Physica B: Condensed Matter 550 (2018) 207–216

Contents lists available at ScienceDirect

Physica B: Condensed Matter

journal homepage: www.elsevier.com/locate/physb

Predictors Generation by Partial Least Square Regression for microwave characterization of dielectric materials

H. Sadou[a,*], T. Hacib[a], Y. Le Bihan[b], O. Meyer[b], H. Acikgoz[c]

[a] Laboratoire L2EI, Faculté des sciences et de la technologie, Université de Jijel, Jijel, Algeria
[b] Laboratoire GeePs, CentraleSupelec, UMR 8507 CNRS, Sorbonne Université, Université Paris-Sud, Gif-sur-Yvette, France
[c] KTO Karatay University, Konya, Turkey

ARTICLE INFO

Index Terms:
Microwave characterization
Open-ended coaxial line
Complex dielectric permittivity
Finite elements method
Extended ***X*** bloc method
PG-PLSR

ABSTRACT

In this paper, the microwave characterization of dielectric materials using open-ended coaxial line probe is proposed. The measuring cell is a coaxial waveguide terminated by a dielectric sample. The study consists in extracting the real and imaginary part of the relative dielectric permittivity ($\varepsilon = \varepsilon' - j\varepsilon''$) of the material under test from the measurements of the probe admittance ($Y_{mes}(f) = G_{mes}(f) + jB_{mes}(f)$) on a broad band frequency (f between 1 MHz and 1.8 GHz), hence a direct and inverse problems have to be solved. In order to build a database, the direct problem is solved using Finite Elements Method (FEM) for the probe admittance ($Y(f) = G(f) + jB(f)$). Concerning the inverse problem, Partial Least Square (PLS) Regression (PLSR) is investigated as a fast, simple and accurate inversion tool. It is a dimensionality reduction method which aims to model the relationship between the matrix of independent variables (predictors) ***X*** and the matrix of dependant variables (response) ***Y***. The purpose of PLS is to find the Latent Variables (LV) having the higher ability of prediction by projecting original predictors into a new space of reduced dimension. The original inverse model has only three predictors (f, G and B) but is nonlinear, so inspired by the extended ***X*** bloc method, more predictors have been created mathematically from the original ones (for example: $1/f^2$, B/f^2, GB, $1/B$, G/f, f^2G, fG^2B, $f^2G^2B^2$, … etc) in order to take into account the nonlinearity, whence the appellation Predictors Generation Partial Least Square Regression (PG-PLSR). Inversion results of experimental measurements for liquid (ethanol, water) and solid (PEEK (Polyether-ether-ketone)) samples have proved the applicability and efficiency of PG-PLSR in microwave characterization. Moreover, the comparison study in the last section has proved the superiority of PG-PLSR on Multi-Layer Perceptron Neural Network (MLP-NN) in terms of rapidity, simplicity and accuracy.

1. Introduction

Nowadays, measurement of the dielectric and magnetic properties of lossy solids and liquids becomes more and more interesting. Usually, material characterization is used in dielectric measurements of biological tissues for cancer research, building materials, negative index materials, electromagnetic shielding and propagation of wireless signals. Many methods are developed for measuring electromagnetic permeability and permittivity. These techniques contain free-space methods, open-ended coaxial-probe techniques, cavity resonators, dielectric-resonator techniques, and transmission line techniques [1].

A review of extensive research work in the area of microwave characterization of dielectric materials reveals that the transmission line reflection method appears to be an efficient and convenient approach, especially when used with an open-ended coaxial line probe. This technique is based on inserting a sample of the material to be tested at the extremity of the coaxial waveguide, and the material permittivity is then extracted from the measurement of the reflection coefficient or the measuring cell admittance [2].

Inverse problems for the determination of dielectric materials properties are primarily solved by iterative methods which involve fast analytical solutions of the forward problem [3–6]. However physical phenomena are rarely solved by analytical methods hence numerical solutions became necessary and iterative inversion can became very time consuming. As a consequence, methods based on artificial intelligence such as Artificial Neural Networks (ANN) [7,8], Support Vector Machines (SVM) [9–11] and Adaptive Network based Fuzzy Inference System (ANFIS) [12] have occurred. Although, these inversion methods give, after training, a real time response, a prior optimization study is necessary. Concerning ANN, the determination of hidden layers and hidden

* Corresponding author.
E-mail address: sshakimjijel@gmail.com (H. Sadou).

https://doi.org/10.1016/j.physb.2018.08.037
Received 11 March 2018; Received in revised form 20 August 2018; Accepted 21 August 2018
Available online 07 September 2018

neurons is very important to avoid the over-fitting phenomenon. This problem is often fixed using Cross Validation (CV) which is a time consuming procedure. In addition, it needs the adjustment of the parameters of the back propagation learning algorithm, and the problem of small samples impedes the generalization ability of ANN modelling in engineering applications. Regarding SVM, a prior study involving an optimization algorithm is required to determine its hyper parameters. In the case of ANFIS, first of all, the number of Membership Functions (MF) per input, the MF type and the parameters of the learning algorithm have to be fixed. Furthermore, the problem of small samples impedes the generalization ability of ANN and SVM modelling in engineering applications [7–11], contrary to ANFIS which needs a relatively small database (comparable to that used here with PG-PLSR).

Nowadays, electromagnetic characterization of materials has become of strong interest, and finding faster techniques is more and more occupant for researchers. In this paper, Partial Least Square Regression (PLSR) as a fast inversion tool is invested. It models the relationship between a multivariate response and predictors by extracting the Latent Variables (LV) using the projection into a new space of reduced dimension. The forward problem is solved using Finite Elements Method (FEM) for the complex cell admittance ($Y(f) = G(f) + jB(f)$) on a broad-band frequency f (from 1 MHz to 1.8 GHz) involving a dielectric sample of a complex permittivity ($\varepsilon = \varepsilon'\text{-}j\varepsilon''$) to build a learning database. Therefore, in the inverse problem the three vectors (f, G and B) are the predictors, whereas ε' and ε'' are the responses. In order to integrate the nonlinearity of predictors (f, G and B) compared to responses (ε' and ε'') and to take into account the possible correlated influence predictors, additional predictors are generated mathematically from the original ones (f, G and B) following a technique inspired by the extended ***X*** bloc method [13], i.e. $1/f^2$, B/f^2,G^2, fB, $1/f$, 1/B, 1/G, B^2/G^2, f^2G, fG^2B, $f^2G^2B^2$, … etc, whence the appellation Predictors Generation Partial Least Square Regression (PG-PLSR) appears, hence the problem becomes suitable for PLSR known as a powerful method for predicting a set of dependent variables from a large set of independent variables (even noisy and highly correlated) [14].

2. Measurement setup and numerical method

The characterization cell, called SuperMit, consists in a junction between a coaxial waveguide and a circular waveguide which is filled by the dielectric material under test and short-circuited at its end (electric wall). In order to study liquids, a coaxial tight window is inserted between the two guides. The whole device is connected to a network analyser (Fig. 1).

The measurement protocol SuperMit was first developed by Belhadj-Tahar et al. [3] for isotropic and homogeneous dielectrics. The protocol requires the propagation of a Transverse Electromagnetic (TEM) mode in the circular coaxial line, from the analyser to the discontinuity (plane of measurement), where the wave is reflected by the coaxial guide-circular guide discontinuity. In the case of liquid samples, the permittivity and the thickness of the tight window are selected in such a way that only incident and reflected TEM modes exist at the interface coaxial guide-window. In order to avoid a high reflexion at the interface coaxial guide-window, the real part of window permittivity must be as low as possible. Meanwhile, to avoid losses in the window, the imaginary part must be also low. Here the probe is held so that the liquid sample is placed in the upper side to avoid air bubbles effect [7]. This configuration may be used from low frequencies until 19.6 GHz (cut-off frequency) using the coaxial waveguide APC7 (inner diameter = 3.04 mm, outer diameter = 7 mm).

The forward problem is expressed in terms of the electric field $\overrightarrow{E}$ which satisfies the following harmonic wave equation:

$$\nabla \times \left(-\frac{1}{j\omega\mu} \nabla \times \overrightarrow{E} \right) = j\omega\varepsilon\overrightarrow{E} \tag{1}$$

where ω is the pulsation, $\varepsilon(\varepsilon = \varepsilon_0(\varepsilon' - j\varepsilon''))$ and μ are the permittivity and the permeability, respectively. ε_0 is the permittivity of the free space, and ε', ε'' are respectively real and imaginary part of the relative permittivity.

