A new method to simultaneously estimate the radius of a cylindrical object and the wave propagation velocity from GPR data

Aleksandar Vaso Ristic*, Dusan Petrovacki, Miro Govedarica
Faculty of technical sciences, University of Novi Sad, Trg Dositeja Obradovica 6, 21000 Novi Sad, Serbia

ARTICLE INFO
Article history:
Received 4 January 2008
Received in revised form
22 October 2008
Accepted 14 January 2009

Keywords:
Radar scan
Hyperbolic reflection
Least squares fitting
Cylindrical object radius
Propagation velocity

ABSTRACT

We present a new method to simultaneously estimate cylindrical object radius (R) and electromagnetic (EM) wave propagation velocity (v) from ground penetrating radar (GPR) data. R estimation methods have been investigated since the middle of the previous decade, but studies have become more intensive and important over the last several years since they increase the utility of GPR data and enable new GPR applications. Since existing methods, according to the author's best knowledge, are based on a priori known v, the proposed method has an advantage: it eliminates the measurement of v and its influence on R estimation quality. Estimating v accurately results in better soil characterisation.

Three steps are used to simultaneously estimate v and R. First, using the extracted raw data, the coordinates of the hyperbola apex (x_0, t_0) are estimated. Second, the boundary speed (v_0) is estimated, based on the previous results. In the final step, v is reduced from v_0 to a predefined v_min. From the analysis of propagation velocity choice criterion, an optimal v is chosen, which is used to calculate a unique R. This proposed method is a nonlinear least squares fitting procedure. The method is implemented and verified, using data collected under real conditions, in a Matlab environment.

A comparison of the proposed and existing methods shows that the new method is significantly more accurate and robust with regard to noise and the amount of raw data.

© 2009 Elsevier Ltd. All rights reserved.

1. Introduction

Ground penetrating radar (GPR) refers to a range of electromagnetic (EM) techniques designed primarily to locate objects or interfaces buried beneath the earth’s surface or located within a visually opaque structure (Daniels, 2004). GPR techniques can be divided into shallow surface and deep borehole surveying. This paper analyses a proposed method to interpret raw data from radar scans that are acquired using the shallow surface surveying technique.

Radar scans of cylindrical underground objects have specific hyperbolic shapes (hyperbolic reflections). Their geometries contain information on various parameters, such as depth, radius, spatial orientation and the relative permittivity of the soil (ε_R) surrounding the object. The process of estimating parameters from radar scans first requires that the raw data be incorporated into mathematical models of hyperbolic reflection. Thereafter, the relevant parameters are calculated based on fitted hyperbolic geometries.

Although rules for radar scan generation are well-known, procedures are still being developed to automate data interpretation. It is particularly important to characterise cylindrical underground objects in terms of R because these objects include pipelines (gas, water, oil, sewage, etc.), cylindrical tanks and cables (energy, optical, signal). Since the cadastre of underground utilities is very important in developing complex urban communities, automating the interpretation of GPR raw data is key. If v is known, it is possible to characterise the soil in the scanning zone in terms of volumetric moisture content (θ [m^3/m^3]) according to the equation from Topp (Topp et al., 1980; Huisman et al., 2003). In addition, we can compute the spatial movement of moisture through the soil, thickness of soil layers (necessary in land consolidation), thickness of mud layers in canals, rivers and lakes, status of flooding banks, etc.

This paper aims to more completely characterise detected cylindrical objects independent of the operator’s subjectivity and assumptions about the value of v. The goal of this paper is to construct a method that is more accurate and robust to noise and that is independent of the amount of raw data.

We make a few assumptions in this paper. The first is related to estimating the average v of a medium on top of a cylindrical underground object. We do this by considering a localised object in a homogeneous medium. We also assume a straight ray path when modelling hyperbolic reflection. Since in real conditions a cylindrical underground object is in heterogeneous soil, the v of the surrounding medium can vary significantly, and EM waves...
may reflect and/or refract from soil layers depending on the frequency emitted by the antenna. These circumstances may affect estimation results. We note that the authors of all the methods analysed in this paper used these assumptions. The reasons and conditions for these assumptions are explicitly defined by Al-Nuaimy et al. (2000) and Borgioli et al. (2008).

In subsequent sections, we verify the characteristics of the method mentioned above with data collected under real conditions, and we compare them to results obtained using existing methods.

