874$
5056-D	$r = 0.40$ $AUC = 0.556$ $acc_{bal} = 60.9\%$ $TPR = 30.6\%$ $FPR = 8.8\%$	$r = 0.40$ $AUC = 0.493$ $acc_{bal} = 41.0\%$ $TPR = 50.8\%$ $FPR = 68.9\%$	$r = 0.42$ $AUC = 0.637$ $acc_{bal} = 57.1\%$ $TPR = 54.8\%$ $FPR = 40.6\%$

Table C.49: Classification and correlation results with mean frequency, bandwidth and integral of bandwidth as BWE-equivalent on scan sets 5056A–D. Decision threshold to achieve $TPR \approx TNR$ determined on 3 sets, results with this threshold on remaining set; results of all three scans of a set combined.

results of specimen/ scan	F_{centre}	B_{3dB}	A_{bw3}
5056-A	threshold to achieve $TPR \approx TNR$ on 5056-B/C/D		
	$thres = 6.3498$	$thres = 0.7613$	$thres = 375.7461$
	$r = 0.01$ $AUC = 0.418$ $acc_{bal} = 41.3\%$ $TPR = 17.8\%$ $FPR = 35.2\%$	$r = 0.36$ $AUC = 0.699$ $acc_{bal} = 62.3\%$ $TPR = 60.1\%$ $FPR = 35.5\%$	$r = 0.52$ $AUC = 0.801$ $acc_{bal} = 73.2\%$ $TPR = 71.9\%$ $FPR = 25.6\%$
5056-B	threshold to achieve $TPR \approx TNR$ on 5056-A/C/D		
	$thres = 6.3900$	$thres = 0.7749$	$thres = 382.7892$
	$r = 0.46$ $AUC = 0.732$ $acc_{bal} = 54.6\%$ $TPR = 83.0\%$ $FPR = 73.8\%$	$r = 0.46$ $AUC = 0.793$ $acc_{bal} = 70.3\%$ $TPR = 77.2\%$ $FPR = 36.6\%$	$r = 0.82$ $AUC = 0.970$ $acc_{bal} = 88.8\%$ $TPR = 95.3\%$ $FPR = 17.6\%$
5056-C	threshold to achieve $TPR \approx TNR$ on 5056-A/B/D		
	$thres = 6.3644$	$thres = 0.7652$	$thres = 373.3488$
	$r = 0.06$ $AUC = 0.477$ $acc_{bal} = 49.2\%$ $TPR = 43.9\%$ $FPR = 45.5\%$	$r = 0.43$ $AUC = 0.751$ $acc_{bal} = 67.7\%$ $TPR = 68.0\%$ $FPR = 32.5\%$	$r = 0.59$ $AUC = 0.828$ $acc_{bal} = 75.5\%$ $TPR = 71.4\%$ $FPR = 20.3\%$
5056-D	threshold to achieve $TPR \approx TNR$ on 5056-A/B/C		
	$thres = 6.3705$	$thres = 0.7581$	$thres = 374.6841$
	$r = 0.41$ $AUC = 0.562$ $acc_{bal} = 49.4\%$ $TPR = 51.8\%$ $FPR = 52.9\%$	$r = 0.48$ $AUC = 0.741$ $acc_{bal} = 65.0\%$ $TPR = 60.5\%$ $FPR = 30.5\%$	$r = 0.67$ $AUC = 0.875$ $acc_{bal} = 77.5\%$ $TPR = 75.0\%$ $FPR = 20.0\%$

Table C.50: Classification and correlation results with linear regression on amplitude spectrum from 0.66 MHz to 7 MHz for natural porosity area 5056A–D. Training (including decision threshold to achieve $TPR \approx TNR$) determined on 3 areas (scan sets), test results on scans of the remaining scan set.

results of specimen/ scan	decision threshold	test results		
	determined on 5056-B/C/D	scan A1	scan A2	scan A3
5056-A	*thres*= 13.2260	$r = 0.86$ $AUC = 0.992$ $acc_{bal} = 94.2\%$ $TPR = 97.1\%$ $FPR = 8.8\%$	$r = 0.85$ $AUC = 0.988$ $acc_{bal} = 92.4\%$ $TPR = 95.6\%$ $FPR = 10.8\%$	$r = 0.88$ $AUC = 0.992$ $acc_{bal} = 95.0\%$ $TPR = 96.6\%$ $FPR = 6.7\%$
	determined on 5056-A/C/D	scan B1	scan B2	scan B3
5056-B	*thres*= 13.4920	$r = 0.93$ $AUC = 1.000$ $acc_{bal} = 99.6\%$ $TPR = 99.7\%$ $FPR = 0.5\%$	$r = 0.95$ $AUC = 1.000$ $acc_{bal} = 100.0\%$ $TPR = 100.0\%$ $FPR = 0.1\%$	$r = 0.95$ $AUC = 1.000$ $acc_{bal} = 99.9\%$ $TPR = 99.9\%$ $FPR = 0.0\%$
	determined on 5056-A/B/D	scan C1	scan C2	scan C3
5056-C	*thres*= 13.5516	$r = 0.74$ $AUC = 0.949$ $acc_{bal} = 87.1\%$ $TPR = 80.4\%$ $FPR = 6.1\%$	$r = 0.88$ $AUC = 0.981$ $acc_{bal} = 94.4\%$ $TPR = 97.4\%$ $FPR = 8.7\%$	$r = 0.88$ $AUC = 0.982$ $acc_{bal} = 94.8\%$ $TPR = 98.3\%$ $FPR = 8.6\%$
	determined on 5056-A/B/C	scan D1	scan D2	scan D3
5056-D	*thres*= 13.4756	$r = 0.77$ $AUC = 0.945$ $acc_{bal} = 83.2\%$ $TPR = 72.4\%$ $FPR = 6.0\%$	$r = 0.86$ $AUC = 0.983$ $acc_{bal} = 92.8\%$ $TPR = 93.3\%$ $FPR = 7.6\%$	$r = 0.94$ $AUC = 1.000$ $acc_{bal} = 97.0\%$ $TPR = 100.0\%$ $FPR = 5.9\%$

Table C.51: Classification and correlation results with linear regression on amplitude spectrum 0 MHz to 23 MHz for natural porosity area 5056A–D. Hann window applied on intermediate echo time series prior to Fourier transform. Training (including decision threshold to achieve $TPR \approx TNR$) determined on 3 areas (scan sets), test results on scans of the remaining scan set

results of specimen/ scan	decision threshold	test results		
	determined on 5056-B/C/D	scan A1	scan A2	scan A3
5056-A	*thres*= 12.4702	$r = 0.91$ $AUC = 1.000$ $acc_{bal} = 99.3\%$ $TPR = 98.8\%$ $FPR = 0.3\%$	$r = 0.91$ $AUC = 1.000$ $acc_{bal} = 99.5\%$ $TPR = 99.2\%$ $FPR = 0.1\%$	$r = 0.89$ $AUC = 0.997$ $acc_{bal} = 97.1\%$ $TPR = 94.7\%$ $FPR = 0.5\%$
	determined on 5056-A/C/D	scan B1	scan B2	scan B3
5056-B	*thres*= 12.2740	$r = 0.92$ $AUC = 1.000$ $acc_{bal} = 99.3\%$ $TPR = 99.8\%$ $FPR = 1.3\%$	$r = 0.94$ $AUC = 1.000$ $acc_{bal} = 99.7\%$ $TPR = 99.9\%$ $FPR = 0.5\%$	$r = 0.91$ $AUC = 1.000$ $acc_{bal} = 99.5\%$ $TPR = 99.8\%$ $FPR = 0.8\%$
	determined on 5056-A/B/D	scan C1	scan C2	scan C3
5056-C	*thres*= 12.8960	$r = 0.92$ $AUC = 0.999$ $acc_{bal} = 98.5\%$ $TPR = 98.8\%$ $FPR = 1.7\%$	$r = 0.88$ $AUC = 0.992$ $acc_{bal} = 96.2\%$ $TPR = 96.4\%$ $FPR = 4.0\%$	$r = 0.90$ $AUC = 0.993$ $acc_{bal} = 96.5\%$ $TPR = 97.0\%$ $FPR = 4.0\%$
	determined on 5056-A/B/C	scan D1	scan D2	scan D3
5056-D	*thres*= 12.3712	$r = 0.90$ $AUC = 0.999$ $acc_{bal} = 98.3\%$ $TPR = 97.5\%$ $FPR = 0.8\%$	$r = 0.92$ $AUC = 1.000$ $acc_{bal} = 99.3\%$ $TPR = 99.4\%$ $FPR = 0.9\%$	$r = 0.94$ $AUC = 1.000$ $acc_{bal} = 99.5\%$ $TPR = 100.0\%$ $FPR = 1.0\%$

C.3.4 Test results on thinner specimens

Results on 6 mm thick specimens, reduced intermediate echo gate

Table C.52: Classification and correlation results with linear regression on amplitude spectrum from 0 MHz to 22.67 MHz for natural porosity area 5056A–D; reduced IE gate from 1 µs to 2.5 µs. Hann window applied on intermediate echo time series prior to Fourier transform. Training (including decision threshold to achieve $TPR \approx TNR$) determined on 3 areas (scan sets), test results on scans of the remaining scan set.

results of specimen/ scan	decision threshold	test results		
	determined on 5056-B/C/D	scan A1	scan A2	scan A3
5056-A	*thres*= 12.9156	$r = 0.89$ $AUC = 0.998$ $acc_{bal} = 97.3\%$ $TPR = 98.9\%$ $FPR = 4.4\%$	$r = 0.87$ $AUC = 0.995$ $acc_{bal} = 96.4\%$ $TPR = 97.2\%$ $FPR = 4.5\%$	$r = 0.87$ $AUC = 0.985$ $acc_{bal} = 96.2\%$ $TPR = 94.2\%$ $FPR = 1.9\%$
	determined on 5056-A/C/D	scan B1	scan B2	scan B3
5056-B	*thres*= 12.8619	$r = 0.93$ $AUC = 1.000$ $acc_{bal} = 99.8\%$ $TPR = 99.9\%$ $FPR = 0.4\%$	$r = 0.94$ $AUC = 1.000$ $acc_{bal} = 99.4\%$ $TPR = 100.0\%$ $FPR = 1.2\%$	$r = 0.93$ $AUC = 0.997$ $acc_{bal} = 98.9\%$ $TPR = 98.4\%$ $FPR = 0.5\%$
	determined on 5056-A/B/D	scan C1	scan C2	scan C3
5056-C	*thres*= 13.2200	$r = 0.83$ $AUC = 0.975$ $acc_{bal} = 91.1\%$ $TPR = 86.6\%$ $FPR = 4.3\%$	$r = 0.84$ $AUC = 0.978$ $acc_{bal} = 92.4\%$ $TPR = 91.9\%$ $FPR = 7.1\%$	$r = 0.87$ $AUC = 0.983$ $acc_{bal} = 94.7\%$ $TPR = 95.4\%$ $FPR = 6.1\%$
	determined on 5056-A/B/C	scan D1	scan D2	scan D3
5056-D	*thres*= 13.1192	$r = 0.85$ $AUC = 0.988$ $acc_{bal} = 93.9\%$ $TPR = 91.1\%$ $FPR = 3.2\%$	$r = 0.91$ $AUC = 0.998$ $acc_{bal} = 98.1\%$ $TPR = 98.9\%$ $FPR = 2.7\%$	$r = 0.93$ $AUC = 1.000$ $acc_{bal} = 98.3\%$ $TPR = 100.0\%$ $FPR = 3.4\%$

Test on 4 mm thin specimens

Table C.53: Classification and correlation results with linear regression on amplitude spectrum from 0 MHz to 22.67 MHz for natural porosity area 4955A/B; IE gate from 1 µs to 2.5 µs. Hann window applied on intermediate echo time series prior to Fourier transform. Training (including decision threshold to achieve $TPR \approx TNR$) determined on scan sets 5056A/B/D, IE gate 1 µs–2.5 µs.

results of specimen/ scan	decision threshold	test results		
	determined on 5056-A/B/D	scan A1	scan A2	scan A3
4955-A	*thres*= 13.2200	$r = 0.88$ $AUC = 0.994$ $acc_{bal} = 95.7\%$ $TPR = 96.2\%$ $FPR = 4.9\%$	$r = 0.85$ $AUC = 0.977$ $acc_{bal} = 94.0\%$ $TPR = 89.4\%$ $FPR = 1.4\%$	$r = 0.85$ $AUC = 0.988$ $acc_{bal} = 94.5\%$ $TPR = 93.9\%$ $FPR = 5.0\%$
	determined on 5056-A/B/D	scan B1	scan B2	scan B3
4955-B	*thres*= 13.2200	$r = 0.76$ $AUC = 0.941$ $acc_{bal} = 89.0\%$ $TPR = 92.0\%$ $FPR = 14.0\%$	$r = 0.77$ $AUC = 0.964$ $acc_{bal} = 91.1\%$ $TPR = 93.5\%$ $FPR = 11.3\%$	$r = 0.76$ $AUC = 0.933$ $acc_{bal} = 86.6\%$ $TPR = 93.7\%$ $FPR = 20.4\%$

C.4 Unidirectional specimens; non-resonance case

C.4.1 Recurrence quantification analysis

Euclidean distance

Table C.54: Classification and correlation results, RQA, with *RR* as BWE-equivalent, using Euclidean distance, for natural porosity area 2438A–D. Optimum RQA parameters and decision threshold determined on 3 areas (scan sets), test results with these parameters on scans of the remaining scan set.

results of specimen/ scan	parameters	test results		
	determined on 2438-B/C/D	scan A1	scan A2	scan A3
2438-A	$d = 3$ $\tau = 9$ $\epsilon = 1.2$ *thres*= 0.1963	$r = 0.65$ *AUC*= 0.893 $acc_{bal} = 81.5\%$ *TPR* = 78.9% *FPR* = 16.0%	$r = 0.55$ *AUC*= 0.841 $acc_{bal} = 77.0\%$ *TPR* = 69.3% *FPR* = 15.3%	$r = 0.55$ *AUC*= 0.842 $acc_{bal} = 76.5\%$ *TPR* = 69.8% *FPR* = 16.8%
	determined on 2438-A/C/D	scan B1	scan B2	scan B3
2438-B	$d = 3$ $\tau = 9$ $\epsilon = 1.2$ *thres*= 0.2009	$r = 0.42$ *AUC*= 0.760 $acc_{bal} = 68.9\%$ *TPR* = 75.0% *FPR* = 37.1%	$r = 0.43$ *AUC*= 0.768 $acc_{bal} = 69.4\%$ *TPR* = 69.8% *FPR* = 31.1%	$r = 0.42$ *AUC*= 0.760 $acc_{bal} = 69.5\%$ *TPR* = 74.8% *FPR* = 35.8%
	determined on 2438-A/B/D	scan C1	scan C2	scan C3
2438-C	$d = 3$ $\tau = 9$ $\epsilon = 0.8$ *thres*= 0.0885	$r = 0.71$ *AUC*= 0.940 $acc_{bal} = 86.7\%$ *TPR* = 87.6% *FPR* = 14.2%	$r = 0.67$ *AUC*= 0.915 $acc_{bal} = 83.3\%$ *TPR* = 86.5% *FPR* = 20.0%	$r = 0.67$ *AUC*= 0.919 $acc_{bal} = 82.3\%$ *TPR* = 88.5% *FPR* = 23.9%
	determined on 2438-A/B/C	scan D1	scan D2	scan D3
2438-D	$d = 3$ $\tau = 9$ $\epsilon = 0.9$ *thres*= 0.1129	$r = 0.66$ *AUC*= 0.936 $acc_{bal} = 86.4\%$ *TPR* = 84.8% *FPR* = 12.0%	$r = 0.68$ *AUC*= 0.933 $acc_{bal} = 84.9\%$ *TPR* = 84.3% *FPR* = 14.4%	$r = 0.69$ *AUC*= 0.933 $acc_{bal} = 84.9\%$ *TPR* = 88.1% *FPR* = 18.4%

Table C.55: Classification and correlation results, RQA, with *DIV* as BWE-equivalent, using Euclidean distance, for natural porosity area 2438A–D. Optimum RQA parameters and decision threshold determined on 3 areas (scan sets), test results with these parameters on scans of the remaining scan set.

results of specimen/ scan	parameters	test results		
	determined on 2438-B/C/D	scan A1	scan A2	scan A3
2438-A	$d = 2$	$r = -0.20$	$r = -0.22$	$r = -0.20$
	$\tau = 6$	$AUC = 0.647$	$AUC = 0.638$	$AUC = 0.630$
	$\epsilon = 1.5$	$acc_{bal} = 52.7\%$	$acc_{bal} = 59.2\%$	$acc_{bal} = 58.0\%$
		$TPR = 36.6\%$	$TPR = 44.9\%$	$TPR = 46.4\%$
	$thres = 0.0040$	$FPR = 31.2\%$	$FPR = 26.5\%$	$FPR = 30.5\%$
	determined on 2438-A/C/D	scan B1	scan B2	scan B3
2438-B	$d = 8$	$r = -0.07$	$r = -0.06$	$r = -0.01$
	$\tau = 10$	$AUC = 0.564$	$AUC = 0.566$	$AUC = 0.510$
	$\epsilon = 1.8$	$acc_{bal} = 54.1\%$	$acc_{bal} = 57.8\%$	$acc_{bal} = 51.7\%$
		$TPR = 68.4\%$	$TPR = 67.6\%$	$TPR = 66.2\%$
	$thres = 0.0053$	$FPR = 60.2\%$	$FPR = 52.0\%$	$FPR = 62.7\%$
	determined on 2438-A/B/D	scan C1	scan C2	scan C3
2438-C	$d = 4$	$r = -0.32$	$r = -0.16$	$r = -0.03$
	$\tau = 2$	$AUC = 0.706$	$AUC = 0.559$	$AUC = 0.519$
	$\epsilon = 2.0$	$acc_{bal} = 67.8\%$	$acc_{bal} = 55.5\%$	$acc_{bal} = 49.0\%$
		$TPR = 57.5\%$	$TPR = 52.1\%$	$TPR = 52.1\%$
	$thres = 0.0039$	$FPR = 22.0\%$	$FPR = 41.1\%$	$FPR = 54.1\%$
	determined on 2438-A/B/C	scan D1	scan D2	scan D3
2438-D	$d = 3$	$r = -0.22$	$r = -0.21$	$r = -0.18$
	$\tau = 3$	$AUC = 0.673$	$AUC = 0.662$	$AUC = 0.635$
	$\epsilon = 1.9$	$acc_{bal} = 65.0\%$	$acc_{bal} = 62.4\%$	$acc_{bal} = 57.3\%$
		$TPR = 58.1\%$	$TPR = 57.6\%$	$TPR = 60.7\%$
	$thres = 0.0039$	$FPR = 28.2\%$	$FPR = 32.8\%$	$FPR = 46.0\%$

Table C.56: Classification and correlation results, RQA, with *TND* as BWE-equivalent, using Euclidean distance, for natural porosity area 2438A–D. Optimum RQA parameters and decision threshold determined on 3 areas (scan sets), test results with these parameters on scans of the remaining scan set.

results of specimen/ scan	parameters	test results		
	determined on 2438-B/C/D	scan A1	scan A2	scan A3
2438-A	$d = 4$ $\tau = 9$ $\epsilon = 1.0$ $thres = -0.6587$	$r = -0.47$ $AUC = 0.796$ $acc_{bal} = 71.6\%$ $TPR = 72.7\%$ $FPR = 29.4\%$	$r = -0.52$ $AUC = 0.840$ $acc_{bal} = 75.3\%$ $TPR = 76.8\%$ $FPR = 26.3\%$	$r = -0.52$ $AUC = 0.836$ $acc_{bal} = 74.2\%$ $TPR = 78.4\%$ $FPR = 29.9\%$
	determined on 2438-A/C/D	scan B1	scan B2	scan B3
2438-B	$d = 4$ $\tau = 9$ $\epsilon = 1.0$ $thres = -0.6771$	$r = -0.39$ $AUC = 0.751$ $acc_{bal} = 64.9\%$ $TPR = 79.5\%$ $FPR = 49.8\%$	$r = -0.34$ $AUC = 0.727$ $acc_{bal} = 65.3\%$ $TPR = 71.8\%$ $FPR = 41.3\%$	$r = -0.31$ $AUC = 0.701$ $acc_{bal} = 63.8\%$ $TPR = 75.2\%$ $FPR = 47.6\%$
	determined on 2438-A/B/D	scan C1	scan C2	scan C3
2438-C	$d = 4$ $\tau = 9$ $\epsilon = 0.9$ $thres = -0.5050$	$r = -0.61$ $AUC = 0.899$ $acc_{bal} = 81.8\%$ $TPR = 75.6\%$ $FPR = 12.0\%$	$r = -0.54$ $AUC = 0.871$ $acc_{bal} = 79.0\%$ $TPR = 76.5\%$ $FPR = 18.5\%$	$r = -0.54$ $AUC = 0.863$ $acc_{bal} = 77.7\%$ $TPR = 74.0\%$ $FPR = 18.6\%$
	determined on 2438-A/B/C	scan D1	scan D2	scan D3
2438-D	$d = 4$ $\tau = 9$ $\epsilon = 1.0$ $thres = -0.6395$	$r = -0.63$ $AUC = 0.910$ $acc_{bal} = 84.0\%$ $TPR = 81.0\%$ $FPR = 12.9\%$	$r = -0.52$ $AUC = 0.825$ $acc_{bal} = 74.1\%$ $TPR = 63.1\%$ $FPR = 14.9\%$	$r = -0.55$ $AUC = 0.853$ $acc_{bal} = 77.4\%$ $TPR = 70.0\%$ $FPR = 15.2\%$

Table C.57: Classification and correlation results, RQA, with *COR* as BWE-equivalent, using Euclidean distance, for natural porosity area 2438A–D. Optimum RQA parameters and decision threshold determined on 3 areas (scan sets), test results with these parameters on scans of the remaining scan set.

results of specimen/ scan	parameters	test results		
	determined on 2438-B/C/D	scan A1	scan A2	scan A3
2438-A	$d = 3$ $\tau = 9$ $\epsilon = 0.4$ $thres = -0.3634$	$r = -0.65$ $AUC = 0.898$ $acc_{bal} = 82.0\%$ $TPR = 79.3\%$ $FPR = 15.3\%$	$r = -0.59$ $AUC = 0.856$ $acc_{bal} = 78.2\%$ $TPR = 72.7\%$ $FPR = 16.3\%$	$r = -0.61$ $AUC = 0.876$ $acc_{bal} = 79.8\%$ $TPR = 73.1\%$ $FPR = 13.6\%$
	determined on 2438-A/C/D	scan B1	scan B2	scan B3
2438-B	$d = 3$ $\tau = 9$ $\epsilon = 0.4$ $thres = -0.3635$	$r = -0.56$ $AUC = 0.841$ $acc_{bal} = 77.3\%$ $TPR = 75.1\%$ $FPR = 20.4\%$	$r = -0.50$ $AUC = 0.814$ $acc_{bal} = 74.9\%$ $TPR = 68.6\%$ $FPR = 18.9\%$	$r = -0.55$ $AUC = 0.836$ $acc_{bal} = 76.3\%$ $TPR = 70.3\%$ $FPR = 17.6\%$
	determined on 2438-A/B/D	scan C1	scan C2	scan C3
2438-C	$d = 3$ $\tau = 9$ $\epsilon = 0.4$ $thres = -0.3720$	$r = -0.70$ $AUC = 0.924$ $acc_{bal} = 83.5\%$ $TPR = 88.2\%$ $FPR = 21.2\%$	$r = -0.64$ $AUC = 0.890$ $acc_{bal} = 79.2\%$ $TPR = 86.6\%$ $FPR = 28.2\%$	$r = -0.66$ $AUC = 0.899$ $acc_{bal} = 78.9\%$ $TPR = 88.9\%$ $FPR = 31.0\%$
	determined on 2438-A/B/C	scan D1	scan D2	scan D3
2438-D	$d = 3$ $\tau = 9$ $\epsilon = 0.4$ $thres = -0.3684$	$r = -0.66$ $AUC = 0.915$ $acc_{bal} = 83.9\%$ $TPR = 83.5\%$ $FPR = 15.8\%$	$r = -0.69$ $AUC = 0.932$ $acc_{bal} = 85.3\%$ $TPR = 87.1\%$ $FPR = 16.5\%$	$r = -0.68$ $AUC = 0.919$ $acc_{bal} = 83.7\%$ $TPR = 88.6\%$ $FPR = 21.3\%$

Table C.58: Classification and correlation results, RQA, with *DET* as BWE-equivalent, using Euclidean distance, for natural porosity area 2438A–D. Optimum RQA parameters and decision threshold determined on 3 areas (scan sets), test results with these parameters on scans of the remaining scan set.

results of specimen/ scan	parameters	test results		
	determined on 2438-B/C/D	scan A1	scan A2	scan A3
2438-A	$d = 7$	$r = -0.79$	$r = -0.61$	$r = -0.65$
	$\tau = 9$	$AUC = 0.956$	$AUC = 0.862$	$AUC = 0.894$
	$\epsilon = 1.0$	$acc_{bal} = 88.9\%$	$acc_{bal} = 74.6\%$	$acc_{bal} = 75.7\%$
	$l_{min} = 6$	$TPR = 81.7\%$	$TPR = 58.3\%$	$TPR = 56.6\%$
	$thres = 0.4528$	$FPR = 3.9\%$	$FPR = 9.2\%$	$FPR = 5.2\%$
	determined on 2438-A/C/D	scan B1	scan B2	scan B3
2438-B	$d = 8$	$r = -0.71$	$r = -0.67$	$r = -0.73$
	$\tau = 9$	$AUC = 0.917$	$AUC = 0.889$	$AUC = 0.929$
	$\epsilon = 1.1$	$acc_{bal} = 79.9\%$	$acc_{bal} = 78.3\%$	$acc_{bal} = 80.9\%$
	$l_{min} = 5$	$TPR = 65.0\%$	$TPR = 65.6\%$	$TPR = 66.6\%$
	$thres = 0.5577$	$FPR = 5.1\%$	$FPR = 9.0\%$	$FPR = 4.8\%$
	determined on 2438-A/B/D	scan C1	scan C2	scan C3
2438-C	$d = 7$	$r = -0.72$	$r = -0.69$	$r = -0.75$
	$\tau = 9$	$AUC = 0.925$	$AUC = 0.886$	$AUC = 0.926$
	$\epsilon = 0.9$	$acc_{bal} = 77.0\%$	$acc_{bal} = 76.0\%$	$acc_{bal} = 75.5\%$
	$l_{min} = 4$	$TPR = 96.0\%$	$TPR = 93.1\%$	$TPR = 98.0\%$
	$thres = 0.5536$	$FPR = 42.0\%$	$FPR = 41.1\%$	$FPR = 46.9\%$
	determined on 2438-A/B/C	scan D1	scan D2	scan D3
2438-D	$d = 7$	$r = -0.59$	$r = -0.74$	$r = -0.78$
	$\tau = 9$	$AUC = 0.879$	$AUC = 0.942$	$AUC = 0.950$
	$\epsilon = 1.1$	$acc_{bal} = 80.5\%$	$acc_{bal} = 84.0\%$	$acc_{bal} = 83.1\%$
	$l_{min} = 8$	$TPR = 82.6\%$	$TPR = 98.2\%$	$TPR = 98.5\%$
	$thres = 0.3320$	$FPR = 21.7\%$	$FPR = 30.1\%$	$FPR = 32.3\%$

