Identification of Horizontal Circular Arc from Spatial Data Sources

IDENTIFICATION OF HORIZONTAL CIRCULAR ARC FROM SPATIAL DATA SOURCES

Marco Bassani* Ph.D P.Eng., Giuseppe Marinelli Ph.D P.Eng., Marco Piras Ph.D P.Eng.

Marco Bassani (* = Corresponding author)
Associate Professor
Politecnico di Torino, Department of Environment, Land and Infrastructures Engineering (DIATI)
24, corso Duca degli Abruzzi, Torino, I, 10129
Phone: +39 011 090 5635, E-mail: marco.bassani@polito.it

Giuseppe Marinelli
Research Associate
Politecnico di Torino, Department of Environment, Land and Infrastructures Engineering (DIATI)
24, corso Duca degli Abruzzi, Torino, I, 10129
Phone: +39 011 090 5635, E-mail: giuseppe.marinelli@polito.it

Marco Piras
Assistant Professor
Politecnico di Torino, Department of Environment, Land and Infrastructures Engineering (DIATI)
24, corso Duca degli Abruzzi, Torino, I, 10129
Phone: +39 011 090 7661, E-mail: marco.piras@polito.it

ABSTRACT

Curves represent the most interesting element to be analyzed when evaluating the operational and safety effects of road geometrics. Despite the existence of several static and dynamic survey techniques, the correct identification of curves and the interpretation of spatial data from horizontal road alignment surveys still represent a challenge.

The paper presents the results of an investigation into the effects of both accuracy and sampling frequency (i.e., the interval between data points) of alignment spatial data on the identification of the radius and center of curvature. In pursuit of this objective, the authors adopted an original approach consisting of the generation of dispersed spatial data points from circular arcs with known characteristics. The Least Squares, Huber and Landau methods were used to back-calculate the radius and the coordinates of the center of curvature, and to assess the effects of accuracy and interval between data points.

The results obtained quantified the effects of accuracy, interval between points, radius magnitude, and length of circular arc on the identification of circular arcs. In support of future survey activities, considerations regarding the accuracy of measurements, the sampling frequency, and arc characteristics (i.e., length, radius) are discussed in the paper.

INTRODUCTION

The identification of geometric features of the horizontal alignment of roads is of interest in many fields, including those of cadastral and surveying, quality control and assurance (QC/QA) in construction, safety and driver behavioral studies, and lately, automotive engineering. In the field of highway engineering, of the basic geometric elements that form the horizontal alignment (i.e., tangents, circular arcs and, when present, spirals), it is the circular arcs which are of particular interest. In fact, they are estimated by highway agencies when fixing curve speed limits and performing highway safety analysis (Carlson et al. 2005). The segment with constant curvature represents the most critical element since it is characterized by the highest crash frequency (Hauer 1999) as a consequence of stability issues for vehicles and visibility issues for drivers.

A field survey of curves is necessary when no alignment data are available, as tends to be the case for most existing roads and in the QC/QA of new constructions. Although alignment data are sometimes available when they have been included in project drawings, a lot of time and effort is required to retrieve them in a numerical format useful for the above-mentioned applications.

Hence, in the recent past, several research efforts have aimed to derive curve data characteristics from a number of different sources (Harkey et al. 2004). In particular, several attempts were made to derive alignment information from aerial and satellite georeferenced images (Easa et al. 2005; Mena and Malpica 2005; Deng et al. 2007), from GIS roadway maps (Li et al. 2012), or from data collected with positioning sensors on mobile systems and vehicles such as GPS/GNSS devices, laser scanners, LiDAR, inertial units, and hybrid tools/technologies (Carlson et al. 2005; Drakopoulos and Örnek 2000; Choi and Sung 2007; Imran et al. 2006; Bassani et al. 2012; Jimenez 2011; Lee et al. 2012).

All these technologies are characterized by different levels of precision and accuracy; therefore, when employed in the evaluation of the geometric characteristics of horizontal alignments, different results can be expected. The evaluation of circular arcs poses a particular

challenge since deviation from the correct/expected value depends on the precision and accuracy of the survey technology employed, the spatial distribution of surveyed points denoting the alignment (i.e., sampling frequency), the geometric characteristics of the arc (i.e., length, radius, and angle of deviation/curvature), and the mathematical algorithms employed for the identification and characterization of the elements (Kasa 1976; Chan 1965; Berman and Culpin 1986).

Starting from the state of the art and knowledge available in literature, the authors have endeavored to make a contribution to the discussion on this topic with their investigations into the effects of accuracy, and sampling frequency of alignment spatial data on the evaluation of circular arc characteristics.

BACKGROUND

In the past (and also today, albeit rarely), engineers designed the horizontal alignment by filleting successive tangents with circular arcs (Figure 1a). Nowadays, the standard policy with the geometric design of highways and streets is to recommend or, sometimes, impose the use of spirals between tangents and arcs forming combined curves (Figure 1b), which significantly improve drivability, comfort, road esthetics and safety. The parametric equation of spiral used in road design is:

$$R \cdot L_S = A^2 \qquad \text{(eq. 1)}$$

where R is the radius of the circular arc, L_S is the length of the spiral, and A is its scale factor.

As can be seen in Figure 1b, when spiral transitions are used in the design of horizontal alignments, the two tangents and the curve must be separated by a distance ΔR (also named offset) and the two spirals included between the two sets of points TS (tangent to spiral) and SC (spiral to curve), CS (curve to spiral) and ST (spiral to tangent).

On existing alignments, spirals can be identified only if tangents and circular arcs are detected first. In fact, when the offset ΔR and radius R are known, the following set of approximate

equations can be used to calculate the geometric characteristics of a spiral, starting from its scale parameter A:

$$A = \sqrt[4]{\Delta R \cdot R^3} \quad \text{(eq. 2)}$$

the abscissa X_C (which is fundamental for the positioning of the coordinate system - X,Y):

$$X_C = \frac{L_S}{2} \quad \text{(eq. 3)}$$

the coordinates of the final point (X_S,Y_S):

$$X_S = L_C \quad \text{(eq. 4)}$$

$$Y_S = \frac{L_S^3}{6 \cdot A^2} \quad \text{(eq. 5)}$$

and the deviation angle (θ_S):

$$\theta_S = \frac{A^2}{2 \cdot R^2} \quad \text{(eq. 6)}$$

The calculation of the geometric parameters of spirals assumes knowledge of the radius R and the center coordinates (X_C, Y_C) of the circular arc. Hence, any algorithm developed for the analysis of horizontal alignments and which may or may not include spiral transitions, must start with a complete and precise identification of circular arcs.

The geometry of a horizontal alignment can be extracted (manually or automatically) from satellite or aerial imagery, or by performing a field survey with a variety of devices and methods, each one characterized by specific levels of precision, accuracy and sampling frequency.

There are several geospatial techniques dedicated to the acquisition of data through terrestrial or aerial measurements. These techniques can be distinguished, one from the other, by their accuracy, data quality, survey costs, acquisition and processing times, etc. Starting from literature (Kraus 1997; Gomarasca 2009; Novak 1995; McGlone 2013; Maclean 1994; Hofmann-Wellenhof et al. 2008), the current various techniques and their characteristics are summarized in Table 1.

The methods listed in Table 1 are used to produce multi-purpose maps and geospatial models. Each of them has advantages and disadvantages in terms of costs and times for acquisition and processing, and the most appropriate one to be used depends on the environment morphology (i.e., urban canyon, open sky, countryside, etc.), the size of the area to be surveyed, and the desired accuracy of the final product.

At present, mobile mapping is the most widespread technology used to collect 3D data of road geometrics, since it leads to the fast acquisition of accurate data. Although aerial photogrammetry with the use of Unmanned Aerial Vehicles (UAV) also produces data of comparable quality, it faces regulation problems when flying over critical areas, such as populated zones.

Rasdorf et al. (2010) evaluated three GIS curve identification methods for the prediction of the radius of precisely drawn curves from field curve data. While the benchmark for the three applications may be considered satisfactory when applied to points extracted from the circular curves, the results obtained from an interpretation of the field data measured were less accurate as a consequence of the poor quality of the GIS line work that models those curves in the GIS. Only 12 to 24% of radii were within a 10% range of the actual field data, thus demonstrating the importance of having high-quality spatial data.

Drakopoulos and Örnek (2000) presented the results of an investigation in which several digital devices including a distance measuring device (odometer), a vertical gyroscope (lateral slope and gradient), and a gyro-compass (azimuth) were used simultaneously. From the combination of measurements, a panel of 3D raw data was gathered with points spaced at constant intervals of 16.1 m. An algorithm developed to provide the horizontal curve length (L_C), degree of curve, and deviation angle (α), was validated on two-lane highways. Comparisons with as-built information show high percentage deviations in both curve length (between -48% and +55%) and deflection angle, (between -34% and +58%), with more reliable results in the case of curves longer than 300 m. A lower percentage deviation of between -17% and +3% was observed for the degree of

curve. The point of curvature (from tangent to curve, TC) and point of tangency (from curve to tangent, CT) were detected with an accuracy of ±2 observations (see Figure 1a).

Imran et al. (2006) illustrated an integrated GPS/GIS-based procedure for the identification of the horizontal alignment based on the tracking of the path of an instrumented vehicle moving in the two directions. The data recorded at three different frequencies (10, 2 and 1 Hz) were processed using an algorithm working on GIS software which was able to identify tangents, spirals, and circular curves. The algorithm was validated by comparing the following data from circular curves: radius values ranging from 349.00 to 873.20 m and length values from 162.40 to 783.64 m. Results showed a limited difference in radius estimation (in a range between -4.74% and +6.16%) only when the average values from different speed surveys in each direction were evaluated. In the case of curve length (L_C), larger differences were observed between the estimated and real values (average increase of 4.6% and absolute average difference of 16.7%).

Finally, Tsai et al. (2009) proposed the use of photographic images taken using low-cost digital cameras. They developed an algorithm using vision technology to acquire and analyze horizontal roadway curvature data from roadway images. More recently, Ai and Tsai (2014) presented an automated method to extract horizontal curve radii from GPS data. In the investigation, radii ranging from 71.01 m to 1731.87 m derived from a fitting algorithm were compared using the in-field chord method. The relative average error was equal to 4.83% ± 1.8%. Unfortunately, no other data describing the characteristics of the investigated curves (i.e., length, angle of deflection) were available.

From the literature review based on civil and transportation engineering studies, it is worth noting that previous contributions have focused mainly on the evaluation and/or comparison of existing methods and technologies in the interpretation of spatial data.