2. Problem overview

The first assumptions in estimating \( R \) were proposed by Stolte and Nick (1994). These authors explored the functional dependence between \( R \) and the hyperbolic eccentricity. Shihab et al. (2004) showed that hyperbolic eccentricity is only a function of \( \nu \). Three different approaches to data analysis have been proposed in the literature:

1. The first concept is based on directly estimating \( R \) from a representative sample set of raw data, using the generalised Hough transform (Borgioli et al., 2008, Windsor et al., 2005a, b). This approach assumes the existence of four variables: \( R \), \( \nu \), and the hyperbola apex coordinates \( X_0 \) and \( t_0 \). The system of four-specific hyperbola equations with these four variables must be solved. The four points that best represent the raw data are extracted using a statistical condition (e.g., coordinate differentials). The presence of stochastic in correlating the variables, the influence of noise and the amount of raw data are such that, according to Borgioli et al. (2008), Windsor et al. (2005a, b), the robustness and quality of the results cannot be guaranteed. It is often impossible to estimate all four variables. If \( \nu \) is known a priori, then it is possible to estimate \( R \) in an acceptable range for problems with low noise and a larger input dataset. The error in determining \( \nu \) has a significant influence on the estimation error of \( R \), as will be presented in this paper.

2. The second concept is based on an interactive interpretation of hyperbolic reflection from radar scans. Olhoeft (2000) thoroughly describes this interactive procedure, and suggests further improvements and estimation results using an approach that is independent of the mathematical details of the problem. The principle of this method involves visual overlap of a predefined hyperbola with the previously processed scan. The exact depth and \( \nu \) are determined according to the angle between hyperbola asymptotes. \( R \) is estimated according to the curvature hyperbola in the apex and given its width. According to Olhoeft (2000), the processing results do not have to be unique, and a high-quality scan is crucial. In addition, it is important to determine the exact hyperbola centre position, since it is unknown. This problem was not thoroughly analysed until Shihab and Al-Nuaimy (2005), who showed that the \( R \) estimation error caused by an undefined hyperbola centre position may be as much as 40–50%. Yufryakov and Linnikov (2006) also use the interactive procedure, which, with a detailed geometrical description, estimates \( \nu \) from a 3D scan according to a measured initial value. It also measures \( R \) and spatial orientation parameters. This concept expands the list of estimated values, but problems of scan quality and determining the exact hyperbola centre position remain.

3. The third concept is based on representing raw data with a mathematically described fitted hyperbola and further hyperbola geometry interpretations in the context of estimating \( R \). The results of Shihab et al. (2004) serve as a foundation for further research into all three concepts. These authors presented the concept of raw data fitting with a second-order polynomial using direct least squares fitting (Fitzgibbon et al., 1996, 1999), modified in terms of hyperbola-forced constraints. The final \( R \) is estimated assuming that \( \nu \) is known a priori or measured. Together with other analyses mentioned in this section, Shihab and Al-Nuaimy (2005) introduced robustness-related improvements to the algorithm proposed by Halir and Flusser (1998) and O’Leary and Zsombor-Murray (2004). In this paper and one by Shihab and Al-Nuaimy (2006), the fitting concept is abandoned because it does not include the impact of \( \nu \). Dolgiy et al. (2006) proposed a hyperbola-specific fitting technique using weighted least squares with respect to the mathematical error expectation of two-way travel time. The technique is based on previous measurement-based knowledge of \( \nu \). They proposed using the common midpoint (CMP) or point reflector analysis methods. By analysing various test site scans, the authors discuss the quality of the fitted data. The use of estimation results in soil characterisation was analysed by Petrovacki and Ristic (2007).

3. Data collection

In GPR, a transmitting antenna sends polarised, high-frequency EM waves into the soil. Various soil inhomogeneities cause reflection, while other parts of the EM waves refract through deeper layers until they become too weak (Daniels, 2004). The time needed for EM waves to travel the distance \( r_0 \) (m) from the transmitting antenna to a border surface and back to a receiver is the two-way travel time \( t_0 \) (ns). Fig. 1 shows how a radar scan is generated: the X-axis represents the antenna’s trajectory \( x_0 \) (N is the scanning resolution (scan/m)), while the Y-axis represents \( t_0 \), the two-way travel time of the reflected EM wave. A regular hyperbolic reflection originates when the antenna’s trajectory is orthogonal to the object axis (transversal scan).

The distance between the antenna and the object changes as the antenna moves. Changes in the distances of points \( x_0 \), \( x_0 \), ..., \( x_0 \) are marked with lines \( r_0 \), ..., \( r_0 \), ..., \( r_0 \) (Fig. 1a). If these lines are orthogonal to the antenna’s trajectory, connecting the successive end points of the lines produces the geometric shape of a hyperbola (Fig. 1a).

The shortest line \( r_0 \) related to the hyperbola apex, represents the relative depth of the object. The radar scan also contains information about the amplitude of a reflected signal. This value is a searching criterion for the scan and the object’s location and depth.

Considering that the antenna’s EM beam has an angle between 35° and 45°, it is possible to detect a cylindrical object even if it is
not directly beneath the antenna. The transmitting antenna sends EM waves orthogonally to a surface such that the angle between the hyperbola axis and the antenna’s trajectory is 90°. Fig. 1b represents an ideal radar scan of an object in homogeneous soil while the antenna moves orthogonally to the object’s axis. Under real-world conditions, raw data contain a high level of noise due to the soil’s inhomogeneities and hyperbolic reflections originating from other objects.

