An adaptive sampling strategy for quasi real time crack characterization on eddy current testing signals

Shamim Ahmed, Christophe Reboud, Pierre-Emile Lhuillier, Pierre Calmon, Roberto Miorelli

An Adaptive Sampling Strategy for Quasi Real Time Crack Characterization on Eddy Current Testing Signals

Shamim AHMED[a,*], Christophe Reboud[a], Pierre-Emile LHUILLIER[b], Pierre CALMON[a], Roberto MIORELLI[a]

[a] *CEA LIST, Centre de Saclay, F- 91191 Gif-sur-Yvette, France*
[b] *EDF RD, Department MMC, EDF Lab Les Renardieres, 77818, Moret-sur-Loing, France*

Abstract

This work describes quasi real time flaw(s) characterization in conductive plate(s) through the inversion of eddy current testing (ECT) signals using learning by examples (LBE) paradigm. Within the framework of LBE, a fast and accurate learning model is fitted on an optimal training set based on simulated eddy current testing data and the corresponding set of parameters during a preliminary *offline* phase. More specially, the optimal training set has been generated in the *offline* phase by adopting an adaptive sampling strategy through exploiting Partial Least Squares (PLS) feature extraction and output space filling (OSF). Subsequently, a non linear model is fitted on the training set and used to predict the set of parameters associated to an unknown (possibly large) test set during the so-called *online* phase. Different models, i.e., learning algorithms, such as Support Vector Regression (SVR), Kernel Ridge Regression (KRR), Relevance Vector Regression (RVR) and

*Principal Corresponding author
Email address: shamimshawon@yahoo.com (Shamim AHMED)

Augmented Radial Basis Function (A-RBF) have been adopted in order to build different accurate predictors. Afterwards, quasi real-time inversion has been performed on unknown test set by utilizing the corresponding trained models. Comparative results are reported through numerical and experimental data sets to assess the inversion performance of different learning algorithms based on the PLS-OSF sampling strategy.
Keywords: Eddy Current Testing, Real Time Inversion, Learning By Examples, Partial Least Squares, Output-Space-Filling, Support Vector Regression, Kernel Ridge Regression, Relevance Vector Regression, Augmented Radial Basis Function.

---

1. Introduction

Nowadays, machine learning and statistical methods are becoming pervasive data analysis tools in many scientific fields ranging from signal and image processing, chemometrics, biophysics to biomedical imaging and many more. In the very recent years, non-destructive testing scientific community is starting also to profit of the most advanced machine learning methods in order to enhance the capability of already established procedures. That is, automatic classification and regression tasks can be performed in a very efficient way by enabling almost real-time feedback to on field operators during inspection campaigns. In this context, the coming machine learning tools could be applied in order to enhance the accuracy in inspections and mitigate, to some extents, decision of errors due to the so called human factor.

In this paper, we studied and implemented a machine learning methodology targeting regression tasks in view of crack characterization. Inverse

problem solutions in electromagnetic non-destructive testing (E-NDT) can be solved in iterative [1], [2] and non-iterative [3], [4], [5] approaches. Due to higher computational complexity of the iterative solutions for minimizing certain cost function related to the discrepancy between simulation and observation, non-iterative approaches are gaining more interest for real time inversion. Application of machine learning algorithms started to appear in E-NDT almost three decades ago. Authors in [6] showed eddy current testing (ECT) signals interpretation dealing with classification problem. Longitudinal and transverse surface-breaking notches detection and characterization are shown in [7] by using time-domain parameters as the functions of digitized in-phase and quadrature components of probe impedance. Reconstruction of stress map in the strained metallic plates by using electromagnetic measures have been exploited in [8] through support vector regression (SVR). Combination of principal component analysis (PCA) feature extraction and k-mean algorithms is used [9] for cracks detection and classification by applying pulse eddy current testing. Defect localization, and classification tasks are explored in [10] by applying Fisher linear discriminant analysis (FLDA) feature extraction with support vector machine (SVM) by applying pulse eddy current testing. In [11], Tikhonov regularization has been applied as an inversion algorithm for visualization of the geometric profile and size of the cracks through the determination of the 2D image of the eddy current density.

In this work, an adaptive sampling strategy has been addressed for ECT crack characterization problem in order to increase the prediction ability of machine learning algorithms (refers as learning algorithms hereafter) within the framework of Learning by Examples (LBE) [5], [12], [13] paradigm. LBE

can be formulated into two phases approach. A fast and accurate model is built on a training set of input-output (I/O) pairs at the preliminary *offline* phase by regression/prediction techniques (i.e., learning algorithms). During the second phase (*online* phase), the developed model from *offline* phase is used to predict the output associated to an unknown test sample. The performance of different learning techniques for crack detection, localization and characterization depends on the particular choice of the training set generation. Especially, the choice of optimal features selection/extraction from ECT signal space and selection of the corresponding crack parameters from parameter space have impacts on inversion accuracy. Moreover, in case of very time consuming simulations/experiments, one needs to face limitations of available data (e.g., samples). Towards this end, this work describes an adaptive-sampling strategy which combines Partial Least Squares (PLS) [14] feature extraction and modified version of output space filling (OSF) [15]. This strategy aims to uniformly explore the extracted feature space that collects enough information during training set generation phase [16], [17], [18]. Different learning algorithms such as Support Vector Regression (SVR) [19], Kernel Ridge Regression (KRR) [20], Relevance Vector Regression (RVR) [21] and Augmented Radial Basis Function (A-RBF) [22] have been utilized for accurate training model creation and real time flaw characterization by the corresponding trained model. Rather than finding the optimal learning algorithm, the main objective and contribution of this work is to show the impact and robustness of the mentioned PLS-OSF adaptive sampling approach on inversion by applying different learning algorithms in presence of corrupted (additive synthetic noise) test sets as well as experimental data.

Therefore, initially the performance evaluation of the mentioned prediction approaches on the data set obtained by PLS-OSF sampling strategy will be explored in presence of noisy test set and compared with more standard sampling approach (i.e., full factorial grid (GRID) [12]). Finally, a set of unknown samples consisting in real experimental data are utilized for evaluating the performance of crack characterization. The paper is organized as follows. In Section 2, a brief summary on the mathematical formulation of the eddy current signal modelling is provided. Thus an exhaustive description of the adaptive sampling strategy and the addressed learning methods employed in the paper are described. In Section 3, training and test set configurations are provided for both numerical and experimental cases. Moreover, a comparative analysis of the treated LBE solutions on numerical and experimental data are detailed in Section 4. Concluding remarks and further research directions are outlined in Section 5.

2. Mathematical formulation

2.1. Problem definition

Let us consider a homogeneous conductive specimen inspected by two coils working in emitting-receiving mode as the one shown in Fig.1. The presence discontinuities within the inspected medium (e.g., flaws, voids, inclusions, etc.) perturb the eddy current flow compared to the flawless specimen. The differential signal between eddy currents in healthy and flawed medium represents the so called crack signature or ECT signal signature. Magnitude and phase of the ECT signals are changed in function of the crack parameters e.g., crack geometry, conductivity, position in the specimen, probe geome-

try, specimen characteristics etc. Numerical modelling of ECT signals aims at simulating inspection problems in which all these quantities vary. For sake of simplicity we introduce the modelling based on the inspection of a plate structure affected by a three dimensional crack with rectangular section. The extension of the model described hereafter has been provided for planar multi-layered structures affected by multiple cracks elsewhere [23].

Let us describe the presence of the crack as a local variation of conductivity described as $\delta\sigma\left(\underline{\mathrm{r}}\right) = \sigma_c\left(\underline{r}\right) - \sigma$, where σ stands for the plate conductivity and $\sigma_c\left(\underline{r}\right)$ stands for the crack conductivity at the point $\underline{\mathrm{r}} = (x, y)$ which represents a coordinate in a three dimension Cartesian space. The eddy currents are induced within the specimen by a driven coil and having the incident electric field within the specimen in absence of flaw(s) described by $\underline{\mathrm{E}}^{inc}\left(\underline{\mathrm{r}}\right)$. Considering the Volume Integral Method (VIM) formalism [24], one can show that the total field within the crack zone $\underline{\mathrm{E}}^{tot}\left(\underline{\mathrm{r}}\right)$ is obtained in terms of the following Fredholm equation of second kind

$$\underline{\mathrm{E}}^{tot}\left(\underline{\mathrm{r}}\right) = \underline{\mathrm{E}}^{inc}\left(\underline{\mathrm{r}}\right) + i\omega\mu_0 \int_{V_f} \underline{\underline{\mathrm{G}}}\left(\underline{\mathrm{r}}, \underline{\mathrm{r}}'\right) \delta\sigma\left(\underline{\mathrm{r}}'\right) \underline{\mathrm{E}}^{tot}\left(\underline{\mathrm{r}}'\right) dV_f. \tag{1}$$