The available case studies have not covered the full spectrum of combinations obtainable from changing the magnitude of radius, the sampling frequency and the length of the investigated

circular element (in other terms, the number and the spatial extension of data points available), and the accuracy of the survey methodology. Consequently, at present there are no clear indications as to the most appropriate (i.e., fastest, cheapest and most accurate) method for the estimation of circular arc radii of road alignments. Furthermore, the alignment identification techniques proposed in literature do not consider any possible effects associated with the selection of the fitting algorithm.

OBJECTIVES AND METHODOLOGY

To address this gap in knowledge and try to understand the degree to which the radius estimation process is affected by the accuracy of original measurements and by the amount of spatial data available, the authors opted for an alternative approach to those used by previous authors. Instead of using raw survey data of circular arcs as the various authors mentioned in the Background did, an algorithm, able to generate different arcs in terms of radius, length and deviation angle, was constructed first, and then used to:

a) discretize the arcs generated into a sequence of spaced points,
b) modify the points from their original position by adding a Gaussian noise,
c) modify a specific percentage of points (i.e., 5, 10 and 20% of the total) to generate large outliers (e.g., to simulate the effects associated with the use of GPS, which may be affected either by location errors, temporary occlusions, or unusually high ionospheric delays).

The discretized circular arcs were then back-calculated using different fitting algorithms to derive the original curve characteristics, and finally the initial values were then compared with the back-calculated ones.

In keeping with this procedure, sequences of points generated from circular arcs with known radius and length were converted into data populations with controlled accuracy, similar to those

derived from field or aerial surveys, or from GIS data. With these populations, three fitting methods were used and tested to back-calculate the geometric characteristics of the original curves.

The authors are aware that in-field measurements are affected by different disturbance phenomena such as thermal, flicker, white, and electronic interference noises, which derive from the electronic components of sensors and related devices. In this investigation, the authors considered the effects of only white and Gaussian noises since the others are difficult to identify and are strongly dependent on the quality of sensors. Moreover, such sources of noise may be filtered and reduced after post-processing analyses.

In the following, both traditional statistical and robust methods have been considered to understand their behavior when used with datasets affected by noise and gross errors.

FITTING METHODS

The basic equation for the circumference or segment of a circular arc in the (x,y) plane is:

$$x^2 + y^2 + ax + by + c = 0 \qquad \text{(eq. 7)}$$

where a, b and c are the parameters that fully define its position and characteristics. The radius (r), and the coordinates of the curve center (x_c,y_c), can be derived from the following equations:

$$x_c = -\frac{a}{2},\ y_c = -\frac{b}{2} \qquad \text{(eq. 8)}$$

$$r = \sqrt{x_c^2 + y_c^2 - c} \qquad \text{(eq. 9)}$$

Least Squares Method

The most common method, one that has a rigorous and sufficiently accurate approach, used to solve a system of equations is the Least Squares (LS) method (Wolberg 2005). This method is suggested when the data have a good redundancy and when the number of gross errors (or outliers) is limited.

The LS method is based on the overall solution obtained by minimizing the sum of squared residuals, where the residual is the difference between an observed value and the fitted value provided by a model. The LS solution is defined by starting from the residuals *v*, which are equal to:

$$\boldsymbol{Ax} - \boldsymbol{T} = \boldsymbol{v} \qquad \text{(eq. 10)}$$

where $\boldsymbol{A}$ is the design matrix, $\boldsymbol{x}$ is the vector of the unknown terms, and $\boldsymbol{T}$ is the vector of the known terms. The design matrix represents the matrix of the coefficients of the unknown parameters obtained with a linearization of the problem. Considering the case of the estimation of the three parameters of a circular arc, it is possible to start from the following system of equations:

$$\begin{cases} ax_1 + by_1 + c = -x_1^2 - y_1^2 \\ ax_2 + by_2 + c = -x_2^2 - y_2^2 \\ ax_3 + by_3 + c = -x_3^2 - y_3^2 \\ \vdots \\ ax_n + by_n + c = -x_n^2 - y_n^2 \end{cases} \qquad \text{(eq. 11)}$$

The LS solution (*x*) is obtained by means of the following equation:

$$\boldsymbol{x} = \boldsymbol{N}^{-1}\boldsymbol{T}_N \qquad \text{(eq. 12)}$$

in which $\boldsymbol{N}$, the Normal matrix, is equal to $\boldsymbol{A}^T\boldsymbol{A}$ $\boldsymbol{T}_N$ is the normalized unknown term, which is equal to $\boldsymbol{A}^T$.

To appreciate the quality and precision of the estimated solution *(a, b, c)*, the variance-covariance matrix can be determined with the following equation:

$$\boldsymbol{C}_{xx} = \sigma_0^2\boldsymbol{N}^{-1} = \begin{bmatrix} \sigma_a^2 & \sigma_{ab} & \sigma_{ac} \\ \sigma_{ba} & \sigma_b^2 & \sigma_{bc} \\ \sigma_{ca} & \sigma_{cb} & \sigma_c^2 \end{bmatrix} \qquad \text{(eq. 13)}$$

where the variances of *a, b,* and *c* are to be found in the diagonal, while the covariances are outside the diagonal.

As noted before, the geometric parameters are not directly estimated, therefore their precision has to be assessed by applying the variance propagation law (Jeudy 1988). If σ_a^2, σ_b^2 and σ_c^2 are deemed to be the variances of the estimated parameters *a, b* and *c* respectively, the variances of the radius and center coordinates are:

$$\sigma_{xc}^2 = \frac{1}{4}\sigma_a^2 \quad \text{(eq. 14)}$$

$$\sigma_{yc}^2 = \frac{1}{4}\sigma_b^2 \quad \text{(eq. 15)}$$

$$\sigma_r^2 = \frac{x_C^2}{x_C^2+y_C^2-c}\sigma_{xc}^2 + \frac{y_C^2}{x_C^2+y_C^2-c}\sigma_{yc}^2 + \frac{1}{x_C^2+y_C^2-c}\sigma_C^2 \quad \text{(eq. 16)}$$

While one benefit of this approach is the accuracy of the estimated solution, the final solution is affected by the existence of gross errors, hence it is fundamental to anticipate the use of some statistical tests devoted to outlier detection or data analysis (Baarda 1967, 1968).

Least Median of Squares Method

Statistical methods are sometimes preferred to the LS method when the quality of data is not particularly high or when the risk of having a gross error is high. From literature, the various methods can be divided into the following three groups (Huber 1981):

a. methods for problems with outliers in the y-direction, where the most common procedures for solving these types of problems are based on the Huber estimator;
b. methods applied where there is a moderate percentage of outliers in the coordinate space, also known as the leverage point; the estimators (also known as bounded influence estimators) employed to solve this type of problem are the Mallows (Mallows, C. L. "On some topics in robustness". Unpublished memorandum, Bell Telephone Laboratories, Murray Hill, NJ) and Schweppe type (Handschin et al. 1975) estimators;
c. methods used when the frequency of outliers in both the *x* and *y* directions is high; in this case high breakdown estimators are used.

In all these methods, the object function of the Least Median of Squares (LMS) minimizes the median value of the squared residuals. The LMS estimator has a breakdown point equal to 50%, which means that this method can provide a solution even when 50% of the data are gross errors.

One disadvantage, however, is its poor asymptotic efficiency, which renders the estimator less efficient in terms of precision when compared to LS.

For this investigation, the Huber estimator was adopted (Huber 1981). In the Matlab® language (Mathworks 2011) the statistic toolbox denominated *robustfit* was used for this purpose, with a weight function that was set equal to 1.345 which derives from a specific tuning of the algorithm.

Landau Method

Alternative methods could come from other disciplines in which the estimation of geometric elements is a fundamental aim as in computer vision where 3D modeling reconstruction is a regular activity.

One example is the Landau method (Landau 1987; Thomas and Chan 1989), in which a series of numerical variables is used to estimate the properties of circular arcs. The procedure, which initiates with a definition of the numerical p_i variables, is briefly summarized as follows:

$$\begin{cases} p_1 = \sum_{i=1}^{N} x_i \\ p_2 = \sum_{i=1}^{N} x_i^2 \\ p_3 = \sum_{i=1}^{N} x_i\, y_i \\ p_4 = \sum_{i=1}^{N} y_i \\ p_5 = \sum_{i=1}^{N} y_i^2 \\ p_6 = \sum_{i=1}^{N} x_i^3 \\ p_7 = \sum_{i=1}^{N} x_i y_i^2 \\ p_8 = \sum_{i=1}^{N} y_i^3 \\ p_9 = \sum_{i=1}^{N} x_i^2\, y_i \end{cases} \qquad \text{(eq. 17)}$$

where N is the number of points, and (x,y) the coordinates of points. Working from the definition of p_i, other parameters dependent on p_i and N are derived:

$$\begin{cases} a_1 = 2 \cdot (p_1^2 - N \cdot p_2) \\ b_1 = 2 \cdot (p_1 \cdot p_4 - N \cdot p_3) \\ a_2 = b_1 \\ b_2 = 2 \cdot (p_4^2 - N \cdot p_5) \\ c_1 = p_1 \cdot p_2 - N \cdot p_6 + p_1 \cdot p_5 - N \cdot p_7 \\ c_2 = p_2 \cdot p_4 - N \cdot p_8 + p_4 \cdot p_5 - N \cdot p_9 \end{cases} \qquad \text{(eq. 18)}$$

Finally, the radius and coordinates of the center are estimated according to the following three equations:

$$x_C = \frac{c_1 \cdot b_2 - c_2 \cdot b_1}{a_1 \cdot b_2 - a_2 \cdot b_1} \quad \text{(eq. 19)}$$

$$y_C = \frac{a_1 \cdot c_2 - a_2 \cdot c_1}{a_1 \cdot b_2 - a_2 \cdot b_1} \quad \text{(eq. 20)}$$

$$r = \sqrt{\frac{p_2 - 2 \cdot p_1 \cdot x_C + N \cdot x_C^2 + p_5 - 2 \cdot p_4 \cdot y_C + N \cdot y_C^2}{N}} \quad \text{(eq. 21)}$$

DISCRETIZATION AND FITTING ALGORITHM

The processes of discretization and fitting of data points have been performed on some reference arcs. In all the analyses, original dedicated routines in the Matlab® language (Matworks 2011) were used.

As previously indicated, the investigation was carried out to highlight the effects of the accuracy of spatial points, their interval (which is related to the sampling frequency), and the length of circular arcs from which they (the points) were extracted.