Thanks to the axial symmetry of the system, only an angular sector (5°) of the geometry is meshed (Fig. 2). Second-order 3D tetrahedral edge finite elements are used. The system is meshed so that there are at least 10 elements per wavelengths.

At the input of the coaxial waveguide, a port of excitation was applied. It allows generating an incident TEM electromagnetic field and at the same time to absorb the same reflected mode. Perfect electric conductor (PEC) boundary conditions are applied on the surface of the waveguide conductors and at the end of the measurement cell, whereas, a perfect magnetic conductor (PMC) boundary condition is applied to both symmetry planes. These boundary conditions are given by:

$\vec{n} \times \overrightarrow{E} = 0$, on conductors.

$\vec{n} \times \overrightarrow{E} = \vec{n} \times \overrightarrow{E_0}$, on the excitation plane with $\overrightarrow{E_0}$a source field.

$\vec{n} \times \overrightarrow{H} = 0$, on symmetry planes (cutting planes).

The relationship which links the complex admittance Y to the reflexion coefficient at the discontinuity Γ is:

$$Y = Y_0 \frac{1-\Gamma}{1+\Gamma} = G + jB \tag{2}$$

where $Y_0 = 0.02S$ is the characteristic admittance of the coaxial line (50Ω).

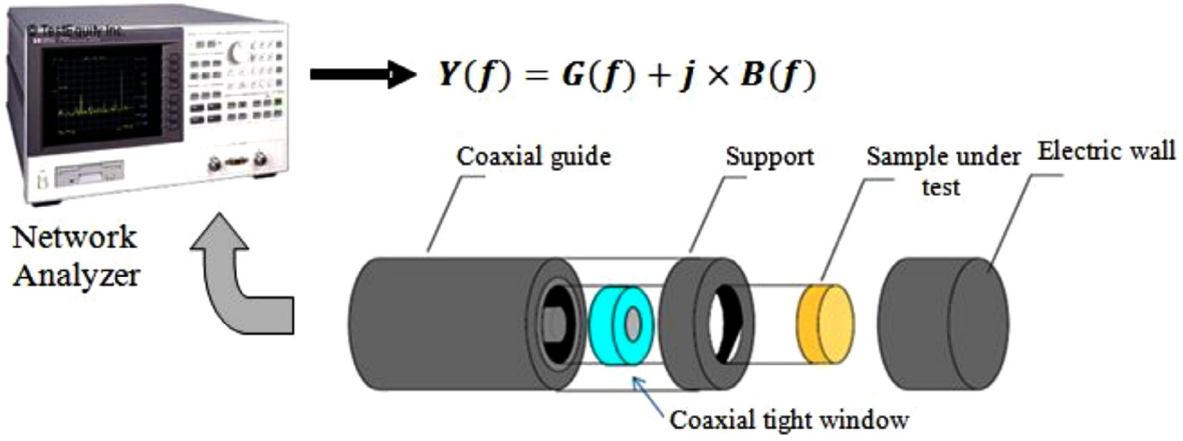


Fig. 1. Measurement setup with SuperMit cell for liquids.

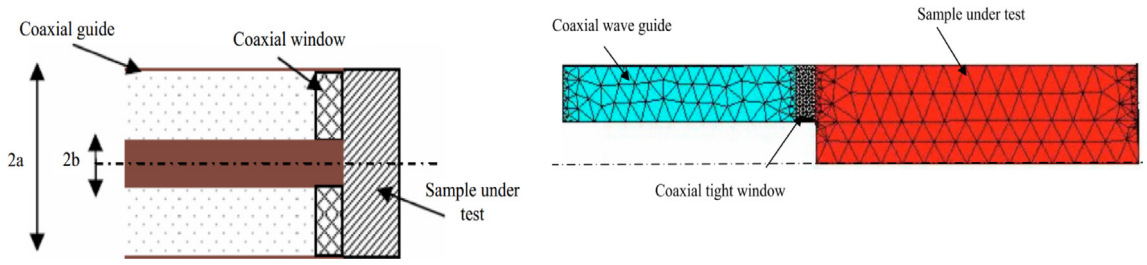


Fig. 2. SuperMit measuring cell for liquids, (a) 2D geometry, (b) cut of the meshed geometry.

3. PLSR for inverse problem solving

PLS was initially developed by Wold [15] and became popular in the fields of chemometrics, physiology, marketing, geophysics, supervision … etc. PLSR is a technique that combines the main features from Principal Component Analysis and Multiple Linear Regressions. The PLSR model is used to establish a relationship between a predictors matrix $\boldsymbol{X}$ ($x_1, x_2, \ldots, x_P$) and a response matrix $\boldsymbol{Y}$ ($y_1, y_2, \ldots, y_R$), where: x_P is the P^{th} independent variable put in the form of a column vector of N observations, y_R is the R^{th} dependent variable put in the same form as x_P.

The prediction is achieved by extracting from $\boldsymbol{X}$ a set of orthogonal (uncorrelated) components called latent variables (or principal components) which have the best predictive power by maximizing the covariance between the projection of the predictors in a subspace of reduced dimension and the responses. In other words, PLS finds the best axes that simultaneously explain variations in $\boldsymbol{X}$ and predict $\boldsymbol{Y}$ [16,17].

The power of PLSR is seen when dealing with multivariate response problems ($\boldsymbol{Y}$ is a multidimensional output matrix) with noisy data and correlated independent variables ($\boldsymbol{X}$) under limited number of observations (even less than predictors number) [14,18,19]. Additionally, it automatically performs variable selection that is easy to implement, statistically very efficient and computationally very fast, which renders it practical for application to large data sets [20].

The model accuracy is strongly depending on the latent components retained. When the latter is insufficient, some useful information will be lost (under-fitting). On the other hand, if we take an over-many number of components, the model risks over-fitting phenomena and brings an unnecessary noise signal which affects the robustness of the model [19,21]. Usually, CV is adopted to fix this problem [22].

Theoretically, PLS method assumes the existence of a linear relationship between independent variables $\boldsymbol{X}$ and dependant variables $\boldsymbol{Y}$. Let us consider a given predictors matrix $\boldsymbol{X}$ ($N \times P$) of N observations parameterized by P parameters (inputs) and an output variable $\boldsymbol{Y}$ ($N \times R$) assumed to be both fitted by a linear combination of a reduced number n of LV.

PLS decomposes $\boldsymbol{X}$ and $\boldsymbol{Y}$ into the following form [16,20]:

$$\boldsymbol{X} = \boldsymbol{T}\boldsymbol{P}' + \boldsymbol{E} \tag{3}$$

$$\boldsymbol{Y} = \boldsymbol{T}\boldsymbol{Q}' + \boldsymbol{F} \tag{4}$$

In this PLS formulation, the predictors and responses are controlled by a single score matrix $\boldsymbol{T}$, just such formulation is suitable for regression problems where the predictors are high dimensional but the response is low dimensional ($R \ll P$) or even univariate ($R = 1$).Where $\boldsymbol{T}$ is an ($N \times n$) matrix of the n extracted score vectors (components, latent vectors) for both predictors and responses, the ($P \times n$) matrix $\boldsymbol{P}$ and the ($R \times n$) matrix $\boldsymbol{Q}$ represent matrices of loadings, and the ($N \times P$) matrix $\boldsymbol{E}$ and the ($N \times R$) matrix $\boldsymbol{F}$ are the matrices of residuals.

The first LV (first column of $\boldsymbol{T}$) $t_1 = \boldsymbol{X}r_1$ is extracted by maximizing $r'_1\boldsymbol{X}'\boldsymbol{Y}\boldsymbol{Y}'\boldsymbol{X}r_1$ with the constraint $||r_1|| = 1$. This leads to calculate the dominant eigenvector r_1 of the matrix $\boldsymbol{X}'\boldsymbol{Y}\boldsymbol{Y}'\boldsymbol{X}$. In the same manner, the second LV (second column of $\boldsymbol{T}$) $t_2 = \boldsymbol{X}r_2$ is determined by maximizing $r'_2\boldsymbol{X}'\boldsymbol{Y}\boldsymbol{Y}'\boldsymbol{X}r_2$ with the double constraint $||r_2|| = 1$ and mutually orthogonal score vectors $t'_2t_1 = 0$.