Table C.59: Classification and correlation results, RQA, with *RATIO* as BWE-equivalent, using Euclidean distance, for natural porosity area 2438A–D. Optimum RQA parameters and decision threshold determined on 3 areas (scan sets), test results with these parameters on scans of the remaining scan set.

results of specimen/ scan	parameters	test results		
	determined on 2438-B/C/D	scan A1	scan A2	scan A3
2438-A	$d = 4$	$r = -0.66$	$r = -0.60$	$r = -0.59$
	$\tau = 9$	$AUC = 0.975$	$AUC = 0.907$	$AUC = 0.924$
	$\epsilon = 1.0$	$acc_{bal} = 91.4\%$	$acc_{bal} = 81.1\%$	$acc_{bal} = 81.8\%$
	$l_{min} = 4$	$TPR = 85.4\%$	$TPR = 68.4\%$	$TPR = 67.7\%$
	$thres = 6.7428$	$FPR = 2.5\%$	$FPR = 6.2\%$	$FPR = 4.1\%$
	determined on 2438-A/C/D	scan B1	scan B2	scan B3
2438-B	$d = 5$	$r = -0.58$	$r = -0.58$	$r = -0.59$
	$\tau = 9$	$AUC = 0.897$	$AUC = 0.897$	$AUC = 0.917$
	$\epsilon = 1.3$	$acc_{bal} = 82.2\%$	$acc_{bal} = 81.2\%$	$acc_{bal} = 85.0\%$
	$l_{min} = 4$	$TPR = 76.9\%$	$TPR = 75.3\%$	$TPR = 79.4\%$
	$thres = 5.9453$	$FPR = 12.5\%$	$FPR = 12.9\%$	$FPR = 9.3\%$
	determined on 2438-A/B/D	scan C1	scan C2	scan C3
2438-C	$d = 4$	$r = -0.69$	$r = -0.69$	$r = -0.68$
	$\tau = 9$	$AUC = 0.973$	$AUC = 0.955$	$AUC = 0.964$
	$\epsilon = 1.0$	$acc_{bal} = 86.6\%$	$acc_{bal} = 84.4\%$	$acc_{bal} = 82.1\%$
	$l_{min} = 4$	$TPR = 96.9\%$	$TPR = 95.8\%$	$TPR = 98.2\%$
	$thres = 6.1056$	$FPR = 23.8\%$	$FPR = 27.0\%$	$FPR = 34.1\%$
	determined on 2438-A/B/C	scan D1	scan D2	scan D3
2438-D	$d = 4$	$r = -0.62$	$r = -0.71$	$r = -0.73$
	$\tau = 9$	$AUC = 0.954$	$AUC = 0.980$	$AUC = 0.976$
	$\epsilon = 1.0$	$acc_{bal} = 88.2\%$	$acc_{bal} = 89.3\%$	$acc_{bal} = 86.8\%$
	$l_{min} = 4$	$TPR = 89.2\%$	$TPR = 97.8\%$	$TPR = 97.7\%$
	$thres = 6.2525$	$FPR = 12.9\%$	$FPR = 19.3\%$	$FPR = 24.0\%$

Table C.60: Classification and correlation results, RQA, with L_{nor} as BWE-equivalent, using Euclidean distance, for natural porosity area 2438A–D. Optimum RQA parameters and decision threshold determined on 3 areas (scan sets), test results with these parameters on scans of the remaining scan set.

results of specimen/ scan	parameters	test results		
	determined on 2438-B/C/D	scan A1	scan A2	scan A3
2438-A	$d = 7$	$r = 0.78$	$r = 0.61$	$r = 0.66$
	$\tau = 8$	$AUC = 0.952$	$AUC = 0.864$	$AUC = 0.895$
	$\epsilon = 1.1$	$acc_{bal} = 87.0\%$	$acc_{bal} = 75.0\%$	$acc_{bal} = 75.9\%$
	$l_{min} = 2$	$TPR = 80.0\%$	$TPR = 61.3\%$	$TPR = 59.7\%$
	$thres = 0.3350$	$FPR = 6.0\%$	$FPR = 11.2\%$	$FPR = 8.0\%$
	determined on 2438-A/C/D	scan B1	scan B2	scan B3
2438-B	$d = 8$	$r = 0.68$	$r = 0.62$	$r = 0.71$
	$\tau = 8$	$AUC = 0.901$	$AUC = 0.868$	$AUC = 0.917$
	$\epsilon = 1.2$	$acc_{bal} = 79.5\%$	$acc_{bal} = 77.1\%$	$acc_{bal} = 80.7\%$
	$l_{min} = 2$	$TPR = 68.1\%$	$TPR = 66.8\%$	$TPR = 68.8\%$
	$thres = 0.3050$	$FPR = 9.0\%$	$FPR = 12.6\%$	$FPR = 7.3\%$
	determined on 2438-A/B/D	scan C1	scan C2	scan C3
2438-C	$d = 7$	$r = 0.72$	$r = 0.70$	$r = 0.75$
	$\tau = 9$	$AUC = 0.922$	$AUC = 0.890$	$AUC = 0.923$
	$\epsilon = 1.1$	$acc_{bal} = 77.2\%$	$acc_{bal} = 78.1\%$	$acc_{bal} = 76.6\%$
	$l_{min} = 2$	$TPR = 96.1\%$	$TPR = 94.3\%$	$TPR = 98.1\%$
	$thres = 0.4284$	$FPR = 41.7\%$	$FPR = 38.1\%$	$FPR = 45.0\%$
	determined on 2438-A/B/C	scan D1	scan D2	scan D3
2438-D	$d = 8$	$r = 0.64$	$r = 0.69$	$r = 0.74$
	$\tau = 8$	$AUC = 0.899$	$AUC = 0.924$	$AUC = 0.939$
	$\epsilon = 1.2$	$acc_{bal} = 82.0\%$	$acc_{bal} = 83.4\%$	$acc_{bal} = 83.7\%$
	$l_{min} = 2$	$TPR = 83.6\%$	$TPR = 94.0\%$	$TPR = 95.5\%$
	$thres = 0.3213$	$FPR = 19.6\%$	$FPR = 27.1\%$	$FPR = 28.2\%$

Table C.61: Classification and correlation results, RQA, with ENT as BWE-equivalent, using Euclidean distance, for natural porosity area 2438A–D. Optimum RQA parameters and decision threshold determined on 3 areas (scan sets), test results with these parameters on scans of the remaining scan set.

results of specimen/ scan	parameters	test results		
	determined on 2438-B/C/D	scan A1	scan A2	scan A3
2438-A	$d = 8$	$r = 0.59$	$r = 0.42$	$r = 0.39$
	$\tau = 2$	$AUC = 0.850$	$AUC = 0.751$	$AUC = 0.730$
	$\epsilon = 3.5$	$acc_{bal} = 76.8\%$	$acc_{bal} = 68.7\%$	$acc_{bal} = 67.0\%$
	$l_{min} = 6$	$TPR = 70.9\%$	$TPR = 55.4\%$	$TPR = 53.7\%$
	$thres = 4.0651$	$FPR = 17.4\%$	$FPR = 18.1\%$	$FPR = 19.7\%$
	determined on 2438-A/C/D	scan B1	scan B2	scan B3
2438-B	$d = 9$	$r = 0.29$	$r = 0.26$	$r = 0.27$
	$\tau = 2$	$AUC = 0.668$	$AUC = 0.660$	$AUC = 0.665$
	$\epsilon = 4.0$	$acc_{bal} = 63.5\%$	$acc_{bal} = 62.4\%$	$acc_{bal} = 62.7\%$
	$l_{min} = 5$	$TPR = 58.5\%$	$TPR = 55.9\%$	$TPR = 60.1\%$
	$thres = 4.1622$	$FPR = 31.6\%$	$FPR = 31.1\%$	$FPR = 34.8\%$
	determined on 2438-A/B/D	scan C1	scan C2	scan C3
2438-C	$d = 5$	$r = -0.51$	$r = -0.52$	$r = -0.60$
	$\tau = 9$	$AUC = 0.794$	$AUC = 0.781$	$AUC = 0.833$
	$\epsilon = 0.7$	$acc_{bal} = 69.5\%$	$acc_{bal} = 69.6\%$	$acc_{bal} = 70.6\%$
	$l_{min} = 2$	$TPR = 89.3\%$	$TPR = 86.0\%$	$TPR = 93.0\%$
	$thres = 1.4399$	$FPR = 50.3\%$	$FPR = 46.8\%$	$FPR = 51.8\%$
	determined on 2438-A/B/C	scan D1	scan D2	scan D3
2438-D	$d = 10$	$r = 0.37$	$r = 0.39$	$r = 0.40$
	$\tau = 10$	$AUC = 0.778$	$AUC = 0.766$	$AUC = 0.757$
	$\epsilon = 8.0$	$acc_{bal} = 69.1\%$	$acc_{bal} = 68.0\%$	$acc_{bal} = 64.8\%$
	$l_{min} = 2$	$TPR = 74.1\%$	$TPR = 75.6\%$	$TPR = 78.1\%$
	$thres = 3.3156$	$FPR = 35.9\%$	$FPR = 39.6\%$	$FPR = 48.5\%$

Angular distance

Table C.62: Classification and correlation results, RQA, with *RR* as BWE-equivalent, using angular distance, for natural porosity area 2438A–D. Optimum RQA parameters and decision threshold determined on 3 areas (scan sets), test results with these parameters on scans of the remaining scan set.

results of specimen/ scan	parameters	test results		
	determined on 2438-B/C/D	scan A1	scan A2	scan A3
2438-A	$d = 9$	$r = -0.61$	$r = -0.48$	$r = -0.52$
	$\tau = 2$	$AUC = 0.924$	$AUC = 0.797$	$AUC = 0.825$
	$\epsilon = 0.100\pi$	$acc_{bal} = 83.9\%$	$acc_{bal} = 68.5\%$	$acc_{bal} = 68.8\%$
		$TPR = 77.0\%$	$TPR = 49.4\%$	$TPR = 47.7\%$
	$thres = 0.0113$	$FPR = 9.2\%$	$FPR = 12.3\%$	$FPR = 10.2\%$
	determined on 2438-A/C/D	scan B1	scan B2	scan B3
2438-B	$d = 9$	$r = -0.47$	$r = -0.47$	$r = -0.53$
	$\tau = 2$	$AUC = 0.793$	$AUC = 0.785$	$AUC = 0.831$
	$\epsilon = 0.100\pi$	$acc_{bal} = 64.9\%$	$acc_{bal} = 65.8\%$	$acc_{bal} = 69.7\%$
		$TPR = 37.7\%$	$TPR = 40.8\%$	$TPR = 46.9\%$
	$thres = 0.0117$	$FPR = 7.9\%$	$FPR = 9.3\%$	$FPR = 7.5\%$
	determined on 2438-A/B/D	scan C1	scan C2	scan C3
2438-C	$d = 10$	$r = -0.41$	$r = -0.34$	$r = -0.35$
	$\tau = 7$	$AUC = 0.811$	$AUC = 0.743$	$AUC = 0.795$
	$\epsilon = 0.125\pi$	$acc_{bal} = 71.2\%$	$acc_{bal} = 70.2\%$	$acc_{bal} = 69.2\%$
		$TPR = 97.8\%$	$TPR = 92.8\%$	$TPR = 99.4\%$
	$thres = 0.0071$	$FPR = 55.3\%$	$FPR = 52.4\%$	$FPR = 61.0\%$
	determined on 2438-A/B/C	scan D1	scan D2	scan D3
2438-D	$d = 10$	$r = -0.22$	$r = -0.53$	$r = -0.50$
	$\tau = 7$	$AUC = 0.795$	$AUC = 0.905$	$AUC = 0.889$
	$\epsilon = 0.125\pi$	$acc_{bal} = 76.6\%$	$acc_{bal} = 80.4\%$	$acc_{bal} = 75.9\%$
		$TPR = 80.1\%$	$TPR = 100.0\%$	$TPR = 99.9\%$
	$thres = 0.0072$	$FPR = 26.9\%$	$FPR = 39.2\%$	$FPR = 48.2\%$

Table C.63: Classification and correlation results, RQA, with *DIV* as BWE-equivalent, using angular distance, for natural porosity area 2438A–D. Optimum RQA parameters and decision threshold determined on 3 areas (scan sets), test results with these parameters on scans of the remaining scan set.

results of specimen/ scan	parameters	test results		
	determined on 2438-B/C/D	scan A1	scan A2	scan A3
2438-A	$d = 10$	$r = 0.45$	$r = 0.32$	$r = 0.40$
	$\tau = 8$	$AUC = 0.906$	$AUC = 0.764$	$AUC = 0.764$
	$\epsilon = 0.250\pi$	$acc_{bal} = 84.6\%$	$acc_{bal} = 70.0\%$	$acc_{bal} = 64.2\%$
		$TPR = 77.7\%$	$TPR = 52.9\%$	$TPR = 43.9\%$
	$thres = 0.0046$	$FPR = 8.4\%$	$FPR = 12.9\%$	$FPR = 15.4\%$
	determined on 2438-A/C/D	scan B1	scan B2	scan B3
2438-B	$d = 9$	$r = 0.48$	$r = 0.45$	$r = 0.55$
	$\tau = 8$	$AUC = 0.816$	$AUC = 0.839$	$AUC = 0.883$
	$\epsilon = 0.250\pi$	$acc_{bal} = 66.6\%$	$acc_{bal} = 76.5\%$	$acc_{bal} = 76.0\%$
		$TPR = 43.2\%$	$TPR = 64.0\%$	$TPR = 63.9\%$
	$thres = 0.0045$	$FPR = 10.1\%$	$FPR = 11.0\%$	$FPR = 11.8\%$
	determined on 2438-A/B/D	scan C1	scan C2	scan C3
2438-C	$d = 9$	$r = 0.49$	$r = 0.50$	$r = 0.52$
	$\tau = 6$	$AUC = 0.879$	$AUC = 0.823$	$AUC = 0.881$
	$\epsilon = 0.125\pi$	$acc_{bal} = 73.2\%$	$acc_{bal} = 73.6\%$	$acc_{bal} = 71.8\%$
		$TPR = 98.6\%$	$TPR = 95.0\%$	$TPR = 99.6\%$
	$thres = 0.0179$	$FPR = 52.2\%$	$FPR = 47.9\%$	$FPR = 56.1\%$
	determined on 2438-A/B/C	scan D1	scan D2	scan D3
2438-D	$d = 9$	$r = 0.42$	$r = 0.44$	$r = 0.49$
	$\tau = 6$	$AUC = 0.824$	$AUC = 0.939$	$AUC = 0.924$
	$\epsilon = 0.125\pi$	$acc_{bal} = 76.9\%$	$acc_{bal} = 81.4\%$	$acc_{bal} = 77.7\%$
		$TPR = 78.6\%$	$TPR = 99.9\%$	$TPR = 99.8\%$
	$thres = 0.0172$	$FPR = 24.9\%$	$FPR = 37.1\%$	$FPR = 44.4\%$

Table C.64: Classification and correlation results, RQA, with *TND* as BWE-equivalent, using angular distance, for natural porosity area 2438A–D. Optimum RQA parameters and decision threshold determined on 3 areas (scan sets), test results with these parameters on scans of the remaining scan set.

results of specimen/ scan	parameters	test results		
	determined on 2438-B/C/D	scan A1	scan A2	scan A3
2438-A	$d = 10$ $\tau = 5$ $\epsilon = 0.100\pi$ $thres = -0.0487$	$r = 0.67$ $AUC = 0.967$ $acc_{bal} = 85.8\%$ $TPR = 74.1\%$ $FPR = 2.4\%$	$r = 0.56$ $AUC = 0.841$ $acc_{bal} = 63.1\%$ $TPR = 31.9\%$ $FPR = 5.7\%$	$r = 0.61$ $AUC = 0.887$ $acc_{bal} = 64.2\%$ $TPR = 32.8\%$ $FPR = 4.5\%$
	determined on 2438-A/C/D	scan B1	scan B2	scan B3
2438-B	$d = 10$ $\tau = 10$ $\epsilon = 0.075\pi$ $thres = -0.0512$	$r = 0.55$ $AUC = 0.831$ $acc_{bal} = 67.8\%$ $TPR = 44.3\%$ $FPR = 8.8\%$	$r = 0.54$ $AUC = 0.823$ $acc_{bal} = 68.0\%$ $TPR = 46.2\%$ $FPR = 10.1\%$	$r = 0.60$ $AUC = 0.866$ $acc_{bal} = 71.0\%$ $TPR = 49.8\%$ $FPR = 7.9\%$
	determined on 2438-A/B/D	scan C1	scan C2	scan C3
2438-C	$d = 10$ $\tau = 5$ $\epsilon = 0.100\pi$ $thres = -0.0395$	$r = 0.65$ $AUC = 0.896$ $acc_{bal} = 70.8\%$ $TPR = 98.7\%$ $FPR = 57.2\%$	$r = 0.61$ $AUC = 0.831$ $acc_{bal} = 72.3\%$ $TPR = 95.4\%$ $FPR = 50.9\%$	$r = 0.54$ $AUC = 0.879$ $acc_{bal} = 70.4\%$ $TPR = 99.9\%$ $FPR = 59.2\%$
	determined on 2438-A/B/C	scan D1	scan D2	scan D3
2438-D	$d = 10$ $\tau = 5$ $\epsilon = 0.100\pi$ $thres = -0.0416$	$r = 0.44$ $AUC = 0.823$ $acc_{bal} = 77.7\%$ $TPR = 80.5\%$ $FPR = 25.1\%$	$r = 0.69$ $AUC = 0.942$ $acc_{bal} = 81.8\%$ $TPR = 99.9\%$ $FPR = 36.3\%$	$r = 0.69$ $AUC = 0.938$ $acc_{bal} = 77.2\%$ $TPR = 100.0\%$ $FPR = 45.6\%$

Table C.65: Classification and correlation results, RQA, with *COR* as BWE-equivalent, using angular distance, for natural porosity area 2438A–D. Optimum RQA parameters and decision threshold determined on 3 areas (scan sets), test results with these parameters on scans of the remaining scan set.

results of specimen/ scan	parameters	test results		
	determined on 2438-B/C/D	scan A1	scan A2	scan A3
2438-A	$d = 8$	$r = 0.50$	$r = 0.15$	$r = 0.23$
	$\tau = 1$	$AUC = 0.787$	$AUC = 0.580$	$AUC = 0.643$
	$\epsilon = 0.025\pi$	$acc_{bal} = 73.3\%$	$acc_{bal} = 56.2\%$	$acc_{bal} = 58.3\%$
		$TPR = 61.1\%$	$TPR = 32.7\%$	$TPR = 30.3\%$
	$thres = -0.1281$	$FPR = 14.5\%$	$FPR = 20.4\%$	$FPR = 13.6\%$
	determined on 2438-A/C/D	scan B1	scan B2	scan B3
2438-B	$d = 8$	$r = 0.17$	$r = 0.14$	$r = 0.21$
	$\tau = 2$	$AUC = 0.579$	$AUC = 0.563$	$AUC = 0.597$
	$\epsilon = 0.025\pi$	$acc_{bal} = 58.4\%$	$acc_{bal} = 56.7\%$	$acc_{bal} = 60.0\%$
		$TPR = 36.3\%$	$TPR = 35.8\%$	$TPR = 40.7\%$
	$thres = -0.0001$	$FPR = 19.6\%$	$FPR = 22.4\%$	$FPR = 20.7\%$
	determined on 2438-A/B/D	scan C1	scan C2	scan C3
2438-C	$d = 8$	$r = 0.36$	$r = 0.33$	$r = 0.38$
	$\tau = 10$	$AUC = 0.679$	$AUC = 0.653$	$AUC = 0.684$
	$\epsilon = 0.025\pi$	$acc_{bal} = 67.5\%$	$acc_{bal} = 64.7\%$	$acc_{bal} = 67.6\%$
		$TPR = 65.7\%$	$TPR = 64.5\%$	$TPR = 74.6\%$
	$thres = -0.0066$	$FPR = 30.8\%$	$FPR = 35.1\%$	$FPR = 39.5\%$
	determined on 2438-A/B/C	scan D1	scan D2	scan D3
2438-D	$d = 8$	$r = 0.29$	$r = 0.47$	$r = 0.46$
	$\tau = 10$	$AUC = 0.652$	$AUC = 0.758$	$AUC = 0.737$
	$\epsilon = 0.025\pi$	$acc_{bal} = 64.8\%$	$acc_{bal} = 74.4\%$	$acc_{bal} = 72.4\%$
		$TPR = 56.4\%$	$TPR = 83.1\%$	$TPR = 78.9\%$
	$thres = -0.0066$	$FPR = 26.9\%$	$FPR = 34.4\%$	$FPR = 34.2\%$

Table C.66: Classification and correlation results, RQA, with *DET* as BWE-equivalent, using angular distance, for natural porosity area 2438A–D. Optimum RQA parameters and decision threshold determined on 3 areas (scan sets), test results with these parameters on scans of the remaining scan set.

results of specimen/ scan	parameters	test results		
	determined on 2438-B/C/D	scan A1	scan A2	scan A3
2438-A	$d = 5$	$r = -0.75$	$r = -0.60$	$r = -0.57$
	$\tau = 8$	$AUC = 0.948$	$AUC = 0.863$	$AUC = 0.858$
	$\epsilon = 0.2\pi$	$acc_{bal} = 81.0\%$	$acc_{bal} = 65.2\%$	$acc_{bal} = 63.7\%$
	$l_{min} = 6$	$TPR = 65.0\%$	$TPR = 35.5\%$	$TPR = 31.8\%$
	$thres = 0.3393$	$FPR = 2.9\%$	$FPR = 5.1\%$	$FPR = 4.4\%$
	determined on 2438-A/C/D	scan B1	scan B2	scan B3
2438-B	$d = 6$	$r = -0.18$	$r = NaN$	$r = -0.22$
	$\tau = 6$	$AUC = 0.602$	$AUC = 0.396$	$AUC = 0.630$
	$\epsilon = 0.1\pi$	$acc_{bal} = 55.4\%$	$acc_{bal} = 45.0\%$	$acc_{bal} = 56.9\%$
	$l_{min} = 6$	$TPR = 24.2\%$	$TPR = 73.2\%$	$TPR = 27.2\%$
	$thres = 0.3973$	$FPR = 13.4\%$	$FPR = 83.2\%$	$FPR = 13.3\%$
	determined on 2438-A/B/D	scan C1	scan C2	scan C3
2438-C	$d = 5$	$r = -0.73$	$r = -0.69$	$r = -0.75$
	$\tau = 8$	$AUC = 0.927$	$AUC = 0.877$	$AUC = 0.917$
	$\epsilon = 0.2\pi$	$acc_{bal} = 83.0\%$	$acc_{bal} = 78.6\%$	$acc_{bal} = 79.3\%$
	$l_{min} = 6$	$TPR = 92.8\%$	$TPR = 86.6\%$	$TPR = 97.0\%$
	$thres = 0.3306$	$FPR = 26.8\%$	$FPR = 29.3\%$	$FPR = 38.5\%$
	determined on 2438-A/B/C	scan D1	scan D2	scan D3
2438-D	$d = 4$	$r = -0.63$	$r = -0.79$	$r = -0.80$
	$\tau = 8$	$AUC = 0.897$	$AUC = 0.973$	$AUC = 0.964$
	$\epsilon = 0.2\pi$	$acc_{bal} = 80.3\%$	$acc_{bal} = 88.5\%$	$acc_{bal} = 87.8\%$
	$l_{min} = 6$	$TPR = 71.7\%$	$TPR = 99.2\%$	$TPR = 98.3\%$
	$thres = 0.2485$	$FPR = 11.1\%$	$FPR = 22.2\%$	$FPR = 22.7\%$

Table C.67: Classification and correlation results, RQA, with *RATIO* as BWE-equivalent, using angular distance, for natural porosity area 2438A–D. Optimum RQA parameters and decision threshold determined on 3 areas (scan sets), test results with these parameters on scans of the remaining scan set.

results of specimen/ scan	parameters	test results		
	determined on 2438-B/C/D	scan A1	scan A2	scan A3
2438-A	$d = 4$	$r = -0.78$	$r = -0.68$	$r = -0.65$
	$\tau = 8$	$AUC = 0.964$	$AUC = 0.903$	$AUC = 0.895$
	$\epsilon = 0.6\pi$	$acc_{bal} = 89.5\%$	$acc_{bal} = 81.1\%$	$acc_{bal} = 77.5\%$
	$l_{min} = 10$	$TPR = 85.7\%$	$TPR = 71.9\%$	$TPR = 63.9\%$
	$thres = 1.0315$	$FPR = 6.7\%$	$FPR = 9.7\%$	$FPR = 9.0\%$
	determined on 2438-A/C/D	scan B1	scan B2	scan B3
2438-B	$d = 3$	$r = -0.62$	$r = -0.62$	$r = -0.66$
	$\tau = 8$	$AUC = 0.910$	$AUC = 0.902$	$AUC = 0.935$
	$\epsilon = 0.2\pi$	$acc_{bal} = 55.3\%$	$acc_{bal} = 53.9\%$	$acc_{bal} = 57.6\%$
	$l_{min} = 6$	$TPR = 99.9\%$	$TPR = 100.0\%$	$TPR = 99.8\%$
	$thres = 2.6083$	$FPR = 89.3\%$	$FPR = 92.3\%$	$FPR = 84.5\%$
	determined on 2438-A/B/D	scan C1	scan C2	scan C3
2438-C	$d = 5$	$r = -0.69$	$r = -0.68$	$r = -0.70$
	$\tau = 8$	$AUC = 0.923$	$AUC = 0.905$	$AUC = 0.913$
	$\epsilon = 0.6\pi$	$acc_{bal} = 79.8\%$	$acc_{bal} = 79.4\%$	$acc_{bal} = 78.5\%$
	$l_{min} = 10$	$TPR = 94.7\%$	$TPR = 92.2\%$	$TPR = 96.1\%$
	$thres = 1.0114$	$FPR = 35.1\%$	$FPR = 33.5\%$	$FPR = 39.1\%$
	determined on 2438-A/B/C	scan D1	scan D2	scan D3
2438-D	$d = 5$	$r = -0.64$	$r = -0.67$	$r = -0.74$
	$\tau = 8$	$AUC = 0.910$	$AUC = 0.943$	$AUC = 0.955$
	$\epsilon = 0.6\pi$	$acc_{bal} = 84.1\%$	$acc_{bal} = 86.1\%$	$acc_{bal} = 86.9\%$
	$l_{min} = 10$	$TPR = 82.7\%$	$TPR = 94.6\%$	$TPR = 95.3\%$
	$thres = 1.0200$	$FPR = 14.5\%$	$FPR = 22.3\%$	$FPR = 21.4\%$