The dataset of points generated in the simulation is extracted along circular arcs (Figure 2), following the selection of parameters such as the interval between points *s*, i.e., the average distance measured along the curve from points extracted from the arc, the radius *r* and the central angle α, as per the algorithm reported in Figure 3a. Using the algorithm, points are generated by varying α from a minimum (α_{min}) to a maximum value (α_{max}). The α_{min} is set at the level required to avoid numerical instability of the fitting process. A minimum level of numerical redundancy in the equation is guaranteed, thus a minimum of four points (N_{pts} = 4) are considered in the fitting process (one more than the three research parameters *a*, *b* and *c*).

Afterwards, each single point of the sequence is moved slightly from its original position by applying two noise terms (Δx and Δy) through the implementation of a sub-routine in the algorithm described in Figure 3a. In particular, if *N* is the size of the original extracted points, the sub-routine

generates a series of normally distributed errors inside a specific boundary of the same size *N*, with the mean equal to zero and a standard deviation equal to 1 (i.e., Gaussian distribution). The errors are then scaled to a specified accuracy (or maximum error), indicated as e_{max}.

A different sub-code, explained in Figure 3b, was then used to determine the new position of those points that were artificially modified to become outliers. The number was selected a priori ($N_{outliers}$), the shifting direction (θ), the selected points, and their new positions were defined randomly (e.g., on a curve composed of 50 points, if the outliers represent 10% of the total, 5 points were randomly selected and then moved from their initial position). Then, a loop process was inserted into the code to move the selected outliers from their original positions to the new shifted coordinates.

Then, following the generation of the matrixes (*A*, *T*) the fitting process is launched. At the end, the radius and center point of the fitted circle is stored in the database, and this process is looped until the angle reaches the maximum value of 90°. Every arc is treated with values of e_{max} equal to 2, 5, 10, 20, and 40 cm, and different intervals between points equal to 0.5, 1, 2, 5, 10, 15, and 20 m. In the case with the outliers, a maximum error of 2 m was adopted.

Figure 4 contains an example of the noise generated with the code for a combination of two maximum errors (5 and 40 cm) and two intervals (0.5 and 20 cm), for an arc of length equal to 120 m. In the figure, the origin of the coordinated axes corresponds to the initial (exact) position of points when they were extracted from the original curve. In the algorithm (Figure 3b), eq. 7 was rearranged into the system of *N* equations:

$$\begin{cases} x_1^2 + y_1^2 + ax_1 + by_1 + c = 0 \\ x_2^2 + y_2^2 + ax_2 + by_2 + c = 0 \\ x_3^2 + y_3^2 + ax_3 + by_3 + c = 0 \\ \vdots \\ x_N^2 + y_N^2 + ax_N + by_N + c = 0 \end{cases} \quad \text{(eq. 22)}$$

and then solved by means of statistical and numerical methods previously described.

EFFECTS OF FITTING METHODS ON IDENTIFICATION

Figure 5 contains an example of a circular arc composed of 55 points spaced 5 m apart, and determined to within an accuracy of 10 cm, generated according to the procedure described in the "Objectives and Methodology" section. The arc has been fitted with three different curves determined according to the LS, LMS Huber and Landau methods.

No disparities in the results are evident from an analysis of the graph, while the data reported in Table 2 for 0% of outliers evidence the differences in the estimation of the *a*, *b* and *c* parameters, the radius and the coordinates of the center of curvature. Of the three solutions, the LS and Landau ones are practically identical, while the LMS Huber one is slightly different and closer to the original circle. The graph in Figure 6 contains the fitting of the same data points that were previously modified including 5 (Figure 6a), 10 (Figure 6b) and 20% (Figure 6c) of outliers, while Table 2 reports the complete set of results for a presence of outliers equal to 5, 10 and 20%.Tables 2 highlights the benefits deriving from the use of a more robust fitting technique as opposed to the ordinary LS method, or the equivalent Landau method. In particular, the differences indicated under the symbol $\Delta(\cdot)$ in the tables, and expressed as a percentage of the reference value of the original arc (included in Table 2), become noticeable in the presence of outliers. If, in the case without outliers, the radius is underestimated with an error value greater than 3.2 (LMS Huber method) or 3.5% (LS and Landau methods), then, in the cases with outliers, an increase in the percentage of outliers will result in increased errors together with underestimated radii values. In the case of 5% of outliers the percentage error with the LS and Landau methods is 14.8%, while with the LMS Huber method it is equal to 7.6%; when the percentage of outliers is 10% the errors are 31.6 and 13.3% respectively; finally, when the percentage of outliers is 20% the errors are 40.0 and 30.6% respectively.

It is also worth noting that when the level of disturbance in the data (i.e., quantity of outliers) increases, the center of the fitted circular arcs tends to move towards the curve, and this can explain the significant decrease in radius estimates.

EFFECTS OF ACCURACY, INTERVAL, AND RADIUS MAGNITUDE

Figure 7 reports the results obtained for the same target value of 550 m. The x-axis represents the value of the subtended arc (α), while the y-axis reports the radius (r) of the estimated arc as per the LS method. The graph includes ten different results generated by running the algorithm ten times using an interval of 0.5 m and 40 cm of accuracy. The numerical solutions pertaining to the same computation have been linked to the same line. The results are distributed along an S-shaped curve comprising the two horizontal asymptotes of 550 m for $\alpha \to 90°$, and of 0 m for $\alpha \to 1°$.

This indicates that when data are affected by casual errors with known levels of accuracy, radii estimated with fitting algorithms are frequently underestimated (in Figure 7 the estimated radius is above the target value of 550 m in only a few cases). Hence, a minimum length of the curve (or the subtended angle α) is necessary to get a good approximation of the expected radius. With respect to the parameters used for the case in Figure 7, the best results when estimating the radius can be obtained for central angles greater than 20°. Results also depend on the causality in the distribution of errors between points from the same curve.

Figure 8 illustrates the results obtained from the algorithm assuming a radius of 550 m. The first graph (Figure 8a) reports the radius estimates for different interval values between points with an accuracy of 5 cm. The results confirm that good radius values can only be obtained when the fitting algorithm elaborates a sufficient number of points. In this case, the interval (between points) does not seem to be of critical importance for the estimation of the radius. It is only when high interval values (> 5 m) are considered that the number of points available are lower than 4 for some α amplitudes and thus not sufficient to run the fitting algorithm. This happens in cases where the

central angle α is very low. For this specific set of parameters, good results can be obtained when the amplitude of the central angle is greater than 10°.

The curves in the second graph (Figure 8b) have been obtained for an interval of 0.5 m and an accuracy ranging from 2 to 40 cm. Results indicate that the minimum acceptable value of α increases when the accuracy of the dataset of points decreases. Hence, when the value of α is low (short arcs), a good quality of fitting can be expected only when data are collected with a very high level of accuracy.

In Figure 9, the effects of a change in the radius magnitude from 25 to 5500 m are plotted for two accuracy levels (2 and 40 cm), and points generated with a fixed interval of 0.5 m. In the figure, the lines represent the fitting radii obtained by starting with circular arcs of known radius, which were then discretized and modified with the Gaussian noise as per the indications in the accompanying text.

In the case of small radii, results show that either an increased arc length (or angle) or greater accuracy in measurements is necessary to attain a satisfactory estimate of the radius. For example, when the radius is 25 m, in the case of high accuracy (2 cm) the minimum arc angle required for a good estimate of the fitting radius is around 14°. For the same arc, in the case of low accuracy (40 cm) the minimum arc angle required to obtain satisfactory results increases to 25°.

Minor problems are encountered when the radius increases. When the radius is 5500 m, no problem of fitting occurs when there is high sampling frequency and accuracy. With low accuracy (40 cm), poor results in the fitting radius are to be expected for angles lower than 3°.

These results confirm the difficulties faced by some researchers when fitting the arcs with small radii, since a significant quantity of data points is needed, especially when the accuracy is poor (Rasdorf et al. 2010).

Starting from the data in Figure 9, in Figure 10 each point from the two graphs with a bi-logarithmic scale represents a combination of minimum angle and arc radius that differs more

than 2% from the radius of the arc from which the spatial points were generated. Figure 10 exhibits the regression curves that follow the equation:

$$R = m \cdot \alpha^n \quad \text{(eq. 23)}$$

with intervals between data points of 0.5 m (Figure 10a), and 15 m (Figure 10b). In both figures, the accuracy value ranges from 2 to 40 cm. Table 3 synthetized the values of the coefficients m and n included in eq. 23 with values of 0.5 and 15 m considered for the interval between surveyed points.

In Figure 10a an example with the reference radius of 400 m is considered. The horizontal dashed line intersects the five curves for angle values that go from 4.2° for e_{max} equal to 2 cm, to 16° for e_{max} equal to 40 cm. Figure 10b for the same radius shows that the angle values pass from 11° for e_{max} equal to 2 cm, to 29° for e_{max} equal to 40 cm.

Looking at the two graphs, it is evident that the minimum central angle which ensures good estimation of radius values increases when the accuracy and interval decrease, thus suggesting the adoption of appropriate and precise surveying tools together with a short interval (high frequency of sampling) between spatial data points.

According to these results and referring to Table 1, photogrammetry, mobile mapping, and remote sensing are the most suitable road survey techniques for the identification of road alignment. These techniques while cost effective are not cheap, but they provide a great deal of information (i.e., data points) over long road segments. The costs could be reduced if unconventional systems like the UAV are used, which can operate on segments of limited extensions only.

It is worth noting that in the field of surveying, technologies evolve continuously and quickly with the introduction of new sensors with greater accuracy and lower costs. In perspective, the acquisition tool that presents the best compromise in terms of cost and performance today might not be the best in the future.

CONCLUSIONS

In the paper the authors sought to test the reliability of some identification algorithms for circular arcs in terms of location (x_C,y_C) and radius, and discuss the algorithms for the identification of geometric elements as a function of the type of source data. The use of a reliable, affordable, and consistent analytical tool and algorithm is fundamental for such applications.

Several algorithms exist for the identification of circular arcs belonging to road geometric elements, but the paper demonstrates that their effectiveness depends essentially on the geometric characteristics of the elements that have to be estimated. In particular, the authors have quantified the effects of certain factors such as the accuracy of data points surveyed, the length of the curve, the magnitude of the radius, and the fitting method used.

The results obtained demonstrate that, under normal conditions, where the points describing the circular curve are sufficient in number to allow coverage of a large part of the circular arc, and when the random errors that affect each measurement are small, then traditional and robust fitting methods exhibit similar behaviors. When large outliers affect the data, robust methods are more efficient in the identification of the center of curvature and the radius value. Furthermore, the results highlight that when only a small number of points are available with respect to radius size, the classical statistical methods are not efficient.

