

Document downloaded from:

<http://hdl.handle.net/10251/54484>

This paper must be cited as:

Carrión García, A.; Miralles Ricós, R.; Lara Martínez, G. (2014). Measuring Predictability in Ultrasonic Signals: An Application to Scattering Material Characterization. *Ultrasonics*. 54(7):1904-1911. doi:10.1016/j.ultras.2014.05.008.



The final publication is available at

<http://dx.doi.org/10.1016/j.ultras.2014.05.008>

Copyright Elsevier

Additional Information

Measuring Predictability in Ultrasonic Signals: An Application to Scattering Material Characterization.

Alicia Carrión, Ramón Miralles, Guillermo Lara

*Instituto de Telecomunicaciones y Aplicaciones Multimedia (iTEAM)
Universidad Politécnica de Valencia, Spain.*

Abstract

In this paper, we present a novel and completely different approach to the problem of scattering material characterization: measuring the degree of predictability of the time series. Measuring predictability can provide information of the signal strength of the deterministic component of the time series in relation to the whole time series acquired. This relationship can provide information about coherent reflections in material grains with respect to the rest of incoherent noises that typically appear in non-destructive testing using ultrasonics. This is a non-parametric technique commonly used in chaos theory that does not require making any kind of assumptions about attenuation profiles. In highly scattering media (low SNR), it has been shown theoretically that the degree of predictability allows material characterization. The experimental results obtained in this work with 32 cement probes of 4 different porosities demonstrate the ability of this technique to do classification. It has also been shown that, in this particular application, the measurement of predictability can be used as an indicator of the percentages of porosity of the test samples with great accuracy.

Keywords: Ultrasonic Signal Modality, Ultrasonic NDT, Multiple scattering, Noise, Determinism, Chaos Theory, Higher order statistics

1. Introduction

When scattering materials are subject to ultrasonic non-destructive testing (NDT), the ultrasonic pulse undergoes some variations that are related to the internal grain microstructure of the specimen. Each grain behaves like a scattering center, producing an echo that when superimposed on other echoes coming from other grains can even hide the echoes produced by a possible defect. Similar situations are found in other related fields such as ultrasound B-mode scans (where the grain noise is called speckle) and in radar with clutter [1].

In the literature, a wide range of solutions has been proposed to enhance the detection of small cracks or defects and to reduce (or even eliminate) the effect discussed in each of the above situations. Some of these solutions are signal averaging, auto- and cross-correlation, matched filtering, frequency spectrum analysis [2], spectral correlation [3], and wavelet transformations [4]. The aforementioned analyses discard the information encoded in the grain noise; however, this information can be used to recognize potential differences among materials, tissues, or surfaces. This approach has been employed to characterize materials by extracting temporal signal statistics [1], the resonance frequency [5], and even the penetration depth [6]. Our work continues with this line of thought and proposes a new approach that attempts to extract information about the nature of the signal, thereby characterizing the signal modality. Signal modality characterization is a key concept of a multidisciplinary research topic that includes different concepts such as signal linearity, stationarity, and stochasticity [7]. The applications of signal modality characterization are becoming more and more relevant in signal processing and machine learning. Specifically, signal modality has been tested in the analysis of EEG data and weather information, but it has never been used in ultrasonic testing.

In order to apply signal modality algorithms to ultrasonic signals, a brief review of the typical noise sources that appear in ultrasonic inspection must be made. The vast majority of noise sources can be approximated by random processes (thermal noise, quantification, etc.), but some of the sources have a distinctly deterministic pattern (vibrations, grain noise, or speckle). Fig. 1 illustrates how coherent and incoherent noises are combined as a function of the

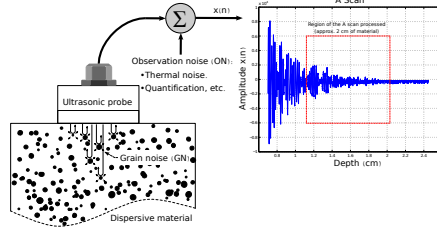


Figure 1: Pulse-echo model of an ultrasonic inspection of scattering materials and an example of the resulting signals.

inner material structure in a typical ultrasonic inspection of scattering materials. Measuring the predictability of the resulting signals can give information about this blend and, thus, about the material characteristics.

The remainder of this work is organized as follows. Section II proposes an alternative index based on higher-statistics which avoids some of the drawbacks when measuring predictability in ultrasonic signals. Section III describes and mathematically formulates three algorithms that are typically employed in chaos theory when studying determinism. These will be used for comparison with the proposed predictability index. All these algorithms are tested in Section IV with a theoretical ultrasound model. In Section V, we focus on a real ultrasonic application where different kinds of scattering materials are classified by measuring their degree of predictability. Finally, the conclusions are summarized.

2. Analysing the predictability of ultrasonic scattered signals using higher order statistics

Ultrasonic scattered signals are not simply deterministic or stochastic, but rather a combination of both [8]. Predictability can be viewed as the signal strength of the deterministic component of the time series with respect to the whole time series. A feasible alternative for characterizing the predictability of the signal is to use the correlation between points. Second-order correlation (autocorrelation) cannot give enough information about higher order interactions among parameters that govern the model that leads to the time series. Due to the complex physical interactions taking place in the ultrasonic inspection of a highly scattering material, we propose using higher order statistics to measure the predictability of the time series.