The antenna used to acquire data transmits a pulse with a certain polarity: a positive peak first, then a negative peak (possibly followed by a second positive). Since the reflected signal is a copy of the transmitted pulse, in the case of a metal cylindrical underground object, the polarity remains the same. When the cylindrical underground object is made of a non-metal and is filled with air, polarity inversion occurs (phase change of π) because of the low dielectric constant associated with air. The reflected signal may contain repeated reflections following the first, strongest reflection signal. These reflections are caused by reverberation of radiated EM waves between the underground object and the antenna. They may also emanate from the side (between layers adjacent to the underground object). Therefore, the raw radar scan must be pre-processed in order to remove undesired system and ground effects. This usually involves time zero detection, background clutter removal and low pass filtering (Al-Nuaimy et al., 2000). Raw data extraction was performed in each scan by picking the highest reflected intensity from the zone that features hyperbolic patterns. Fig. 2 shows a greyscale radar linescan of a steel gas pipe with a diameter of 35.56 cm (14”), where white bands indicate positive and black bands represent negative peaks. The black dots are raw data points extracted from the white hyperbolic reflection (maximum of reflected signal-positive peak). Extracted point coordinates \((x_i, t_i)\) are used as input data for the fitting operation.

As shown in Al-Nuaimy et al. (2000) and Lossani and Gamba (2000), raw data extraction can be performed using pattern recognition and image processing techniques. An input picture is divided into layers, which are then classified and extracted.

In our proposed method, after raw data extraction there are no overlapping hyperbolic reflections that originate from closely separated cylindrical underground objects. If there are overlapping hyperbolic reflections, Borgioli et al. (2008) and Windsor et al. (2005a, b) showed that the Hough transform provides a possible solution.

4. Model description

This paper presents an application of a third concept, described in Section 2. This method uses a canonical hyperbola equation to model raw data generated by the hyperbolic reflection \((x_i, t_i)\) in the case of cylindrical objects. The canonical hyperbola equation allows dependencies to be created between the fitted hyperbola geometry (Fig. 3) and the estimated parameters \(v\) and \(R\). The raw hyperbolic reflection data can also be represented using a second-order polynomial with a hyperbola-forced constraint. This approach is not appropriate to take into account the influence of \(v\), as will be shown at the end of this section.

According to our analyses of the EM wave trajectories, it is possible to define the dependence between the fitted hyperbola geometry and estimated parameters (Fig. 3). The distance from the antenna centre at an arbitrary position \(x_i\) \((m)\) \((-N, 0, \ldots, N)\) to the cylindrical object is denoted \(r_i\). The distance at the time when it is exactly above the cylindrical object’s axis is given by \(r_0\) and \(R\) is the radius. When a triangle is formed, it follows that:

\[
(r_i + R)^2 = (r_0 + R)^2 + (x_i - x_0)^2.
\]

If the arbitrary distance from the antenna’s centre to the pipe \(r_i\) is expressed with \(v\) and \(t_i\) as \(r_i = (vt_i)/2\), i.e., the relative cylindrical object depth \(r_0\) as \(r_0 = (vt_0)/2\), then using expression (1), it is possible to write the hyperbola equation as a function of \(R, v, x_0\) and \(t_0\):

\[
t_i = \frac{2}{R} \sqrt{\frac{(vt_0^2)}{2} + R} + (x_i - x_0)^2 - R.
\]

Each pair of raw data \((x_i, t_i)\), extracted from a hyperbolic reflection of finite \(R\) satisfies Eq. (2), depending on the noise level. Eq. (2) can be rewritten as the canonical hyperbola equation, centred around \((x_0, -2R/v)\):

\[
\frac{(t_i + (2R/v))^2}{(t_0 + (2R/v))^2} - \frac{(x_i - x_0)^2}{(2R/v) + R} = 1.
\]

From Eq. (3), the hyperbola semi axes are:

\[
a = t_0 + \frac{2R}{v},
\]

\[
b = \frac{v^2}{2} a.
\]

The angle between the hyperbola asymptotes (Fig. 4) is defined using Eqs. (4) and (5):

\[
tg \varphi = \frac{b}{a} = \frac{v}{2}.
\]

Eq. (6) shows that the angle \(\varphi\) is directly proportional to \(v\). This means that with increasing \(v\), the angle \(\varphi\) increases (for the same \(R\)) and vice versa.

Consistent with the above facts, when a cylindrical object of finite \(R\) and a point reflector \((R=0)\) are scanned under the same conditions, hyperbola asymptotes form the same angle for both objects, since they have the same \(v\) (Fig. 4). Modifying Eq. (3) for a
point reflector gives a hyperbola centred around \((x_0, 0)\)

\[
\left( \frac{t_0}{t} \right)^2 - \frac{2(x - x_0)}{vt_0} = 1.
\]

(7)

It is clear from Eqs. (3) and (7) that the position of the hyperbola centre is different for a point reflector than for a cylindrical object with finite \(R\). Using model (7) for a cylindrical object with finite \(R\) produces an artificially high \(v\).