Table C.68: Classification and correlation results, RQA, with L_{nor} as BWE-equivalent, using angular distance, for natural porosity area 2438A–D. Optimum RQA parameters and decision threshold determined on 3 areas (scan sets), test results with these parameters on scans of the remaining scan set.

results of specimen/ scan	parameters	test results		
	determined on 2438-B/C/D	scan A1	scan A2	scan A3
2438-A	$d = 5$	$r = 0.77$	$r = 0.60$	$r = 0.58$
	$\tau = 8$	$AUC = 0.952$	$AUC = 0.860$	$AUC = 0.856$
	$\epsilon = 0.2\pi$	$acc_{bal} = 80.2\%$	$acc_{bal} = 64.2\%$	$acc_{bal} = 62.8\%$
	$l_{min} = 2$	$TPR = 62.9\%$	$TPR = 32.5\%$	$TPR = 29.8\%$
	$thres = 0.4750$	$FPR = 2.4\%$	$FPR = 4.0\%$	$FPR = 4.1\%$
	determined on 2438-A/C/D	scan B1	scan B2	scan B3
2438-B	$d = 6$	$r = NaN$	$r = NaN$	$r = NaN$
	$\tau = 6$	$AUC = 0.601$	$AUC = 0.604$	$AUC = 0.633$
	$\epsilon = 0.1\pi$	$acc_{bal} = 56.3\%$	$acc_{bal} = 56.3\%$	$acc_{bal} = 59.1\%$
	$l_{min} = 2$	$TPR = 34.4\%$	$TPR = 38.5\%$	$TPR = 39.2\%$
	$thres = 0.4131$	$FPR = 21.9\%$	$FPR = 25.9\%$	$FPR = 21.0\%$
	determined on 2438-A/B/D	scan C1	scan C2	scan C3
2438-C	$d = 4$	$r = 0.74$	$r = 0.68$	$r = 0.75$
	$\tau = 8$	$AUC = 0.932$	$AUC = 0.875$	$AUC = 0.917$
	$\epsilon = 0.2\pi$	$acc_{bal} = 84.7\%$	$acc_{bal} = 79.5\%$	$acc_{bal} = 80.2\%$
	$l_{min} = 2$	$TPR = 92.9\%$	$TPR = 86.7\%$	$TPR = 96.4\%$
	$thres = 0.5250$	$FPR = 23.5\%$	$FPR = 27.6\%$	$FPR = 36.1\%$
	determined on 2438-A/B/C	scan D1	scan D2	scan D3
2438-D	$d = 4$	$r = 0.60$	$r = 0.78$	$r = 0.79$
	$\tau = 8$	$AUC = 0.883$	$AUC = 0.966$	$AUC = 0.958$
	$\epsilon = 0.2\pi$	$acc_{bal} = 79.6\%$	$acc_{bal} = 87.2\%$	$acc_{bal} = 86.8\%$
	$l_{min} = 2$	$TPR = 71.7\%$	$TPR = 99.1\%$	$TPR = 98.9\%$
	$thres = 0.5215$	$FPR = 12.6\%$	$FPR = 24.8\%$	$FPR = 25.3\%$

Table C.69: Classification and correlation results, RQA, with *ENT* as BWE-equivalent, using angular distance, for natural porosity area 2438A–D. Optimum RQA parameters and decision threshold determined on 3 areas (scan sets), test results with these parameters on scans of the remaining scan set.

results of specimen/ scan	parameters	test results		
	determined on 2438-B/C/D	scan A1	scan A2	scan A3
2438-A	$d = 3$	$r = -0.74$	$r = -0.63$	$r = -0.62$
	$\tau = 8$	$AUC = 0.943$	$AUC = 0.869$	$AUC = 0.876$
	$\epsilon = 0.1\pi$	$acc_{bal} = 86.1\%$	$acc_{bal} = 76.3\%$	$acc_{bal} = 74.2\%$
	$l_{min} = 3$	$TPR = 77.4\%$	$TPR = 62.5\%$	$TPR = 56.2\%$
	$thres = 1.6072$	$FPR = 5.2\%$	$FPR = 10.0\%$	$FPR = 7.7\%$
	determined on 2438-A/C/D	scan B1	scan B2	scan B3
2438-B	$d = 2$	$r = -0.36$	$r = -0.35$	$r = -0.43$
	$\tau = 8$	$AUC = 0.710$	$AUC = 0.695$	$AUC = 0.751$
	$\epsilon = 0.2\pi$	$acc_{bal} = 58.7\%$	$acc_{bal} = 58.2\%$	$acc_{bal} = 61.0\%$
	$l_{min} = 2$	$TPR = 24.8\%$	$TPR = 31.0\%$	$TPR = 31.6\%$
	$thres = 1.8614$	$FPR = 7.4\%$	$FPR = 14.7\%$	$FPR = 9.5\%$
	determined on 2438-A/B/D	scan C1	scan C2	scan C3
2438-C	$d = 8$	$r = -0.66$	$r = -0.52$	$r = -0.58$
	$\tau = 6$	$AUC = 0.895$	$AUC = 0.787$	$AUC = 0.834$
	$\epsilon = 0.6\pi$	$acc_{bal} = 78.5\%$	$acc_{bal} = 69.7\%$	$acc_{bal} = 69.8\%$
	$l_{min} = 4$	$TPR = 93.4\%$	$TPR = 87.6\%$	$TPR = 95.8\%$
	$thres = 3.6385$	$FPR = 36.3\%$	$FPR = 48.2\%$	$FPR = 56.3\%$
	determined on 2438-A/B/C	scan D1	scan D2	scan D3
2438-D	$d = 8$	$r = -0.47$	$r = -0.67$	$r = -0.69$
	$\tau = 6$	$AUC = 0.849$	$AUC = 0.921$	$AUC = 0.914$
	$\epsilon = 0.6\pi$	$acc_{bal} = 79.6\%$	$acc_{bal} = 82.4\%$	$acc_{bal} = 77.7\%$
	$l_{min} = 4$	$TPR = 84.2\%$	$TPR = 98.5\%$	$TPR = 98.0\%$
	$thres = 3.6498$	$FPR = 25.1\%$	$FPR = 33.6\%$	$FPR = 42.7\%$

C.4.2 Evaluation in time domain

Table C.70: Classification and correlation results with x_{max} as BWE-equivalent for natural porosity area 2438A–D. Optimum decision threshold determined on 3 areas (scan sets), test results with these parameters on scans of the remaining scan set.

results of specimen/ scan	decision threshold	test results		
	determined on 2438-B/C/D	scan A1	scan A2	scan A3
2438-A	*thres*= 9.8633	$r = -0.68$ $AUC = 0.890$ $acc_{bal} = 83.1\%$ $TPR = 74.3\%$ $FPR = 8.1\%$	$r = -0.09$ $AUC = 0.570$ $acc_{bal} = 48.7\%$ $TPR = 11.2\%$ $FPR = 13.8\%$	$r = -0.38$ $AUC = 0.771$ $acc_{bal} = 53.6\%$ $TPR = 19.2\%$ $FPR = 11.9\%$
	determined on 2438-A/C/D	scan B1	scan B2	scan B3
2438-B	*thres*= 10.5225	$r = -0.28$ $AUC = 0.685$ $acc_{bal} = 50.0\%$ $TPR = 4.5\%$ $FPR = 4.6\%$	$r = -0.20$ $AUC = 0.644$ $acc_{bal} = 47.5\%$ $TPR = 4.6\%$ $FPR = 9.7\%$	$r = -0.39$ $AUC = 0.747$ $acc_{bal} = 49.5\%$ $TPR = 6.2\%$ $FPR = 7.1\%$
	determined on 2438-A/B/D	scan C1	scan C2	scan C3
2438-C	*thres*= 7.7881	$r = -0.33$ $AUC = 0.650$ $acc_{bal} = 60.6\%$ $TPR = 97.3\%$ $FPR = 76.1\%$	$r = 0.00$ $AUC = 0.572$ $acc_{bal} = 43.9\%$ $TPR = 8.9\%$ $FPR = 21.1\%$	$r = 0.02$ $AUC = 0.581$ $acc_{bal} = 42.2\%$ $TPR = 3.7\%$ $FPR = 19.3\%$
	determined on 2438-A/B/C	scan D1	scan D2	scan D3
2438-D	*thres*= 8.2153	$r = -0.35$ $AUC = 0.749$ $acc_{bal} = 68.5\%$ $TPR = 80.7\%$ $FPR = 43.7\%$	$r = -0.44$ $AUC = 0.751$ $acc_{bal} = 76.9\%$ $TPR = 99.3\%$ $FPR = 45.5\%$	$r = -0.35$ $AUC = 0.689$ $acc_{bal} = 67.2\%$ $TPR = 99.3\%$ $FPR = 64.9\%$

Table C.71: Classification and correlation results with s^2 as BWE-equivalent for natural porosity area 2438A–D. Optimum decision threshold determined on 3 areas (scan sets), test results with these parameters on scans of the remaining scan set.

results of specimen/ scan	decision threshold	test results		
	determined on 2438-B/C/D	scan A1	scan A2	scan A3
2438-A	*thres*= 3.3631	$r = -0.75$ *AUC*= 0.936 $acc_{bal} = 86.9\%$ *TPR* = 78.7% *FPR* = 4.8%	$r = -0.28$ *AUC*= 0.672 $acc_{bal} = 55.5\%$ *TPR* = 19.0% *FPR* = 8.0%	$r = -0.57$ *AUC*= 0.864 $acc_{bal} = 62.5\%$ *TPR* = 28.0% *FPR* = 2.9%
	determined on 2438-A/C/D	scan B1	scan B2	scan B3
2438-B	*thres*= 3.6549	$r = -0.34$ *AUC*= 0.697 $acc_{bal} = 53.3\%$ *TPR* = 8.4% *FPR* = 1.8%	$r = -0.33$ *AUC*= 0.690 $acc_{bal} = 53.5\%$ *TPR* = 9.4% *FPR* = 2.4%	$r = -0.49$ *AUC*= 0.799 $acc_{bal} = 55.4\%$ *TPR* = 12.3% *FPR* = 1.5%
	determined on 2438-A/B/D	scan C1	scan C2	scan C3
2438-C	*thres*= 2.5704	$r = -0.51$ *AUC*= 0.761 $acc_{bal} = 63.6\%$ *TPR* = 98.3% *FPR* = 71.2%	$r = -0.11$ *AUC*= 0.488 $acc_{bal} = 58.2\%$ *TPR* = 93.2% *FPR* = 76.8%	$r = -0.09$ *AUC*= 0.464 $acc_{bal} = 60.0\%$ *TPR* = 97.3% *FPR* = 77.3%
	determined on 2438-A/B/C	scan D1	scan D2	scan D3
2438-D	*thres*= 2.7209	$r = -0.50$ *AUC*= 0.837 $acc_{bal} = 74.7\%$ *TPR* = 88.2% *FPR* = 38.8%	$r = -0.53$ *AUC*= 0.797 $acc_{bal} = 78.1\%$ *TPR* = 99.6% *FPR* = 43.3%	$r = -0.46$ *AUC*= 0.741 $acc_{bal} = 69.1\%$ *TPR* = 99.2% *FPR* = 61.0%

Table C.72: Classification and correlation results with QCD as BWE-equivalent for natural porosity area 2438A–D. Optimum decision threshold determined on 3 areas (scan sets), test results with these parameters on scans of the remaining scan set.

results of specimen/ scan	decision threshold	test results		
	determined on 2438-B/C/D	scan A1	scan A2	scan A3
2438-A	*thres*= 0.0933	$r = 0.17$ $AUC = 0.607$ $acc_{bal} = 59.3\%$ $TPR = 48.9\%$ $FPR = 30.3\%$	$r = -0.02$ $AUC = 0.515$ $acc_{bal} = 50.7\%$ $TPR = 69.5\%$ $FPR = 68.2\%$	$r = 0.11$ $AUC = 0.563$ $acc_{bal} = 54.6\%$ $TPR = 29.8\%$ $FPR = 20.5\%$
	determined on 2438-A/C/D	scan B1	scan B2	scan B3
2438-B	*thres*= 0.0934	$r = 0.11$ $AUC = 0.570$ $acc_{bal} = 57.2\%$ $TPR = 35.1\%$ $FPR = 20.6\%$	$r = 0.34$ $AUC = 0.719$ $acc_{bal} = 68.0\%$ $TPR = 57.6\%$ $FPR = 21.6\%$	$r = 0.26$ $AUC = 0.667$ $acc_{bal} = 64.4\%$ $TPR = 44.0\%$ $FPR = 15.1\%$
	determined on 2438-A/B/D	scan C1	scan C2	scan C3
2438-C	*thres*= 0.1014	$r = -0.15$ $AUC = 0.646$ $acc_{bal} = 56.7\%$ $TPR = 35.6\%$ $FPR = 22.2\%$	$r = -0.16$ $AUC = 0.639$ $acc_{bal} = 56.6\%$ $TPR = 40.2\%$ $FPR = 27.1\%$	$r = -0.20$ $AUC = 0.678$ $acc_{bal} = 60.2\%$ $TPR = 45.4\%$ $FPR = 24.9\%$
	determined on 2438-A/B/C	scan D1	scan D2	scan D3
2438-D	*thres*= 0.1008	$r = -0.18$ $AUC = 0.579$ $acc_{bal} = 55.7\%$ $TPR = 51.3\%$ $FPR = 39.9\%$	$r = 0.18$ $AUC = 0.633$ $acc_{bal} = 57.4\%$ $TPR = 76.4\%$ $FPR = 61.7\%$	$r = 0.06$ $AUC = 0.570$ $acc_{bal} = 51.9\%$ $TPR = 72.3\%$ $FPR = 68.4\%$

Table C.73: Classification and correlation results with linear regression in time domain for natural porosity area 2438A–D. Training (including decision threshold to achieve $TPR \approx TNR$) determined on 3 areas (scan sets), test results on scans of the remaining scan set.

results of specimen/ scan	decision threshold	test results		
	determined on 2438-B/C/D	scan A1	scan A2	scan A3
2438-A	*thres*= 22.3215	$r = 0.65$ $AUC = 0.880$ $acc_{bal} = 78.4\%$ $TPR = 62.3\%$ $FPR = 5.5\%$	$r = 0.63$ $AUC = 0.819$ $acc_{bal} = 74.8\%$ $TPR = 69.4\%$ $FPR = 19.8\%$	$r = 0.70$ $AUC = 0.900$ $acc_{bal} = 79.2\%$ $TPR = 62.3\%$ $FPR = 4.0\%$
	determined on 2438-A/C/D	scan B1	scan B2	scan B3
2438-B	*thres*= 22.4868	$r = 0.60$ $AUC = 0.850$ $acc_{bal} = 77.5\%$ $TPR = 69.1\%$ $FPR = 14.1\%$	$r = 0.63$ $AUC = 0.864$ $acc_{bal} = 76.8\%$ $TPR = 83.1\%$ $FPR = 29.4\%$	$r = 0.76$ $AUC = 0.944$ $acc_{bal} = 88.2\%$ $TPR = 83.4\%$ $FPR = 6.9\%$
	determined on 2438-A/B/D	scan C1	scan C2	scan C3
2438-C	*thres*= 21.9709	$r = 0.68$ $AUC = 0.915$ $acc_{bal} = 81.5\%$ $TPR = 89.8\%$ $FPR = 26.7\%$	$r = 0.72$ $AUC = 0.904$ $acc_{bal} = 83.8\%$ $TPR = 79.1\%$ $FPR = 11.4\%$	$r = 0.75$ $AUC = 0.927$ $acc_{bal} = 83.1\%$ $TPR = 91.6\%$ $FPR = 25.4\%$
	determined on 2438-A/B/C	scan D1	scan D2	scan D3
2438-D	*thres*= 22.9054	$r = 0.08$ $AUC = 0.491$ $acc_{bal} = 48.4\%$ $TPR = 38.2\%$ $FPR = 41.5\%$	$r = 0.80$ $AUC = 0.947$ $acc_{bal} = 72.5\%$ $TPR = 98.4\%$ $FPR = 53.4\%$	$r = 0.82$ $AUC = 0.954$ $acc_{bal} = 79.2\%$ $TPR = 98.9\%$ $FPR = 40.5\%$

C.4.3 Fourier analysis

Table C.74: Classification and correlation results with maximum frequency, left and right cutoff frequency as BWE-equivalent on scan sets 2438A–D. Decision threshold to achieve $TPR \approx TNR$ determined on 3 sets, results with this threshold on remaining set; results of all three scans of a set combined.

results of specimen/ scan	F_{peak}	F_{cl}	F_{cr}
2438-A	threshold to achieve $TPR \approx TNR$ on 2438-B/C/D		
	$thres = 4.3333$	$thres = 4.1933$	$thres = 4.8313$
	$r = 0.34$ $AUC = 0.712$ $acc_{bal} = 61.9\%$ $TPR = 42.9\%$ $FPR = 19.1\%$	$r = 0.33$ $AUC = 0.711$ $acc_{bal} = 67.1\%$ $TPR = 53.8\%$ $FPR = 19.6\%$	$r = 0.31$ $AUC = 0.699$ $acc_{bal} = 67.9\%$ $TPR = 56.4\%$ $FPR = 20.5\%$
2438-B	threshold to achieve $TPR \approx TNR$ on 2438-A/C/D		
	$thres = 4.3333$	$thres = 4.2169$	$thres = 4.8314$
	$r = 0.30$ $AUC = 0.684$ $acc_{bal} = 59.3\%$ $TPR = 40.6\%$ $FPR = 21.9\%$	$r = 0.29$ $AUC = 0.682$ $acc_{bal} = 66.7\%$ $TPR = 56.0\%$ $FPR = 22.6\%$	$r = 0.27$ $AUC = 0.678$ $acc_{bal} = 65.9\%$ $TPR = 54.4\%$ $FPR = 22.6\%$
2438-C	threshold to achieve $TPR \approx TNR$ on 2438-A/B/D		
	$thres = 5.0000$	$thres = 4.8226$	$thres = 5.3871$
	$r = 0.53$ $AUC = 0.797$ $acc_{bal} = 77.7\%$ $TPR = 87.0\%$ $FPR = 31.5\%$	$r = 0.51$ $AUC = 0.783$ $acc_{bal} = 77.7\%$ $TPR = 88.1\%$ $FPR = 32.8\%$	$r = 0.52$ $AUC = 0.804$ $acc_{bal} = 77.5\%$ $TPR = 87.6\%$ $FPR = 32.6\%$
2438-D	threshold to achieve $TPR \approx TNR$ on 2438-A/B/C		
	$thres = 4.6667$	$thres = 4.7303$	$thres = 5.2464$
	$r = 0.58$ $AUC = 0.841$ $acc_{bal} = 81.1\%$ $TPR = 87.3\%$ $FPR = 25.1\%$	$r = 0.57$ $AUC = 0.833$ $acc_{bal} = 81.1\%$ $TPR = 87.8\%$ $FPR = 25.6\%$	$r = 0.57$ $AUC = 0.844$ $acc_{bal} = 80.9\%$ $TPR = 87.9\%$ $FPR = 26.0\%$

Table C.75: Classification and correlation results with mean frequency, bandwidth and integral of bandwidth as BWE-equivalent on scan sets 2438A–D. Decision threshold to achieve $TPR \approx TNR$ determined on 3 sets, results with this threshold on remaining set; results of all three scans of a set combined.

results of specimen/ scan	F_{centre}	B_{3dB}	A_{bw3}
2438-A	threshold to achieve $TPR \approx TNR$ on 2438-B/C/D		
	$thres = 4.4927$	$thres = 0.5363$	$thres = 60.6866$
	$r = 0.32$ $AUC = 0.705$ $acc_{bal} = 67.6\%$ $TPR = 55.0\%$ $FPR = 19.7\%$	$r = -0.03$ $AUC = 0.521$ $acc_{bal} = 51.5\%$ $TPR = 53.2\%$ $FPR = 50.2\%$	$r = -0.27$ $AUC = 0.658$ $acc_{bal} = 60.3\%$ $TPR = 61.1\%$ $FPR = 40.5\%$
2438-B	threshold to achieve $TPR \approx TNR$ on 2438-A/C/D		
	$thres = 4.5664$	$thres = 0.5362$	$thres = 62.5793$
	$r = 0.28$ $AUC = 0.679$ $acc_{bal} = 66.8\%$ $TPR = 56.1\%$ $FPR = 22.5\%$	$r = -0.09$ $AUC = 0.542$ $acc_{bal} = 51.9\%$ $TPR = 53.6\%$ $FPR = 49.9\%$	$r = -0.24$ $AUC = 0.617$ $acc_{bal} = 57.5\%$ $TPR = 50.7\%$ $FPR = 35.8\%$
2438-C	threshold to achieve $TPR \approx TNR$ on 2438-A/B/D		
	$thres = 5.0484$	$thres = 0.5336$	$thres = 58.8382$
	$r = 0.52$ $AUC = 0.795$ $acc_{bal} = 77.6\%$ $TPR = 87.8\%$ $FPR = 32.7\%$	$r = 0.37$ $AUC = 0.705$ $acc_{bal} = 63.9\%$ $TPR = 60.8\%$ $FPR = 33.1\%$	$r = 0.25$ $AUC = 0.640$ $acc_{bal} = 57.5\%$ $TPR = 49.5\%$ $FPR = 34.5\%$
2438-D	threshold to achieve $TPR \approx TNR$ on 2438-A/B/C		
	$thres = 4.9913$	$thres = 0.5536$	$thres = 61.7765$
	$r = 0.57$ $AUC = 0.840$ $acc_{bal} = 81.0\%$ $TPR = 87.9\%$ $FPR = 25.9\%$	$r = 0.41$ $AUC = 0.734$ $acc_{bal} = 64.1\%$ $TPR = 70.1\%$ $FPR = 41.9\%$	$r = 0.24$ $AUC = 0.645$ $acc_{bal} = 55.2\%$ $TPR = 58.6\%$ $FPR = 48.2\%$

Table C.76: Classification and correlation results with linear regression on amplitude spectrum from 0 MHz to 30 MHz for natural porosity area 2438A–D. Hann window applied on intermediate echo time series prior to Fourier transform. Training (including decision threshold to achieve $TPR \approx TNR$) determined on 3 areas (scan sets), test results on scans of the remaining scan set.

results of specimen/ scan	decision threshold	test results		
	determined on 2438-B/C/D	scan A1	scan A2	scan A3
2438-A	*thres*= 21.4479	$r = 0.68$ $AUC = 0.891$ $acc_{bal} = 82.4\%$ $TPR = 72.4\%$ $FPR = 7.5\%$	$r = 0.67$ $AUC = 0.876$ $acc_{bal} = 79.5\%$ $TPR = 69.0\%$ $FPR = 10.1\%$	$r = 0.68$ $AUC = 0.893$ $acc_{bal} = 79.4\%$ $TPR = 65.4\%$ $FPR = 6.5\%$
	determined on 2438-A/C/D	scan B1	scan B2	scan B3
2438-B	*thres*= 21.9638	$r = 0.77$ $AUC = 0.963$ $acc_{bal} = 88.2\%$ $TPR = 81.2\%$ $FPR = 4.8\%$	$r = 0.79$ $AUC = 0.963$ $acc_{bal} = 89.0\%$ $TPR = 84.2\%$ $FPR = 6.3\%$	$r = 0.79$ $AUC = 0.969$ $acc_{bal} = 89.1\%$ $TPR = 82.4\%$ $FPR = 4.1\%$
	determined on 2438-A/B/D	scan C1	scan C2	scan C3
2438-C	*thres*= 20.9806	$r = 0.79$ $AUC = 0.959$ $acc_{bal} = 86.5\%$ $TPR = 94.0\%$ $FPR = 21.0\%$	$r = 0.79$ $AUC = 0.961$ $acc_{bal} = 88.2\%$ $TPR = 93.7\%$ $FPR = 17.2\%$	$r = 0.81$ $AUC = 0.976$ $acc_{bal} = 89.9\%$ $TPR = 96.4\%$ $FPR = 16.6\%$
	determined on 2438-A/B/C	scan D1	scan D2	scan D3
2438-D	*thres*= 22.1091	$r = 0.73$ $AUC = 0.928$ $acc_{bal} = 84.2\%$ $TPR = 89.6\%$ $FPR = 21.2\%$	$r = 0.75$ $AUC = 0.946$ $acc_{bal} = 86.7\%$ $TPR = 89.5\%$ $FPR = 16.2\%$	$r = 0.79$ $AUC = 0.969$ $acc_{bal} = 88.9\%$ $TPR = 94.7\%$ $FPR = 17.0\%$

Table C.77: Classification and correlation results with linear regression on amplitude spectrum from 1 MHz to 30 MHz for natural porosity area 2438A–D. Hann window applied on intermediate echo time series prior to Fourier transform. Training (including decision threshold to achieve $TPR \approx TNR$) determined on 3 areas (scan sets), test results on scans of the remaining scan set.

results of specimen/ scan	decision threshold	test results		
	determined on 2438-B/C/D	scan A1	scan A2	scan A3
2438-A	*thres*= 21.5577	$r = 0.67$ $AUC = 0.882$ $acc_{bal} = 81.7\%$ $TPR = 76.1\%$ $FPR = 12.7\%$	$r = 0.69$ $AUC = 0.899$ $acc_{bal} = 81.9\%$ $TPR = 80.2\%$ $FPR = 16.4\%$	$r = 0.68$ $AUC = 0.899$ $acc_{bal} = 81.9\%$ $TPR = 78.1\%$ $FPR = 14.4\%$
	determined on 2438-A/C/D	scan B1	scan B2	scan B3
2438-B	*thres*= 21.9187	$r = 0.75$ $AUC = 0.955$ $acc_{bal} = 86.3\%$ $TPR = 77.9\%$ $FPR = 5.3\%$	$r = 0.76$ $AUC = 0.947$ $acc_{bal} = 85.8\%$ $TPR = 77.7\%$ $FPR = 6.1\%$	$r = 0.76$ $AUC = 0.951$ $acc_{bal} = 85.2\%$ $TPR = 76.4\%$ $FPR = 6.0\%$
	determined on 2438-A/B/D	scan C1	scan C2	scan C3
2438-C	*thres*= 20.9627	$r = 0.76$ $AUC = 0.948$ $acc_{bal} = 86.1\%$ $TPR = 92.3\%$ $FPR = 20.1\%$	$r = 0.76$ $AUC = 0.955$ $acc_{bal} = 87.9\%$ $TPR = 94.0\%$ $FPR = 18.2\%$	$r = 0.76$ $AUC = 0.962$ $acc_{bal} = 88.9\%$ $TPR = 92.7\%$ $FPR = 15.0\%$
	determined on 2438-A/B/C	scan D1	scan D2	scan D3
2438-D	*thres*= 22.1555	$r = 0.73$ $AUC = 0.937$ $acc_{bal} = 84.7\%$ $TPR = 90.4\%$ $FPR = 20.9\%$	$r = 0.70$ $AUC = 0.916$ $acc_{bal} = 83.8\%$ $TPR = 83.2\%$ $FPR = 15.5\%$	$r = 0.76$ $AUC = 0.957$ $acc_{bal} = 87.7\%$ $TPR = 91.4\%$ $FPR = 15.9\%$

