Non-intrusive microwave technique for direct detection of concrete compressive strength monitoring by multivariate modeling

Non-Intrusive Microwave Technique for Direct Detection of Concrete Compressive Strength Monitoring by Multivariate Modeling

Leonardo Franceschelli[1], Eleonora Iaccheri[2,3], Elisa Franzoni[4], Annachiara Berardinelli[5,6], Luigi Ragni[2,3], Claudio Mazzotti[4] and Marco Tartagni[1,]

[1]School Department of Electrical, Electronic and Information Engineering "Guglielmo Marconi"—University of Bologna,Via dell'Università, 50, 47521 Cesena, Italy

[2]Department of Agricultural and Food Sciences, Alma Mater Studiorum, University of Bologna,Piazza Goidanich 60, 47521 Cesena, Italy

[3]Interdepartmental Center for Industrial Agri-Food Research, University of Bologna, Via Q. Bucci 336,47521 Cesena, Italy

[4]Department of Civil, Chemical, Environmental and Materials Engineering (DICAM), University of Bologna, Via Terracini 28, 40131 Bologna, Italy

[5]Department of Industrial Engineering, University of Trento, Via Sommarive, 9, 38123 Povo, Italy

[6]Centre Agriculture Food Environment, University of Trento, Via E. Mach, 1, 38010 S. Michele all'Adige, Italy

Corresponding author:
Leonardo Franceschelli, School Department of Electrical, Electronic and Information Engineering "Guglielmo Marconi"—University of Bologna, Via dell'Università, 50, 47521 Cesena, Italy
Email: leonar.franceschell2@unibo.it

Abstract

This paper investigates using microwaves from 1.5 to 14 GHz for monitoring concrete compressive strength in real-time to implement a non-intrusive sensor. The approach is based on acquiring reflected microwave spectra by the concrete irradiated through a rectangular cavity antenna connected to a Vector Network Analyzer (VNA) and investigating their relationship with strength. Since the information is hidden in the spectrum, we used the multivariate statistical analysis technique to create both classification and regression models. The acquired Voltage Standing Wave Ratio (VSWR) spectra were statistically inferred with compressive strength from destructive tests. The experimentation was performed on 16 concrete samples, lasted 28 days, and the acquired data consisted in 409 spectra where the reflected spectra are correlated with the compressive strength during curing time. The inference processes show high predictability, with an *F1* score from 0.91 to 0.99 for classification and a cross-validation value of the parameter R^2 of 0.997.

Keywords

concrete, machine learning, multivariate statistical analysis, microwave, non-invasive, sensor.

1. Introduction

Concrete is the most used material in the civil infrastructure and construction industry. It is composed of a mix of cement powder, sand (fine aggregates), gravel (coarse aggregates), and water. The ratio of these components influences the mechanical properties of concrete. In particular, the water/cement (w/c) ratio is critical for the strength and durability of concrete due to the role of water in the hydration process and the microstructure evolution [1]. So, it is crucial to monitor the conditions and characteristics of concrete during the hydration process, especially in the first hours. Nowadays, several tests are used as the gold standard for the evaluation of hardened concrete properties, like the water absorption test (ASTM C: 642-81), rapid chloride ion penetration test (ASTM C1202), impact strength test (ACI 544.2R-89), and compressive strength test (ASTM C39/C39M). However, even if they allow a satisfactory evaluation during construction, all these methods are time-consuming and expensive. Moreover, they require an invasive approach and the creation of additional samples [2] with different mechanical characteristics for the on-site batch due to other manufacturing processes and materials affecting the results [3], [4].

For these reasons, non-destructive techniques (NDT) have been recently studied with several different technologies: acoustic waves [5], [6], ultrasonic wave propagation [7], Ground Penetrating Radar (GPR) wave attenuation [8], [9], piezoelectric materials [10], [11], gamma scattering [12], electrical resistivity [13], [14], and many others. Among them, a particular interest is aroused by microwaves, which allow penetrating the

material and gathering information on the internal water content. The microwave spectra are highly influenced by the hydration of concrete and the consequent changes in the water content because of its significant influence on the dielectric properties. Different techniques were used to determine these dielectric properties of concrete and their changes: the first tests were performed by Bhargava and Lundberg [15] in 1972 with a microwave resonant cavity, obtaining a linear relationship between the output of the instrument and the moisture content of concrete. Three years later, Wittman and Schlude [16] studied microwave absorption using concrete disk samples with a free wave technique. The results showed that the wave attenuation decreases and then reaches a plateau due to the absorption of the free water contained in the sample. In 1982 Gorur et al. [17] introduced a different measurement method, focusing on using a two-port (a waveguide section full of concrete) and using the scattering parameters S to compute the complex dielectric constant. Results showed that the dielectric permittivity decreases with time and increases with the w/c ratio of the concrete. More recently, the same results were obtained by Haddad and Al-Quady [18], that used a coaxial transmission line in the range of 100 MHz to 1 GHz, whereas Buyukozturk et al. determined the complex permittivity and the loss factor of several materials from the transmission coefficient and the Time Difference of Arrival (TDOA), using a network analyzer in the range of 8-18 GHz, obtaining similar results. More in general, in the last two decades, several different techniques have been used to study and monitor the dielectric properties of concrete, like horn lens antennas [19]–[21], Ground Penetrating Radar (GPR) [22], [23], time-domain reflectometry [24], and several types of waveguides [25]–[28]. However, all the above techniques only focused on the variations of dielectric properties without trying to correlate them with the changes in the concrete physical properties. Furthermore, their outputs are often hard to interpret and interfaced with quantitative measurements and lack robust predictive models. Different from conventional spectra-based approaches, the main goal of this work is to introduce a sensing technique that operates by extracting hidden information from raw data, even if this is not evident at first sight [29].

The proposed approach uses a Vector Network Analyzer (VNA) connected with a rectangular open cavity antenna emitting an RF signal towards the concrete material. The VNA allowed us to grab scattering coefficients (S-parameters) and derive them in a spectrum between 1.5GHz and 14GHz, a range where the water component presents a higher dielectric constant (from 88 to 27) versus other concrete constituents. Many concrete cubic samples were cast simultaneously using the same material (sample twins) from which spectra were acquired for 28 days. Part of these samples was used in destructive tests to determine compressive strength. Then, the acquired spectra and the compressive strength data were used as input for a multivariate statistical analysis referred to as the Soft Independent Method of Class Analogy (SIMCA). This algorithm allows for resolving classification problems, where new samples are assigned to existing groups of similar samples, thanks to modeling the common properties of the classes [30]. Finally, a subset of spectra was used to create a preliminary regression model, thanks to a technique called Partial Least Square Regression (PLSR), which allows for predicting compressive strength values from newly acquired spectra.

Since the predictive model could be easily implemented in low computationally demanding digital platforms such as a microcontroller, the present research aims at exploring a compact, non-destructive and cost-effective instrumental chain for concrete hardening monitoring.

2. Preliminary investigations

The task of the preliminary tests is twofold. First, we must determine the best microwave spectrum that copes with the application. Second, we need to understand the most suitable spectrum for building the dataset of the predictive model.