From the previous discussion, it follows that a second-order polynomial with a hyperbola-forced constraint does not provide exact values for the semi axes given by Eqs. (4) and (5). Another problem with this interpretation is that the forced hyperbola is not Euclidean distance. This does not include the influence of \(v\) on the noise level. Consistent with the sensitivity as well as the complexity of the algorithm, it is more efficient to apply a robust weighted least squares procedure similar to Euclidean distance minimisation.

Fig. 5 represents the implementation flowchart of simultaneous \(v\) and \(R\) estimation, which is performed using three steps. These steps are described in detail later in this section, including approaches necessary to overcome previously mentioned impacts.

5. Method implementation

Eq. (2) represents the nonlinear estimation model with four variables: \(x_0, t_0, R\) and \(v\). The estimation results depend on:

(a) Correlations between estimated variables.

- Functional dependencies between variables are strong and well-defined consistent with Eq. (2). The presence of errors in real raw data along with correlations may produce an unsatisfactory solution, but when the errors for the data pairs are sufficiently low, precise solutions can be calculated.
- Experiments show that \(v\) has the strongest influence on \(R\) estimation. This functional dependence is nonlinear. It is possible to determine \(R = R(v)\) through a number of experiments appropriate to soil type, relative object depth, value of \(R\), etc. Defining additional functional dependencies and criteria would overcome any adverse impacts on the quality of the correlation.

(b) The impact of soil structure on the noise level.

- Raw data acquisition is performed under real-world conditions with unknown and inhomogeneous soil layers above the pipe. This means that the change in \(v\) is stochastic, and that the noise level in the raw data is high. In order to use this method, a more robust fitting algorithm and/or outlier removal procedure must be applied.

(c) Fitting procedure characteristics.

- Procedures (Fitzgibbon et al., 1996; Halir and Flusser, 1998; O’Leary and Zsombor-Murray, 2004) that minimise the algebraic distance are insufficiently robust to noise, due to the characteristics of the algorithm. Using a second-order polynomial with a hyperbola-forced constraint does not account for the impact of \(v\). These problems can be solved using a nonlinear model (2) with a robust fitting algorithm.
- Mathematical error expectation ("statistical" distance) minimisation procedures (Kanatani, 1994; Zhang, 1997; Halir\(^1\)) correct the fitting results in terms of the equalising impact of points from low and high curvature sections. The results of this method are slightly better than in previous procedures, since they are designed for larger sets of raw data and are optimal only in the sense of being unbiased. The statistical analysis does not include the impact of \(v\).
- Euclidean distance minimisation procedures (e.g., Boggs et al., 1992) strongly depend on initial values and on the noise level. Consistent with the sensitivity as well as the complexity of the algorithm, it is more efficient to apply a robust weighted least squares procedure similar to Euclidean distance minimisation.

5.1. Step I: estimation of hyperbola apex coordinates \((x_0, t_0)\)

\((x_0, t_0)\) is estimated using a modified Levenberg–Marquardt method (Marquardt, 1963; Moré and Sorensen, 1983). The applied algorithm is robust and was adapted to solve nonlinear problems using the least squares method.

The basic task of the first step is to decrease the number of correlations between the estimated parameters, i.e., to reduce the problem dimensionality from four to two correlated parameters. The optimality criterion is now

\[
\min_{x_0, t_0} \sum_{i=1}^{n} \left( f(x_0, t_0, x_{data(i)} - t_{data(i)})^2, \right.
\]

where \([x_0, t_0]^T\) is the vector of estimated coefficients, \(x_{data}\) and \(t_{data}\) are raw row data coordinate vectors of \(n\) points, and \(f(x_0, t_0, x_{data})\) is the nonlinear model from Eq. (2). Expression (8) shows that the function minimises the sum of the squared residuals, i.e., the algebraic distance.

All procedures mentioned in this chapter can be more or less successfully used to estimate \((x_0, t_0)\). Section 6 provides a comparative presentation of estimation results. The estimated values \((x_0, t_0)\) are the input data for the next step.

5.2. Step II: estimation of boundary speed \(v_0 = v_{\text{max}}\)

Since \(v\) is unknown (or approximately known, which is insufficient to accurately estimate \(R\), an additional condition to simultaneously estimate \(v\) and \(R\) is defined as the velocity range \([v_{\text{max}} - v_{\text{min}}]\).