Let us model the ultrasonic register as a discrete stationary stochastic process $\{\tilde{x}(n)\}$. We can define what will be called a displacement sequence,

$$\{\tilde{S}_m(n)\} = \{\tilde{x}(n)\} - \{\tilde{x}(n+m)\} = \{\tilde{x}(n)\} * h(n) \quad (1)$$

where m is the time delay, $*$ denotes the discrete convolution of the stochastic process $\{\tilde{x}(n)\}$ with the linear time-invariant system $h(n) = \delta(n) - \delta(n+m)$, and $\delta(n)$ is the discrete Dirac delta function.

The statistical properties of $\{\tilde{S}_m(n)\}$ can be analyzed as a function of the statistics of the input process $\{\tilde{x}(n)\}$ and the impulse response of the filter $h(n)$, previously defined. We can obtain the 4th order cumulant of $\{\tilde{S}_m(n)\}$ ($c_4^{S_m}(k_1, k_2, k_3) = E[S_m(n)S_m(n+k_1)S_m(n+k_2)S_m(n+k_3)]$) using filter relationships of a colored process [9]:

$$c_4^{S_m}(k_1, k_2, k_3) = \sum_{m_1} \sum_{m_2} \sum_{m_3} c_4^x(k_1 - m_1, k_2 - m_2, k_3 - m_3) \cdot c_4^h(m_1, m_2, m_3) \quad (2)$$

where

$$c_4^h(m_1, m_2, m_3) = \sum_l h(l)h(l+m_1)h(l+m_2)h(l+m_3) \quad (3)$$

In Equation (2), k_1, k_2 and k_3 are the time lags where the 4th order cumulant of the process $\{\tilde{S}_m(n)\}$ is computed and m_1, m_2, m_3 are also time lags varying in the region of support of the 4th order cumulants. If Equation (2) is particularized for $k_1 = k_2 = k_3 = 0$, the kurtosis ($\gamma_4^{S_m}$) of the displacement sequence is obtained (see Equation (4)).

$$c_4^{S_m}(0, 0, 0) = \gamma_4^{S_m} = \sum_{m_1} \sum_{m_2} \sum_{m_3} c_4^x(-m_1, -m_2, -m_3) \cdot c_4^h(m_1, m_2, m_3) \quad (4)$$

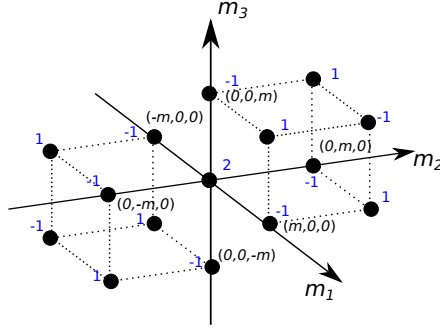


Figure 2: Graphical representation of the 4th order cumulants of $h(n)$. The black dots indicate unique values where $c_4^h(m_1, m_2, m_3)$ is non zero (the blue numbers next to the dots indicate the amplitude of the cumulants at these lags).

The deterministic 4th order cumulants $c_4^h(m_1, m_2, m_3)$ are only non zero for some specific lags as illustrated in Fig. 2. Using this property and also taking into account the symmetry properties of the 4th order cumulants, Equation (4) can be simplified into Equation (5).

$$\begin{aligned} \gamma_4^{S_m} &= 2 \cdot \left(c_4^x(0, 0, 0) - c_4^x(m, 0, 0) - c_4^x(0, 0, m) - c_4^x(0, m, 0) + \right. \\ &\quad \left. c_4^x(m, m, 0) + c_4^x(m, 0, m) + c_4^x(0, m, m) - c_4^x(m, m, m) \right) \\ &= 2 \cdot \left(c_4^x(0, 0, 0) - c_4^x(m, m, m) + 3 \cdot c_4^x(m, m, 0) - 3 \cdot c_4^x(m, 0, 0) \right) \end{aligned} \quad (5)$$

Thus, evaluating the 4th order central moment (kurtosis) of the displacement sequence $\{\tilde{S}_m(n)\}$ provides a simple way of measuring how the 4th order cumulants of $\{\tilde{x}(n)\}$ change with the time lag m . This time lag defines the side of the cube whose vertexes are the points where $c_4^x(k_1, k_2, k_3)$ is evaluated. In order to obtain a clear indication of how the higher order correlations evolve as the time lag increases, the vertex of this cube must increase. In this work, we propose using an averaged kurtosis of the displacement sequence obtained for different time lags m as an estimator of the predictability:

$$DAKDS = \frac{1}{M} \sum_{m=1}^M \gamma_4^{S_m} \quad (6)$$

where M is the maximum lag introduced whose choice depends on the sample frequency and the size of the observation window; however, this choice is not critical. DAKDS are the acronyms of Determinism based on the Average Kurtosis of the Displacement Sequence. It is worth noting that the 3rd order statistic moment (skewness) has not been applied since its equivalent expression to Equation (5) would be null for all kind of signals, whether they are deterministic or stochastic. The 4th order cumulant is the lowest statistic which allows predictability to be measured taking into account higher order interactions.