C.4.4 Extended recurrence quantification analysis

Table C.78: Classification and correlation results, RQA, with *DET* as BWE-equivalent, using Euclidean distance, for natural porosity area 2438A–D. Computation with fixed RR between 1% and 30%. Optimum RQA parameters and decision threshold determined on 3 areas (scan sets), test results with these parameters on scans of the remaining scan set.

results of sample/ scan	parameters	test results		
	determined on 2438-B/C/D	scan A1	scan A2	scan A3
2438-A	$d = 4$ $\tau = 7$ $RR = 2\%$ $l_{min} = 4$ $thres = 0.5301$	$r = -0.74$ $AUC = 0.930$ $acc_{bal} = 85.1\%$ $TPR = 79.6\%$ $FPR = 9.3\%$	$r = -0.58$ $AUC = 0.843$ $acc_{bal} = 74.2\%$ $TPR = 62.9\%$ $FPR = 14.6\%$	$r = -0.63$ $AUC = 0.878$ $acc_{bal} = 76.0\%$ $TPR = 61.8\%$ $FPR = 9.8\%$
	determined on 2438-A/C/D	scan B1	scan B2	scan B3
2438-B	$d = 2$ $\tau = 7$ $RR = 4\%$ $l_{min} = 3$ $thres = 0.4585$	$r = -0.59$ $AUC = 0.859$ $acc_{bal} = 75.9\%$ $TPR = 62.8\%$ $FPR = 10.9\%$	$r = -0.60$ $AUC = 0.860$ $acc_{bal} = 77.1\%$ $TPR = 64.9\%$ $FPR = 10.7\%$	$r = -0.66$ $AUC = 0.895$ $acc_{bal} = 79.2\%$ $TPR = 67.0\%$ $FPR = 8.7\%$
	determined on 2438-A/B/D	scan C1	scan C2	scan C3
2438-C	$d = 2$ $\tau = 8$ $RR = 3\%$ $l_{min} = 3$ $thres = 0.3056$	$r = -0.76$ $AUC = 0.939$ $acc_{bal} = 81.1\%$ $TPR = 96.0\%$ $FPR = 33.8\%$	$r = -0.72$ $AUC = 0.894$ $acc_{bal} = 78.4\%$ $TPR = 94.4\%$ $FPR = 37.6\%$	$r = -0.77$ $AUC = 0.926$ $acc_{bal} = 76.8\%$ $TPR = 98.1\%$ $FPR = 44.6\%$
	determined on 2438-A/B/C	scan D1	scan D2	scan D3
2438-D	$d = 4$ $\tau = 8$ $RR = 3\%$ $l_{min} = 2$ $thres = 0.7995$	$r = -0.61$ $AUC = 0.903$ $acc_{bal} = 82.2\%$ $TPR = 84.6\%$ $FPR = 20.2\%$	$r = -0.69$ $AUC = 0.962$ $acc_{bal} = 84.8\%$ $TPR = 98.0\%$ $FPR = 28.4\%$	$r = -0.75$ $AUC = 0.963$ $acc_{bal} = 83.4\%$ $TPR = 98.4\%$ $FPR = 31.7\%$

Table C.79: Classification and correlation results, RQA, with *RATIO* as BWE-equivalent, using Euclidean distance, for natural porosity area 2438A–D. RQA based on cross recurrence plots, optimum RQA parameters and decision threshold determined on 3 areas (scan sets), test results with these parameters on scans of the remaining scan set. Reference for cross recurrence is the mean of intermediate echo time series out of 24-α1 of the area $x = 28$–30, $y = 20$–22

results of sample/ scan	parameters	test results		
	determined on 2438-B/C/D	scan A1	scan A2	scan A3
2438-A	$d = 4$	$r = -0.60$	$r = -0.56$	$r = -0.54$
	$\tau = 9$	$AUC = 0.966$	$AUC = 0.887$	$AUC = 0.903$
	$\epsilon = 1.1$	$acc_{bal} = 90.1\%$	$acc_{bal} = 79.8\%$	$acc_{bal} = 81.6\%$
	$l_{min} = 4$	$TPR = 84.0\%$	$TPR = 65.9\%$	$TPR = 68.2\%$
	$thres = 4.5655$	$FPR = 3.8\%$	$FPR = 6.3\%$	$FPR = 4.9\%$
	determined on 2438-A/C/D	scan B1	scan B2	scan B3
2438-B	$d = 4$	$r = -0.57$	$r = -0.57$	$r = -0.58$
	$\tau = 8$	$AUC = 0.881$	$AUC = 0.882$	$AUC = 0.903$
	$\epsilon = 1.2$	$acc_{bal} = 80.5\%$	$acc_{bal} = 79.2\%$	$acc_{bal} = 82.1\%$
	$l_{min} = 5$	$TPR = 76.0\%$	$TPR = 72.8\%$	$TPR = 77.1\%$
	$thres = 3.8861$	$FPR = 15.0\%$	$FPR = 14.4\%$	$FPR = 12.9\%$
	determined on 2438-A/B/D	scan C1	scan C2	scan C3
2438-C	$d = 4$	$r = -0.58$	$r = -0.57$	$r = -0.49$
	$\tau = 9$	$AUC = 0.962$	$AUC = 0.941$	$AUC = 0.949$
	$\epsilon = 0.8$	$acc_{bal} = 86.1\%$	$acc_{bal} = 83.7\%$	$acc_{bal} = 82.0\%$
	$l_{min} = 3$	$TPR = 94.6\%$	$TPR = 93.5\%$	$TPR = 97.1\%$
	$thres = 9.9512$	$FPR = 22.4\%$	$FPR = 26.1\%$	$FPR = 33.2\%$
	determined on 2438-A/B/C	scan D1	scan D2	scan D3
2438-D	$d = 4$	$r = -0.46$	$r = -0.56$	$r = -0.55$
	$\tau = 9$	$AUC = 0.950$	$AUC = 0.974$	$AUC = 0.963$
	$\epsilon = 0.8$	$acc_{bal} = 87.8\%$	$acc_{bal} = 89.1\%$	$acc_{bal} = 85.9\%$
	$l_{min} = 3$	$TPR = 88.7\%$	$TPR = 96.4\%$	$TPR = 96.2\%$
	$thres = 10.0717$	$FPR = 13.1\%$	$FPR = 18.2\%$	$FPR = 24.5\%$

Table C.80: Classification and correlation results, RQA, with *RATIO* as BWE-equivalent, using Euclidean distance, for natural porosity area 2438A–D. RQA based on joint recurrence plots, optimum RQA parameters and decision threshold determined on 3 areas (scan sets), test results with these parameters on scans of the remaining scan set. Reference for cross recurrence is the mean of intermediate echo time series out of 24-α1 of the area $x = 28$–30, $y = 20$–22

results of sample/ scan	parameters	test results		
	determined on 2438-B/C/D	scan A1	scan A2	scan A3
2438-A	$d = 3$	$r = -0.67$	$r = -0.61$	$r = -0.59$
	$\tau = 9$	$AUC = 0.965$	$AUC = 0.909$	$AUC = 0.918$
	$\epsilon = 1.0$	$acc_{bal} = 89.5\%$	$acc_{bal} = 83.1\%$	$acc_{bal} = 82.6\%$
	$l_{min} = 5$	$TPR = 84.9\%$	$TPR = 74.4\%$	$TPR = 73.1\%$
	$thres = 6.0092$	$FPR = 5.9\%$	$FPR = 8.3\%$	$FPR = 8.0\%$
	determined on 2438-A/C/D	scan B1	scan B2	scan B3
2438-B	$d = 3$	$r = -0.56$	$r = -0.58$	$r = -0.57$
	$\tau = 9$	$AUC = 0.851$	$AUC = 0.869$	$AUC = 0.864$
	$\epsilon = 1.1$	$acc_{bal} = 77.5\%$	$acc_{bal} = 78.2\%$	$acc_{bal} = 78.6\%$
	$l_{min} = 4$	$TPR = 75.3\%$	$TPR = 71.7\%$	$TPR = 77.3\%$
	$thres = 6.0056$	$FPR = 20.3\%$	$FPR = 15.2\%$	$FPR = 20.2\%$
	determined on 2438-A/B/D	scan C1	scan C2	scan C3
2438-C	$d = 2$	$r = -0.72$	$r = -0.68$	$r = -0.70$
	$\tau = 9$	$AUC = 0.958$	$AUC = 0.931$	$AUC = 0.937$
	$\epsilon = 0.6$	$acc_{bal} = 88.1\%$	$acc_{bal} = 83.7\%$	$acc_{bal} = 82.4\%$
	$l_{min} = 4$	$TPR = 91.2\%$	$TPR = 90.8\%$	$TPR = 94.6\%$
	$thres = 9.4186$	$FPR = 15.0\%$	$FPR = 23.4\%$	$FPR = 29.7\%$
	determined on 2438-A/B/C	scan D1	scan D2	scan D3
2438-D	$d = 3$	$r = -0.64$	$r = -0.69$	$r = -0.72$
	$\tau = 9$	$AUC = 0.948$	$AUC = 0.964$	$AUC = 0.962$
	$\epsilon = 1.0$	$acc_{bal} = 87.8\%$	$acc_{bal} = 88.1\%$	$acc_{bal} = 87.3\%$
	$l_{min} = 4$	$TPR = 86.7\%$	$TPR = 92.6\%$	$TPR = 94.2\%$
	$thres = 7.2683$	$FPR = 11.0\%$	$FPR = 16.3\%$	$FPR = 19.6\%$

Table C.81: Classification and correlation results, RQA, with *RATIO* as BWE-equivalent, using Euclidean distance, for natural porosity area 2438A–D. RQA based on difference recurrence plots, optimum RQA parameters and decision threshold determined on 3 areas (scan sets), test results with these parameters on scans of the remaining scan set. Reference for cross recurrence is the mean of intermediate echo time series out of 24-α1 of the area $x = 28$–30, $y = 20$–22

results of sample/ scan	parameters	test results		
	determined on 2438-B/C/D	scan A1	scan A2	scan A3
2438-A	$d = 5$	$r = -0.66$	$r = -0.60$	$r = -0.61$
	$\tau = 9$	$AUC = 0.955$	$AUC = 0.899$	$AUC = 0.906$
	$\epsilon = 0.2$	$acc_{bal} = 88.8\%$	$acc_{bal} = 79.0\%$	$acc_{bal} = 80.0\%$
	$l_{min} = 2$	$TPR = 85.1\%$	$TPR = 64.9\%$	$TPR = 66.7\%$
	$thres = 4.2399$	$FPR = 7.6\%$	$FPR = 6.8\%$	$FPR = 6.6\%$
	determined on 2438-A/C/D	scan B1	scan B2	scan B3
2438-B	$d = 5$	$r = -0.53$	$r = -0.57$	$r = -0.56$
	$\tau = 9$	$AUC = 0.857$	$AUC = 0.874$	$AUC = 0.876$
	$\epsilon = 0.2$	$acc_{bal} = 78.9\%$	$acc_{bal} = 79.5\%$	$acc_{bal} = 80.2\%$
	$l_{min} = 2$	$TPR = 74.9\%$	$TPR = 70.0\%$	$TPR = 74.8\%$
	$thres = 4.2035$	$FPR = 17.0\%$	$FPR = 11.1\%$	$FPR = 14.5\%$
	determined on 2438-A/B/D	scan C1	scan C2	scan C3
2438-C	$d = 4$	$r = -0.73$	$r = -0.66$	$r = -0.69$
	$\tau = 9$	$AUC = 0.972$	$AUC = 0.919$	$AUC = 0.940$
	$\epsilon = 0.4$	$acc_{bal} = 89.7\%$	$acc_{bal} = 81.1\%$	$acc_{bal} = 79.7\%$
	$l_{min} = 3$	$TPR = 95.3\%$	$TPR = 91.6\%$	$TPR = 97.1\%$
	$thres = 2.1436$	$FPR = 15.9\%$	$FPR = 29.4\%$	$FPR = 37.7\%$
	determined on 2438-A/B/C	scan D1	scan D2	scan D3
2438-D	$d = 5$	$r = -0.61$	$r = -0.77$	$r = -0.78$
	$\tau = 9$	$AUC = 0.926$	$AUC = 0.983$	$AUC = 0.980$
	$\epsilon = 0.4$	$acc_{bal} = 85.1\%$	$acc_{bal} = 87.0\%$	$acc_{bal} = 85.3\%$
	$l_{min} = 3$	$TPR = 87.6\%$	$TPR = 99.9\%$	$TPR = 99.6\%$
	$thres = 2.2061$	$FPR = 17.4\%$	$FPR = 25.9\%$	$FPR = 29.0\%$

C.4.5 Test results on thinner specimens

Table C.82: Classification and correlation results, RQA, with *RATIO* as BWE-equivalent, using Euclidean distance, for natural porosity area 2438A–D; reduced intermediate echo gate of 1 µs to 2.5 µs. Optimum RQA parameters and decision threshold determined on 3 areas (scan sets), test results with these parameters on scans of the remaining scan set.

results of specimen/ scan	parameters	test results		
	determined on 2438-B/C/D	scan A1	scan A2	scan A3
2438-A	$d = 5$	$r = -0.61$	$r = -0.49$	$r = -0.54$
	$\tau = 8$	$AUC = 0.920$	$AUC = 0.822$	$AUC = 0.857$
	$\epsilon = 4.0$	$acc_{bal} = 86.2\%$	$acc_{bal} = 75.0\%$	$acc_{bal} = 77.5\%$
	$l_{min} = 5$	$TPR = 84.1\%$	$TPR = 58.7\%$	$TPR = 62.6\%$
	$thres = 2.0301$	$FPR = 11.7\%$	$FPR = 8.6\%$	$FPR = 7.7\%$
	determined on 2438-A/C/D	scan B1	scan B2	scan B3
2438-B	$d = 5$	$r = -0.34$	$r = -0.42$	$r = -0.45$
	$\tau = 8$	$AUC = 0.710$	$AUC = 0.813$	$AUC = 0.833$
	$\epsilon = 4.0$	$acc_{bal} = 68.0\%$	$acc_{bal} = 75.6\%$	$acc_{bal} = 77.6\%$
	$l_{min} = 5$	$TPR = 63.0\%$	$TPR = 69.6\%$	$TPR = 72.7\%$
	$thres = 2.0051$	$FPR = 27.0\%$	$FPR = 18.5\%$	$FPR = 17.6\%$
	determined on 2438-A/B/D	scan C1	scan C2	scan C3
2438-C	$d = 5$	$r = -0.62$	$r = -0.52$	$r = -0.52$
	$\tau = 8$	$AUC = 0.942$	$AUC = 0.834$	$AUC = 0.874$
	$\epsilon = 4.0$	$acc_{bal} = 88.0\%$	$acc_{bal} = 74.4\%$	$acc_{bal} = 73.3\%$
	$l_{min} = 5$	$TPR = 88.1\%$	$TPR = 81.7\%$	$TPR = 89.3\%$
	$thres = 1.9473$	$FPR = 12.0\%$	$FPR = 32.9\%$	$FPR = 42.6\%$
	determined on 2438-A/B/C	scan D1	scan D2	scan D3
2438-D	$d = 5$	$r = -0.58$	$r = -0.66$	$r = -0.63$
	$\tau = 8$	$AUC = 0.909$	$AUC = 0.962$	$AUC = 0.963$
	$\epsilon = 4.0$	$acc_{bal} = 83.9\%$	$acc_{bal} = 86.0\%$	$acc_{bal} = 83.8\%$
	$l_{min} = 5$	$TPR = 85.7\%$	$TPR = 97.8\%$	$TPR = 98.1\%$
	$thres = 1.9413$	$FPR = 17.9\%$	$FPR = 25.8\%$	$FPR = 30.5\%$

Table C.83: Classification and correlation results, RQA, with *RATIO* as BWE-equivalent, using Euclidean distance, for natural porosity area 2636A/B; intermediate echo gate from 1 µs to 2.5 µs. Optimum RQA parameters and decision threshold determined on 2438A/C/D.

results of specimen/ scan	parameters	test results		
	determined on 2438-A/C/D	scan A1	scan A2	scan A3
2636-A	$d = 5$	$r = -0.46$	$r = -0.45$	$r = -0.46$
	$\tau = 8$	$AUC = 0.681$	$AUC = 0.719$	$AUC = 0.743$
	$\epsilon = 4.0$	$acc_{bal} = 63.7\%$	$acc_{bal} = 66.4\%$	$acc_{bal} = 63.0\%$
	$l_{min} = 5$	$TPR = 58.4\%$	$TPR = 56.4\%$	$TPR = 48.2\%$
	$thres = 2.0051$	$FPR = 31.0\%$	$FPR = 23.6\%$	$FPR = 22.3\%$
	determined on 2438-A/C/D	scan B1	scan B2	scan B3
2636-B	$d = 5$	$r = -0.41$	$r = -0.57$	$r = -0.40$
	$\tau = 8$	$AUC = 0.715$	$AUC = 0.758$	$AUC = 0.699$
	$\epsilon = 4.0$	$acc_{bal} = 66.7\%$	$acc_{bal} = 68.3\%$	$acc_{bal} = 66.1\%$
	$l_{min} = 5$	$TPR = 55.3\%$	$TPR = 60.8\%$	$TPR = 43.0\%$
	$thres = 2.0051$	$FPR = 21.9\%$	$FPR = 24.1\%$	$FPR = 10.9\%$

Table C.84: Classification and correlation results with linear regression on amplitude spectrum from 0.67 MHz to 30 MHz for natural porosity area 2438A–D; reduced IE gate from 1 µs to 2.5 µs. Hann window applied on intermediate echo time series prior to Fourier transform. Training (including decision threshold to achieve *TPR*≈*TNR*) determined on 3 areas (scan sets), test results on scans of the remaining scan set.

results of specimen/ scan	decision threshold	test results		
	determined on 2438-B/C/D	scan A1	scan A2	scan A3
2438-A	*thres*= 21.8253	$r = 0.50$ *AUC*= 0.791 $acc_{bal} = 73.6\%$ *TPR* = 60.2% *FPR* = 13.0%	$r = 0.47$ *AUC*= 0.752 $acc_{bal} = 69.2\%$ *TPR* = 55.6% *FPR* = 17.3%	$r = 0.40$ *AUC*= 0.726 $acc_{bal} = 67.9\%$ *TPR* = 51.5% *FPR* = 15.8%
	determined on 2438-A/C/D	scan B1	scan B2	scan B3
2438-B	*thres*= 22.0910	$r = 0.49$ *AUC*= 0.791 $acc_{bal} = 71.0\%$ *TPR* = 60.4% *FPR* = 18.4%	$r = 0.60$ *AUC*= 0.847 $acc_{bal} = 77.4\%$ *TPR* = 72.0% *FPR* = 17.3%	$r = 0.58$ *AUC*= 0.841 $acc_{bal} = 77.0\%$ *TPR* = 71.2% *FPR* = 17.1%
	determined on 2438-A/B/D	scan C1	scan C2	scan C3
2438-C	*thres*= 21.6902	$r = 0.56$ *AUC*= 0.829 $acc_{bal} = 72.5\%$ *TPR* = 82.5% *FPR* = 37.6%	$r = 0.56$ *AUC*= 0.807 $acc_{bal} = 71.6\%$ *TPR* = 79.0% *FPR* = 35.9%	$r = 0.67$ *AUC*= 0.896 $acc_{bal} = 77.7\%$ *TPR* = 90.3% *FPR* = 35.0%
	determined on 2438-A/B/C	scan D1	scan D2	scan D3
2438-D	*thres*= 22.2703	$r = 0.55$ *AUC*= 0.817 $acc_{bal} = 72.5\%$ *TPR* = 84.5% *FPR* = 39.5%	$r = 0.76$ *AUC*= 0.952 $acc_{bal} = 81.4\%$ *TPR* = 95.6% *FPR* = 32.9%	$r = 0.76$ *AUC*= 0.947 $acc_{bal} = 81.8\%$ *TPR* = 96.0% *FPR* = 32.4%

Table C.85: Classification and correlation results with linear regression on amplitude spectrum from 0.67 MHz to 30 MHz for natural porosity area 2636A/B; IE gate from 1 µs to 2.5 µs. Hann window applied on intermediate echo time series prior to Fourier transform. Training (including decision threshold to achieve *TPR*≈*TNR*) determined on scan sets 2438A/B/D, IE gate 1 µs–2.5 µs.

results of specimen/ scan	decision threshold	test results		
	determined on 2438-A/B/D	scan A1	scan A2	scan A3
2636-A	*thres*= 21.8253	$r = 0.38$ *AUC*= 0.677 $acc_{bal} = 63.8\%$ *TPR* = 54.3% *FPR* = 26.7%	$r = 0.47$ *AUC*= 0.715 $acc_{bal} = 66.0\%$ *TPR* = 60.7% *FPR* = 28.7%	$r = 0.44$ *AUC*= 0.674 $acc_{bal} = 64.1\%$ *TPR* = 57.1% *FPR* = 29.0%
	determined on 2438-A/B/D	scan B1	scan B2	scan B3
2636-B	*thres*= 21.8253	$r = 0.46$ *AUC*= 0.719 $acc_{bal} = 64.2\%$ *TPR* = 48.1% *FPR* = 19.7%	$r = 0.53$ *AUC*= 0.749 $acc_{bal} = 68.0\%$ *TPR* = 64.9% *FPR* = 28.9%	$r = 0.58$ *AUC*= 0.818 $acc_{bal} = 72.0\%$ *TPR* = 55.0% *FPR* = 10.9%

Appendix D

Code MATLAB® and Python

The author coded the calculation of linear features (cf. sect. 3.1 and 3.2) and of recurrence features (sect. 3.3) for correlation on fabric specimens (sect. 5.3.1) with MATLAB; and computation for recurrence features for classification on all materials (sect. 5.3.1, 5.5.1 and 5.6.1) in Python, including calculation of fixed recurrence rate (RR), cross, joint and difference recurrence plots.

This appendix contains the key MATLAB (R2016B) functions and scripts (.m files) and a Python 3.7 file with several functions.

The time series treated in this thesis are rather short (301 data points) and thus allow for two particular implementations without exceeding typical computer memory of 8 to 16 GB: a global matrix containing all delays can be created, and RQA features can be calculated for a large variation of RQA parameters and stored in one array, containing e.g. 3000 parameter variations for 11310 time series each.

D.1 MATLAB code

RQA calculation for ultrasonic data obtained for correlation on fabric material and computation of linear features for all data was performed in MATLAB.

D.1.1 Creating a versatile matrix with embedded time series

The input of this function takes a cell, in which each entry represents one A-scan. Furthermore, start and end for the intermediate echo gate and the back-wall echo gate and embedding dimension d times time delay τ has to be defined.

This function takes each time series as input and creates versatile time embedded matrices as output; each matrix corresponds to one entry in the output cell. Every column of one matrix presents one state space vector at one instant in time with delay $\tau = 1$ and embedding dimension of e.g. $d = 100$ if d times τ has been set to 100. With this value, any embedding matrix up to $\tau = 10$ and $d = 10$ can be created by taking certain rows out of this matrix.

```
1 function[cell_delayembvecs, array_BWE, array_SE, trigIEgate_start,...
2     trigIEgate_end, trigBWEgate_start, trigBWEgate_end]=...
3     ssrec (volscans, trigIEgate_start,...
4     trigIEgate_end, trigBWEgate_start, trigBWEgate_end, dtimestau)
```

```
5
6  %% function for state space reconstruction:
7  % Creating vectors of state space reconstruction and maximum back wall
8  % echo and maximum surface echo out of volume scans given in file
9  % "volscan". Done in a versatile way using a time delay of 1 with an
10 % embedding dimension of 100.
11
12 % The delay embedding vectors are contained in one matrix for one
13 % measurement point, and these matrices are contained in one cell.
14 % For further creation of delay embedded vectors different to tau=1 and
15 % d=100, e.g. for tau=2 and d=3, first, third and fifth row has to be
16 % taken.
17
18 % input data:
19 % "volscan" to be given as supercell with coordinates of scan x and y
20 % and different volume scans, "IEgate_start" and "IEgate_end" as begin
21 % and end of the intermediate echo gate, "BWEgate_start" and
22 % "BWEgate_end"  as begin and end of back-wall echo gate
23
24 %28.02.2018
25
26 tic
27
28 %determining size of supercell with data "volscan"
29 [kappa,lambda,nScans] = size(volscans);
30
31 %pre-defining arrays for back-wall echoes (BWE), surface echoes (SE)
32 array_BWE = zeros(kappa,lambda,nScans);
33 array_SE = zeros(kappa,lambda,nScans);
34
35 %%creating delay-embedded vectors in one matrix per A-scan;
36 %%3D cells, one point being one matrix
37
38 %pre-defining cells for the matrices with delay-embedded vectors
39 cell_delayembvecs = cell(kappa,lambda,nScans);
40
41 %loop for number of scan
42 for m = 1:nScans
43     for k = 1:kappa
44         for l = 1:lambda
45             %calculating BWE
46             array_BWE(k,l,m) =...
47                max(abs(volscans{k,l,m}...
48            (trigBWEgate_start:trigBWEgate_end)));
49             %calculating surface echo SE as maximum echo of entire A-scan
50             array_SE(k,l,m) =...
51                 max(abs(volscans{k,l,m}));
52
53             %loop for embedding dimension
54             %create matrices with delay-embedded vectors in cell
55             for embdim = 1:dtimestau
56                 cell_delayembvecs{k,l,m}(embdim,1:...
57                 (trigIEgate_end-trigIEgate_start+2-embdim))...
58                     = volscans{k,l,m}...
```

```
                    ((trigIEgate_start-1+embdim):trigIEgate_end);
            end
        end
    end
end

toc

end
```

D.1.2 Calculating RQA features out of versatile embedding matrix

These functions create RQA features out of the versatile embedding cell (sect. D.1.1), for Euclidean and for angular distance, respectively. The calculation of RQA features beyond *RR* and *DET* was not used in the calculations presented, because these are performed with Python (sect. D.2); the code below served as a starting point for calculations in Python. Following calculations include the main diagonal of recurrence plots (LOI).

```
function [ RR, DET, RATIO, Lave, DIV, TND, TND2, ENT, disttoNforTND, ...
    embdimmin, embdimmax, taumin, taumax, taustep, ...
    epsmin, epsmax, epsstep, lminmin, lminmax, lminstep ] =...
    calcRQAfeatures ( cell_delayembvecs, ...
    embdimmin, embdimmax, taumin, taumax, taustep, epsmin, epsmax,...
    epsstep, lminmin, lminmax, lminstep, disttoNforTND, Euc0ang1)
%%Calculation of RQA features for different RQA parameters
%recurrence rates RR, determinism DET, RATIO (DET/RR), Lave average
%line length, DIV divergence (1/maximum diagonal line length), TND trend,
%ENT entropy
%for
% - different embedding dimensions,
% - different time delays tau,
% - different epsilons (recurrence thresholds),
% - different minimum line length to be taken into account for DET, Lave,
%and without saving similarity matrices (distance matrices);
%similarity matrices calculated with Euclidean distance if Euc0ang1 = 0,
%with angular distance if Euc0ang1 = 1

%pre-requisite: data 'cell_delayembvecs' previously calculated

%input data:
%"array_BWE" to give the correlation between BWE equivalent and BWE,
%"cell_delayembvecs" to be given as cell with delay embedding vectors

%03.04.2018

tic

%%pre-define arrays for RQA features