The procedure is repeated until extracting all LV: $t_k = \boldsymbol{X}r_k$ ($1 < k < n$), and the weight matrix $\boldsymbol{R}$ constituted of vectors r_k is used to establish the projection of $\boldsymbol{X}$ into the new space of reduced dimension: $\boldsymbol{T} = \boldsymbol{X}\boldsymbol{R}$. After that, the loading matrix $\boldsymbol{Q}$ is computed using an ordinary least squares regression of $\boldsymbol{Y}$ on $\boldsymbol{T}$: $\boldsymbol{Q} = (\boldsymbol{T}'\boldsymbol{T})^{-1}\boldsymbol{T}'\boldsymbol{Y}$.

Finally, the complete PLSR model can be established from equation (4) assuming that the error projection is minimum (close to zero) so that the residual matrix ($\boldsymbol{F}$) is negligible as:

$$\boldsymbol{Y} = \boldsymbol{X}\boldsymbol{B}_{\mathbf{PLS}} + \boldsymbol{F} \tag{5}$$

with $\boldsymbol{B}_{PLS} = \boldsymbol{R}\boldsymbol{Q}'$ is the PLS regression coefficient matrix of $P+1$ rows and R columns.

There are a number of algorithms proposed for the practical estimation of PLS coefficients (t_k, r_k, … etc) such as SIMPLS and NIPALS. In the present work, the SIMPLS algorithm is used since it ensures lower computational load and faster convergence [23,24].

Usually, the number of extracted score vectors n is limited by the convergence criterion of the iterative procedure. In this paper, the maximum number of LV (n) is chosen to be the same as the number of predictors (P) and the inversion of experimental measurements will be done using an optimal number of LV fixed by CV procedure.

4. Database generation and preprocessing

The purpose of this paper is to elaborate an inverse model based on PLSR which links the complex permittivity of the sample under test to the probe admittance measurements. Therefore a database is created by solving numerically (FEM) the forward problem to get the measuring cell admittance ($Y(f) = G(f) + jB(f)$) in function of frequency (f) for a given material of a complex relative permittivity: $\varepsilon' - j\varepsilon''$. Consequently, PLSR inputs are G, B and f, outputs are ε' and ε''. Database is created by sweeping all possible values of the relative dielectric permittivity of the sample under test, whereas the measurement frequency is from 1 MHz to 1.8 GHz. The learning set elements (ε', ε'') are selected in such a way that they are uniformly distributed on the domain of interest, while the frequency points are distributed following to a logarithmic law. Before, entering the database generation routine, all the N values of the generated vectors ε', ε'' and f are permuted randomly.

Two databases are created; one to elaborate a PLSR model able to predict the permittivity of liquid samples (ethanol and water) by solving the direct problem involving the SuperMit cell with coaxial tight window, and another for solid samples (SuperMit cell without coaxial tight window). For liquid samples, the database is created by varying ε' between 1 and 100, whereas ε'' is from 0 to 100. In the case of solid samples, the material under test (PEEK) has a low relative permittivity (especially for imaginary part), consequently, a denser distribution of the examples is required, so the permittivity range covered by the data set is reduced ($1 \leq \varepsilon' \leq 10$ and $0 \leq \varepsilon'' \leq 10$).

The procedure of database generation is illustrated by Table 1.

Table 1
Procedure of database generation.

N° of xample	Inputs of the direct problem				Outputs of the direct problem
	ε'	ε''	f	G	B
1	ε'_1	ε''_1	f_1	G_1	B_1
2	ε'_2	ε''_2	f_2	G_2	B_2
…	…	…	…	…	…
…	…	…	…	…	…
N	ε'_N	ε''_N	f_N	G_N	B_N
	outputs of the inverse problem		Inputs of the inverse problem		

Matrices $\boldsymbol{X}$ and $\boldsymbol{Y}$ are written as follows:

$$\boldsymbol{X} = \begin{bmatrix} f_1 & G_1 & B_1 \\ f_2 & G_2 & B_2 \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ f_N & G_N & B_N \end{bmatrix}_{(N\times 3)} \quad \boldsymbol{Y} = \begin{bmatrix} \varepsilon'_1 & \varepsilon''_1 \\ \varepsilon'_2 & \varepsilon''_2 \\ \cdot & \cdot \\ \cdot & \cdot \\ \varepsilon'_N & \varepsilon''_N \end{bmatrix}_{(N\times 2)}$$

Where N is the number of observations (examples).

Inputs and outputs of the databases present a great difference in numerical values; G and B are very low, f *is* very high whereas ε' and ε" are middle values. To avoid the impact of numerical representation and make each variable to play the same role, each variable is scaled to unit variance by dividing them by their standard deviations, and centred to zero mean by subtracting their averages (for all columns of $\boldsymbol{X}$ and $\boldsymbol{Y}$).

5. Predictors Generation Partial Least Square Regression (PG-PLSR)

PLS is particularly useful when predicting a set of dependent variables from a (very) large set of independent variables (predictors) [18]. The aim of PLS is the dimensionality reduction by iterative deflation. In other words, replacing correlated predictors by some LV which maximize the covariance with Y. Given the nonlinear relationship between the inputs (f, G and B) and the outputs (ε' and ε"), it is obvious that a linear model like PLSR is unable to fill this task with these three input variables. So, the use of a non-linear PLSR model is obligatory.

There are lot of methods to integrate the non-linearity in PLSR [25–28]. In this paper, a technique inspired by the extended $\boldsymbol{X}$ bloc method is used. Proposed by Berglund and Wold [13], this method consists of expanding the predictors matrix $\boldsymbol{X}$ with higher-order terms, e.g. quadratic, cross-product and sometimes cubic terms. The expanded set of $\boldsymbol{X}$-variables is then used as predictors of $\boldsymbol{Y}$ and the model is fitted by ordinary PLSR. Since our problem has only three predictors (f, G and B), the extended $\boldsymbol{X}$ bloc method has led to poor results.

Now the idea is to expand the $\boldsymbol{X}$ bloc differently, where the number of predictors (columns of $\boldsymbol{X}$) is raised mathematically by adding new combinations of original inputs (f, G and B), in order to create a model able to take into account the non-linearity and the possible correlations between predictors. Therefore, the appellation PG-PLSR appears.

Each column vector $\{x_m\}_{m=1}^{P}$ of the new matrix of predictors is written as follows:

$$x_m = f^i G^j B^k,\ i = -2,\ \ldots,2,\ j = 0,\ \ldots,2,\ k = 0,\ \ldots,2,\ \text{(except the combination when } i = j = k = 0) \tag{6}$$

Consequently, a total of 44 new predictors is created. For example:$1/f^2$, B/f^2, B^2/f^2, …, G^2, B^2, fB, …, $1/f$, 1/B, 1/G, …,B^2/G^2,f^2G, fG^2B, …,$f^2G^2B^2$.

Now, the new matrix of predictors $\boldsymbol{X}$ takes the following form ($\boldsymbol{Y}$ keeps the same shape):

$$\boldsymbol{X} = \begin{bmatrix} (1/f^2)_1 & (B/f^2)_1 & \cdots & (B)_1 & \cdots & (G)_1 & \cdots & (f^2G^2B)_1 & (f^2G^2B^2)_1 \\ (1/f^2)_2 & (B/f^2)_2 & \cdots & (B)_2 & \cdots & (G)_2 & \cdots & (f^2G^2B)_2 & (f^2G^2B^2)_2 \\ \vdots & \vdots & \cdots & \vdots & \cdots & \vdots & \cdots & \vdots & \vdots \\ (1/f^2)_N & (B/f^2)_N & \cdots & (B)_N & \cdots & (G)_N & \cdots & (f^2G^2B)_N & (f^2G^2B^2)_N \end{bmatrix}_{(N\times 44)}$$

6. Implementation of PG-PLSR

Before implementing the PG-PLSR, the data set is randomly divided into two different subsets; a calibration set (2/3) used for developing the calibration model and performing the CV, and a prediction (test) set (1/3) used for assessing the performances of the PG-PLSR model [29].