As far as the first objective is concerned, we need to investigate the penetration depth of microwaves inside the concrete to understand the best operating spectra bandwidth. This data is important since the proposed technique aims to extract material information from a sufficiently large volume where the surface can differ from the inner material conditions. Therefore, the radio frequency penetration depth was estimated from experimental permittivity measurements. The penetration depth (d_p) is defined as the distance where an incident electromagnetic wave penetrates the material under test with an intensity (power) falling to $1/e$ [31] and it can be calculated by permittivity as follows

$$d_p = \frac{c}{2\sqrt{2}\,\pi f \left\{ \varepsilon_r' \left[\sqrt{1+\left(\frac{\varepsilon_r''}{\varepsilon_r'}\right)^2} - 1 \right] \right\}^{\frac{1}{2}}}, \quad (1)$$

where c is the speed of light in vacuum, 2.998×10^8 m/s, f is the frequency (Hz), ε_r' and ε_r'' are, respectively, the real and imaginary part of the relative complex permittivity $\varepsilon_r^* = \varepsilon_r' + j\varepsilon_r''$, depending on the frequency. It should be pointed out that equation (1) is related to the matter interaction with plane waves and in far-field conditions [32], which is only an approximation in our case. However, the calculation is useful to get a rough estimation of the spectrum range used for this application. Preliminary EM simulations considering the true geometrical features of the entire system show penetration depths longer by a factor between two and three. However, these analyses will be objective in a future paper since they are away the focus of the present one.

The complex parameters were assessed using an open-ended coaxial instrumental chain (Speag, DAK 3.5, Swiss) and reflectometer (R140, Copper Mountain, USA), depicted in Fig. 1.

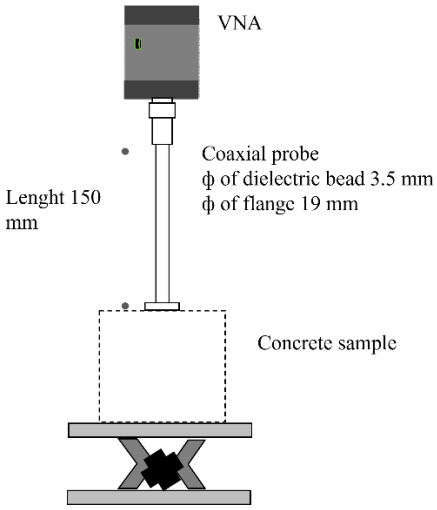

Figure 1. Layout of the instrumental setup used for the preliminary calculation of penetration depth.

The frequency range investigated was 1.5-14 GHz, and each spectrum was composed of 10000 spectral points. Measurements were conducted after 2.2 hours of the concrete preparation (wet), after 23 hours (semi-dried), and at advanced hardening (104 hours). The latter time is considered sufficient to achieve the drying condition of the concrete uppermost layer that can be assumed (from the dielectric point of view) as representative of the hardening state of the inner part of the sample after about 28 days. On the other hand, due to its conformation, the coaxial probe performs only superficial dielectric measurements (in the order of a few mm), so the measurements can be taken as indicative of the sample's inner dielectric properties. As a reference, the real and imaginary values of the complex electric permittivity (ε_r' and ε_r'') for a frequency of about 2 GHz were the following: 27.3+j7.1 (wet), 11.2+j1.7 (semi-dried), 7.3+j0.89 (dried). The penetration depth (d_p) calculated using (1) within the instrument spectrum bandwidth is shown in Fig. 2.

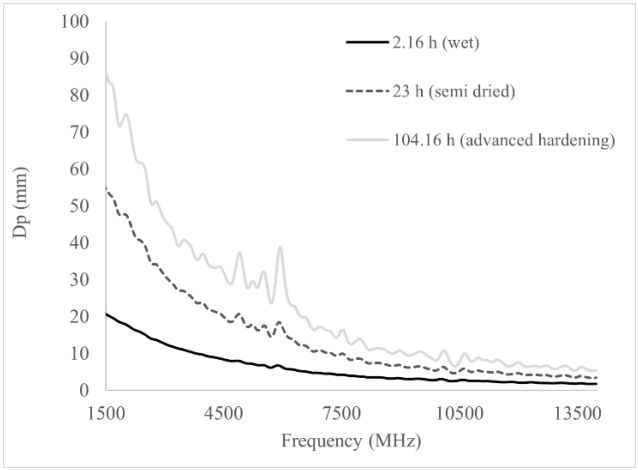

Figure 2. Penetration depth (d_p) as a function of frequency (MHz) for wet, semi-dried, and dried concrete.

In conclusion, we can say that the spectra frequency range and the concrete's hydration level directly affect the penetration depth. Using the bandwidth we investigated, we can grab information from the concrete under test with a d_p in the order of centimeters scale, especially for frequencies below 7 GHz. It should be pointed out that the penetration depth is a reference parameter that indicates the distance at which the surface intensity decays of a factor 1/e. However, the reading depth, i.e., the depth at which the VNA can detect the reflected signal with an acceptable signal-to-noise ratio, can be much greater than the penetration depth. This mainly depends on the output power and dynamic range of the instrument. In our application, tests (not discussed in this paper) conducted by placing a reflecting surface under the concrete sample found a reading depth up to about three times greater than the penetration depth. Therefore, we can still detect reflected signals at further greater depths using higher incident power and/or instrumental chains with lower noise floor. So, data from the calculation and preliminary tests were used to design the microwave apparatus. It was chosen as a probe, a short open section of a waveguide optimized for a TE10 mode at 2GHz with a cutoff frequency of 1.56GHz (Fig. 3a) with the antenna inserted at a ¼ of the guided wavelength.

As far as the second objective of preliminary investigations is concerned, i.e., the choice of the kind of spectrum, we proceeded as follows. We connected the designed open waveguide, as shown in Fig. 3a, and used the setup as detailed in the following "Measurement Setup" section. Then, we acquired several spectra based on S-parameters, e.g., real and imaginary parts of S_{11} and VSWR spectra, in the frequency range from 1.5 to 14 GHz). The relationship between S_{11} and VSWR is given by

$$VSWR = \frac{1+|S_{11}|}{1-|S_{11}|}, \tag{2}$$

where

$$|S_{11}| = \sqrt{\left(\mathrm{Re}(S_{11})\right)^2 + \left(\mathrm{Im}(S_{11})\right)^2}, \tag{3}$$

We noticed that they all show variations versus the compressive strength state due to the evolution of the chemical process. However, to understand what spectrum is better suited for building the dataset, we compared the Root Mean Square Error in Cross-Validation (RMSECV) values of the predictive model (see the "Result" section) for the datasets built on different spectra. Since we did not get any appreciable differences, we decided to use VSWR mainly for its better simplicity of implementation [33] in a future prototype. An example of a VSWR spectrum grabbed with the designed setup is shown in Fig.3b.