The paper provides graphs (Figure 10) that help to identify the threshold of the minimum central angle as a function of the interval between points and the accuracy of the measurement that allows a good estimation of the radius. The graph facilitates the selection of the survey system accuracy level which enables a precise estimation of the center and radius of circular arcs or, alternatively, it gives an idea of the possible accuracy that can be obtained when the performance of the acquisition system, and both the length and magnitude of the radius are known. The paper also provides useful information on the effects of the interval length between surveyed spatial data.

The use of simulated data to investigate the effects of error distribution on the consistency and reliability of results served to facilitate comprehension of the fitting issues arising from the use of spatial data points of road alignments. Any difference between the results presented here and previous results from literature depends on the different nature (random and/or systematic), magnitude and spatial distribution of errors with those artificially generated and treated in this study.

ACKNOWLEDGEMENTS

The research presented in the paper refers to part of the activities carried out in the **Pro-Vision Project** (Title: *Sistema Prototipale per le Verifiche di Visibilità su Infrastrutture di Trasporto Esistenti in Ambito di Smart City,* in Italian), which was funded by the Regione Piemonte (F.E.S.R. 2007/2013).

REFERENCES

Ai, C. and Tsai, J.. (2014). "An automated horizontal curve radii measurement method for roadway safety analysis using GPS data." *93rd Annual Meeting,* Transportation Research Board, No. 14-3833, Washington D.C.

Baarda, W. (1967), "Statistical concepts in geodesy." *Publ. on Geodesy*, 2(4), Netherlands Geodetic Commission, Delft, The Netherlands.

Baarda, W. (1968). "A testing procedure for use in geodetic networks." *Publ. on Geodesy*, 2(5), Netherlands Geodetic Commission, Delft, The Netherlands.

Banks, J.H (1998). *Introduction to Transportation Engineering*. McGraw-Hill.

Bassani, M., Lingua, A., Piras, M., De Agostino, M., Marinelli, G. and Petrini, G. (2012). "Alignment data collection of highways using mobile mapping and image analysis

techniques." *91st Annual Meeting,* Transportation Research Board, No. 12-0312 Washington D.C.

Berman, M. and Culpin, D. (1986). "The statistical behaviour of some least squares estimators of the centre and radius of a circle." *J. of Royal Stat. Soc.: Series B, Stat. Meth.*, 48, 183–196.

Carlson, P.J., Burris, M., Black, K. and Rose, E.R. (2005). "Comparison of radius-estimating techniques for horizontal curves." *Transp. Res. Rec.*, 1918, 76-83.

Chan, N.N. (1965). "On circular functional relationships." *J. of Royal Stat. Soc.: Series B, Stat. Meth.*, 27, 45–56.

Choi, S. and Sung, J. (2007). "Data generalization algorithm for the extraction of road horizontal alignment design elements using the GPS/INS data." *Adv. in Hybrid Inf. Tech.*, 4413, 51-62.

Dong, H., Easa, S.M. and Li, J. (2007). "Approximate extraction of spiralled horizontal curves from satellite imagery." *J. of Surveying Eng.*, 10.1061/(ASCE)0733-9453(2007)133:1(36), 36-40.

Drakopoulos, A. and Örnek, E. (2000). "Use of vehicle-collected data to calculate existing roadway geometry." *J. of Transportation Eng.*, 10.1061/(ASCE)0733-947X(2000)126:2(154), 154-160.

Easa, S.M., Dong, H.M. and Li, J. (2005). "Use of satellite imagery for establishing road horizontal alignments." *J. of Surveying Eng.*, 10.1061/(ASCE)0733-9453(2007)133:1(29), 29–35.

Gomarasca, M. (2009). *Basic of Geomatics*. Springer ed., London – New York.

Handschin, E., Schweppe, F. C., Kohlas, J., and Fiechter, A. (1975). "Bad data analysis for power system state estimation." *IEEE Trans. on Power App. and Syst.*, PAS-94, 329-337.

Harkey, D.L., Chang, Y. and Feaganes, J. (2004). "Evaluation and validation of automated in-vehicle data collection system for developing roadway alignments." *Transp. Res. Rec.*, 1897, 164-172.

Hauer, E. (1999). "Safety and the choice of degree of curve." *Transp. Res. Rec.*, 1665, 22-27.

Hofmann-Wellenhof, B., Lichtenegger, H. and Wasle, E. (2008). *GNSS – Global Navigation Satellite Systems*. Springer.

Huber, P. J. (1981). *Robust Statistics*. Hoboken, NJ: John Wiley & Sons, Inc.

Imran, I., Hassan, Y. and Patterson, D. (2006). "GPS-GIS-based procedure for tracking vehicle path on horizontal alignments." *Computer-Aided Civ. and Infr. Eng.*, 21, 383-394.

Kasa, I. (1976). "A circle fitting procedure and its error analysis." *IEEE Trans. on Instr. and Meas.*, 25, 8–14.

Kraus, K. (1997). *Photogrammetry*. *vol. 1 Fundamentals and Standard Processes, vol. 2 Advanced Methods and Applications*. Walter De Gruyter ed., Berlin – New York.

Jeudy, L.M.A. (1988). "Generalized variance-covariance propagation law formulae and application to explicit least-squares adjustments." *Bull. Géodésique*, 62 (2), 113-124.

Jimenez, F. (2011). "Improvements in road geometry measurement using inertial measurement systems in datalog vehicles." *Measurement*, 44 (1), 102-112.

Landau, U.M. (1987). "Estimation of a circular arc center and its radius." *Comp. Vision, Graphics and Image Proc.*, 38, 317–326.

Lee, J., Yun, D., Sung, J., Jung, Y. and Son, J. (2012). "The GPS-IMU performance test for road alignment surveying." *91st Annual Meeting*, Transportation Research Board, No. 12-1585, Washington D.C.

Li, Z., Chitturi, M.V., Bill, A.R. and Noyce, D.A. (2012). "Automated identification and extraction of horizontal curve information from GIS roadway maps." *Transp. Res. Rec.*, 2291, 80-92.

Maclean, A. (1994). *Remote Sensing & GIS*. American Society for Photogrammetry and Remote Sensing.

MathWorks (2011). *MATLAB User Manual*. Natik, Mass.

McGlone, J. C. (2013). *Manual of Photogrammetry*. American Society for Photogrammetry and Remote Sensing, 6th edition.

Mena, J.B. and Malpica, J.A. (2005). "An automatic method for road extraction in rural and semi-urban areas starting from high resolution satellite imagery." *Patt. Rec. Let.*, 26, 1201–1220.

Novak, K. (1995). "Mobile mapping technology for GIS data collection." *Phot. Eng. and Remote Sensing*, 61(5), 493-501.

Rasdorf, W., Findley, D.J., Zegeer, C.V., Sundstrom, C.A. and Hummer, J.E. (2010). "Evaluation of GIS applications for horizontal curve data collection." *J. of Transportation Eng.*, 10.1061/(ASCE)CP.1943-5487.0000127, 191-203.

Thomas, S.M. and Chan, Y.T. (1989). "A simple approach for the estimation of circular arc center and its radius." *Comp. Vision, Graphics and Image Proc.*, 45(3), 362-370.

Tsai, Y., Wu, J., Wang, Z., and Hu, Z. (2009). "Horizontal roadway curvature computation algorithm using vision technology." *Computer-Aided Civil and Infrastructure Engineering*, 25, 78–88.

Wolberg, J. (2005). *Data analysis using the method of least squares: Extracting the most information from experiments*. Springer.

LIST OF FIGURE CAPTIONS

Fig. 1. Simple (a) and combined (b) curves and their geometric characteristics

Fig. 2. Generation of a sequence of points from a circular arc

Fig. 3. Flow chart of the algorithms employed for the generation and fitting of dispersed data points; a) random normally distributed errors, b) outliers

Fig. 4. Error distribution around the true point for sampling distances (s) equal to 0.5 and 20m, with accuracy levels (e_{max}) equal to 5 and 40 cm for a circular arc of 108 m in length

Fig. 5. LS, LMS Huber, and Landau fitting of data points spaced 5 m apart, accuracy (e_{max}) of 10 cm, and pertaining to a curve of 108 m in length and an original radius of 550 m (with 0% of outliers)

Fig. 6. LS, LMS Huber, and Landau fitting of data points spaced 5 m apart, accuracy (e_{max}) of 10 cm, 20% of outliers, and pertaining to a curve of 108 m in length and an original radius of 550 m. (a) 5%, (b) 10%, and (c) 20% of outliers.

Fig. 7. Sequences of fitting solutions obtained for a 550 m circular arc as a function of the central angle (α) for an interval between points equal to 0.5 m, and accuracy (e_{max}) equal to 40 cm

Fig. 8. Values of fitting radii as a function of the maximum value of the central angle (α) for different intervals (a), and accuracy values (b) for a target curve of 550 m in radius

Fig. 9. Values of fitting radii as a function of the radius magnitude (accuracy values - e_{max} - of 2 and 40 cm, interval of 0.5 m)

Fig. 10. Threshold curves that guarantee a radius estimation error of 2%: (a) interval of 0.5 m, and (b) interval of 15 m