The calibration set is divided into 10 equal subsets. The PG-PLSR model is implemented for different numbers of LV for 10 times involving 9 subset as training set and checked by the 10th one (validation set) until all subsets were left out at once. Therefore, 180 examples are used for training, 20 for validation and 100 for testing. The optimal number of LV is the one corresponding to the min average (on 10 subsets) of the Cross Validation-Root Mean Squared Error (*CV-RMSE*) computed on the validation set, or the one after which the *CV-RMSE* does not show a significant decrease. Besides the Root Mean Squared Error (*RMSE*) (Equation (7)) which indicates the accuracy of the prediction, the performance of the regression methods is also measured by the coefficient of determination R^2 (Equation (8)) quantifying the strength of the model [30].

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N} \|y_i - \hat{y}_i\|^2} \tag{7}$$

$$R^2 = \frac{\left(\sum_{i=1}^{N}(y_i - \overline{y}_i)(\hat{y}_i - \overline{\hat{y}}_i)\right)^2}{\sum_{i=1}^{N}\|y_i - \overline{y}_i\|^2 \sum_{i=1}^{N}\|\hat{y}_i - \overline{\hat{y}}_i\|^2} \tag{8}$$

where y_i is the value of material properties, $\hat{y}_i$ is the value of the PG-PLSR prediction, $\overline{y}_i$ and $\overline{\hat{y}}_i$ are the mean of y_i and $\hat{y}_i$ respectively and N is the number of examples in the data set.

In what follows, we define: the Root Mean Squared Error of Calibration (*RMSEC*), the Root Mean Squared Error of Prediction (*RMSEP*), the determination coefficient of calibration (R^2_{cal}) and the determination coefficient of prediction (R^2_{pred}). Good results means a very low *RMSEP* and R^2_{pred} being approximately one.

7. Results and discussions

7.1. Application on simulated data

As discussed previously, latent component number has a crucial role on the accuracy of PG-PLSR model. Fig. 3 shows the average *CV-RMSE* evolution on the 10 validation sets in function of the number of LV retained using the database dedicated for liquid samples (curve with *) and the one for solid samples (curve with circles). Following this graph, 26 and 29 components have been taken to provide simultaneously ε' and ε'' ($R = 2$) for solid and liquid samples, respectively. Beyond these numbers, the two models enter clearly in the over-fitting phenomenon. In addition, it can be observed the high level of prediction (very low *CV-RMSE*) reached with reduced number of LV.

In order to illustrate the choice of LV number, the ratio of the variance of $\boldsymbol{Y}$ dividing the one of its PLS approximation ($\boldsymbol{TQ'}$), i.e. Var ($\boldsymbol{TQ'}$)/Var($\boldsymbol{Y}$), in function of the number of LV retained is shown in Fig. 4. The graph presents two similar curves showing that the maximum (≅100%) of the explained variance in $\boldsymbol{Y}$ is widely explained by the LV retained, before, using CV procedure for liquid (29 LV) and solid (26 LV) samples, and more than 97% of the variance is explained by the first five LV for both cases. Moreover, from Fig. 5, the explained variance in $\boldsymbol{X}$ (Var ($\boldsymbol{TP'}$)/Var ($\boldsymbol{X}$)) realized by PLS analysis has reached its

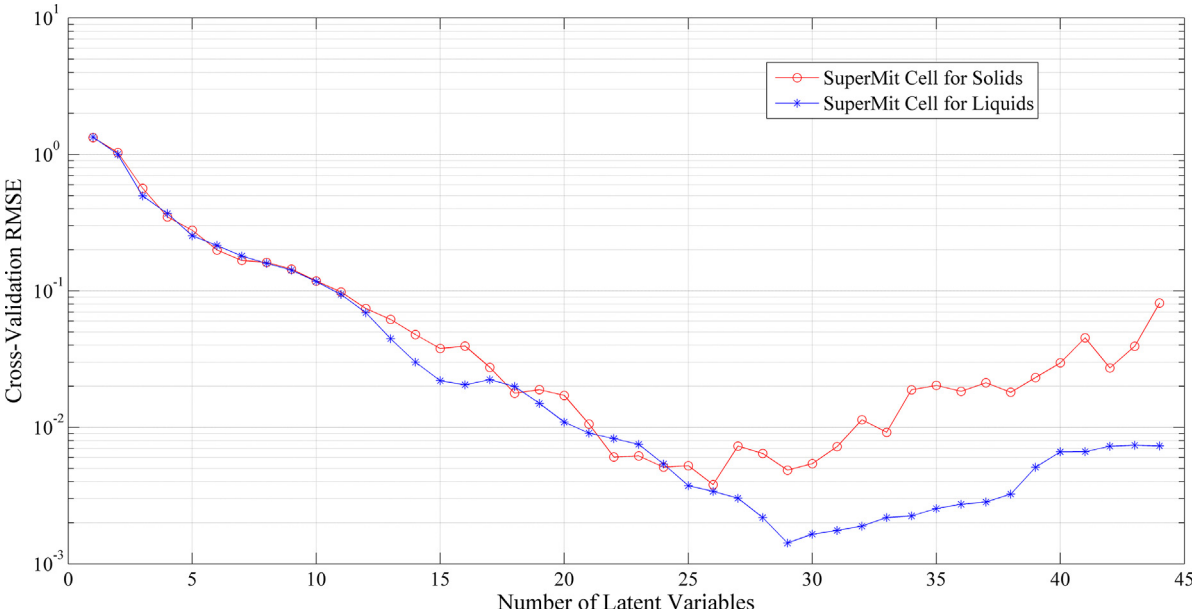


Fig. 3. The average of *RMSE* on the validation sets in function of the number of PLS components.

maximum (≅100%) using the optimal number of LV for both cases, beyond this number the contribution of high ranked components is negligible and they could be largely corrupted by noise when considering measurements. Usually, this kind of figure gives an idea of the LV which can be reasonably extracted with the PLS procedure, contrary to Fig. 4 which indicates the appropriate number of LV required for measurements inversion.

In order to assess the performances of the PG-PLSR models, these latter is applied to predict the permittivity values included in the test set for liquids and solids. In Figs. 6 and 7 the predicted values of the permittivity (outputs of PG-PLSR model) is depicted in function of the desired ones (true values) for real and imaginary parts. The points in each figure represent a straight line with a slope of one, which means they match well with a function of equation: $\hat{y} = y$ (estimated permittivity = desired permittivity). Moreover, the *RMSEP* and R^2_{pred} calculated on the test sets (Table 2) indicate the very good accuracy obtained. In addition, some researchers consider a good model when *RMSEC* and *RMSEP* are very closes which is the case here [29].

Table 2 summarizes the performances of each model.

7.2. Application on experimental data

Since the PG-PLSR model has proved its accuracy on simulated data, experimental data can be inverted. Measurements have been carried out by using an impedance analyser Agilent 4291Aon a dielectric sample whose dielectric characteristics are known. The coaxial guide is an APC7 standard (a = 3.5 mm, b = 1.52 mm). The measurement frequency band is from 1 MHz to 1.8 GHz. The thickness of the sample under test is 13.8 mm for liquids and 0.7 mm for solids. The tight coaxial window used to characterise liquids (ethanol and water) is made of Plexiglas and has a thickness of 0.82 mm and a relative dielectric constant of 2.7. Fig. 8 shows the experimental measurements of the real (G_{mes}) and imaginary part (B_{mes}) of the measuring cell admittance with the operating frequency f_{mes} for ethanol, water and PEEK. To avoid contact effects the solid samples are metallized on contact surfaces with line conductors.

