A Linear Prediction and Support Vector Regression-Based Debonding Detection Method Using Step-Frequency Ground Penetrating Radar

Cédric Le Bastard, Jingjing Pan, *Student Member, IEEE*, Yide Wang, *Member, IEEE*, Meng Sun,
Shreedhar Savant Todkar, Vincent Baltazart, Nicolas Pinel, Amine Ihamouten,
Xavier Dérobert, and Christophe Bourlier

Abstract—In the field of civil engineering, ground penetrating radar (GPR) is a highly efficient nondestructive testing tool for sustainable management of pavement infrastructures. GPR allows to evaluate the structure of the roadway over large distances (with contactless configurations) and to detect significant subsurface defects. This letter presents a new method to detect thin debondings within pavement structures with the step-frequency GPR. The proposed method enables us to carry out the detection with only a small number of frequency samples and A-scans. It is based on the linear prediction and support vector regression theories. Two experimental results show its effectiveness.

Index Terms—Debondings, detection, linear prediction (LP), pavements, step-frequency ground penetrating radar (SF-GPR), support vector regression (SVR).

I. INTRODUCTION

GROUND penetrating radar (GPR) is a widely used tool for evaluating the structure and quality of the road pavement [1], [2] (and the references therein). It allows

rapid nondestructive probing of roadways (the roadways are assumed to be horizontally stratified). The vertical pavement structure can be deduced by backscattered echoes, and more precisely by the estimated time delays and amplitudes of the measured GPR signals. Amplitude estimation is used to retrieve the wave speed within each layer. This letter focuses on the detection of thin interlayer debondings within pavement structures. These debondings have a substantial influence on the residual life of the pavement, and thus, early detection is important for pavement maintenance. Nowadays, a GPR allows to detect significant interlayer debondings [3]–[5]. However, the detection of thin debondings still remains unresolved due to the very small thicknesses of the debonding layers. This challenge remains important for end users (clients or awarding authorities) to optimize the road structures' maintenance once such defects are detected at an early age. This letter presents a new method to detect the thin debondings within pavement structures with the step-frequency ground penetrating radar (SF-GPR).

In the context of thin interlayer debondings, i.e., over-lapped echoes, electromagnetic inversion [1], subspace methods [6], [7], compressive sensing methods [8], deconvolution methods [9], and supervised machine learning-based regression methods [7], [10] are some of the methods that are able to estimate the amplitudes and time delays of backscattered echoes. However, estimation methods, such as subspace methods or supervised machine learning-based regression methods, assume, either explicitly or implicitly, that the number of echoes (or the number of interfaces) is known beforehand. Some criteria (such as Akaike's information criterion or minimum description length [11]) can be used to detect the number of echoes. Nevertheless, in practice, these criteria are not efficient enough due to the high coherence of the echoes. Compressive sensing methods are also a family of promising methods, but their behavior and performance depend on the dictionaries used and the regularization parameters. The behavior of deconvolution methods also depends on the regularization parameters.

Furthermore, buried object detection methods using linear prediction (LP) are proposed in [12]–[14]. LP can be used

to predict the next GPR signal using previous observations. An object is detected when the measured signal is different from its prediction. Yoldemir and Sezgin [14] proposed a least-squares-based LP approach, which assumes that the number of time samples is larger than the number of spatial samples (number of A-scans). This assumption can be perfectly valid in the case of pulse GPR in the time domain. Unfortunately, this assumption may not hold true in the case of SF-GPR. Recently, Pan *et al.* [15] proposed a new forward–backward linear prediction (FBLP) method to estimate the time delays of backscattered GPR echoes with a limited number of frequency samples and A-scans. To achieve this, they proposed to combine the FBLP and support vector regression (SVR) theories. Based on the structural risk minimization principle, the SVR serves as a good machine learning algorithm for sparse modeling [16]. Due to the potential advantages, it is of great interest to extend the detection method of [14] by using the techniques in [15]. Thus, the proposed detection method requires only a small number of frequency samples and A-scans.

The rest of this letter is organized as follows. Section II presents the adopted signal model. Section III briefly presents the conventional LP and the proposed LP-SVR detection methods. Section IV shows the performance of the proposed method with experimental data. Section V gives the concluding remarks.