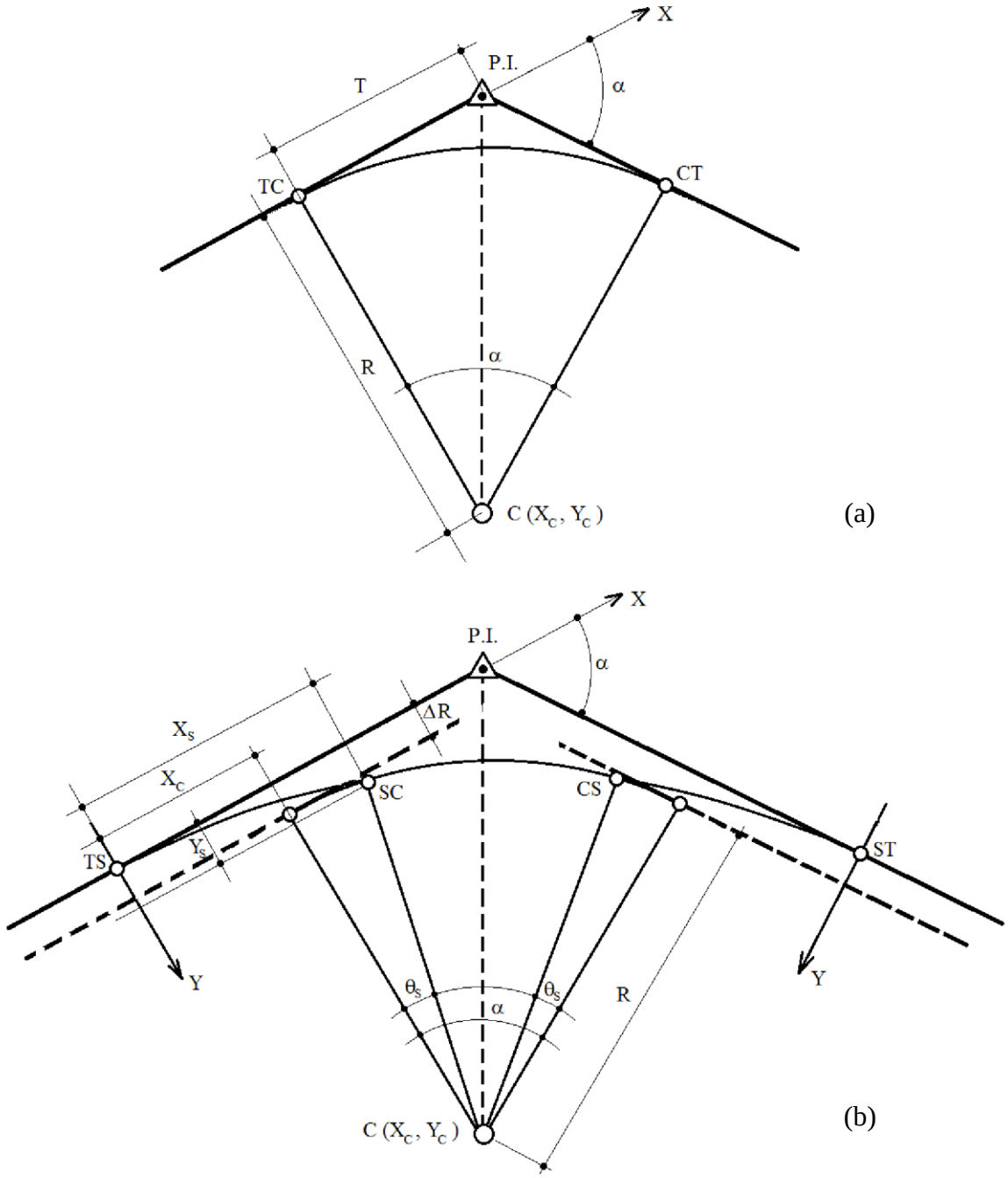


Fig. 1. Simple (a) and combined (b) curves and their geometric characteristics

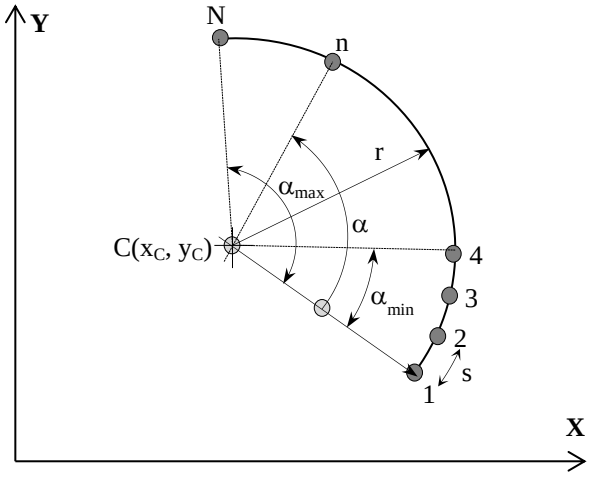
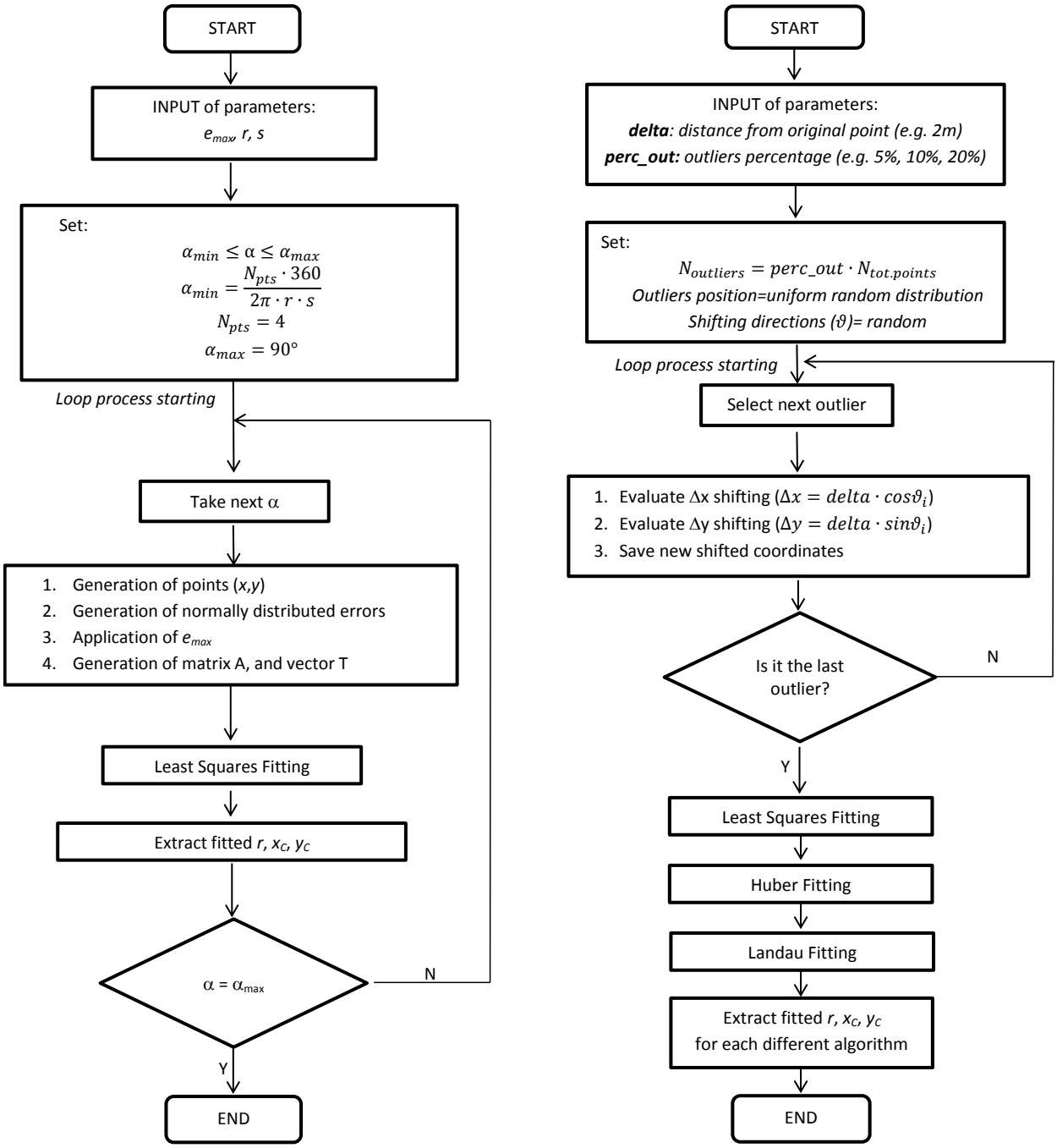


Fig. 2. Generation of a sequence of points from a circular arc



(a) (b)

Fig. 3. Flow chart of the algorithms employed for the generation and fitting of dispersed data points; a) random normally distributed errors, b) outliers

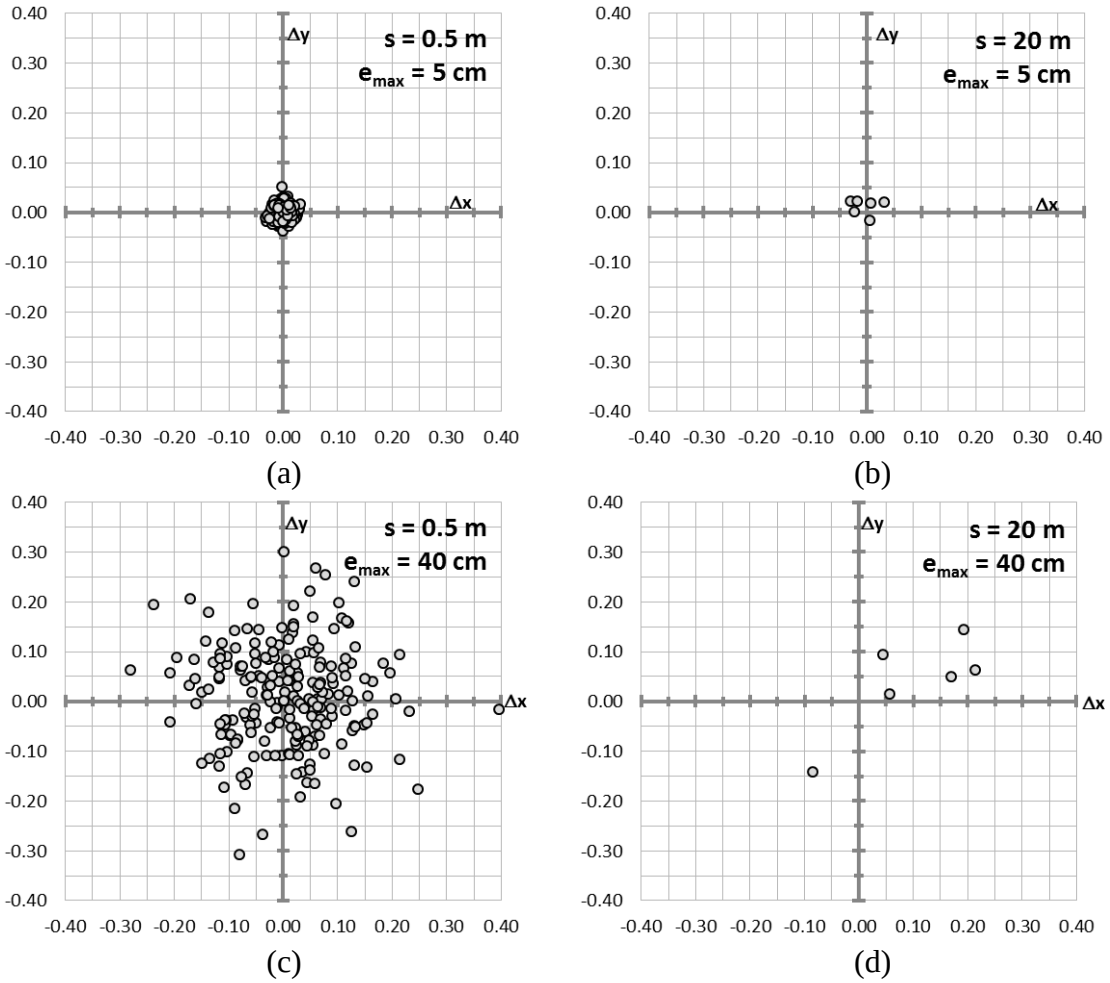


Fig. 4. Error distribution around the true point for sampling distances (s) equal to 0.5 and 20m, with accuracy levels (e_{max}) equal to 5 and 40 cm for a circular arc of 108 m in length

