The use of machine learning algorithms (SVR) to link volumetric water content and complex permittivity of concretes

Amine IHAMOUTEN(1), Cedric LE BASTARD(1), Jean-Noël GAUDIN(1), Géraldine VILLAIN(2), and Xavier DEROBERT(2)

(1) CEREMA, DTO/DLRCA, ERA17-CND, Angers, France
(2) LUNAM Université, IFSTTAR Centre de Nantes, F- 44344 Bouguenais, France
*corresponding author, e-mail address: amine.ihamouten@cerema.fr, tel: 33 2 41 79 1328

Abstract
This paper deals with the development and the validation of an innovative method for estimating the volumetric water content in various concrete mixtures. A statistical resolution of the inverse problem was done using a non-deterministic optimization (Support Vector Regression – SVR) of the in-lab calibration curves that link the controlled water content in various concrete mixtures and the frequency complex permittivity issued from an electromagnetic coaxial/cylindrical transition line, in the frequency range of the ground-penetrating radar (GPR) technique. An extrapolation procedure using a frequency power law model was developed and validated for the estimation of the complex permittivity on a broad frequency bandwidth.

The two-step estimation procedure (extrapolation and SVR methods) proposed in this paper is validated on a broad array of concrete specimens hydrically controlled in laboratory. This allows us to build calibration curves relying the complex effective permittivity and the water content in various concretes and taking into account the frequency dependence.

Keywords
Durability, complex permittivity, Support Vector Regression, Jonscher’s model, coaxial/cylindrical transition line, water content, concrete.

1. Introduction
Excluding the natural aging of the concrete, the life-time of a concrete structure may be limited by the penetration of aggressive agents such as chlorides or by freezing-thawing cycles that lead to cracks and a scaling of surfaces exposed to deicing salts.

A common factor of all these pathologies is the water content. For an in situ assessment of the progression of these degradations, it is essential to adjust the models of water transfer in concrete according to the durability indicators as well as the observables issued from non-destructive (ND) measurements.

Numerous research works have shown the possibility to link the durability indicators and ND measurement methods based on the electromagnetic (EM) wave propagation [1-4]. Indeed the EM methods make it thus possible to assess the mean water content of the investigated volume under the condition to be liberated from the mix design. For each studied structure, if the concrete mix design changes, the calibration curve linking, for example, dielectric permittivity and water content changes.

In front of inverse physical modelling constraints, we have opted in this research work for a statistical resolution of the inverse problem with a non-deterministic optimization of the in-lab calibration curves that link the controlled water content in various concrete mixtures and the complex dielectric permittivity issued from EM characterization device.

In this paper, we propose to use the Support Vector Regression method to estimate the calibration curves determined on small cores from the studied concretes, at controlled water
contents in laboratory, by using an EM coaxial/cylindrical transition line [5]. This EM cell
developed in IFSTTAR’s laboratory [5] is combined with the 4p variant of Jonscher’s model
[6] to extrapolate the permittivity dispersion curves at the required frequency band [1 MHz–2
GHz]. The use of this extrapolation method enables taking into account various physical
phenomena affecting the dielectric behaviour of concrete according to the frequency.
The two-step estimation procedure (extrapolation and SVR methods) proposed in this paper is
validated on a broad array of concrete specimens hydrically controlled in laboratory. This
enables to build calibration curves relying complex effective permittivity and water content in
various concretes and taking into account the frequency dependence.
The reminder of this paper will be organized as follows. In Section II, we introduce the
general principle behind SVR method. In Section III, the EM experimental device is detailed.
In Section IV, the experimental validations of this method on various concrete mixtures will
be carried out. Lastly, a number of conclusions will be drawn in Section V.

2. Support Vector Regression

In this study, we propose to use v-SVR method to estimate the water content from frequency
complex permittivity measurements in a context of regression [7]. For that, several parameters
(v, C) and kernel parameters must be set. Here, v is set at v= 0.54 [7] and the other parameters
are determined by the well-known k-fold cross validation on training data [7]. In k-fold cross-
validation, the data used are partitioned into k equally sized subsets like shown in Fig. 1. For
each fold, the data are taken randomly. k iterations of training and validation are performed.
Within each iteration a different fold of the data is held-out for validation while the remaining
k–1 folds are used for learning. The performance is measured by the mean square error
(MSE) on the output Wc from the i-th subset that is used as validation set. The procedure is
repeated k times and the average of MSE is calculated. Through searching different
parameters (C, γ, or other) by a grid, the different parameters which provide the best average
MSE in this grid will be chosen out.