3. Measurement setup

A total of 16 cubic samples (15×15×15 cm^3) were prepared for the acquisition of electromagnetic spectra

and compressive strength evaluation. All the samples were cast (compacted by needle vibrator) from the same batch, with a *w/c* ratio of 0.45, and the following components: 40 kg of cement, 18 L of water, 17 kg of fine sand 0-1 mm, 74 kg of sand 0-5 mm and 85 kg of gravel 5-15 mm. After the first day, when the concrete hardened with progressive reduction of the free water content, all the samples were stored in a climatic chamber with a temperature of 20±2°C and an RH=60±5%. Electromagnetic spectra were acquired with the designed cavity antenna (internal dimensions, in mm: 96 × 46 × 118) connected to a VNA (R140, Copper Mountain, USA), as shown in Fig.3. The VNA was calibrated using a calibration kit (N1801, Copper Mountain, USA). The instrument provides spectra of the real and imaginary parts of the scattering parameters. At the end of the acquisition period, which lasted 28 days, a total of 409 VSWR spectra were acquired, each calculated as the mean of 5 consecutive acquisitions. Measurements frequency decreased starting with a period of 10 minutes up to 6 hours, as reported in Table I.

To take the maximum information on the evolution of the process, spectra acquisition intervals should be set to have an equal increment of the compressive strength. Therefore, sampling intervals should not be equally spaced since the curing process is fast at the beginning and slows down to an asymptotic value at the final state. For the above reason, we have set the spectra acquisition times as summarized in Tab.1, where intervals are denser at the beginning compared to later times.

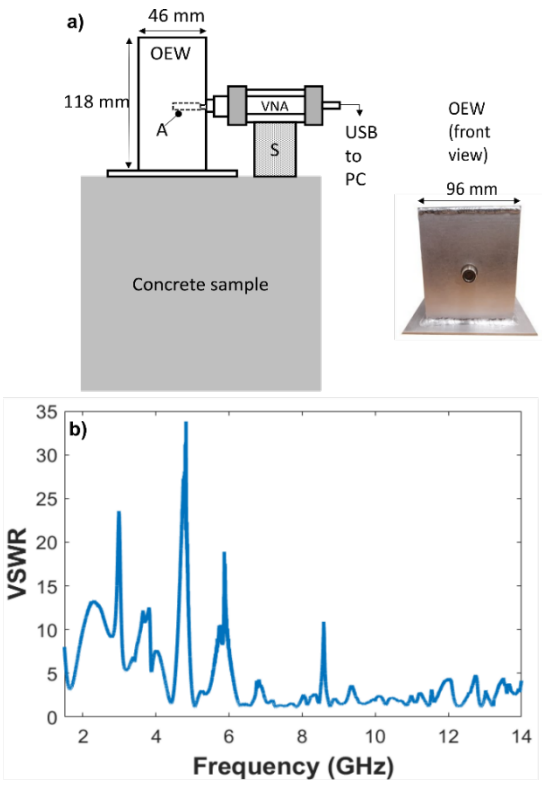

Figure 3. a) Schematic layout for the acquisition of electromagnetic spectra. Legend: OEW, open-ended waveguide; A, antenna; VNA, vector network analyzer; S, NVA support. The drawing is not to scale. b) Example of a VSWR spectrum.

Table 1. Time interval and number of acquisitions of the electromagnetic spectra on concrete

Interval of time [days]	*Time between acquisitions [min]*	*Number of acquisitions*
0 - 0.09	10	14
0.09 - 0.96	20	59
0.96 - 3.24	30	102
3.24 – 8.98	60	127
8.98 – 14.2	180	50
14.2 - 28	300	62

4. Data acquisition

The data acquisition procedure was conducted as follows. Sixteen equal concrete samples were cast at the same time. Among them, two were used for acquiring VSWR spectra and temperature measurements at increasing curing time, while others were subject to destructive compressive tests, according to the schedule of Tab. 2.

4.1 Destructive resistance tests

During the first 4 hours, when concrete was starting to develop its mechanical properties, its penetration resistance was measured by a Concrete pocket dial penetrometer Matest (diameter of the penetration plunger 6.4 mm, load 0-50 kg/cm^2), providing the load necessary to plunge a probe of the known area into concrete to a fixed depth. From this parameter, a proper calibration coefficient provided the corresponding compressive strength. Two cubic concrete samples were used for this purpose, and the penetration resistance was determined as the average of three different plungings. After the first hardening, the strength was measured at significant curing times by compression tests (2 samples for each condition) in a universal testing machine with 4000 kN capacity. Tab. 2 reports all the compressive strength values, including both those coming from the early penetration resistance values and those measured by the destructive tests. As shown in Fig. 4, the compression strength values closely follow a predictive model available in the literature [34] described by the following equation

$$Res = e^{s*\left(1-\sqrt{\frac{28}{t}}\right)} * R_{c,28}, \quad (4)$$

where s = 0.2, t is the time from cast expressed in days, and $R_{c,28}$ = 27.8 MPa is the 28 days compressive strength.

Table 2. Time of acquisition and results of the compression strength of concrete.

Time [day]	*Strength [MPa]*	*Instrument*
0.045	0.67	Concrete penetrometer Matest
0.062	0.81	
0.083	2.13	
0.104	2.67	
0.125	3.60	
0.146	3.90	
0.167	5.00	
1	11.7	Universal Testing machine
2	17.7	
3	18.9	
7	23.4	
10	23.7	
14	25.7	
28	27.8	

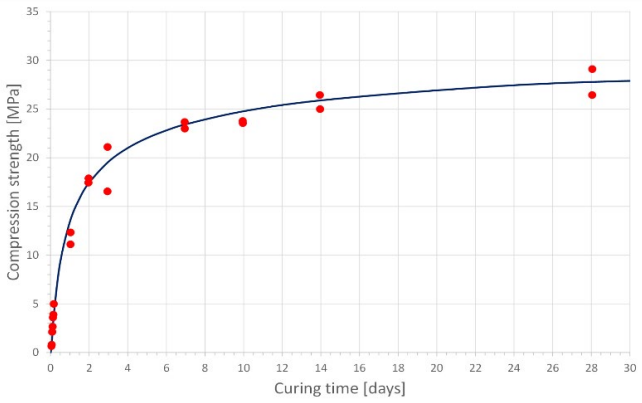

Figure 4. Measured resistance (dots) and theoretical value according to (4) (line).

4.2 Spectra acquisitions

Fig. 5 shows all the 409 acquired spectra, colored by the order, from blue (first measure) to yellow (last measure). It is easy to see that in certain regions, there is a clear trend in the spectra, where the VSWR values gradually change over time in a monotonous way. This trend is a good indication for the creation of statistical models: already, from the observation of these data, it is plausible to assume that the spectra contain useful information about the dielectric properties of the concrete and, consequently, about its free water content and compressive strength. Moreover, this phenomenon is more pronounced in the first part of the spectrum (1.5-6 GHz), which is conveniently also the part of the spectrum with the higher penetration depth, as reported previously.