The conventional way of measuring the kurtosis consists of assuming ergodicity and computing the time averages of $\frac{E[(S_m(n) - \mu)^4]}{\sigma^4} - 3$, where μ is its mean and σ is its standard deviation of $\{\tilde{S}_m(n)\}$. Kurtosis can be interpreted as a measure of the dispersion of a distribution around the values $\mu \pm \sigma$, which can be large when the probability mass is concentrated either near the mean or in the tails of the distributions. To solve this, a robust alternative to estimate the kurtosis was proposed in [10] and will be used here:

$$RK(S_m(n)) = \frac{(E_7 - E_5) + E_3 - E_1}{E_6 - E_2} - 1.23 \quad (7)$$

where E_j is the j^{th} octile of the distribution of a realization $\{\tilde{S}_m(n)\}$. The estimator RK is consistent and unbiased. A detailed analysis of this estimator for different types of distributions is done in [11]. Additionally, RK it is not even

influenced by the presence of outliers since it is based on octiles, which converge to the true values reasonably quickly. Later, in Section IV, we present a simulation showing the variance evolution of the estimator for an ultrasonic model that simulates a scattering material.

3. Review of typical algorithms that are commonly employed to measure predictability

Typically, the study of a deterministic signal has relied on the concept of “phase space”, which is a vector space that collects all the possible system states that are useful for determining the future evolution of the signal. For a time series $x(n)$, the “phase space” would be defined by

$$\begin{aligned}\vec{X}_n &= [x(n), x(n+L), \dots, x(n+(E-1)\cdot L)], \\ n &= 1, \dots, N-(E-1)\cdot L\end{aligned}\quad (8)$$

where N is the total number of points, L is the delay between points of the time series, and E is the embedding dimension ([12], [13]).

The time series $x(n)$ is said to be deterministic if its “phase space” can be modeled as a continuous function. This approach leads to the appropriate selection of the parameters L and E , which is crucial to be able to distinguish between deterministic and random signals. In order to choose the suitable parameters, some authors propose using correlation dimension, Lyapunov exponents, mutual information theory, or different entropy quantities; however, a valid conclusion has not been reached for all cases.

3.1. Modified Kaplan-Glass Method

Wayland et al. [14] developed a method that detects determinism by testing the continuity in the reconstructed “phase space”. This algorithm finds the nearest neighbors and computes the average distance between points. Let \vec{X}_{n_0} be a fixed vector that is computed as in Equation (8), and let $\vec{X}_{n_1}, \vec{X}_{n_2}, \dots, \vec{X}_{n_K}$ be its K nearest neighbors. If the time series under study is deterministic, the displacement vector defined as

$$\vec{V}_{n_k} = \vec{X}_{n_{k+1}} - \vec{X}_{n_k} \quad (9)$$

for each one of the K neighbors and its consecutive vectors is expected to be almost the same.

To quantify this similarity, we compute the average of the translation vectors \vec{V}_{n_k}

$$\langle \vec{V} \rangle = \frac{1}{K+1} \sum_{k=0}^K \vec{V}_{n_k} \quad (10)$$

The translation error that is used as the predictability index is computed as shown in Equation (11)

$$DMKG_{n_0} = \frac{1}{K+1} \sum_{k=0}^K \frac{\|\vec{V}_{n_k} - \langle \vec{V} \rangle\|^2}{\|\langle \vec{V} \rangle\|^2} \quad (11)$$

where $\|\cdot\|$ denotes the Euclidean norm. This calculation is extended to a subset $\vec{X}_p \subset \vec{X}_n$ of N_{res} random centers of the phase space (a quarter of the total number of points is an accurate solution). The median of these values will be called $DMKG$ and it provides a robust measure of the continuity of the “phase space”.

$$DMKG = \text{Median}_{\vec{X}_{n_0} \in \vec{X}_p} [DMKG_{n_0}] \quad (12)$$

With this algorithm, we are able to quantify the spread of the phase space based on the vector translations. If the time series is predictable, \vec{V}_{n_k} will be nearly equal and the continuity index using the Modified Kaplan-Glass Method ($DMKG$) will be very small. Otherwise, \vec{V}_{n_k} will be different and the resulting continuity index will be larger.

3.2. Recurrence Plots

Eckmann et al. [15] introduced a tool called Recurrence Plots (RP), which can visualize the recurrence of states \vec{X}_n in a phase space. One of the main advantages of the RP is that they allow the E -dimensional phase-space trajectory to be investigated through a binary two-dimensional representation of the recurrences of the states. Additionally, the quantification of the number and duration of the recurrences allows us to study the degree of predictability [16] as well as some other characteristics of the underlying dynamic systems: laminar phases [17], unstable periodic orbits [18], etc.

Among the different variations of computing the RP, the most common way is using Equation (13),

$$R_{n_1, n_2} = \Theta(\varepsilon - \|\vec{X}_{n_1} - \vec{X}_{n_2}\|_\infty), \quad n_1, n_2 = 1, \dots, N_s \quad (13)$$

where N_s is the number of considered states \vec{X}_n , ε is a threshold distance, $\|\cdot\|_\infty$ is the maximum norm, and $\Theta(\cdot)$ is the Heaviside function.