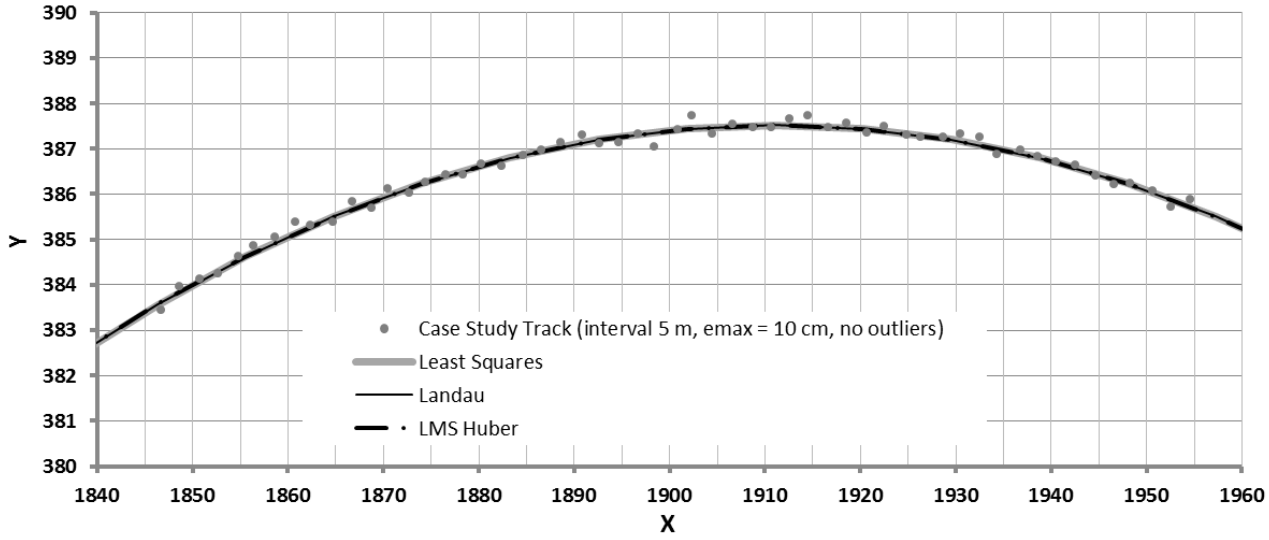
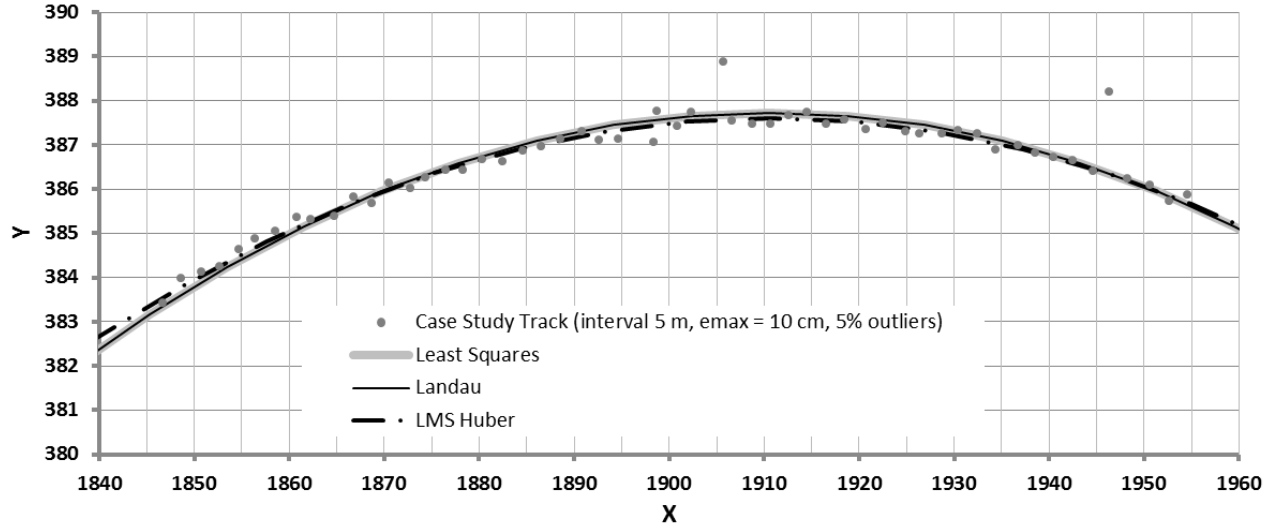
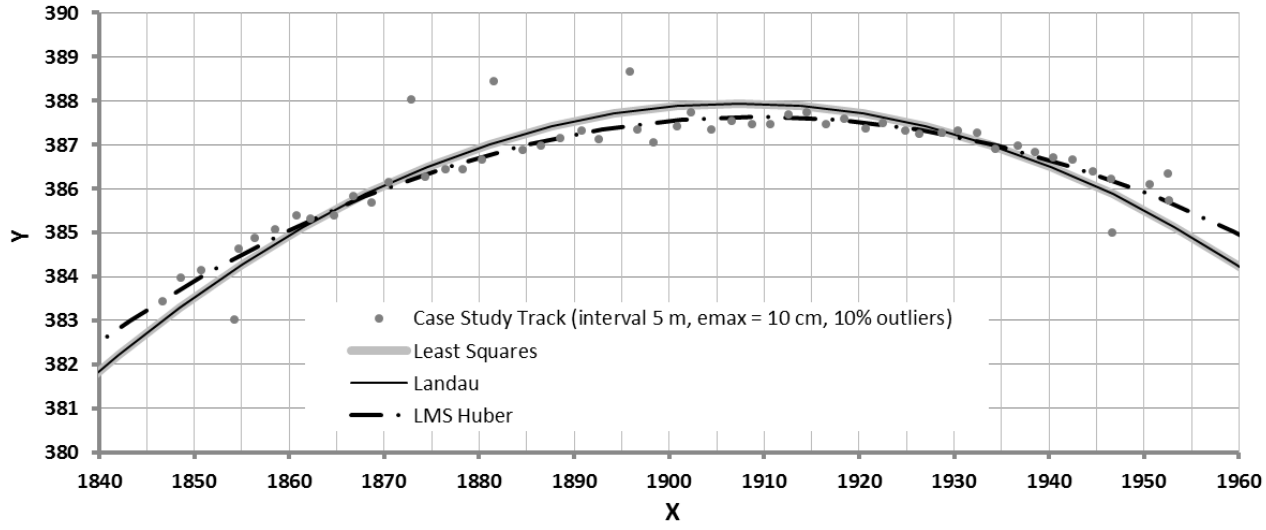


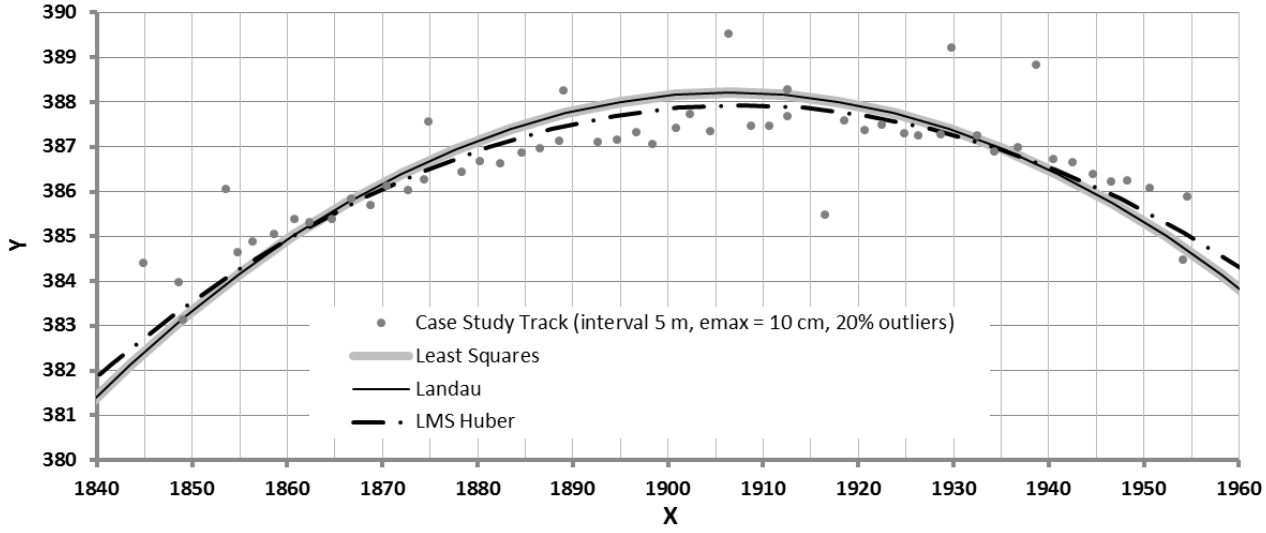
Fig. 5. LS, LMS Huber, and Landau fitting of data points spaced 5 m apart, accuracy (e_{max}) of 10 cm, and pertaining to a curve of 108 m in length and an original radius of 550 m (with 0% of outliers)



(a)



(b)



(c)

Fig. 6. LS, LMS Huber, and Landau fitting of data points spaced 5 m apart, accuracy (e_{max}) of 10 cm, and pertaining to a curve of 108 m in length and an original radius of 550 m. (a) 5%, (b) 10%, and (c) 20% of outliers.