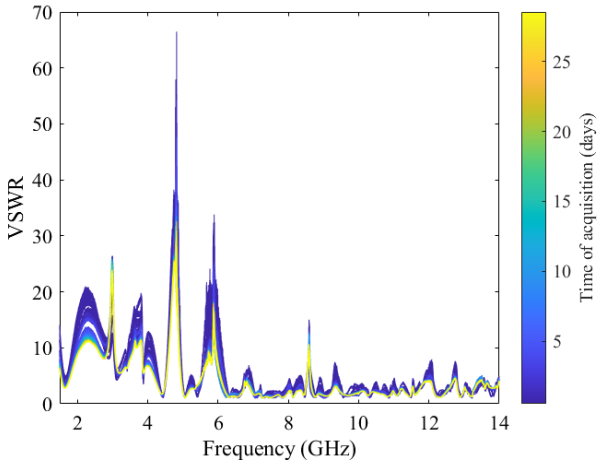

Figure 5 Visualization of all the 409 acquired spectra, colored following the acquisition time, from blue (day 1) to yellow (day 29).

5. Statistical inference and predictive model setup

The objective of the present section is to present the theory behind the statistical analyses used to set up the predictive models, inferring the compression strength from spectra. Firstly, a cluster analysis was carried out by dividing the full scale of values into several ranges of interest. Then, VSWR spectra were used with the measured compressive strength values to create a classification model that estimates the actual strength range from spectral variations. In general, a spectrum could be seen as a series of K variables; this means that a single point could describe it in a *K-dimensional* space. In the present case, the variables are the values of VSWR at different wavelengths. The spectral data acquired on the concrete sample created the first experimental dataset, which was utilized as input for the multivariate analyses. Among them, the first one used on these data was the Principal Component Analysis (PCA) to analyze the primary relationships between all the variables. PCA finds a number $A < K$ of new variables, called Principal Components (PCs), that contain most of the original data variance. From a geometrical point of view, PCA extracts an A-dimensional subspace of the original K-dimensional space that still contains most of the original information of the spectra. Thus, following the classical linear regression analysis approach [35], the original dataset matrix $\boldsymbol{X}$ of dimensions $N \times K$ (N spectra, each defined by K frequency values) could be defined as

$$\boldsymbol{X} = \boldsymbol{T}\boldsymbol{W}' + \boldsymbol{E}, \tag{5}$$

where $\boldsymbol{T}$ is called the *score* matrix of dimensions $N \times A$ and contains the spectral acquisitions already present in $\boldsymbol{X}$ but recalculated in the new PCs variables, whereas $\boldsymbol{W}'$, of dimensions $A \times K$, is the transpose of the weight (or *loadings*) matrix, and represents how each of the original K variables contributes to the calculation of the values of $\boldsymbol{T}$ starting from $\boldsymbol{X}$. Finally, $\boldsymbol{E}$ is the residual matrix. Returning to the geometrical viewpoint, the PCA

algorithm calculates the values of the loading that maximize the projections in the PC's subspace, thus minimizing the residual errors. Therefore, $\boldsymbol{T}$ represents a "good summary" of the original dataset matrix $\boldsymbol{X}$ since it maintains most of the information in a reduced dimension.

This analysis helps to identify clusters in the data and allows us to divide them into classes that could maximize the interclass difference and, consequently, the classification accuracy. This "classified" dataset was used as input for a second analysis, called Soft Independent Method of Class Analogy (SIMCA), to solve a classification problem and identify the more representative class for future acquisitions. SIMCA approach could be divided into two subsequent steps. In the first one, a PCA analysis is carried out separately on the spectra of each cluster class (Fig. 5), finding a series of A PCs for each of them, all contained in the original K-data space (A could be different for each class). From the geometrical point of view, projection subspaces are defined for each class. Then, for each new acquisition submitted to the model, SIMCA calculates for each class two statistical parameters, called Q residuals and Hotelling's T^2, describing how well the new sample belongs to them[36]. In particular, the Q residual, also called Squared Prediction Error (SPE), measures the squared error of the projection of the new sample in the PC spaces [37]. Thus, it indicates the amount of information lost in the process, measured as the distance from the PC's subspace [38]. In general, for an i-th observation, the Q-residual related to a given cluster is calculated as

$$Q_i = e_i e_i', \tag{6}$$

where e_i is the residual of the projection of the observation in the PCA models calculated by the SIMCA analysis, and it is a row of the $\boldsymbol{E}$ matrix for the related cluster. On the other side, Hotelling T^2 measures the difference between the new observation and the mean value of the model. From a geometrical point of view, it measures the Euclidean distance between the center of the PCs subspace and the projection of the observation, and it is calculated as

$$T_i^2 = \frac{t_i t_i'}{\lambda}, \tag{7}$$

where t_i is the i-th observation *score*, that is, the projection value of the observation along all the considered PC variables (it is a row of the $\boldsymbol{T}$ matrix); λ represents the distribution variance [37]. Note that T_i and Q_i are orthogonal contributions. For a given new observation, SIMCA combines these two parameters in a final parameterR_i, with the following formula

$$R_i = \sqrt{(Q_i^2 + T_i^2)}. \tag{8}$$

The observation is finally assigned to the class with the lowest R_i value: it is the class that better describes the new spectrum. Fig. 6 shows a geometrical visualization of the prediction process for a new sample between two classes.

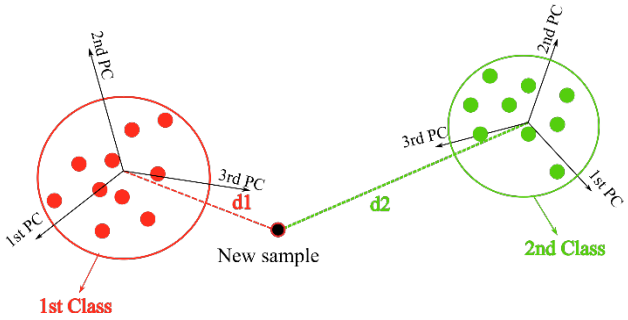

Figure 6. Visual representation of the prediction of a new sample with a SIMCA algorithm. Every point represents a spectrum, whereas PCs create a subspace in the k-dimensional space of the original variables (frequencies). In the present case, d is defined by eq. 8.

For our application, 80% of the original spectral data (266) were used to build the SIMCA model, creating the calibration test. The remaining 20% (66), randomly chosen and, in percentage, equally distributed between all the classes, were used as a test set to perform the validation phase: these spectra were assigned to the different classes by the model, and the results compared to the actual classes, obtaining a confusion table and a confusion matrix, useful to compare the predictive robustness to different SIMCA models.

To obtain more informative results, 10 SIMCA models were created, with different test sets (always chosen at random), and a mean value of the obtained statistical parameters was calculated. Moreover, we explored the use of the regression analysis called PLSR for the quantitative prediction of concrete resistance. PLSR is based on PCA, but it goes a step further with the introduction of a second dataset, $\boldsymbol{Y}$ $(N \times M)$, that contains the values of M variables of interest (for this work, the compressive strength, thus $M = 1$) related to every spectrum contained in the dataset $\boldsymbol{X}$. Also the PLSR algorithm founds new directions in the dataspace, similar but not equal to the PCs: these new directions, called Latent Variables (LVs), are the one that at the same time contains most of the $\boldsymbol{X}$ data variance, of the $\boldsymbol{Y}$ data variance, and the covariance between the two datasets [35].