When a model for \((R > 0)\) is used, the calculated \(v_0\) is higher than the real \(v\) represented in Fig. 4. A propagation velocity higher than \(v_0\) does not make sense, because it produces negative values of \(R\). In the final step, it is then possible to use a refining procedure for the specified velocity range, and with proper criteria, to choose an optimally estimated \(v\). The value of \(v\) decreases, starting from \(v_0 = v_{\text{max}}\) in steps of \(\Delta v\).

The boundary velocity is estimated iteratively by varying \( \nu \). The value closest to satisfying the condition \( R \geq 0 \) (i.e., \( a \approx t_o \) according to (4)) is accepted as \( t_o \). The estimated \( t_o \) includes a nonlinear model (2) with an optimality criterion that minimises the sum of absolute errors.

\[
\min_{t_o} \sum_{i=1}^{n} |f(t_o, x_{data}) - t_{data}|.
\]

5.3. Step III: simultaneous estimation of \( \nu \) and \( R \)

5.3.1. Propagation velocity choice criterion

An optimal \( \nu \) is chosen from the finite set of possible solutions using the velocity choice criterion. With this method, a first-order optimality criterion (10) of vector \( \nu \) is chosen, defined as

\[
\text{foo} = \|g^T p\|_{\infty}.
\]

Where the symbol \((.)^T\) represents the member-wise product of two vectors, e.g., \( \{a_i\} \times \{b_i\} = \{a_i \times b_i\} \). The infinity (Chebyshev or maximum) norm of a vector \( x \in \mathbb{R}^n \) is defined as \( \|x\|_{\infty} = \max_{1 \leq i \leq n} |x_i| \), where \( x_i \) is the \( i \)th component of the vector (Boyd and Vandenberghe, 2004).

In expression (10), \( g \) represents the gradient (11) of the optimality criterion (12) for variable \( w = [\nu, R] \)

\[
g_i = \frac{\partial f(w, x_{data})}{\partial w_i},
\]

\[
\min_{w} \sum_{i=1}^{n} |f(w, x_{data}) - t_{data}|^2.
\]

If the optimality criterion (12) is represented as a minimisation problem with constraints

\[
\min_{p(w_i)} \left\{ g_i < 0(w_i \downarrow) \land u_i < \infty \Rightarrow p_i = w_i - u_i \right\}
\]

\[
\begin{array}{l}
g_i < 0 \land u_i = \infty \Rightarrow p_i = w_i - l_i \\
g_i > 0 \land u_i = \infty \Rightarrow p_i = 1 \\
g_i > 0 \land l_i = \infty \Rightarrow p_i = 1
\end{array}
\]

For problems of type (13), the necessary condition of optimality is that \( \text{foo}(10) \) with \( p(w_i) \) defined in Eq. (14) becomes zero (Coleman and Li, 1994, 1996). When the solution is within the feasible region, the condition is satisfied because the gradient \( g \) (11) vanishes; on the other hand, when the solution is on the boundary, \( \text{foo} \) vanishes due to the definition of vector \( p(w_i) \) (14). As a result, the value of \( \text{foo} \) can be taken as a measure of the accuracy of the obtained solution.

The vectors of the lower and upper bounds can be formed consistent with the physical properties of the estimated variables \( \nu \) and \( R \) (e.g., \( 1 < \nu < 28 \text{ cm/ns} \); \( 1 < R < 100 \text{ cm} \)).

When expression (12) is used with model (2) to simultaneously estimate \( \nu \) and \( R \), \( \text{foo} \) represents the sum of absolute values of the first derivation by \( w = [\nu, R] \) for \( n \) raw data points

\[
\text{foo} = \sum_{i=1}^{n} |f(w, x_{data})|.
\]

For optimally estimated \( \nu \), Eq. (4) is applied to directly calculate \( R \).

5.3.2. Stopping criterion \( \nu_{\text{min}} \)

When analysing our experiments, we noticed that changes in \( \text{foo}(\nu) \) exhibited a certain regularity. Disregarding outliers, \( \text{foo}(\nu) \) exponentially rises in the range \([\nu_0 - \nu_{\text{min}}] \). Rapid changes or saddle points of the optimality criterion function (12) cause a minimum of \( \text{foo}(\nu) \) to appear. After the first minimum of \( \text{foo}(\nu) \), this function behaviour repeats periodically. Experimental analysis shows that the first minimum always has the smallest root mean square error (RMSE), which in terms of algebraic distance minimisation means that the first minimum is global. All other minima result in physically irregular values of \( R \), confirming that the first minimum is global.

Using this conclusion, it is possible to define a stopping criterion \( \nu_{\text{min}} \) as the value of \( \nu \) after the first minimum where \( \text{foo}(\nu) \) starts to rise rapidly.