In equation (1), V_f stands for the the volume occupied by the flaw and $\underline{\underline{\mathrm{G}}}\left(\underline{\mathrm{r}}, \underline{\mathrm{r}}'\right)$ is the electric-electric dyadic Green function which is employed to describe the effects of a electric dipole source located at $\underline{\mathrm{r}}'$ on the electric field at the position $\underline{\mathrm{r}}$. Equation (1) is solved by using the method of moments thus the unknown of the problem i.e., the equivalent electric dipole density in the crack zone $\mathrm{d}\left(\underline{\mathrm{r}}\right) = \underline{\mathrm{E}}^{tot}\left(\underline{\mathrm{r}}'\right) \delta\sigma\left(\underline{\mathrm{r}}'\right)$, is retrieved. Consequently, the impedance variation on the receiving coil at the k-th scanning position with $k = 1, ..., K$,

can be calculated via the reciprocity theorem as

$$\Delta Z_k = \frac{1}{I^2} \int_{V_f} \delta\sigma \left(\underline{\mathrm{r}}'_k\right) \underline{\underline{\mathrm{E}}}^{inc} \left(\underline{\mathrm{r}}'_k\right) \cdot \underline{\underline{\mathrm{E}}}^{tot} \left(\underline{\mathrm{r}}'_k\right) dV_f. \tag{2}$$

In this paper, in order to describe the different crack configurations, we have chosen to define the volume V_f with a set of $Q = 3$ descriptors defined through the vector $\underline{\mathrm{p}} = (l_c, h_c, w_c)$ where the different components represent the crack length, height and width, respectively. Crack center and direction are assumed to be a priori for the described problem. However, in case of other problems (e.g., crack localization or crack characterization and localization), crack position parameters can be defined by considering them within vector $\underline{\mathrm{p}}$. For sake of brevity, hereafter we refer to equation (2) as $\Delta Z_k = \mathcal{F}\left\{\underline{\mathrm{r}}'_k, \underline{\mathrm{p}}\right\}$ where $\mathcal{F}$ stands for the employed black box forward solver. In this work we have employed the CIVA software [25] in order to carry out the simulation campaign. Due to their complex valued nature, ECT signals are laying in a $2K$-dimensional feature space which can be very large when a large number of coil positions are needed to properly describe the signal in space. Furthermore, ECT signals may not vary smoothly in function of the different crack parameter combinations thus a possibly high number of samples is needed in order to generate the suitable training set. These two issues make the learning problem very cumbersome to address in the native $2K$-dimensional feature space, this problem is known in the literature as curse of dimensionality (COD). In order to overcome this issue, this paper proposed a strategy in which the ECT signals are projected first onto a latent space, i.e. the extracted features space, via PLS having a dimension $J << 2K$. Subsequently, we propose an adaptive algorithm called PLS-OSF aiming at homogeneously fill the the projected feature space. We will show that this

strategy leads to an uneven sample filling of the parameter space. The samples locations obtained by the adaptive algorithm are such that an higher emphasis is provided on the set of parameters "carrying" more information content i.e., the ones for which the outputs vary more.

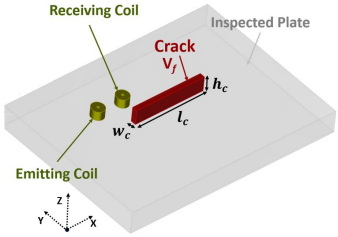

Figure 1: Example of the studied plate geometry.

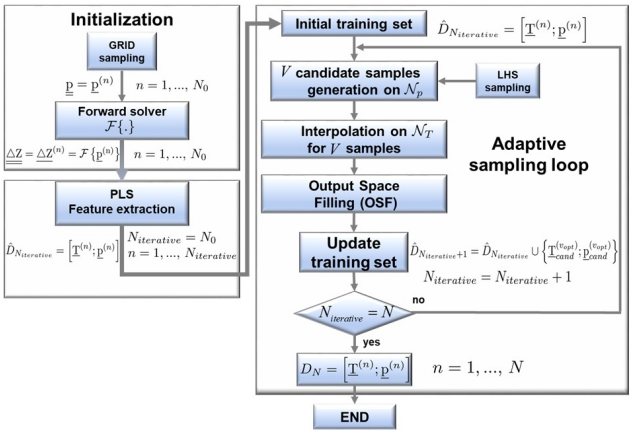

Figure 2: Flow chart representation of PLS-OSF sampling.

2.2. Adaptive sampling through feature extraction

The sampling strategy apply, as a first step, the PLS feature extraction on the ECT signal space $\mathcal{N}_{\triangle Z}$ to reduce the dimension of $\mathcal{N}_{\triangle Z}$. Subsequently, an adaptive sampling strategy is performed directly in the reduced extracted

feature space $\mathcal{N}_T$ in order to retrieve the lowest number of training samples during training phase. The following steps describe the iterative procedure of PLS-OSF sampling approach summed-up by the flow chart in Fig. 2. It is worth to be mentioned that this preliminary and mandatory stage, also known as *offline* phase, enable to build suitable I/O pairs for building optimal training model.

i **Initialization**- Generate N_0 number of initial samples by using a uniform GRID (i.e., full factorial grid) sampling [12] within the parameter space $\mathcal{N}_p$. Thus, a matrix of defect parameters $\underline{\underline{\mathrm{p}}} = (\underline{\mathrm{p}}^{(n)}; n = 1, ..., N_0)$ having $(N_0 \times Q)$ dimension is formed, where $\underline{\mathrm{p}}^{(n)}$ is the n-th row of $\underline{\underline{\mathrm{p}}}$. By using $\mathcal{F}\{.\}$, generate the associated ECT coil signals having $(N_0 \times 2K)$ feature matrix $\underline{\underline{\triangle \mathrm{Z}}} = \left(\underline{\triangle \mathrm{Z}}^{(n)}; n = 1, ..., N_0\right)$, where n-th signal is represented by $\underline{\triangle \mathrm{Z}}^{(n)}$. The n-th signal vector of $2K$ ECT features in the $\mathcal{N}_{\triangle Z}$, is represented by $\underline{\triangle \mathrm{Z}}^{(n)} = \left\{\left(\Re\left\{\triangle Z_k^{(n)}\right\}; \Im\left\{\triangle Z_k^{(n)}\right\}\right); k = 1, ..., K\right\}$.

ii **PLS Feature Extraction**- In this step, the $2K$-dimension ECT features associated to the ECT signal are reduced to J-dimension extracted features where $J << 2K$. Considering a data set composed by N_0 samples, we refer to the predictors as the set of whole ECT signals $\underline{\underline{\triangle \mathrm{Z}}}^* = \left[\Re\left\{\underline{\triangle \mathrm{Z}}\right\}, \Im\left\{\underline{\triangle \mathrm{Z}}\right\}\right]$ with $\underline{\underline{\triangle \mathrm{Z}}}^* \in \mathbb{R}^{N_0 \times 2K}$ and to the responses as the set of $\underline{\underline{\mathrm{p}}}^*$ parameters having size $\mathbb{R}^{N_0 \times Q}$. $\underline{\underline{\triangle \mathrm{Z}}}^*$ and $\underline{\underline{\mathrm{p}}}^*$ are the zero mean versions of $\underline{\underline{\triangle \mathrm{Z}}}$ and $\underline{\underline{\mathrm{p}}}$, respectively. That is, each of the columns of $\underline{\underline{\triangle \mathrm{Z}}}$ and $\underline{\underline{\mathrm{p}}}$ has been centered by subtracting to the mean value of that column. The underlying concept in PLS algorithm assume that predictors and responses can be properly described by a linear combination of a subset of J variables. These variables are also known as latent variables

or extracted features which maximize the covariance between the predictors and responses. PLS decomposes the predictors $\underline{\underline{\triangle Z}}^*$ and responses $\underline{\underline{p}}^*$ such that [14]:

$$\begin{aligned} \underline{\underline{\triangle Z}}^* &= \underline{\underline{T P}}^\top + \underline{\underline{E}} \\ \underline{\underline{p}}^* &= \underline{\underline{T C}}^\top + \underline{\underline{F}} \end{aligned}, \tag{3}$$

where $\underline{\underline{P}} \in \mathbb{R}^{2K \times J}$ is the predictors loading matrix and $\underline{\underline{E}} \in \mathbb{R}^{N_0 \times 2K}$ is the matrices of residuals associated to the predictors. $\underline{\underline{C}} \in \mathbb{R}^{Q \times J}$ and $\underline{\underline{F}} \in \mathbb{R}^{N_0 \times Q}$ the response loading and the matrices of residuals associated responses, respectively. Finally, $\underline{\underline{T}} \in \mathbb{R}^{N_0 \times J}$ is the score matrix and superscript "$\top$" stands for transpose operation. One can show that the score matrix can be defined as $\underline{\underline{T}} = \underline{\underline{\triangle Z}}^* \times \underline{\underline{W}}$ with $\underline{\underline{W}} \in \mathbb{R}^{N_0 \times J}$ is the so-called weights matrix and it is calculated such that each column of $\underline{\underline{T}}$ is orthogonal to each other. The columns of T represent the latent variables associated to a given ECT signal. T is obtained by multiplying $\underline{\underline{\triangle Z}}^*$ with W that maximizes the covariance between $\underline{\underline{\triangle Z}}^*$ and $\underline{\underline{p}}^*$. In this sense, T can be seen as the projection of $\underline{\underline{\triangle Z}}^*$ signal having 2K-dimensions onto a J-dimensional space. In this paper the weights have been calculated by using SIMPLS algorithm [26] as detailed in Annex 1. Assign the number of training sample $N_{iterative} = N_0$ and construct an initial training set $\hat{D}_{N_{iterative}} = \left\{ \left(\underline{T}^{(n)}; \underline{p}^{(n)} \right); n = 1, ..., N_{iterative} \right\}$ for the adaptive step.