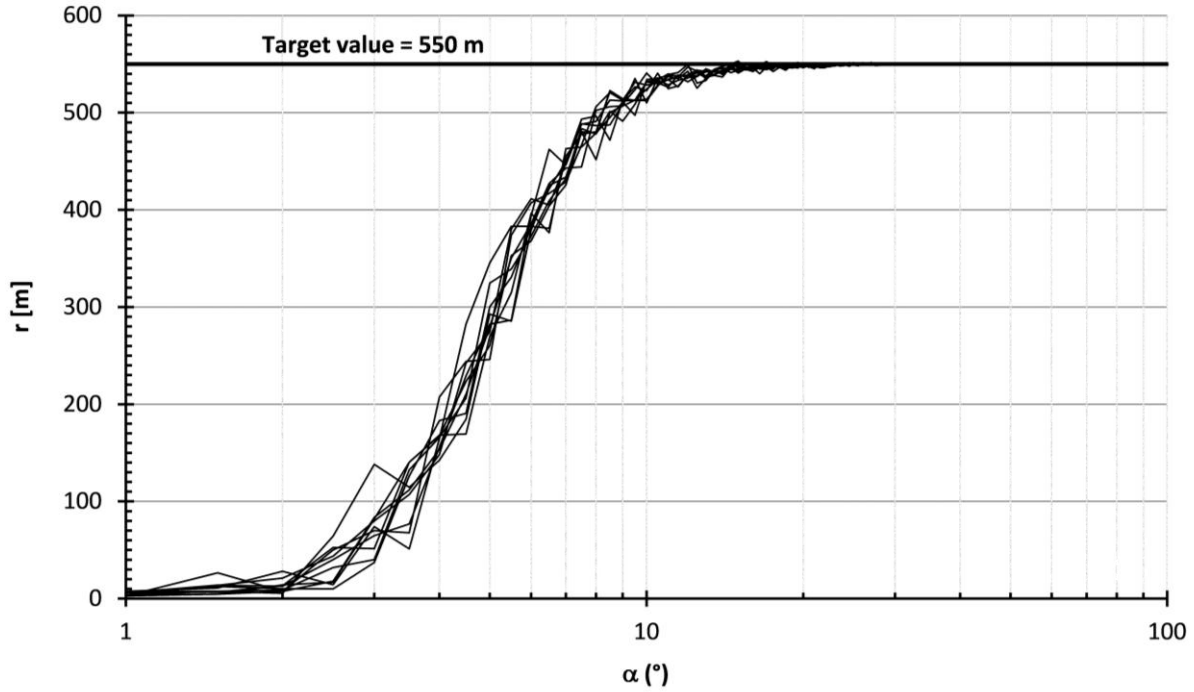
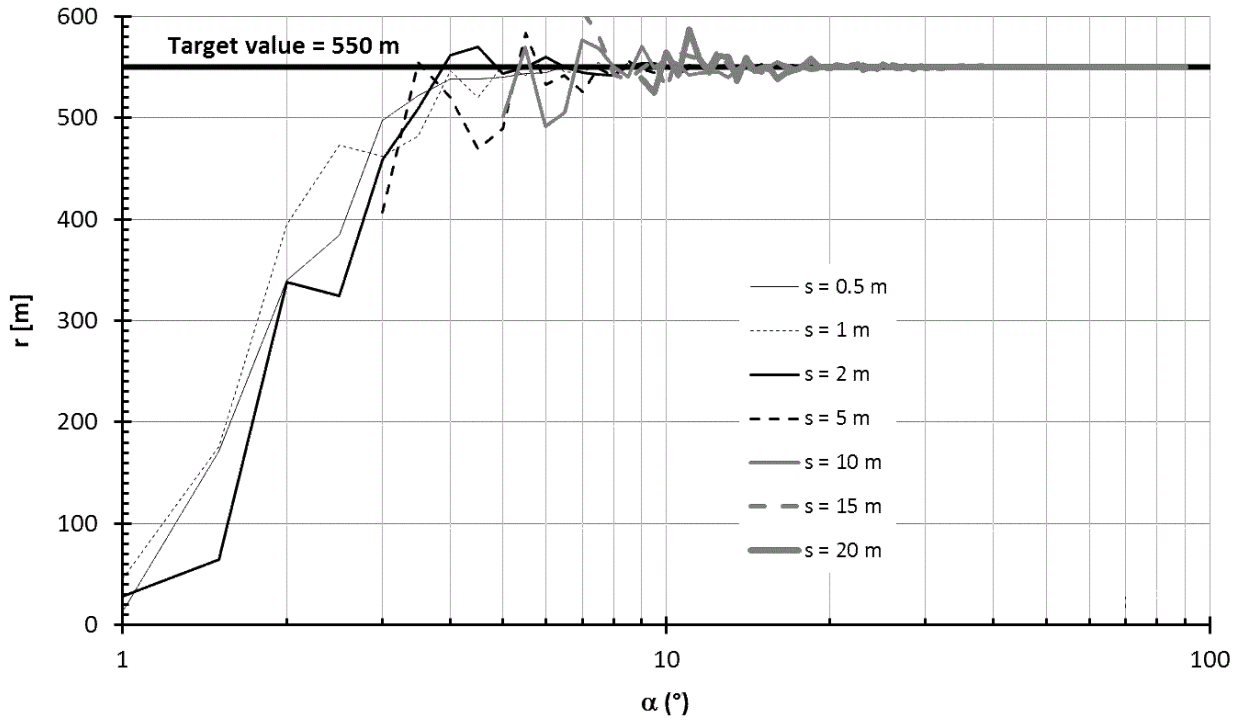
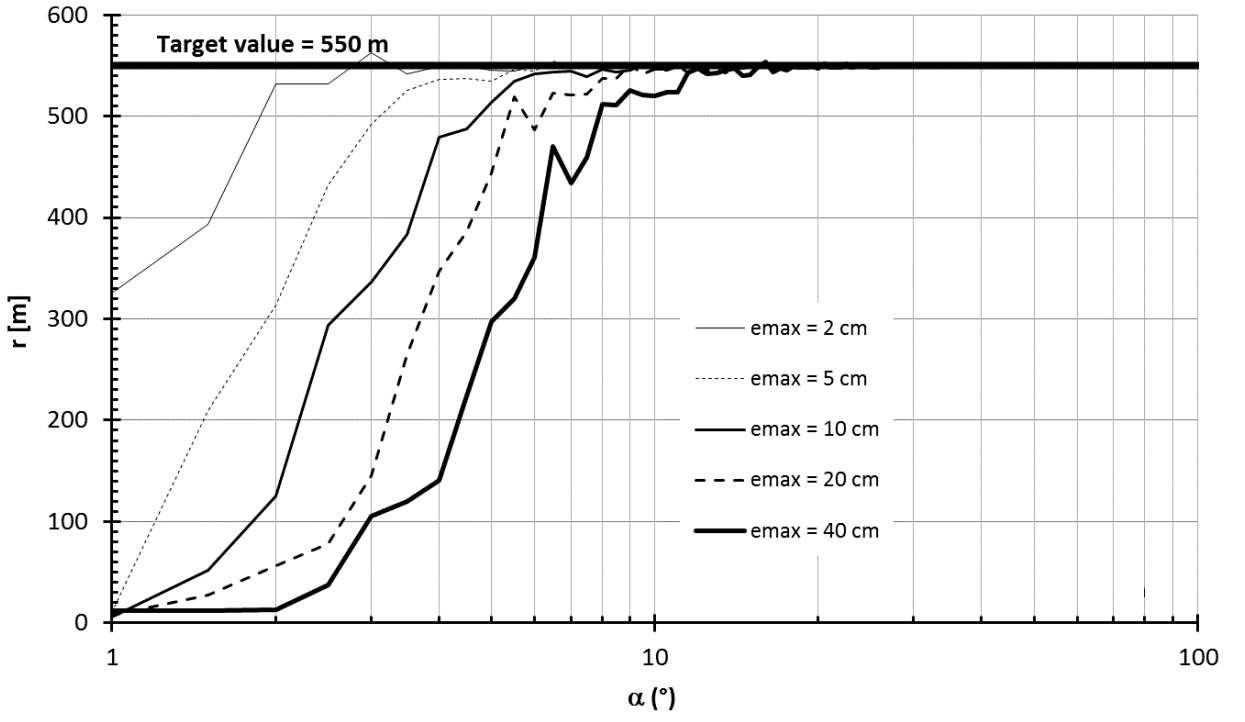


Fig. 7. Sequences of fitting solutions obtained for a 550 m circular arc as a function of the central angle (α), for an interval between points equal to 0.5 m, and accuracy (e_{max}) equal to 40 cm



(a)



(b)

Fig. 8. Values of fitting radii as a function of the maximum value of the central angle (α) for different intervals (a), and accuracy values (b) for a target curve of 550 m in radius