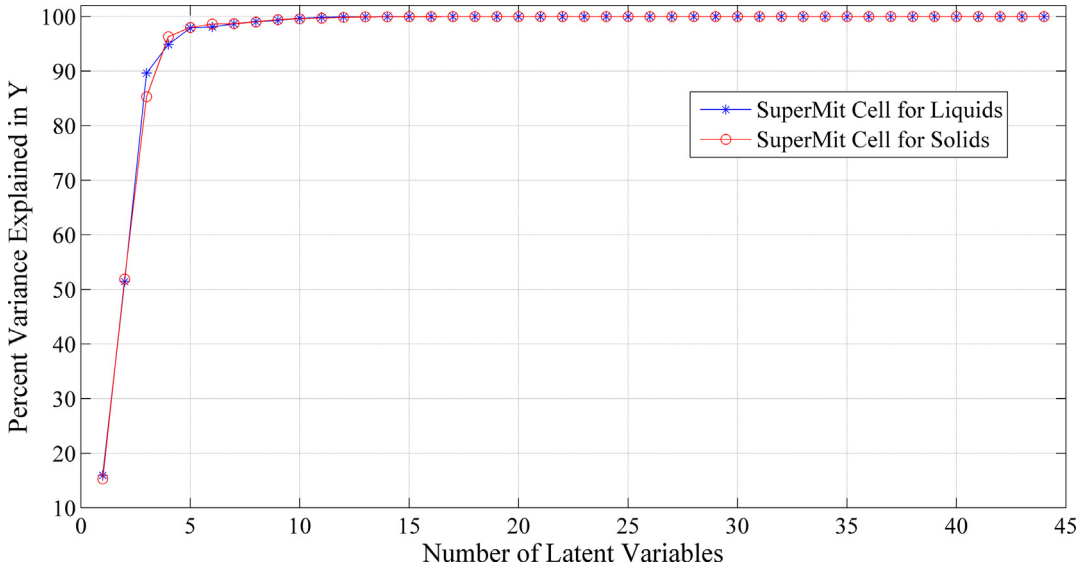


Fig. 4. Percentage of the variance of ***Y*** explained by the PLS analysis.

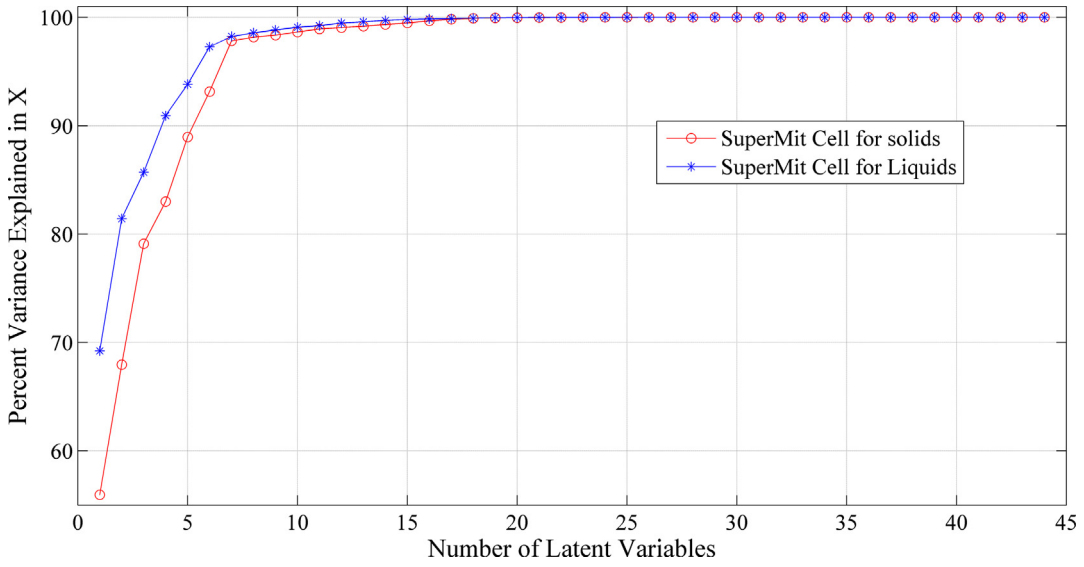


Fig. 5. Percentage of the variance of ***X*** explained by the PLS analysis.

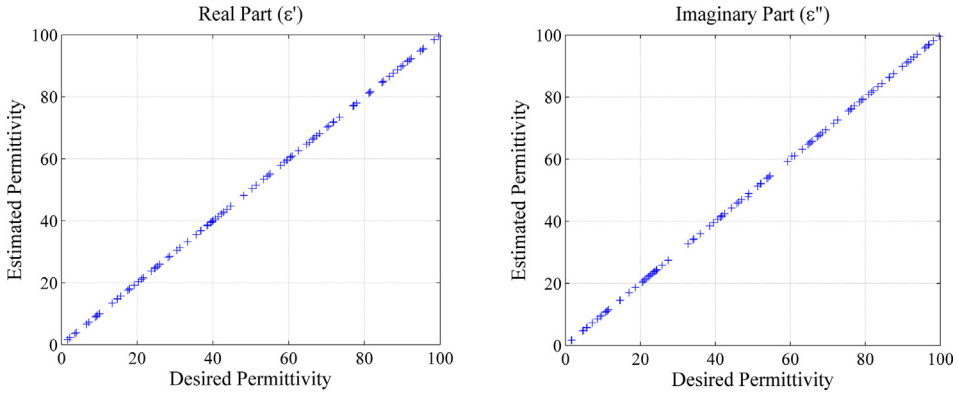


Fig. 6. Comparison between the permittivity provided by PG-PLSR and the permittivity contained in the test set (data base dedicated for liquids), (a) real part ε′, (b) imaginary part ε".

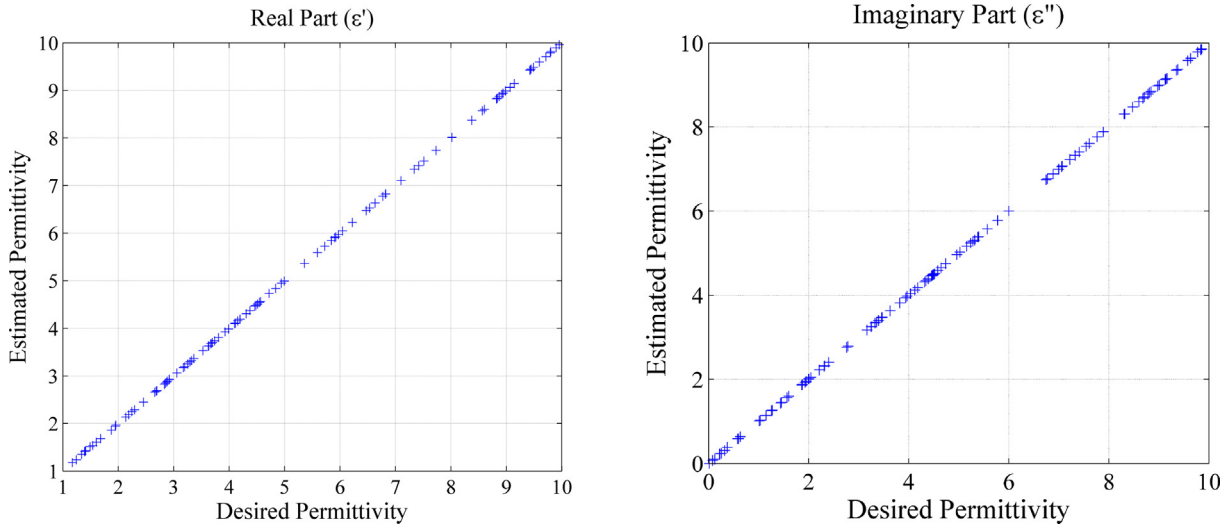


Fig. 7. Comparison between the permittivity provided by PG-PLSR and the permittivity contained in the test set (data base dedicated for solids), (a) real part ε′, (b) imaginary part ε".

Table 2
Performances of the two PG-PLSR models.