iii **Adaptive Sampling-** Generate V candidate samples within the parameter space (i.e., $\mathcal{N}_p$) by $\underline{p}_{cand}^{(v)} = \left(p_{cand,q}^{(v)}; q = 1, ..., Q \right)$ through Latin Hypercube Sampling (LHS) strategy where $v = 1, ..., V$. An estimation of the J-dimensional set of extracted features corresponding to each v-th candidate, $\underline{\tilde{T}}_{cand}^{(v)}$ is retrieved by applying a multi-dimensional lin-

ear interpolator on $\hat{D}_{N_{iterative}}$. Select the optimal $v = v_{opt}$ candidate (i.e., $\underline{\mathrm{p}}_{cand}^{(v_{opt})}$) from V such that the minimum distance between the obtained extracted features $\underline{\tilde{\mathrm{T}}}_{cand}^{(v_{opt})}$ and all the available extracted features $\underline{\mathrm{T}}^{(n)}$ $(n = 1, ..., N_{iterative})$ within $\hat{D}_{N_{iterative}}$ is maximized [i.e., $v_{opt} = \arg\left(\max_{v=1,...,V}\left\{\min_{n=1,...,N}\left[d_{vn}\right]\right\}\right)$]. d_{vn} is the Euclidean distance between $\underline{\tilde{\mathrm{T}}}_{cand}^{(v)}$ $(v = 1, ..., V)$ and $\underline{\mathrm{T}}^{(n)}(n = 1, ..., N_{iterative})$, which can be described by $d_{vn} = \sqrt{\sum_{j=1}^{J}\left(\tilde{T}_{cand,j}^{(v)} - T_j^{(n)}\right)^2}$. The ECT features $\underline{\underline{\triangle \mathrm{Z}}}_{cand}^{(v_{opt})} = \left\{\left(\Re\left\{\triangle Z_{cand,k}^{(v_{opt})}\right\}; \Im\left\{\triangle Z_{cand,k}^{(v_{opt})}\right\}\right); k = 1, ..., K\right\}$ associated to the selected candidate sample is computed by utilizing $\mathcal{F}\{.\}$. Finally, the set of extracted features is obtained by $\underline{\mathrm{T}}_{cand}^{(v_{opt})} = \left(\underline{\underline{\triangle \mathrm{Z}}}_{cand}^{(v_{opt})}\right)^* \times \underline{\underline{\mathrm{W}}}$. Finally, update the training set with $\hat{D}_{N_{iterative}+1} = \hat{D}_{N_{iterative}} \cup \left\{\underline{\mathrm{T}}_{cand}^{(v_{opt})}; \underline{\mathrm{p}}_{cand}^{(v_{opt})}\right\}$ and update $N_{iterative} = N_{iterative} + 1$. We shall refer to this adaptive sampling schema a customized OSF version (i.e., candidate parameters are chosen such that features are uniformly distributed).

iv **Stop Criterion**- The adaptive sampling step adds new sample iteratively until $N_{iterative} = N$ (N is the desired/feasible training size).

Figure 3 represents the exploration of adaptive training samples generation in parameter space as well as in extracted feature space. In Fig. 3, feature space is represented for the first two extracted features (out of five) based on the obtained training set. At this stage, different predictions techniques have been utilized to train separately q-th set of I/O pairs $\hat{D}_{N,q} = \left\{\left(\underline{\mathrm{T}}^{(n)}; p_q^{(n)}\right); n = 1, ..., N\right\}$ on the generated training set for each q-th parameter $(q = 1, ..., Q)$ of the crack.

It is worth to mention that the test set containing $2K$ ECT features associ-

ated to previously-unseen crack parameter configurations are also projected to the J-dimensional PLS-extracted feature space. Thus, the m-th unknown test sample $\underline{\triangle \mathrm{Z}}^{(m)}$ having $2K$ ECT features for the defect parameters $\underline{\mathrm{p}}^{(m)}$ is projected to $\mathcal{N}_T$ through $\underline{\underline{\mathrm{W}}}$ [i.e., $\underline{\mathrm{T}}^{(m)} = \left(\underline{\triangle \mathrm{Z}}^{(m)}\right)^* \times \underline{\underline{\mathrm{W}}}$]. $\left(\underline{\triangle \mathrm{Z}}^{(m)}\right)^*$ is obtained by using the mean values obtained from the already described PLS Feature Extraction step (i.e., from offline phase). Finally, $\underline{\mathrm{T}}^{(m)}$ is given as input to the corresponding trained model in order to estimate the q-th parameter of the crack, $\tilde{p}_q^{(m)}$ $(q = 1, ..., Q)$.

2.3. Learning algorithms

Given a data set $\hat{D}_{N,q} = \left\{\left(\underline{\mathrm{T}}^{(n)}; p_q^{(n)}\right)\right\}$ through adaptive sampling detailed in the previous section (Subsec. iii), the aim of the learning algorithm is to find a function $\tilde{\Theta}_q\{.\}$ that estimates the relationship between $\underline{\mathrm{T}}^{(n)}$ and $p_q^{(n)}$ and provides a predicted/estimated output which is close to $p_q^{(n)}$ for all n. The obtained $\tilde{\Theta}_q\{.\}$ is then used to estimate the q-th crack parameter, $\tilde{p}_q^{(m)} = \tilde{\Theta}_q(\underline{\mathrm{T}}^{(m)})$ for an unknown m-th test sample $\underline{\mathrm{T}}^{(m)}$ where, $q = 1, ..., Q; m = 1, ..., M$. This section briefly describes the formulation of different learning algorithms that are treated within the scope of LBE.

2.3.1. Augmented Radial Basis Function (A-RBF)

A-RBF is the extended version of actual Radial Basis Function (RBF) interpolator, where a polynomial term is added with the standard RBF interpolator. The prediction of the p_q-th crack parameter by applying A-RBF can be formulated by

$$\tilde{p}_q^{(m)} = \sum_{s=1}^{S} z_s(\underline{\mathrm{T}}^{(n)})\rho_s + \sum_{n=1}^{N} \kappa(\underline{\mathrm{T}}^{(n)}, \underline{\mathrm{T}}^{(m)})\omega_q^{(n)}. \tag{4}$$

where $z_s(\underline{\mathrm{T}}^{(n)}) = 1,\ T_1^n, ...,\ T_J^n, (T_1^n)^2, (T_1^n T_2^n), ...,\ (T_J^n)^2, ...$ defines the polynomial function with polynomial coefficient ρ_s. $\omega_q^{(n)}$ is the RBF weight for the n-th sample that is obtained during training model development. κ is the kernel function whose value depends on the radial distance between $\underline{\mathrm{T}}^{(n)}$ and $\underline{\mathrm{T}}^{(m)}$. Different types of kernel function such as linear, Gaussian, thin plate spline, multi-quadratic etc. can be used as a kernel function. To be consistent with all other learning methods, we have applied Gaussian kernel that can be expressed by

$$\kappa(\underline{\mathrm{T}}^{(n)},\ \underline{\mathrm{T}}^{(m)}) = exp\left(-\gamma_q \left|\left|\underline{\mathrm{T}}^{(n)} - \underline{\mathrm{T}}^{(m)}\right|\right|^2\right) \tag{5}$$

where γ_q is the Gaussian coefficient that needs to be calibrated to obtain optimal training model in *offline* phase.

2.3.2. Support Vector Regression (SVR)

Analogously to [12, 27], by employing ε-SVR, prediction of p_q-th parameter for the m-th unknown test sample can be derived by

$$\tilde{p}_q^{(m)} = \sum_{n=1}^{N} (\varsigma_q^{(n)} - \iota_q^{(n)}) \kappa(\underline{\mathrm{T}}^{(n)},\ \underline{\mathrm{T}}^{(m)}) + b_q. \tag{6}$$

where $\varsigma_q^{(n)}, \iota_q^{(n)} \in (0, C_q)$ are Lagrange multipliers that are determined by solving dual optimization problem [27]. C_q is the user defined constant parameter that controls error larger than ε_q for N number of training samples. b_q is the bias coefficient that is computed by fulfilling Karush-Kuhn-Tucker (KKT) conditions [28]. Gaussian kernel κ has been utilized which is controlled by Gaussian coefficient γ_q. Thus, the learning ability of SVR depends on the appropriate value of ε_q, C_q and γ_q.