A diagonal line in the RP appears when a segment of the trajectory runs parallel to another segment. The length of this diagonal line is determined by the duration that the two trajectory segments have a similar evolution. As a result, the presence of diagonal lines that run parallel to the mean diagonal (line of identity) indicates that the evolution of states is similar at different times and that the process could be deterministic. The percentage of recurrence points that form diagonal lines can be used to measure predictability. This percentage can be computed as follows:

$$DRP = \frac{\sum_{j=j_{min}}^{N_s} j \cdot P(j)}{\sum_{n_1, n_2=1}^{N_s} R_{n_1, n_2}} \quad (14)$$

where $P(j)$ is the number of diagonal lines of length j and j_{min} is the minimum number of points to be considered as a diagonal line (in this work, $j_{min} = 2$ is used).

3.3. The Method of Jeong et al.

Jeong et al. [19] also computed the “phase space”, proposing a new predictability measure that is based on the angles between consecutive vectors, eliminating the need to search for neighboring points.

Let

$$\vec{V}_n = \vec{X}_{n+1} - \vec{X}_n \quad (15)$$

be the tangent between consecutive vectors in the phase space; the angles are computed as

$$\cos(\theta_n) = \frac{\vec{V}_{n+1} \cdot \vec{V}_n}{\|\vec{V}_{n+1}\| \|\vec{V}_n\|} \quad (16)$$

where $\vec{V}_{n+1} \cdot \vec{V}_n$ denotes the scalar product of vectors \vec{V}_{n+1} and \vec{V}_n . The predictability index is obtained by averaging the computed angles in the whole phase space:

$$DJM = \frac{1}{S} \sum_{n=1}^S \cos(\theta_n) \quad (17)$$

where $S = (N - (E - 1) \cdot L) - 2$.

If the time series is deterministic, the trajectory in the state space is smooth and most of the cosine values are close to 1 (or larger than 0.7). Otherwise, if the time series is stochastic, cosine values are near 0 (or smaller than 0.3). The intermediate case (0.3-0.7) is known to sometimes arise from time series having different degrees of predictability.

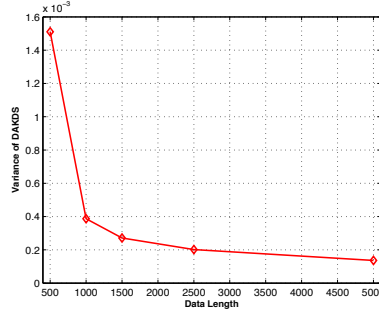


Figure 3: Variance of DAKDS for the ultrasonic model computed for five different data lengths. 1000 Monte Carlo realizations are used for each data length. The ultrasonic model parameters are: carrier frequency (w_0) 10 MHz, attenuation 10 dB/cm, percentage of scatterers 40% and K -Type amplitude distributed reflections.

4. Application to the ultrasound model

This section analyzes the results achieved by applying the four methods described above on simulated ultrasound signals. The simulated signals were obtained using the model presented in [1] and reproduced in Equation (18). This equation simulates a backscattered ultrasonic signal at time t for a given transducer position \mathbf{z} .

$$x(t, \mathbf{z}) = \sum_{i=1}^{N_e(\mathbf{z})} A_i(\mathbf{z}) \cdot p(t - \tau_i(\mathbf{z})) \cdot \cos(w_0(t - \tau_i(\mathbf{z}))) + \{\tilde{n}(t)\} \quad (18)$$

where

$$p(t - \tau_i(\mathbf{z})) = p_0(t - \tau_i(\mathbf{z})) \cdot e^{-\alpha \cdot \tau_i(\mathbf{z})} \quad (19)$$

Equation (18) computes the backscattered signal (A-scan) as the sum of all contributions from the $N_e(\mathbf{z})$ scatters characterized by their reflectivity $A_i(\mathbf{z})$ placed randomly ($\tau_i(\mathbf{z})$) inside the sample, with $p_0(t)$ being the pulse envelope at the origin, and w_0 being the carrier frequency. Note that, due to attenuation α (dB/cm), the amplitude of $p(t - \tau_i(\mathbf{z}))$ will decrease exponentially as the ultrasonic pulse propagates deeply into the material (Equation (19)). The stochastic process $\{\tilde{n}(t)\}$ models the Observation Noise (ON), and it is incorporated to model the instrumentation noise or other effects that are not related to material properties (Fig. 1). Thus, the resulting simulated signals are the sum of a material dependent grain noise (GN) plus a random ON. Due to the different grain echoes and the attenuation with depth, the registered A-scans show characteristic waveforms whose amplitude decays until GN collapses below the ON level. As a result, the signals have a decreasing GOR (Grain-to-Observation-Noise-Ratio), which is characteristic of the sample being tested. When two specimens with different amounts of grain are compared, the nature of the reflected signal will be different. In the case of fewer grain echoes, the reflections or echoes will be similar to the incident signal. However, in the presence of more scatters, the energy will be dispersed and there will be a large number of contributions which will give rise to different levels of predictability. In order to prevent predictability measures from being the result of attenuation and signal amplitude changes, the signal resulting from the model have been normalized in amplitude without varying the rest of the properties.