So, from a mathematical point of view, $\boldsymbol{X}$ can be defined with eq (5), where the spectra contained in $\boldsymbol{T}$ are calculated alongside the LVs. Regarding $\boldsymbol{Y}$, thanks to the assumptions of PLSR, it can also be described with the score matrix $\boldsymbol{T}$:

$$\boldsymbol{Y} = \boldsymbol{T}\boldsymbol{C}' + \boldsymbol{F}; \tag{9}$$

Merging (5) and (9), we obtain:

$$\boldsymbol{Y} = \boldsymbol{X}\boldsymbol{W}\boldsymbol{C}' + \boldsymbol{F} = \boldsymbol{X}\boldsymbol{B} + \boldsymbol{F}; \tag{10}$$

defining a direct and linear relationship between an input spectrum $\boldsymbol{X}$ and the corresponding variable of interest $\boldsymbol{Y}$ [39].

For the creation of the PLSR model, we needed a strict correlation between the spectra contained in dataset ***X*** and the respective values of compressive strength contained in ***Y***: to obtain this, we selected the three acquisitions closest in time with every one of the destructive tests reported in Table 1, for a total of 33 spectra, and linked the respective resistance values to them, obtaining the ***X*** and ***Y*** calibration datasets. This selection was performed to create a regression model with the spectra strongly related to the experimental resistance value. The prediction ability of the model was tested with a process named Cross-Validation (CV): it divides the datasets into a series of subsets, then iteratively selects one of them and uses the other ones to create a model. The selected subset is used to test its prediction ability, evaluating the error between the predicted and the actual ***Y*** subset values. With this technique, a coefficient of determination of cross-validation (R^2_{CV}) was obtained: the closer the value of this parameter is to 1, the better the predictive abilities of the model.

5.1 Data preprocessing and explorative PCA

Before being used as input for SIMCA analysis, all the VSWR spectra were preprocessed with two standard techniques: the Savitsky-Golay smoothing and autoscale, which performs a mean centering and scales each variable to unit standard deviation. This preprocessing phase, and all the following statistical analyses, were performed thanks to PLS_Toolbox [40], a software developed by EVRI in a MATLAB environment. The first statistical analysis performed on the preprocessed spectra was an explorative PCA (with $A=5$ PCs) to visualize clusters in the dataset and divide them into the classes mentioned above, which is necessary for the following statistical analyses. We empirically divide the data into 4 classes based on the obtained score plot, shown in Fig. 7. This plot shows the data in the space defined by the first 2 of the 5 total PCs, which contains the majority of data variance (62.95 % the first PC and 17.37 % the second one, for a total of 80.32%), allowing us to easily visualize the data trends and clusters. Then, we linked these classes to ranges of compressive strength. To do so, we evaluated the time of acquisition of the first and last acquisition of each subset: if it is close to one of the destructive tests reported in Table 2, we used the experimental value as a reference; in the other case, we evaluated the theoretical strength with (4). With this method, we obtained these references for the compressive strength: <12 MPa for class 1, 12 - 18.5 MPa for class 2, 18.5 – 23.4 MPa for class 3, and 23.4-28 MPa for class 4.

Finally, for creating the PLSR model, a different set of preprocessing algorithms was used. After the same Savitsky-Golay algorithm, we applied a Multiplicative Scatter Correction (MSC) that performs a weighted normalization, a baseline correction of the spectra, and a Mean Centering.

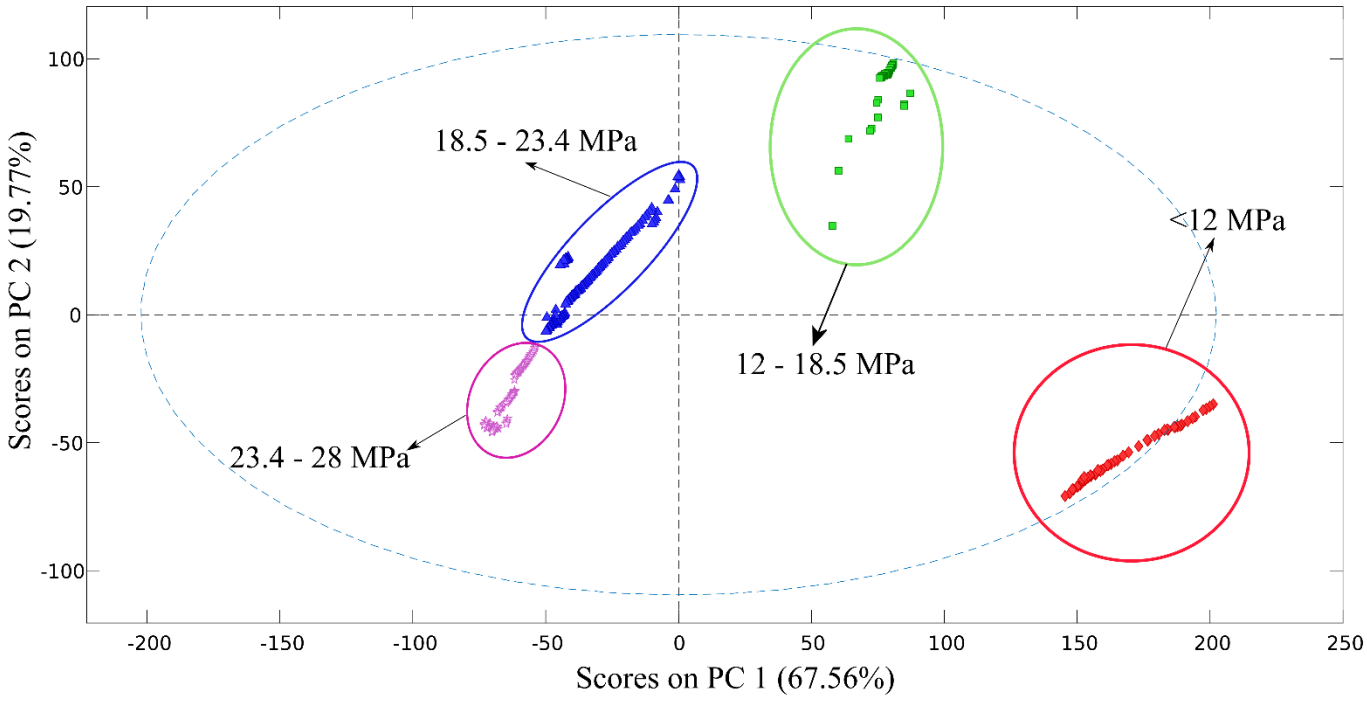

Figure 7. PCA score plot. The 4 classes used as input for the classification algorithms are highlighted.

6. Results

As reported in section 5, the SIMCA analysis utilizes two parameters, Q and T^2, to assign new samples to the more representative class (the one with the lowest value of both parameters). Fig. 8 shows the Q-T^2 plots of one among the 10 SIMCA models as an example. It can be observed that for the first two classes (<12 and 12-18.5 MPa), there is a considerable difference between the samples of the correct class to the other ones, while for class 3 (18.5 – 23.4 MPa) and 4 (23.4 – 28 MPa) the difference is less noticeable. However, as shown in Figs 8e and 8f, there are still enough differences between the values of Q and T^2 to correctly classify the test samples in most cases.