2.3.3. Kernel Ridge Regression (KRR)

Kernel ridge regression combines the well known ridge regression with kernel trick [20]. Through applying KRR, the prediction of p_q-th parameter for the m-th unknown test sample can be obtained by

$$\tilde{p}_q^{(m)} = \underline{\mathrm{p}}_q^{\top} \left(\kappa \left(\underline{\mathrm{T}}^{(n)}, \underline{\mathrm{T}}^{(o)} \right) + a_q \underline{\underline{\mathrm{I}}} \right)^{-1} \kappa \left(\underline{\mathrm{T}}^{(n)}, \underline{\mathrm{T}}^{(m)} \right) \tag{7}$$

where $n = 1, ..., N; o = 1, ..., N$. a_q is the fixed positive constant that is also known as regularization parameter. For using Gaussian kernel κ, accurate KRR model development is depended on the appropriate choice of Gaussian coefficient γ_q and user defined parameter a_q.

2.3.4. Relevance Vector Regression (RVR)

Relevance vector regression adopts Bayesian probabilistic framework for developing regression model. Having the model parameters, for an unknown m-th test sample, the p_q-th parameter can be obtained through the posterior predictive density [21]

$$p \left(\tilde{p}_q^{(m)} \mid \underline{\mathrm{p}}_q, \underline{\alpha}_{ML}, \beta_{ML} \right) = \mathcal{N} \left(\tilde{p}_q^{(m)} \mid y_q^{(m)}, (\sigma^{(m)})^2 \right) \tag{8}$$

where $\mathcal{N}$ specifies a Gaussian distribution over $\tilde{p}_q^{(m)}$ with $y_q^{(m)} = \underline{\mu}_q^{\top} \varphi \left(\underline{\mathrm{T}}^{(m)} \right)$ is the posterior predictive mean that is obtained by the basis vector weighted by a sparse vector of mean weights $\underline{\mu}_q$. $\underline{\alpha}_{ML}$ and β_{ML} define a vector of hyper-parameters and the precision on the noise, respectively that are computed by maximum likelihood estimation. $(\sigma^{(m)})^2 = \beta_{ML}^{-1} + \varphi \left(\underline{\mathrm{T}}^{(m)} \right)^{\top} \underline{\underline{\Sigma}} \varphi \left(\underline{\mathrm{T}}^{(m)} \right)$ is the variance on the predictions. The posterior co-variance matrix is denoted by $\underline{\underline{\Sigma}} = \left(\beta \underline{\underline{\phi}}^{\top} \underline{\underline{\phi}} + \underline{\underline{\mathrm{A}}} \right)^{-1}$ with $[N \times N+1]$ design matrix $\underline{\underline{\phi}} = \left[\varphi \left(\underline{\mathrm{T}}^{(1)} \right), ..., \varphi \left(\underline{\mathrm{T}}^{(N)} \right) \right]^{\top}$. $\underline{\mathrm{A}} = diag(\alpha_0, \alpha_1, ..., \alpha_N)$ is the precision vector on the weight vector $\underline{w}$. The

basis function vector $\varphi\left(\underline{\mathrm{T}}^{(m)}\right) = \left[1, \kappa\left(\underline{\mathrm{T}}^{(m)}, \underline{\mathrm{T}}^{(1)}\right), ..., \kappa\left(\underline{\mathrm{T}}^{(m)}, \underline{\mathrm{T}}^{(N)}\right)\right]^{\top}$ is defined by the input observations through kernel function κ. Due to the use of Gaussian kernel, the accurate training model depends on the choice of Gaussian coefficient γ_q which must be properly selected.

3. Training and test set configuration

This section describes the numerical and experimental configurations that are addressed in this paper. We consider a homogeneous plate of thickness 1.27 mm and conductivity 1.02 MS/m affected by a single sub-surface crack (crack is placed opposite side of the inspecting coil probe) as shown in Fig. 1. The plate is inspected by 2 coils in emitting and receiving mode working at a frequency of $f = 200$ kHz. More details on coil parameters can be found in [29]. $Q = 3$ crack parameters (length l_c, height h_c, and width w_c) will be investigated for the rest of the paper. CIVA simulator has been utilized as forward operator $\mathcal{F}\{.\}$ for obtaining ECT signals. Real life inspections suffer for various types of noise sources (e.g., set up noise, man-made noise etc.). Therefore, before dealing with real experimental data, it is a good practice to validate the trained models on noisy test set. To partially consider noise effects, the test set response signal collected by the coil is blurred by an Additive White Gaussian Noise (AWGN) with a signal-to-noise ratio (SNR), where $SNR = 10\log_{10}\{\frac{\Sigma_{k=1}^{K}|\triangle Z_k|^2}{\Sigma_{k=1}^{K}|\zeta_k|^2}\}$. ζ_k denotes the complex valued additive noise which corrupts the ECT signal $\triangle Z_k$ at k-th measurement point. The performance of the inversion scheme is analyzed by computing the normalized mean error, $NME_q = \frac{1}{M}\sum_{m=1}^{M}\left(\left|p_q^{(m)} - \tilde{p_q}^{(m)}\right| / p_q^{(m)}\right)$ and relative prediction error, $\xi_q^{(m)} = \left(\left|p_q^{(m)} - \tilde{p_q}^{(m)}\right| / p_q^{(m)}\right)$ for numerical and experimental evalua-

tion, respectively. $p_q^{(m)}$ and $\tilde{p}_q^{(m)}$ are the actual and predicted q-th parameter $(q = 1, ..., Q)$ of the m-th $(m = 1, ..., M)$ unknown test sample respectively.

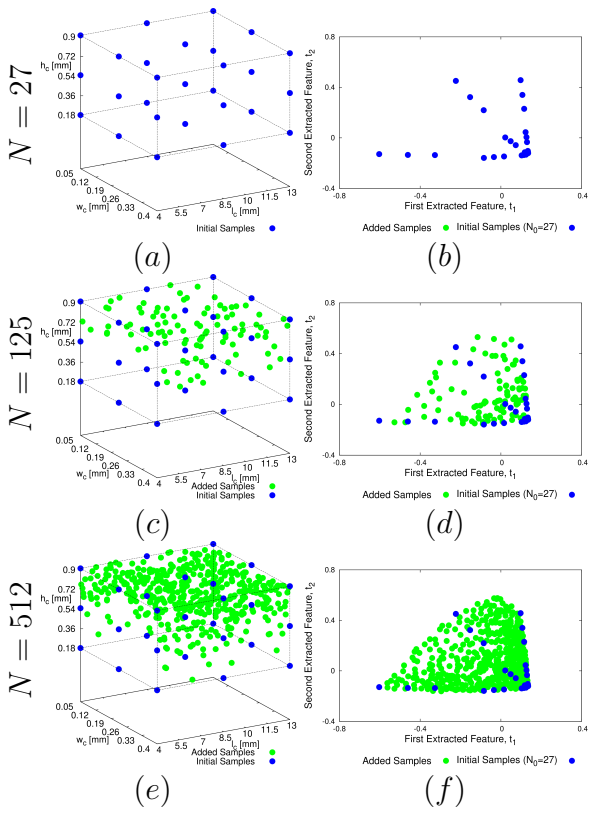

Figure 3: Disposition of training samples obtained by PLS-OSF in $(a)(c)(e)$ crack parameter space and $(b)(d)(f)$ extracted feature space for $(a)(b)$ $N = 27$, $(c)(d)$ $N = 125$, $(e)(f)$ $N = 512$.

3.1. Numerical set up

A set of $N \in [27, 512]$ samples has been built by utilizing GRID and PLS-OSF sampling methods separately by changing the crack dimensions within the range $l_c \in [3,\ 14]$ mm, $h_c \in [0.189,\ 0.931]$ mm and $w_c \in [0.05,\ 0.4]$ mm

within the selected ranges. In case of PLS-OSF approach, $N = 512$ maximum samples have been obtained adaptively starting from $N_0 = 27$ initial samples and $N = [27, 64, 125, 216, 343, 512]$ has been considered for GRID approach. ECT signals are generated for each of these training samples and collected from $K = 644$ inspected points by $\mathcal{F}\{.\}$. Hence each sample point is associated to $2K = 1288$ ECT features. Six different training sets have been obtained for three crack parameters (l_c, h_c, w_c) for each sampling method (i.e., GRID, PLS-OSF) using N number of training samples. An unknown test set of $M = 1000$ samples for 3 crack parameters has been generated by using LHS design. 1288 ECT features are also considered for each test sample for treating the training models obtained by GRID sampling. Whereas, $J = 5$ features are extracted from 1288 actual ECT features for both training and test set generation as described in Sec. 2.2 for PLS-OSF sampling strategy. The choice of optimal selection of J with other calibration parameters are described in the following section.