II. SIGNAL MODEL

The signal model from [6] and [15] is adopted in this letter. It represents the backscattered radar echoes (time-shifted and attenuated replicas of the transmitted signal) from lossless stratified media. In order to use the principle of LP methods, the whitening process that divides each received frequency sample by the corresponding radar pulse is applied. Thus, the signal model used within frequency bandwidth B with N discrete frequencies can be written as follows:

$$r(f_n) = \frac{g(f_n)}{e(f_n)} = \sum_{k=1}^{K} s_k \exp(-j2\pi f_n t_k) + u(f_n) \quad (1)$$

with $u(f_n) = (m(f_n)/e(f_n))$; K represents the number of echoes; $e(f_n)$ and $m(f_n)$ are the GPR pulse and an additive Gaussian white noise with zero mean and variance σ^2 at frequency f_n, respectively. Frequency f_n is defined as $f_n = f_1 + (n-1)\Delta f$, with $n = 1, 2 \ldots, N$, f_1 the lowest frequency, and Δf the frequency shift. s_k and t_k are the amplitude and time delay of the kth backscattered echo, respectively. For N discrete frequencies within the bandwidth B, the whitened received signal $\mathbf{r}$ can be written in the following vector form:

$$\mathbf{r} = \mathbf{As} + \mathbf{u} \quad (2)$$

where $\mathbf{r} = [r(f_1), r(f_2), \ldots, r(f_N)]^T$ is the data vector representing the measurements from a step frequency radar to which the whitening procedure is applied; the superscript T denotes the transpose operation; $\mathbf{A} = [\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_K]$ is the

mode matrix and $\mathbf{a}_k$, the mode vector, is defined as: $\mathbf{a}_k = [e^{-j2\pi f_1 t_k}, e^{-j2\pi f_2 t_k}, \ldots, e^{-j2\pi f_N t_k}]^T$; $\mathbf{s} = [s_1, s_2, \ldots, s_K]^T$ is the vector composed of the reflected echoes' amplitudes; $\mathbf{u} = [u(f_1), u(f_2), \ldots, u(f_N)]^T$ is the complex noise vector.

For a B-scan composed of L A-scans, we can define matrix $\mathbf{R}$, in the frequency domain, as follows:

$$\mathbf{R} = [\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_L] \quad (3)$$

with $\mathbf{r}_i$ the ith A-scan in the frequency domain and $i = 1, 2 \ldots, L$.

III. LINEAR PREDICTION DETECTION METHOD

A. Conventional LP

The LP-based detection method of [14] developed in the time domain is briefly presented here in the frequency domain. It is used to predict the next whitened GPR signal $\mathbf{y} = [y(f_1), y(f_2), \ldots, y(f_N)]^T$ from the previous observations $\mathbf{R}$. Debonding is detected when the whitened measured signal, $\mathbf{r}_i = \mathbf{z} = [z(f_1), z(f_2), \ldots, z(f_N)]^T$, with $i > L$, is different from the prediction $\mathbf{y}$. LP estimates the unknown samples with a linear combination of the known observations by minimizing the mean square prediction error, $\mathbf{e}$, as follows:

$$\mathbf{R}\omega = \mathbf{y} \quad (4)$$

$$\mathbf{e} = \mathbf{z} - \mathbf{y} \quad (5)$$

where $\mathbf{y} \in \mathbb{C}^{N \times 1}$, $\mathbf{R} \in \mathbb{C}^{N \times L}$, and $\boldsymbol{\omega} \in \mathbb{C}^{L \times 1}$. The weighting coefficient vector $\boldsymbol{\omega}$ can be estimated by the least-squares approach

$$\boldsymbol{\omega} = (\mathbf{R}^H \mathbf{R})^{-1} \mathbf{R}^H \mathbf{z} \quad (6)$$

where the superscript H denotes the conjugate transpose operation. The Euclidean norm of the error, $\|\mathbf{e}\|_2$, indicates the possible occurrence of a defect (debonding). The larger the error is, the higher the chance of detection will be.