Table III reports the mean confusion matrix, whereas Table IV is the mean of the classification parameters measured by PLS_Toolbox [41], both obtained from the mean of the 10 SIMCA models. *N* is the number of spectra used as a test set for every class. The values *TPR* (sensitivity), *FPR* (false positive rate), *TNR* (specificity), *FNR* (false negative rate), Err (misclassification error), *PPV* (precision), and *F1* (F1 score) are given by the following relationships

$$TPR = \frac{TP}{TP + FN} \tag{8}$$

$$FPR = \frac{FP}{FP + TN} \tag{9}$$

$$TNR = \frac{TN}{TN + FP} \tag{10}$$

$$FNR = \frac{FN}{FN + TP} \tag{11}$$

$$Err = \frac{FP + FN}{TP + TN + FP + FN} \tag{12}$$

$$PPV = \frac{TP}{TP + FP} \tag{13}$$

$$F1 = \frac{2}{(TPR)^{-1} + (PPV)^{-1}} \tag{14}$$

where *TP* = true positive, *FP* = false positive, *TN* =true negative and *FN* =false negative. Note that one of the most significant parameters for the classification performance is *F1*, which is the harmonic mean of sensitivity and precision, thus emphasizing the smaller of the two. Since precision is related to the displacement of the decision threshold towards the *negative result* distribution and sensitivity is related to threshold displacement for the *positive results*, *F1* is a representative indicator of the threshold placement between classes.

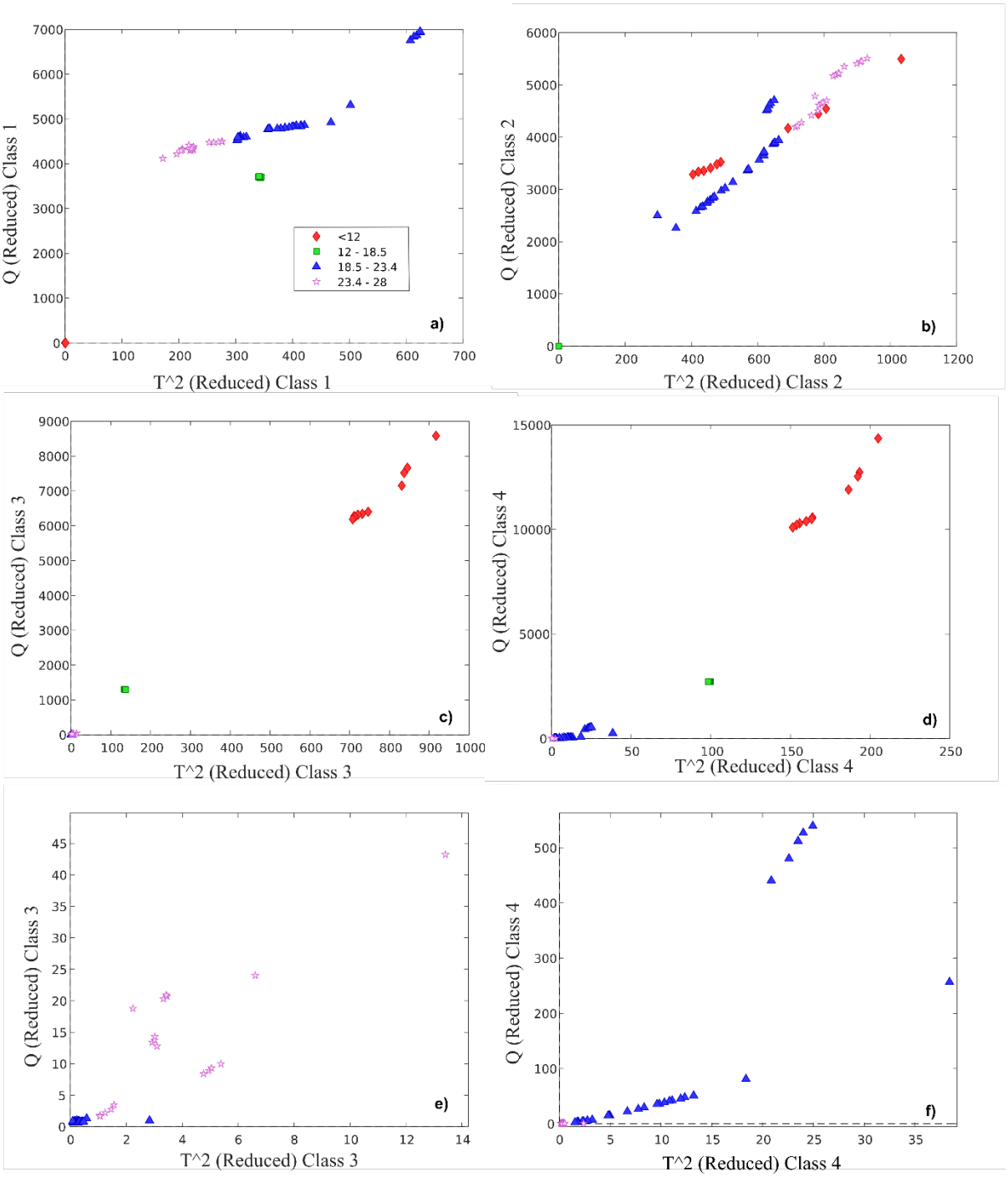

Figure 8. a) - d) Plots of T2 vs. Q for each of the 4 classes. e)-f) Zoom for classes 3 and 4 to better visualize the differences between the test samples.

Table 3. Mean confusion matrix

	Observed class				
Predicted class		*<12*	*12-18.5*	*18.5-23.4*	*23.4-28*
	<12	10	0.2	1.4	0.4
	12-18.5	0	6.7	0	0
	18.5-23.4	0	0	27.3	0
	23.4-28	0	0	0	18.3

Table 4. Mean statistical parameters, calculated from the confusion matrix

Class	N	TPR	FPR	TNR	FNR	Err	PPV	F1
<12	10	1.00	0.04	0.96	0.00	0.03	0.84	0.91
12-18.5	7	0.97	0.00	1.00	0.03	0.003	1	0.98
18.5-23.4	30	0.95	0.00	1.00	0.05	0.02	1	0.97
23.4-28	19	0.97	0.00	1.00	0.03	0.01	1	0.99

TPR = True Positive Ratio, FPR = False Positive Ratio, TNR = True Negative Ratio, FNR = False Negative Ratio, Err = Error, PPV = Precision, F1 = F1 score

The most interesting and promising results are the low misclassification error *Err* and the high *F1* score obtained for each class: the SIMCA models have a good prediction ability concerning the four classes. The "limit" values (0 and 1) obtained in some cases from TPR, FPR, TNR, and PPV could be explained by the fact that, given the low number of test samples and the good general clustering of data, no test samples were classified to a wrong class.