Before using the proposed DAKDS parameter, the variance of the estimator should be analyzed. This analysis was conducted for the ultrasonic backscattered signal of Equation (18). The parameters employed in the simulation were: $w_0 = 2 \cdot \pi \cdot 10$ MHz, attenuation 10 dB/cm, percentage of scatterers 40%, and K -Type amplitude distributed reflections. The estimated DAKDS was computed as described in Equations (6) and (7) for 1000 Monte Carlo realizations for each data length using M equal to 400 as maximum time lag. Fig. 3 shows that the variance of the DAKDS estimator decreases as the data length increases.

The four methods described in Section 3 have also been evaluated using the above model with the same parameters. In this case, the data length is 1500 points, and 50 Monte Carlo realizations are used for each attenuation level and each model. We present the Fig. 4 to analyze the behaviour of the predictability indexes against different attenuation values. Fig. 4a shows that $DMKG$ (Eq. 12) and DJM (Eq. 17) do not allow the level of signal attenuation to be

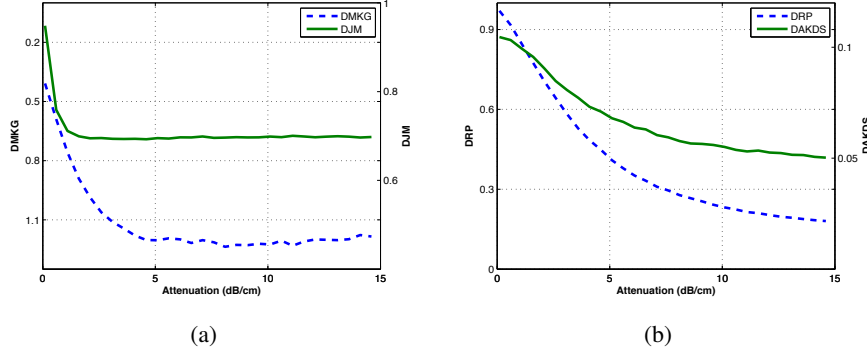


Figure 4: Degree of predictability as a function of the attenuation values (dB/cm) for the ultrasonic model. The data length is 1500 points, and 50 Monte Carlo realizations are used for each attenuation level and each model. The ultrasonic model parameters are: carrier frequency w_0 10 MHz, reflectivity 40% and the statistic of the reflectivity is K -Type. a) DMKG (left) and RJM (right), b) DRP (left) and DAKDS (right).

identified from predictability since the curves reach constant values above a certain level of attenuation (1 dB/cm and 4.5 dB/cm, respectively). When attenuation increases, the maximum GOR decreases; therefore, both methods fail to provide a characteristic measure that allows predictability identification for high noise levels. However, Fig. 4b shows the resulting curves for DRP (Eq. 14) and $DAKDS$ (Eq. 6) that allow us to determine the value of the characteristic attenuation based on the predictability measured. Both curves show that despite high attenuation levels, both methods are able to distinguish signals even if they are random. This results are justified because of the binarization process that the RP plots undergo through the Heaviside function (DRP parameter) and because of the insensitivity of the higher order statistics to Gaussian noise ($DAKDS$ parameter).

5. Application to the characterization of cement pastes

We applied the proposed parameters in a real application to measure the degree of porosity of different cement paste probes. This is an important problem because cement paste is the main component of mortar and concrete [20]. The vulnerability (and, thus, the durability) of these construction materials to external agents is very dependent on the porosity of the cement [21].

We consider cement paste to be composed of a hydrated gel matrix with pore cavities that are incrustated within and not occupied by gel. Porosity is dependent upon the water/cement ratio (w/c) and cement composition. In this study, we chose two types of cement with the same mineral composition and identical percentage of added limestone (about 17% of substitution), but with a different specific surface of grain. Therefore, we had two different mechanical compression resistances, namely 32.5 and 42.5 MPa. We used the notation CEM 32.5 and CEM 42.5 and considered two different w/c ratios (0.4 and 0.5), so that there were four cement paste types.

Table 1: Porosity mean values corresponding to the four types of test samples.

Type of cement (MPa)	Water/Cement ratio	
	0.4	0.5
32.5	30.73 %	37.63 %
42.5	27.36 %	33.65 %

We built 32 test samples (8 for each type of cement paste), which were prisms of a size $16 \times 4 \times 4$ (cm). Four of them were used for measuring (in a destructive manner) the porosity of each type of cement paste following the method described in [22]. Table 1 shows the porosity mean value obtained (the average of the four test samples) for each type of cement paste. The other four test samples were used to perform ultrasonic non-destructive testing. In each sample, we collected 20 A-scans of backscattering noise along the two sides of the prism that were uniformly spaced. One of

this A-Scan is shown in Fig. 1. The ultrasonic equipment and the most significant acquisition data were: ultrasound pulser-receiver card (IPR-100, Physical Acoustics, Princeton, USA); transducer (Krautkramer KBA-10 MHz, Huerth, Germany); digitalisation (Oscilloscope Tektronix TDS-3012, Wilsonville, USA); and sampling frequency (250 MHz).

The selection of the 10 MHz transducer was justified by the need to obtain enough grain noise in the collected records. Using the mercury intrusion method, we verified that, at 10 MHz, $E[D]/\lambda$ varied between 0.026 and 0.071 for the different types of cement paste considered, where λ is the wavelength and D is the mean grain size. Therefore, we are working in the Rayleigh region [23], and we have a moderate, but sufficient, amount of grain noise. When the frequency increases, it produces an attenuation that is too large; when the frequency decreases, it reduces the level of GN (the ultrasound pulse propagated better, but there were no echoes from the inner microstructure). Having enough information about the inner composition of the probes is paramount in classifying the probes as a function of the predictability measured on the echo pulses.