![Figure 1: Principle of K-fold cross validation.](image)

For the validation on experimental data, the total database (TD) is composed of M samples
corresponding to the hole concrete mixtures at M controlled water contents (or relative
humidity RH) and for which the frequency complex permittivity is obtained from the EM
cylindrical transition line. For each sample Wc (%), the feature vector is composed of 2N=2000
elements. The frequency bandwidth used contains N=1000 frequencies (after extrapolation).
The total database is partitioned into two databases: training database (TGD) and test database
(TTB). The TGD and TTB are composed of 2M/3 and M/3 samples respectively. In the
following, the symbols $\bar{W}_c$and $\bar{W}_c$ represent the quantity Wc in the training and test phases
respectively. The value \( \bar{W}_c \) means the estimated value of \( W_c \).

In order to set the various parameters, the k-fold cross validation is used on the TGD with \( k = 5 \). In this study, we have used ν-SVM of the LIBSVM [8] and all the results will be presented in Section IV.

3. Experimental validation

Concrete mixtures and hydric conditioning

The following set of requirements has been adopted in order to design six concrete mixtures from which samples will be cast for further EM testing; this design process includes:

- 2 water-to-cement ratios (w/c) of approx. 0.35 and 0.65, covering the range of medium to high strength concretes classically encountered in bridge construction,
- 3 maximum aggregate diameters \( D_{\text{max}} \): 4 mm, 10 mm and 20 mm. In all, four concretes, labeled B1 to B4, and two mortars, B5 and B6, were mixed as part of this design process (see Table I).

The six concrete mixtures were studied at five values of relative humidity (RH), i.e. 30%, 50%, 70%, 80%, 94% and 100%, which provided 6x36 volumetric water contents (\( W_c \)) (with six \( W_c \) equal to zero corresponding to RH=0%), ranging from 1.9% to 26.4% (Table II).

<table>
<thead>
<tr>
<th>Water-to-cement ratio (w/c)</th>
<th>B1</th>
<th>B2</th>
<th>B3</th>
<th>B4</th>
<th>B5</th>
<th>B6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maximum aggregate diameter (mm)</td>
<td>20</td>
<td>20</td>
<td>10</td>
<td>10</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>Mean saturated density by gamma dens. (( \rho_{\text{sat}} ) (kg/m(^3))</td>
<td>2406 ± 9</td>
<td>2312 ± 13</td>
<td>2370 ± 10</td>
<td>2275 ± 11</td>
<td>2257 ± 4</td>
<td>2152 ± 10</td>
</tr>
<tr>
<td>Mean saturated density by water saturation (( \rho_{\text{sat}} ) (kg/m(^3))</td>
<td>2437 ± 5</td>
<td>2346 ± 1</td>
<td>2396 ± 14</td>
<td>2302 ± 9</td>
<td>2290 ± 4</td>
<td>2176 ± 12</td>
</tr>
<tr>
<td>Bulk porosity by water saturation (( \phi ) (%)</td>
<td>12.5 ± 0.2</td>
<td>16.6 ± 0.3</td>
<td>13.5 ± 0.3</td>
<td>17.5 ± 0.6</td>
<td>19.5 ± 0.3</td>
<td>26.4 ± 0.5</td>
</tr>
<tr>
<td>Paste volume (l/m(^3))</td>
<td>374.6</td>
<td>351.4</td>
<td>369.2</td>
<td>350.0</td>
<td>531.2</td>
<td>503.1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Concretes</th>
<th>RH = 30 %</th>
<th>RH = 50 %</th>
<th>RH = 70 %</th>
<th>RH = 80 %</th>
<th>RH = 94 %</th>
<th>RH = 100 %</th>
</tr>
</thead>
<tbody>
<tr>
<td>B2</td>
<td>1.958</td>
<td>3.365</td>
<td>4.26</td>
<td>5.638</td>
<td>7.665</td>
<td>16.6</td>
</tr>
<tr>
<td>B4</td>
<td>3.679</td>
<td>5.08</td>
<td>5.953</td>
<td>7.288</td>
<td>9.384</td>
<td>17.5</td>
</tr>
<tr>
<td>B5</td>
<td>4.836</td>
<td>5.786</td>
<td>6.966</td>
<td>8.768</td>
<td>12.101</td>
<td>19.5</td>
</tr>
<tr>
<td>B6</td>
<td>4.958</td>
<td>6.749</td>
<td>7.962</td>
<td>9.96</td>
<td>12.615</td>
<td>26.4</td>
</tr>
</tbody>
</table>