5.3.3. Simultaneous estimation of \( \nu \) and \( R \)—an overview

Simultaneous estimation of \( \nu \) and \( R \) is based on a nonlinear model (2), with an optimality criterion that minimises the sum of the squared residuals (12), with known \( (x_o, t_o) \) and \( \nu \) varied in the range \([\nu_0 - \nu_{\text{min}}] \). For an optimally estimated \( \nu \), a choice criterion is
6.2. Estimation of \((x_0, t_0)\) — step I

\((x_0, t_0)\) was estimated using our proposed method and was compared with existing fitting procedures. Procedures were as listed in Section 5. Tables 3 and 4 contain the results of the comparative analysis.

- The Fitzgibbon procedure was not successful in 5 out of 10 examples, due to:
  - high noise levels (examples 3, 8, 10),
  - a small number of points located on the hyperbola apex (example 6), or
  - high SNR (example 4).

Since the success rate was only 50% and convergence problems were always present, this procedure can only be used for initial analyses.

- The Halir procedure (modified Fitzgibbon) was unsuccessful in 4 out of 10 examples. This procedure was only successful in removing convergence problems caused by high SNR (example 4). For examples 8 and 10, the procedure converged, but the fitted hyperbola was stretched. Since raw data with high SNR are rare, the problems cited for the previous method remain.

- The O’Leary procedure is in the same form as the Halir procedure, with a modified implementation, and the estimation results are not significantly different.

- The Kanatani procedure does not solve the problems mentioned above, although the results for the \(t_0\) coordinate are slightly better.

- The results of the Boggs procedure depend strongly on the initial values. Because the previous procedure was not completely successful, the results of the proposed method were used as initial values. The results of the Boggs procedure were similar to the proposed method, since the initial values were close to the exact solution. Another problem is that this procedure does not take into account the influence of \(v\).

- Results in Tables 3 and 4 show that the proposed method is successful for all examples (higher/lower SNR, few points, poor disposition of points and hyperbola segment absence). Outliers were not excluded in these tests. Overall, the proposed method works better than the others in two ways:
  - It estimates \((x_0, t_0)\) well in cases that feature very few points, which are common in real-world field scanning conditions. This is also an advantage as it avoids the need for raw data correction by means of manual point addition, thus eliminating the possible influence of these points on the final estimation results.
  - It estimates \((x_0, t_0)\) well in cases of higher noise, which are common in real-world field scanning conditions.

6.3. Estimation of \(v_0\) — step II

Table 5 shows the final estimation results for steps II and III of the proposed method. The second column contains estimated \(v_0\)
Table 2
Raw data and fitting results.

<table>
<thead>
<tr>
<th>Example 1</th>
<th>Example 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>72 points, 400MHz GSSI 5103 antenna</td>
<td>125 points, 400MHz GSSI 5103 antenna</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Example 3</th>
<th>Example 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>40 points, 400MHz GSSI 5103 antenna</td>
<td>39 points, 400MHz GSSI 5103 antenna</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Example 5</th>
<th>Example 6</th>
</tr>
</thead>
<tbody>
<tr>
<td>20 points, 400MHz GSSI 5103 antenna</td>
<td>11 points, 400MHz GSSI 5103 antenna</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Example 7</th>
<th>Example 8</th>
</tr>
</thead>
<tbody>
<tr>
<td>85 points, 400MHz GSSI 5103 antenna</td>
<td>95 points, 800MHz Zond-12c antenna</td>
</tr>
</tbody>
</table>
### Table 3
Comparative estimation of \((x_0, t_0)\) results—examples 1–5.