Regarding PLSR, we created a series of models with different datasets as input, as already stated in the "Preliminary investigations" section. We inspected the parameter Root Mean Square Error in Cross-Validation (RMSECV), which indicates how closely the models predict the measured values: it can have any values starting from 0, and the smaller, the better: 0.846 (1.5-14 GHZ) – 0.743 (1.4-6 GHz) for the real component of S11, 0.814 (1.5-14 GHZ) - 0.835 (1.4-6 GHz) for the imaginary component of S11, 0.964 (1.5-14 GHZ) - 0.984 (1.4-6 GHz). Therefore, as stated in the "Preliminary investigations" section, there is no real advantage of using a kind of spectrum versus others, and we have chosen VSWR for future implementation ease reasons. Fig. 9 shows the results of the Cross-Validation for the chosen dataset (VSWR, 1.5-14 GHz), where for the purpose, we used N = 33. For all the $\boldsymbol{X}$ spectra (represented by the blue dots), the prediction of the compressive strength was close to the $\boldsymbol{Y}$ values (the red line represents the ideal fit between the prediction and the observed values). This result is also supported by the excellent value of the parameters RMSECV, 0.986, and R^2_{CV}, equal to 0.997.

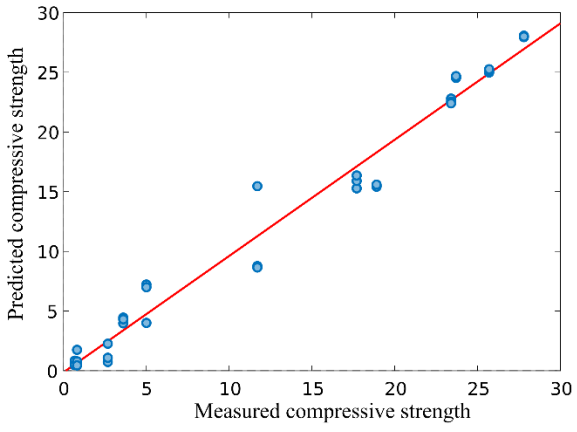

Figure 9. Predicted vs. observed plot of the PLSR model. The blue dots are the compressive strength predicted by the model (Y-axis) versus the observed values (X-axis). The red line is the bisector, representing the ideal prediction

7. Conclusions

A novel non-invasive approach for concrete resistance monitoring during curing time was explored; it consists of the acquisition of reflected microwave spectra coupled with a statistical predictive model. The technique is based on a VNA governed by a PC-based graphical user interface coupled with an open waveguide. This system was used to acquire a total of 409 VSWR spectra in the frequency range from 1.5 to 14 GHz over 28 days: in several frequency sub-ranges, these spectra present a clear descending trend concerning the acquisition time, showing that they contain good information about the amount of free water in the concrete samples. First, we divided measurements into 4 classes thanks to an explorative PCA. At the same time, a series of 14 tests for compression resistance was performed (7 with the penetrometer and 7 with the universal testing machine), and the obtained values were used to link the classes to 4 different concrete resistance ranges, up to 28 MPa. These classified data were the input for two types of predictive models: the first one, SIMCA, was able to classify new acquisitions in one of the 4 resistance classes, whereas the PLSR, applied to a subset of data, predicted the precise values of resistance. The statistical results obtained from the SIMCA models show a relatively low error of misclassification and a high *F1* score. Also, the regression prediction ability of the PLSR model was good, with a very high value of the parameter R^2_{CV}. For future development, it will surely be necessary to test the model accuracy with spectra acquired on concrete samples with different parameters (water content ratios, components, etc.) and under different environmental conditions. The parameters of the acquisition process could be refined in future studies: the results show that the most informative region is a subset of the one we explored (from 1.5 to 6 GHz). Moreover, other reflection parameters than the VSWR, like the phase, the impedance, or the simpler S11 parameter (real or imaginary), could contain information to predict the concrete compressive strength. Finally, regarding

classification analyses, different classes and resistance ranges could be tested to find any accuracy improvements. In conclusion, this initial approach is promising for developing future autonomous and non-invasive monitoring devices for curing time monitoring since the predictive models, once developed offline on a dataset, could be easily implemented in a microcontroller in real time due to its low level of computation: to predict a new compressive strength value, $\hat{Y}$ related to an acquired spectrum, it is sufficient to multiply $\hat{X}$ by the coefficient array B (see (10)), resulting in a linear complexity *O(K)*.

Acknowledgments

The authors acknowledge Alessandra Di Florio Di Renzo for helpful suggestions and hints on EM modeling and interpretation.

References

[1] S. Popovics, *Concrete Materials - Properties, Specification and Testing*. 1992.

[2] A. Wahab, M. M. A. Aziz, A. R. M. Sam, K. Y. You, A. Q. Bhatti, and K. A. Kassim, "Review on microwave nondestructive testing techniques and its applications in concrete technology," *Constr. Build. Mater.*, vol. 209, pp. 135–146, 2019, doi: 10.1016/j.conbuildmat.2019.03.110.

[3] X. Shi, N. Xie, K. Fortune, and J. Gong, "Durability of steel reinforced concrete in chloride environments: An overview," *Constr. Build. Mater.*, vol. 30, pp. 125–138, 2012, doi: 10.1016/j.conbuildmat.2011.12.038.

[4] S. Ismail, W. H. Kwan, and M. Ramli, "Mechanical strength and durability properties of concrete containing treated recycled concrete aggregates under different curing conditions," *Constr. Build. Mater.*, vol. 155, pp. 296–306, 2017, doi: 10.1016/j.conbuildmat.2017.08.076.

[5] K. Van Den Abeele, W. Desadeleer, G. De Schutter, and M. Wevers, "Active and passive monitoring of the early hydration process in concrete using linear and nonlinear acoustics," *Cem. Concr. Res.*, vol. 39, no. 5, pp. 426–432, 2009, doi: 10.1016/j.cemconres.2009.01.016.

[6] Z. Song, T. Frühwirt, and H. Konietzky, "Fatigue characteristics of concrete subjected to indirect cyclic tensile loading: Insights from deformation behavior, acoustic emissions and ultrasonic wave propagation," *Constr. Build. Mater.*, vol. 302, no. December 2020, p. 124386, 2021, doi: 10.1016/j.conbuildmat.2021.124386.

[7] N. Tareen, J. Kim, W. K. Kim, and S. Park, "Comparative analysis and strength estimation of fresh concrete based on ultrasonicwave propagation and maturity using smart temperature and PZT sensors," *Micromachines*, vol. 10, no. 9, pp. 1–17, 2019, doi: 10.3390/mi10090559.

[8] Kaplanvural, E. Pekşen, and K. Özkap, "Volumetric water content estimation of C-30 concrete using GPR," *Constr. Build. Mater.*, vol. 166, pp. 141–146, 2018, doi: 10.1016/j.conbuildmat.2018.01.132.

[9] İ. Kaplanvural, K. Özkap, and E. Pekşen, "Influence of water content investigation on GPR wave attenuation for early age concrete in natural air-drying condition," *Constr. Build. Mater.*, vol. 297, 2021, doi: 10.1016/j.conbuildmat.2021.123783.

[10] Y. Y. Lim, K. Z. Kwong, W. Y. H. Liew, and C. K. Soh, "Practical issues related to the application of piezoelectric based wave propagation technique in monitoring of concrete curing," *Constr. Build. Mater.*, vol. 152, pp. 506–519, 2017, doi: 10.1016/j.conbuildmat.2017.06.163.

[11] Y. Y. Lim, S. T. Smith, and C. K. Soh, "Wave propagation based monitoring of concrete curing using piezoelectric materials: Review and path forward," *NDT E Int.*, vol. 99, no. June, pp. 50–63, 2018, doi: 10.1016/j.ndteint.2018.06.002.