However, the inversion of $\mathbf{R}^H \mathbf{R}$ requires the inequality constraint: $N > L$. As discussed in Section I, the assumption made in [14] (i.e., $N > L$) is not necessarily valid when using SF-GPR.

B. Proposed Method: LP-SVR

The objective of LP is to estimate the weighting vector $\boldsymbol{\omega}$. When the weighting vector $\boldsymbol{\omega}$ is estimated, we can predict the whitened A-scan $\mathbf{y}$ using (4). In the context where the number of A-scans (L) is larger than the number of frequency samples, the weighting vector cannot be well-estimated and the original LP method cannot work. Notice that (4) represents a typical form of SVR. Thus, we propose to use the SVR theory to estimate the weighting vector $\boldsymbol{\omega}$ without the inequality constraint ($N > L$). The SVR method aims to find a hyperplane that fits the data within a deviation less than a given value ϵ. The ϵ-insensitive loss function is used here [17]. The optimization

problem can be written as follows [15]:

$$\min_{\boldsymbol{\omega},\xi^r,\hat{\xi}^r,\xi^i,\hat{\xi}^i} \left(\frac{1}{2}\|\boldsymbol{\omega}\|^2 + C\sum_{n=1}^{N}\left(\xi_n^r + \hat{\xi}_n^r + \xi_n^i + \hat{\xi}_n^i\right) \right)$$

$$\text{s.t.} \begin{cases} \text{Re}\left(y_n - \mathbf{x}_n^H\boldsymbol{\omega}\right) \leqslant \epsilon + \xi_n^r, & n = 1,\dots,N \\ \text{Re}\left(-y_n + \mathbf{x}_n^H\boldsymbol{\omega}\right) \leqslant \epsilon + \hat{\xi}_n^r, & n = 1,\dots,N \\ \text{Im}\left(y_n - \mathbf{x}_n^H\boldsymbol{\omega}\right) \leqslant \epsilon + \xi_n^i, & n = 1,\dots,N \\ \text{Im}\left(-y_n + \mathbf{x}_n^H\boldsymbol{\omega}\right) \leqslant \epsilon + \hat{\xi}_n^i, & n = 1,\dots,N \\ \xi_n^r, \hat{\xi}_n^x, \xi_n^i, \hat{\xi}_n^i \geqslant 0, & n = 1,\dots,N \end{cases}$$

$$(7)$$

where $\mathbf{x}_n$ is the nth column of $\mathbf{R}^H$ and $y_n = y(f_n)$; ξ_n^r and $\hat{\xi}_n^r$ stand, respectively, for positive and negative errors in the real part of the output y_n, while ξ_n^i and $\hat{\xi}_n^i$ stand for the corresponding imaginary part; Re$(\cdot)$ and Im$(\cdot)$ denote the real and imaginary parts, respectively; the value C controls the tradeoff between the structural and empirical errors. To solve (7), we employ the Lagrangian multipliers and the Karush–Kuhn–Tucker theorem with Wirtinger's calculus on the complex variable $\boldsymbol{\omega}$ [15]. Thus, the weighting vector $\boldsymbol{\omega}$ can be written as follows:

$$\boldsymbol{\omega} = \sum_{n=1}^{N}\left((a_n - \hat{a}_n) + i(b_n - \hat{b}_n)\right)\mathbf{x}_n \qquad (8)$$

with a_n, $\hat{a}_n$, b_n, and $\hat{b}_n$, the Lagrangian multipliers.

Define $\boldsymbol{a} = [a_1,\dots,a_N]^T$, $\hat{\boldsymbol{a}} = [\hat{a}_1,\dots,\hat{a}_N]^T$, $\boldsymbol{b} = [b_1,\dots,b_N]^T$, and $\hat{\boldsymbol{b}} = [\hat{b}_1,\dots,\hat{b}_N]^T$. The Lagrange multiplier vectors $\boldsymbol{a}$, $\hat{\boldsymbol{a}}$, $\boldsymbol{b}$, and $\hat{\boldsymbol{b}}$ can be computed from the maximization problem as follows:

$$\max_{\boldsymbol{a},\hat{\boldsymbol{a}},\boldsymbol{b},\hat{\boldsymbol{b}}} -\frac{1}{2}\begin{bmatrix}\boldsymbol{a}-\hat{\boldsymbol{a}}\\\boldsymbol{b}-\hat{\boldsymbol{b}}\end{bmatrix}^T \begin{bmatrix}\text{Re}(\mathbf{R}\mathbf{R}^H) & -\text{Im}(\mathbf{R}\mathbf{R}^H)\\ \text{Im}(\mathbf{R}\mathbf{R}^H) & \text{Re}(\mathbf{R}\mathbf{R}^H)\end{bmatrix}\begin{bmatrix}\boldsymbol{a}-\hat{\boldsymbol{a}}\\\boldsymbol{b}-\hat{\boldsymbol{b}}\end{bmatrix}$$
$$-\epsilon\mathbf{1}^T(\boldsymbol{a}+\hat{\boldsymbol{a}}+\boldsymbol{b}+\hat{\boldsymbol{b}})$$
$$+\text{Re}(\mathbf{y}^T)(\boldsymbol{a}-\hat{\boldsymbol{a}})+\text{Im}(\mathbf{y}^T)(\boldsymbol{b}-\hat{\boldsymbol{b}})$$
$$\text{s.t. } 0 \leq a_n, \hat{a}_n, b_n, \hat{b}_n \leq C, \quad n = 1,\dots,N \qquad (9)$$

where $\mathbf{1}$ is an all-one column vector with N elements. In the calculation, a small identity term $\gamma\mathbf{I}$ is added in the cost function in the case of ill-conditional inaccuracies [17]. Thus, the weighting vector $\boldsymbol{\omega}$ can be obtained.

Compared with the conventional LP, the proposed LP-SVR needs additional computation. However, the number of frequencies and A-scans is limited. Thus, the increase of computational burden is very low. A comparison of the average execution time between LP and LP-SVR will be given in Section IV.

IV. EXPERIMENTAL RESULTS

In this section, the proposed method is tested with two experimental data sets from IFSTTAR carousel [18], [19]. Two media are probed with SF-GPR. The first probed area is composed of two and three layers, as shown in Fig. 1. Layers 1 and 3 are asphalt surface pavements and layer 2 is made of sand. Layer 2 represents artificial debondings with a thickness about 0.5–1 cm. The measurements have been carried out at 10-, 50-, and 200-Kcycles loading stages [18].
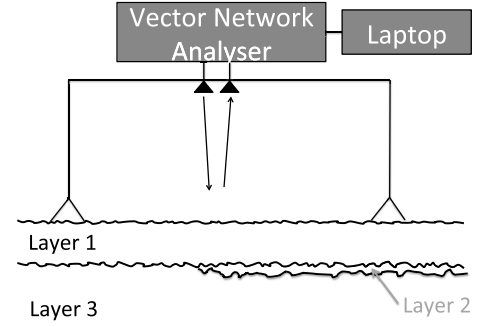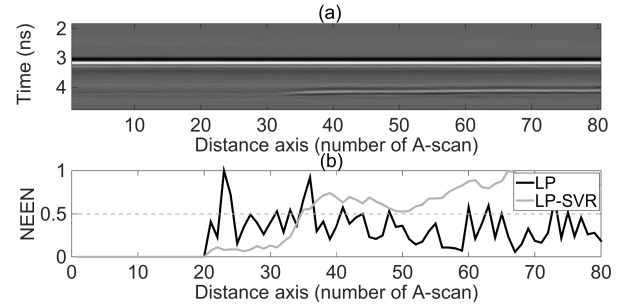

Fig. 1. Horizontal stratified medium.

Fig. 2. (a) B-scan from SF-GPR in the time domain, 10 Kcycles; $B = [0.8 - 10.8]$ GHz. (b) Detection results from LP and LP-SVR; $L = 20$ and $N = 11$.
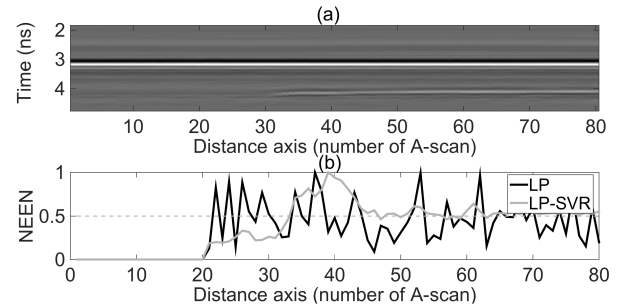

Fig. 3. Experimental results for the same parameters as in Fig. 2, excepted for the number of Kcycles; here, 50 Kcycles.