	Outputs	Calibration set (200 examples)		Test set (100 examples)	
		RMSEC	R^2_{cal}	*RMSEP*	R^2_{pred}
PG-PLSR for liquids	ε' and ε''	1.2354×10^{-2}	0.9999998	3.5334×10^{-2}	0.9999983
PG-PLSR for solids	ε' and ε''	5.0037×10^{-3}	0.9999966	6.0690×10^{-3}	0.9999953

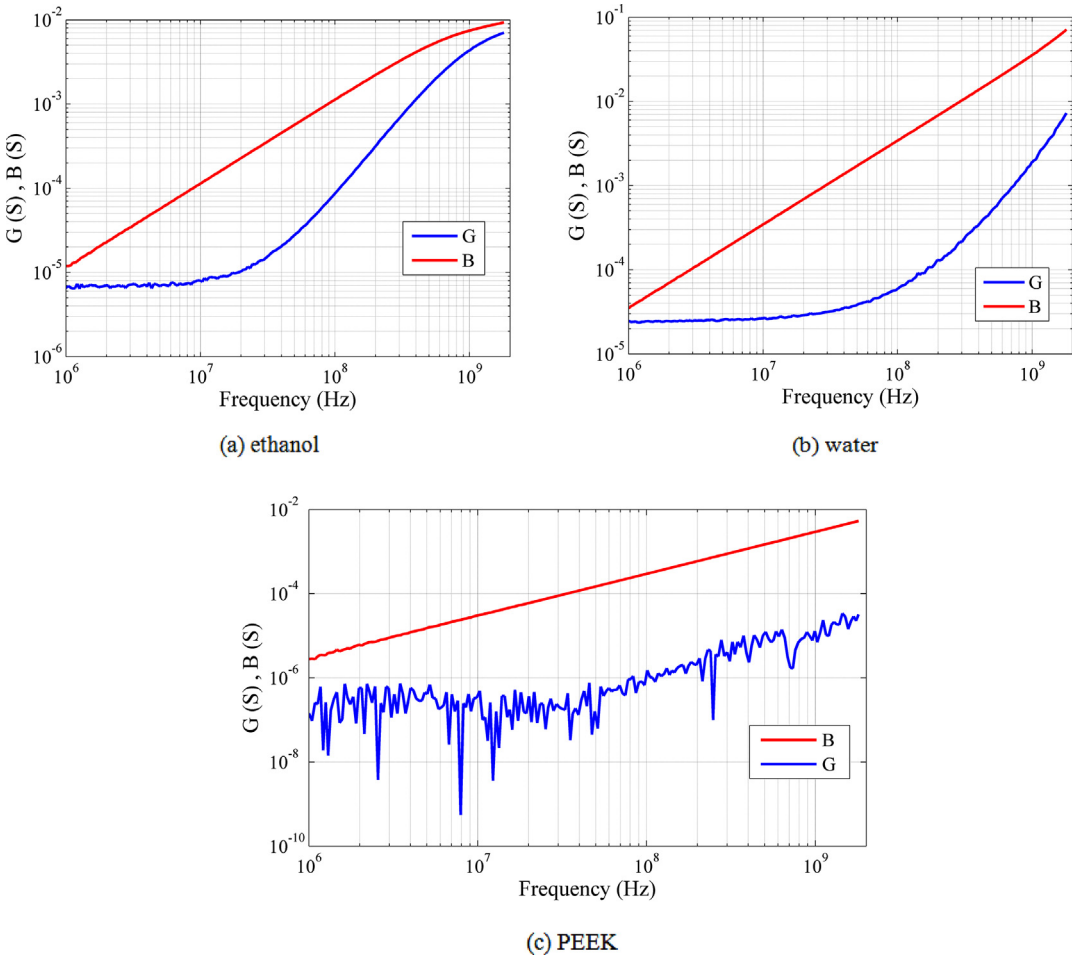


Fig. 8. Experimental measurements of real (G) and imaginary part (B) of the measuring cell admittance.

It should be noticed that the matrix of predictors resulting from measurements ($\boldsymbol{X}_{\mathbf{mes}}$) has gone through the same transformation as matrix $\boldsymbol{X}$ of simulated data (discussed in section 5). Hence, it changed from 3 to 44 centred and normalized predictors x_{mes} (columns of $\boldsymbol{X}_{\mathbf{mes}}$):

$$x_{\mathrm{mes}} = f_{\mathrm{mes}}^{i} G_{mes}^{j} B_{mes}^{k}, \; i = -2, \; ..., 2, \; j = 0, \; ..., 2 \text{ and } k = 0, \; ..., 2 \tag{9}$$

Here again, the combination when $i = j = k = 0$ is excluded.

Now, the new matrix of experimental predictors $\boldsymbol{X}_{\mathbf{mes}}$ has the following form:

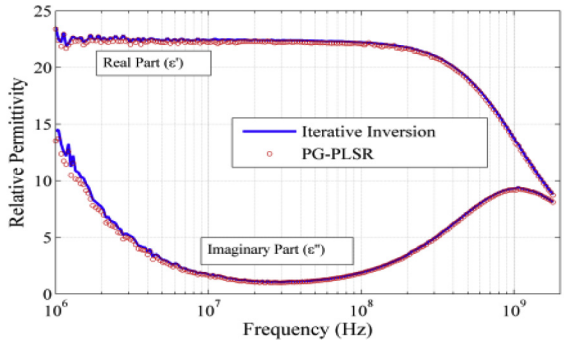
$$\boldsymbol{X}_{mes} = \begin{bmatrix} (1/f_{mes}^2)_1 & (B_{mes}/f_{mes}^2)_1 & \cdots & (B_{mes})_1 & \cdots & (f_{mes}^2 G_{mes}^2 B_{mes}^2)_1 \\ (1/f_{mes}^2)_2 & (B_{mes}/f_{mes}^2)_2 & \cdots & (B_{mes})_2 & \cdots & (f_{mes}^2 G_{mes}^2 B_{mes}^2)_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ (1/f_{mes}^2)_{N_{mes}} & (B_{mes}/f_{mes}^2)_{N_{mes}} & \cdots & (B_{mes})_{N_{mes}} & \cdots & (f_{mes}^2 G_{mes}^2 B_{mes}^2)_{N_{mes}} \end{bmatrix}_{(N_{mes} \times 44)}$$

Where N_{mes} is the number of measurement frequencies.

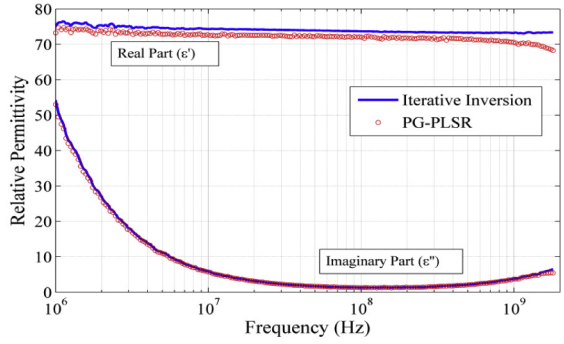
The reason why indexes j and k begin from 0, in contrary to index i, is to avoid big values of predictors when dividing by $G_{\mathrm{mes}}{}^2$, G_{mes}, $B_{\mathrm{mes}}{}^2$ and B_{mes} during the phase of predictors generation (Equations (6) and (9)) especially in the case of PEEK which is known as a lossless material (G nearly null), hence reduced sizes of predictors matrices and less number of LV to be computed, so PG-PLSR model becomes simpler and faster. Notice that all these 44 predictors are multiplied by $\boldsymbol{B}_{\mathbf{PLS}}$ to get the measurements inversion: $\boldsymbol{Y}_{\mathbf{mes}} = \boldsymbol{X}_{\mathbf{mes}}\boldsymbol{B}_{\mathbf{PLS}} + \boldsymbol{F}$, regardless the optimal number of LV retained by CV.

Fig. 9 shows the inversion results of experimental measurements using PG-PLSR for ethanol, water and PEEK besides those of a model-based iterative inversion. This last method consists in inserting a direct model in an iterative procedure. First, from a couple (ε', ε'') taken arbitrary, the cell admittance is calculated analytically according to the mode-matching method [3,4]. After that, an iterative method derived from the gradient method [5,6] is used to reduce the difference between the measured admittance and the calculated one. Finally, the couple (ε', ε'') that give the smallest difference is retained as solution for the considered frequency.

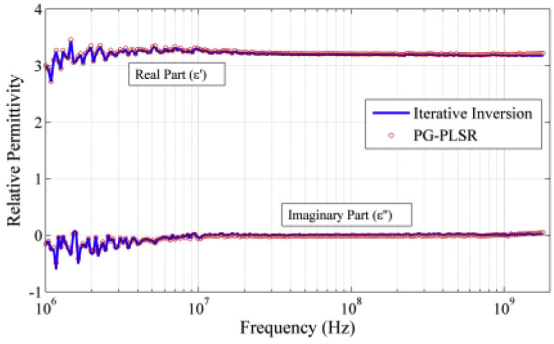
A good agreement between the two methods can be observed. Although the coincidence of results, PG-PLSR is very fast compared with iterative inversion, which is a computationally expensive method



(a) Ethanol



(b) Water



(c) PEEK

Fig. 9. Permittivity evolution obtained by PG-PLSR and Iterative inversion.

since it involves the multiple resolution of the forward problem [4–6]. At approximately 1.1 GHz, the graph of Fig. 9(a) shows the relaxation phenomenon of ethanol by opposition to Fig. 9(b) and (c) which don't show any relaxation since it appears at higher frequency (about 17 GHz) for water [5] and non-existent for PEEK. Moreover, inversion results fit well with theoretical low frequency permittivity, such as, in reference [7] the theoretical low frequency permittivity of water and ethanol at 25 °C are 79 and 24 respectively. Concerning PEEK known as a lossless material ($\varepsilon'' \approx 0$), the dielectric constant is 3.2–3.3 between 50 Hz and 10 kHz following reference [31]. Finally, results fit well with other literature references [7–11,32–35].