As mentioned in Section IV, the degree of predictability must be proportional to the number of pores in the material. Fig. 5 compares the “phase space” of ultrasound signals measured in two different specimens. It must be highlighted that the trajectories followed similar paths, but the plot on the left is smoother than one on the right. If there is a higher number of pores, the signal is more irregular, which is a clear sign of randomness.

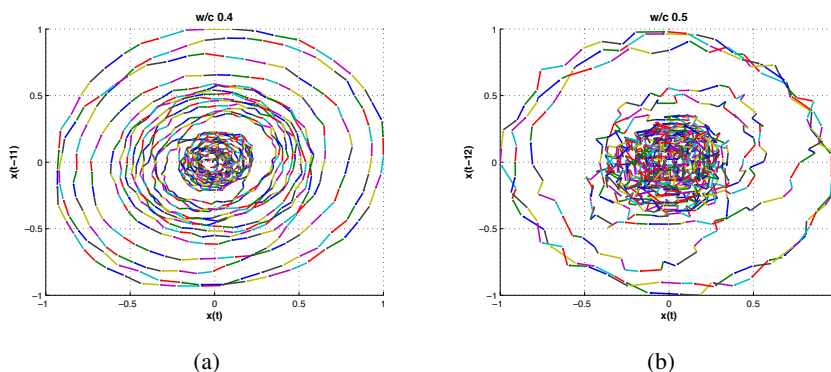


Figure 5: a) “Phase space” of an ultrasound signal measured in a w/c 0.4 sample. b) “Phase space” of an ultrasound signal measured in a w/c 0.5 sample. $E=2$ and $M=11$.

Accordingly, we must be able to estimate the percentage of pores in each of the resulting samples of different materials or different manufacturing processes by applying the reviewed methods. For this purpose, the four algorithms were applied on the 20 signals measured in each specimen. For each record, we considered an interval of 1300 points, removing the initial transient and long end tails, which had been subjected to a process of normalization in amplitude. Those points covered a depth distance of 2 cm (an average propagation velocity of 3500 m/s was used to convert time scale into depth scale).

For each method, the parameters introduced in the mathematical formulation were configured as follows. For the reviewed methods (*DMKG*, *DRP*, and *DJM*), the embedding dimension E was equal to 2 and the time lag L was the first null of the non-bias autocorrelation function of the signal [24]. Moreover, in *DMKG*, $K = 5$ neighbors and $N_{res} = 100$. In *DRP*, ε was equal to 60% of the mean Euclidean distance of the phase space vectors [25]. From the above, it can be deduced that the reviewed methods require a greater number of input parameters than the proposed method which eliminates having to calculate the autocorrelation, and the result was weighted to a maximum lag M equal to 400 points.

Fig. 6 shows the boxplots obtained for each test sample and each method. Note that we tested 4 specimens of each class indicated on the x-axis. The red marks represent the median value of the 20 measures, and the blue box boundaries are the 25th and 75th percentiles. At a glance, it can be observed that the four methods group each specimen sample so that a cluster can be identified for each type of material. If we compare the predictability values obtained and the results using the theoretical model (Fig. 4), it can be concluded that all recorded signals had low attenuation levels (below 5 dB/cm), and, therefore, excellent GOR relations. Then, all four methods were able to identify the different materials in this interval. Moreover, the trend coincided in all of the methods and indicated that the specimens with w/c ratio 0.4 were classified with greater predictability values; however, the specimens of w/c

ratio equal to 0.5 were more random, which is correlated with the higher percentage of porosity presented in Table 1. Note that the values of predictability obtained for the proposed index (the DAKDS parameter) are three orders higher than the variance obtained in Fig. 3 for the data length equal to 1300 points.