A quasi-monostatic step frequency radar with transmitter (Tx) and receiver (Rx) antennas close to each other is used. The distance between Tx and Rx is 20 cm and it is constant during the measurements. For each experiment, the height of the antennas is set to 70 cm above the ground level. A frequency bandwidth of 10 GHz, $B = [0.8–10.8]$ GHz, with $N = 11$ samples, is used. Before using the proposed method, we apply two preprocessing techniques. First, a time filtering is used to delete the air wave. Second, data whitening by the pulse is applied [18]. The radar pulse is measured from a metal plane. The first time delay (first interface) is considered as constant in the B-scan. The number of A-scans (L) used to estimate the weighting vector $\boldsymbol{\omega}$ is set to 20. The SVR-related parameters are set like in [15], i.e., $\epsilon = 0$, $C = 1$, and $\gamma = 10^{-6}$.

Figs. 2(a), 3(a), and 4(a) show the B-scans in the time domain obtained from SF-GPR at various Kcycle loading stages. Figs. 2(b), 3(b), and 4(b) show the normalized error
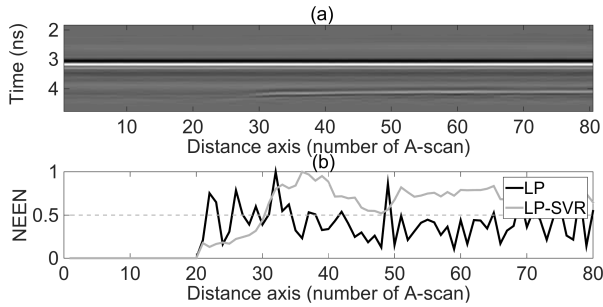
Fig. 4.   Experimental results for the same parameters as in Fig. 2, excepted for the number of Kcycles; here, 200 Kcycles.
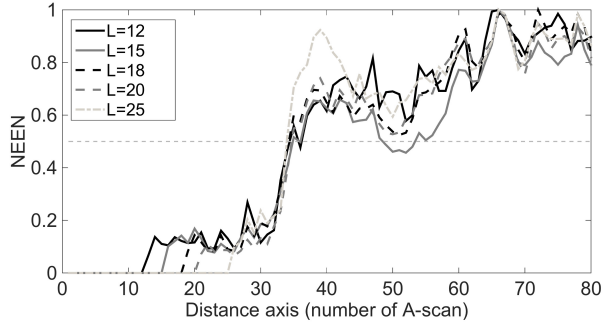

Fig. 5.   LP-SVR results from different number of previous A-scans (L); 10 Kcycles, $B = [0.8 - 10.8]$ GHz, and $N = 11$.

Fig. 6.   Detection results from LP-SVR and LP; 10 Kcycles, $B = [0.8 - 10.8]$ GHz, $N = 30$, and $L = 20$.

Fig. 7.   (a) B-scan from SF-GPR, $B = [0.8 - 10.8]$ GHz, tack-free interface. (b) Detection results from LP and the proposed method; $L = 20$, $B = [1 - 6]$ GHz, and $N = 11$.

Euclidean norm (NEEN), defined as $\|\mathbf{e}\|_2 / \max(\|\mathbf{e}\|_2)$, with LP and LP-SVR. The method detects the debonding when the NEEN is greater than a certain threshold. Following various measurements, we have observed that the detection threshold should be set between 0.4 and 0.5. Thus, we have chosen to set the detection threshold to 0.5, which is represented in Figs. 2(b), 3(b), 4(b), 5, 6, and 7(b) by the dotted horizontal gray line. In Figs. 2–4, L is superior to N and as a consequence LP cannot work. Only LP-SVR is able to detect the debonding presence. We can see that LP-SVR allows to detect the presence of the debonding from the 35th A-scan, the 34th A-scan, and the 31th A-scan for 10-, 50-, and 200-Kcycle loading stages, respectively. These results have been obtained by a computer with a processor unit (CPU) at 2.3 GHz and 8GB RAM. The average execution time for one A-scan is about 0.022 s for Figs. 2(b), 3(b), and 4(b).