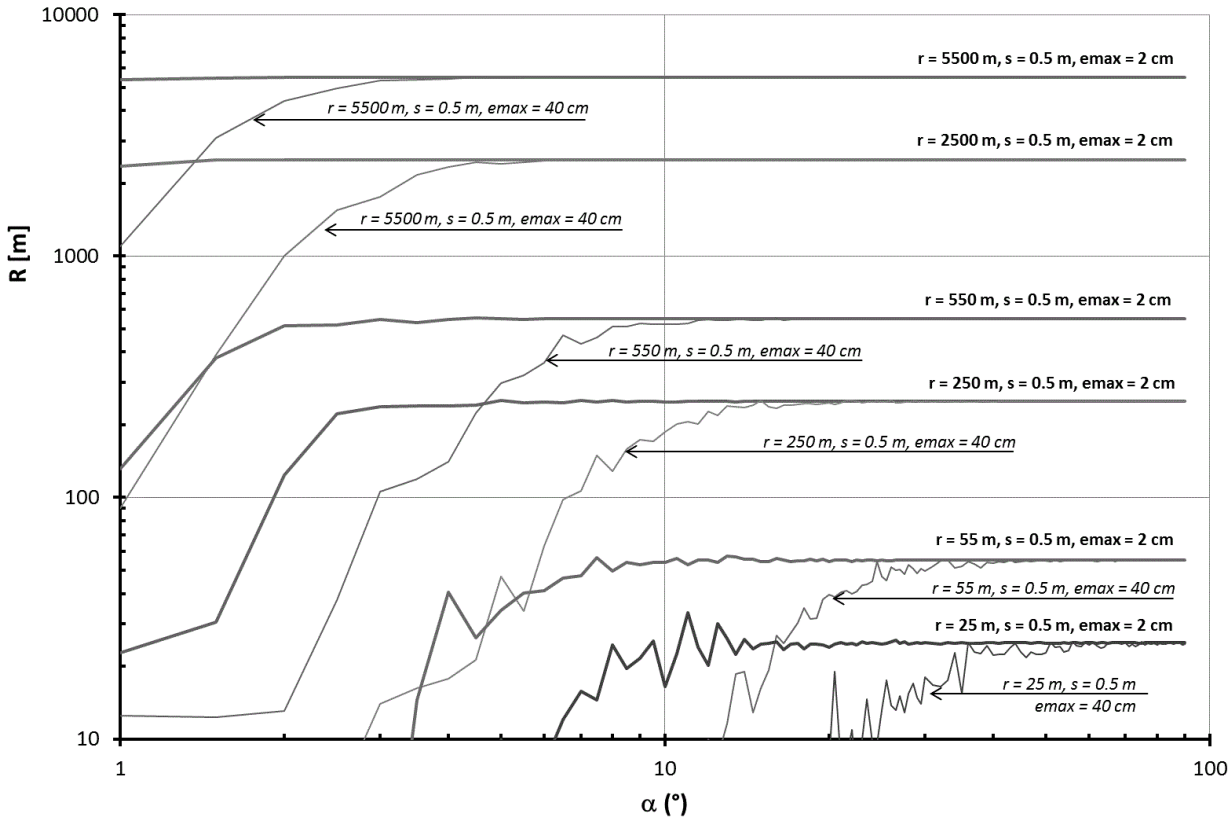
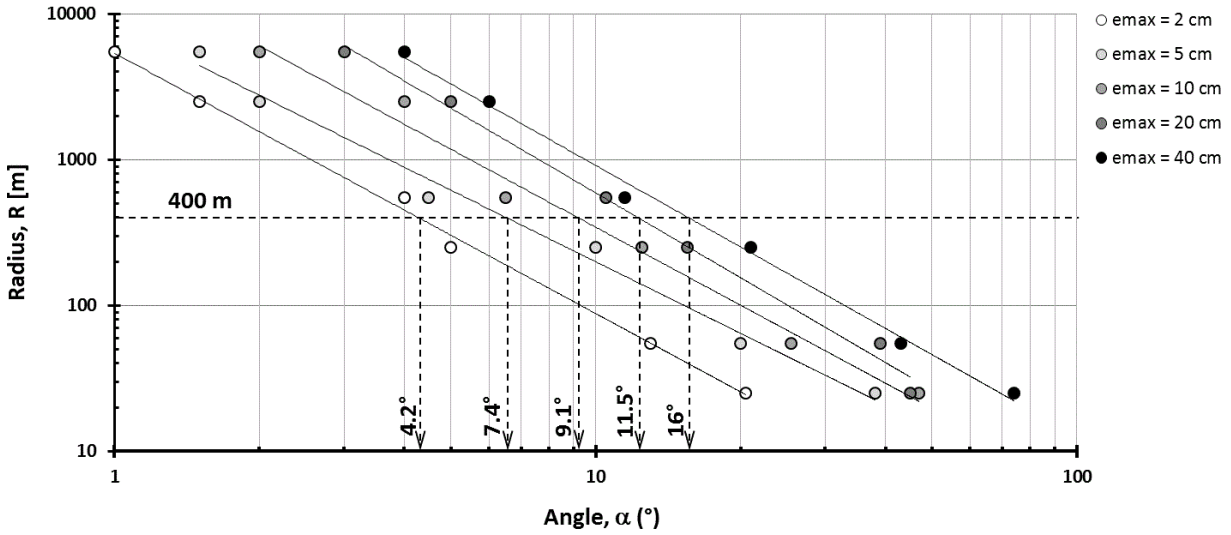
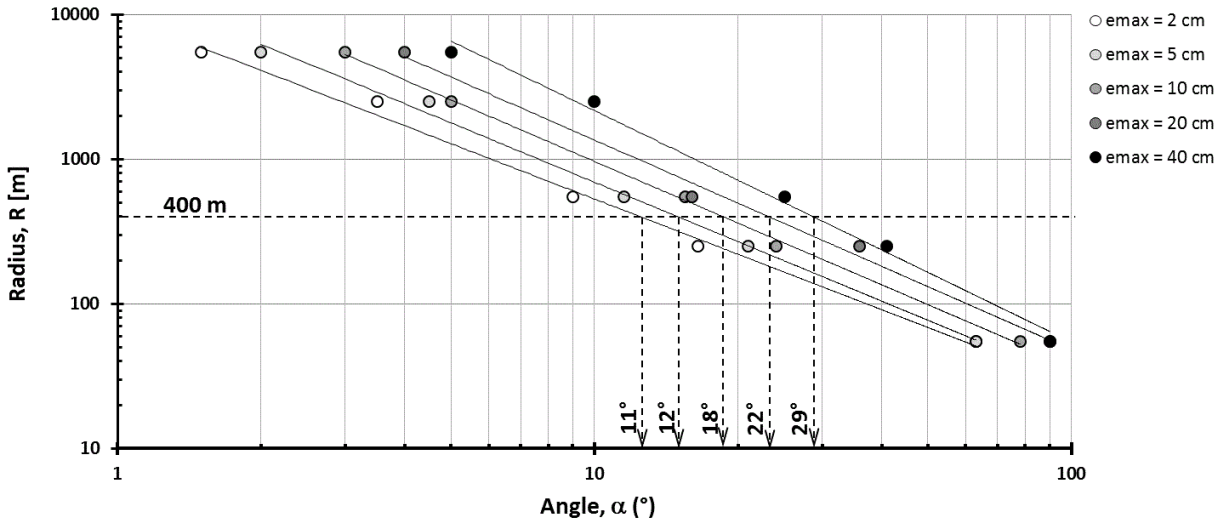


Fig. 9. Values of fitting radii as a function of the radius magnitude (accuracy values - e_{max} - of 2 and 40 cm, interval of 0.5 m)



(a)



(b)

Fig. 10. Threshold curves that guarantee a radius estimation error of 2%: (a) interval of 0.5 m, and (b) interval of 15 m

Table 1. Description, accuracy, and characteristics of current road survey techniques

Technique (and description)	Accuracy	Time ([a])	Cost ([b])	Final product
Photogrammetry (terrestrial or aerial acquisition of stereopairs of images)	0,5 cm – 1 m (with UAV or traditional planes)	4	4–5	point clouds, 3D models, and orthophotos
Real Time Kinematics (RTK) Global Navigation Satellite System (GNSS) (measurement of single point or kinematic trajectory using GNSS receiver)	2 – 4 cm	4	2	single point with coordinates
Mobile mapping system (MMS) (terrestrial vehicle equipped with on board sensors such as digital camera, laser scanner, GNSS, Inertial Measurement Units)	1 cm – 1 m (depending on the distance between MMS and object)	2	4	point clouds, 3D models, and orthophotos
Terrestrial Laser Scanner (acquisition with LiDAR)	1 cm – 50 cm (depending on the distance between LiDAR and object)	3	3	3D clouds with millions of points
Remote sensing (satellite images)	0,6 m – 5 m (depending on the type of satellite)	1	3	orthophotos and 3D models
Total station (measurement of single point or kinematic trajectory using total station)	1 – 4 cm	5	2	single point with coordinates

Note: ([a]) 1 faster – 5 slower, ([b]) 1 cheaper – 5 more expensive

Table 2. Synthesis of results of fitting for the circular arc of Fig. 5 (with 0% of outliers), and Fig. 6 (with 5, 10, and 20% of outliers)

Parameters		**a** (**Δa**)	**b** (**Δb**)	**c** (**Δc**)	**x_c** (**$Δx_c$**)	**y_c** (**$Δy_c$**)	**r** (**Δr**)
u.m.		**m** (**%**)	**m** (**%**)	**m** (**%**)	**m** (**%**)	**m** (**%**)	**m** (**%**)
Original arc		-3,822.48	324.929	3,376,728.91	1,911.24	-162.465	550
0% of outliers (Fig. 5)	LS	-3,822.03	286.581	3,390,751.72	1,911.02	-143.291	530.816
		(0.0)	(11.8)	(-0.4)	(0.0)	(11.8)	(3.5)
	Landau	-3,822.03	286.581	3,390,751.72	1,911.02	-143.291	530.816
		(0.0)	(11.8)	(-0.4)	(0.0)	(11.8)	(3.5)
	LMS Huber	-3,822.04	289.288	3,389,719.06	1,911.02	-144.644	532.169
		(0.0)	(11.0)	(-0.4)	(0.0)	(11.0)	(3.2)
5% of outliers (Fig. 6)	LS	-3,821.10	161.939	3,437,075.51	1,910.55	-80.969	468.704
		(0.0)	(50.2)	(-1.8)	(0.0)	(50.2)	(14.8)
	Landau	-3,821.10	161.939	3,437,075.51	1,910.55	-80.969	468.704
		(0.0)	(50.2)	(-1.8)	(0.0)	(50.2)	(14.8)
	LMS Huber	-3,821.34	241.181	3,406,947.48	1,910.67	-120.59	508.188
		(0.0)	(25.8)	(-0.9)	(0.0)	(25.8)	(7.6)
10% of outliers (Fig. 6)	LS	-3,814.70	-23.043	3,496,439.79	1,907.35	11.522	376.411
		(0.2)	(107.1)	(-3.5)	(0.2)	(107.1)	(31.6)
	Landau	-3,814.70	-23.043	3,496,439.79	1,907.35	11.522	376.411
		(0.2)	(107.1)	(-3.5)	(0.2)	(107.1)	(31.6)
	LMS Huber	-3,819.13	178.041	3,427,153.36	1,909.56	-89.02	476.662
		(0.1)	(45.2)	(-1.5)	(0.1)	(45.2)	(13.3)
20% of outliers (Fig. 6)	LS	-3,812.93	-116.378	3,529,087.50	1,906.47	58.189	330.019
		(0.2)	(135.8)	(-4.5)	(0.2)	(135.8)	(40.0)
	Landau	-3,812.93	-116.378	3,529,087.50	1,906.47	58.189	330.019
		(0.2)	(135.8)	(-4.5)	(0.2)	(135.8)	(40.0)
	LMS Huber	-3,815.37	-12.777	3,493,727.53	1,907.69	6.388	381.541
		(0.2)	(103.9)	(-3.5)	(0.2)	(103.9)	(30.6)

Table 3. Regression parameters of eq. 23

Interval [m]	Accuracy [cm]	m	n	R^2
0.5	2	5379	-1.786	0.996
	5	8557	-1.632	0.990
	10	20601	-1.777	0.988
	20	50756	-1.933	0.993
	40	65804	-1.857	0.995
15	2	9912	-1.271	0.996
	5	16096	-1.365	0.990
	10	23703	-1.399	0.988
	20	29376	-1.384	0.993
	40	86233	-1.600	0.995