```

```
%determine size of scan
[kappa,lambda,nScans] = size(cell_delayembvecs);
%number of steps of epsilon
amountStepseps = int8((epsmax-epsmin+epsstep)/epsstep);
%pre-define array for recurrence rate RR
RR = zeros(kappa,lambda,nScans,embdimmax,taumax,amountStepseps,'single');
%pre-define array for determinism DET in same way as RR
DET = zeros(kappa,lambda,nScans,embdimmax,taumax,amountStepseps,...
    lminmax,'single');
%pre-define array for RATIO - determinism DET divided by RR in same way
%as RR and DET
RATIO = zeros(kappa,lambda,nScans,embdimmax,taumax,amountStepseps,...
    lminmax,'single');
%pre-define average diagonal line length
Lave = zeros(kappa,lambda,nScans,embdimmax,taumax,amountStepseps,...
    lminmax,'single');
%pre-define divergence, the inverse of the maximum diagonal line length
DIV = zeros(kappa,lambda,nScans,embdimmax,taumax,amountStepseps,...
    'single');
%pre-define trend
TND = zeros(kappa,lambda,nScans,embdimmax,taumax,amountStepseps,...
    'single');
TND2 = zeros(kappa,lambda,nScans,embdimmax,taumax,amountStepseps,...
    'single');
%pre-define entropy
ENT = zeros(kappa,lambda,nScans,embdimmax,taumax,amountStepseps,...
    lminmax,'single');

%%calculation of recurrence rate and determinism DET

%loop for several volume scans
for currnrscan = 1:nScans
    %loop for embedding dimension
    for embdim=embdimmin:embdimmax
        %loop for time delay tau
        for tau = taumin:taustep:taumax
            tau
            %number of embedding vectors = length of sequence (time
            %series) of embedded vectors, reduced accordingly to the
            %number of embedding dimensions and tau
            nDelayembvecs=length(cell_delayembvecs{1,1,currnrscan})-...
                (embdim-1)*tau;
            for k=1:kappa %x-direction of volume scan
                for l=1:lambda %y-direction of volume scan
                    %temporary similarity (distance) matrix (discarded
                    %after calculation of RQA features
                    %for all epsilon)
                    if Euc0ang1 == 0 %Euclidean distance
                        tempdistmat=squareform(pdist((...
                        cell_delayembvecs{k,l,currnrscan}(1:tau:...
                        (1+(embdim-1)*tau),1:nDelayembvecs))'));
                    else %angular distance
                        tempdistmat=acos(1-squareform(pdist(...
                        (cell_delayembvecs{k,l,currnrscan}...
```

```
86                          (1:tau:(1+(embdim-1)*tau),1:nDelayembvecs)...
87                          )','cosine')));
88                      end
89                      %loop for different epsilon to calculate recurrence
90                      %matrix and RQA features
91                      for nStepseps=1:amountStepseps
92                          epsilon=epsmin+(nStepseps-1)*epsstep;
93                          %temporary recurrence matrix (discarded after
94                          %calculation of recurrence rate and determinism)
95                          temprecmat=tempdistmat<epsilon;
96                          %Recurrence rate
97                          RR(k,l,currnrscan,embdim,tau,nStepseps) = ...
98                              mean(temprecmat(:));
99                          %trend
100                         %number of diagonals used
101                         amountdiagTND = length(temprecmat)-disttoNforTND;
102                         %loop to go through all diagonals
103                         %temporary variable with RRs per diagonal
104                         RRdiag = zeros(amountdiagTND,1);
105                         %vector with numbers of diagonals
106                         vecnrdiag = [1:amountdiagTND];
107                         for nDiag=1:amountdiagTND
108                             %temporary variable with diagonal
109                             tempdiag=diag(temprecmat,nDiag);
110                             RRdiag(nDiag) = mean(tempdiag);
111                         end
112                         TND(k,l,currnrscan,embdim,tau,nStepseps) = ...
113                             (vecnrdiag-amountdiagTND/2)*...
114                             (RRdiag-mean(RRdiag))/...
115                             sum((vecnrdiag-amountdiagTND/2).^2);
116                         %as a correlation coefficient
117                         cc = corrcoef(vecnrdiag,RRdiag);
118                         TND2(k,l,currnrscan,embdim,tau,nStepseps) = ...
119                         cc(1,2); %later named `COR'
120                         %calculate vector with amount of lines (1st entry
121                         %amount of separated points not in any diagonal
122                         %line, 2nd entry all points which are in diagonal
123                         %lines of exact length of 2 points etc.)
124                         %main diagonal (line of identity) left out
125                         [ tempamountlines ] = ...
126                             calctempamountlines( temprecmat );
127
128                         %now determining features based on diagonal lines
129                         %for different minimum line length out of
130                         %tempamountlines
131                         for lmin=lminmin:lminstep:lminmax
132                             %determinism
133                             DET(k,l,currnrscan,embdim,tau,nStepseps,...
134                             lmin) = (2*(tempamountlines(lmin:end)*...
135                                 (lmin : nDelayembvecs )')...
136                                 + nDelayembvecs)/(nDelayembvecs^2)/...
137                                 RR(k,l,currnrscan,embdim,tau,nStepseps);
138                             %ratio (DET/RR)
139                             RATIO...
```

```
                                (k,l,currnrscan,embdim,tau,nStepseps,lmin)...
                                    =DET(k,l,currnrscan,embdim,tau,...
                                    nStepseps,lmin)/...
                                    RR(k,l,currnrscan,embdim,tau,nStepseps);
                                %average line length
                                Lave...
                        (k,l,currnrscan,embdim,tau,nStepseps,lmin)...
                                    = tempamountlines(lmin:end)*...
                                    (lmin : nDelayembvecs )'/...
                                    sum(tempamountlines(lmin:end));
                                %entropy
                                %probability of occurrence of lines
                                %to determine entropy
                                p = tempamountlines(lmin:end)/...
                                    sum(tempamountlines(lmin:end));
                                ENT(k,l,currnrscan,embdim,tau,nStepseps,...
                                lmin) = -log(p(p>0))*p(p>0)';
                            end
                            %determining divergence (inverse of maximum line
                            %length - independent of minimum line length
                            %if loop for the case that there is no recurrence
                            %point except the line of identity
                            if sum(tempamountlines) > 0
                                DIV(k,l,currnrscan,embdim,tau,nStepseps) =...
                                    1/find(tempamountlines,1,'last');
                            else
                                DIV(k,l,currnrscan,embdim,tau,nStepseps) = 0;
                            end
                            clear temprecplot
                        end
                    clear tempdistmat
                    end
                end
            end
        end

end

%convert NaNs in Lave (created if no diagonal line with
%length equal to or greater than minimum line length in
%calculation) to 0
Lave(isnan(Lave)) = 0;
%multiplying TND by 1000
TND = 1000*TND;

toc

end
```

In above function a C code is called. This has been generated with the MATLAB C code generation out of following function.

```
function [ tempamountlines ] = calctempamountlines( temprecmat )
```

```
%core calculation of vector tempamountlines, out of which C-code
%is generated

%pre-define temporary variable that contains the number of lines of
%different length, including the separated 1s, i.e. diagonal "lines" of
%length 1
tempamountlines=zeros(1,length(temprecmat));

%loop to go through all diagonals
for nDiag=1:(length(temprecmat)-2)
    %temporary variable with diagonal
    tempdiag=diag(temprecmat,nDiag);
    %temporary variable to sum up neighboured 1s
    tempsum1s=0;

    %loop for w as variable to go through whole diagonal "tempdiag"
    %and count all 1s in line of minimum length line
    for w=1:length(tempdiag)
        if tempdiag(w)==1
            tempsum1s=tempsum1s+1;
        elseif tempsum1s > 0
            tempamountlines(tempsum1s)=tempamountlines(tempsum1s)+1;
            tempsum1s=0;
        end
    end
    if tempsum1s>0
        tempamountlines(tempsum1s)=tempamountlines(tempsum1s)+1;
    end
end

end
```

D.1.3 Calculating features in time and in frequency domain

Following features have been calculated (cf. sect. 3.1 and 3.2)

- Maximum absolute intermediate echo,
- Variance of the intermediate echo time series,
- Quartile coefficient of dispersion,
- Linear regression of time series values,
- Fourier transform
  - Sum of (subset of) frequency values,
  - Largest frequency
  - Bandwidth around largest frequency
  - Linear regression on (portion of) frequency values.

Most of these features are standard functions built into MATLAB and not presented. There are two exceptions:

Quartile coefficient of dispersion

```
function [qcd_array] = QCD (iets)
%%Calculation of quartile coefficient of dispersion QCD with an offset
%out of intermediate echo time series iets

%09.02.2017

%alternative - minimum of each time series
min_iets = min(iets);
iets_sort = sort(iets);
qcd_array = ...
    ((iets_sort(ceil(3*end/4))-min_iets)-(iets_sort(ceil(end/4))...
        -min_iets))/...
    ((iets_sort(ceil(3*end/4))-min_iets)+(iets_sort(ceil(end/4))...
        -min_iets));

end
```

Bandwidth around largest frequency (including peak frequency, left and right cut-off frequency and integral according to bandwidth)

```
function [ maxf, meanf,leftf,rightf,bw,integ ] = ...
    bandwdth( absfftiets, dBlim)
%calculation peak frequency (frequency of maximum amplitude),
%of the bandwidth, according lower and upper cut-off frequency,
%mean frequency and integral in area of bandwidth
%dblim has to be given (in general this will be -3 dB and -6 dB
%(-3 dB used for evaluation in this thesis)

%factor for calculating frequency in MHz from FFT values, 100 MHz
%sampling frequency
factf = 100/2/floor(length(absfftiets)/2);
%max frequency in first half of FFT (symmetric because time series is
%real)
%first setting first value of fft to zero to avoid DC part being counted
%as max frequency
absfftiets(1) = 0;
[abscomplexfmax,posfmax]=max(absfftiets(1:ceil(length(absfftiets)/2)));
maxf = posfmax*factf;%maxf in MHz;
%dB as linear ratio
linlim = 10^(-dBlim/20);

%searching for position of left cut-off of bandwidth
o = posfmax;
while absfftiets(o)>abscomplexfmax*linlim
    o = o-1;
end

```

```
28 %searching for position of right cut-off of bandwidth
29 u = posfmax;
30 while absfftiets(u)>abscomplexfmax*linlim
31     u = u+1;
32 end
33
34 %covering situation when no drop to dBlim is present; in that case set
35 %bw, cut-off and integ to zero
36 if o > 1 && u < ceil(length(absfftiets)/2)
37     %interpolation
38     %distance frequency value above and below limit left
39     distabl = absfftiets(o+1)-absfftiets(o);
40     %distance frequency value below and abscomplexfmax*linlim
41     distl = abscomplexfmax*linlim-absfftiets(o);
42     leftf = (o+distl/distabl)*factf;
43
44     %distance frequency value above and below limit right
45     distabr = absfftiets(u-1)-absfftiets(u);
46     %distance frequency value below and abscomplexfmax*linlim
47     distr = abscomplexfmax*linlim-absfftiets(u);
48     rightf = (u-distr/distabr)*factf;
49
50     %mean frequency as mean of left and right cut-off
51     meanf = (leftf+rightf)/2;
52
53     %bandwidth
54     bw = rightf-leftf;
55
56     %integral using Matlab function 'trapz', result in MHz^-1
57     %prepare frequency vector in MHz
58     ft = [o+distl/distabl,o+1:u-1,u-distr/distabr]*factf;
59     %prepare vector with absolute values of frequencies
60     fa = [abscomplexfmax*linlim,absfftiets(o+1:u-1), ...
61     abscomplexfmax*linlim];
62     integ = trapz(ft,fa);
63 else
64     meanf = 0;
65     bw = 0;
66     integ = 0;
67     leftf = 0;
68     rightf = 0;
69 end
70
71 end
```