EM device and processing

The EM coaxial transition line used in this study enables measuring the complex permittivity of cylindrical samples of heterogeneous materials with large aggregate dimensions (up to 25 mm) over a frequency range depending on the value of the permittivity real part. Since this EM cell is limited to the minimum frequency of 50 MHz and to the maximum frequency of 600 MHz, then we developed a new extrapolation method to estimate the
complex permittivity on a broad frequency bandwidth. This method is based on the extrapolation of the EM cell data using the 4p variant of the Jonscher’s model [6].

4. Results and discussion

Concerning the SVR processing, in this study TD is composed of M=192 samples of which the water content is between 1.9 % and 26.4 %. For each sample \( W_c (\%) \), the feature vector \( \mathbf{v}_i \) is composed of 2N=2000 elements. The frequency bandwidth used after extrapolation is [0.1-2] Ghz with N=1000 frequencies.

The total database is partitioned into two databases: training database (TGD) and test database (TTB). The TGD and TTB are composed of 128 and 64 samples respectively. As discussed below, the k-fold cross validation is used on the TGD with \( k = 5 \).

Various kernel are been tested with the k-fold cross-validation: linear, RBF and sigmoid kernels. The best results were obtained by RBF kernel with \( C = 78.79 \) and \( \gamma = 0.088 \).

After the training phase, the test phase is carried out as shown in Fig. 2. For each sample, the method performs an estimation of water content. The “square” marker represents the real water content, whereas the “cross” marker represents the estimated water content.

Fig. 2 shows the \( \hat{W}_c \) from the TTB. The results show that the \( \hat{W}_c \) are generally well estimated. The SVR method estimates quite correctly the water content with a relatively low error, around 1 % in average.

From all these results we can deduce the good correspondence between experimental data and the SVR inverse model data. They also highlight the capability of the 4p Jonscher’s model to extend the frequency bandwidth at low frequencies where dispersion phenomena appear (Maxwell-Wagner effects) which allow us to take it into account in the SVR processing.

As a consequence, we can consider that the EM cell associated with the Jonscher’s extrapolation constitutes a valuable means to obtain the calibration curve relating the dielectric constant with the water content on cores.

5. Conclusions

In this paper, we have studied an innovative method for estimating the volumetric water concrete in various concrete mixtures. A statistical resolution of the inverse problem was done using a non-deterministic optimization (machine learning algorithms) of the in-lab calibration curves that link the controlled water content in various concrete mixtures and the complex dielectric permittivity issued from EM characterization device. An extrapolation procedure using the 4p variant of Jonscher’s model was developed and validated for the estimation of the complex permittivity at very low and very high frequencies. The use of this extrapolation method allows us to take into account various physical phenomena (i.e. polarizations vs.
water content) affecting in general the dielectric behaviour of concrete according to the frequency. 
The two-step estimation procedure (extrapolation and SVR methods) proposed in this paper is validated on a broad array of concrete specimens hydrically controlled in laboratory. This allows us to build calibration curves relying the complex effective permittivity and the water content in various concretes and taking into account the frequency dependence. Although the prediction of the volumetric water content obtained via the proposed method is very satisfying (maximum error<3.5% and mean error around 1 %) in front of needs expressed by the infrastructure owners, it is essential to optimize the inverse approach with a new settings for kernels taking into account various EM wave dispersion phenomena associated with polarizations of free ions present in the concrete interstitial solution.

References