Fig. 5 shows the NEEN with LP-SVR for 10-Kcycle loading stage with respect to the number of previous A-scans, L. We can see that the results are almost the same with different values of L. Debonding is detected from the 35th or 36th A-scan.

Fig. 6 shows the NEEN from LP and LP-SVR for 10-Kcycle loading stage. Here, we consider the case where $N > L$. The behavior of both the methods is almost the same in this context. Debonding is detected from the 35th A-scan. Furthermore, the average execution time for one A-scan is 2.4×10^{-4} and 0.08 s, for LP and LP-SVR, respectively.

To analyze the behavior of LP-SVR for thin debondings, the second experiment is carried out on a medium composed of two and three layers, as shown in Fig. 1. In this case, layer 2 is a thin air gap (occurring due to the uncoat interface
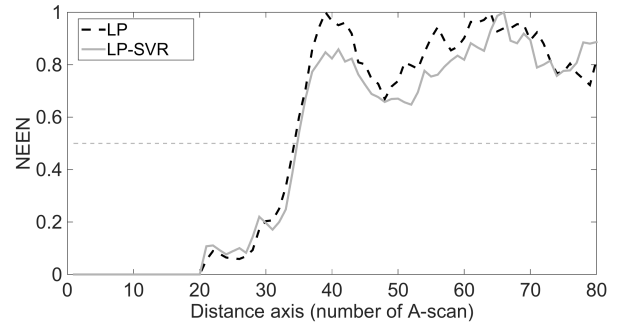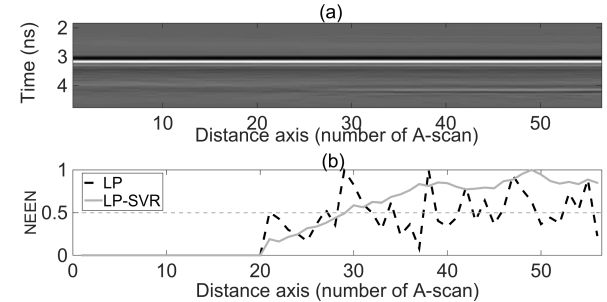
between layers 1 and 3). Fig. 7(a) shows the B-scan obtained from SF-GPR in the time domain with a frequency bandwidth $B = [0.8 - 10.8]$ GHz. Fig. 7(b) shows the NEEN from LP and LP-SVR. The frequency bandwidth used here for LP and LP-SVR is $B = [1 - 6]$ GHz, with $N = 11$ samples. The number of A-scans used for the estimation is $L = 20$. Fig. 7(b) shows that only LP-SVR is able to detect the debonding presence from the 30th A-scan.

The proposed LP-SVR method has also been tested on simulated data for a medium composed of three interfaces of which the first one is horizontal and the last two ones are modeled by a wedge model. The results indicate the effectiveness of the proposed method in detecting debondings with different layer thicknesses.

V. CONCLUSION

In this letter, we propose a new method based on LP and SVR theories to detect thin debondings within pavement structures using SF-GPR. Unlike the conventional LP method, the proposed method is not constrained by the number of frequency samples and A-scans. In the particular context where the number of frequency samples is smaller than the number of A-scans ($N < L$), the proposed method outperforms the conventional LP method. The experimental results have shown the efficiency of the proposed method. In perspective, future work will be focused on the theoretical evaluation of the detection threshold.

REFERENCES

[1] A. Benedetto and L. Pajewski, *Civil Engineering Applications of Ground Penetrating Radar*. New York, NY, USA: Springer, 2015.

[2] W. W.-L. Lai, X. Dérobert, and P. Annan, "A review of ground penetrating radar application in civil engineering: A 30-year journey from locating and testing to imaging and diagnosis," *NDT&E Int.*, vol. 96, pp. 58–78, Jun. 2018.

[3] T. Saarenketo and T. Scullion, "Road evaluation with ground penetrating radar," *J. Appl. Geophys.*, vol. 43, nos. 2–4, pp. 119–138, 2000.