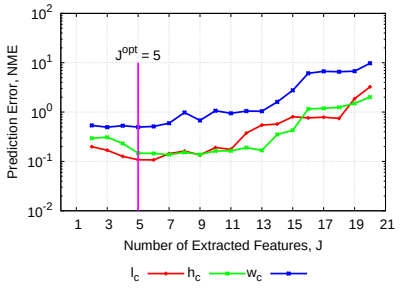

Figure 4: Prediction error, NME_q vs. number of extracted features J representation while performing inversion on the test set at $SNR = 10$ dB using PLS regression technique for $N = 512$ and $M = 1000$ obtained by PLS-OSF sampling strategy.

3.2. Parameters selection

Dealing with feature extraction and learning algorithms, it is essential and fare to use their own parameter in their optimal state. Our main goal is to show comparative analysis instead of showing superiority of each learning algorithm by utilizing as much as lower number of tuning parameters. As a consequence, the choice of optimal number of features (i.e., J) to be extracted from actual ECT signal space is a crucial issue. There is no particular criteria available for selecting optimal J value. Nevertheless, it is well known in machine learning framework that the lower the number of features to be learned the more robust learning model can be obtained. That is, the number of extracted features should be kept as low as possible to minimize the complexity of the learning algorithm in training phase. However, it should be large enough to ensure that all the essential information from ECT data is compressed within the extracted features. Within this guideline, it is worth pointing out that the PLS technique ranks the features from the most to the less important ones accordingly to the corresponding amount of information on the unknown quantity to be predicted [14]. So, an effective and easy-to-implement strategy for choosing the optimal value of J is to build a training set with a maximum number of extracted features J_{max}. Then, iteratively removing the higher-order (i.e., less informative) ones such that an optimum error is found. Figure 4 shows the behavior of the error index for predicting each crack parameter for $J \in [2,\ 20]$ with a step size 1, while processing noisy test data (i,e., $SNR = 10$ [dB]) for applying PLS as a regression technique [14]. By averaging the optimal values obtained for each defect parameter, $J^{opt} = 5$ has been chosen.

Using ε-SVR from LibSVM [30], ε is fixed for 10^{-2}. Whereas, Dlib [31] has been adopted for KRR regression. For initialization, $\alpha = 10^{12}$ has been assigned as par the default value while using the RVR library described in [32], which is updated during training model development. All learning techniques are exploited by Gaussian kernel. Thus, the choice of the kernel hyper-parameter γ_q for all the learning algorithms and C_q for SVR are obtained by performing 5-folds cross-validation during *offline* phase. Table 1 reports the optimal parameters obtained for different learning techniques for the training sets at $N = 512$ obtained by both PLS-OSF and GRID sampling strategies.

Table 1: Optimal setup for the calibration parameters

PLS-OSF	SVR		KRR	RVR	A-RBF
Parameters	C^{opt}	γ^{opt}	γ^{opt}	γ^{opt}	γ^{opt}
l_c	10^5	1	1	5	1
h_c	10^2	10^{-1}	1	1	10
w_c	10^2	1	1	10	10
GRID	SVR		KRR	RVR	A-RBF
Parameters	C^{opt}	γ^{opt}	γ^{opt}	γ^{opt}	γ^{opt}
l_c	10^4	10^{-2}	10^{-2}	10^{-2}	10^{-5}
h_c	10^4	10^{-3}	10^{-2}	10^{-3}	10^{-4}
w_c	10^4	10^{-2}	10^{-1}	10^{-5}	10^{-4}

Training models are obtained separately by different learning algorithms on the obtained training sets with the corresponding calibration parameters in *offline* phase. Finally, the estimation of crack parameters on the unknown test set is evaluated (*online* phase) by the obtained trained models.

Table 2: Crack dimension in experimental test set

Crack No.	1	2	3	4	5	6
l_c [mm]	6	7	8	12	9	9
h_c [mm]	0.508	0.508	0.508	0.508	0.254	0.381
w_c [mm]	0.2	0.2	0.2	0.2	0.2	0.2

3.3. Experimental set up

A homogeneous conductive plate having the same properties described in Sec. 3 is affected by an external crack for 6 different crack configurations (Fig. 5). The actual dimensions of these cracks are indicated in Tab. 2. All other inspection parameters remain unchanged. ECT response signals are collected from $K = 644$ inspected points. As a consistency check, Fig. 6 compares the real and the imaginary parts of experimental ECT signals with the CIVA simulated signals along the x-axis of the cracks. For proper adjustment with the simulated data, a calibration factor $[\Re = -0.0049; \Im = 0.0135]$ and bias factor $[\Re = 0.004; \Im = 0.005]$ have been applied on the experimental data. This is worthy to mention that the discrepancy between experimental and simulated data may encounter due to the presence of uncertainties between normal and real parameters (e.g., probe lift-off, probe tilt/pitch, plate thickness etc.) employed in both cases. Moreover, non negligible noise components can also corrupt the realistic measurement. Fig. 5 (b) and Fig. 5 (e) contain the signal maps of two cracks which are partially corrupted during real experiment. Similar to the numerical test set, $J^{opt} = 5$ optimal features are extracted for each test sample (i.e., ECT signal map for each

crack) for applying PLS-OSF sampling. The obtained training models for $N = 512$ samples by PLS-OSF sampling with the corresponding optimal parameters (Tab. 1) of the learning algorithms are applied for predicting the experimental cracks parameters.

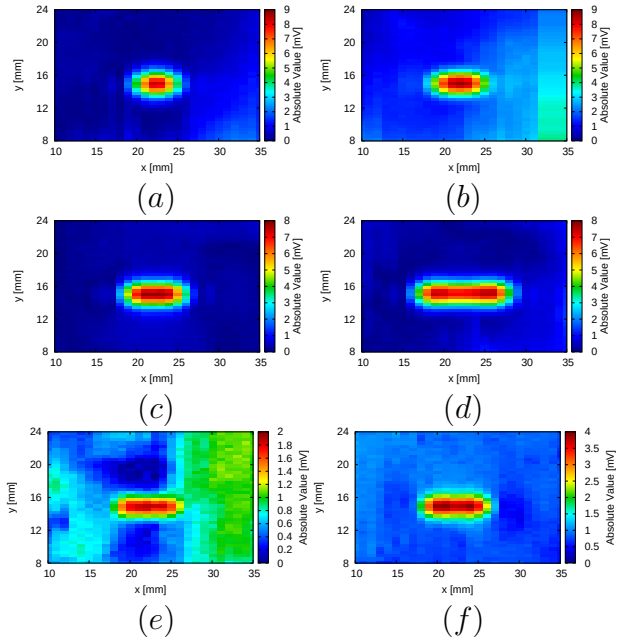

Figure 5: Referring to Table 2, ECT signal maps of (a) Crack 1, (b) Crack 2, (c) Crack 3, (d) Crack 4, (e) Crack 5 and (f) Crack 6.

4. Results and discussion

4.1. Numerical assessment

As a reference and the further discussion, let us first analyze the prediction accuracy of the learning algorithms by applying GRID sampling strategy. In case of higher number of ECT features (i.e., $F = 2K$) while dealing with GRID sampling, higher N is required for solving COD problem. Hence, prediction accuracy improves significantly by each learning algorithm with

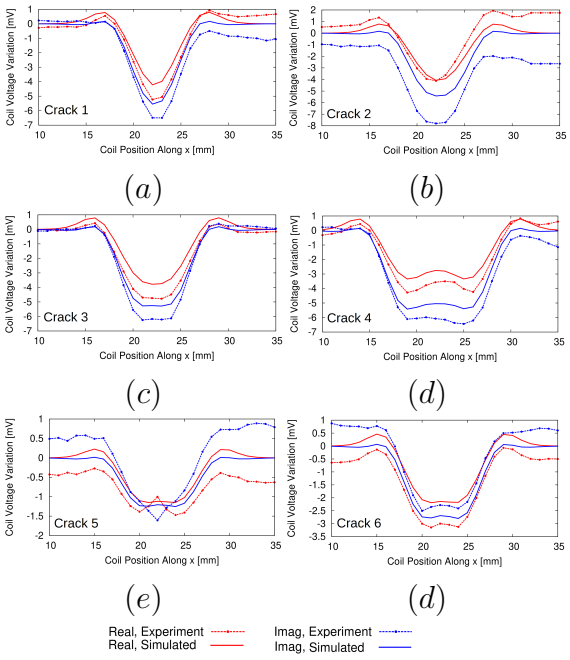

Figure 6: Referring to Tab. 2, real and imaginary parts of the coil voltage variation experimental/CIVA simulated for the coil position along x direction at $y = 15$ [mm].

the increment of N while dealing with $Noiseless$ test set (Fig. 7(a), 7(c), 7(e)). Due to the use of polynomial function with RBF interpolator, A-RBF shows higher prediction accuracy with respect to other algorithm for $Noiseless$ data. The employed A-RBF is an interpolation technique that commonly do not assess any regularization, consequently, mostly affected in presence of noisy data and shows so higher error to be visible within the Fig. 7(a), 7(c), and 7(e) for the training set obtained by GRID sampling. This is a sign of over-fitting phenomenon, where the trained model does not work at all in case of very unknown/corrupted noisy test set. On the other hand, having regularization parameters, the applied regression methods (i.e.,