The slight difference between inversion results and theoretical values is due to the presence of impurities in liquids and experimental inaccuracies. Whereas, Dispersed and negative values at low frequencies for Fig. 9(c) is due to the very low values of the admittance (especially for the real part G which is the image of ε") combined with the inaccuracy of the analyser at frequencies below 10 MHz.

8. Comparison between PG-PLSR and MLP-NN

In order to prove the high performances of PG-PLSR in microwave characterization using open-ended coaxial line, a comparison study with MLP-NN is achieved. Both PG-PLSR and MLP-NN models are used to predict simultaneously ε' and ε'' ($R = 2$) of ethanol using the database dedicated for liquids. The comparison is achieved on the optimal PG-PLSR and MLP-NN models found by using CV procedure. The optimal 30 hidden neurons (one hidden layer) MLP-NN was trained using Levenberg-Marquardt algorithm [36,37] with a nonlinear activation function (logsig) and linear output function, while the optimal PG-PLSR model (29 LV) is the one described in section 7. Since the goal error cannot be imposed in PG-PLSR, the *RMSEC* found previously (Table 2) is used as a stopping criterion for MLP-NN.

The comparison results presented in Table 3 show the large differences in the total CPU time spent for the calculations and the generalization ability traduced by *RMSEP* and R^2_{pred}. Concerning the database size used by PG-PLSR, it's well known that PLSR can predict the responses even when observations are less than predictors, but the larger sample size can result in higher performance of the model [30]. Indeed, Table 3 shows that the larger sample size results in lower *RMSEP* and bigger R^2_{pred}. Meanwhile, from Fig. 10 it can be observed that 1/10 of the database used for calibrating MLP-NN model (200 examples) is widely sufficient for PG-PLSR, whereas an MLP-NN model with 200 examples (Fig. 10 (c)) is totally useless (without talking about a model MLP-NN with 45 examples).

In summary, a very fast PG-PLSR model with 200 examples (Fig. 10 (b)) and a very time consuming MLP-NN model with 2000 examples (Fig. 10 (e)) are very satisfactory with an explained variance of more than 99.99% ($R^2_{\text{pred}} \approx 1$). Moreover, reduced database size is a very important advantage in the cases where available data are limited such as modelling of survival in medical researches or in the cases where experimental measurements are required to create databases.

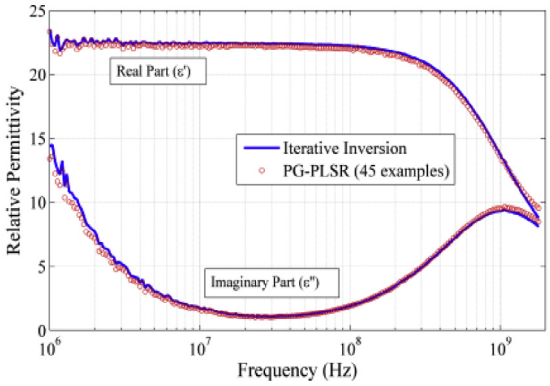
Table 3 summarizes the performances of each model.

9. Conclusion

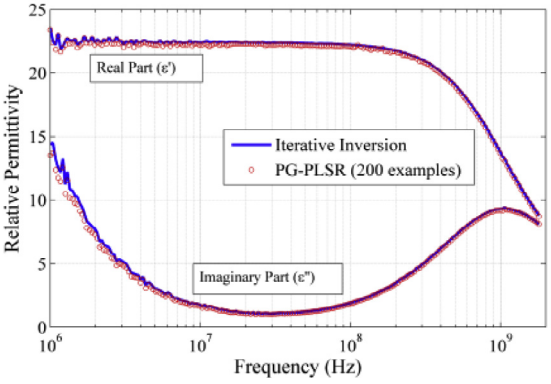
In this work, the applicability and efficiency of the PG-PLSR method for microwave characterization of dielectric materials is presented. Usually, PLS create a regression model linking a matrix of predictors $\boldsymbol{X}$

Table 3
Performances of the two models.

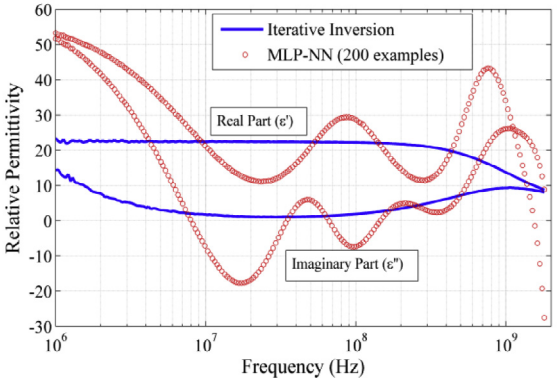
	CPU Time (s)	*RMSEC*	R^2_{cal}	Calibration set	*RMSEP*	R^2_{pred}	Test set
PG-PLSR	0.019	5.212×10^{-3}	0.9999999	45	4.132×10^{-1}	0.9997870	1000
	0.022	1.235×10^{-2}	0.9999998	200	5.196×10^{-2}	0.9999966	1000
	0.060	2.038×10^{-2}	0.9999994	2000	2.509×10^{-2}	0.9999992	1000
MLP-NN	4.59	7.2297	0.93674	200	17.833	0.71027	1000
	451	4.270×10^{-2}	0.9999977	2000	4.944×10^{-2}	0.9999970	1000



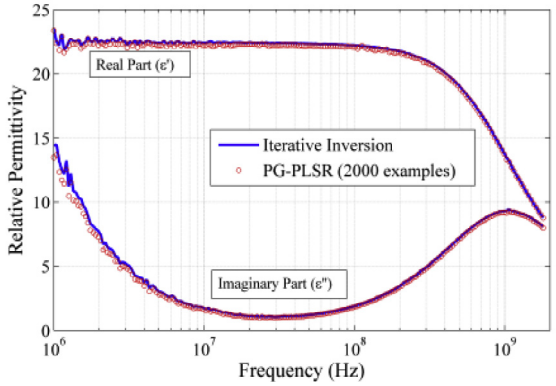
(a) PG-PLSR with 45 examples



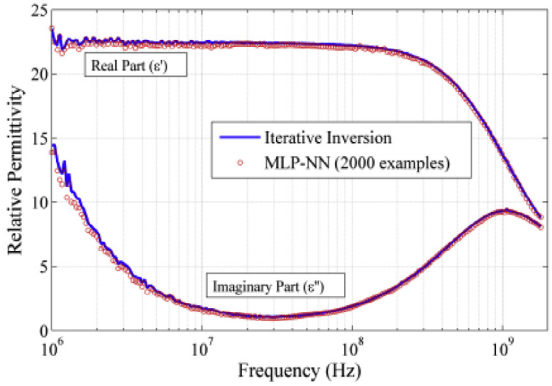
(b) PG-PLSR with 200 examples



(c) MLP-NN with 200 examples



(d) PG-PLSR with 2000 examples



(e) MLP-NN with 2000 examples

Fig. 10. Permittivity evolution obtained by PG-PLSR and MLP-NN.

and a matrix of responses $\boldsymbol{Y}$ using a dimensionality reduction procedure, in contrary to this paper where more predictors are first created mathematically from original ones, inspired by the extended $\boldsymbol{X}$ bloc method, before implementing PLS model. Results of measurements inversion show the power of PG-PLSR as a fast, simple and accurate inversion tool unless it requires more observations than predictors. PG-PLSR gives a real time response contrary to iterative inversion which is a very time consuming method. PG-PLSR is implemented with a relatively small database contrary to other inversion tools like ANN and SVM used in previous similar works [7–11]. Moreover, the comparison with MLP-NN shows its superiority in terms of rapidity, simplicity and accuracy for multivariate ($R = 2$) response prediction.