[4] L. Krysiński and J. Sudyka, "Typology of reflections in the assessment of the interlayer bonding condition of the bituminous pavement by the use of an impulse high-frequency ground-penetrating radar," *Nondestruct. Test. Eval.*, vol. 27, no. 3, pp. 219–227, 2012.

[5] V. Marecos, M. Solla, S. Fontul, and V. Antunes, "Assessing the pavement subgrade by combining different non-destructive methods," *Construct. Building Mater.*, vol. 135, pp. 76–85, Mar. 2017.

[6] L. Qu, Q. Sun, T. Yang, L. Zhang, and Y. Sun, "Time-delay estimation for ground penetrating radar using ESPRIT with improved spatial smoothingtechnique," *IEEE Geosci. Remote Sens. Lett.*, vol. 11, no. 8, pp. 1315–1319, Aug. 2014.

[7] S. Meschino, L. Pajewski, M. Pastorino, A. Randazzo, and G. Schettini, "Detection of subsurface metallic utilities by means of a SAP technique: Comparing MUSIC- and SVM-based approaches," *J. Appl. Geophys.*, vol. 97, pp. 60–68, Oct. 2013.

[8] J. Li, C. Le Bastard, Y. Wang, G. Wei, B. Ma, and M. Sun, "Enhanced GPR signal for layered media time-delay estimation in low-SNR scenario," *IEEE Geosci. Remote Sens. Lett.*, vol. 13, no. 3, pp. 299–303, Mar. 2016.

[9] S. Zhao and I. L. Al-Qadi, "Development of regularization methods on simulated ground-penetrating radar signals to predict thin asphalt overlay thickness," *Signal Process.*, vol. 132, pp. 261–271, Mar. 2017.

[10] C. Le Bastard, Y. Wang, V. Baltazart, and X. Derobert, "Time delay and permittivity estimation by ground-penetrating radar with support vector regression," *IEEE Geosci. Remote Sens. Lett.*, vol. 11, no. 4, pp. 873–877, Apr. 2014.

[11] A. Badawy, T. Salman, T. Elfouly, T. Khattab, A. Mohamed, and M. Guizani, "Estimating the number of sources in white Gaussian noise: Simple eigenvalues based approaches," *IET Signal Process.*, vol. 11, no. 6, pp. 663–673, Aug. 2017.

[12] P. van Genderen and I. Nicolaescu, "Imaging of stepped frequency continuous wave GPR data using the Yule–Walker parametric method," in *Proc. IEEE Radar Conf. (EURAD)*, Oct. 2005, pp. 77–80.

[13] K. C. Ho and P. D. Gader, "A linear prediction land mine detection algorithm for hand held ground penetrating radar," *IEEE Trans. Geosci. Remote Sens.*, vol. 40, no. 6, pp. 1374–1384, Jun. 2002.

[14] A. B. Yoldemir and M. Sezgin, "A least squares approach to buried object detection using ground penetrating radar," *IEEE Sensors J.*, vol. 11, no. 6, pp. 1337–1341, Jun. 2011.

[15] J. Pan, C. Le Bastard, Y. Wang, and M. Sun, "Time-delay estimation using ground-penetrating radar with a support vector regression-based linear prediction method," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 5, pp. 2833–2840, May 2018.

[16] V. N. Vapnik, *Statistical Learning Theory*. New York, NY, USA: Wiley, 1998.

[17] M. Martínez-Ramón and C. Christodoulou, *Support Vector Machines for Antenna Array Processing and Electromagnetics* (Lectures on Computational Electromagnetics) vol. 1. San Rafael, CA, USA: Morgan & Claypool, 2005, pp. 1–120.

[18] M. Sun *et al.*, "Time delay and interface roughness estimation using modified ESPRIT with interpolated spatial smoothing technique," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 3, pp. 1475–1484, Mar. 2018.

[19] J.-M. Simonin *et al.*, "Case study of detection of artificial defects in an experimental pavement structure using 3D GPR systems," in *Proc. 15th Int. Conf. Ground Penetrating Radar*, Brussels, Belgium, Jun./Jul. 2014, pp. 847–851.