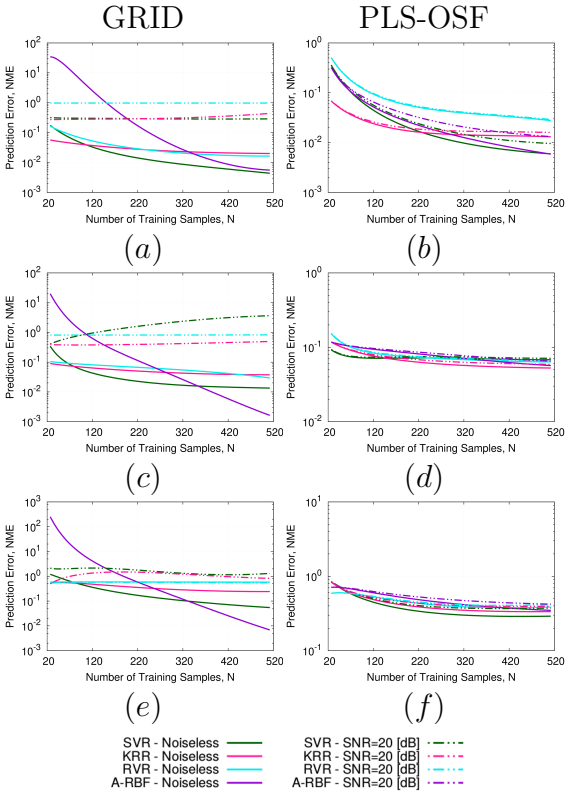

Figure 7: Prediction error, *NME* versus training size N representation at $(a)(b)\, p_1 = l_c$, $(c)(d)\, p_2 = h_c$ and $(e)(f)\, p_3 = w_c$ when processing noiseless and noisy test set ($SNR = 20$ [dB]) by GRID and PLS-OSF.

SVR, KRR, RVR) can slightly deal with noisy data which show, off-course, higher prediction error with respect $Noiseless$ test case. For a given number of samples $N = 512$, GRID sampling strategy could not generate optimal training sets to build robust training models by the regression methods to handle noisy test set. Comparatively, through applying PLS feature extraction, actual ECT features from $\mathcal{N}_{\triangle Z}$ are transformed to $J = 5$ dimensional extracted feature space $\mathcal{N}_T$ with most significant features. This overcomes

the COD problem. Consequently, it requires lower number of training samples to have more accurate training models by the learning algorithms. By applying PLS-OSF, $\mathcal{N}_T$ is uniformly filled such a way that represents optimal (i.e., complete and non-redundant) feature space. Accordingly, optimal training sets are obtained for lower N. This improves the learning ability of the treated algorithms and provides higher prediction accuracy on $Noiseless$ as well as noisy data by means of lower prediction error (i.e., NME) for crack dimension estimation (Fig. 7(b), 7(d), 7(f)). Though, PLS-OSF can provide optimal training set, prediction accuracy of different crack parameters are dependent on different prediction techniques. We can observe this behavior from Fig. 7 for l_c, h_c and w_c estimation. For a given number of training samples $N = 512$, PLS-OSF/SVR and PLS-OSF/A-RBF show lowest prediction error for l_c estimation for $Noiseless$ as well as on noisy data (Fig. 8). However, all learning algorithms have comparable prediction error for h_c estimation for applying PLS-OSF sampling. Due to the narrower crack gap (i.e., crack width w_c) ranges with respect to the probe diameter, ECT signal for the variation of crack width is suffered for spatial resolution problem which becomes more difficult for dealing with external crack. Moreover, eddy current does not significantly change due to the variation of crack width. This impacts on ECT signals and consequently w_c estimation that leads to higher *NME*. One of the significant outcomes of PLS-OSF sampling is that the obtained training sets significantly improve the learning ability which is robust on noisy test set. Indeed, PLS-OSF/A-RBF shows better prediction accuracy on noisy test sets (i.e., $SNR = 10$ [dB]), while GRID/A-RBF could not predict any defect parameter on noisy data. The overall prediction accu-

racy has shown higher prediction accuracy (Fig. 8) for PLS-OSF sampling compare to GRID sampling by applying all the learning algorithms.

Figure 9 depicts the true vs. predicted plots of crack dimension estimation by applying PLS-OSF sampling and different learning algorithms, testing on noisy data ($SNR = 20$ [dB]). Almost all prediction methods can estimate crack length l_c and height h_c with higher accuracy. Similar to higher estimation error, crack width w_c estimation has shown poorest prediction on noisy test set for all the learning methods. In short, by properly treating ECT data through PLS-OSF sampling, almost all learning algorithms show similar prediction accuracy on synthetic noisy test set data.

4.2. Analysis of computational time efficiency

Concerning quasi real time inversion we are focusing on testing time (*online*). All simulations are done on a standard laptop having 4-core, 2.20 GHz CPU and 8 GB of RAM. A-RBF and SVR take highest (0.16s) and lowest (0.03s) CPU simulation time, respectively for testing on 1000 test samples. SVR deals with the retrieval of number of support vectors during training phase, which is lower than the actual number of training samples. For examples, developing a training model for l_c estimation on a training set of $N = 512$ samples obtained by PLS-OSF sampling strategy, 232 support vectors are retrieved. This makes the learned model more coarser to deal with lower number of kernel function during testing. Thence, SVR takes lowest prediction time. The prediction solution provided by RVR is much sparser than SVR, which retrieved 95 relevance vectors from $N = 512$ samples. Thus, the prediction time comprises the weight multiplication with the test set of 1000 samples that takes only 0.007s. However, RVR needs to calculate de-

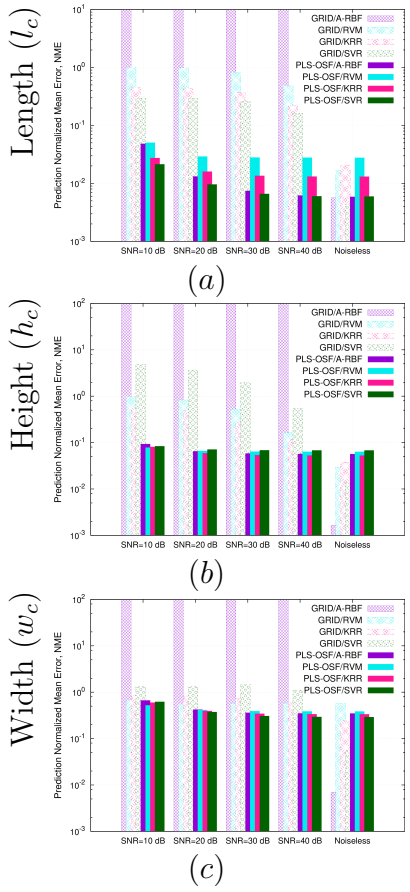

Figure 8: Prediction error, NME versus SNR representation at $(a)\, p_1 = l_c$, $(b)\, p_2 = h_c$ and $(c)\, p_3 = w_c$ when processing noiseless and noisy test set by GRID and PLS-OSF.

sign matrix $\underline{\underline{\phi}}$ based on the training and test set, which takes extra time and increases the total testing time to 0.12s. Whereas, KRR takes 0.07s, that need to deal with all the weights obtained during training phase. Moreover, we can also remark that all prediction techniques take lower prediction time, ranging from 0.03s - 0.16s. This also confirms that all predictors are suitable for quasi real time inversion.

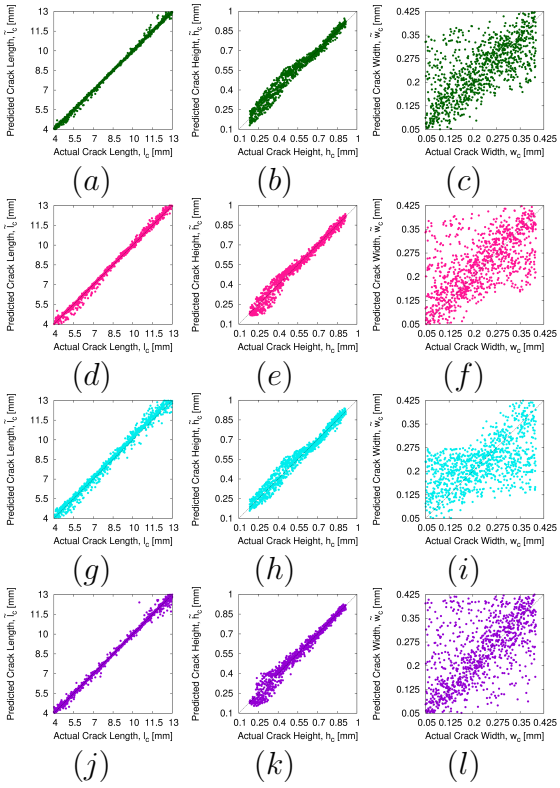

Figure 9: Actual versus predicted values of $p_1 = l_c$ at $(a)(d)(g)(j)$, $p_2 = h_c$ at $(b)(e)(h)(k)$, and $p_3 = w_c$ at $(c)(f)(i)(l)$ when processing noisy test set ($SNR = 20$ [dB]) for training size $N = 512$ through $(a) - (c)$ PLS-OSF/SVR, $(d) - (f)$ PLS-OSF/KRR, $(g) - (i)$ PLS-OSF/RVR and $(j) - (l)$ PLS-OSF/A-RBF.