D.2 Python code for RQA feature calculation

RQA calculation for unidirectional material was performed with Python.

```
1 # -*- coding: utf-8 -*-
2 """
```

```
Calculation of RQA features RR, DET, RATIO, Lave, DIV, TND, ENT and
normalised Lnor and TNRnor.

Includes cross, joint and difference recurrence plots as well as in
addition to fixed epsilon fixed RR and locally fixed recurrence rate

Computation without and with LOI and Theiler window

last update 31.10.2019

@author: Carsten Brandt

Python 3.7.2

"""

import numpy as np #numpy 1.16.2
from scipy import spatial #scipy 1.2.1
import timeit
from numba import jit #numba 0.43.0

# ignoring division by zero
np.seterr(divide='ignore', invalid='ignore')

# state space reconstruction via delay embedding
def ssrecfunc(volscan, trigIEgate_start,
          trigIEgate_end, trigBWEgate_start, trigBWEgate_end, dtimestau):
    """
    State space reconstruction out of given time series in array volscan.
    Because of relatively short time series, it is possible to create an
    array with all possible delayed vectors, out of which the correct
    vectors are taken for calculation of recurrence matrices for RR etc.
    """

    # pre-define arrays for back-wall echo, surface echo, delay embedding
    array_BWE = np.zeros\
    ((volscan.shape[0], volscan.shape[1], volscan.shape[2]))
    array_SE = np.zeros\
    ((volscan.shape[0], volscan.shape[1], volscan.shape[2]))

    array_delayembvecs = np.zeros\
    ((volscan.shape[0], volscan.shape[1], volscan.shape[2],\
    dtimestau, trigIEgate_end-trigIEgate_start+1))
    # Attention for different indexing, starting with 0, in Python, and
    # for the different kind to access parts of arrays:
    # indexing into time series from 100:400 in Python will give
    # 1.00 mics to 3.99 mics (instead of 0.99 mics to 3.99 mics) -- all
    # values included. In general I use (for 6 mm specimens) start 100
    # and end 401 (thus 400 included) to include 4mics into calculation.

    # loops for number of scans, x and y
    for currnrscan in range(volscan.shape[0]):
        for x in range(volscan.shape[1]):
            for y in range(volscan.shape[2]):
```

```
57                  # back-wall echo
58                  array_BWE[currnrscan, x, y] = \
59                  np.abs(volscan[currnrscan, x, y, trigBWEgate_start:\
60                                 trigBWEgate_end+1]).max()  # +1 to account
61                                  # for the fact that the value itself is
62                                  # not inside in Python
63                  # computing surface echo SE as max echo of entire A-scan
64                  array_SE[currnrscan, x, y] = \
65                  np.abs(volscan[currnrscan, x, y, :]).max()
66                  # loop for embedding
67                  for embdim in range(dtimestau):
68                      array_delayembvecs[currnrscan, x, y, embdim, 0:\
69                                         trigIEgate_end-trigIEgate_start+1\
70                                        -embdim] = \
71                      volscan\
72                      [currnrscan, x, y, \
73                       (trigIEgate_start+embdim):trigIEgate_end+1]
74
75      return (array_delayembvecs, array_BWE, array_SE, trigIEgate_start,\
76      trigIEgate_end, trigBWEgate_start, trigBWEgate_end)
77
78
79  # following functions as subfunctions of main RQA calculation function
80
81  # calculation of distance matrix (unthresholded recurrence matrix)
82  def calcdistmat(delembvecs, metric, RPtype, nDelayembvecs, \
83                  currnrscan, x, y, embdim, tau, delembvecsreffinal, \
84                  LOI, Theiler, recurrence):
85      # 2nd distance matrix pre-defined to be able to return something
86      # even if not needed
87      tempdistmat2 = np.zeros((nDelayembvecs, nDelayembvecs))
88      #change names for metric in calculation of pdist/cdist (dist. matrix)
89      if metric == 'Euclidean':
90          metricdist = 'euclidean'
91      elif metric == 'angular':
92          metricdist = 'cosine'
93      else:
94          print("Metric has to be given as 'Euclidean' or 'angular'.")
95
96      if RPtype == 'RP' or RPtype == 'JRP' or RPtype == 'JRP2':
97          tempdistmattemp = spatial.distance.squareform(\
98          spatial.distance.pdist\
99          (delembvecs[currnrscan, x, y, \
100                      0:1+embdim*(tau+1):(tau+1), \
101                      0:nDelayembvecs].transpose(), metric=metricdist))
102         #for angular distance arcus cosine necessary
103         if metric == 'angular':
104             if embdim: # calculation only makes sense for
105                 # embedding dimension > 1
106                 tempdistmat = np.arccos\
107                 (1 - tempdistmattemp)
108             else:
109                 tempdistmat = np.zeros((nDelayembvecs, nDelayembvecs), \
110                                        dtype=bool)
```

```
111         else: # Euclidean distance
112             tempdistmat = tempdistmattemp
113     elif RPtype == 'CRP':
114         tempdistmattemp = spatial.distance.cdist(delembvecs \
115           [curnrscan, x, y, 0:1+embdim*(tau+1):(tau+1), \
116            0:nDelayembvecs].transpose(), \
117            delembvecsreffinal.transpose(), metric=metricdist)
118         if metric == 'angular':
119             if embdim: # embedding dimension > 1
120                 tempdistmat = np.arccos\
121                 (1 - tempdistmattemp)
122             else:
123                 tempdistmat = np.zeros((nDelayembvecs, nDelayembvecs), \
124                                        dtype=bool)
125         else: # Euclidean distance
126             tempdistmat = tempdistmattemp
127     else:
128         print("'RPtype' has to be given as RP, CRP, JRP or JRP2. \
129               Please correct.")
130     # additional calculation for JRP and JRP2
131     if RPtype == 'JRP' or RPtype == 'JRP2':
132         tempdistmat2 = spatial.distance.squareform(\
133         spatial.distance.pdist\
134         (delembvecsreffinal.transpose(), \
135             metric=metricdist))
136         #for angular distance arcus cosinus necessary
137         if metric == 'angular':
138             if embdim: # embedding dimension > 1
139                 tempdistmat2 = np.arccos\
140                 (1 - tempdistmat2)
141             else:
142                 tempdistmat2 = np.zeros((nDelayembvecs, nDelayembvecs), \
143                                        dtype=bool)
144     if RPtype == 'JRP2': # here, the absolute value of differences of
145         # distances are built
146         tempdistmat = np.absolute(tempdistmat - tempdistmat2)
147     # excluding LOI and Theiler window , if applicable, to exclude them
148     # from calculation of epsilon
149     if recurrence != 'fixedepsilon':
150         if LOI: # == 1
151             np.fill_diagonal(tempdistmat, np.nan)
152         if Theiler: # > 0
153             for inde in range(Theiler):
154                 # setting diagonal lines below main diagonal
155                 # (origin upper left) to NaN
156                 tempdistmat[range(inde+1, nDelayembvecs),
157                             range(nDelayembvecs-inde-1)] = np.nan
158                 # setting diagonal lines above main diagonal
159                 # (origin upper left) to NaN
160                 tempdistmat[range(nDelayembvecs-inde-1),
161                             range(inde+1, nDelayembvecs)] = np.nan
162     return tempdistmat, tempdistmat2
163
164 # calculation of recurrence matrix
```

```
165 def calcRP(tempdistmat, tempdistmat2, recurrence, \
166             epsilonorperc, nDelayembvecs, RPtype):
167     # determining epsilon
168     if recurrence == 'fixedepsilon':
169         epsilon = epsilonorperc
170         epsilon2 = epsilonorperc
171     elif recurrence == 'fixedrecglobal':
172         # determination of epsilon out of distance distribution
173         epsilon = np.nanpercentile(tempdistmat, epsilonorperc)
174         epsilon2 = np.nanpercentile(tempdistmat2, epsilonorperc)
175     elif recurrence == 'fixedreclocal': # epsilon as vector to provide
176         # values for every column
177         epsilon = np.nanpercentile(tempdistmat, epsilonorperc, axis=0)
178         epsilon2 = np.nanpercentile(tempdistmat2, epsilonorperc, axis=0)
179     elif recurrence == 'fixedneighbours': # vector of epsilons computed
180         # with percentile generated from number of fix neighbours
181         epsilon = np.nanpercentile\
182         (tempdistmat, epsilonorperc*100/nDelayembvecs, axis=0)
183         epsilon2 = np.nanpercentile\
184         (tempdistmat2, epsilonorperc*100/nDelayembvecs, axis=0)
185     # RP, CRP, JRP or JRP2
186     temprecmat = tempdistmat < epsilon # with this boolean operation, the
187      # NaN values of LOI change to False, i.e. if LOI shall not be taken
188      # into account, LOI contains 0s---this is of no effect, since
189      # further processing of RP is only in function calc_histdiaglines
190      # below for creation of histogram of diagonal lines, where main
191      # diagonal of RP is excluded if LOI = 'without'
192     if RPtype == 'JRP':
193         temprecmat2 = tempdistmat2 < epsilon2
194         temprecmat = temprecmat*temprecmat2 # element-wise calculation
195         # for the two matrices; JRP contains recurrence points only if
196         # both recurrence plots contain recurrence points
197     # now determine mean of all local epsilon if recurrence =
198     # 'fixedreclocal' or 'fixedneighbours' to give information about this
199     if recurrence == 'fixedreclocal' or \
200     recurrence == 'fixedneighbours':
201         epsilon = np.mean(epsilon)
202     return(temprecmat, epsilon)
203
204 # function to create the histogram of diagonal lines, which is used for
205     # all features
206 @jit #just-in-time compilation, increases speed for this function
207 def calc_histdiaglines(temprecmat, nDelayembvecs, RRdiag, \
208                                hist_diaglines, LOI, Theiler):
209     # number of diagonals running from 0, i.e. main diagonal (line of
210     # identity LOI), for LOI='with';
211     # or from 1 (or -1), i.e. next to main diagonal, or, when using a
212     # Theiler window, Theiler larger, (Theiler is taken into account
213     # here, because it is also relevant for hist_diaglines) and ending
214     # at last diagonal hist_diaglines is here only calculated for one
215     # triangle of the matrix;
216     # for non-symmetric RPs, it is called in main function twice for both
217     # triangles; for symmetric RPs, it is doubled within main function
218     """
```

```
219     # LOI = 'with' = 0
220     # LOI = 'without' = 1
221     if LOI == 0 and nDelayembvecs>0:
222         plusTheilerstart = 0
223     elif LOI == 0 and nDelayembvecs<0:
224         # going through the lower diagonal matrix, starting at - 1 (plus
225         # Theiler) not to count the main diagonal twice
226         plusTheilerstart = -1
227     elif LOI == 1:
228         plusTheilerstart = np.sign(nDelayembvecs) # if nDelayembvecs is
229         # given as neg. value for this function, the lower diagonal part
230         # of the dist. matrix shall be calculated (for non-symmetric RPs)
231     """
232     plusTheilerstart=np.sign(nDelayembvecs)*(np.sign(nDelayembvecs)<=LOI)
233     # This equation covers the above if-loops in a faster way.
234 
235     for nDiag in range(Theiler+plusTheilerstart, nDelayembvecs, \
236                        np.sign(nDelayembvecs)): #
237         # negative values for nDiag if looping through lower diagonal
238         # part of recurrence matrix, exclude Theiler window around line
239         # of identity, default is 0 (nothing excluded)
240         # temporary variable with diagonal
241         tempdiag = np.diag(temprecmat, nDiag)
242         # however, fill 'RRdiag' for positive values (and 0 if LOI=
243         # 'with', taking LOI, i.e. main diagonal, into account),
244         # because no distinguishing upper/lower part of recurrence matrix
245         RRdiag[np.sign(nDelayembvecs)*nDiag] = np.mean(tempdiag)
246 
247         # temporary variable to sum up neighboured 1s
248         tempsum1s = 0
249         # loop for w as variable to go through whole diagonal "tempdiag"
250         # and count all 1s in line of minimum length line
251         for w in np.ndenumerate(tempdiag):
252             if w[1]: # 1 instead of 0 (true instead false)
253                 tempsum1s += 1
254             elif tempsum1s: # greater 0
255                 hist_diaglines[tempsum1s-1] += 1 # fill histogram of
256                 # diagonal lines with one more on the position of the
257                 # length of the diagonal line (taking 0 indexing into
258                 # account, e.g. lines of length 3 at index 2)
259                 tempsum1s = 0
260         if tempsum1s: # greater 0
261             hist_diaglines[tempsum1s-1] += 1
262     return(hist_diaglines, RRdiag)
263 
264 # 3 features are calculated externally, because in trials found out that
265 # it is faster for these features (whereas for the other features it is
266 # not faster to calculate in extra functions)
267 # external calculation of RR with histogram of diagonal lines, covers
268 # cases with/without LOI and with/without Theiler window
269 @jit
270 def calcRR(LOI, Theiler, hist_diaglines, nDelayembvecs):
271     RR = (np.dot(hist_diaglines, np.arange(1, nDelayembvecs+1)))/ \
272     (nDelayembvecs**2 - LOI*nDelayembvecs - \
```

```
273       2*(np.arange(nDelayembvecs-1, nDelayembvecs-LOI-Theiler, -1)).sum())
274      # the latter 2 terms subtract, if applicable, line of identity and
275      # twice the Theiler window, respectively; the numerator has this
276      # taken into account, since hist_diaglines has it incorporated
277      return RR
278
279  # trend TND
280  @jit
281  def calcTND\
282  (vecnrdiag, maxindexdiagTNDCOR, meandist, LOI, Theiler, RRdiag):
283      TND = np.dot((vecnrdiag-meandist), \
284          (RRdiag[LOI+Theiler:maxindexdiagTNDCOR+1] - \
285           np.mean(RRdiag[LOI+Theiler:maxindexdiagTNDCOR+1])))/ \
286           np.square(vecnrdiag-meandist).sum()
287      return TND
288
289  # trend as correlation coefficient COR
290  def calcCOR\
291  (vecnrdiag, maxindexdiagTNDCOR, meandist, LOI, Theiler, RRdiag):
292      COR = np.dot((vecnrdiag-meandist), \
293            (RRdiag[LOI+Theiler:maxindexdiagTNDCOR+1]-\
294             RRdiag[LOI+Theiler:maxindexdiagTNDCOR+1].mean()))/ \
295         np.sqrt(np.sum(\
296              (vecnrdiag-meandist)**2)* \
297              np.sum((RRdiag[LOI+Theiler:maxindexdiagTNDCOR+1]- \
298              RRdiag[LOI+Theiler:maxindexdiagTNDCOR+1].mean())**2))
299      return COR
300
301  # actual calculation of RQA features
302  def calcRQAfeatures(delembvecs, \
303      embdimmin, embdimmax, taumin, taumax, taustep, epsmin, epsmax, \
304      epsstep, lminmin, lminmax, lminstep, \
305      trigIEgate_start, trigIEgate_end, trigBWEgate_start,trigBWEgate_end,\
306      metric, recurrence, LOI='without', Theiler=0, RPtype='RP', \
307      disttoNforTNDCOR=2, \
308      refscanCJRP='int', refscanint='samescan', \
309      refscanxstart=1000, refscanxend=1000, \
310      refscanystart=1000, refscanyend=1000, \
311      refscanstart=1000, refscanend=1000, delembvecsref=0):
312      """
313      Calculation of recurrence rate, determinism and several other
314      'standard' RQA features
315      with options for 'metric' (Euclidean or angular); 'recurrence': fixed
316      epsilon or fixed recurrence rate (globally or locally); calculating
317      with or without line of identity 'LOI'; excluding a Theiler window if
318      'Theiler' > 0;
319      and choosing 'RPtype': RP (recurrence plot), CRP (cross recurrence
320      plot), JRP (joint recurrence plot) or JRP2 (creating a distance
321      matrix as difference of 2 distance matrices and calculating RP out of
322      that) -- now DRP (difference recurrence plot)
323      """
324      start_timeall = timeit.default_timer() # start calculate running time
325
326      # vector with steps of epsilon (or percentage recurrence rate if
```

```
    # recurrence='fixedrecglobal' or 'fixedreclocal', so that epsilon is
    # adapted to achieve fixed recurrence rates; or number of neighbours
    # if recurrence='fixedneighbours')
    epsvec = np.arange(epsmin, epsmax+0.5*epsstep, epsstep)
    # vector with steps for lmin
    lminvec = np.arange(lminmin, lminmax+lminstep, lminstep)
    # pre-define numpy arrays for features
    if recurrence == 'fixedepsilon': # for fixed epsilon
        # recurrence rate
        RR = np.zeros((delembvecs.shape[0], delembvecs.shape[1], \
                       delembvecs.shape[2], embdimmax, taumax, \
                       len(epsvec)), dtype=np.float32)
        # trend
        TND = np.zeros((delembvecs.shape[0], delembvecs.shape[1], \
                       delembvecs.shape[2], embdimmax, taumax, \
                       len(epsvec)), dtype=np.float32)
        # trend as real correlation coefficient, i.e. normalised, lying
        # between -1 and 1---now correlation COR
        COR = np.zeros((delembvecs.shape[0], delembvecs.shape[1], \
                       delembvecs.shape[2], embdimmax, taumax, \
                       len(epsvec)), dtype=np.float32)
        # divergence, the inverse of the maximum diagonal line length
        DIV = np.zeros((delembvecs.shape[0], delembvecs.shape[1], \
                       delembvecs.shape[2], embdimmax, taumax, \
                       len(epsvec)), dtype=np.float32)
    elif recurrence == 'fixedrecglobal' or recurrence == 'fixedreclocal'\
    or recurrence == 'fixedneighbours':
        # for fixed recurrence, one dimension is added to RR, TND, COR
        # and DIV to include information about the used global epsilon or
        # mean of local epsilons if epsilon is adapted
        # recurrence rate
        RR = np.zeros((delembvecs.shape[0], delembvecs.shape[1], \
                       delembvecs.shape[2], embdimmax, taumax, \
                       len(epsvec), 2), dtype=np.float32)
        # trend
        TND = np.zeros((delembvecs.shape[0], delembvecs.shape[1], \
                       delembvecs.shape[2], embdimmax, taumax, \
                       len(epsvec), 2), dtype=np.float32)
        # trend as real correlation coefficient, i.e. normalised, lying
        # between -1 and 1
        COR = np.zeros((delembvecs.shape[0], delembvecs.shape[1], \
                       delembvecs.shape[2], embdimmax, taumax, \
                       len(epsvec), 2), dtype=np.float32)
        # divergence, the inverse of the maximum diagonal line length
        DIV = np.zeros((delembvecs.shape[0], delembvecs.shape[1], \
                       delembvecs.shape[2], embdimmax, taumax, \
                       len(epsvec), 2), dtype=np.float32)
    else:
        print("The argument 'recurrence' has to be given as \
              'fixedepsilon', 'fixedrecglobal' (recurrence rate fixed \
              for whole RP, as percentage of distances) or \
              'fixedreclocal'/'fixedneighbours' (for each point in time,\
              i.e. for each column in RP same recurrence rate in \
              percentage of distances (fixedreclocal, independent of \
```

```
381                 size of RP, i.e. d and tau) or as fixed number of \
382                 neighbours (fixedneighbours).")
383         # convert LOI to number, especially for function calc_histdiaglines
384         if LOI == 'without':
385             LOI = 1
386         elif LOI == 'with':
387             LOI = 0
388         else:
389             print("'LOI' (line of identity) has to be given as 'with' or \
390             'without'. Please correct.")
391 
392         # for other features with dependence of lmin, the epsilon is given at
393         # the place where lmin = 0 would be, thus no need for different size
394         # of arrays determinism
395         DET = np.zeros((delembvecs.shape[0], delembvecs.shape[1], \
396                         delembvecs.shape[2], embdimmax, taumax, len(epsvec), \
397                         lminmax), dtype=np.float32)
398         # RATIO - determinism DET divided by RR
399         RATIO = np.zeros((delembvecs.shape[0], delembvecs.shape[1], \
400                         delembvecs.shape[2], embdimmax, taumax, len(epsvec), \
401                         lminmax), dtype=np.float32)
402         # average diagonal line length
403         Lave = np.zeros((delembvecs.shape[0], delembvecs.shape[1], \
404                         delembvecs.shape[2], embdimmax, taumax, len(epsvec), \
405                         lminmax), dtype=np.float32)
406         # average diagonal line length, normalised with minimum line
407         # length by dividing this minimum diagonal line by Lave
408         Lnor = np.zeros((delembvecs.shape[0], delembvecs.shape[1], \
409                         delembvecs.shape[2], embdimmax, taumax, len(epsvec), \
410                         lminmax), dtype=np.float32)
411         # entropy
412         ENT = np.zeros((delembvecs.shape[0], delembvecs.shape[1], \
413                         delembvecs.shape[2], embdimmax, taumax, len(epsvec), \
414                         lminmax), dtype=np.float32)
415 
416         # start calculation
417         # loops for scan, dimension, embedding dimension etc.
418         for currnrscan in range(delembvecs.shape[0]):
419             print('Scan being calculated is: ')
420             print(currnrscan+1)
421             print('Embedding dimension being calculated is: ')
422             for embdim in range(embdimmin-1, embdimmax):
423                 print(embdim+1)
424                 for tau in np.arange(taumin-1, taumax, taustep):
425                     # number of embedding vectors = length of sequence (time
426                     # series) of embedded vectors, reduced accordingly to the
427                     # number of embedding dimensions and tau, (d-1)*tau
428                     nDelayembvecs = delembvecs.shape[4]-embdim*(tau+1)
429                     ## variables for calculation within epsilon loop
430                     # maximum diagonal used for TND and COR, i.e. the
431                     # diagonal with maxindexdiagTNDCORth distance to LOI is
432                     # the last taken into account for calculation of TND and
433                     # COR (this coincides here with 0 indexing, since 0
434                     # corresponds to main diagonal)
```

```
435                 maxindexdiagTNDCOR = nDelayembvecs - disttoNforTNDCOR - 1
436                 # mean of distances to LOI, not taking LOI and Theiler
437                 # window into account if applicable; if LOI and Theiler
438                 # window and thus whole RP is calculated, distances start
439                 # at 0 (main diagonal LOI)
440                 meandist = (maxindexdiagTNDCOR+LOI+Theiler)/2
441                 # vector with numbers of diagonals, i.e. distance
442                 # to the main diagonal, starting with 0, or higher if LOI
443                 # and a Theiler window shall be taken into account (0 is
444                 # included to include LOI if LOI='with') and ending at
445                 # maxindexdiagTNDCOR (keeping in mind that the end value
446                 # of slicing in Python is one before the value given)
447                 vecnrdiag = \
448                 np.arange(LOI+Theiler, maxindexdiagTNDCOR+1)
449                 # temporary variable with RRs per diagonal
450                 # RRdiag has full size and will be sliced when used for
451                 # TND. It also contains 0 at index 0, standing for the
452                 # main diagonal, which is used if LOI is set to 'with',
453                 # which may be done for non-symmetric RPs.
454                 # pre-define RR per diagonal used for TND and COR
455                 RRdiag = np.zeros(nDelayembvecs)
456
457                 # reference delay vectors for CRP, JRP or JRP2
458                 # out of array delembvecsref with matrix containing all
459                 # delayed embedding vectors as being done in 'ssrecfun'
460                 if RPtype != 'RP':
461                     if refscanCJRP == 'int': # out of the same volume
462                         # scan use the same computed array with delay
463                         # embedded vectors as reference as the one taken
464                         # for main calculation
465                         delembvecsref = delembvecs
466                         # if refscanCJRP = 'ext', use delembvecsref that
467                         # is given, so do nothing
468                     if refscanCJRP == 'int' or refscanCJRP == 'ext':
469                         # Same calculation for both cases
470                         # The area is given with arguments
471                         # 'refscanxstart' etc.; first checking whether
472                         # this have been given correctly (default values
473                         # are outside the volume scan, so they have to
474                         # be given explicitly)
475                         if refscanxstart > refscanxend or \
476                         refscanystart > refscanyend or \
477                         refscanxend > delembvecsref.shape[1] or \
478                         refscanyend > delembvecsref.shape[2]:
479                             print('Start and end for the reference scan \
480                                   for CRP or JRP are outside the bounds \
481                                   of the volume scan, start is higher \
482                                   than end or they have not been defined\
483                                    at all.\n Please correct.')
484                         else:
485                             if refscanint == 'samescan': # out of the
486                                 # same currnrscan
487                                 delembvecsreffinal = \
488                                 np.mean( delembvecsref[currnrscan, \
```

```
                                    refscanxstart-1:refscanxend, \
                                    refscanystart-1:refscanyend, \
                                  0:1+embdim*(tau+1):(tau+1), \
                                  0:nDelayembvecs], axis = (0, 1)) # mean
                                    # over several scans if start and end
                                    # differs; first axis fixed with
                                    # curnrscan, so mean over first and
                                    # second varying axis (x and y)
                            elif refscanint == 'fixed': # always from the
                                # same scan(s)
                                # check whether given scans are valid
                                if refscanstart <= refscanend and \
                                refscanend <= delembvecsref.shape[0]:
                                    delembvecsreffinal = \
                                    np.mean( delembvecsref\
                                        [refscanstart-1:refscanend, \
                                        refscanxstart-1:refscanxend, \
                                        refscanystart-1:refscanyend, \
                                      0:1+embdim*(tau+1):(tau+1), \
                                      0:nDelayembvecs], axis = (0, 1, 2))
                                        # mean over three axes, scan
                                        # itself can now differ as well
                                else:
                                    print('Either the given first scan \
                                          is larger than the last given \
                                          scan or the scans are outside \
                                          the bounds of the volume scan \
                                          or they have not been defined \
                                          at all. \nPlease correct.')
                        else:
                            print("The argument 'refscanCJRP' has to be \
                                  given as 'int' (reference for CRP or JRP \
                                  is from same volume scan than the actual \
                                  calculations are performed on) or 'ext' \
                                  (reference for CRP or JRP is from an \
                                  external scan; not implemented yet). \
                                  Please give correct argument. Default is \
                                  'int'. For using only one scan, define \
                                  start and end as same value. Ignore zero \
                                  indexing of Python, give e.g. first scan \
                                  as 1.")

                    else: # standard RP calculation
                        delembvecsreffinal = 0 # define as dummy because
                        # needed in later calculations for RP
                    # checking whether correct argument for 'RPtype' has been
                    # given will be done later
                    # check whether Theiler has been given correctly
                    if isinstance(Theiler, str):
                        print("'Theiler' has to be given as non-negative \
                              number (0 corresponds to no Theiler window, \
                              >0 to window). \nPlease correct.")

                    # start of calculation features --
```

```
543                 # loop for different A-scans in one volume scan
544                 for x in range(delembvecs.shape[1]):
545                     for y in range(delembvecs.shape[2]):
546                         # creation of distance matrix 'tempdistmat'
547                         # (and tempdistmat2 in case of JRP; this is
548                         # however always created as dummy)
549                         # (unthresholded recurrence plot) in extra
550                         # function (matrix is discarded after calculation
551                         # of RQA features for all epsilon)
552                         (tempdistmat, tempdistmat2) = \
553                         calcdistmat(delembvecs, metric, RPtype, \
554                         nDelayembvecs, curnrscan, x, y, embdim, tau, \
555                         delembvecsreffinal, LOI, Theiler, recurrence)
556                         # loop for different epsilon (or percentage
557                         # recurrence etc.) to calculate recurrence matrix
558                         # and RQA features
559                         for nStepseps in range(len(epsvec)):
560                             epsilonorperc = epsvec[nStepseps]
561                             # temporary recurrence matrix (discarded
562                             # after calculation of features)
563                             # calculated in extra function
564                             (temprecmat, epsilon) = calcRP \
565                               (tempdistmat, tempdistmat2, recurrence, \
566                               epsilonorperc, nDelayembvecs, RPtype)
567                             # pre-define temporary variable that contains
568                             # the number of lines of different length,
569                             # including the separated 1s, i.e.
570                             # diagonal "lines" of length 1:
571                             # histogram of diagonal lines
572                             hist_diaglines = np.zeros(nDelayembvecs)
573                             # running to length of RP, because, when
574                             # using LOI, this is included (i.e. for
575                             # symmetric RPs the longest line will be LOI)
576                             (hist_diaglines, RRdiag) = \
577                             calc_histdiaglines\
578                             (temprecmat, nDelayembvecs, RRdiag, \
579                              hist_diaglines, LOI, Theiler)
580                             # double this for symmetric recurrence
581                             # matrices (but not the last entry, the LOI
582                             # (in case it is used)), i.e. for RP or JRP
583                             # and global recurrence...
584                             if RPtype != 'CRP' and (recurrence == \
585                             'fixedepsilon' or recurrence == \
586                             'fixedrecglobal'):
587                                 hist_diaglines = hist_diaglines + \
588                                 np.concatenate((hist_diaglines[:-1], \
589                                                 range(1)))
590                             # ...or let it run again with negative values
591                             # for non-symmetric RPs (with fixed local
592                             # recurrence or fixed amount of neighbours)
593                             # leaving out the LOI in 2nd calculation
594                             # below (because it has been taken into
595                             # account in first run) is taken care for in
596                             # function calc_histdiaglines
```

```
                else:
                    # 2nd temporary variable with RRs per
                    # diagonal
                    RRdiag2 = np.zeros(nDelayembvecs)
                    # hist_diaglines from prior calculation
                    # taken and filled up with values per
                    # line length
                    (hist_diaglines, RRdiag2) = \
                    calc_histdiaglines\
                    (temprecmat, -nDelayembvecs, RRdiag2, \
                     hist_diaglines, LOI, -Theiler)
                    # calculate RRdiag for non-symmetric
                    # recurrence matrices as mean of both
                    # RRdiags
                    RRdiag = (RRdiag+RRdiag2)/2.
                # RR calculation after calculation of
                # histogram of diagonal lines (RR calculation
                # is based here for consistency on the
                # histogram, though it could be done by
                # simple summing, however leaving out LOI
                # and Theiler window if applicable)
                RR[curnrscan, x, y, embdim, tau, nStepseps] \
                = calcRR(LOI, Theiler, hist_diaglines, \
                       nDelayembvecs)

                # distinguishing whether or not to store
                # epsilon for TND and COR
                # for TND, Theiler window need not taken into
                # account here, because it was taken into
                # account when calculating RRdiag
                # external calculation
                if recurrence == 'fixedepsilon':
                    TND[curnrscan, x, y, embdim, tau, \
                     nStepseps] = \
                     calcTND(vecnrdiag, maxindexdiagTNDCOR, \
                             meandist, LOI, Theiler, RRdiag)
                    COR[curnrscan, x, y, embdim, tau, \
                       nStepseps] = \
                       calcCOR(vecnrdiag,maxindexdiagTNDCOR,\
                               meandist, LOI, Theiler,RRdiag)
                else:
                    TND[curnrscan, x, y, embdim, tau, \
                     nStepseps, 0] = \
                     calcTND(vecnrdiag, maxindexdiagTNDCOR,\
                             meandist, LOI, Theiler, RRdiag)
                    COR[curnrscan, x, y, embdim, tau, \
                       nStepseps, 0] = \
                       calcCOR(vecnrdiag,maxindexdiagTNDCOR,\
                               meandist, LOI, Theiler,RRdiag)
                    TND[curnrscan, x, y, embdim, tau, \
                       nStepseps, 1] = epsilon
                    COR[curnrscan, x, y, embdim, tau, \
                       nStepseps, 1] = epsilon
                    # store epsilon also for RR
```

```
651                                     RR[currnrscan, x, y, embdim, tau, \
652                                        nStepseps, 1] = epsilon
653
654                                 # now determining features based on diagonal
655                                 # lines for different minimum line length out
656                                 # of hist_diaglines
657                                 # indexing in Python
658                                 for lmin in lminvec: # attention: d, tau and
659                                     # epssteps are run through in Python zero
660                                     # indexing; lmin are 'real' values
661                                     # determinism
662                                     DET[currnrscan, x, y, embdim, tau, \
663                                         nStepseps, lmin-1]\
664                                         = (np.dot(hist_diaglines[lmin-1:],\
665                                         np.arange(lmin, nDelayembvecs+1)))/ \
666                                         (np.dot(hist_diaglines,\
667                                         np.arange(1, nDelayembvecs+1)))
668                                     # RATIO = DET/RR
669                                     # if loop to distinguish RR with or
670                                     # without epsilon information
671                                     if recurrence == 'fixedepsilon':
672                                         RATIO[currnrscan, x, y, embdim, tau, \
673                                         nStepseps, lmin-1]\
674                                         = DET[currnrscan, x, y, embdim, tau, \
675                                             nStepseps, lmin-1]/ \
676                                             RR[currnrscan, x, y, embdim, tau,\
677                                             nStepseps]
678                                     else:
679                                         RATIO[currnrscan, x, y, embdim, tau, \
680                                         nStepseps, lmin-1]\
681                                         = DET[currnrscan, x, y, embdim, tau, \
682                                             nStepseps, lmin-1]/ \
683                                             RR[currnrscan, x, y, embdim, tau,\
684                                             nStepseps, 0]
685                                     # ave. line length (in dissertation 'L')
686                                     Lave[currnrscan, x, y, embdim, tau, \
687                                         nStepseps, lmin-1] = np.dot\
688                                          (hist_diaglines[lmin-1:], np.arange\
689                                           (lmin, nDelayembvecs+1))/\
690                                             hist_diaglines[lmin-1:].sum()
691                                     # normalised average line length Lnor
692                                     Lnor[currnrscan, x, y, embdim, tau, \
693                                         nStepseps, lmin-1]\
694                                         = np.float32(lmin)/Lave[currnrscan, \
695                                                     x, y, embdim, tau, \
696                                         nStepseps, lmin-1] # float32 to get
697                                         # inf as result instead of error; inf
698                                         # values are changed to 0 at the end
699                                     # entropy ENT
700                                     # probability of occurrence of lines
701                                     # to determine entropy
702                                     p = hist_diaglines[lmin-1:]/ \
703                                         hist_diaglines[lmin-1:].sum()
704                                     ENT[currnrscan, x, y, embdim, tau, \
```

```
                                    nStepseps, lmin-1] = \
                                    -np.dot(np.log(p[p>0]), p[p>0])
                                    # logarithm to the base 2 # to get
                                    # values to base e, has to be
                                    # multiplied with 1/np.log(2)
                                # storing epsilon
                                if recurrence != 'fixedepsilon':
                                    DET[currnrscan, x, y, embdim, tau, \
                                      nStepseps, 0] = epsilon
                                    RATIO[currnrscan, x, y, embdim, tau, \
                                      nStepseps, 0] = epsilon
                                    Lave[currnrscan, x, y, embdim, tau, \
                                      nStepseps, 0] = epsilon
                                    Lnor[currnrscan, x, y, embdim, tau, \
                                      nStepseps, 0] = epsilon
                                    ENT[currnrscan, x, y, embdim, tau, \
                                      nStepseps, 0] = epsilon
                            # determining divergence (inverse of maximum
                            # line length, independent of minimum
                            # line length)
                            # checking whether or not to store epsilon
                            if recurrence == 'fixedepsilon':
                                # try - exception for the case that there
                                # is no recurrence point and thus only 0
                                # entries in histogram of diagonal lines
                                try:
                                    DIV[currnrscan, x, y, embdim, tau, \
                                     nStepseps] = 1/ \
                                      (np.nonzero(hist_diaglines)[0].\
                                       max()+1)
                                except ValueError:
                                    DIV[currnrscan, x, y, embdim, tau, \
                                        nStepseps] = 0
                            else:
                                try:
                                    DIV[currnrscan, x, y, embdim, tau, \
                                        nStepseps, 0] = 1/ \
                                        (np.nonzero(hist_diaglines)[0].\
                                         max()+1)
                                    DIV[currnrscan, x, y, embdim, tau, \
                                        nStepseps, 1] = epsilon
                                except ValueError:
                                    DIV[currnrscan, x, y, embdim, tau, \
                                        nStepseps, 0] = 0
                            # DIV is given trivially as the length of the
                            # LOI (of symmetric RP and JRP if fixed
                            # epsilon is used), if LOI is not excluded
                            # (i.e. if LOI='with')

    # convert NaNs in DET and RATIO, which are created by having only
    # zeros in (relevant entries of) hist_diaglines, to 0
    DET[np.isnan(DET)] = 0
    RATIO[np.isnan(RATIO)] = 0

```

```
759     # convert NaNs in Lave (created if no diagonal line with length equal
760     # to or greater than minimum line length in calculation) to 0
761     Lave[np.isnan(Lave)] = 0
762 
763     # do the same for Lnor, when Lave has values 0
764     Lnor[np.isnan(Lnor)] = 0
765 
766     # same for COR
767     COR[np.isnan(COR)] = 0
768 
769     # change 'inf' to zero
770     COR[COR == np.inf] = 0
771     Lnor[Lnor == np.inf] = 0
772 
773     # multiplying TND by 1000
774     TND = 1000*TND
775 
776     elapsedall = timeit.default_timer() - start_timeall
777     print('overall running time was:')
778     print(elapsedall)
779 
780     #create dictionary with all chosen values
781     info = {'disttoNforTNDCOR': disttoNforTNDCOR, \
782     'trigIEgate_start': trigIEgate_start, 'trigIEgate_end': \
783     trigIEgate_end, 'trigBWEgate_start': trigBWEgate_start, \
784     'trigBWEgate_end': trigBWEgate_end, 'metric': metric, 'recurrence': \
785     recurrence, 'LOI': LOI, 'Theiler': Theiler, 'RPtype': RPtype, \
786     'refscanCJRP': refscanCJRP, 'refscanint': refscanint, \
787     'refscanxstart': refscanxstart, 'refscanxend': refscanxend, \
788     'refscanystart': refscanystart, 'refscanyend': refscanyend, \
789     'refscanstart': refscanstart, 'refscanend': refscanend, \
790     'running time [s]': elapsedall}
791 
792     return(RR, DET, RATIO, Lave, Lnor, DIV, TND, COR, ENT, \
793            disttoNforTNDCOR, embdimmin, embdimmax, taumin, taumax, \
794            taustep, epsvec, lminvec, info, trigIEgate_start, \
795            trigIEgate_end, trigBWEgate_start, trigBWEgate_end, \
796            metric, recurrence, LOI, Theiler, RPtype, refscanCJRP, \
797            refscanint, refscanxstart, refscanxend, refscanystart, \
798            refscanyend, refscanstart, refscanend, delembvecsref)
```

Bibliography

All URLs and DOIs have been checked on 26 January 2020 except where otherwise stated.

[Abraham et al. 1989]
Abraham, N. B., A. M. Albano, A. Passamante & P. E. Rapp, eds. (1989). *Measures of Complexity and Chaos*. Boston: Springer.

[Abu-Mostafa et al. 2012]
Abu-Mostafa, Y. S., M. Magdon-Ismail & H.-T. Lin (2012). *Learning From Data: a Short Course*. [United States]: AMLBook.com. (p. 36)

[Adler et al. 1986]
Adler, L., C. Mobley & J. H. Rose (1986). 'Ultrasonic method to determine gas porosity in aluminum alloy castings: Theory and experiment'. In: *Journal of Applied Physics.- [S.l.]* DOI: 10.1063/1.336689. (p. 8)

[Aeyels 1981]
Aeyels, D. (1981). 'Generic observability of differentiable systems'. In: *SIAM Journal on Control and Optimization* 19.5, pp. 595–603. DOI: 10.1137/0319037. (p. 25)

[Alessio 2016]
Alessio, S. M. (2016). *Digital Signal Processing and Spectral Analysis for Scientists: Concepts and Applications*. Cham: Springer. (pp. 1, 18, 19)

[Altman et al. 2003]
Altman, D. G., D. Machin, T. N. Bryant & M. J. Gardner (2003). *Statistics with confidence: confidence intervals and statistical guidelines*. 2nd ed., repr. London: BMJ. (p. 40)

[Anderson 1994]
Anderson, T. W. (1994). *The Statistical Analysis of Time Series*. New York: Wiley. (p. 15)

[Argyris et al. 2015]
Argyris, J., G. Faust, M. Haase & R. Friedrich (2015). *An Exploration of Dynamical Systems and Chaos*. Completely revised and enlarged 2nd ed. Berlin: Springer. (pp. 24, 46)

[Arrowsmith & Place 1992]
Arrowsmith, D. K. & C. M. Place (1992). *Dynamical Systems: differential equations, maps and chaotic behaviour*. London: Chapman & Hall. (p. 24)

[ASTM 2014]
ASTM E-1065 / E-1065M - 14 Standard Guide for Evaluating Characteristics of Ultrasonic Search Units (2014). ASTM. West Conshohocken, PA, USA: ASTM Int. DOI: 10.1520/E1065_E1065M-14. (p. 21)

[Babloyantz 1989]
Babloyantz, A. (1989). 'Some Remarks on Nonlinear Data Analysis of Physiological Time Series'. In: *Measures of Complexity and Chaos*. Ed. by N. B. Abraham, A. M. Albano, A. Passamante & P. E. Rapp. Boston: Springer, pp. 51–62. DOI: 10.1007/978-1-4757-0623-9_4. (p. 30)

[Barreira 2003]
Barreira, L. (2003). 'Poincaré recurrence: old and new'. In: XIVth International Congress on Mathematical Physics (Lisboa, Portugal, 28th July–2nd Aug. 2003). World Scientific. DOI: 10.1142/9789812704016_0039. (p. 27)

[Barut et al. 2012]
Barut, S., V. Bissauge, G. Ithurralde & W. Claassens (2012). 'Computer-aided analysis of ultrasound data to speed-up the release of aerospace CFRP components'. In: 18th World Conference on Nondestructive Testing (Durban, South Africa, 16th–20th Apr. 2012). URL: http://www.ndt.net/article/wcndt2012/papers/667_wcndtfinal00669.pdf. (p. 52)

[Beals 1987]
Beals, R. (1987). *Advanced Mathematical Analysis: Periodic Functions and Distributions, Complex Analysis, Laplace Transform and Applications*. 3rd corrected printing. New York: Springer. (pp. 15, 17)

[Berthelot 1999]
Berthelot, J.-M. (1999). *Composite Materials: Mechanical Behavior and Structural Analysis*. New York: Springer. (p. 3)

[Birleanu 2012]
Birleanu, F.-M. (2012). 'Design of RPA representation spaces for the analysis of transient signals'. PhD thesis. Université de Grenoble, France, pp. 1–133. URL: https://tel.archives-ouvertes.fr/tel-00762428v1/document. (p. 27)

[Birt & Smith 2004]
Birt, E. A. & R. A. Smith (2004). 'A review of NDE methods for porosity measurement in fibre-reinforced polymer composites'. In: *Insight - Non-Destructive Testing and Condition Monitoring* 46.11, pp. 681–686. DOI: 10.1784/insi.46.11.681.52280. (p. 14)

[Blitz 1967]
Blitz, J. (1967). *Fundamentals of Ultrasonics*. London: Butterworths. (pp. 5, 8)

[Blitz & Simpson 1996]
Blitz, J. & G. Simpson (1996). *Ultrasonic Methods of Non-destructive Testing*. London: Chapman & Hall. (pp. 6–8, 10–13, 21, 41)

[Bonett 2006]
Bonett, D. G. (July 2006). 'Confidence interval for a coefficient of quartile variation'. In: *Computational Statistics & Data Analysis* 50.11, pp. 2953–2957. DOI: 10.1016/j.csda.2005.05.007. (p. 119)

[Brandt & Maaß 2015]
Brandt, C. & P. Maaß (2015). 'A State Space Approach for the Non-Destructive Evaluation of CFRP with Ultrasonic Testing'. In: 7th International Symposium on NDT in Aerospace 2015 (Bremen, Germany, 16th–18th Nov. 2015). URL: http:

//www.ndt.net/article/aero2015/papers/Mo_3_A_1_Brandt_v2.pdf. (pp. 2, 13, 117)

[Brandt & Maaß 2016]
Brandt, C. & P. Maaß (2016). 'Recurrence Quantification Analysis for Non-Destructive Evaluation with an Application in Aeronautic Industry'. In: 19th World Conference on Non-Destructive Testing 2016 (Munich, Germany, 13th–17th June 2016). URL: http://www.ndt.net/article/wcndt2016/papers/we1i4.pdf. (pp. 4, 10, 13)

[Brandt 2010]
Brandt, C. (2010). 'Schichtdickenmessung auf CFK-Oberflächen'. In: DGZfP-Jahrestagung 2010 (annual conference of Germany society for NDT) (Erfurt, Germany, 10th–12th May 2010). URL: http://www.ndt.net/article/dgzfp2010/Inhalt/mi2c2.pdf. (p. 1)

[Brandt 2014]
Brandt, C. (31st July 2014). 'Technik der zerstörungsfreien Prüfung'. German pat. req. DE 10 2014 011 424 A1 2016.02.04. URL: https://depatisnet.dpma.de/DepatisNet/depatisnet?action=pdf&docid=DE102014011424A1&xxxfull=1.