References

[1] J. Baker-Jarvis, M.D. Janezic, B. Riddle, R.T. Johnk, P. Kabos, C.L. Holloway, R.G. Geyer, C.A. Grosvenor, Measuring the Permittivity and Permeability of Lossy Materials: Solids, Liquids, Metals, Building Materials, and Negative-index Materials,

Technical report National Institute of Standards and Technology, 2004 NIST TN1536.
[2] H. Zheng, C.E. Smith, Permittivity measurements using a short open-ended coaxial line probe, IEEE Microw. Guid. Wave Lett. 1 (1991) 337–339.
[3] N.-E. Belhadj-Tahar, A. Fourrier-Lamer, Broad-band analysis of a coaxial discontinuity used for dielectric measurements, IEEE Trans. Microw. Theor. Tech. 34 (1986) 346–350.
[4] N.-E. Belhadj-Tahar, A. Fourrier-Lamer, H. De Chanterac, Broad-band simultaneous measurement of complex permittivity and permeability using a coaxial discontinuity, IEEE Trans. Microw. Theor. Tech. 38 (1990) 1–7.
[5] N. Belhadj-Tahar, O. Meyer, A. Fourrier-Lamer, Broad-band microwave characterization of bilayered materials using a coaxial discontinuity with applications for thin conductive films for microelectronics and material in air-tight cell, IEEE Trans. Microw. Theor. Tech. 45 (1997) 260–267.
[6] N.-E. Belhadj-Tahar, O. Dubrunfaut, A. Fourrier-Lamer, Broad-band microwave characterization of a tri-layer structure using a coaxial discontinuity with applications for magnetic liquids and films, IEEE Trans. Microw. Theor. Tech. 46 (1998) 2109–2116.
[7] H. Acikgoz, Y. Le Bihan, O. Meyer, L. Pichon, Neural network for broad-band evaluation of complex permittivity using a coaxial discontinuity, Eur. Phys. J. Appl. Phys. 39 (2007) 197–201.
[8] H. Acikgoz, Y. Le Bihan, O. Meyer, L. Pichon, Microwave characterization of dielectric material using bayesian neural networks, Prog. Electromagn. Res. C 3 (2008) 169–182.
[9] T. Hacib, H. Acikgoz, Y. Le Bihan, M.R. Mekideche, O. Meyer, L. Pichon, Support vector machines for measuring dielectric properties of materials, COMPEL Int. J. Comput. Math. Electr. Electron. Eng. 29 (2010) 1081–1089.
[10] T. Hacib, Y. Le Bihan, M.R. Mekideche, H. Acikgoz, O. Meyer, L. Pichon, Microwave characterization using least-square Support vector machines, IEEE Trans. Magn. 46 (2010) 2811–2814.
[11] T. Hacib, Y. Le Bihan, M.K. Smail, M.R. Mekideche, O. Meyer, L. Pichon, Microwave characterization using ridge polynomial neural networks and least-square Support vector machines, IEEE Trans. Magn. 47 (2011) 990–993.
[12] H. Sadou, T. Hacib, H. Acikgoz, Y. Le-Bihan, O. Meyer, M.R. Rachid Mekideche, An approach based on ANFIS and input selection procedure for microwave characterization of dielectric materials, COMPEL Int. J. Comput. Math. Electr. Electron. Eng. 37 (2018) 799–813.
[13] A. Berglund, S.I.N.L.R. Wold, Implicit non-linear latent variable regression, J. Chemometr. 11 (1997) 141–156.
[14] Ruihua Li, Guoxiang Meng, Naikui Gao, Hengkun Xie, Combined use of partial least-squares regression and neural network for residual life estimation of large generator stator insulation, Meas. Sci. Technol. 18 (2007) 2074–2082.
[15] H. Wold, In encyclopedia of statistical sciences, in: S. Kotz, N.L. Johnson (Eds.), Wiley, New York, 1985, pp. 581–591.
[16] Y. Le Bihan, J. Pavo, C. Marchand, Partial least square regression: an analysis tool for quantitative non-destructive testing, Eur. Phys. J. Appl. Phys. 67 (2014) 30901(7pp).
[17] M. Ruiz, L.E. Mujica, X. Berjaga, J. Rodellar, Partial least square/projection to latent structures (PLS) regression to estimate impact localization in structures, Smart Mater. Struct. 22 (2013) 025028 (11pp).
[18] H. Abdi, Partial least squares regression and projection on latent structure regression (PLS Regression) Wiley Interdiscip, Rev. Comput. Stat. 2 (2010) 97–106.
[19] A. Alin, Comparison of PLS algorithms when number of objects is much larger than number of variables, Stat. Pap. 50 (2009) 711–720.
[20] A.L. Boulesteix, K. Strimmer, Partial least squares: a versatile tool for the analysis of high-dimensional genomic data, Briefings Bioinf. 8 (2006) 32–44.
[21] Y. Li, Z. Zhencai, L. Jianyi, T. Hailong, Modeling on the drum loading performance of the helix shearer based on partial least square regression. Advanced technology of design and manufacture (ATDM 2010), International Conference on, 23-25 Nov. 2010, pp. 134–137 (Beijing, China).
[22] S. Wold, M. Sjostrom, L. Eriksson, PLS-regression: a basic tool of chemometrics, Chemometr. Intell. Lab. Syst. 58 (2001) 109–130.
[23] S. De Jong, SIMPLS: an alternative approach to partial least squares regression, Chemometr. Intell. Lab. Syst. 18 (1993) 251–263.
[24] E. Zamprogna, M. Barolo, D.E. Seborg, Estimating product composition profiles in batch distillation via partial least squares regression, Contr. Eng. Pract. 12 (2004) 917–929.
[25] T. Naes, Tomas Isaksson, Tom Fearn, Tony Davies, Multivariate Calibration and Classification, NIR Publications Chichester, UK, 2004.
[26] G.E.P. Box, W.G. Hunter, J.S. Hunter, Statistics for Experimenters, Wiley, Chichester (, 1978.
[27] O.M. Kvalheim, Latent-variable regression models with higher-order terms: an extension of response modelling by orthogonal design and multiple linear regression, Chemometr. Intell. Lab. Syst. 8 (1990) 59–67.
[28] O.M. Kvalheim, Model building in chemistry, a unified approach, Anal. Chim. Acta 223 (1989) 53–73.
[29] A. Rahman, N. Kondo, Y. Ogawa, T. Suzuki, K. Kanamori, Determination of K value for fish flesh with ultraviolete-visible spectroscopy and interval partial least squares (iPLS) regression method, Biosyst. Eng. 141 (2016) 12–18.
[30] D. Ashourloo, H. Aghighi, A.A. Matkan, M.R. Mobasheri, A.M. Rad, An investigation into machine learning regression techniques for the leaf rust disease detection using hyperspectral measurement, IEEE J. Sel. Top. Appl. Earth Obs. Rem. Sens. 9 (2016) 4344–4351.
[31] http://www.goodfellow.com/F/Polyetherethercetone.HTML, accessed in February 2018.
[32] K. Kupfer, Electromagnetic Aquametry, Electromagnetic Wave Interaction with Water and Moist Substances, Springer-Verlag, Berlin: Heidelberg, 2005.
[33] F. Buckley, A. Maryott, Tables of Dielectric Dispersion Data for Pure Liquids and Dilute Solutions, U.S. Dept. Of Comerce, National Bureau of Standards, Washington, D. C., November, 1958.
[34] N. Wagner, M. Schwing, A. Scheuermann, Numerical 3-D FEM and experimental analysis of the open-ended coaxial line technique for microwave dielectric spectroscopy on soil, IEEE Trans. Geosci. Rem. Sens. 52 (2014) 880–893.
[35] S. Rajesh, K.P. Murali, V. Priyadarsini, S.N. Potty, R. Ratheesh, P. Mohanan, Microwave Dielectric Properties of Rutile Filled PEEK Composites.Polymer-plastics Technology and Engineering vol. 47, (2008), pp. 242–246.
[36] D.W. Marquardt, An algorithm for least-squares estimation of nonlinear parameters, J. Soc. Ind. Appl. Math. 11 (1963) 431–441.
[37] M.T. Hagan, M. Menhaj, Training feed-forward networks with the Marquardt algorithm, IEEE Trans. Neural Network. 5 (1994) 989–993.