<table>
<thead>
<tr>
<th>Method</th>
<th>Example 1</th>
<th>Example 2</th>
<th>Example 3</th>
<th>Example 4</th>
<th>Example 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fitzgibbon</td>
<td>(x_0 = 133.585)</td>
<td>(x_0 = 515.402)</td>
<td>Not OK (cannot fit)</td>
<td>Not OK (cannot fit)</td>
<td>(x_0 = 122.271)</td>
</tr>
<tr>
<td></td>
<td>(t_0 = 14.8454)</td>
<td>(t_0 = 18.5341)</td>
<td></td>
<td>(t_0 = 10.3303)</td>
<td></td>
</tr>
<tr>
<td>Halir</td>
<td>(x_0 = 133.585)</td>
<td>(x_0 = 515.402)</td>
<td>Not OK (cannot fit)</td>
<td>(x_0 = 446.662)</td>
<td>(x_0 = 122.271)</td>
</tr>
<tr>
<td></td>
<td>(t_0 = 14.8454)</td>
<td>(t_0 = 18.5341)</td>
<td></td>
<td>(t_0 = 8.5622)</td>
<td>(t_0 = 10.3303)</td>
</tr>
<tr>
<td>O'Leary</td>
<td>(x_0 = 133.631)</td>
<td>(x_0 = 515.402)</td>
<td>Not OK (cannot fit)</td>
<td>(x_0 = 446.803)</td>
<td>(x_0 = 122.404)</td>
</tr>
<tr>
<td></td>
<td>(t_0 = 14.8452)</td>
<td>(t_0 = 18.5342)</td>
<td></td>
<td>(t_0 = 8.5622)</td>
<td>(t_0 = 10.3303)</td>
</tr>
<tr>
<td>Halir (Kanatani)</td>
<td>(x_0 = 133.882)</td>
<td>(x_0 = 515.402)</td>
<td>Not OK (cannot fit)</td>
<td>(x_0 = 446.753)</td>
<td>(x_0 = 122.271)</td>
</tr>
<tr>
<td></td>
<td>(t_0 = 14.8561)</td>
<td>(t_0 = 18.5461)</td>
<td></td>
<td>(t_0 = 8.5622)</td>
<td>(t_0 = 10.3303)</td>
</tr>
<tr>
<td>Boggs</td>
<td>(x_0 = 133.722)</td>
<td>(x_0 = 515.422)</td>
<td>(x_0 = 165.237)</td>
<td>(x_0 = 446.522)</td>
<td>(x_0 = 122.312)</td>
</tr>
<tr>
<td></td>
<td>(t_0 = 14.9015)</td>
<td>(t_0 = 18.6496)</td>
<td>(t_0 = 7.5330)</td>
<td>(t_0 = 8.6061)</td>
<td>(t_0 = 10.3554)</td>
</tr>
<tr>
<td>Proposed method</td>
<td>(x_0 = 135.000)</td>
<td>(x_0 = 515.610)</td>
<td>(x_0 = 165.370)</td>
<td>(x_0 = 446.610)</td>
<td>(x_0 = 123.200)</td>
</tr>
<tr>
<td></td>
<td>(t_0 = 14.8962)</td>
<td>(t_0 = 18.6500)</td>
<td>(t_0 = 7.5330)</td>
<td>(t_0 = 8.5840)</td>
<td>(t_0 = 10.3768)</td>
</tr>
</tbody>
</table>

### Table 4
Comparative estimation of \((x_0, t_0)\) results—examples 6–10.

<table>
<thead>
<tr>
<th>Method</th>
<th>Example 6</th>
<th>Example 7</th>
<th>Example 8</th>
<th>Example 9</th>
<th>Example 10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fitzgibbon</td>
<td>Not OK (cannot fit) parabola</td>
<td>(x_0 = 145.733)</td>
<td>(t_0 = 17.1012)</td>
<td>(x_0 = 442.571)</td>
<td>(t_0 = 8.8532)</td>
</tr>
<tr>
<td>Halir</td>
<td>Not OK (cannot fit)</td>
<td>(x_0 = 145.632)</td>
<td>(t_0 = 17.1022)</td>
<td>(x_0 = 442.571)</td>
<td>(t_0 = 8.8532)</td>
</tr>
<tr>
<td>O'Leary</td>
<td>Not OK (cannot fit)</td>
<td>(x_0 = 145.831)</td>
<td>(t_0 = 17.1022)</td>
<td>(x_0 = 442.571)</td>
<td>(t_0 = 8.8532)</td>
</tr>
<tr>
<td>Halir (Kanatani)</td>
<td>Not OK (cannot fit)</td>
<td>(x_0 = 145.604)</td>
<td>(t_0 = 17.1001)</td>
<td>(x_0 = 442.571)</td>
<td>(t_0 = 8.8532)</td>
</tr>
<tr>
<td>Boggs</td>
<td>(x_0 = 543.193)</td>
<td>(x_0 = 145.642)</td>
<td>(t_0 = 17.2541)</td>
<td>(x_0 = 444.285)</td>
<td>(t_0 = 9.0737)</td>
</tr>
<tr>
<td>Proposed method</td>
<td>(x_0 = 544.040)</td>
<td>(x_0 = 145.853)</td>
<td>(t_0 = 17.3218)</td>
<td>(x_0 = 443.910)</td>
<td>(t_0 = 9.8767)</td>
</tr>
</tbody>
</table>
values. By comparing these with \( v_{\text{EST}} \) in column 5, two ranges can be determined:

- For smaller \( R \) (5–12 cm), the difference between \( v_{\text{EST}} \) and \( v_0 \) is between 0.27 and 0.9 cm/ns. This means that for this range, it is almost impossible to use the measured values, given the accuracy of existing methods for measuring \( v \).
- For bigger \( R \) (12–36 cm), the difference between \( v_{\text{EST}} \) and \( v_0 \) is between 1 and 2 cm/ns. For this range, the difference is less than 2 cm/ns, regardless of \( R \) and the soil structure.

### 6.4. Simultaneous estimation of \( v \) and \( R \)—step III

Column 3 shows that values of \( v_0 \) are close to zero for optimally estimated \( v \) and \( R \). The differences between estimated values of the semi-axis \( a \) in column 4 and coordinate \( t_0 \) of the hyperbola apex (Tables 3 and 4) determine the ratio \( 2R/v \), as in Eq. (4). The radius estimation error \( \text{err}_R \) is between 2% and 10% in all examples. In all cases, the maximum error does not impact our ability to estimate standard values of \( R \). This is an important verification of the accuracy with which our proposed method can estimate \( R \).