[Brandt 2015]
Brandt, C. (28th July 2015). 'Technique for Non-destructive Testing'. U.S. pat. req. US 2016/0034422 A1. URL: https://depatisnet.dpma.de/DepatisNet/depatisnet?action=pdf&docid=US020160034422A1&xxxfull=1.

[Brandt 2016]
Brandt, C. (2016). 'Recurrence Quantification Analysis as an Approach for Ultrasonic Testing of Porous Carbon Fibre Reinforced Polymers'. In: *Recurrence Plots and Their Quantifications: Expanding Horizons. Proceedings of the 6th International Symposium on Recurrence Plots, Grenoble, France, 17-19 June 2015*. Ed. by C. L. Webber Jr., N. Marwan & I. Cornel. Vol. 180. Springer Proceedings in Physics. Cham: Springer. Chap. 19, pp. 355–377. DOI: 10.1007/978-3-319-29922-8_19.
© Springer International Publishing Switzerland 2016. (pp. 56, 114)

[Brandt 2018]
Brandt, C. (2018). *Recurrence Quantification Analysis for Non-Destructive Testing of Porosity in Carbon Fibre Reinforced Polymers*. URL: https://www.researchgate.net/publication/323695223_Recurrence_Quantification_Analysis_for_Non-Destructive_Testing_of_Porosity_in_Carbon_Fibre_Reinforced_Polymers.

[Brandt et al. 2019]
Brandt, C., M. Hamann & J. Leuschner (2019). *Regression Models for Ultrasonic Testing of Carbon Fiber Reinforced Polymers*. URL: http://www.math.uni-bremen.de/zetem/cms/media.php/262/BerichtTechnomathematik1901english.pdf. (pp. 21, 65, 88, 111)

[Brockwell & Davis 1991]
Brockwell, P. J. & R. A. Davis (1991). *Time Series: Theory and Methods*. New York: Springer. (pp. 15, 17)

[Brodersen et al. 2010]
Brodersen, K. H., C. S. Ong, K. E. Stephan & J. M. Buhmann (2010). 'The Balanced Accuracy and Its Posterior Distribution'. In: *2010 20th International Conference on Pattern Recognition*, pp. 3121–3124. DOI: 10.1109/ICPR.2010.764. (p. 38)

[Broer & Takens 2011]
Broer, H. W. & F. Takens (2011). *Dynamical Systems and Chaos*. New York: Springer. (pp. 15, 22–25)

[Bunsell & Renard 2005]
Bunsell, A. R. & J. Renard (2005). *Fundamentals of fibre reinforced composite materials*. Bristol: Inst. of Physics Publ. (p. 3)

[Cacciola et al. 2007]
Cacciola, M., D. Costantino, C. M. Francesco & M. Versaci (Aug. 2007). 'Dynamical analysis for flaw shape identification in non linear eddy current tests'. In: *COMPEL* 26.4, pp. 1081–1094. DOI: 10.1108/03321640710756401. (p. 46)

[Carathéodory 1919]
Carathéodory, C. (1919). 'Über den Wiederkehrsatz von Poincaré'. In: *Sitzungsberichte der Preussischen Akademie der Wissenschaften* 1, pp. 580–584. (p. 27)

[Carrión Garcia et al. 2017]
Carrión Garcia, A., V. Genovés Gómez, R. Miralles Ricós, J. J. Payá Bernabeu & J. Gosálbez Castillo (2017). 'An Advanced Ultrasonic Method based on Signal Modality for Structural Damage Characterization on Concrete: The Cube Problem'. In: *2017 IEEE International Ultrasonics Symposium (IUS)*, pp. 1–4. DOI: 10.1109/ULTSYM.2017.8091944. (p. 47)

[Carrión et al. 2015a]
Carrión, A., G. Lara, R. Miralles, J. Gosalbez & I. Bosch (2015a). *On the use of Recurrence Quantification Analysis for Signal Modality Characterization: two applications*. URL: http://www.iteam.upv.es/wp-content/uploads/pdf_articles/83.pdf. (pp. 27, 47)

[Carrión et al. 2014]
Carrión, A., R. Miralles & G. Lara (Sept. 2014). 'Measuring predictability in ultrasonic signals: An application to scattering material characterization'. In: *Ultrasonics* 54.7, pp. 1904–1911. DOI: 10.1016/j.ultras.2014.05.008. (pp. 47, 112, 114)

[Carrión et al. 2015b]
Carrión, A., R. Miralles & G. Lara (2015b). 'Scattering material characterization based on Recurrence Plots Quantification Analysis (RQA)'. not published. Sixth International Symposium on Recurrence Plots (Grenoble, France, 17th–19th June 2015). (p. 27)

[Chawla 2013]
Chawla, K. K. (2013). *Composite materials: science and engineering*. New York: Springer. (pp. 3, 4)

[Cheeke 2002]
Cheeke, J. D. N. (2002). *Fundamentals and applications of ultrasonic waves*. Boca Raton, Fla: CRC Press. (pp. 5, 7, 8, 10)

[Chen et al. 2016]
Chen, Y., C. Yang, X. Zhou, Z. Li & H. Zheng (Feb. 2016). 'CFRP voids 3D identification and location method based on the process of backscattered signal'. In: *Journal of Wuhan University of Technology-Mater.Sci.Ed.* 31.1, pp. 172–177. DOI: 10.1007/s11595-016-1348-1. (pp. 8, 42, 44, 111)

[Chillingworth 1976]
Chillingworth, D. R. J. (1976). *Differential topology with a view to applications*. London: Pitman. (pp. 25, 26)

[Clopper & Pearson 1934]
Clopper, C. J. & E. S. Pearson (1934). 'The Use of Confidence or Fiducial Limits Illustrated in the Case of the Binomial'. In: *Biometrika* 26.4, pp. 404–413. URL: http://www.barestatistics.nl/uploads/1/1/7/9/11797954/clopper__pearson_1934.pdf. (p. 40)

[Collins Cobuild 1999]
Collins Cobuild English Dictionary (1999). London: Harper Collins publ. (p. 26)

[Colloquium 2018]
The Dynamics Group, Hamburg (2018). *Colloquium Irregular Engineering Oscillations and Signal Processing*. URL: https://cgi.tu-harburg.de/~dynwww/cgi-bin/colloquium-2018/. (p. 1)

[Damaschke 1996]
Damaschke, F. (1996). 'Quantitative Bewertung der inneren Struktur von Faserverbundwerkstoffen mittels Ultraschallsignalanalyse'. Diplomarbeit, pp. 1–136. URL: http://gso.gbv.de/DB=2.1/SET=1/TTL=1/MAT=/NOMAT=T/CLK?IKT=8062%5C&TRM=Quantitative+Bewertung+der+inneren+Struktur+von+Faserverbundwerkstoffen+mittels+Ultraschallsignalanalyse. (p. 49)

[Daniel et al. 1992]
Daniel, I. M., S. C. Wooh & I. Komsky (1992). 'Quantitative porosity characterization of composite materials by means of ultrasonic attenuation measurements'. In: *Journal of Nondestructive Evaluation*. DOI: 10.1007/BF00566012. (p. 14)

[Dep. of Defense 1998]
Joint Service Specification Guide JSSG-2006: Aircraft Structures (1998). Department of Defense USA. URL: http://everyspec.com/USAF/USAF-General/JSSG-2006_10206/. (p. 38)

[DeSanto 1992]
DeSanto, J. A. (1992). *Scalar wave theory: Green's functions and applications*. Berlin: Springer. (p. 7)

[Digulescu et al. 2016]
Digulescu, A., I. Murgan, C. Ioana, I. Candel & A. Serbanescu (2016). 'Applications of Transient Signal Analysis Using the Concept of Recurrence Plot Analysis'. In: *Recurrence Plots and Their Quantifications: Expanding Horizons. Proceedings of the 6th International Symposium on Recurrence Plots, Grenoble, France, 17-19*

June 2015. Ed. by C. L. Webber Jr., N. Marwan & I. Cornel. Vol. 180. Springer Proceedings in Physics. Cham: Springer, pp. 19–38. DOI: 10.1007/978-3-319-29922-8_2. (p. 26)

[Diks 1999]
Diks, C. (1999). *Nonlinear Time Series Analysis: Methods and Applications*. Vol. 4. Nonlinear Time Series and Chaos. Singapore; River Edge, NJ: World Scientific. (p. 22)

[DIN EN 2015]
DIN EN 1330-1:2015-05 Non-destructive testing - terminology - part 1: List of general terms (trilingual version) (2015). 2nd ed. Deutsches Institut für Normung e.V. Berlin, Germany: Beuth. DOI: 10.31030/2270074. (p. 8)

[DIN EN 2016]
DIN EN 12668-2:2010-06 Non-destructive testing — Characterization and verification of ultrasonic examination equipment — Part 2: Probes (2016). British Standards Institution. Berlin, Germany: Beuth Verlag. DOI: 10.31030/1545296. (pp. 10, 21)

[Dominguez 2005]
Dominguez, N. (2005). 'Modélisation de la propagation ultrasonore en milieu complexe : application au contrôle non destructif et à la caractérisation de la porosité dans les matériaux composites stratifié'. PhD thesis. Université de Toulouse III, France, 1 o. (180 p.) URL: http://www.theses.fr/2006TOU30235. (p. 8)

[Dominguez & Mascaro 2006]
Dominguez, N. & B. Mascaro (2006). 'Ultrasonic Non-Destructive Inspection of Localised Porosity in Composite Materials'. In: ECNDT 2006 - 9th European Conference on NDT 2006 (Berlin, Germany, 18th–19th Sept. 2006). URL: http://www.ndt.net/article/ecndt2006/doc/Tu.2.1.4.pdf. (pp. 8, 42–44, 111)

[Drinkwater & Wilcox 2006]
Drinkwater, B. W. & P. D. Wilcox (2006). 'Ultrasonic arrays for non-destructive evaluation: A review'. In: *NDT and E International* 39.7, pp. 525–541. DOI: 10.1016/j.ndteint.2006.03.006. (p. 10)

[Duda et al. 2001]
Duda, R. O., D. G. Stork & P. E. Hart (2001). *Pattern Classification*. New York: Wiley. (pp. 36, 37)

[Eckmann et al. 1987]
Eckmann, J.-P., S. Oliffson Kamphorst & D. Ruelle (1987). 'Recurrence Plots of Dynamical Systems'. In: *EPL (Europhysics Letters)* 4.9, p. 973. URL: http://stacks.iop.org/0295-5075/4/i=9/a=004. (pp. 27, 28, 30, 45)

[Eckmann et al. 1986]
Eckmann, J.-P., S. Oliffson Kamphorst, D. Ruelle & S. Ciliberto (Dec. 1986). 'Liapunov exponents from time series'. In: *Phys.Rev.A* 34.6, pp. 4971–4979. DOI: 10.1103/PhysRevA.34.4971. (p. 28)

[Ehrhart 2016]
Ehrhart, B. (2016). 'Quality assessment of bonded primary CFRP structures by

means of laser proof testing'. PhD thesis. Universität Bremen, Germany. URL: https://nbn-resolving.de/urn:nbn:de:gbv:46-00105320-18. (p. 8)

[Engl et al. 2003]
Engl, G., R. Meier & C. Brandt (2003). 'Testing large aerospace CFRP components by ultrasonic multichannel conventional and phased array pulse-echo techniques'. In: *Insight - Non-Destructive Testing and Condition Monitoring* 45.3, pp. 202–205. DOI: 10.1784/insi.45.3.202.53160. (p. 12)

[Fahr 2014]
Fahr, A. (2014). *Aeronautical Applications of Non-destructive Testing*. Lancaster, PA: DEStech Publ. (pp. 7, 8, 12, 38, 44, 49)

[Flach 2012]
Flach, P. (2012). *Machine Learning: The Art and Science of Algorithms that Make Sense of Data*. Cambridge: Cambridge Univ. Press. (pp. 35–39)

[Freedman 2008]
Freedman, D. (2008). *Statistical Models: Theory and Practice*. Cambridge: Cambridge Univ. Press. (p. 16)

[Furstenberg 1981]
Furstenberg, H. (1981). *Recurrence in ergodic theory and combinatorial number theory*. Princeton, NJ.: Princeton Univ. Press. (pp. 26, 27)

[Galton 1886a]
Galton, F. (1886a). 'Hereditary Stature'. In: *Nature*. DOI: 10.1038/033317b0. (p. 35)

[Galton 1886b]
Galton, F. (1886b). 'Regression Towards Mediocrity in Hereditary Stature.' In: *The Journal of the Anthropological Institute of Great Britain and Ireland* 15, pp. 246–263. DOI: 10.2307/2841583. (p. 35)

[Galton 1888]
Galton, F. (1888). 'Co-Relations and Their Measurement, Chiefly from Anthropometric Data'. In: *Proceedings of the Royal Society of London* 45, pp. 135–145. URL: https://www.jstor.org/stable/114860. (p. 35)

[Gausmann 1995]
Gausmann, A. (1995). 'Quantitative Bewertung von Poren in Faserverbundwerkstoffen mittels Ultraschalltechnik'. Diplomarbeit. Fachhochschule Osnabrück, Germany. (p. 49)

[Goebbels 1980]
Goebbels, K. (1980). 'Structure Analysis by Scattered Ultrasonic Radiation'. In: *Research techniques in nondestructive testing*. Ed. by R. S. Sharpe. Vol. 4. Research Techniques in Nondestructive Testing. London: Acad. Press, pp. 87–157. (p. 41)

[Grager et al. 2018]
Grager, J.-C., H. Mooshofer & C. U. Große (2018). 'Evaluation of the imaging performance of a CFRP-adapted TFM algorithm'. In: *ECNDT 2018 - 12th European Conference on NDT 2018 Online-Proceedings on ndt.net* (11th–15th June 2018).

Gothenburg, Sweden. URL: https://www.ndt.net/article/ecndt2018/papers/ecndt-0183-2018.pdf. (p. 11)

[Grager et al. 2016]
Grager, J.-C., M. Schrapp, H. Mooshofer, M. G. R. Sause, A.-M. Zelenyak & C. U. Große (2016). 'Ultrasonic Imaging of Carbon Fiber-Reinforced Plastics Using the Full Matrix Capture Data Acquisition Technique'. In: *19th World Conference on Non-Destructive Testing 2016 Online-Proceedings on ndt.net* (13th–17th June 2016). Munich, Germany. URL: https://www.ndt.net/article/wcndt2016/papers/fr1d4.pdf. (p. 11)

[Grolemund & Tsai 1992]
Grolemund, D. & C. S. Tsai (Oct. 1992). 'Statistical moments of Rayleigh scattered fields in porous fiber reinforced composites'. In: *IEEE 1992 Ultrasonics Symposium Proceedings*, 815–818 o.2. DOI: 10.1109/ULTSYM.1992.275871. (pp. 41, 43)

[Grolemund & Tsai 1995]
Grolemund, D. & C. S. Tsai (1995). 'Phase-Insensitive Detection of Backscattered Ultrasound for Void Content Determination in Fiber Reinforced Composites'. In: ed. by J. P. Jones. Acoustical Imaging. Boston, MA: Springer US, pp. 661–666. DOI: 10.1007/978-1-4615-1943-0_73. (p. 41)

[Grolemund & Tsai 1998]
Grolemund, D. & C. S. Tsai (1998). 'Statistical moments of backscattered ultrasound in porous fiber reinforced composites'. In: *Front Cover* 50.2, pp. 295–304. DOI: 10.1109/58.660140. (pp. 41–43, 111)

[Große 2016]
Große, C. U. (June 2016). 'Modules for an Efficient and Holistic Application of NDT Methods for Fiber Reinforced Polymers'. In: *ECCM17 17th European Conference on Composite Materials*. Munich, Germany. (p. 8)

[Guckenheimer & Holmes 2002]
Guckenheimer, J. & P. Holmes (2002). *Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields*. New York: Springer. (pp. 24, 26)

[Halmshaw 1987]
Halmshaw, R. (1987). *Non-destructive Testing*. London: Arnold. (pp. 5, 8)

[Hand 1981]
Hand, D. J. (1981). *Discrimination and classification*. Chichester: Wiley. (p. 36)

[Hastie et al. 2009]
Hastie, T., R. Tibshirani & J. H. Friedman (2009). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. 2nd ed. New York: Springer. (pp. 16, 17, 36)

[He et al. 2018]
He, X., S. Jin & L. Lin (2018). 'Simulation on non-destructive evaluation of CFRP void distribution with recurrence quantification analysis of ultrasonic back-scatter signals'. In: *Fuhe Cailiao Xuebao/ Acta Materiae Compositae Sinica* 35.10, pp. 2753–2759. DOI: 10.13801/j.cnki.fhclxb.20180105.001. (p. 45)

[Henneke 1990]
Henneke II, E. G. (1990). 'Ultrasonic Nondestructive Evaluation of Advanced

Composites'. In: *Non-Destructive Testing of Fibre-Reinforced Plastics Composites*. Ed. by J. Summerscales. Vol. 2. London: Elsevier Applied Science. Chap. 2, pp. 55–159. (pp. 8, 10)

[Hillger & Elze 2002]
Hillger, W. & S. Elze (2002). *Determination of Porosity in Aerospace Structures by Ultrasonic Pulse Echo Technique*. URL: `http://www.dlr.de/fa/Portaldata/17/Resources/dokumente/institut/2002/u2002_2.pdf`. (pp. 42, 43)

[Hillger et al. 2002]
Hillger, W., R. Henrich & S. Elze (2002). 'Ultraschallprüftechnik zur Porositätsbestimmung in CFK-Bauteilen'. In: *DGZfP-Jahrestagung 2002 (annual conference of Germany society for NDT)*. Weimar, Germany. URL: `https://www.ndt.net/article/dgzfp02/papers/p02/p02.htm`. (p. 43)

[Hinrichsen & Pritchard 2010]
Hinrichsen, D. & A. J. Pritchard (2010). *Mathematical Systems Theory*. 1, corrected softcover printing. Vol. 48. Texts in Applied Mathematics. Heidelberg: Springer. (pp. 23, 27)

[Hogan & Lakey 2005]
Hogan, J. A. & J. D. Lakey (2005). *Time-Frequency and Time-Scale Methods: Adaptive Decompositions, Uncertainty Principles, and Sampling*. Boston: Birkhäuser. (p. 21)

[Hsu 2013]
Hsu, D. K. (2013). 'Non-destructive evaluation (NDE) of aerospace composites: ultrasonic techniques'. In: *Non-destructive evaluation (NDE) of polymer matrix composites: Techniques and applications*. Ed. by V. M. Karbhari. Cambridge: Woodhead Pub., pp. 397–422. DOI: `10.1533/9780857093554.3.397`. (p. 108)

[Hsu 2014]
Hsu, H. P. (2014). *Signals and Systems*. 3rd ed. New York: McGraw-Hill. (p. 1)

[Huke 2006]
Huke, J. P. (2006). *Embedding nonlinear dynamical systems: A guide to Takens' theorem*. URL: `http://eprints.ma.man.ac.uk/175/1/embed.pdf`. (pp. 1, 25, 26)

[Hull & Clyne 1996]
Hull, D. & T. W. Clyne (1996). *An Introduction to Composite Materials*. Cambridge: Cambridge Univ. Press. (p. 3)

[Int. Committee f. NDT 2019]
International Committee for Non-Destructive Testing, The (2019). *ICNDT - The World Organisation for NDT*. URL: `http://www.icndt.org/`. (p. 10)

[Ioana et al. 2014]
Ioana, C., A. Digulescu, A. Serbanescu, I. Candel & F.-M. Birleanu (Jan. 2014). 'Recent Advances in Non-stationary Signal Processing Based on the Concept of Recurrence Plot Analysis'. In: *Translational Recurrences. From Mathematical Theory to Real-World Examples*. Ed. by N. Marwan, M. Riley, A. Giuliani & C. L. Webber Jr. Vol. 103. Springer Proceedings in Mathematics and Physics. Cham: Springer. Chap. 5, pp. 75–93. DOI: `10.1007/978-3-319-09531-8_5`. (p. 27)

[ISO 2005]
ISO 21347:2005(E): Space systems - Fracture and damage control (2005). Version 1. International Organization of Standardization. Geneva, Switzerland. URL: https://www.iso.org/obp/ui/#iso:std:iso:21347:ed-1:v1:en. (p. 38)

[ISO/TS 2005]
ISO/TS 18173:2005 Technical Specification: Non-destructive testing - General terms and definitions (2005). 1st ed. International Organization of Standardization. Geneva, Switzerland. URL: https://www.iso.org/obp/ui/#iso:std:iso:ts:18173:ed-1:v1:en. (p. 8)

[James et al. 2013]
James, G., D. Witten, T. Hastie & R. Tibshirani (2013). *An Introduction to Statistical Learning: with Applications in R*. New York: Springer. (pp. 35–37)

[Jeong & Hsu 1995]
Jeong, H. & D. K. Hsu (1995). 'Experimental analysis of porosity-induced ultrasonic attenuation and velocity change in carbon composites'. In: *Ultrasonics* 33.3, pp. 195–203. DOI: 10.1016/0041-624X(95)00023-V. (p. 14)

[Jeong et al. 2002]
Jeong, J., J. C. Gore & B. S. Peterson (Nov. 2002). 'A Method for Determinism in Short Time Series, and its Application to Stationary EEG'. In: *IEEE Transactions on Biomedical Engineering* 49.11, pp. 1374–1379. DOI: 10.1109/TBME.2002.804581. (pp. 47, 112)

[Jin et al. 2019]
Jin, S. J., X. C. He, J. Chen, S. Q. Shi & L. Lin (Dec. 2019). *Recurrence quantitative analysis for porosity characterization of CFRP with complex void morphology*. DOI: 10.1016/j.compstruct.2019.111383. (pp. 42, 45, 108, 111)

[Just & Schuster 2005]
Just, W. & H. G. Schuster (2005). *Deterministic Chaos: An Introduction*. Weinheim: Wiley-VCH. (p. 22)

[Kak et al. 1983]
Kak, A. C., M. Azimi & M. Slaney (1983). 'On the Estimation of Porosity in Composites by Oblique Angle Illumination'. In: *Review of Progress in Quantitative Nondestructive Evaluation*. Ed. by D. O. Thompson & D. E. Chimenti. Vol. 2A. Boston, MA: Springer. DOI: 10.1007/978-1-4613-3706-5_54. (p. 41)

[Kantz & Schreiber 2005]
Kantz, H. & T. Schreiber (2005). *Nonlinear Time Series Analysis*. 2nd ed., repr. Cambridge: Cambridge Univ. Press. (pp. 22, 23, 25, 26, 45, 112, 114)

[Kaplan & Glass 1992]
Kaplan, D. T. & L. Glass (Jan. 1992). 'Direct test for determinism in a time series'. In: *Physical Review Letters* 68.4, pp. 427–430. DOI: 10.1103/PhysRevLett.68.427. (p. 47)

[Karabutov & Podymova 2013]
Karabutov, A. A. & N. B. Podymova (Sept. 2013). 'Nondestructive Porosity Assessment of CFRP Composites with Spectral Analysis of Backscattered Laser-

Induced Ultrasonic Pulses'. In: *Journal of Nondestructive Evaluation* 32.3, pp. 315–324. DOI: 10.1007/s10921-013-0184-x. (pp. 42, 44, 111)

[Kennel & Isabelle 1992]
Kennel, M. B. & S. Isabelle (Sept. 1992). 'Method to distinguish possible chaos from colored noise and to determine embedding parameters'. In: *Phys.Rev.A* 46.6, pp. 3111–3118. DOI: 10.1103/PhysRevA.46.3111. (p. 112)

[Kim et al. 2013]
Kim, K.-B., D. K. Hsu & D. J. Barnard (June 2013). 'Estimation of porosity content of composite materials by applying discrete wavelet transform to ultrasonic backscattered signal'. In: *NDT & E International* 56, pp. 10–16. DOI: 10.1016/j.ndteint.2013.01.014. (pp. 21, 42, 43, 110, 111)

[Kokoska & Zwillinger 2000]
Kokoska, S. & D. Zwillinger (2000). *Standard probability and statistics tables and formulae*. Boca Raton: Chapman & Hall-CRC. (p. 16)