4.3. Experimental assessment

Referring to Tab. 2, Fig. 10 describes the experimental assessment of crack parameters (l_c, h_c and w_c) estimation through different learning techniques using PLS-OSF sampling strategy by means of relative error, ξ_q ($q = 1, ..., Q$). Among all the prediction techniques, SVR shows highest prediction accuracy for predicting crack 4 (largest crack with $l_c = 12$ mm) length and height in terms of lowest error with relative error $\xi^{(4)}_{l_c(PLS-OSF/SVR)} \approx 0.7\%$

and $\xi^{(4)}_{h_c(PLS-OSF/SVR)} \approx 1\%$, respectively where, RVR also shows lower error for $\xi^{(4)}_{l_c(PLS-OSF/RVR)} \approx 4\%$ and $\xi^{(4)}_{h_c(PLS-OSF/RVR)} \approx 5\%$. KRR and A-RBF exhibits reasonable prediction accuracy on l_c and h_c estimation on the same crack. Though RVR shows highest prediction accuracy ($\xi^{(4)}_{w_c(PLS-OSF/RVR)} \approx 11\%$), similar to numerical evaluation, other learning algorithms have shown higher prediction error for w_c estimation. Let's consider the prediction performance on the smallest crack (i.e., crack 1). KRR and RVR depict higher prediction accuracy with $\xi^{(1)}_{l_c(PLS-OSF/KRR)} \approx 1\%$ and $\xi^{(1)}_{h_c(PLS-OSF/RVR)} \approx 6\%$ respectively. Having regularization capability, SVR also shows lower prediction error for both l_c and h_c estimation ($\xi^{(1)}_{l_c(PLS-OSF/SVR)} \approx 10\%$ and $\xi^{(1)}_{h_c(PLS-OSF/RVR)} \approx 9\%$). The training models have also been treated on some corrupted experiential test samples (e.g., crack 2, 5, 6). Considering the most corrupted crack (i.e., crack 5) shown in Fig. 5, we can observe that almost all learning methods can predict crack height for relative error $\approx 10\%$ (Fig. 10 (e)). However, due to corrupted signal, crack length estimation has been highly affected. In this case, we can also observe that SVR is showing lowest error ($\xi^{(5)}_{l_c(PLS-OSF/SVR)} \approx 27\%$) for crack length estimation. Similarly, SVR outperforms for l_c estimation on crack 2 and 6 with estimation error $\xi^{(2)}_{l_c(PLS-OSF/SVR)} \approx 16\%$ and $\xi^{(6)}_{l_c(PLS-OSF/SVR)} \approx 2\%$. Whereas, KRR and RVR have lowest estimation error on h_c estimation for crack 2 and crack 6, respectively ($\xi^{(2)}_{h_c(PLS-OSF/KRR)} \approx 11\%$, $\xi^{(6)}_{h_c(PLS-OSF/RVR)} \approx 8\%$). Being an interpolator A-RBF fails to predict crack length of all the experimental crack which are effected by noise. Finally, it is quite visible that, through obtaining optimal training set by PLS-OSF sampling, regression methods (i.e., SVR, KRR, RVR) are able to deal with noisy experimental data for crack

characterization. Sparser, regression models (SVR, RVR) are preferable for CPU efficiency. Nevertheless, we can refer that SVR has promising inversion accuracy of experimental data in presence of noise, while RVR shows better accuracy for smaller crack.

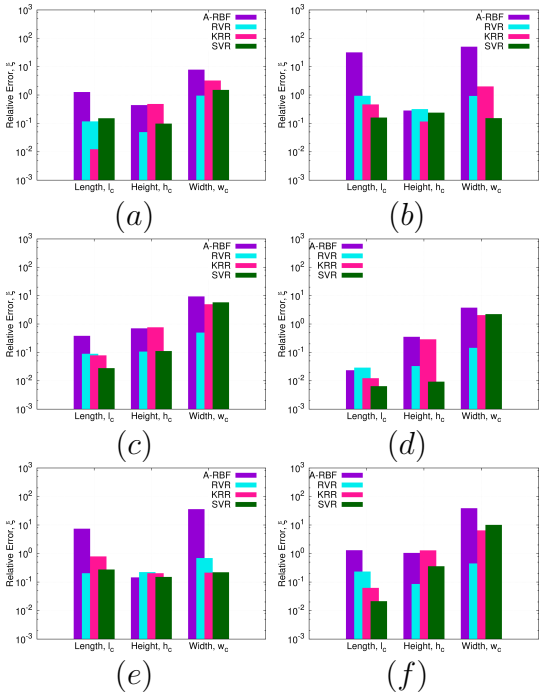

Figure 10: Referring to Tab. 2, relative error, ξ on crack parameters estimation of (a) Crack 1, (b) Crack 2, (c) Crack 3, (d) Crack 4, (e) Crack 5 and (f) Crack 6 for $N = 512$, $F = 1288$ $J = 5$ processed by PLS-OSF sampling through A-RBF, RVR, KRR and SVR algorithms.

5. Conclusion

In this paper, we have shown an adaptive sampling strategy, called PLS-OSF, within the framework of LBE for enhancing the learning abilities of different learning algorithms (e.g., SVR, KRR, RVR and A-RBF). A test

case, consisting a single crack in a conductive plate has been treated for solving crack characterization problem by means of inversion. Comparative analysis of learning algorithms for performing inversion has been numerically performed on the training set obtained by PLS-OSF sampling and compared with more standard GRID sampling approach. Quasi-real time inversion has also been demonstrated by predicting $M = 1000$ test samples within 0.03s to 0.16s during *online* phase on a standard laptop. By applying PLS-OSF sampling strategy, an optimal training set can be retrieved which implies on higher prediction accuracy in presence of noisy data. Moreover, PLS-OSF sampling strategy increases the learning capability of all prediction methods. This strategy also demonstrates higher prediction accuracy for crack dimension estimation on real experimental data. Due to the generalization capability SVR, RVR and KRR exhibit higher prediction accuracy during experimental assessment. Other linear feature extraction techniques such as principle component analysis, linear discriminant analysis, canonical correlation analysis and non linear extensions based on the use of kernels can be straightforwardly applied within the framework of LBE. As different learning algorithms can be useful for estimating different crack parameters by means of optimal prediction accuracy, multi-regression by combining the treated regression techniques, i.e. the framework known under the name of ensemble learning in the machine learning community, may improve the overall prediction accuracy during crack characterization. In authors' opinion, the presented LBE strategy can also be extended to other NdT methods such as ultrasounds testing (UT) and infra-red thermography testing. In particular, due to the characteristics of the treated data, it would be interesting to

apply PLS-OSF on the UT signals issued by post processing (i.e. scattering matrix), collected by ultrasonic arrays [33].

Acknowledgment

This work has been partially supported by the SIRENA project (2014-2017) funded by DIGITEO (France) under the Call for Chairs 2014, the French project ANR-ByPASS.

Annex 1: SIMPLS Algorithm

1. A matrix $\underline{\underline{\mathrm{R}}} = (\underline{\underline{\triangle \mathrm{Z}^*}})^\top \underline{\underline{\mathrm{p}}}^*$ is computed.
2. Run a for loop over J- For $j = 1, ..., J$

- Apply the Singular Value Decomposition to $\underline{\underline{\mathrm{R}}}$.
  If $j > 1$, $\underline{\underline{\mathrm{R}}} = \underline{\underline{\mathrm{R}}} - \left(\underline{\underline{\mathrm{P}}}\right)^\top \left(\underline{\underline{\mathrm{P}}} \left(\underline{\underline{\mathrm{P}}}\right)^\top\right)^{-1} \underline{\underline{\mathrm{PR}}}$;
- The j-th column of $\underline{\underline{\mathrm{W}}}$ denoted by $\overline{\mathrm{W}}_j = (W_{jk};\, k = 1, ..., 2K)$, is set equal to the first left singular vector, where $\underline{\underline{\mathrm{W}}} = \left(\overline{\mathrm{W}}_j;\, j = 1, ..., J\right)$.
- The j-th $\triangle$Z- score denoted by $\overline{\mathrm{T}}_j = \underline{\underline{\triangle \mathrm{Z}^*}} \overline{\mathrm{W}}_j$ is computed while, $\underline{\underline{\mathrm{T}}} = \left(\overline{\mathrm{T}}_j;\, j = 1, ..., J\right)$, $\overline{\mathrm{T}}_j = (T_{nj};\, n = 1, ..., N)$ and $\overline{\mathrm{W}}_j = (\mathrm{W}_{jk};\, k = 1, ..., 2K)$.