#### 6.4.1. Propagation velocity choice criterion \( \text{foo} \)

Figs. 6 and 7 illustrate the calculation of \( R(v) \) and \( \text{foo}(v) \) in example 1.

Fig. 6 shows that \( R \) is sensitive to \( v \), as the function \( R(v) \). This means that estimation methods based on \( a \) priori known \( v \) (measured) cannot generate an acceptable tolerance for \( R \) (e.g., up to 10%). It is also clear that \( R(v) \) is nonlinear. Fig. 7 shows the segment of the curve \( \text{foo}(v) \) containing the minimum repeats in cycles. Its shape is similar across all the examples.

#### 6.4.2. Stopping criterion \( v_{\text{min}} \)

The velocity range in example 1 \((v_0 = 14.657 \text{ to } v = 12 \text{ cm/ns})\) contains two minima for \( \text{foo}(v) \). The first, for \( v_{\text{EST}} = 13.57 \text{ cm/ns} \), has \( \text{RMSE} = 0.1209 \), and the second, for \( v_{\text{EST}} = 12.41 \text{ cm/ns} \), has \( \text{RMSE} = 0.1249 \). Since the first \( \text{RMSE} \) is the smallest, it is a global minimum. The second minimum results in physically irregular values of \( R \) (39.62 cm), verifying that the first minimum is global.
In addition, $v_{\text{EST}}$ for the second minimum is 2.247 cm/ns lower than $v_0$ and is outside the range proposed in Section 6.3. Consistent with these facts, we chose the stopping criterion $v_{\text{min}}$ as the value of $v$ after the first minimum when $f_{\text{oo}}(v)$ starts to rise rapidly. For instance, in example 1, the chosen value is $v_{\text{min}} = 13$ cm/ns.

### 6.5. Comparison with other methods

$R$ estimation results are compared for examples 8–10. For existing methods, $v$ is a priori known, while for our proposed method, it is estimated from the raw data as previously described. Data for examples 8 and 9 were taken from Dolgiy et al. (2006) and those for example 10 were from Shihab and Al-Nuaimy (2005). These data are for metal pipes and were collected from test sites where $v$ was either measured or known.

The data in Table 6 show that the proposed method produces better results than do the methods in the papers mentioned above. Using some standard optimisation procedures (e.g., recursive Kalman filter, maximum likelihood, Nelder–Mead) on the fitted points does not produce a more accurate estimate for $R$.

### Table 6

| Example | $R_{\text{REAL}}$ (cm) | $R_{\text{EST}}$ (cm) | $|\Delta R|$ (cm) | $\text{err}_R$ (%) | $R_{\text{EST}}$ (cm) | $|\Delta R|$ (cm) | $\text{err}_R$ (%) | $R_{\text{EST}}$ (cm) | $|\Delta R|$ (cm) | $\text{err}_R$ (%) |
|---------|------------------|-----------------|-----------------|------------------|------------------|-----------------|------------------|------------------|-----------------|------------------|
| 8       | $R = 8$          | 9.41            | 1.41            | 17.62            | -                | -               | -                | 7.38             | 0.62            | 7.75             |
| 9       | $R = 26.5$       | 29.94           | 3.44            | 12.98            | -                | -               | -                | 25.83            | 0.67            | 2.53             |
| 10      | $R = 5.0$        | -               | -               | -                | 5.44             | 0.44            | 8.80             | 4.78             | 0.22            | 4.40             |

Simultaneous $v$ and $R$ estimation was achieved using three steps. The final processing results are optimally estimated values of $v$ and $R$ with an estimation error of up to 10%. Our proposed method uses a nonlinear model with an optimality criterion that minimises the sum of the squares of the residuals. The choice criterion for an optimally estimated $v$ is the first minimum of the first-order optimality parameter in the function $f_{\text{oo}}(v)$. The optimally estimated $v$ determines a unique value of $R$.

Comparison with existing procedures shows that this new method is significantly more accurate and robust in terms of noise and the amount of raw data.

The values of $R$ estimated using this method could be used to improve the completeness of the underground utility cadastre as well as to facilitate field excavation work. A low estimation error allows $R$ to be accurately predicted for utility pipes with standard radii.

Values of $v$ estimated using our proposed method can be used to determine the volumetric moisture content in soil, the spatial disposition of moisture and other characteristics of the soil surrounding the buried utility.

### References


Shihab, S., Al-Nuaimy, W., 2006. Hyperbola fitter for characterisation of cylindrical targets in GPR data. In: Proceedings 5th International Conference on Ground Penetrating Radar, Columbus, Ohio, USA. 4.1 Utility detection, UTL4 on CD.