[Krämer et al. 2018]
Krämer, K. H., R. V. Donner, J. Heitzig & N. Marwan (2018). 'Recurrence threshold selection for obtaining robust recurrence characteristics in different embedding dimensions'. In: *Chaos* 28.8, p. 085720. DOI: 10.1063/1.5024914. (p. 28)

[Krautkrämer & Krautkrämer 1990]
Krautkrämer, J. & H. Krautkrämer (1990). *Ultrasonic Testing of Materials*. 4th ed. Berlin: Springer. (pp. 5, 6, 8, 10, 12, 13, 21)

[Kreyszig 1970]
Kreyszig, E. (1970). *Introductory Mathematical Statistics: Principles and Methods*. New York: Wiley. (pp. 17, 35, 40)

[Kreyszig 1989]
Kreyszig, E. (1989). *Introductory Functional Analysis with Applications*. reprint from 1978. New York: Wiley. (p. xiv)

[Kreyszig et al. 2011]
Kreyszig, E., H. Kreyszig & E. J. Norminton (2011). *Advanced engineering mathematics*. 10th ed. Hoboken, NJ: Wiley. (p. 18)

[Kurths & Herzel 1987]
Kurths, J. & H. Herzel (1987). 'An Attractor in Solar Time Series'. In: *Physica D* 25, pp. 165–172. DOI: 10.1016/0167-2789(87)90099-6. (p. 22)

[Kurths et al. 1995]
Kurths, J., A. Voss, A. Witt, P. Saparin, H. J. Kleiner & N. Wessel (1995). 'Quantitative Analysis of Heart Rate Variability'. In: *NLD Preprint*. URL: https://publishup.uni-potsdam.de/opus4-ubp/frontdoor/index/index/docId/1227. (p. 22)

[Kurz & Rabe 2016]
Kurz, J. H. & U. Rabe (16.-18.11.2015 2016). 'Porosity determination of carbon fiber reinforced plastics (CFRP) in aviation applications using ultrasound without a back wall echo'. In: *19th World Conference on Non-Destructive Testing 2016 Online-Proceedings on ndt.net*. Munich, Germany. URL: https://www.ndt.net/article/wcndt2016/papers/we2a2.pdf. (pp. 42, 44)

[León et al. 2011]
León, F. P., U. Kiencke & H. Jäkel (2011). *Signale und Systeme*. München: Oldenbourg. (p. 19)
[Loistl & Betz 1993]
Loistl, O. & I. Betz (1993). *Chaostheorie: Zur Theorie nichtlinearer dynamischer Systeme*. München: Oldenbourg. (p. 45)
[Louis et al. 1997]
Louis, A. K., P. Maaß & A. Rieder (1997). *Wavelets: Theory and Applications*. Chichester: Wiley. (p. 21)
[Lozak et al. 2014]
Lozak, A., C. Boller, A. Bulavinov, R. Pinchuk, J. Kurz & D. Sednev (2014). 'Phase Statistics and Spectral Analysis of Ultrasonic Signals for CFRP Component Assessment'. In: *7th European Workshop on Structural Health Monitoring July 8-11, 2014. Nantes, France*. 7th European Workshop on Structural Health Monitoring (8th–11th July 2014). Nantes, France. URL: https://hal.inria.fr/file/index/docid/1022992/filename/0442.pdf. (pp. 2, 44)
[Lu & Michaels 2006]
Lu, Y. & J. E. Michaels (2006). 'State Space Feature Extraction Applied to Diffuse Ultrasonic Signals Using Simulated Chaotic Excitations'. In: vol. 820. AIP Conference Proceedings. American Institute of Physics, pp. 625–632. DOI: 10.1063/1.2184585. (p. 46)
[Ludyk 1995]
Ludyk, G. (1995). *Theoretische Regelungstechnik 1*. 1st ed. Berlin: Springer. (p. 1)
[Mallick 2008]
Mallick, P. K. (2008). *Fiber-Reinforced Composites: Materials, Manufacturing, and Design*. Boca Raton, Fla.: CRC Press. (pp. 3, 4)
[Mann 2011]
Mann, P. S. (2011). *Introductory Statistics*. 7th edition. Hoboken, NJ: Wiley. (p. 35)
[Marwan & Kurths 2009]
Marwan, N. & J. Kurths (2009). 'Comment on "Stochastic analysis of recurrence plots with applications to the detection of deterministic signals" by Rohde et al. [Physica D 237 (2008) 619-629]'. In: *Physica D: Nonlinear Phenomena* 238.16, pp. 1711–1715. DOI: 10.1016/j.physd.2009.04.018. (p. 46)
[Marwan et al. 2007]
Marwan, N., M. C. Romano, M. Thiel & J. Kurths (2007). 'Recurrence plots for the analysis of complex systems'. In: *Physics Reports* 438.5-6, pp. 237–329. DOI: 10.1016/j.physrep.2006.11.001. (pp. 1, 27–29, 31, 33, 34, 114, 122)
[Marwan 2003]
Marwan, N. (2003). 'Encounters With Neighbours: Current Developments Of Concepts Based On Recurrence Plots And Their Applications'. PhD thesis. URL: http://nbn-resolving.de/urn:nbn:de:kobv:517-0000856. (p. 34)
[Marwan 2011]
Marwan, N. (2011). 'How to avoid potential pitfalls in recurrence plot based data

analysis'. In: *International Journal of Bifurcation and Chaos* 21.04, pp. 1003–1017. DOI: 10.1142/S0218127411029008. (p. 116)

[Marwan 2015]
Marwan, N. (2015). *Cross Recurrence Plot Toolbox for Matlab, Ver. 5.18 (R29.3)*. URL: http://tocsy.pik-potsdam.de/crp.php. (p. 114)

[Marwan & Webber 2015]
Marwan, N. & C. L. Webber Jr. (Jan. 2015). 'Mathematical and Computational Foundations of Recurrence Quantifications'. In: *Recurrence Quantification Analysis: Theory and Best Practices*. Ed. by C. L. Webber Jr. & N. Marwan. Understanding Complex Systems. Cham: Springer Int. Publ., pp. 3–43. DOI: 10.1007/978-3-319-07155-8_1. (pp. 26–29, 31, 32, 67)

[Mathworks 1993]
The MathWorks, Inc. (1993). *MATLAB HIgh-Performance Numeric Computation and Visualization Software: Reference Guide*. Natick, MA: MathWorks, Inc. (p. 17)

[Mayer-Kress & Hübler 1989]
Mayer-Kress, G. & A. Hübler (1989). 'Time Evolution of Local Complexity Measures and Aperiodic Perturbations of Nonlinear Dynamical Systems'. In: *Measures of Complexity and Chaos*. Ed. by N. B. Abraham, A. M. Albano, A. Passamante & P. E. Rapp. Boston: Springer, pp. 155–171. DOI: 10.1007/978-1-4757-0623-9_18. (p. 30)

[Mehdikhani et al. 2019]
Mehdikhani, M., L. Gorbatikh, I. Verpoest & S. V. Lomov (2019). 'Voids in fiber-reinforced polymer composites: A review on their formation, characteristics, and effects on mechanical performance'. In: *Journal of Composite Materials* 53.12, pp. 1579–1669. DOI: 10.1177/0021998318772152. (pp. 4, 14, 41)

[MIL-HDBK 2002]
MIL-HDBK-17-3F: Composite Materials Handbook Volume 3: Polymer Matrix Composites (2002). URL: http://everyspec.com/MIL-HDBK/MIL-HDBK-0001-0099/MIL_HDBK_17_3F_216/. (p. 4)

[Mitchell 1997]
Mitchell, T. M. (1997). *Machine Learning*. New York: McGraw-Hill. (pp. 32, 36)

[Mitra 2011]
Mitra, S. K. (2011). *Digital Signal Processing: A Computer-Based Approach*. New York: McGraw-Hill. (pp. 1, 43)

[Morse & Ingard 1968]
Morse, P. M. & K. U. Ingard (1968). *Theoretical Acoustics*. New York: McGraw-Hill. (p. 7)

[Nagy et al. 1987]
Nagy, P. B., D. V. Rypien & L. Adler (1987). 'Ultrasonic Attenuation Measurement by Backscattering Analysis'. In: *Review of Progress in Quantitative Nondestructive Evaluation*. Ed. by D. O. Thompson & D. E. Chimenti. Vol. 6A. Boston, MA: Springer, pp. 1411–1418. DOI: 10.1007/978-1-4613-1893-4_159. (p. 41)

[NIST 2017]
National Institute of Standards and Technology (2017). *Quartile Coefficient of*

Dispersion. URL: https://www.itl.nist.gov/div898/software/dataplot/refman2/auxillar/qcd.htm. (p. 16)

[Nemytskii & Stepanov 1972]
Nemytskii, V. V. & V. V. Stepanov (1972). *Qualitative Theory of Differential Equations*. Princeton, NJ: Princeton Univ. Press. (p. 27)

[Newcombe 1998]
Newcombe, R. G. (1998). 'Two-sided confidence intervals for the single proportion: comparison of seven methods. Stat. Med'. In: *Statistics in Medicine* 17, pp. 857–872. URL: http://stats.org.uk/statistical-inference/Newcombe1998.pdf. (p. 40)

[Nichols et al. 2006]
Nichols, J. M., S. T. Trickey & M. Seaver (2006). 'Damage detection using multivariate recurrence quantification analysis'. In: *Mechanical Systems and Signal Processing* 20.2, pp. 421–437. DOI: 10.1016/j.ymssp.2004.08.007. (p. 46)

[Niu 2000]
Niu, M. C.-Y. (2000). *Composite Airframe Structures: Practical Design Information and Data*. Hong Kong: Conmilit Press. (p. 4)

[Noakes 1991]
Noakes, L. (1991). 'The Takens Embedding Theorem'. In: *International Journal of Bifurcation and Chaos* 01.04, pp. 867–872. DOI: 10.1142/S0218127491000634. (p. 25)

[Nomura et al. 2013]
Nomura, Y., D. Morimoto, T. Kusaka & H.Furuta (2013). 'Structural damage localization based on recurrence plots of chaotic response attractor'. In: *Transactions of the Japan Society of Mechanical Engineers* 79.807, pp. 4210–4222. DOI: 10.1299/kikaic.79.4210. (p. 46)

[Oppenheim et al. 2014]
Oppenheim, A. V., A. S. Willsky & S. H. Nawab (2014). *Signals and Systems*. Harlow, Essex: Pearson Education. (p. 1)

[Ott 1993]
Ott, E. (1993). *Chaos in Dynamical Systems*. Cambridge: Cambridge Univ. Press. (pp. 22–24)

[Ott et al. 1994]
Ott, E., T. Sauer & J. A. Yorke (1994). *Coping With Chaos: Analysis of Chaotic Data and the Exploitation of Chaotic Systems*. New York: Wiley. (pp. 23, 26)

[Pilyugin 2012]
Pilyugin, S. Y. (2012). *Spaces of Dynamical Systems*. Vol. 3. De Gruyter Studies in Mathematical Physics. Berlin: Walter de Gruyter GmbH Co.KG. (p. 23)

[Place & Arrowsmith 1991]
Place, C. M. & D. K. Arrowsmith (1991). *An introduction to Dynamical Systems*. Cambridge: Cambridge Univ. Press. (p. 25)

[Poincaré 1890]
Poincaré, H. (1890). 'Sur le problème des trois corps et les équations de la dy-

namique'. In: *Acta Mathematica* 13, pp. 1–270. URL: http://henripoincarepapers.univ-nantes.fr/chp/hp-pdf/hp1890am.pdf. (p. 26)

[Prado & West 2010]
Prado, R. & M. West (2010). *Time Series: Modeling, Computation, and Inference*. Boca Raton, Fla.: Chapman & Hall/CRC Press. (p. 15)

[Press et al. 2007]
Press, W. H., S. A. Teukolsky, W. T. Vetterling & B. P. Flannery (2007). *Numerical Recipes: The Art of Scientific Computing*. Cambridge: Cambridge Univ. Press. (pp. 35, 114, 115)

[Priestley 1981]
Priestley, M. B. (1981). *Spectral Analysis and Time Series*. Vol. 1. London: Acad. Press. (pp. 15, 17, 22)

[Priestley 1988]
Priestley, M. B. (1988). *Non-linear and Non-stationary Time Series Analysis*. London: Acad. Press. (p. 22)

[Qu & Achenbach 1987]
Qu, J. & J. D. Achenbach (1987). 'Analytical Treatment of Polar Backscattering from Porous Composites'. In: *Review of Progress in Quantitative Nondestructive Evaluation*. Ed. by D. O. Thompson & D. E. Chimenti. Vol. 6A. Boston, MA: Springer, pp. 1137–1146. DOI: 10.1007/978-1-4613-1893-4_130. (p. 41)

[Richardson 1962]
Richardson, E. G. (1962). *Ultrasonic Physics*. Amsterdam: Elsevier. (p. 5)

[Richert & Coelho 2013]
Richert, W. & L. P. Coelho (2013). *Building machine learning systems with Python: master the art of machine learning with Python and build effective machine learning systems with this intensive hands-on guide*. Birmingham: Packt Publ. (p. 36)

[Roberts 1987]
Roberts, R. A. (1987). 'Porosity Characterization in Fiber-Reinforced Composites by Use of Ultrasonic Backscatter'. In: *Review of Progress in Quantitative Nondestructive Evaluation*. Ed. by D. O. Thompson & D. E. Chimenti. Vol. 6A. Boston, MA: Springer. DOI: https://doi.org/10.1007/978-1-4613-1893-4_131. (p. 41)

[Robinson 2010]
Robinson, J. C. (2010). 'The Takens Time-Delay Embedding Theorem'. In: *Dimensions, embeddings, and attractors*. Dimensions, Embeddings, and Attractors. Cambridge: Cambridge Univ. Press, pp. 145–159. DOI: 10.1017/CBO9780511933912.016. (pp. 25, 26)

[Robinson 1999]
Robinson, R. C. (1999). *Dynamical Systems: Stability, Symbolic Dynamics, and Chaos*. Boca Raton, Fla.: CRC Press. (p. 24)

[Rohde et al. 2008]
Rohde, G. K., J. M. Nichols, B. M. Dissinger & F. Bucholtz (2008). 'Stochastic analysis of recurrence plots with applications to the detection of deterministic

signals'. In: *Physica D: Nonlinear Phenomena* 237.5, pp. 619–629. DOI: 10.1016/j.physd.2007.10.008. (pp. 45, 46)

[Rose 1999]
Rose, J. L. (1999). *Ultrasonic Waves in Solid Media*. Cambridge: Cambridge Univ. Press. (pp. 7, 8)

[Rose 2014]
Rose, J. L. (2014). *Ultrasonic Guided Waves in Solid Media*. Cambridge: Cambridge Univ. Press. (pp. 5, 10)

[Rubin & Jerina 1994]
Rubin, A. M. & K. L. Jerina (1994). 'Evaluation of porosity in composite aircraft structures'. In: *Mechanics of composite materials* 30.1. DOI: 10.1007/BF00821276. (p. 4)

[Sauer 2006]
Sauer, T. (2006). *Attractor reconstruction. Scholarpedia, 1(10):1727*. URL: http://www.scholarpedia.org/article/Attractor_reconstruction. (p. 26)

[Sauer et al. 1991]
Sauer, T., J. A. Yorke & M. Casdagli (1991). 'Embedology'. In: *Journal of statistical physics* 64.3/4, pp. 579–616. DOI: 10.1007/BF01053745. (pp. 25, 26)

[Scheinkman & LeBaron 1989]
Scheinkman, J. A. & B. LeBaron (1989). 'Nonlinear dynamics and GNP data'. In: *Economic complexity: Chaos, Sunspots, Bubbles, and Nonlinearity. Proceedings of the Fourth International Symposium in Economic Theory and Econometrics*. Ed. by W. A. Barnett, J. Geweke & K. Shell. Cambridge: Elsevier Science, pp. 213–227. (p. 30)

[Schmerr 2014]
Schmerr Jr., L. W. (2014). *Fundamentals of Ultrasonic Phased Arrays*. Cham: Springer Int. Publ. (pp. 10, 11)

[Schmerr & Song 2007]
Schmerr Jr., L. W. & S.-J. Song (2007). *Ultrasonic Nondestructive Evaluation Systems: Models and Measurements*. New York, NY: Springer. (pp. 5, 10)

[Schnars & Henrich 2006]
Schnars, U. & R. Henrich (Sept. 2006). 'Applications of NDT Methods on Composite Structures in Aerospace Industry'. In: CDCM 2006 - Conference on Damage in Composite Materials (Stuttgart, Germany, 18th–19th Sept. 2006). URL: http://www.ndt.net/article/cdcm2006/papers/schnars.pdf. (pp. 2, 12, 13)

[Schnars & Kück 2009a]
Schnars, U. & A. Kück (2009a). In: *4th European-American Workshop on Reliability of NDE*. Application of POD Analysis at Airbus (Berlin, Germany, 24th–26th June 2009). URL: http://www.ndt.net/article/reliability2009/Inhalt/we3a1.pdf. (p. 38)

[Schnars & Kück 2009b]
Schnars, U. & A. Kück (2009b). 'Application of POD Analysis at Airbus'. In: 4th European-American Workshop on Reliability of NDE, Berlin, Germany (Berlin,

Germany, 24th–26th June 2009). URL: http://www.ndt.net/article/reliability2009/Inhalt/we3a1.pdf. (p. 40)

[Smith 2010]
Smith, R. A. (2010). 'Use of 3D ultrasound data sets to map the localised properties of fibre-reinforced composites'. PhD thesis. University of Nottingham. URL: http://eprints.nottingham.ac.uk/31253/. (pp. 8, 14)

[Sokolova & Lapalme 2009]
Sokolova, M. & G. Lapalme (2009). 'A systematic analysis of performance measures for classification tasks'. In: *Information Processing and Management* 45.4, pp. 427–437. DOI: 10.1016/j.ipm.2009.03.002. (p. 38)

[Spokoiny & Dickhaus 2015]
Spokoiny, V. G. & T.-I. Dickhaus (2015). *Basics of Modern Mathematical Statistics*. Berlin: Springer. (pp. 16, 17)

[Stade 2005]
Stade, E. (2005). *Fourier Analysis*. Hoboken, NJ: Wiley. (pp. 18, 19, 21)

[Šutilov 1988]
Šutilov, V. A. (1988). *Fundamental physics of ultrasound*. New York: Gordon and Breach. (p. 5)

[Takens 1981]
Takens, F. (1981). 'Detecting strange attractors in turbulence'. In: ed. by D. Rand & L.-S. Young. Vol. 898. Lecture Notes in Mathematics: Dynamical Systems and Turbulence, Warwick 1980 : Proceedings of a Symposium Held at the University of Warwick 1979/80. Berlin: Springer, pp. 366–381. DOI: 10.1007/BFb0091924. (p. 25)

[Takens 2010]
Takens, F. (2010). 'Chapter 7 - Reconstruction Theory and Nonlinear Time Series Analysis'. In: Broer, H. W., B. Hasselblatt & F. Takens. *Handbook of Dynamical Systems: Volume 3*. Amsterdam: Elsevier Science, pp. 345–377. DOI: 10.1016/S1874-575X(10)00315-2. (pp. 24, 26)

[Teng et al. 2019]
Teng, G., X. Zhou, C. Yang & X. Zeng (2019). 'A Nonlinear Method for Characterizing Discrete Defects in Thick Multilayer Composites'. In: *Applied Sciences* 9.6, p. 1183. URL: https://doi.org/10.3390/app9061183. (p. 47)

[Teschl 2012]
Teschl, G. (2012). *Ordinary Differential Equations and Dynamical Systems*. Vol. 140. Graduate Studies in Mathematics. Providence, RI: American Math. Soc. (p. 24)

[Theiler 1986]
Theiler, J. (Sept. 1986). 'Spurious dimension from correlation algorithms applied to limited time-series data'. In: *Phys.Rev.A* 34.3, pp. 2427–2432. DOI: 10.1103/PhysRevA.34.2427. (p. 31)

[Trulla et al. 1996]
Trulla, L. L., A. Giuliani, J. P. Zbilut & C. L. Webber Jr. (1996). 'Recurrence quantification analysis of the logistic equation with transients'. In: *Physics Letters A* 223.4, pp. 255–260. DOI: 10.1016/S0375-9601(96)00741-4. (p. 30)

[Wang et al. 2019a]
Wang, Z., C. L. Yang, X. J. Zhou & Y. H. Teng (Mar. 2019a). 'Identification of Localized Void Defects in Composite by Recurrence Quantification Analysis of Ultrasonic Backscattered Signal'. In: *Russian Journal of Nondestructive Testing* 55.3, pp. 192–201. DOI: 10.1134/S1061830919030112. (pp. 42, 45, 111)

[Wang et al. 2019b]
Wang, Z., C. Yang, X. Zhou & G. Teng (2019b). 'Identification method for CFRP local pore defects based on recursive quantitative analysis of ultrasonic backscattering signal'. In: *Zhendong yu Chongji/ Journal of Vibration and Shock* 38.21, pp. 229–235. DOI: 10.13465/j.cnki.jvs.2019.21.032. (p. 45)

[Wayland et al. 1993]
Wayland, R., D. Bromley, D. Pickett & A. Passamante (Feb. 1993). 'Recognizing determinism in a time series'. In: *Physical Review Letters* 70.5, pp. 580–582. DOI: 10.1103/PhysRevLett.70.580. (pp. 47, 112)

[Webber 1991]
Webber Jr., C. L. (1991). 'Rhythmogenesis of Deterministic Breathing Patterns'. In: *Rhythms in Physiological Systems: Proceedings of the International Symposium at SchloßElmau, Bavaria, October 22–25, 1990.* Ed. by H. P. Haken Hermann; Koepchen. Vol. 55. Springer Series in Synergetics. Berlin: Springer, pp. 177–191. DOI: 10.1007/978-3-642-76877-4_14. (p. 30)

[Webber 2014]
Webber Jr., C. L. (2014). 'Recapitulation of Recurrence Theory and Practice'. In: CSIB meeting (Rome, Italy, 29th–30th May 2014). DOI: 10.13140/2.1.1898.7842. (p. 30)

[Webber & Marwan 2015]
Webber Jr., C. L. & N. Marwan, eds. (2015). *Recurrence Quantification Analysis: Theory and Best Practices.* Understanding Complex Systems. Cham: Springer. (pp. 114, 117)

[Webber et al. 2016]
Webber Jr., C. L., N. Marwan & I. Cornel, eds. (2016). *Recurrence Plots and Their Quantifications: Expanding Horizons. Proceedings of the 6th International Symposium on Recurrence Plots, Grenoble, France, 17-19 June 2015.* Vol. 180. Springer Proceedings in Physics. Cham: Springer. (p. 22)

[Webber et al. 2009]
Webber Jr., C. L., N. Marwan, A. Facchini & A. Giuliani (2009). 'Simpler methods do it better: Success of Recurrence Quantification Analysis as a general purpose data analysis tool'. In: *Physics Letters A* 373.41, pp. 3753–3756. DOI: 10.1016/j.physleta.2009.08.052. (p. 30)

[Webber & Zbilut 1994]
Webber Jr., C. L. & J. P. Zbilut (Feb. 1994). 'Dynamical assessment of physiological systems and states using recurrence plot strategies'. In: *Journal of applied physiology* 76.2, pp. 965–973. DOI: 10.1152/jappl.1994.76.2.965. (pp. 30, 32, 34)

[Website RP Bibliography 2020]
Nonlinear Dynamics Group, University of Potsdam, Germany (2020). *A Comprehensive Bibliography About RPs, RQA And Their Applications*. URL: `http://recurrence-plot.tk/bibliography.php` (visited on 29th Feb. 2020). (pp. 1, 30)

[Wiener 1956]
Wiener, N. (1956). *I am a Mathematician*. London: Victor Gollance Ltd. (p. iii)

[Witten et al. 2011]
Witten, I. H., E. Frank & M. A. Hall (2011). *Data Mining: Practical Machine Learning Tools and Techniques*. Amsterdam: Elsevier/Morgan Kaufmann. (pp. 36, 37)

[World Conf. of NDT 2019]
The Korean Society for Nondestructive Testing (2019). *20th World Conference of Non-Destructive Testing 2020*. URL: `http://www.wcndt2020.com/`. (p. 10)

[Yang et al. 2019]
Yang, C.-l., W. Feng, C.-l. Bian, X.-j. Zhou & J.-y. Chai (2019). 'Micro defects detection in metallic materials based on recurrence analysis of ultrasonic backscattering signal'. In: *Guangxue Jingmi Gongcheng/ Optics and Precision Engineering* 27.4, pp. 932–944. DOI: `10.3788/ope.20192704.0932`. (p. 47)

[Yap & Rozell 2010]
Yap, H. L. & C. J. Rozell (2010). 'Stable Takens' Embeddings for Linear Dynamical Systems'. In: URL: `http://arxiv.org/abs/1010.5938`. (p. 26)

[Zbilut 1991]
Zbilut, J. P. (1991). 'Power Laws, Transients, Attractors, and Entropy: Possible Implications for Cardiovascular Dynamics'. In: *Rhythms in Physiological Systems: Proceedings of the International Symposium at SchloßElmau, Bavaria, October 22–25, 1990*. Ed. by H. Haken & H. P. Koepchen. Vol. 55. Springer Series in Synergetics. Berlin: Springer, pp. 139–152. DOI: `10.1007/978-3-642-76877-4_11`. (p. 30)

[Zbilut et al. 1990]
Zbilut, J. P., M. Koebbe, H. Loeb & G. Mayer-Kress (1990). 'Use of recurrence plots in the analysis of heart beat intervals'. In: *Computers in Cardiology 1990, Proceedings*. Pp. 263–266. DOI: `10.1109/CIC.1990.144211`. (p. 30)

[Zbilut & Marwan 2008]
Zbilut, J. P. & N. Marwan (2008). 'The Wiener-Khinchin theorem and recurrence quantification'. In: *Physics Letters A* 372.44, pp. 6622–6626. DOI: `10.1016/j.physleta.2008.09.027`. (pp. 43, 45)

[Zbilut & Webber 1992]
Zbilut, J. P. & C. L. Webber Jr. (1992). 'Embeddings and delays as derived from quantification of recurrence plots'. In: *Physics Letters A* 171.3-4, pp. 199–203. DOI: `10.1016/0375-9601(92)90426-M`. (pp. 26, 30, 114)

[Zbilut et al. 2002]
Zbilut, J. P., J.-M. Zaldivar-Comenges & F. Strozzi (2002). 'Recurrence quantification based Liapunov exponents for monitoring divergence in experimental data'.

In: *Physics Letters A* 297.3-4, pp. 173–181. DOI: 10.1016/S0375-9601(02)00436-X. (p. 59)

[Zbilut & Webber 2006]
Zbilut, J. P. & C. L. Webber Jr. (2006). 'Recurrence Quantification Analysis of Nonlinear Dynamical Systems'. In: Wiley Encyclopedia of Biomedical Engineering. Wiley. DOI: 10.1002/9780471740360.ebs1355. (p. 59)

Index