[12] P. Priyada, R. Ramar, and Shivaramu, "Determining the water content in concrete by gamma scattering method," *Ann. Nucl. Energy*, vol. 63, pp. 565–570, 2014, doi: 10.1016/j.anucene.2013.07.049.

[13] N. Sabbağ and O. Uyanık, "Determination of the reinforced concrete strength by apparent resistivity depending on the curing conditions," *J. Appl. Geophys.*, vol. 155, pp. 13–25, 2018, doi: 10.1016/j.jappgeo.2018.03.007.

[14] L. Xiao and Z. Li, "Early-age hydration of fresh concrete monitored by non-contact electrical resistivity measurement," *Cem. Concr. Res.*, vol. 38, no. 3, pp. 312–319, 2008, doi: 10.1016/j.cemconres.2007.09.027.

[15] J. Bhargava and K. Lundberg, "Determination of moisture content of concrete by microwave-resonance method," *Matériaux Constr.*, vol. 5, no. 3, pp. 165–168, 1972, doi: 10.1007/BF02539259.

[16] F. H. Wittman and F. Schlude, "MICROWAVE ABSORPTION OF HARDENED CEMENT PASTE," *Cem. Concr. Res.*, vol. 5, pp. 63–71, 1975.

[17] K. Gorur, M. K. Smit, and F. H. Wittmann, "Microwave study of hydrating cement paste at early age," *Cem. Concr. Res.*, vol. 12, no. 4, pp. 447–454, 1982, doi: 10.1016/0008-8846(82)90059-X.

[18] R. Haddad and I. Al-Qadi, "Characterization of Portland Cement Concrete Using," *Cem. Concr. Res.*, vol. 28, no. 10, pp. 1379–1391, 1998.

[19] S. N. Kharkovsky, M. F. Akay, U. C. Hasar, and C. D. Atis, "Measurement and monitoring of microwave reflection and transmission properties of cement-based specimens," *IEEE Trans. Instrum. Meas.*, vol. 51, no. 6, pp. 1210–1217, 2002, doi: 10.1109/TIM.2002.808081.

[20] M. Jamil, M. K. Hassan, H. M. A. Al-Mattarneh, and M. F. M. Zain, "Concrete dielectric properties investigation using microwave nondestructive techniques," *Mater. Struct. Constr.*, vol. 46, no. 1–2, pp. 77–87, 2013, doi: 10.1617/s11527-012-9886-2.

[21] T. T. Dinh *et al.*, “Dielectric material characterization of concrete in GHz range in dependence on pore volume and water content,” *Constr. Build. Mater.*, vol. 311, no. October, 2021, doi: 10.1016/j.conbuildmat.2021.125234.

[22] G. Klysz, J. P. Balayssac, and X. Ferrières, “Evaluation of dielectric properties of concrete by a numerical FDTD model of a GPR coupled antenna-Parametric study,” *NDT E Int.*, vol. 41, no. 8, pp. 621–631, 2008, doi: 10.1016/j.ndteint.2008.03.011.

[23] P. Shen and Z. Liu, “Study on the hydration of young concrete based on dielectric property measurement,” *Constr. Build. Mater.*, vol. 196, pp. 354–361, 2019, doi: 10.1016/j.conbuildmat.2018.11.150.

[24] N. E. Hager and R. C. Domszy, “Monitoring of cement hydration by broadband time-domain-reflectometry dielectric spectroscopy,” *J. Appl. Phys.*, vol. 96, no. 9, pp. 5117–5128, 2004, doi: 10.1063/1.1797549.

[25] X. Zhang, X. Z. Ding, T. H. Lim, C. K. Ong, B. T. G. Tan, and J. Yang, “Microwave study of hydration of slag cement blends in early period,” *Cem. Concr. Res.*, vol. 25, no. 5, pp. 1086–1094, 1995, doi: 10.1016/0008-8846(95)00103-J.

[26] P. Juan-García and J. M. Torrents, “Measurement of mortar permittivity during setting using a coplanar waveguide,” *Meas. Sci. Technol.*, vol. 21, no. 4, 2010, doi: 10.1088/0957-0233/21/4/045702.

[27] K. Chung and S. Kharkovsky, “Measurements of microwave reflection properties of early-Age concrete and mortar specimens,” *Conf. Rec. - IEEE Instrum. Meas. Technol. Conf.*, pp. 1295–1300, 2014, doi: 10.1109/I2MTC.2014.6860954.

[28] A. Wahab, M. M. A. Aziz, A. R. MohdSam, and K. Y. You, “Application of microwave waveguide techniques for investigating the effect of concrete dielectric and reflection properties during curing,” *J. Build. Eng.*, vol. 38, no. January, p. 102209, 2021, doi: 10.1016/j.jobe.2021.102209.

[29] M. Tartagni, *Electronic Sensor Design Principles*. Cambridge (UK): Cambridge University Press, 2022.

[30] S. WOLD and M. SJÖSTRÖM, “SIMCA: A Method for Analyzing Chemical Data in Terms of Similarity and Analogy,” 1977, pp. 243–282.

[31] Y. Wang, T. D. Wig, J. Tang, and L. M. Hallberg, “Dielectric properties of foods relevant to RF and microwave pasteurization and sterilization,” *J. Food Eng.*, vol. 57, no. 3, pp. 257–268, 2003, doi: 10.1016/S0260-8774(02)00306-0.

[32] M. P. Robinson, J. Clegg, and D. A. Stone, “A novel method of studying total body water content using a resonant cavity: Experiments and numerical simulation,” *Phys. Med. Biol.*, vol. 48, no. 1, pp. 113–125, 2003, doi: 10.1088/0031-9155/48/1/308.

[33] EMC Technology Inc., “A VSWR Meter Using a Power Sensing Termination,” *Microw. J.*, 2000.

[34] International Federation for Structural Concrete, *Model Code 2010 - Final draft, Volume 1*. 2012.

[35] K. Dunn, *Process Improvement Using Data*. 2021.

[36] I. Eigenvector Research, “T-Squared Q residuals and Contributions,” 2012. https://www.wiki.eigenvector.com/index.php?title=T-Squared_Q_residuals_and_Contributions (accessed Sep. 08, 2021).

[37] T. Kourti, “Application of latent variable methods to process control and multivariate statistical process control in industry,” *Int. J. Adapt. Control Signal Process.*, vol. 19, no. 4, pp. 213–246, 2005, doi: 10.1002/acs.859.

[38] H. Oddan, “Multivariate Statistical Condition Monitoring,” Norwegian University of Science and Technology, 2017.

[39] S. Wold and M. Sjostrom, “PLS-regression : a basic tool of chemometrics,” *Chemom. Intell. Lab. Syst.*, vol. 58, pp. 109–130, 2001.

[40] Eigenvector Research, Inc., “PLS_Toolbox 8.9.2 (2021).” Manson, WA USA 98831.

[41] Eigenvector Research Inc., “Confusionmatrix,” 2018. https://www.wiki.eigenvector.com/index.php?title=Confusionmatrix#Synopsis.