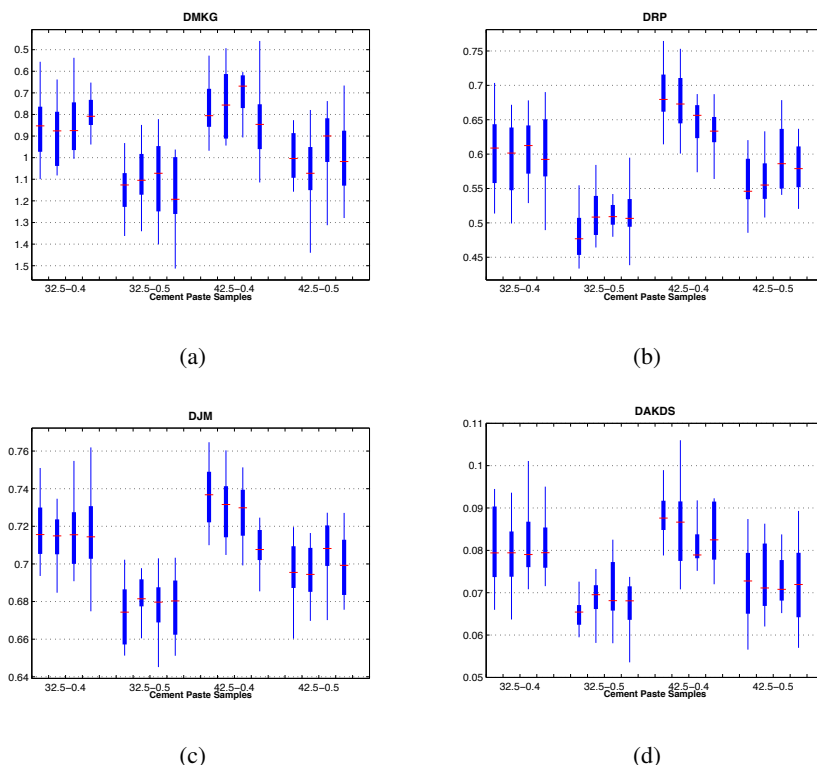


Figure 6: Plot of the predictability measures versus the different cement paste samples. The red marks represent the median value and the blue box boundaries are the 25th and the 75th percentiles. a) DMKG, b) DRP, c) DJM, d) DAKDS.

In order to further analyze the relation between predictability and the percentage of pores, the results were adjusted to a linear equation whose representation is plotted in Fig. 7. The plot illustrates the linear approximation estimated for each method as well as the median predictability index obtained for each specimen of each class (blue marks). Furthermore, the median value for each class (green diamonds) is also shown. Table 2 highlights the excellent linear fitting obtained in all cases regarding the regression error. To compute this quality measure, the resulting linear approximations and the measures were normalized to the interval between 0 and 1. The proposed robust DAKDS algorithm obtained results that are comparable to other published methods.

Table 2: Linear regression error corresponding to the four algorithms.

	DMKG	DRP	DJM	DAKDS
Error	0.3559	0.1378	0.2978	0.2875

6. Conclusions and Future Lines of Research

In this paper, we have considered the general problem of ultrasonic characterization of scattering materials from a viewpoint that is different from those seen to date. We propose the study of the nature of the signal through predictability and using this parameter as a source of information of the inner material structure.

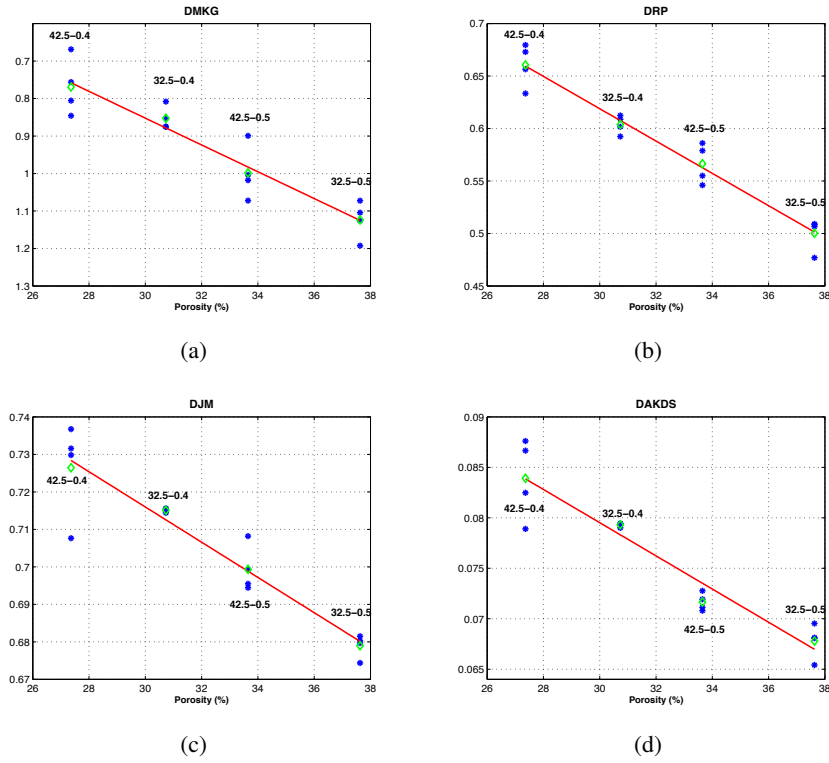


Figure 7: Plot of the predictability measures versus the porosity for each type of cement paste. The blue marks represent the median value for each sample, the green marks represent the median value for each class, and the red lines represent the linear approximation computed for each method. a)DMKG, b) DRP, c) DJM, d) DAKDS.

We have reviewed some of the methods that are available in the literature to measure predictability, and we have also proposed an alternative algorithm based on higher order statistics that is easier to implement and that achieves equivalent results. The predictability index has been used to evaluate a theoretical ultrasound model as well as to estimate the porosity of real cement probes. The theoretical results have shown that the proposed algorithm is more robust against low SNR signals than the rest of the methods and is also less parameter-dependent. When measuring real signals, determinism is strongly related to the degree of ultrasonic scattering; in fact, it has been shown that a linear relation can be an excellent approximation for calculating the porosity of the cement paste probes based on predictability. This new approach can complement or even be used as an alternative to other techniques when extracting information about the internal structure of scattering materials.

As future lines of research, the proposed algorithm could be used in other applications of characterization of scattering materials using ultrasound. This technique should also give good results in mortar or even in concrete as long as the inspection frequency is lowered to maintain the $E[D]/\lambda$ relationship.