- Compute the j-th row of $\underline{\underline{\mathrm{P}}}$ such that
  $\underline{\mathrm{P}}_j = \left(\left(\left(\underline{\underline{\triangle \mathrm{Z}^*}}\right)^\top \overline{\mathrm{T}}_j\right)^\top / \left(\overline{\mathrm{T}}_j\right)^\top \overline{\mathrm{T}}_j\right)$

References

[1] M. Salucci, L. Poli, N. Anselmi, A. Massa, Multifrequency particle swarm optimization for enhanced multiresolution gpr microwave imaging, IEEE Transactions on Geoscience and Remote Sensing 55 (3) (2017) 1305–1317.

[2] M. Salucci, G. Oliveri, A. Massa, Gpr prospecting through an inverse-scattering frequency-hopping multifocusing approach, IEEE Transactions on Geoscience and Remote Sensing 53 (12) (2015) 6573–6592. doi:10.1109/TGRS.2015.2444391.

[3] T. Henriksson, M. Lambert, D.Lesselier, Non-iterative music-type algorithm for eddy-currentusa nondestructive evaluation of metal plates, Studies in Applied Electromagnetics and Mechanics: Elec- tromagnetic Nondestructive Evaluation (XIV) 35 (2011) 22–29.

[4] A. Tamburrino, A. Vento, S. Ventre, A. Maffucci, Monotonicity imaging method for flaw detection in aeronautical applications studies in applied, Electromagnetics and Mechanics: Electromagnet ic Non- destructive Evaluation (XIX) 41 (2016) 284–292.

[5] A. Massa, G. Oliveri, M. Salucci, N. Anselmi, P. Rocca, Learning-by-examples techniques as applied to electromagnetics, Journal of Electromagnetic Waves and Applications 32 (4) (2018) 516–541. doi:10.1080/09205071.2017.1402713.

[6] L. Udpa, S. Udpa, Eddy current defect characterization using neural networks: Materials evaluation, NDT International 23 (6) (1990) 358. doi:https://doi.org/10.1016/0308-9126(90)90883-P.

[7] B. Rao, B. Raj, T. Jayakumar, P. Kalyanasundaram, An artificial neural network for eddy current testing of austenitic stainless steel welds, NDT & E International 35 (6) (2002) 393 – 398. doi:https://doi.org/10.1016/S0963-8695(02)00007-5.

[8] M. Cacciola, F. L. Foresta, F. C. Morabito, M. Versaci, Advanced use of soft computing and eddy current test to evaluate mechanical integrity of metallic plates, NDT & E International 40 (5) (2007) 357 – 362. doi:https://doi.org/10.1016/j.ndteint.2006.12.011.

[9] J. Kim, G. Yang, L. Udpa, S. Udpa, Classification of pulsed eddy current gmr data on aircraft structures, NDT & E International 43 (2) (2010) 141 – 144. doi:https://doi.org/10.1016/j.ndteint.2009.10.003.

[10] X. Chen, D. Hou, L. Zhao, P. Huang, G. Zhang, Study on defect classification in multi-layer structures based on fisher linear discriminate analysis by using pulsed eddy current technique, NDT & E International 67 (Supplement C) (2014) 46 – 54. doi:https://doi.org/10.1016/j.ndteint.2014.07.003.

[11] D. Pasadas, A. Ribeiro, T. Rocha, H. Ramos, 2d surface defect images applying tikhonov regularized inversion and ect, NDT & E International 80 (Supplement C) (2016) 48 – 57.

[12] S. Ahmed, R. Miorelli, M. Salucci, A. Massa, Real-time flaw characterization through learning- by-examples techniques: A comparative study applied to ect, Studies in Applied Electromagnetics and Mechanics: Electromagnetic Nondestructive Evaluation (XX) 42 (2017) 228–235.

[13] M. Salucci, J. Vrba, I. Merunka, A. Massa, Real-time brain stroke detection through a learning-by-examples technique—an experimental assessment, Microwave and Optical Technology Letters 59 (11) (2017) 2796–2799. doi:10.1002/mop.30821.

[14] S. Wold, M. Sjostrom, L. Eriksson, Pls-regression: a basic tool of chemometrics, Chemometrics and Intelligent Laboratory Systems 58 (2001) 109–130. doi:10.1016/S0169-7439(01)00155-1.

[15] S. Bilicz, M. Lambert, S. Gyimothy, Kriging-based generation of optimal databases as forward and inverse surrogate models, Inverse Problems 26 (7) (2010) 074012.

[16] M. Salucci, S. Ahmed, A. Massa, An adaptive learning-by-examples strategy for efficient eddy current testing of conductive structures, in: 10th European Conference on Antennas and Propagation (EuCAP), 2016, pp. 1–4.

[17] M. Salucci, N. Anselmi, G. Oliveri, P. Calmon, R. Miorelli, C. Reboud, A. Massa, Real-time ndt-nde through an innovative adaptive partial least squares svr inversion approach, IEEE Transactions on Geoscience and Remote Sensing 54 (11) (2016) 6818–6832. doi:10.1109/TGRS.2016.2591439.

[18] S. Ahmed, M. Salucci, R. Miorelli, N. Anselmi, G. Oliveri, P. Calmon, C. Reboud, A. Massa, Real time groove characterization combining partial least squares and svr strategies: application to eddy current testing, Journal of Physics: Conference Series 904 (1) (2017) 012017.

[19] V. N. Vapnik, The Nature of Statistical Learning Theory, 2nd Edition, Springer, 1999.

[20] C. Saunders, A. Gammerman, V. Vovk, Ridge regression learning algorithm in dual variables, in: Proceedings of the Fifteenth International

Conference on Machine Learning, ICML '98, Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 1998, pp. 515–521.

[21] M. E. Tipping, Sparse bayesian learning and the relevance vector machine, J. Mach. Learn. Res. 1 (2001) 211–244. doi:10.1162/15324430152748236.

[22] Sui, Yun-Kang, Li, Shan-Po, Guo, Ying-Qiao, An efficient global optimization algorithm based on augmented radial basis function, Int. J. Simul. Multidisci. Des. Optim. 2 (1) (2008) 49–55. doi:10.1051/smdo:2008006.

[23] R. Miorelli, C. Reboud, T. Theodoulidis, N. Poulakis, D. Lesselier, Efficient modeling of ect signals for realistic cracks in layered half-space, IEEE Transactions on Magnetics 49 (6) (2013) 2886–2892. doi:10.1109/TMAG.2012.2236102.

[24] D. Lesselier, A. Razek, Eddy current scattering and inverse scattering, Green's integral and variational formulations., in: E. Pike, P. C. Sabatier (Eds.), Scattering. Scattering and Inverse Scattering in Pure and Applied Science: Part 1 – Scattering of Waves by Macroscopic Targets,, Academic Press, 2002, pp. 486–507.

[25] CEA-LIST, CIVA: Simulation and Analysis for NDT, http://www-civa.cea.fr/en/, [Online; accessed Nov. 2017].

[26] S. de Jong, Simpls: An alternative approach to partial least squares regression, Chemometrics and Intelligent Laboratory Systems 18 (3) (1993) 251 – 263. doi:https://doi.org/10.1016/0169-7439(93)85002-X.

[27] S. Caorsi, D. Anguita, E. Bermani, A. Boni, M. Donelli, A. Massa, A comparative study of nn and svm-based electromagnetic inverse scattering approaches to on-line detection of buried objects, Applied Computational Electromagnetics Society Journal 18 (2) (2003) 1–11.

[28] C. Cortes, V. Vapnik, Support-vector networks, Machine Learning 20 (3) (1995) 273–297. doi:10.1023/A:1022627411411.

[29] Q. Mistral, T. Charret, J. Martens, Array probe implementation (smx) on edf steam generator tubes, in: 10th International Conference on NDE in Relation to Structural Integrity for Nuclear and Pressurizeds, 2013, pp. 329–333.

[30] C.-C. Chang, C.-J. Lin, LIBSVM: A library for support vector machines, ACM Transactions on Intelligent Systems and Technology 2 (2011) 27:1–27:27.

[31] D. E. King, Dlib-ml: A machine learning toolkit, Journal of Machine Learning Research 10 (2009) 1755–1758.

[32] A. Thayananthan, R. Navaratnam, B. Stenger, P. H. S. Torr, R. Cipolla, Pose estimation and tracking using multivariate regression, Pattern Recogn. Lett. 29 (9) (2008) 1302–1310.

[33] L. Bai, A. Velichko, B. W. Drinkwater, Characterization of defects using ultrasonic arrays: a dynamic classifier approach, IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control 62 (12) (2015) 2146–2160.