Acknowledgements

This work has been supported by the Spanish Administration under grant TEC2011-23403. We also wish to acknowledge Dr. Jorge Gosálbez for providing the ultrasound data of the cement paste probes.

References

- [1] L. Vergara-Dominguez, J. M. Páez-Borrillo, Backscattering grain noise modelling in ultrasonic non-destructive testing, *Waves in Random Media* 1 (1991) 81–92.
- [2] X. Li, N. Bilgutay, R. Murthy, Spectral histogram using the minimization algorithm-theory and applications to flaw detection, *Ultrasonics, Ferroelectrics and Frequency Control*, IEEE Transactions on 39 (1992) 279–284.

- [3] K. Donohue, H. Y. Cheah, Spectral correlation filters for flaw detection, *Ultrasonics Symposium, 1995. Proceedings.*, 1995 IEEE 1 (1995) 725–728 vol.1.
- [4] A. Abbate, J. Koay, J. Frankel, S. Schroeder, P. Das, Signal detection and noise suppression using a wavelet transform signal processor: application to ultrasonic flaw detection, *Ultrasonics, Ferroelectrics and Frequency Control*, IEEE Transactions on 44 (1997) 14–26.
- [5] L. Vergara, J. Gosálbez, J. Fuente, R. Miralles, I. Bosch, Measurement of cement porosity by centroid frequency profiles of ultrasonic grain noise, *Signal Processing* 84 (2004) 2315 – 2324.
- [6] L. Vergara, J. Fuente, J. Gosálbez, R. Miralles, I. Bosch, Processing of ultrasonic grain noise signals for the estimation of depth- and frequency-dependent attenuation, *Measurement Science and Technology* 14 (2003) 1018.
- [7] B. Jelfs, S. Javidi, P. Vayanos, D. Mandic, Characterisation of signal modality: Exploiting signal nonlinearity in machine learning and signal processing, *Journal of Signal Processing Systems* 61 (2010) 105–115.
- [8] C. Courtney, S. Neild, P. Wilcox, B. Drinkwater, Application of the bispectrum for detection of small nonlinearities excited sinusoidally, *Journal of Sound and Vibration* 329 (2010) 4279 – 4293.
- [9] J. Mendel, Tutorial on higher-order statistics (spectra) in signal processing and system theory: theoretical results and some applications, *Proceedings of the IEEE* 79 (1991) 278–305.
- [10] J. J. A. Moors, A quantile alternative for kurtosis, *Journal of the Royal Statistical Society, Series D (The Statistician)* 37 (1988) 25–32.
- [11] T.-H. Kim, H. White, On more robust estimation of skewness and kurtosis, *Finance Research Letters* 1 (2004) 56 – 73.
- [12] N. H. Packard, J. Crutchfield, J. D. Farmer, R. S. Shaw, Geometry from a time series, *Phys. Rev. Lett.* 45 (1980) 52–56.
- [13] D. A. Rand, L. S. Young, *Dynamical Systems and Turbulence*, Springer, Berlin, 1981.
- [14] R. Wayland, D. Bromley, D. Pickett, A. Passamante, Recognizing determinism in a time series, *Phys. Rev. Lett.* 70 (1993) 580–582.
- [15] J.-P. Eckmann, S. O. Kamphorst, D. Ruelle, Recurrence plots of dynamical systems, *Europhysics Letters (EPL)* 4 (1987) 973–977.
- [16] L. Trulla, A. Giuliani, J. Zbilut, C. Webber Jr., Recurrence quantification analysis of the logistic equation with transients, *Physics Letters A* 223 (1996) 255 – 260.
- [17] N. Marwan, N. Wessel, U. Meyerfeldt, A. Schirdewan, J. Kurths, Recurrence-plot-based measures of complexity and their application to heart-rate-variability data, *Phys. Rev. E* 66 (2002) 026702.
- [18] D. P. Lathrop, E. J. Kostelich, Characterization of an experimental strange attractor by periodic orbits, *Physical Review A* 40 (1989).
- [19] J. Jeong, J. C. Gore, B. S. Peterson, A method for determinism in short time series, and its application to stationary EEG, *IEEE Transactions on Biomedical Engineering* 49 (2002) 1374–1379.
- [20] M. Hernández, M. Izquierdo, A. Ibez, J. Anaya, L. Ullate, Porosity estimation of concrete by ultrasonic {NDE}, *Ultrasonics* 38 (2000) 531 – 533.
- [21] S. Alekseev, *Durability of reinforced concrete in aggressive media*, A.A. Balkema, Rotterdam, 1991.
- [22] R. A. Olson, C. M. Neubauer, H. M. Jennings, Damage to the pore structure of hardened portland cement paste by mercury intrusion, *Journal of the American Ceramic Society* 80 (1997) 2454–2458.
- [23] J. Saniie, T. Wang, N. Bilgutay, Analysis of homomorphic processing for ultrasonic grain signal characterization, *IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control* 36 (1989) 365–375.
- [24] H. Kantz, T. Schreiber, *Nonlinear Time Series Analysis*, Cambridge University Press, 2004.
- [25] N. Marwan, How to avoid potential pitfalls in recurrence plot based data analysis, *I. J. Bifurcation and Chaos* 21 (2011) 1003–1017.